

Dispense di Comunicazioni Elettriche

per Ingegneria Informatica e Bioingegneria

Prof. Aldo A. Grattarola

tel. 010 353 2987

e-mail aldo.grattarola@unige.it

DITEN. – Università di Genova
via Opera Pia 13
16145 GENOVA

Edizione settembre 2023

Indice

1 Segnali e sistemi	2
1.1 Segnali	2
1.1.1 Classificazione dei segnali	2
1.1.2 Segnali fisicamente realizzabili	3
1.1.3 Media, energia e potenza	3
1.1.4 Segnali periodici	4
1.2 Numeri complessi e fasori	4
1.3 Sistemi di comunicazione	6
1.4 Sistemi a tempo continuo	6
1.5 Sistemi LTI	8
1.5.1 L'impulso di Dirac o funzione delta	8
1.5.2 Risposta all'impulso	9
1.5.3 L'integrale di convoluzione	11
1.5.4 Impulsi e convoluzione	12
1.5.5 La funzione di trasferimento	12
1.6 La trasformata di Fourier	13
1.6.1 Proprietà	14
1.6.2 Trasformate notevoli	18
1.6.3 Trasformate di segnali generalizzati	20
1.6.4 Banda di un segnale	21
1.6.5 Relazioni tempo-frequenza e principio di indeterminazione	21
1.7 Serie di Fourier	23
1.8 Trasformata di Fourier di funzioni periodiche	24
1.9 Spettro di densità di potenza	25
1.10 Spettro di densità di energia	26
1.11 Sistemi LTI: risposta in frequenza	26
1.11.1 Esempi di sistemi LTI	26
1.11.2 Diagrammi a blocchi	27
1.11.3 Correlazione incrociata ingresso-uscita	28
1.12 Filtri	29
1.12.1 Filtri ideali	30
1.12.2 Risposta all'impulso dei filtri ideali	30
1.12.3 Realizzabilità dei filtri	31
1.13 Esempio: FT e filtraggio di un segnale reale	32
1.13.1 Trasformata di Fourier	32
1.13.2 Filtraggio Passa Basso (LPF)	33
1.13.3 Filtraggio Passa Alto (HPF)	34
1.14 Sistemi senza memoria	35
2 Campionamento e quantizzazione	37
2.1 Campionamento	37
2.1.1 Teorema del campionamento	38
2.1.2 Filtro anti-aliasing	39
2.1.3 Ricostruzione	39
2.1.4 Esempio: campionare una sinusoide	40
2.1.5 Sample & hold	41
2.1.6 Chopper sampling	43
2.1.7 Teorema della dimensionalità	43

2.2	Quantizzazione	44
2.2.1	Quantizzazione non uniforme	45
2.3	Pulse Code Modulation (PCM)	46
2.3.1	Esempio: stima grafica della banda del PCM	48
2.3.2	Ricevitore PCM	50
2.4	Pulse Amplitude Modulation (PAM)	50
2.4.1	PAM su canale a banda illimitata	51
2.4.2	PAM su canale a banda stretta	52
2.5	Multiplexing TDM e FDM	57
2.5.1	TDM analogica	58
2.5.2	TDM numerica	59
2.5.3	FDM	60
3	Rappresentazione dei segnali	64
3.1	Spazi di funzioni	64
3.1.1	Prodotto interno, ortogonalità e norma	64
3.1.2	Funzioni linearmente indipendenti	64
3.1.3	Ortogonalità ed indipendenza lineare	65
3.1.4	Rappresentazione vettoriale delle funzioni	66
3.1.5	Basi ortonormali e calcolo dell'immagine di una funzione	67
3.1.6	Cambiamento di Base	68
3.1.7	Procedimento di Gram-Schmidt	68
3.1.8	Basi notevoli: Legendre e Hermite*	71
3.2	Espansione in funzioni ortogonali	72
3.2.1	Teorema di Parseval	73
3.2.2	Serie di Fourier	73
3.2.3	Serie di Fourier per sequenze periodiche	74
3.2.4	Fenomeno di Gibbs	75
3.2.5	Basi complete	77
3.2.6	Teorema del campionamento*	77
3.2.7	Espansioni che mantengono la convoluzione*	80
3.2.8	Esempio: sulla natura di $\delta(t)^*$	80
3.3	Sintesi ed analisi dei segnali	82
3.3.1	Sintesi	82
3.3.2	Analisi	82
3.4	Riassunto sugli spazi di funzioni	84
4	Segnali e sistemi a tempo discreto	86
4.1	Segnali a tempo discreto	86
4.2	Trasformata di Fourier per segnali a tempo discreto	87
4.2.1	Esempi di trasformate di segnali a tempo discreto	87
4.2.2	Trasformata di Fourier di un segnale campionato	89
4.3	Sistemi a tempo discreto	90
4.3.1	Linearità e tempo invarianza	90
4.3.2	Risposta all'impulso	91
4.3.3	Stabilità	92
4.3.4	Funzione di trasferimento e trasformata z	93
4.3.5	Funzione di trasferimento e trasformata di Fourier	94
4.4	Simulazione discreta di sistemi continui	95
4.4.1	Teorema del simulatore	96
4.5	Proprietà delle sommatorie	97
4.5.1	Manipolazione degli indici delle sommatorie infinite	97
4.5.2	Rappresentazione delle sequenze periodiche mediante sommatorie	98
4.6	Trasformata Discreta di Fourier (DFT)	99
4.6.1	Proprietà della DFT	100
4.6.2	Uso della DFT per il calcolo di convoluzioni discrete*	101
4.6.3	Analisi dei segnali mediante DFT	102

5 Teoria della probabilità	106
5.1 Definizioni di probabilità	108
5.1.1 Definizione classica	108
5.1.2 Definizione basata sulla frequenza relativa	108
5.1.3 Definizione assiomatica	108
5.1.4 Interpretazione soggettiva	110
5.2 Definizione di esperimento	110
5.3 Causalità e casualità	111
5.4 Probabilità condizionata	112
5.5 Teoremi della probabilità totale e di Bayes	113
5.5.1 Esempio: BSC	114
5.6 Eventi indipendenti	115
5.6.1 Indipendenza di tre eventi	116
5.7 Esperimenti con infiniti risultati	117
5.7.1 Esempio: treni	117
5.8 Esperimenti congiunti	118
5.8.1 Prodotto cartesiano di due insiemi	119
5.8.2 Prodotto cartesiano di due esperimenti	119
5.8.3 Prodotto cartesiano di N esperimenti	120
5.9 Prove ripetute	121
5.9.1 Prove di Bernoulli	121
5.9.2 Analisi di $p_n(k) = \binom{n}{k} p^k q^{n-k}$	122
5.10 Formula di approssimazione di Gauss	123
5.10.1 Teorema di DeMoivre–Laplace	124
5.10.2 Valutazione approssimata di $P\{k_1 \leq k \leq k_2\}$	124
5.11 La legge dei grandi numeri	126
5.12 Teorema di Poisson	126
5.13 Punti casuali di Poisson	127
5.14 APPENDICE A: valori di erf, G, Q	128
5.15 APPENDICE B: il gioco del lotto	130
5.16 APPENDICE C: simboli della teoria degli insiemi	132
6 Variabili aleatorie	133
6.1 Definizione di variabile aleatoria	133
6.2 Funzione di distribuzione di probabilità di una v.a.	134
6.2.1 Proprietà della funzione di distribuzione di probabilità	134
6.3 Funzione di densità di probabilità di una v.a.	135
6.3.1 Proprietà della funzione di densità di probabilità	136
6.4 Media e varianza	137
6.4.1 Momenti*	138
6.5 V.a. a distribuzione o densità notevoli	138
6.5.1 Teorema di esistenza	138
6.5.2 Uniforme	139
6.5.3 Normale o Gaussiana	139
6.5.4 Binomiale	140
6.5.5 Poisson	141
6.5.6 Rayleigh	141
6.6 Diseguaglianza di Tchebycheff	142
6.7 Distribuzioni e densità di probabilità condizionate	142
6.8 Probabilità totale e teorema di Bayes	144
6.9 Funzioni di una variabile aleatoria	145
6.9.1 Determinazione della distribuzione di $\mathbf{y} = g(\mathbf{x})$	145
6.9.2 Determinazione della densità di $\mathbf{y} = g(\mathbf{x})$	150
6.9.3 Problema inverso	152
6.9.4 Media di $\mathbf{y} = g(\mathbf{x})$	153
6.10 Due variabili aleatorie	154
6.10.1 Distribuzione e densità congiunte	154
6.10.2 Funzioni di due variabili aleatorie	155
6.10.3 Covarianza e correlazione	157
6.10.4 Distribuzioni condizionate	158

6.10.5	Medie condizionate	160
6.10.6	Stima di v.a.	161
6.11	Teorema del limite centrale	163
7	Processi aleatori	165
7.1	Definizioni	165
7.2	Statistiche dei processi aleatori	166
7.2.1	Distribuzioni e densità	166
7.2.2	Media	167
7.2.3	Autocorrelazione e Autocovarianza	167
7.3	Esempi	168
7.3.1	Processo sinusoidale	168
7.3.2	Processo di Poisson	169
7.3.3	Processo telegрафico casuale	170
7.3.4	Rumore bianco	171
7.3.5	Processo gaussiano	171
7.4	Due processi aleatori	171
7.5	Processi stazionari	172
7.5.1	Processi ergodici	173
7.5.2	Rumore bianco stazionario	173
7.6	Sistemi con ingresso aleatorio	174
7.6.1	Sistemi senza memoria	174
7.6.2	Sistemi lineari (LTI)	175
7.7	Densità spettrale di potenza	176
7.7.1	Rumore bianco	177
7.7.2	Trasformata di Hilbert	177
7.7.3	Teorema di Wiener - Khinchin	178
7.7.4	Proprietà	180
7.8	Processo digitale casuale	180
7.9	P.a. combinazione di altri p.a.	181
7.9.1	Sovrapposizione	182
7.9.2	Modulazione	182
7.10	Rumore di quantizzazione	183
8	Rumore e Modulazioni	185
8.1	Guadagno, attenuazione e decibel	185
8.2	Rumore	187
8.2.1	Rumore Termico	187
8.2.2	Rumore bianco e rumore filtrato	188
8.2.3	Rumore additivo e $(S/N)_D$	189
8.2.4	Trasmissione analogica in banda base	190
8.2.5	Ripetitori analogici	192
8.3	Rappresentazione di segnali passa-banda	193
8.3.1	Scomposizione di Rice del rumore gaussiano passa banda	194
8.4	Modulazioni analogiche	195
8.5	Modulazioni Lineari	196
8.5.1	Modulazione DSB	197
8.5.2	Modulazione AM	201
8.5.3	Modulazione SSB	206
8.5.4	Modulazione VSB	209
8.6	Modulazioni Angolari	210
8.6.1	PM a banda stretta - PMNB	211
8.6.2	Spettro della PM/FM tonale*	212
8.6.3	Osservazioni sullo spettro della PM/FM tonale	214
8.6.4	Larghezza di banda per FM e PM	214
8.6.5	Prestazioni dei sistemi di modulazione angolare	216
8.6.6	Prestazioni della PM	217
8.6.7	Prestazioni della FM	218
8.6.8	Preenfasi e deenfasi nella FM *	219
8.6.9	FM stereo e RDS	220

8.6.10 Sintesi sulle modulazioni angolari	222
9 Teoria della ricezione ottima	223
9.1 Definizione di sistema di trasmissione discreto	223
9.2 Canali vettoriali	225
9.2.1 Rumore additivo gaussiano	226
9.3 Trasformazione in canale vettoriale	226
9.3.1 Rumore additivo gaussiano	227
9.3.2 Teorema dell'irrilevanza	228
9.4 Ricevitore ottimo per canali a forme d'onda	229
9.5 Calcolo delle prestazioni di particolari insiemi di segnali	230
9.5.1 Segnali binari	232
9.5.2 Segnali posti ai vertici di un ipercubo	234
9.5.3 Segnali ortogonali	235
9.5.4 Segnali "simplex"	236
9.5.5 Segnali biortogonali	237
10 PAM con rumore	238
10.1 PAM con rumore	239
10.1.1 Probabilità di errore per il caso binario	239
10.1.2 Probabilità di errore per il caso di PAM a M livelli	241
A Formule utili	243
B Nomi illustri	245
C Soluzione agli esercizi proposti	248
C.1 Soluzione dell'esercizio 1.1	248
C.2 Soluzione dell'esercizio 1.5	248
C.3 Soluzione dell'esercizio 7.7	249
C.4 Soluzione dell'esercizio 7.8	249
C.5 Soluzione dell'esercizio su convoluzione	250
C.6 Ancora Lotto	251

Prefazione

Il corso per il quale queste dispense sono state scritte è **Comunicazioni Elettriche**, corso al secondo anno della laurea triennale in Ingegneria Informatica e Bioingegneria dell'Università di Genova. Nella stesura si è dovuto affrontare il problema di trattare argomenti le cui basi sono oggetto di altri corsi. Ove possibile ed opportuno si è scelto di ripetere i concetti fondamentali che lo studente dovrebbe già conoscere, anche se a volte in forma concisa.
I testi consigliati servono come indicazione per eventuali approfondimenti e, a meno di casi particolari, lo studente non è tenuto alla loro consultazione.

Prerequisiti

È **necessario** che lo studente-lettore inizi lo studio disponendo almeno delle conoscenze seguenti.

- **Trigonometria.** Definizione, significato e proprietà (periodicità, pari o dispari...) di seno, coseno, tangente e inverse. Familiarità con la misura degli angoli in gradi ed in radianti. Conoscenza (a memoria) dei valori di seno, coseno e tangente almeno per gli angoli (in radianti): $0, \pi/6, \pi/4, \pi/3, \pi/2$ e corrispondenti negli altri tre quadranti. Conoscenza delle principali formule trigonometriche, alcune delle quali sono riportate in appendice A. Capacità di risolvere semplici equazioni trigonometriche.
- **Integrali.** Definizione e significato degli integrali indefiniti e definiti. Capacità di risolvere per via grafica semplici integrali definiti (ad esempio di funzioni costanti o lineari) e capacità di calcolare integrali definiti delle funzioni più comuni (polinomi di grado qualsiasi, funzioni trigonometriche). Cambio di variabile. Integrali doppi (di superficie). Significato dell'integrazione definita con limiti fissi o variabili: ad esempio è necessario sapere che (se a e b sono numeri) $\int_a^b x(t) dt$ è un numero, $\int_a^b x(s, t) dt$ è una funzione di s , $\int_a^\tau x(s, t) dt$ è una funzione di s e τ . Bisogna saper usare in modo corretto la notazione matematica: non vorrei più vedere cose come $\int_a^t x(t) dt$ o $\int_a^b x(t) dx$.
- **Derivate.** Definizione e significato di derivata. Legame fra derivata e tangente geometrica. Uso della derivata per lo studio di funzioni: legame fra segno della derivata e andamento della funzione (crescente, decrescente, costante). Conoscenza della derivata delle funzioni più comuni (polinomi, funzioni trigonometriche, funzioni esponenziali e logaritmiche).
- **Logaritmi.** Definizione, significato e proprietà dei logaritmi. Cambio di base. Risoluzione di semplici equazioni logaritmiche.
- **Numeri complessi.** Significato, rappresentazione e proprietà dei numeri complessi. Formule di Eulero (richiamate in appendice A) e loro significato.
- **Limiti.** Definizione e significato. Concetto di esistenza di un limite. Calcolo dei limiti delle funzioni più comuni. Risoluzione delle forme indeterminate ($0/0, \infty/\infty$).
- **Serie numeriche.** Legami fra integrali e sommatorie. Serie infinite e convergenza. Serie notevoli.

Capitolo 1

Segnali e sistemi

1.1 Segnali

Un segnale può essere definito come una funzione od una grandezza che **contiene informazione**, in generale riguardo allo stato o al comportamento di un sistema fisico. Anche se i segnali possono essere rappresentati in molti modi, l'informazione è sempre contenuta nelle variazioni di una o più grandezze in qualche dominio. Ad esempio, il segnale può essere costituito dall'insieme delle variazioni di una grandezza nel tempo o nello spazio.

Matematicamente i segnali sono rappresentati come funzioni di una o più variabili indipendenti. Ad esempio la tensione in uscita da un generatore di forme d'onda può essere rappresentata come una funzione del tempo $s(t) = A \cos(2\pi f_0 t + \phi_0)$ ed una fotografia in bianco e nero come una luminosità funzione di due variabili spaziali $I(x, y)$. La maggior parte dei segnali trattati avranno il tempo come variabile indipendente.

1.1.1 Classificazione dei segnali

Segnali a TEMPO CONTINUO / a TEMPO DISCRETO I primi sono i segnali definiti su di un insieme temporale continuo e sono quindi rappresentati da funzioni di variabile continua. I segnali a tempo discreto sono definiti su di un insieme discreto di tempi. Al di fuori degli istanti sui quali è definito un segnale a tempo discreto *non è definito* (attenzione, è errato dire che vale zero). Un segnale a tempo discreto è quindi rappresentabile come una sequenza di numeri.

Segnali ad AMPIEZZA CONTINUA / ad AMPIEZZA DISCRETA I primi sono i segnali che possono assumere un qualsiasi valore compreso in un certo intervallo (anche infinito), mentre i secondi assumono solo un numero finito di valori in un insieme discreto predefinito.

I segnali a tempo continuo ed ampiezza continua sono chiamati anche segnali analogici.

I segnali a tempo discreto ed ampiezza discreta sono detti numerici o anche, con un brutto neologismo, digitali. La classe dei segnali numerici ha assunto negli ultimi decenni una importanza notevole: solo tali segnali possono infatti essere elaborati dai calcolatori elettronici (sia di uso generale che specializzati per le telecomunicazioni). È evidente come sia più semplice elaborare i segnali via *software* che costruire dispositivi *hardware* dedicati. Questo è uno dei motivi per i quali tali segnali hanno tanta diffusione nel mondo delle telecomunicazioni. Un altro motivo, almeno altrettanto importante, è che i segnali numerici hanno la capacità di rappresentare informazioni di qualsiasi tipo (ad esempio voce, dati, immagini, video): questo permette di progettare sistemi di comunicazione in grado di trasmettere segnali diversi, anche etereogenei, in modo flessibile.

Segnali MULTICANALE In alcune applicazioni vengono generati segnali da sensori multipli.

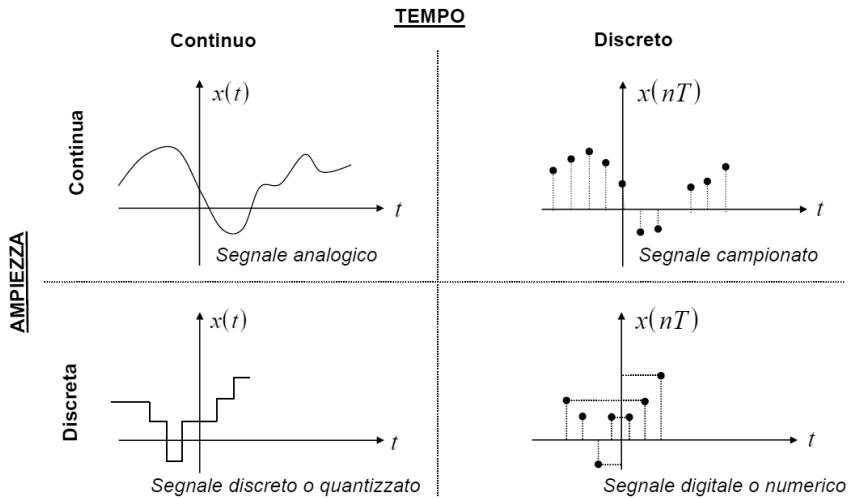
Ad esempio un elettroencefalogramma registrato mediante dodici sensori è composto da dodici segnali analogici: può quindi essere rappresentato in modo vettoriale mediante:

$$S_{12}(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ \dots \\ s_{12}(t) \end{bmatrix}$$

Una sequenza video a colori è composta dalle tre immagini nei colori RGB (rosso, verde e blu) variabili nel tempo.

Quest'ultimo è un esempio di segnale multicanale e multidimensionale.

$$I(x, y, t) = \begin{bmatrix} I_R(x, y, t) \\ I_G(x, y, t) \\ I_B(x, y, t) \end{bmatrix}$$



Segnali CERTI (o DETERMINISTICI) / ALEATORI (o CASUALI) L’analisi e l’elaborazione dei segnali richiedono la disponibilità di una descrizione matematica dei segnali. Questa descrizione, spesso indicata come modello del segnale, porta ad una ulteriore classificazione.

Ogni segnale che possa essere descritto in modo univoco per mezzo di una espressione matematica esplicita, di una tabella di dati o di una regola ben definita è detto certo o deterministico. Questi termini sono usati per indicare che tutti i valori assunti dal segnale, passati, presenti e futuri, sono noti esattamente, senza alcuna incertezza.

Esistono tuttavia segnali che o non possono essere descritti nei modi indicati in precedenza o tali per cui tale descrizione è troppo complicata per essere utilizzabile. La mancanza di un modello esatto implica che tali segnali evolvano nel tempo in un modo che non può essere noto esattamente in anticipo. Questo tipo di segnali è detto aleatorio o casuale. Esempi di segnali aleatori sono: il mormorio di un ruscello, le onde sismiche prodotte da un terremoto, il fruscio di un sistema di amplificazione, il rumore termico. Anche i segnali che viaggiano su di un sistema di trasmissione non possono, ovviamente, essere noti in anticipo e vanno quindi modellati come segnali aleatori, esattamente come i disturbi (rumore) che interferiranno con essi.

I segnali aleatori non possono quindi essere studiati in termini di valori assunti nel tempo, ma ha senso ed è in genere molto utile studiarli in termini statistici. Ciò viene realizzato mediante la teoria della probabilità ed i processi aleatori, come vedremo più avanti nel corso. Da notare come la classificazione di segnali appartenenti al mondo reale in certi o aleatori non sia sempre facile o chiara. A volte entrambe le classificazioni portano a risultati utilizzabili, altre volte sbagliare la classificazione può condurre a conclusioni errate.

1.1.2 Segnali fisicamente realizzabili

I segnali *fisicamente realizzabili*, e cioè osservabili e misurabili in laboratorio, devono soddisfare le seguenti condizioni ([Cou02]):

- assumono valori non nulli su di un intervallo temporale finito;
- possiedono uno spettro a valori non nulli su di un intervallo frequenziale finito;
- sono funzioni continue del tempo;
- hanno valore di picco finito;
- sono a valori reali.

La terza condizione è in realtà conseguenza della seconda, come vedremo nel seguito. Da notare che spesso si usano modelli matematici che violano una o più di tali condizioni (ad esempio, la classica onda quadra), per semplificare l’analisi.

1.1.3 Media, energia e potenza

Nell’analisi dei segnali vi sono alcuni valori di sintesi particolarmente utili.

Si definisce **valor medio** temporale di un segnale $s(t)$ il valore:

$$\overline{s(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t) dt \quad (1.1)$$

Si definisce **energia** di un segnale $s(t)$, se esiste, il valore:

$$E_s = \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (1.2)$$

E_s è ovviamente ≥ 0 . Se $0 < E_s < \infty$, $s(t)$ è un segnale ad energia finita (o anche un segnale di energia). Affinchè $s(t)$ sia ad energia finita occorre che tenda a zero, per $t \rightarrow \infty$, almeno come $1/\sqrt{t}$. Se $s(t)$ è nullo al di fuori dell'intervallo (a, b) , ovvero ha durata limitata, allora $E_s = \int_a^b |s(t)|^2 dt < \infty$, almeno per i segnali fisicamente realizzabili. Se $x(t)$ è la tensione applicata ai capi di una resistenza da 1Ω , allora $E_s = \int_a^b |x(t)|^2 dt$ è l'energia dissipata sulla resistenza nell'intervallo (a, b) : questo giustifica il nome.

Si definisce **potenza media** di un segnale $s(t)$, se esiste, la quantità:

$$P_s = \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} |s(t)|^2 dt \quad (1.3)$$

P_s è ovviamente ≥ 0 . Se P_s esiste ed è $< \infty$, $s(t)$ è un segnale a potenza finita (o anche un segnale di potenza). Un segnale ad energia finita ha potenza media nulla. Analogamente un segnale a potenza finita non ha energia finita.

1.1.4 Segnali periodici

Un segnale $s(t)$ è periodico di periodo T se $s(t) = s(t+T)$ per qualsiasi valore di t , dove T è il più piccolo numero positivo che soddisfa la definizione.

Un segnale periodico non ha energia finita (a meno che non abbia energia nulla).

La potenza media di un segnale periodico $s(t)$, se esiste, si può calcolare su di un qualsiasi intervallo di durata pari al periodo T :

$$P_s = \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} |s(t)|^2 dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |s(t)|^2 dt = \frac{1}{T} \int_a^{a+T} |s(t)|^2 dt \quad (1.4)$$

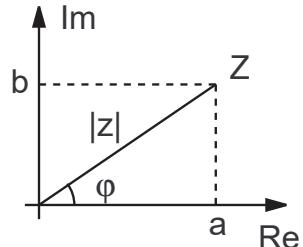
1.2 Numeri complessi e fasori

Un numero complesso z può essere rappresentato in due modi:

$$\begin{aligned} z &= a + jb && a \text{ parte reale, } b \text{ parte immaginaria} \\ z &= |z|e^{j\varphi} && |z| \text{ modulo, } \varphi \text{ fase.} \end{aligned}$$

Le relazioni che legano le due rappresentazioni sono:

$$\begin{aligned} |z| &= \sqrt{a^2 + b^2} & \varphi &= \arctan(\frac{b}{a}) \\ a &= |z| \cos \varphi & b &= |z| \sin \varphi \end{aligned}$$



L'arcotangente deve tener conto dei segni di a e b in modo che φ abbia valori in $[0, 2\pi]$.

Si definisce z complesso coniugato il numero $z^* = a - jb = |z|e^{-j\varphi}$. Ovviamente $z \cdot z^* = |z|^2$.

Da notare come le due rappresentazioni dei numeri complessi, parte reale e parte immaginaria o modulo e fase, siano del tutto equivalenti per quanto riguarda il numero rappresentato, ma non lo siano per la facilità nell'eseguire le operazioni matematiche: con la prima rappresentazione è facile eseguire somme e sottrazioni, con la seconda moltiplicazioni, divisioni ed elevamento a potenza.

Formule di Eulero

Dovrebbero essere ben note allo studente. Esse sono:

$$e^{\pm j\varphi} = \cos \varphi \pm j \sin \varphi \quad \cos \varphi = \frac{e^{j\varphi} + e^{-j\varphi}}{2} \quad \sin \varphi = \frac{e^{j\varphi} - e^{-j\varphi}}{2j} \quad (1.5)$$

Fasori

Una semplice oscillazione armonica è descritta matematicamente dal segnale sinusoidale a tempo continuo

$$v(t) = A \cos(\omega_0 t + \theta) \quad -\infty < t < \infty$$

Questo segnale è completamente descritto da tre parametri:

A che è il valore di picco o ampiezza della sinusoide;
 θ che è l'angolo di fase. Indica che il picco del coseno è spostato dall'origine all'istante $t = -\theta/\omega_0$;
 ω_0 che è la pulsazione o frequenza in radianti al secondo (rad/s).

Il segnale $v(t)$ è periodico di periodo $T_0 = 2\pi/\omega_0$. L'inverso del periodo è la frequenza in Hz, o cicli per secondo:

$$f_0 = \frac{1}{T_0} = \frac{\omega_0}{2\pi} \quad \omega_0 = 2\pi f_0$$

In termini di f_0 , $v(t)$ si può esprimere come $v(t) = A \cos(2\pi f_0 t + \theta)$. Utilizzando le formule di Eulero si ottiene:

$$v(t) = A \cos(\omega_0 t + \varphi) = \frac{A}{2} e^{j(\omega_0 t + \varphi)} + \frac{A}{2} e^{-j(\omega_0 t + \varphi)} \quad (1.6)$$

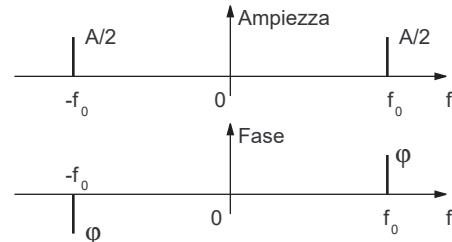
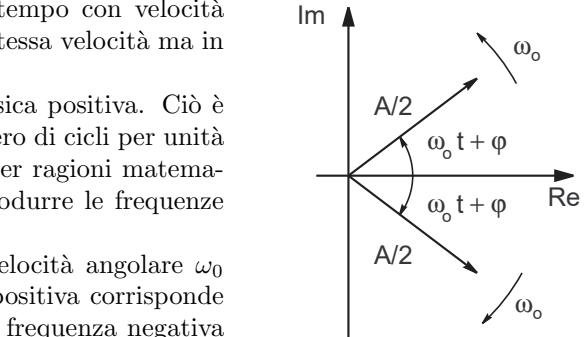
Ciascun termine nella 1.6 può essere interpretato come un vettore nel piano complesso (**fasore**) che ruota nel tempo con velocità angolare ω_0 rad/s. I due vettori ruotano alla stessa velocità ma in direzione opposta.

Per definizione la frequenza è una quantità fisica positiva. Ciò è ovvio se interpretiamo la frequenza come numero di cicli per unità di tempo di un segnale periodico. Tuttavia, per ragioni matematiche, è opportuno (e molto conveniente) introdurre le frequenze negative.

I due fasori ruotano in senso opposto, con velocità angolare ω_0 rad/s, pari a f_0 Hz. Siccome una frequenza positiva corrisponde ad un moto rotatorio in senso antiorario, una frequenza negativa corrisponderà semplicemente ad un moto in senso orario.

La somma dei due fasori giace sull'asse reale, la sua ascissa nel tempo è evidentemente $v(t)$. Ciascun fasore è specificato da tre parametri: ampiezza (modulo), fase e frequenza.

Un fasore può quindi essere descritto nel dominio delle frequenze associando la corrispondente ampiezza e fase alla frequenza f_0 . I due fasori visti possono perciò essere rappresentati anche come in figura:



NOTE

1. L'angolo di fase viene misurato rispetto al coseno 0, in modo equivalente, rispetto all'asse reale positivo del piano complesso (dove ruotano i fasori). Quindi, ad esempio, i seni verranno convertiti in coseni mediante l'identità

$$\sin(\omega t) = \cos(\omega t - 90^\circ) = \cos(\omega t - \frac{\pi}{2})$$

2. L'ampiezza è sempre una quantità positiva. Se compare un segno negativo deve essere "assorbito" dalla fase:

$$-A \cos(\omega t) = A \cos(\omega t \pm 180^\circ) = A \cos(\omega t \pm \pi)$$

Ovviamente -180° e $+180^\circ$ sono lo stesso angolo di fase.

3. Gli angoli di fase possono essere espressi sia in gradi che in radianti. In $(\omega t \pm 180^\circ)$ ωt è in radianti, mentre 180° è in gradi, senza che nascano ambiguità. Analogamente saranno usate sia le frequenze in Hz (cicli al secondo), indicate dalla lettera f , che le pulsazioni ω (in radianti al secondo).

Esercizio 1.1 Rappresentare con un diagramma ampiezza-fase il segnale

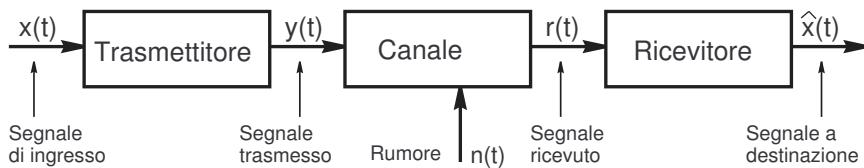
$$w(t) = 7 + 10 \cos(2\pi 20t + 120^\circ) + 4 \cos(2\pi 60t - 90^\circ)$$

Suggerimento:

$$[w(t) = 7 \cos(2\pi 0t) + 10 \cos(2\pi 20t + 120^\circ) + 4 \cos(2\pi 60t - 90^\circ)]$$

1.3 Sistemi di comunicazione

Il diagramma a blocchi seguente presenta l'architettura generale di un sistema di comunicazione.



L'informazione è contenuta nel segnale $x(t)$ in ingresso al sistema. Il segnale $x(t)$ è di tipo elettrico e la sua ampiezza rappresenta la tensione o la corrente (di solito la tensione, almeno in queste dispense). Il messaggio in uscita al ricevitore, indicato da $\hat{x}(t)$, è quello che arriva a destinazione e si auspica contenga ancora l'informazione che si vuole trasmettere. Il segnale a destinazione $\hat{x}(t)$ è diverso dal segnale trasmesso $x(t)$ per effetto dei disturbi presenti nel canale di trasmissione (che sono fisicamente ineliminabili) o a causa di altri fattori di degradazione, ad esempio filtri indesiderati o non linearità. In un sistema di comunicazione, indipendentemente dalla particolare realizzazione, sono sempre presenti tre sottosistemi fondamentali:

Trasmettitore: ha lo scopo di modificare il segnale da trasmettere $x(t)$ in forma adatta a viaggiare sul mezzo trasmissivo (canale) resistendo, per quanto possibile, agli effetti dei disturbi.

Canale: è il mezzo che propaga il segnale fino al ricevitore. Ogni canale reale attenua il segnale trasmesso $y(t)$, la cui ampiezza decresce progressivamente all'allontanarsi dal trasmettitore. Inoltre ogni canale reale introduce un disturbo o rumore che modifica il segnale che viaggia. Il segnale $r(t)$ che giunge al ricevitore è quindi una copia attenuata e degradata di quello che entra nel canale. I canali di trasmissione si dividono in due grandi categorie: via cavo (in inglese *wired*) e senza fili (*wireless*). Esempi di trasmissione via cavo sono le linee telefoniche su doppino in rame, i cavi coassiali, le guide d'onda, le fibre ottiche. Canali senza fili sono l'aria, il vuoto interstellare, il mare. Le regole generali per realizzare un sistema di comunicazione sono sempre le stesse, ma il tipo di canale disponibile vincola molto i parametri di progetto.

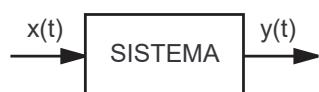
Ricevitore: a partire dal segnale $r(t)$ ricostruisce una copia la più fedele possibile del segnale $x(t)$.

Lo scopo del progetto di sistema è quello di minimizzare la degradazione dell'informazione rispettando un certo numero di vincoli di progetto, tipicamente la potenza in trasmissione, la banda occupata e il costo.

La progettazione e l'analisi di un sistema di trasmissione richiedono l'unione di diverse competenze, le principali sono la teoria dei sistemi, l'analisi in frequenza, la progettazione di filtri analogici o numerici, la teoria dei fenomeni aleatori.

1.4 Sistemi a tempo continuo

Un *sistema* è un dispositivo fisico che accetta in ingresso una funzione del tempo $x(t)$ la cui azione, all'interno del dispositivo, causa un funzione del tempo in uscita $y(t)$. Questo comportamento è simboleggiato dal diagramma a blocchi in figura.



In altre parole un sistema è una regola per assegnare ad una qualsiasi funzione $x(t)$ un'altra funzione $y(t)$ secondo

$$y(t) = T[x(t)]$$

Dove la lettera T indica una generica trasformazione. I sistemi si dividono in diverse classi, a seconda delle loro caratteristiche. Una classe particolarmente importante è la classe dei sistemi LINEARI, per i quali la rappresentazione ed il calcolo dell'uscita dato l'ingresso sono semplici. Spesso, per indicare un sistema lineare, la lettera T viene sostituita da L .

Linearità Un sistema L è lineare se vale il principio di sovrapposizione:

$$L[ax_1(t) + bx_2(t)] = aL[x_1(t)] + bL[x_2(t)] \quad (1.7)$$

per ogni $a, b, x_1(t)$ e $x_2(t)$.

Una seconda divisione dei sistemi riguarda la invarianza temporale. Tali sistemi si comportano allo stesso modo se li considero adesso o fra un'ora. È abbastanza evidente come un sistema che si comporti in modo diverso a seconda dell'istante nel quale lo osservo sia di difficile gestione.

Tempo invarianza Un sistema è tempo invariante (o stazionario) se

$$T[x(t - t_0)] = y(t - t_0) \quad (1.8)$$

per ogni t_0 reale.

La linearità e la tempo invarianza sono proprietà *indipendenti*: un sistema può essere o meno tempo invariante sia che sia lineare sia che non lo sia.

La sottoclasse di maggior interesse è quella dei sistemi sia lineari che tempo invarianti: questi sistemi si indicano con l'acronimo **LTI**.

Esempi

- (a) $y(t) = |x(t)|$ (rettificatore): non lineare, tempo invariante.
- (b) $y(t) = t^2 x(t)$: lineare, tempo variante.
- (c) $y(t) = x(t - a)$ (linea di ritardo o di anticipo): lineare, tempo invariante.
- (d) $y(t) = ax(t)$ (amplificatore se $a > 1$, attenuatore se $a < 1$): lineare, tempo invariante.

Nota: se un sistema è lineare e $x(t) \equiv 0$, allora $y(t) \equiv 0$.

Una ulteriore classificazione dei sistemi riguarda la causalità. Il parola semplici, un sistema è causale se "non risponde in anticipo", ovvero se non prevede il futuro. Questa classificazione ha senso soprattutto per sistemi tempo invarianti, visto che un sistema tempo variante può anche essere causale certe volte e non causale altre.

Causalità Una funzione $x(t)$ è causale se $x(t) = 0$ per $t < 0$. Un *sistema* è causale se ad un ingresso causale corrisponde una uscita causale. Un sistema causale tempo invariante ha la proprietà seguente:

$$\text{Se } x(t) = 0 \text{ per } t \leq t_0 \text{ allora } y(t) = 0 \text{ per } t \leq t_0 \quad (1.9)$$

Corollario: se un sistema è causale e $x_1(t) = x_2(t)$ per $t \leq t_0$ allora $y_1(t) = y_2(t)$ per $t \leq t_0$.

In altri termini: un sistema è causale se la sua risposta all'istante t_0 dipende solo dai valori dell'ingresso per $t \leq t_0$.

Tutti i sistemi fisici sono ovviamente causali se t è il tempo, tuttavia: (i) se la variabile indipendente non è il tempo ma, ad esempio, lo spazio, esistono sistemi fisici non causali; (ii) un sistema non causale nel tempo può essere realizzato fisicamente purchè si accetti un ritardo nell'uscita ed eventualmente un po' di approssimazione (sennò i filtri passa-basso non esisterebbero).

Sistema reale Un sistema è detto reale se la risposta ad una qualsiasi funzione in ingresso reale è una funzione reale. Da tale definizione e se il sistema è lineare segue che, se $x_1(t)$ e $x_2(t)$ sono due funzioni reali e

$$L[x_1(t) + jx_2(t)] = y_1(t) + jy_2(t) \quad \text{allora} \quad y_1(t) = L[x_1(t)] \quad y_2(t) = L[x_2(t)]$$

Stabilità alla BIBO Secondo questa definizione un sistema è stabile se *ogni* ingresso limitato (in ampiezza) produce una uscita limitata.

L'acronimo BIBO viene dall'inglese: Bounded Input Bounded Output. Il concetto di stabilità è fondamentale nello studio dei sistemi, ma è principalmente di interesse in campi diversi dalle comunicazioni elettriche (ad esempio analisi dei sistemi e controlli): altri corsi approfondiranno l'argomento, anche se qualcosa ancora vedremo...

1.5 Sistemi LTI

Limitiamo, per ora, la trattazione ai sistemi LTI che, come già detto, sono quelli di maggior interesse. Nell'analisi dei sistemi è necessario avere una rappresentazione che permetta di descrivere in forma matematica il sistema. Tale rappresentazione deve essere completa, ovvero fornire tutta l'informazione necessaria a specificare le caratteristiche ed il comportamento del sistema: a partire da essa deve quindi essere possibile realizzare fisicamente il dispositivo. Inoltre deve essere possibile calcolare l'uscita per qualsiasi ingresso.

Esistono tre tipi diversi di descrizione per i sistemi lineari tempo invarianti:

1. la risposta all'impulso;
2. la funzione di trasferimento;
3. le equazioni differenziali di diverso tipo.

Nel seguito tratteremo solo le prime due, che sono più in uso nel campo delle comunicazioni.

1.5.1 L'impulso di Dirac o funzione delta

Prima di parlare di risposta all'impulso è necessario definire cosa è un impulso:

Definizione: la funzione impulso $\delta(t)$ o *delta di Dirac* è definita dalla seguente equazione:

$$\int_{-\infty}^{\infty} w(x)\delta(x) dx = w(0) \quad (1.10)$$

dove $w(x)$ è una qualsiasi funzione continua in $x = 0$.

Dalla 1.10 si ricava facilmente che

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (1.11)$$

e, più in generale,

$$\int_a^b \delta(x) dx = \begin{cases} 1 & \text{se l'intervallo } (a,b) \text{ contiene l'origine} \\ 0 & \text{se l'intervallo } [a,b] \text{ non contiene l'origine} \end{cases} \quad (1.12)$$

La precedente equazione vale ovviamente anche se l'intervallo (a,b) è infinitesimo. Quindi $\delta(t)$ deve essere "infinito" nell'origine e zero altrove (in realtà sono ammessi punti diversi da zero, basta che l'integrale sia nullo).

Da questa definizione si capisce che $\delta(t)$ non è una funzione nel senso ordinario del termine: il suo valore in qualsiasi punto dell'asse reale è ignoto! L'impulso di Dirac appartiene alla classe delle *funzioni generalizzate*, studiate in una parte della matematica detta teoria delle distribuzioni e definite, come $\delta(t)$, non da una equazione ma dal loro effetto.

Proprietà

Proprietà 1	$\delta(at) = \frac{1}{ a } \delta(t)$
Proprietà 2	$\int_{-\infty}^{\infty} w(\tau) \delta(t - \tau) d\tau = \int_{-\infty}^{\infty} w(\tau) \delta(\tau - t) d\tau = w(t)$
Proprietà 3	$\int_{-\infty}^{\infty} w(\tau) \delta^{(n)}(\tau) d\tau = (-1)^n w^{(n)}(0)$

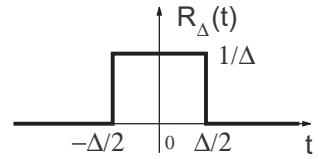
(1.13)

Dove l'apice $^{(n)}$ indica la derivata n -esima ed assumendo che $w^{(n)}$ esista e sia continua in zero. Da notare che queste proprietà NON seguono dalla 1.10: sono assunzioni consistenti con le proprietà formali degli integrali.

Consideriamo la funzione impulso rettangolare di area unitaria mostrata in figura, $R_\Delta(t)$. La sua caratteristica importante è quella di avere area unitaria, ma durata (intervallo diverso da zero) che può essere resa piccola a piacere.

Consideriamo ora il limite di $R_\Delta(t)$ per Δ che tende a zero: l'area rimane unitaria mentre la durata diventa infinitesima, ovvero

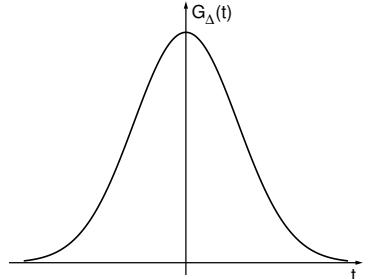
$$\lim_{\Delta \rightarrow 0} R_\Delta(t) = \delta(t) \quad (1.14)$$



Il segno di uguale nella 1.14 deve intendersi nel modo seguente: il fatto che il limite di R_Δ sia un impulso di Dirac non significa che l'impulso di Dirac sia il limite di R_Δ . Ovvero l'uguaglianza significa che il comportamento sotto integrale dei due lati è lo stesso, non che le due funzioni sono le stesse punto per punto.

Esistono altre funzioni che, al limite, si comportano come un impulso di Dirac. Ad esempio la gaussiana centrata nell'origine ad area unitaria in figura, definita da

$$G_\Delta(t) = \frac{1}{\sqrt{2\pi}\Delta} e^{\frac{-t^2}{2\Delta^2}}$$



ha la stessa caratteristica di $R_\Delta(t)$, cioè

$$\lim_{\Delta \rightarrow 0} G_\Delta(t) = \delta(t)$$

Vedremo più avanti altre funzioni che, al limite, si comportano allo stesso modo.

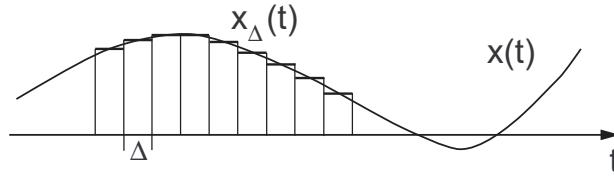
Graficamente l'impulso di Dirac $\delta(t)$ si indica mediante una freccia verticale posizionata in $t = 0$.

1.5.2 Risposta all'impulso

Riprendiamo la funzione impulso rettangolare $R_\Delta(t)$ vista in precedenza. Quando tale funzione viene posta in ingresso ad un sistema LTI si ottiene in uscita una funzione $h_\Delta(t)$ detta risposta del sistema al rettangolo unitario. Ovvimente tale risposta varia al variare di Δ . Consideriamo ora una generica funzione $x(t)$: questa si può approssimare come una sequenza di impulsi rettangolari traslati e scalati, ovvero

$$x(t) \approx x_\Delta(t) = \sum_{k=-\infty}^{\infty} x(k\Delta) \Delta R_\Delta(t - k\Delta) \quad (1.15)$$

Ciascun termine della sommatoria rappresenta un rettangolo centrato in $t = k\Delta$ di altezza $x(k\Delta)\Delta \frac{1}{\Delta} = x(k\Delta)$, ovvero pari al valore della funzione $x(t)$ nell'istante che corrisponde al centro dell'impulso rettangolare traslato. La situazione è illustrata nella figura seguente



Al diminuire di Δ l'approssimazione migliora, tanto che, almeno se $x(t)$ è una funzione continua,

$$x(t) = \lim_{\Delta \rightarrow 0} \sum_{k=-\infty}^{\infty} x(k\Delta) R_\Delta(t - k\Delta) \Delta \quad (1.16)$$

Ricordando quanto visto nei corsi di matematica si può notare come, se $\Delta \rightarrow 0$, la sommatoria diventi, per definizione, un integrale. Inoltre R_Δ diventa l'impulso di Dirac (come visto in precedenza). Quindi

$$x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau \quad (1.17)$$

Consideriamo ora un sistema LTI nel quale entra la funzione $x_\Delta(t)$ definita in 1.15. Chiamiamo $y(t)$ la risposta del sistema quando in ingresso c'è $x(t)$. Siccome il sistema è lineare vale la sovrapposizione

degli effetti, per cui l'uscita relativa a $x_\Delta(t)$ si può scrivere come

$$y_\Delta(t) = \sum_{k=-\infty}^{\infty} x(k\Delta) \Delta h_\Delta(t - k\Delta) \quad (1.18)$$

dove $h_\Delta(t)$ è la risposta del sistema a $R_\Delta(t)$ e si è usato il fatto che il sistema è invariante alle traslazioni. Infatti solo se il sistema è TI è vero che, se $h_\Delta(t)$ è la risposta del sistema a $R_\Delta(t)$, allora $h_\Delta(t - w)$ è la risposta a $R_\Delta(t - w)$.

Se ora facciamo tendere Δ a zero otteniamo:

$$y(t) = \lim_{\Delta \rightarrow 0} \sum_{k=-\infty}^{\infty} x(k\Delta) h_\Delta(t - k\Delta) \Delta \quad (1.19)$$

qui abbiamo sfruttato la 1.16: al tendere di Δ a zero l'ingresso diventa $x(t)$ e l'uscita deve quindi diventare $y(t)$. Scrivendola in forma integrale otteniamo

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau \quad (1.20)$$

La funzione $h(t)$ è ovviamente la risposta del sistema a $R_\Delta(t)$ quando $\Delta \rightarrow 0$, cioè la risposta del sistema all'impulso di Dirac $\delta(t)$:

$$h(t) = L[\delta(t)] \quad (1.21)$$

Se il sistema, oltre che LTI, è anche causale, vale che $h(t) = 0$ per $t < 0$.

Per un generico sistema LTI la risposta all'impulso caratterizza in modo completo il sistema (ovvero non serve conoscere altro per definirlo) e, come indica la 1.20, permette di conoscere l'uscita del sistema LTI per qualsiasi ingresso $x(t)$. L'integrale che compare nella 1.20 si chiama **integrale di convoluzione** o semplicemente convoluzione. È così diffuso da meritarsi una abbreviazione matematica, il simbolo $*$: la relazione che lega l'uscita di un sistema LTI ad un qualsiasi ingresso si può quindi scrivere

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau = \int_{-\infty}^{\infty} x(t - \tau) h(\tau) d\tau \quad (1.22)$$

dove il secondo integrale si ottiene dal primo con un semplice cambio di variabile. Il simbolo $*$ indica quindi entrambi gli integrali, che sono infatti del tutto equivalenti.

Se il sistema è causale si ha che $h(t) = 0$ per $t < 0$, visto che si può assumere $\delta(t) = 0$ per $t < 0$. In questo caso l'uscita si può anche scrivere come

$$y(t) = x(t) * h(t) = \int_{-\infty}^t x(\tau) h(t - \tau) d\tau = \int_0^{\infty} x(t - \tau) h(\tau) d\tau \quad (1.23)$$

Se anche $x(t) = 0$ per $t < 0$ allora anche $y(t) = 0$ per $t < 0$: per $t > 0$ $y(t)$ vale

$$y(t) = x(t) * h(t) = \int_0^t x(\tau) h(t - \tau) d\tau = \int_0^t x(t - \tau) h(\tau) d\tau \quad (1.24)$$

Stabilità Un sistema LTI con risposta all'impulso $h(t)$ è stabile (alla BIBO) se e solo se

$$\int_{-\infty}^{\infty} |h(t)| dt < \infty \quad (1.25)$$

La semplice dimostrazione è omessa (si trova, ad esempio, in [Sak68] pag. 28).

NOTA: l'integrale di convoluzione in 1.22 può anche non esistere, ovvero non fornire valori finiti. Questo può accadere se il sistema non è stabile alla BIBO. L'argomento, che non riguarda questo corso (i sistemi che vedremo sono stabili), verrà approfondito in Teoria dei Sistemi.

1.5.3 L'integrale di convoluzione

L'integrale di convoluzione ha significato anche al di fuori del calcolo dell'uscita di un sistema LTI ed è opportuno dedicargli un po' di spazio.

Date due generiche funzioni $x(t)$ e $y(t)$ si definisce convoluzione di x ed y

$$x(t) * y(t) = \int_{-\infty}^{\infty} x(\tau) y(t - \tau) d\tau = \int_{-\infty}^{\infty} x(t - \tau) y(\tau) d\tau \quad (1.26)$$

Se le funzioni sono "abbastanza semplici" è possibile il calcolo della convoluzione per via grafica. Partendo, ad esempio, da

$$u(t) = x(t) * y(t) = \int_{-\infty}^{\infty} x(\tau) y(t - \tau) d\tau$$

si eseguono i seguenti passaggi:

- si ribalta $y(t)$ attorno all'origine, ottenendo $y(-\tau)$. Si moltiplica per $x(\tau)$ e si integra: il valore trovato è $u(0)$.
- si trasla $y(-\tau)$ in $t_0 \neq 0$ per avere $y(t_0 - \tau)$. Si moltiplica per $x(\tau)$ e si integra: il valore trovato è $u(t_0)$.
- si ripete *per tutti* i valori di t_0 .

Un esempio aiuta a capire il meccanismo.

Esempio 1.2 Siano date due funzioni $x(t)$ e $y(t)$ definite da:

$$x(t) = \begin{cases} A & 0 \leq t \leq a \\ 0 & \text{altrove} \end{cases}$$

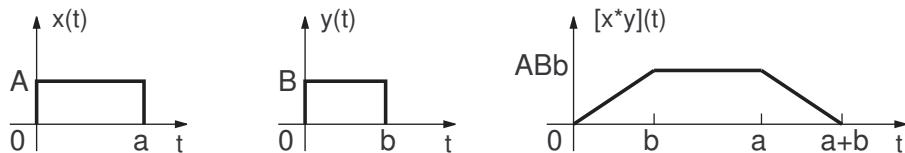
$$y(t) = \begin{cases} B & 0 \leq t \leq b \\ 0 & \text{altrove} \end{cases}$$

assumendo $a > b$. L'integrale di convoluzione è ovviamente

$$x(t) * y(t) = \int_{-\infty}^{\infty} x(\tau) y(t - \tau) d\tau$$

e può essere calcolato a tratti secondo lo schema seguente

$$[x * y] = \begin{cases} t \leq 0 & 0 \\ 0 < t < b & ABt = \int_0^t AB d\tau \\ b \leq t \leq a & ABb = \int_{t-b}^t AB d\tau = AB(t - t + b) \\ a < t < a + b & AB(a + b - t) = \int_{t-b}^a AB d\tau \\ a + b \leq t & 0 \end{cases}$$



La convoluzione fra due rettangoli genera quindi un trapezio avente base maggiore pari alla somma delle durate dei rettangoli, e base minore pari alla differenza. L'altezza del trapezio è pari al prodotto delle altezze dei due rettangoli per la durata del rettangolo più breve. Se i rettangoli sono di pari durata la convoluzione assume la forma di un triangolo.

Il risultato dell'esempio precedente si può generalizzare: la convoluzione di due funzioni qualsiasi diverse da zero rispettivamente in $[t_1, t_2]$ e $[t_3, t_4]$ assume valori diversi da zero in un intervallo di durata massima pari a $|t_2 - t_1| + |t_4 - t_3|$. In altre parole si può dire che la convoluzione di due segnali a supporto compatto ha durata pari alla somma delle durate dei due segnali.

Il calcolo dell'integrale di convoluzione può essere molto difficile, vedremo più avanti come la trasformata di Fourier fornisca un diverso modo di calcolo, spesso più semplice.

1.5.4 Impulsi e convoluzione

Eseguire l'integrale di convoluzione quando una delle due funzione è l'impulso di Dirac porta a risultati particolari, che verranno spesso usati nel seguito.

Dalle equazioni 1.10 e 1.13 si deriva immediatamente che

$$x(t) * \delta(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau = x(t) \quad (1.27)$$

ovvero la convoluzione di una funzione con un impulso centrato nell'origine lascia la funzione inalterata.

La convoluzione con un impulso traslato è

$$x(t) * \delta(t - k) = \int_{-\infty}^{\infty} x(\tau) \delta(t - k - \tau) d\tau = x(t - k) \quad (1.28)$$

risultato che si ottiene facilmente chiamando $t - k = \alpha$ e risolvendo l'integrale come al solito. Quindi la convoluzione di una funzione con un impulso traslato trasla la funzione.

Riassumendo i due importanti risultati ottenuti si ha

$x(t) * \delta(t) = x(t)$	$x(t) * \delta(t - k) = x(t - k)$
---------------------------	-----------------------------------

1.5.5 La funzione di trasferimento

Inviamo in ingresso ad un sistema LTI la funzione esponenziale

$$x(t) = e^{j\omega_0 t}$$

Dalla 1.22 segue che la risposta del sistema è

$$y(t) = \int_{-\infty}^{\infty} e^{j\omega_0(t-\tau)} h(\tau) d\tau = e^{j\omega_0 t} \int_{-\infty}^{\infty} e^{-j\omega_0 \tau} h(\tau) d\tau \quad (1.29)$$

L'ultimo integrale è ovviamente un numero, in generale complesso, che dipende solo da ω_0 e da $h(t)$. Se chiamiamo tale numero $H(\omega_0)$ possiamo scrivere che

$$y(t) = e^{j\omega_0 t} \int_{-\infty}^{\infty} e^{-j\omega_0 \tau} h(\tau) d\tau = e^{j\omega_0 t} H(\omega_0)$$

Quindi anche $y(t)$ è una funzione esponenziale, e precisamente la funzione in ingresso moltiplicata per il valore $H(\omega_0)$ della trasformata di Fourier

$$H(\omega) = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} dt \quad (1.30)$$

della risposta all'impulso $h(t)$. Essendo quest'ultimo risultato vero per ogni ω_0 , concludiamo che

$$L[e^{j\omega t}] = H(\omega) e^{j\omega t} \quad (1.31)$$

Gli esponenziali complessi sono quindi le **autofunzioni** dei sistemi LTI, ovvero quelle funzioni che vengono trasformate dal sistema in se stesse, a meno di un fattore moltiplicativo.

In generale, la funzione di trasferimento di un sistema (in inglese System Function) è una rappresentazione dello stesso in un dominio differente da quello del tempo. A seconda del dominio considerato esistono quindi diverse funzioni di trasferimento. La funzione $H(\omega)$ è la funzione di trasferimento per esponenziali complessi o *risposta in frequenza* del sistema.

L'equazione 1.31 indica anche un modo per determinare la funzione di trasferimento di un sistema LTI incognito.¹ Da notare che se un sistema non è stabile alla BIBO può non essere possibile calcolare $H(\omega)$.

¹Su [Pap77], pag. 20, è indicato come usare un ingresso sinusoidale per determinare, a punti, $H(\omega)$.

1.6 La trasformata di Fourier

Si è visto in sezione 1.2 come una oscillazione sinusoidale possa essere rappresentata mediante ampiezza e fase nel dominio delle frequenze. Per tali segnali il concetto di frequenza è semplice: $f = 1/T$, ovvero la frequenza è l'inverso del periodo.

In molte applicazioni i segnali non sono sinusoidali, e neppure periodici, ma è comunque di grande utilità poter ricavare il loro *contenuto frequenziale*. Questa è la base, ad esempio, per poter inviare più segnali sullo stesso mezzo trasmissivo senza che interferiscano fra di loro, visto che ciascuno di essi occupa bande di frequenza diverse (come nel caso delle trasmissioni radiofoniche o televisive). Il contenuto frequenziale di un segnale si ottiene mediante la *trasformata di Fourier* (a volte indicata come FT).

La notazione $f(t) \leftrightarrow F(f)$ indica che $F(f)$ è la trasformata di Fourier di $f(t)$, cosa che si può anche indicare come $F(f) = \mathcal{F}[f(t)]$. La trasformata di Fourier è definita dall'equazione

$$F(f) = \mathcal{F}[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-j2\pi ft} dt \quad (1.32)$$

assumendo ovviamente che tale integrale esista. La variabile f è la frequenza, che si misura in Hz (Hertz, dal nome del fisico tedesco Heinrich Rudolf Hertz), ovvero 1 / secondo.

La definizione in 1.32 è detta *bilatera*, visto che comprende anche le frequenze negative, come introdotto in sezione 1.2. In generale, sia $F(f)$ che $f(t)$ sono segnali complessi: se $f(t)$ è reale $F(f)$ risulta comunque complessa (salvo casi particolari), visto che $e^{-j2\pi ft}$ è complesso. Siccome $F(f)$ è, in generale, complessa si può esprimere come parte reale e parte immaginaria o come modulo e fase:

$$F(f) = \operatorname{Re}\{F(f)\} + j\operatorname{Im}\{F(f)\} = |F(f)| e^{j\angle F(f)}$$

dove

$$|F(f)| = \sqrt{[\operatorname{Re}\{F(f)\}]^2 + [\operatorname{Im}\{F(f)\}]^2} \quad \angle F(f) = \arctan\left(\frac{\operatorname{Im}\{F(f)\}}{\operatorname{Re}\{F(f)\}}\right)$$

Il modulo della trasformata di Fourier di una funzione, $|F(f)|$, è detto *spettro di ampiezza*: la sua conoscenza permette di stabilire per quali frequenze il segnale $f(t)$ abbia componenti significative. La fase $\angle F(f)$ è ovviamente un angolo, e assume valori in un intervallo di ampiezza 2π , solitamente $[0, 2\pi]$ oppure $[-\pi, \pi]$.

Esempio 1.3 Sia $w(t) = e^{-\alpha t}U(t)$ con $\alpha > 0$, ovvero un impulso esponenziale decrescente.

La funzione $U(t)$ è il gradino unitario, cioè la funzione che vale 1 per $t \geq 0$ e zero altrove.

Quindi $w(t)$ è reale e causale e vale

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

Usando la 1.32 si ottiene

$$W(f) = \mathcal{F}[w(t)] = \int_0^{\infty} e^{-\alpha t} e^{-j2\pi ft} dt = \frac{e^{-(\alpha+j2\pi f)t}}{-(\alpha+j2\pi f)} \Big|_0^{\infty} = \frac{1}{\alpha + j2\pi f} \quad \alpha > 0$$

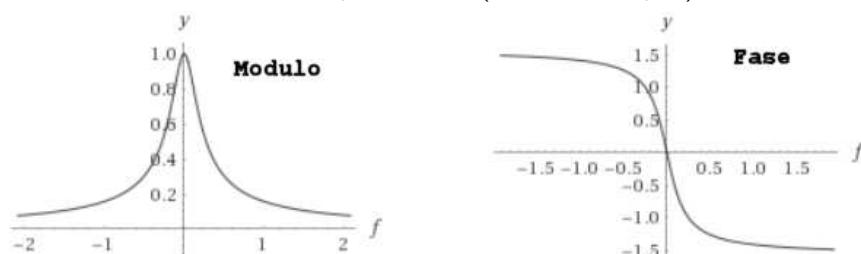
Poichè $W(f)$ è complesso, possiamo esprimere in termini di parte reale e parte immaginaria

$$\operatorname{Re}[W(f)] = \frac{\alpha}{\alpha^2 + (2\pi f)^2} \quad \operatorname{Im}[W(f)] = \frac{-2\pi f}{\alpha^2 + (2\pi f)^2}$$

oppure di modulo e fase:

$$|W(f)| = \sqrt{\frac{1}{\alpha^2 + (2\pi f)^2}} \quad \angle W(f) = -\arctan\left(\frac{2\pi f}{\alpha}\right)$$

L'andamento del modulo e della fase per $\alpha = 1$ è (da *WolframAlpha*)



Trasformata inversa. La trasformata di Fourier è invertibile: la formula per ritornare a $f(t)$ è

$$f(t) = \mathcal{F}^{-1}[F(f)] = \int_{-\infty}^{\infty} F(f) e^{j2\pi ft} df \quad (1.33)$$

Esistenza. Non tutte le funzioni ammettono trasformata di Fourier, per cui le definizioni 1.32 e 1.33 vanno intese "se l'integrale esiste".

Stabilire quali funzioni ammettano trasformata di Fourier è un problema matematicamente complesso. Una condizione sufficiente (ma non necessaria) affinché una funzione la ammetta sono le cosiddette *condizioni di Dirichlet*, ovvero $w(t)$ ammette FT se

- è assolutamente integrabile,

$$\int_{-\infty}^{\infty} |w(t)| dt < \infty$$

- su di un qualsiasi intervallo di lunghezza finita $w(t)$ ha un numero finito di discontinuità con valori limite da destra e sinistra finiti ed ha un numero finito di massimi e minimi.

Una diversa condizione sufficiente (a volte più restrittiva, a volte meno) è chiedere che $w(t)$ sia ad energia finita, cioè

$$\int_{-\infty}^{\infty} |w(t)|^2 dt < \infty \quad (1.34)$$

Tutti i segnali fisicamente realizzabili sono ad energia finita, quindi tutti i segnali che si usano nella pratica hanno trasformata di Fourier.

Vedremo nel seguito esempi di segnali che non soddisfano le due condizioni sopra viste ma hanno trasformata di Fourier (almeno per un ingegnere), come ad esempio $\cos(2\pi t)$.

La FT può essere definita, in modo del tutto equivalente, utilizzando, al posto della variabile frequenza (f in Hz), la variabile pulsazione $\omega = 2\pi f$, che si misura in radianti al secondo. Con un semplice cambio di variabile le 1.32 e 1.33 diventano

$$F(\omega) = F[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (1.35)$$

$$f(t) = F^{-1}[F(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (1.36)$$

Gli studenti sono tenuti ad avere familiarità con entrambe le notazioni: anche per questo motivo nel resto delle dispense verranno usate entrambe.

Legami fra le FT in f e in ω Vediamo che legami ci sono fra le rappresentazioni in frequenza ed in pulsazione della FT. Dalle definizioni si ha che la trasformata del segnale $f(t)$ vale:

$$\begin{aligned} F_f(f) &= \int_{-\infty}^{\infty} f(t) e^{-j2\pi ft} dt \\ F_{\omega}(\omega) &= \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \end{aligned}$$

Se cambio le variabili delle FT ottengo che $F_f(\omega = 2\pi f) = F_{\omega}(\omega)$ e $F_{\omega}[f = \omega/(2\pi)] = F_f(f)$. Quindi esprimere le trasformate in f oppure ω è del tutto equivalente.

Una piccola complicazione nasce dalle formule inverse: se uso la variabile ω devo dividere per 2π , se uso f no. Ciò deriva ovviamente dal fatto che, nelle formule inverse, f ed ω rappresentano la variabile di integrazione: se cambio una nell'altra esce fuori il fattore 2π . Le proprietà che seguono verranno espresse in entrambi i domini, anche se il passaggio da uno all'altro è immediato.

1.6.1 Proprietà

Dalla definizione di trasformata di Fourier discendono importanti proprietà (teoremi).

Linearità La trasformata di Fourier è un operatore lineare: tale proprietà discende direttamente dalla definizione. Quindi

$$\mathcal{F}[a_1 s_1(t) + a_2 s_2(t)] = a_1 \mathcal{F}[s_1(t)] + a_2 \mathcal{F}[s_2(t)] \quad (1.37)$$

FT di un segnale reale Se un segnale $s(t)$ è **reale** la sua FT $S(f)$ ha la proprietà di *simmetria Hermitiana*, ovvero

$$S(-f) = S^*(f) \quad (1.38)$$

Dimostrazione: Dalla 1.32 si ha

$$S(-f) = \int_{-\infty}^{\infty} s(t) e^{j2\pi f t} dt$$

e coniugando la 1.32 si ottiene

$$S^*(f) = \int_{-\infty}^{\infty} s^*(t) e^{j2\pi f t} dt$$

Se $s(t)$ è reale $s(t) = s^*(t)$, quindi $S(-f) = S^*(f)$: **c.v.d.**

La 1.38 significa che $S(f)$ ha *parte reale e modulo pari e parte immaginaria e fase dispari*. Infatti

$$S(-f) = S^*(f) \Rightarrow Re\{S(-f)\} + jIm\{S(-f)\} = Re\{S(f)\} - jIm\{S(f)\}$$

ovvero

$$Re\{S(-f)\} = Re\{S(f)\} \quad Im\{S(-f)\} = -Im\{S(f)\}$$

Scrivendo poi $S(f)$ in forma polare

$$S(f) = |S(f)| e^{j\theta(f)}$$

vale che

$$S(-f) = |S(-f)| e^{j\theta(-f)} \quad S^*(f) = |S(f)| e^{-j\theta(f)}$$

per cui

$$|S(-f)| = |S(f)| \quad \theta(-f) = -\theta(f)$$

Dalla 1.38 segue che la FT di un segnale reale è completamente determinata dai suoi valori per $f \geq 0$: i valori per $f < 0$ si ricavano per simmetria.

FT di un segnale reale pari Se $s(t)$ è sia reale che pari la sua FT $S(f)$ è *pari* e *reale* (ovvero ha parte immaginaria nulla).

Infatti, poichè $e^{-j\omega t} = \cos(\omega t) - j \sin(\omega t)$ e $s(t)$ è reale, la 1.35 può essere scritta come

$$S(\omega) = \int_{-\infty}^{\infty} s(t) [\cos(\omega t) - j \sin(\omega t)] dt = \int_{-\infty}^{\infty} s(t) \cos(\omega t) dt - j \int_{-\infty}^{\infty} s(t) \sin(\omega t) dt \quad (1.39)$$

se $s(t) = s(-t)$, ovvero se è pari, il secondo integrale si annulla (il seno è dispari) e

$$S(\omega) = \int_{-\infty}^{\infty} s(t) \cos(\omega t) dt = 2 \int_0^{\infty} s(t) \cos(\omega t) dt \quad (1.40)$$

Quindi $S(\omega)$ è reale. Deve inoltre essere pari, sia perchè lo abbiamo già dimostrato (vedi 1.38), sia perchè la ω compare solo come argomento del coseno (che è pari).

FT di un segnale reale dispari Se $s(t)$ è sia reale che dispari la sua FT $S(f)$ è *dispari* e *immaginaria pura* (ovvero ha parte reale nulla).

Infatti, se $s(-t) = -s(t)$ la 1.39 diventa

$$S(\omega) = -j \int_{-\infty}^{\infty} s(t) \sin(\omega t) dt = -2j \int_0^{\infty} s(t) \sin(\omega t) dt \quad (1.41)$$

(il meno davanti al secondo integrale è giusto, dato che il prodotto di due funzioni dispari è pari).

Quindi $S(\omega)$ è puramente immaginaria. Deve inoltre essere dispari, sia perchè lo abbiamo già dimostrato (vedi 1.38), sia perchè la ω compare solo come argomento del seno (che è dispari).

Dualità Detta anche *regola di simmetria*. Se $s(t) \leftrightarrow S(\omega)$ allora $S(t) \leftrightarrow 2\pi s(-\omega)$. Dimostrazione: dalla 1.36 segue che

$$2\pi s(-\omega) = \int_{-\infty}^{\infty} S(x) e^{-j\omega x} dx \quad (1.42)$$

In termini di frequenza diventa: se $w(t) \leftrightarrow W(f)$ allora $W(t) \leftrightarrow w(-f)$.

Questa proprietà ha un importante ruolo teorico, ma è anche molto utile nelle applicazioni, in quanto dalla conoscenza di $w(t) \leftrightarrow W(f)$ si può passare automaticamente alla coppia $W(t) \leftrightarrow w(-f)$.

Coniugazione Se $s(t) \leftrightarrow S(\omega)$ allora $s^*(t) \leftrightarrow S^*(-\omega)$. In frequenza è ovviamente lo stesso: $w(t) \leftrightarrow W(f)$ implica $w^*(t) \leftrightarrow W^*(-f)$. La dimostrazione è nuovamente banale (dalla definizione).

Area Se nelle 1.32 e 1.33 si pongono rispettivamente $f = 0$ e $t = 0$ si ottiene:

$$\begin{aligned} S(0) &= \int_{-\infty}^{\infty} s(t) dt = \text{area}(s) \\ s(0) &= \int_{-\infty}^{\infty} S(f) df = \text{area}(S) \end{aligned} \quad (1.43)$$

Quindi, ad esempio, l'area del segnale $s(t)$ coincide col valore in zero della sua FT. Inoltre se un segnale limitato nel tempo ha valor medio (vedi 1.1) nullo la sua FT per $f = 0$ vale zero.

Scalamento Per ogni a reale

$$s(at) \longleftrightarrow \frac{1}{|a|} S\left(\frac{\omega}{a}\right) \quad (1.44)$$

Dimostrazione:

$$\int_{-\infty}^{\infty} s(at)e^{-j\omega t} dt = \frac{1}{|a|} \int_{-\infty}^{\infty} s(x)e^{-j\omega x/a} dx = \frac{1}{|a|} S\left(\frac{\omega}{a}\right)$$

In frequenza

$$w(at) \longleftrightarrow \frac{1}{|a|} W\left(\frac{f}{a}\right) \quad (1.45)$$

Quindi se un segnale si "allarga" nel tempo si "stringe" in frequenza e viceversa.

Traslazione nel tempo Per ogni T_d reale

$$s(t - T_d) \longleftrightarrow e^{-j\omega T_d} S(\omega) \quad (1.46)$$

Dimostrazione:

$$\int_{-\infty}^{\infty} s(t - T_d) e^{-j\omega t} dt = \int_{-\infty}^{\infty} s(x) e^{-j\omega(x+T_d)} dx = e^{-j\omega T_d} S(\omega)$$

In frequenza

$$w(t - T_d) \longleftrightarrow e^{-j2\pi f T_d} W(f) \quad (1.47)$$

Quindi una traslazione nel tempo di un segnale (sia anticipandolo che ritardandolo) altera solo la fase della sua FT, mentre il modulo rimane invariato.

Traslazione in frequenza Per ogni a reale

$$e^{jat} s(t) \longleftrightarrow S(\omega - a) \quad (1.48)$$

Infatti

$$\int_{-\infty}^{\infty} e^{jat} s(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} s(t) e^{-j(\omega-a)t} dt = S(\omega - a)$$

In frequenza:

$$e^{j2\pi at} w(t) \longleftrightarrow W(f - a) \quad (1.49)$$

Modulazione

$$s(t) \cos(\omega_0 t) \longleftrightarrow \frac{1}{2} [S(\omega + \omega_0) + S(\omega - \omega_0)] \quad (1.50)$$

La dimostrazione segue dalla 1.48 visto che

$$\cos(\omega_0 t) = \frac{1}{2} [e^{j\omega_0 t} + e^{-j\omega_0 t}]$$

In frequenza

$$w(t) \cos(2\pi f_c t) \longleftrightarrow \frac{1}{2} [W(f + f_c) + W(f - f_c)] \quad (1.51)$$

e se il coseno è sfasato

$$w(t) \cos(2\pi f_c t + \theta) \longleftrightarrow \frac{1}{2} [e^{-j\theta} W(f + f_c) + e^{j\theta} W(f - f_c)] \quad (1.52)$$

Lo studente è invitato a dimostrare la 1.52 come verifica di apprendimento.

Derivazione Indichiamo la derivata ennesima rispetto al tempo del segnale $s(t)$ con il simbolo $s^{(n)}(t) = \frac{d^n s(t)}{dt^n}$.

$$s^{(n)}(t) \leftrightarrow (j\omega)^n S(\omega) \quad (1.53)$$

La dimostrazione si ottiene derivando n volte la 1.36

$$s^{(n)}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega)^n S(\omega) e^{j\omega t} d\omega$$

In frequenza

$$w^{(n)}(t) \longleftrightarrow (j2\pi f)^n W(f) \quad (1.54)$$

Vale anche la relazione duale

$$(-jt)^n s(t) \longleftrightarrow S^{(n)}(\omega) \quad (1.55)$$

Convoluzione Se $f_1(t) \leftrightarrow F_1(f)$ e $f_2(t) \leftrightarrow F_2(f)$

$$f_1(t) * f_2(t) \longleftrightarrow F_1(f) F_2(f) \quad (1.56)$$

Ovvero la convoluzione di due segnali nel tempo ha FT che è la moltiplicazione delle FT.

Dimostrazione: dalla 1.32,

$$\begin{aligned} \mathcal{F}[f_1(t) * f_2(t)] &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f_1(t-u) f_2(u) du \right] e^{-j2\pi ft} dt \\ &= \int_{-\infty}^{\infty} f_2(u) \int_{-\infty}^{\infty} f_1(t-u) e^{-j2\pi ft} dt du \\ &= \int_{-\infty}^{\infty} f_2(u) F_1(f) e^{-j2\pi fu} du = F_1(f) F_2(f) \end{aligned} \quad (1.57)$$

Nei passaggi si è invertito l'ordine di integrazione e si è usata la proprietà di traslazione nel tempo. La stessa relazione vale in ω : se $f_1(t) \leftrightarrow F_1(\omega)$ e $f_2(t) \leftrightarrow F_2(\omega)$ allora $f_1(t) * f_2(t) \leftrightarrow F_1(\omega) F_2(\omega)$.

Prodotto È la proprietà duale della precedente. Se $f_1(t) \leftrightarrow F_1(f)$ e $f_2(t) \leftrightarrow F_2(f)$

$$f_1(t) f_2(t) \longleftrightarrow F_1(f) * F_2(f) \quad (1.58)$$

Ovvero il prodotto di due segnali nel tempo ha FT che è la convoluzione delle FT.

Dimostrazione: esprimendo $f_2(t)$ come antitrasformata di $F_2(f)$ ottengo

$$\begin{aligned} \mathcal{F}[f_1(t) f_2(t)] &= \int_{-\infty}^{\infty} f_1(t) \int_{-\infty}^{\infty} F_2(\lambda) e^{j2\pi\lambda t} d\lambda e^{-j2\pi ft} dt \\ &= \int_{-\infty}^{\infty} F_2(\lambda) \int_{-\infty}^{\infty} f_1(t) e^{-j2\pi(f-\lambda)t} dt d\lambda \\ &= \int_{-\infty}^{\infty} F_2(\lambda) F_1(f-\lambda) d\lambda = F_1(f) * F_2(f) \end{aligned} \quad (1.59)$$

Nei passaggi si è invertito l'ordine di integrazione e si è usata la proprietà di traslazione in frequenza. La stessa relazione vale in ω : se $f_1(t) \leftrightarrow F_1(\omega)$ e $f_2(t) \leftrightarrow F_2(\omega)$ allora $f_1(t) f_2(t) \leftrightarrow \frac{1}{2\pi} F_1(\omega) * F_2(\omega)$.

Integrazione Posto che $w(t) \leftrightarrow W(f)$ vogliamo esprimere la trasformata del suo integrale

$$y(t) = \int_{-\infty}^t w(\alpha) d\alpha$$

Notiamo che $y(t)$ si può scrivere come

$$y(t) = \int_{-\infty}^{\infty} w(\alpha) U(t-\alpha) d\alpha$$

essendo $U(t)$ la funzione gradino unitario, ovvero, come già visto, la funzione che vale 1 per $t \geq 0$ e zero altrove. La trasformata di Fourier di $U(t)$, come vedremo nel seguito, vale

$$U(t) \longleftrightarrow \frac{1}{2}\delta(f) + \frac{1}{j2\pi f}$$

Applicando ora la proprietà della convoluzione [$y(t) = w(t) * U(t)$] si ottiene

$$y(t) = \int_{-\infty}^t w(\lambda) d\lambda \longleftrightarrow Y(f) = W(f) \left[\frac{1}{2} \delta(f) + \frac{1}{j2\pi f} \right] = \frac{W(f)}{j2\pi f} + \frac{1}{2} W(0) \delta(f) \quad (1.60)$$

In ω

$$\int_{-\infty}^t x(\lambda) d\lambda \longleftrightarrow) = \frac{X(\omega)}{j\omega} + \pi X(0) \delta(\omega)$$

Teorema di Parseval Se $f_1(t) \leftrightarrow F_1(f)$ e $f_2(t) \leftrightarrow F_2(f)$ allora

$$\int_{-\infty}^{\infty} f_1(t) f_2^*(t) dt = \int_{-\infty}^{\infty} F_1(f) F_2^*(f) df \quad (1.61)$$

Dimostrazione: esprimendo $f_1(t)$ come antitrasformata di $F_1(f)$ secondo l'eq. 1.33 si ottiene

$$\begin{aligned} \int_{-\infty}^{\infty} f_1(t) f_2^*(t) dt &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} F_1(f) e^{j2\pi f t} df \right] f_2^*(t) dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_1(f) f_2^*(t) e^{j2\pi f t} df dt \\ &= \int_{-\infty}^{\infty} F_1(f) \left[\int_{-\infty}^{\infty} f_2(t) e^{-j2\pi f t} dt \right]^* df \end{aligned} \quad (1.62)$$

e la 1.61 segue.

In ω il teorema di Parseval è

$$\int_{-\infty}^{\infty} f_1(t) f_2^*(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_1(\omega) F_2^*(\omega) d\omega \quad (1.63)$$

Se $f_1(t) = f_2(t) = f(t)$ il teorema di Parseval diventa il **teorema dell'energia di Rayleigh**:

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |F(f)|^2 df \quad (1.64)$$

Quindi il teorema di Parseval fornisce un modo alternativo per valutare l'energia di un segnale (vedi 1.2) nel dominio della frequenza invece che nel dominio del tempo. Inoltre ci dice che la FT, oltre ad essere lineare, preserva l'energia. Questo fatto conduce alla definizione di **spettro di densità di energia**, o anche *densità spettrale di energia*, che si indica con $S_E(f)$ e vale $S_E(f) = |F(f)|^2$. Torneremo in seguito sull'argomento.

1.6.2 Trasformate notevoli

Vediamo in questa sezione alcune trasformate di segnali "interessanti", sia per le applicazioni che didatticamente.

Esponenziale. Una FT la abbiamo già calcolata nell'Esempio 1.3, l'esponenziale monolatera:

$$\begin{cases} e^{-\alpha t} & \text{per } t \geq 0 \\ 0 & \text{per } t < 0 \end{cases} \longleftrightarrow \frac{1}{\alpha + j2\pi f} \quad \alpha > 0 \quad (1.65)$$

La bilatera è ovvia: $\forall t$

$$e^{-\alpha|t|} \longleftrightarrow \frac{2\alpha}{\alpha^2 + (2\pi f)^2} \quad \alpha > 0 \quad (1.66)$$

Infatti

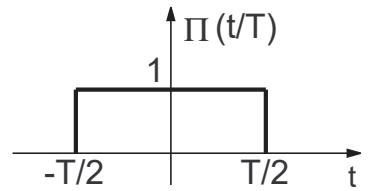
$$\int_{-\infty}^{\infty} e^{-\alpha|t|} e^{-j2\pi f t} dt = \int_0^{\infty} e^{-\alpha t} e^{-j2\pi f t} dt + \int_{-\infty}^0 e^{\alpha t} e^{-j2\pi f t} dt = \frac{1}{\alpha + j2\pi f} + \frac{1}{\alpha - j2\pi f}$$

Rettangolo

Indichiamo con la notazione

$$\Pi\left(\frac{t}{T}\right)$$

la funzione in figura, ovvero un rettangolo alto 1 fra $-T/2$ e $T/2$ e zero altrove.



La sua FT vale

$$\mathcal{F}\left[\Pi\left(\frac{t}{T}\right)\right] = \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-j2\pi f t} dt = \frac{1}{-j2\pi f} [e^{-j2\pi f t}]_{-\frac{T}{2}}^{\frac{T}{2}} = \frac{1}{\pi f} \left(\frac{e^{j2\pi f T/2} - e^{-j2\pi f T/2}}{2j} \right) \quad (1.67)$$

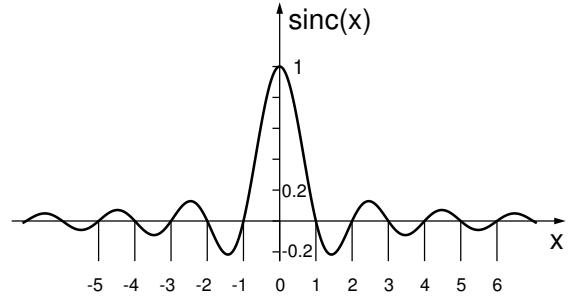
dalle formule di Eulero

$$\mathcal{F}\left[\Pi\left(\frac{t}{T}\right)\right] = \frac{1}{\pi f} \sin\left(2\pi f \frac{T}{2}\right) = T \frac{\sin(\pi f T)}{\pi f T} = T \text{sinc}(fT) \quad (1.68)$$

La funzione **sinc** è definita da

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad (1.69)$$

Il suo andamento sinusoidale smorzato è in figura. Come evidente si annulla per valori interi di x e vale uno nell'origine. Inoltre l'inviluppo dei lobi si attenua come $1/|x|$: il primo massimo del modulo della sinc per x positivi è in 1.5 e vale $1/(1.5\pi)$, il secondo è in 2.5 e vale $1/(2.5\pi)$ eccetera.



Da notare che $T \text{sinc}(fT)$ è reale e pari, come deve essere dato che $\Pi\left(\frac{t}{T}\right)$ è pari e reale. I suoi zeri sono in $f = k/T$, con k intero diverso da zero, ovvero a multipli di $1/T$, ed il valore per $f = 0$ è T , ovvero pari all'area del rettangolo (come deve essere, vedi 1.43).

Dalla 1.68, per la proprietà di dualità, si ottiene, ricordando che il rettangolo è pari,

$$\mathcal{F}[\text{sinc}(\alpha t)] \longleftrightarrow \frac{1}{\alpha} \Pi\left(\frac{f}{\alpha}\right) \quad (1.70)$$

Utilizzando la proprietà dell'area possiamo inoltre calcolare l'area della sinc

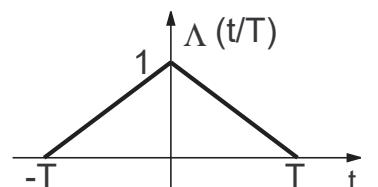
$$\int_{-\infty}^{\infty} \text{sinc}(\alpha t) dt = \frac{1}{\alpha} \Pi\left(\frac{0}{\alpha}\right) = \frac{1}{\alpha} \quad (1.71)$$

Triangolo

Indichiamo con la notazione

$$\Lambda\left(\frac{t}{T}\right)$$

la funzione in figura, ovvero un triangolo alto 1 fra $-T$ e T e zero altrove: da notare che il rettangolo ha base T , il triangolo $2T$, ma entrambi hanno area T .



La sua FT vale

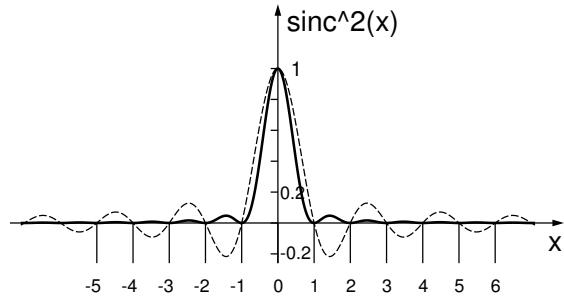
$$\mathcal{F}\left[\Lambda\left(\frac{t}{T}\right)\right] = T \text{sinc}^2(fT) \quad (1.72)$$

Questo risultato si ottiene considerando la convoluzione di due rettangoli alti uno, che produce un triangolo alto T

$$\Pi\left(\frac{t}{T}\right) * \Pi\left(\frac{t}{T}\right) = T \Lambda\left(\frac{t}{T}\right)$$

Utilizzando la proprietà della convoluzione si ottiene subito la 1.72.

La funzione $\text{sinc}^2(x)$ ha l'andamento sinusoidale smorzato in figura. Per confronto viene visualizzata la sinc tratteggiata. Da notare che gli zeri hanno la stessa posizione della sinc(x) e che l'altezza in zero è sempre uno, ma che si smorza più rapidamente, oltre ovviamente ad essere sempre positiva. Infatti l'inviluppo dei lobi si attenua come $1/|x|^2$: il primo massimo della $\text{sinc}^2(x)$ per x positivi è in 1.5 e vale $1/(1.5^2\pi)$, il secondo è in 2.5 e vale $1/(2.5^2\pi)$ eccetera.



1.6.3 Trasformate di segnali generalizzati

Le funzioni generalizzate richiedono una trattazione particolare. In questa sede non entreremo nei dettagli, ma forniremo alcuni esempi.

Impulso di Dirac Considerando le proprietà dell'impulso ideale viste in sezione 1.5.1

$$\mathcal{F}[\delta(t)] = \int_{-\infty}^{\infty} \delta(t) e^{-j2\pi ft} dt = e^{-j2\pi f0} = 1 \quad (1.73)$$

Quindi la trasformata di un impulso è una costante di ampiezza unitaria su tutto l'asse delle frequenze. [Giusto per farvelo notare, questo significa che ha energia infinita!] Dalla proprietà di traslazione nel tempo segue inoltre che

$$\mathcal{F}[\delta(t + t_0)] = e^{j2\pi f t_0} \quad (1.74)$$

Per dualità possiamo ricavare le trasformate della **costante** e dell'**esponenziale complesso**:

$$\mathcal{F}[K] = K \delta(f) \quad \mathcal{F}[e^{j2\pi f_0 t}] = \delta(-f + f_0) = \delta(f - f_0) \quad (1.75)$$

Dalla 1.73 e dalla definizione di FT inversa 1.36 si ha

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega \quad (1.76)$$

o anche, coniugando ambo i membri e cambiando variabile,

$$\int_{-\infty}^{\infty} e^{-j\omega t} dt = 2\pi \delta(\omega) \quad (1.77)$$

risultato che useremo nel seguito.

Funzione segno Definiamo la funzione segno di x come

$$\text{sign}(x) = \begin{cases} 1 & \text{per } x > 0 \\ -1 & \text{per } x < 0 \end{cases} \quad (1.78)$$

Tale funzione non rispetta le condizioni di esistenza. Se proviamo a calcolare la FT secondo definizione si ottiene

$$\begin{aligned} \mathcal{F}[\text{sign}(t)] &= \int_0^{\infty} e^{-j2\pi ft} dt - \int_{-\infty}^0 e^{-j2\pi ft} dt \\ &= \int_0^{\infty} (e^{-j2\pi ft} - e^{j2\pi ft}) dt = \frac{2}{j} \int_0^{\infty} \sin(2\pi ft) dt \end{aligned}$$

Gli integrali sopra *non esistono*, come prevedibile visto che $\text{sign}(x)$ non è infinitesimo al divergere di x . Si può tuttavia calcolare nel senso delle funzioni generalizzate (vedi, ad esempio, l'appendice A di [EB77a]), ottenendo

$$\mathcal{F}[\text{sign}(t)] \Rightarrow \frac{1}{j\pi f} \Rightarrow \frac{2}{j\omega} \quad (1.79)$$

Gradino unitario La funzione gradino unitario $U(t)$ (vista in esempio 1.3) si può esprimere utilizzando la funzione segno come $U(t) = \frac{1}{2} + \frac{1}{2} \text{sign}(t)$.

La sua FT vale quindi:

$$\mathcal{F}[U(t)] \Rightarrow \frac{1}{2}\delta(f) + \frac{1}{j2\pi f} \Rightarrow \pi\delta(\omega) + \frac{1}{j\omega} \quad (1.80)$$

Seno e coseno Per calcolarle le FT di seno e coseno utilizziamo le formule di Eulero

$$\cos(2\pi f_c t) = \frac{1}{2} [e^{j2\pi f_c t} + e^{-j2\pi f_c t}] \quad \sin(2\pi f_c t) = \frac{1}{2j} [e^{j2\pi f_c t} - e^{-j2\pi f_c t}]$$

Dalla 1.75 si ottiene allora

$$\mathcal{F}[\cos(2\pi f_c t)] = \frac{1}{2} [\delta(f - f_c) + \delta(f + f_c)] \quad (1.81)$$

$$\mathcal{F}[\sin(2\pi f_c t)] = \frac{1}{2j} [\delta(f - f_c) - \delta(f + f_c)] \quad (1.82)$$

Da notare che la FT del coseno (pari) è puramente reale e quella del seno (dispari) puramente immaginaria. Inserendo un termine di fase si ottiene la FT di una sinusode generale

$$\mathcal{F}[\cos(2\pi f_c t + \theta)] = \frac{1}{2} e^{j\theta} \delta(f - f_c) + \frac{1}{2} e^{-j\theta} \delta(f + f_c) \quad (1.83)$$

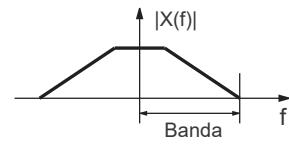
1.6.4 Banda di un segnale

L'intervallo sull'asse delle frequenze occupato da un segnale è una quantità molto importante nei sistemi di telecomunicazione, sia per l'assegnazione dei canali di comunicazione che per la progettazione degli apparati. Questo intervallo prende il nome di *banda del segnale*. Il problema è che esistono diverse definizioni di banda, a seconda che si consideri la definizione di "intervallo occupato sull'asse delle frequenze" in modo assoluto o con approssimazioni. Esistono infatti segnali che hanno valori non nulli su intervalli molto grandi in frequenza (al limite infiniti), ma la cui parte principale è concentrata in un certo sotto-intervallo: per questi segnali si definisce la banda trascurando le parti con contenuto in frequenza "piccolo".

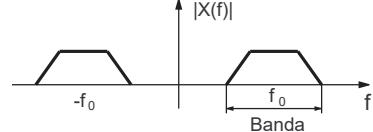
In ogni caso la banda è sempre calcolata considerando il *solo asse positivo* delle frequenze. Ciò si giustifica notando che in telecomunicazioni si considerano segnali reali e filtri (sistemi LTI) aventi risposta impulsiva reale, per i quali la FT ha simmetria hermitiana (quindi la conoscenza dell'andamento sull'asse positivo è sufficiente a ricavare anche la parte per frequenze negative).

In altre parole la banda è sempre una quantità del tipo $f_2 - f_1$, dove $f_2 > f_1 \geq 0$.

Si definiscono **segnali in banda base** quei segnali la cui FT è concentrata attorno all'origine dell'asse delle frequenze. Per questi segnali $f_1 = 0$ e la definizione di banda è esemplificata in figura. **Definizione:** un segnale $x(t)$ in banda base ha **banda** W se la sua FT è tale che $X(f) = 0$ per $|f| \geq W$. L'uguaglianza implica che $X(f) = 0$ per $|f| = W$: normalmente non ha importanza, tranne per segnali che hanno un impulso in $f = W$ (che hanno quindi banda maggiore di W).



Si definiscono **segnali passa-banda** quei segnali la cui FT è concentrata attorno alle frequenze $\pm f_0$ ed ha componenti nulle o trascurabili intorno all'origine.



La *banda assoluta* di un segnale si ottiene considerando solo l'intervallo nel quale $|X(f)| \neq 0$. Altre possibili criteri sono la banda a -3 dB o la banda al 99 % dell'energia, che definiremo nel seguito.

1.6.5 Relazioni tempo-frequenza e principio di indeterminazione

In questa sezione cercheremo le relazioni che legano l'andamento di un segnale $x(t)$ a quello della sua FT $X(f)$. Abbiamo già dimostrato (vedi 1.47) che se un segnale si allarga nel tempo si stringe in frequenza e viceversa. Vediamo ora altre relazioni che approfondiscono l'argomento.

Limiti alle variazioni di un segnale a banda limitata* Sia $x(t)$ un segnale in banda base di banda W : come già visto questo significa che la sua FT è tale che $X(f) = 0$ per $|f| \geq W$. La variazione di $x(t)$ nell'intervallo (t_1, t_2) vale allora (usando la 1.33)

$$|x(t_2) - x(t_1)| = \left| \int_{-W}^W X(f)(e^{j2\pi f t_2} - e^{j2\pi f t_1}) df \right| \leq \int_{-W}^W |X(f)| \cdot |(e^{j2\pi f t_2} - e^{j2\pi f t_1})| df$$

siccome

$$\begin{aligned} e^{j\alpha t_2} - e^{j\alpha t_1} &= \cos(\alpha t_2) + j \sin(\alpha t_2) - [\cos(\alpha t_1) + j \sin(\alpha t_1)] \\ &= \cos(\alpha t_2) - \cos(\alpha t_1) + j[\sin(\alpha t_2) - \sin(\alpha t_1)] \\ &= -2 \sin(\alpha \frac{t_2 + t_1}{2}) \sin(\alpha \frac{t_2 - t_1}{2}) + j2 \cos(\alpha \frac{t_2 + t_1}{2}) \cos(\alpha \frac{t_2 - t_1}{2}) \end{aligned}$$

e

$$\begin{aligned} |e^{j\alpha t_2} - e^{j\alpha t_1}| &= \sqrt{4 \sin^2(\alpha \frac{t_2 + t_1}{2}) \sin^2(\alpha \frac{t_2 - t_1}{2}) + 4 \cos^2(\alpha \frac{t_2 + t_1}{2}) \cos^2(\alpha \frac{t_2 - t_1}{2})} \\ &= \sqrt{4 \sin^2(\alpha \frac{t_2 - t_1}{2}) \left[\sin^2(\alpha \frac{t_2 + t_1}{2}) + \cos^2(\alpha \frac{t_2 + t_1}{2}) \right]} = 2 |\sin(\alpha \frac{t_2 - t_1}{2})| \end{aligned}$$

possiamo scrivere

$$|(e^{j2\pi f t_2} - e^{j2\pi f t_1})| = 2 \left| \sin \frac{2\pi f(t_2 - t_1)}{2} \right| \leq |2\pi f(t_2 - t_1)|$$

e quindi

$$\left| \frac{x(t_2) - x(t_1)}{t_2 - t_1} \right| \leq \int_{-W}^W |X(f)| |2\pi f| df \leq 2\pi W \int_{-W}^W |X(f)| df \quad (1.84)$$

Il primo membro della 1.84 rappresenta la variazione relativa del segnale $x(t)$ nell'intervallo (t_1, t_2) : si vede come essa sia proporzionale alla larghezza di banda W del segnale. L'andamento di un segnale sarà allora tanto più dolce quanto più è stretta la sua banda. Viceversa, affinché un segnale possa avere brusche variazioni è necessario che la sua banda sia larga.

Come esempio possiamo ricordare la FT del gradino unitario, che ha una variazione in zero molto brusca (variazione relativa infinita) e banda infinita. Al contrario le FT di seno e coseno, che sono a variazione limitata, hanno banda finita.

Segnali a supporto limitato Consideriamo un segnale ad energia finita definito (diverso da zero) nell'intervallo $(-T/2, T/2)$, cioè un segnale strettamente limitato nel tempo. Ci chiediamo se esso possa anche essere limitato in banda, ovvero se possa esistere un valore W tale che $X(f) = 0$ per $|f| > W$. Per esaminare tale possibilità osserviamo che $x(t)$ si può esprimere come il prodotto di due segnali

$$x(t) = \prod \left(\frac{t}{T} \right) \cdot z(t)$$

dove $z(t)$ coincide con $x(t)$ in $(-T/2, T/2)$ e può essere come vuole altrove. Allora la FT di $x(t)$, ovvero $X(f)$, si può ottenere come convoluzione di $Z(f)$ con la FT del rettangolo. Poichè la FT del rettangolo ha banda infinita (vedi 1.68) e la convoluzione di due segnali ha durata maggiore delle singole durate, ne risulta che un segnale a durata strettamente limitata nel tempo non può essere strettamente limitato in frequenza.

In modo del tutto analogo si può anche dimostrare la proposizione inversa: ogni segnale strettamente limitato in frequenza non può essere strettamente limitato nel tempo.

Principio di indeterminazione* Rendere quantitative le osservazioni precedenti non è semplice. Il principio di indeterminazione è legato al fatto che non posso definire precisamente il contenuto in frequenza di un segnale se lo osservo per un tempo ridotto. Un modo per capire il concetto è quello di notare che osservare un segnale per poco tempo equivale a moltiplicarlo per una funzione (ad esempio un rettangolo) di breve durata. Questo, come visto, equivale a convolare lo spettro con una funzione "larga" in frequenza e quindi a perdere i dettagli frequenziali.

Per il principio di dualità vale anche l'osservazione inversa: se seleziono un ridotto intervallo di frequenze della FT di un segnale non sono in grado di localizzarle precisamente nel tempo.

Per un approfondimento si può leggere, ad esempio, [EB77a] da pag. 57 o anche [Pap62] da pag. 63.

1.7 Serie di Fourier

La serie di Fourier è un particolare tipo di rappresentazione su base ortogonale con funzioni sinusoidali o esponenziali complesse (vedremo più avanti cosa vuol dire...).

Sia $x(t)$ un segnale ad energia finita definito sull'intervallo $(a, a + T_0)$. Esso può essere rappresentato in tale intervallo dalla serie di Fourier

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad (1.85)$$

dove i coefficienti c_n sono dati da

$$c_n = \frac{1}{T_0} \int_a^{a+T_0} x(t) e^{-jn\omega_0 t} dt \quad (1.86)$$

con $\omega_0 = 2\pi f_0 = 2\pi/T_0$.

La condizione affinchè la serie di Fourier esista è al solito complessa: quella indicata, energia finita, è solo sufficiente, una alternativa, sempre sufficiente, sono le già viste condizioni di Dirichlet.

I coefficienti c_n sono in generale complessi. La 1.85 ci dice che qualunque segnale ad energia finita può essere scomposto in una somma di infiniti esponenziali complessi $\varphi_n(t) = e^{jn\omega_0 t}$: poichè, per le formule di Eulero, $e^{jn\omega_0 t} = \cos(n\omega_0 t) + j \sin(n\omega_0 t)$, $\varphi_n(t)$, per $n \neq 0$, è periodica di periodo T_0/n . Per $n = 0$, $\varphi_0(t)$ è costante e c_0 è pari al valor medio di $x(t)$, risultato che si ottiene confrontando la 1.86 con la 1.1. La frequenza $f_0 = 1/T_0$ è detta fondamentale, mentre le frequenze $n f_0$ con $n > 1$ sono dette armoniche.

Da notare che tutte le $\varphi_n(t)$ sono periodiche con periodo comune minimo pari a T_0 . Quindi, se il segnale $x(t)$ è periodico di periodo T_0 , la rappresentazione mediante serie di Fourier vale su tutto l'asse temporale (cioè per $-\infty < t < \infty$), visto che sia $x(t)$ che le $\varphi_n(t)$ sono periodiche dello stesso periodo. In questo caso la scelta del valore di a è arbitraria: di solito si sceglie $a = 0$ oppure $a = -T_0/2$ (per semplicità).

Teorema di Parseval Dimostreremo più avanti che vale la relazione

$$\frac{1}{T_0} \int_a^{a+T_0} |x(t)|^2 dt = \sum_{n=-\infty}^{\infty} |c_n|^2 \quad (1.87)$$

Quindi se $x(t)$ è aperiodico (definito solo in $(a, a + T_0)$) la sua energia si può calcolare dai coefficienti della serie di Fourier:

$$E_x = T_0 \sum_{n=-\infty}^{\infty} |c_n|^2$$

Se invece è periodico la 1.87 fornisce la potenza media di $x(t)$:

$$P_x = \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} |x(t)|^2 dt = \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} |x(t)|^2 dt = \sum_{n=-\infty}^{\infty} |c_n|^2 \quad (1.88)$$

Dove, essendo $x(t)$ periodico, il secondo integrale si può calcolare su qualsiasi intervallo lungo T_0 .

Proprietà La serie di Fourier possiede diverse proprietà, alcune sono:

1. se $x(t)$ è reale $c_n = c_{-n}^*$;
2. se $x(t)$ è reale e pari $\text{Im}[c_n] = 0$;
3. se $x(t)$ è reale e dispari $\text{Re}[c_n] = 0$;
4. $x(t + t_0) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0(t+t_0)} = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} e^{jn\omega_0 t_0}$
5. $x(t - t_0) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} e^{-jn\omega_0 t_0}$

Da notare come queste proprietà siano analoghe a quelle delle trasformata di Fourier.

1.8 Trasformata di Fourier di funzioni periodiche

Sia $v(t)$ un segnale periodico di periodo T . La sua espansione in serie di Fourier vale

$$v(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\Omega t} \quad t \in (-\infty, \infty) \quad (1.89)$$

dove

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t) e^{-jn\Omega t} dt \quad \Omega = \frac{2\pi}{T} \quad (1.90)$$

Calcoliamo la trasformata di Fourier di $v(t)$ secondo l'eq. 1.35 ed usando l'eq. 1.77:

$$\begin{aligned} V(\omega) &= F[v(t)] = \int_{-\infty}^{\infty} v(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} \left[\sum_{n=-\infty}^{\infty} c_n e^{jn\Omega t} \right] e^{-j\omega t} dt \\ &= \sum_{n=-\infty}^{\infty} c_n \int_{-\infty}^{\infty} e^{jn\Omega t} e^{-j\omega t} dt = 2\pi \sum_{n=-\infty}^{\infty} c_n \delta(\omega - n\Omega) \end{aligned} \quad (1.91)$$

Da ciò si vede che la trasformata di Fourier di un segnale periodico è **a righe**, cioè formata da una successione di impulsi. La distanza fra le righe è costante, pari a $2\pi/T$ in ω (o $1/T$ in f) e quindi inversamente proporzionale al periodo del segnale. Questo risultato è ovvio se si pensa che un segnale periodico di periodo T può contenere solo sinusoidi che abbiano in T un numero di periodi intero, in modo da ripartire esattamente allo stesso modo nel periodo successivo.

E' inoltre evidente il legame tra la trasformata di Fourier ed i coefficienti della serie di Fourier dello stesso segnale periodico.

Se definiamo $v_T(t)$ come il segnale aperiodico consistente in un periodo di $v(t)$, ossia

$$v_T(t) = v(t) \prod\left(\frac{t}{T}\right) = \begin{cases} v(t) & -\frac{T}{2} < t < \frac{T}{2} \\ 0 & \text{altrove} \end{cases}$$

la sua trasformata di Fourier vale

$$V_T(\omega) = F[v_T(t)] = \int_{-\infty}^{\infty} v_T(t) e^{-j\omega t} dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} v_T(t) e^{-j\omega t} dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t) e^{-j\omega t} dt$$

e ricordando che (vedi eq. 1.90)

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t) e^{-jn\Omega t} dt = \frac{1}{T} V_T(n\Omega)$$

si ottiene dall'eq. 1.91

$$V(\omega) = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} V_T(n\Omega) \delta(\omega - n\Omega)$$

ovvero che l'altezza (area) della riga centrata in $n\Omega$ è proporzionale al valore della trasformata di Fourier del segnale $v_T(t)$.

Lo stesso risultato espresso in frequenza anzichè in pulsazioni è

$$V(f) = \frac{1}{T} \sum_{n=-\infty}^{\infty} V_T\left(\frac{n}{T}\right) \delta(f - \frac{n}{T})$$

Specializzando il risultato ottenuto è possibile trovare la trasformata di un treno di impulsi

$$F\left[\sum_{n=-\infty}^{\infty} \delta(t - nT)\right] = \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{T}) \quad (1.92)$$

Questo risultato, che useremo nel seguito, ci dice che una sequenza di infiniti impulsi di Dirac equispaziati nel tempo ha per trasformata una sequenza di infiniti impulsi di Dirac equispaziati in frequenza. Inoltre se avvicino gli impulsi nel tempo si allontanano in frequenza e viceversa.

Da notare un aspetto per nulla intuitivo: un impulso ha trasformata pari ad una costante, infiniti impulsi hanno trasformata pari a infiniti impulsi...

1.9 Spettro di densità di potenza

In sezione 1.7 abbiamo visto che la potenza media del segnale periodico (di periodo T) $v(t)$ vale

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

inoltre abbiamo visto che

$$V(\omega) = 2\pi \sum_{n=-\infty}^{\infty} c_n \delta(\omega - n\Omega)$$

o, in frequenza,

$$V(f) = \sum_{n=-\infty}^{\infty} c_n \delta(f - \frac{n}{T})$$

Definiamo ora **spettro di densità di potenza** la funzione

$$S_v(\omega) = 2\pi \sum_{n=-\infty}^{\infty} |c_n|^2 \delta(\omega - n\Omega) \quad (1.93)$$

che esprime la distribuzione di potenza di $v(t)$ nel dominio di Fourier. Si tratta evidentemente di uno spettro a righe. La potenza media di $v(t)$ si ricava da

$$P_v = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(\omega) d\omega$$

L'antitrasformata di Fourier di $S_v(\omega)$ è

$$R_v(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_v(\omega) e^{j\omega\tau} d\omega = \sum_{n=-\infty}^{\infty} |c_n|^2 e^{jn\Omega\tau} = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t+\tau) v^*(t) dt \quad (1.94)$$

Questo risultato si ottiene come segue: partendo dalla eq. 1.88 si può scrivere che

$$\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t) v^*(t) dt = \sum_{n=-\infty}^{\infty} |c_n|^2 = \sum_{n=-\infty}^{\infty} c_n c_n^*$$

dove c_n e c_n^* sono al solito definiti come

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t) e^{-jn\Omega t} dt \quad c_n^* = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t)^* e^{jn\Omega t} dt$$

L'ultimo passaggio deriva infine dalla proprietà della serie di Fourier per cui

$$v(t \pm t_0) = \sum_{n=-\infty}^{\infty} c_n e^{\pm jn\Omega t_0} e^{jn\Omega t}$$

Infatti:

$$\frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t+t_0) e^{-jn\Omega t} dt =$$

chiamando $\alpha = t + t_0$, da cui $t = \alpha - t_0$,

$$\frac{1}{T} \int_{-\frac{T}{2}+t_0}^{\frac{T}{2}+t_0} v(\alpha) e^{-jn\Omega\alpha} e^{jn\Omega t_0} dt = e^{jn\Omega t_0} c_n$$

Ovviamente lo spostamento di t_0 dei limiti di integrazione non ha effetto, l'importante è che l'integrale venga effettuato su di una durata pari ad un periodo.

La funzione $R_v(\tau)$ è detta **funzione di autocorrelazione** di $v(t)$ ed ha alcune interessanti proprietà:

- a) è periodica di periodo T , come $v(t)$;
- b) è reale e pari;
- c) assume il suo valore massimo nell'origine (anche se possono esserci altri massimi di pari ampiezza), ovvero

$$R_V(0) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |v(t)|^2 dt = P_v \geq |R_v(\tau)| \quad \forall \tau$$

1.10 Spettro di densità di energia

Sia $x(t)$ un segnale aperiodico ad energia finita pari a

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} x(t)x^*(t) dt$$

Se si conosce $X(\omega)$, trasformata di Fourier di $x(t)$, si può scrivere, usando l'eq. 1.36

$$E_x = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(\omega) X^*(\omega') e^{j(\omega-\omega')t} d\omega d\omega' dt$$

integrando in t , ricordando l'eq. 1.77, si ottiene

$$E_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(\omega) X^*(\omega') \delta(\omega - \omega') d\omega d\omega' = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) X^*(\omega) d\omega$$

e quindi

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega$$

che è detta uguaglianza di Parseval (vedi anche eq. 1.64).

Possiamo ora definire lo spettro di densità di energia di $x(t)$ come

$$S_x(\omega) = |X(\omega)|^2$$

Esso ha una antitrasformata di Fourier che si chiama ancora funzione di autocorrelazione di $x(t)$ e si può scrivere come

$$R_x(\tau) = F[S_x(\omega)]^{-1} = \int_{-\infty}^{\infty} x(t + \tau)x^*(t) dt = \int_{-\infty}^{\infty} x(t)x^*(t - \tau) dt$$

La funzione di autocorrelazione $R_x(\tau)$ ha le proprietà **b)** e **c)** già viste per il caso periodico (non ha la proprietà a)). È inoltre evidente che $R_x(0) = E_x$.

1.11 Sistemi LTI: risposta in frequenza

Abbiamo già visto in sez. 1.5.5 la definizione di funzione di trasferimento $H(\omega)$ per esponenziali complessi o risposta in frequenza. A questo punto è ovvio come essa sia la FT della risposta del sistema LTI all'impulso

$$h(t) = L[\delta(t)] \longleftrightarrow H(\omega)$$

L'uscita del sistema al generico ingresso $x(t)$ si ottiene, come visto, mediante convoluzione

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

Per le proprietà della FT posso allora trovare l'uscita anche mediante la relazione

$$Y(f) = H(f)X(f) \tag{1.95}$$

ovvero moltiplicando fra di loro le FT dell'ingresso e della risposta all'impulso. In molti casi è più semplice calcolare l'uscita del sistema mediante la 1.95 che eseguendo l'integrale di convoluzione. Ovviamente per ottenere $y(t)$ devo antitrasformare $Y(f)$.

1.11.1 Esempi di sistemi LTI

Vediamo alcuni esempi di sistemi LTI.

Circuito RC

Consideriamo il circuito RC in figura: poichè $y(t)$ è la tensione ai capi del condensatore (e $x(t)$ la tensione in ingresso), vale che $x(t) = Ri(t) + y(t)$, dove $i(t)$ è la corrente che passa in R (ed in C).

Quindi $x(t) = RC \frac{dy(t)}{dt} + y(t)$. Trasformando si ottiene

$$X(f) = RC j2\pi f Y(f) + Y(f) = Y(f) [RC j2\pi f + 1]$$

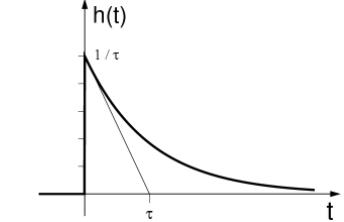
La risposta in frequenza è allora

$$H(f) = \frac{Y(f)}{X(f)} = \frac{1}{1 + j2\pi RC f} = \frac{1}{1 + j2\pi \tau f}$$

avendo chiamato $\tau = RC$ la costante di tempo del circuito. L'antitrasformata fornisce la risposta all'impulso (vedi 1.65)

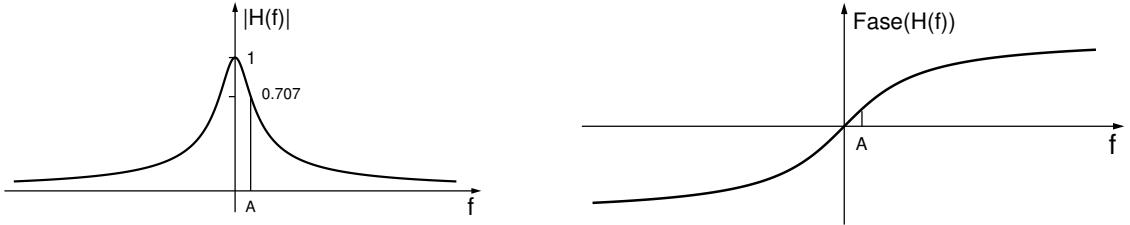
$$h(t) = \frac{1}{RC} e^{-\frac{t}{RC}} U(t) = \frac{1}{\tau} e^{-\frac{t}{\tau}} U(t)$$

Il modulo e la fase di $H(f)$ sono



$$|H(f)| = \sqrt{\frac{1}{1 + (2\pi RC f)^2}} \quad \angle H(f) = \arctan(-2\pi RC f)$$

e sono visualizzati nella figura sotto, dove si è chiamato $A = 1/(2\pi\tau)$ il valore di f per cui il modulo scende al valore $1/\sqrt{2}$, rispetto al massimo che vale uno in $f = 0$. La fase ha asintoti orizzontali pari a $\pm\pi/2$.



Altri esempi Da quanto visto sulla FT si può definire la risposta di alcuni semplici sistemi:

Operazione nel tempo	Funzione di trasferimento	
Moltiplicazione	$y(t) = \pm Kx(t)$	$H(f) = \pm K$
Differenziazione	$y(t) = \frac{dx(t)}{dt}$	$H(f) = j2\pi f$
Integrazione	$y(t) = \int_{-\infty}^t x(\lambda)d\lambda$	$H(f) = \frac{1}{j2\pi f}$
Ritardo	$y(t) = x(t - t_d)$	$H(f) = e^{-j2\pi f t_d}$

1.11.2 Diagrammi a blocchi

Spesso i sistemi di comunicazione sono composti da più sistemi LTI connessi fra loro. Le regole per ricavare la risposta complessiva sono:

Sistemi in cascata Se due sistemi sono connessi in cascata come in Fig. 1.1, la risposta all'impulso e la funzione di trasferimento del sistema risultante sono:

$$h(t) = h_1(t) * h_2(t) \quad H(\omega) = H_1(\omega)H_2(\omega) \quad (1.96)$$

Sistemi in parallelo Se due sistemi sono connessi in parallelo come in Fig 1.2 (a), la risposta all'impulso e la funzione di trasferimento del sistema risultante sono:

$$h(t) = h_1(t) + h_2(t) \quad H(\omega) = H_1(\omega) + H_2(\omega) \quad (1.97)$$

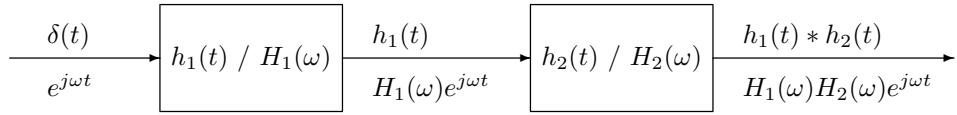


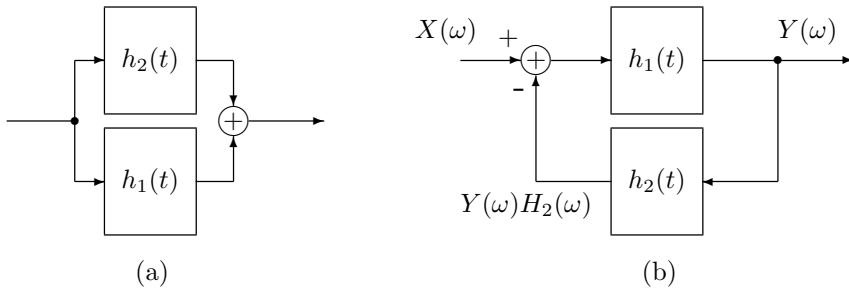
Figura 1.1: Sistemi in cascata.

Sistemi in retroazione Se due sistemi sono connessi in retroazione come in Fig 1.2 (b), la funzione di trasferimento del sistema risultante è:

$$H(\omega) = \frac{H_1(\omega)}{1 + H_1(\omega)H_2(\omega)} \quad (1.98)$$

La risposta all'impulso è difficile da esprimere in funzione delle risposte dei singoli blocchi e viene solitamente ricavata antitrasformando $H(\omega)$.

Da notare che l'esempio è di retroazione negativa, in quanto il segnale di retrazione viene sottratto dall'ingresso.

Figura 1.2: Sistemi in parallelo (a) ed in retroazione (b). La funzione di trasferimento del sistema (b), eq. 1.98, si ottiene da $Y(\omega) = [X(\omega) - Y(\omega)H_2(\omega)]H_1(\omega)$

Esempio 1.4 Il sistema in figura, nel quale il blocco T è un ritardatore ideale di T secondi, ha numerose applicazioni nelle comunicazioni elettriche. Verifichiamo se è un sistema LTI. Per semplice ispezione possiamo trovare la relazione che lega l'ingresso con l'uscita:

$$y(t) = \int_{-\infty}^t [x(\lambda) - x(\lambda - T)] d\lambda \quad \text{Diagram: Input } x(t) \rightarrow \text{Block } T \rightarrow \text{Summing junction} \rightarrow \text{Block } \int_{-\infty}^t \rightarrow \text{Output } y(t)$$

Da questa è facile vedere che il sistema è lineare e tempo invariante. La risposta all'impulso si trova sostituendo $\delta(t)$ a $x(t)$

$$h(t) = \int_{-\infty}^t [\delta(\lambda) - \delta(\lambda - T)] d\lambda = U(t) - U(t - T)$$

essendo $U(t)$ la funzione gradino unitario. La risposta all'impulso è quindi un rettangolo di altezza unitaria e durata T che inizia in $t = 0$. Il sistema è quindi reale, causale e stabile.

La risposta in frequenza si trova per mezzo delle relazioni 1.96 e 1.97, dato che conosciamo le funzioni di trasferimento dei blocchi ritardo ed integratore, e dalle formule di Eulero:

$$H(f) = [1 - e^{-j2\pi fT}] \frac{1}{j2\pi f} = \frac{e^{j\pi fT} - e^{-j\pi fT}}{j2\pi f} e^{-j\pi fT} = \frac{\sin(\pi fT)}{\pi f} e^{-j\pi fT}$$

che è una funzione *sinc* in frequenza e corrisponde ovviamente alla trasformata di $h(t)$.

1.11.3 Correlazione incrociata ingresso-uscita

Si definisce correlazione incrociata di due funzioni a energia finita $v(t)$ e $w(t)$ la funzione

$$R_{vw}(\tau) = \int_{-\infty}^{\infty} v(t) w^*(t - \tau) dt \quad (1.99)$$

Si definisce autocorrelazione della funzione $v(t)$, come visto in sezione 1.10,

$$R_v(\tau) = R_{vv}(\tau) = \int_{-\infty}^{\infty} v(t) v^*(t - \tau) dt \quad (1.100)$$

Analizzando l'equazione 1.99 ci accorgiamo che è un tipo di convoluzione. Chiamando $z(t) = w^*(-t)$ e $t = \lambda$, la parte destra della 1.99 diventa

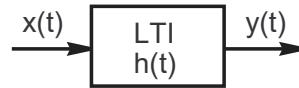
$$\int_{-\infty}^{\infty} v(\lambda) z(\tau - \lambda) d\lambda = v(\tau) * z(\tau)$$

e quindi

$$R_{vw}(\tau) = v(\tau) * w^*(-\tau) \quad (1.101)$$

Analogamente $R_v(\tau) = v(\tau) * v^*(-\tau)$.

Dato il sistema LTI in figura, con risposta all'impulso $h(t)$ ed avente in ingresso la funzione $x(t)$, la cui funzione di autocorrelazione vale $R_x(\tau)$,



la funzione di correlazione incrociata fra l'uscita $y(t)$ del sistema ed il suo ingresso vale

$$R_{yx}(\tau) = h(\tau) * R_x(\tau) = \int_{-\infty}^{\infty} h(\lambda) R_x(\tau - \lambda) d\lambda \quad (1.102)$$

Infatti, usando le equazioni 1.101 e 1.22 si ottiene

$$R_{yx}(\tau) = y(\tau) * x^*(-\tau) = h(\tau) * x(\tau) * x^*(-\tau) = h(\tau) * R_x(\tau)$$

La funzione di autocorrelazione dell'uscita vale

$$R_y(\tau) = h^*(-\tau) * R_{yx}(\tau) = \int_{-\infty}^{\infty} h^*(-\lambda) R_{yx}(\tau - \lambda) d\lambda \quad (1.103)$$

o anche

$$R_y(\tau) = h^*(-\tau) * h(\tau) * R_x(\tau) \quad (1.104)$$

Infatti

$$R_y(\tau) = y(\tau) * y^*(-\tau) = h(\tau) * x(\tau) * h^*(-\tau) * x^*(-\tau) = h(\tau) * h^*(-\tau) * R_x(\tau)$$

1.12 Filtri

Col termine **filtro** si può indicare un qualsiasi sistema LTI, ma normalmente viene applicato a quei sistemi che hanno una particolare risposta in frequenza. Si definisce **filtro ideale** un sistema LTI che lasci passare in modo indistorto le frequenze appartenenti ad un certo intervallo e blocca quelle esterne, ovvero un sistema con risposta in frequenza

$$H(f) = \begin{cases} K e^{-j2\pi f t_d} & f_1 \leq |f| \leq f_2 \\ 0 & \text{altrove} \end{cases} \quad K > 0 \quad (1.105)$$

Dove t_d è il **ritardo** introdotto dal filtro, f_1 e f_2 sono dette **frequenze di taglio** e K è il **guadagno in ampiezza**. Come evidente la fase di tali filtri nell'intervallo $[f_1, f_2]$ è lineare, derivando solo dal ritardo introdotto fra ingresso e uscita: in queste condizioni un segnale avente componenti in frequenza diverse da zero nel solo intervallo $[f_1, f_2]$ non viene modificato dal filtro, ma solo ritardato e, se $K \neq 1$ amplificato ($K > 1$) o attenuato ($K < 1$).

La differenza $B = f_2 - f_1$ si dice **banda passante** del filtro.

I filtri sono normalmente distinti in passa-basso, passa-alto e passa-banda.

1.12.1 Filtri ideali

Filtri passa-basso (LPF) Sono filtri che lasciano passare le frequenze inferiori ad una certa soglia, ovvero definiti da

$$H_{LP}(f) = \begin{cases} K e^{-j2\pi f t_d} & |f| \leq f_2 \quad K > 0 \\ 0 & \text{altrove} \end{cases} \quad (1.106)$$

Il termine LP significa *Low Pass*, ovvero passa basso in inglese.

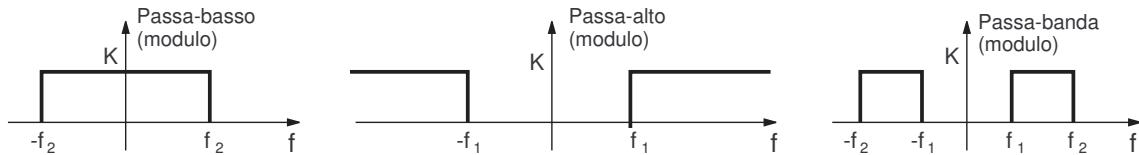
Il circuito RC visto in sezione 1.11.1 è un filtro passa-basso non ideale, visto che non è piatto in banda, che lascia passare le alte frequenze (anche se attenuate) e che la fase non è lineare. I filtri che si possono realizzare sono di questo tipo, cioè non ideali. Per essi vedremo più avanti una definizione di banda passante "approssimata".

Filtri passa-alto (HPF) Sono filtri che lasciano passare le frequenze superiori ad una certa soglia, ovvero definiti da

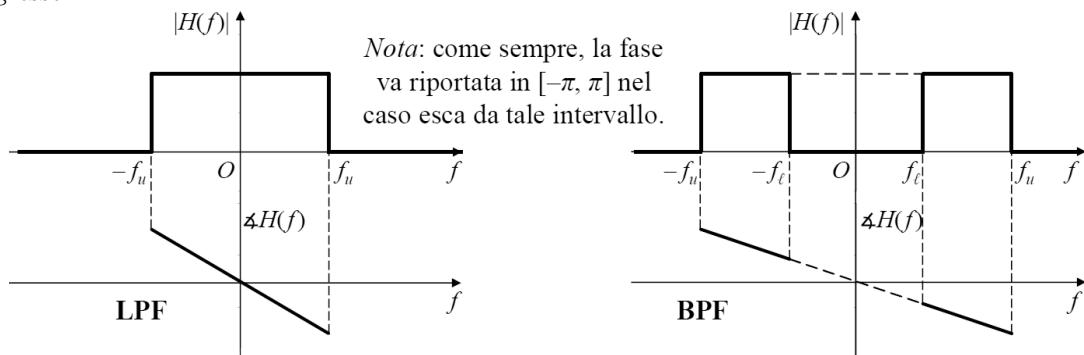
$$H_{HP}(f) = \begin{cases} K e^{-j2\pi f t_d} & f_1 \leq |f| \quad K > 0 \\ 0 & \text{altrove} \end{cases} \quad (1.107)$$

Il termine HP significa *High Pass*, ovvero passa alto in inglese.

Filtri passa-banda (BPF) Sono filtri che lasciano passare le frequenze comprese in un certo intervallo, definiti quindi della 1.105. In inglese si indicano col termine *Band Pass Filter*, ovvero BPF.



Fase di un filtro ideale La fase della funzione di trasferimento di un filtro ideale è **lineare** negli intervalli dell'asse f dove il modulo è diverso da zero (dove il modulo è zero la fase non è definita). Poichè la fase è lineare ($-2\pi f t_d$), le componenti frequenziali che il filtro lascia passare in uscita presentano un ritardo temporale pari a t_d rispetto alle componenti corrispondenti del segnale in ingresso.



1.12.2 Risposta all'impulso dei filtri ideali

Filtri passa-basso o LPF Come visto sono i filtri definiti da

$$H_{LP}(f) = \begin{cases} K e^{-j2\pi f t_d} & |f| \leq f_2 \quad K > 0 \\ 0 & \text{altrove} \end{cases}$$

Un modo diverso per scrivere la loro risposta in frequenza utilizzando la funzione rettangolo è

$$H_{LP}(f) = K \prod \left(\frac{f}{2f_2} \right) e^{-j2\pi f t_d}$$

che ha per antitrasformata la risposta all'impulso

$$h_{LP}(t) = K 2f_2 \operatorname{sinc}(2f_2(t - t_d)) \quad (1.108)$$

Questo risultato si ottiene utilizzando l'equazione 1.70 e la proprietà della traslazione nel tempo di equazione 1.47. Per inciso all'esame dovete essere in grado di eseguire antitrasformate di questo tipo...

L'equazione 1.108 ci dice che il filtro passa-basso ideale non è causale visto che la risposta all'impulso si estende da meno infinito a più infinito, quindi **non è realizzabile**.

Filtri passa-alto o HPF Come visto sono i filtri definiti da

$$H_{HP}(f) = \begin{cases} K e^{-j2\pi f t_d} & f_1 \leq |f| \quad K > 0 \\ 0 & \text{altrove} \end{cases}$$

Possono essere visti come un filtro passa-tutto meno un passa-basso

$$H_{HP}(f) = K \left[1 - \prod \left(\frac{f}{2f_1} \right) \right] e^{-j2\pi f t_d} = K e^{-j2\pi f t_d} - \prod \left(\frac{f}{2f_1} \right) e^{-j2\pi f t_d}$$

che ha per antitrasformata la risposta all'impulso

$$h_{HP}(t) = K \delta(t - t_d) - K 2f_2 \operatorname{sinc}(2f_1(t - t_d)) \quad (1.109)$$

Di nuovo la risposta all'impulso del filtro ideale non è causale e poichè si estende da meno infinito a più infinito il filtro **non è realizzabile**.

Filtri passa-banda o BPF Si possono realizzare mettendo in serie un filtro passa-basso e un filtro passa-alto.

Dall'equazione 1.105 e chiamando $B = f_2 - f_1$ e $f_c = (f_2 + f_1)/2$

$$H_{BP}(f) = K \left[\prod \left(\frac{f - f_c}{B} \right) + \prod \left(\frac{f + f_c}{B} \right) \right] e^{-j2\pi f t_d}$$

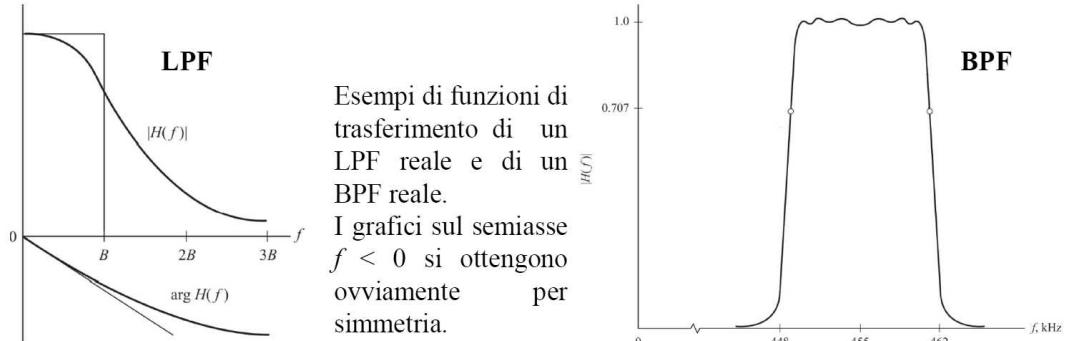
che ha per antitrasformata la risposta all'impulso

$$h_{BP}(t) = K B \operatorname{sinc}(B(t - t_d)) e^{-j2\pi f_c(t - t_d)} + K B \operatorname{sinc}(B(t - t_d)) e^{j2\pi f_c(t - t_d)}$$

Di nuovo la risposta all'impulso del filtro ideale non è causale e poichè si estende da meno infinito a più infinito il filtro **non è realizzabile**.

1.12.3 Realizzabilità dei filtri

Come visto, i filtri ideali non sono realizzabili. Possiamo però approssimarli, ottenendo quelli che si chiamano **filtri reali**: per tali filtri il modulo e la fase della funzione di trasferimento si discostano dall'andamento ideale, causando **distorsione**. Aumentando la complessità circuitale ci si può avvicinare alla risposta di un filtro ideale, al prezzo di un aumento dei costi.



I filtri più semplici da realizzare sono i passa-basso, quelli più complessi sono i passa-banda, che possono essere pensati come un passa-basso più un passa-alto in cascata.

Per i filtri passa-banda la realizzabilità dipende molto dal rapporto fra l'intervallo di frequenze che lasciano passare e la frequenza centrale. Se, con riferimento alla 1.105, definiamo $\Delta_f = f_2 - f_1$

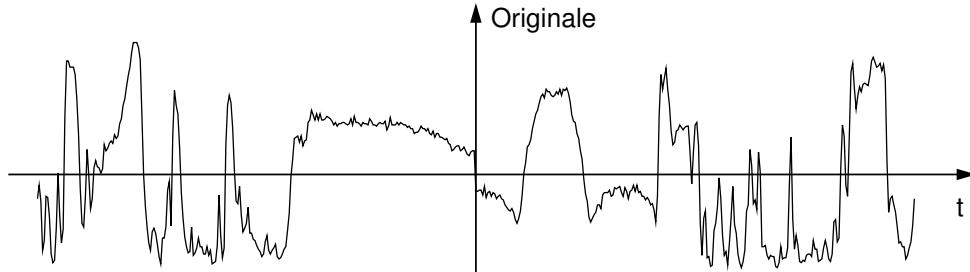
l'intervallo di frequenze e $f_c = (f_1 + f_2)/2$ la frequenza centrale, una regola empirica dice che un filtro è realizzabile se

$$\frac{\Delta_f}{f_c} \geq 0.01$$

Il termine Δ_f/f_c si definisce banda frazionaria. Un modo diverso per dire la stessa cosa è definire il fattore di merito Q come $Q = f_c/\Delta_f$ e dire che $Q \leq 100$.

1.13 Esempio: FT e filtraggio di un segnale reale

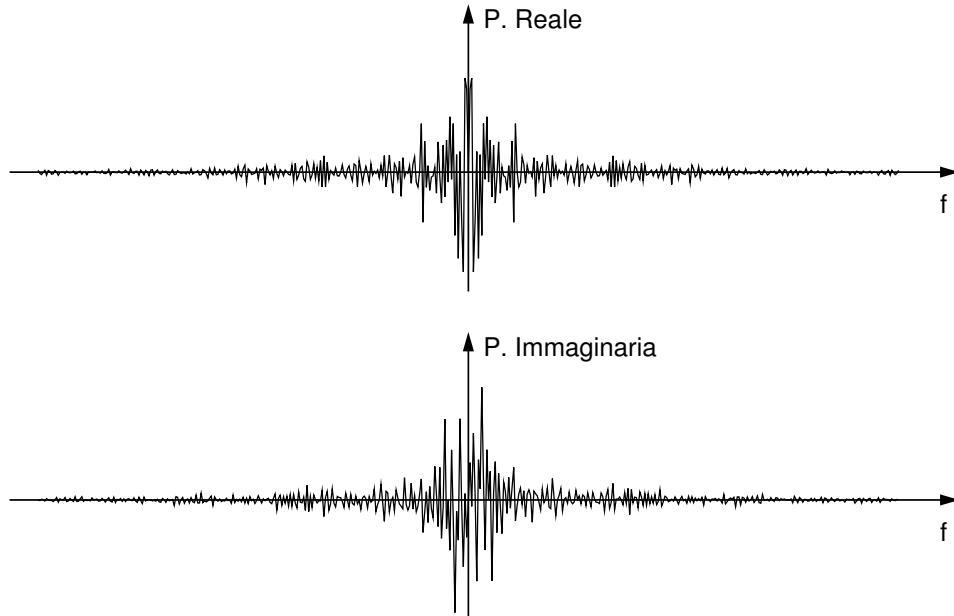
Che aspetto ha la trasformata di un segnale? Che cosa succede quando viene filtrato? Per chiamarlo consideriamo un segnale reale (sia nel senso che ha parte immaginaria nulla sia nel senso che appartiene al mondo reale). La figura che segue lo rappresenta:

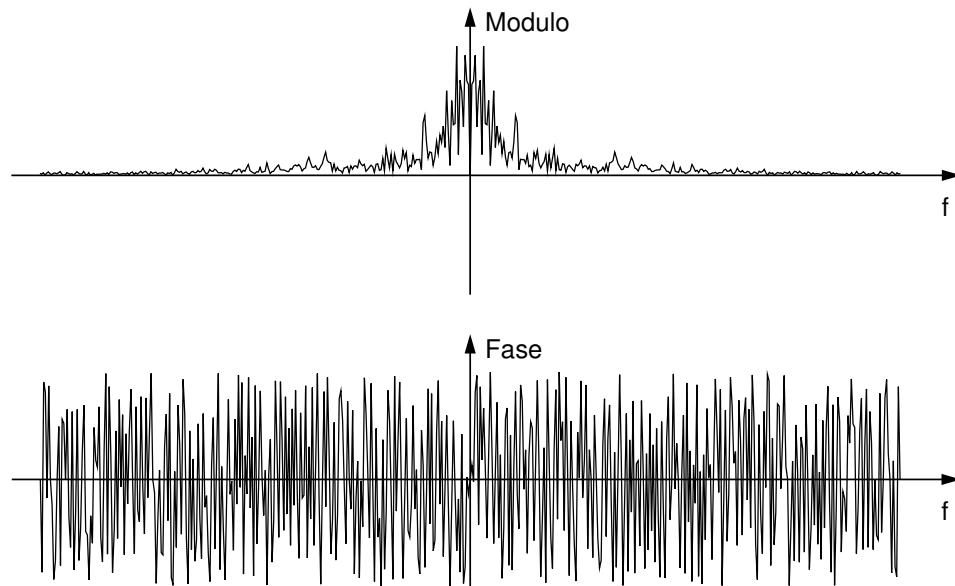


Esso è la riga 266 di una immagine (Lena512.bmp), portata a valor medio nullo. I valori massimo e minimo sono -75.7285 e 107.271 . La variabile indipendente è il tempo. Per i nostri scopi l'intervallo temporale di esistenza non è importante: volendo lo si può definire pari a $52 \mu\text{s}$, in modo che corrisponda alla durata di riga del segnale televisivo. In tal modo la figura rappresenta quanto si potrebbe vedere mediante un oscilloscopio.

1.13.1 Trasformata di Fourier

Le figure che seguono visualizzano la trasformata di Fourier del segnale sia come parte reale e immaginaria che come modulo e fase. A parte per la fase, che ha limiti $[-\pi, \pi]$, si è mantenuta una scala di ampiezza costante. I valori minimo e massimo sono: per la parte reale -5128.4 e 4837.79 , per la parte immaginaria -5788.74 e 5788.74 , per il modulo 0 e 6639.52 . Si intuisce che parte reale e modulo sono pari mentre parte immaginaria e fase sono dispari, come deve essere visto che è la FT di un segnale reale (nel senso di parte immaginaria nulla).



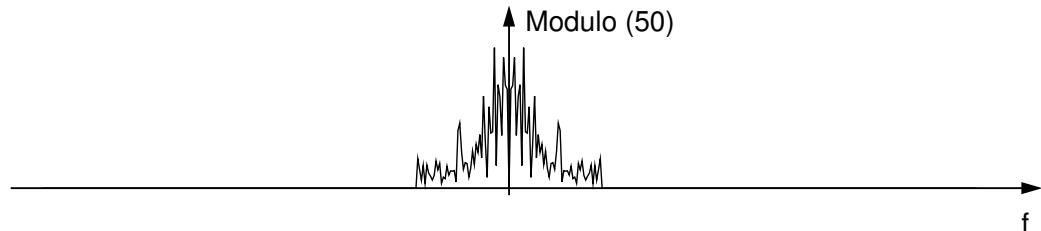


1.13.2 Filtraggio Passa Basso (LPF)

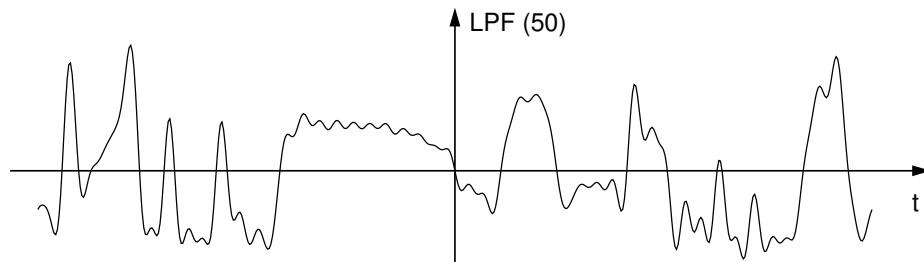
Vediamo cosa succede filtrando il segnale introdotto nella sezione precedente.

Iniziamo da un passa-basso ideale: azzeriamo (portiamo a $0 + j0$) la FT del segnale per $|f| > f_{th}$, essendo f_{th} una frequenza di soglia scelta in modo che sia circa un quinto della banda del segnale (precisamente i $100/512$).

Il suo effetto, ad esempio sul modulo, è ovviamente quello di troncarlo come in figura

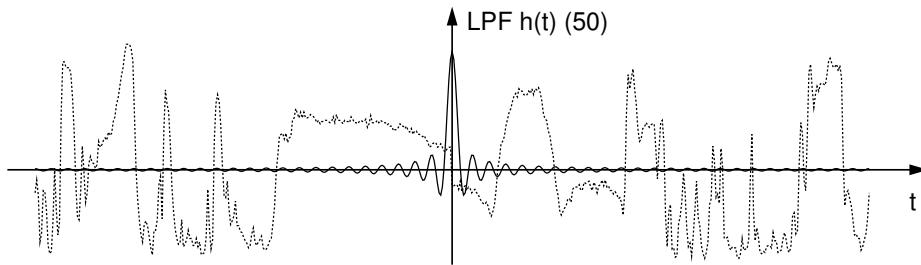


L'effetto su parte reale, parte immaginaria e fase è analogo. Il segnale in uscita al filtro è

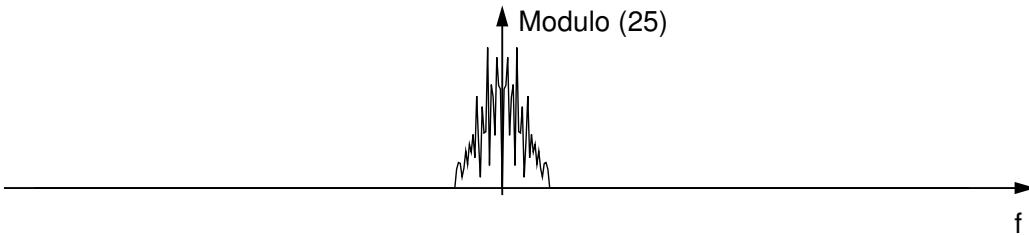


Come evidente, l'eliminazione delle alte frequenze produce un segnale con variazioni temporali più lente, quindi un segnale "più morbido".

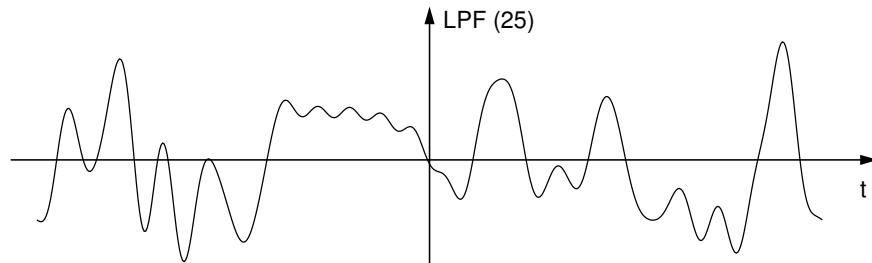
Questo esempio, oltre a mostrare l'effetto di un filtro LP ideale (rettangolare alto uno), serve anche per mostrare l'effetto "ammorbidente" della convoluzione. Infatti il filtro visto si può interpretare sia come un taglio in frequenza (la moltiplicazione per il rettangolo) che come la convoluzione nel tempo fra il segnale originale e l'antitrasformata del filtro (che è una *sinc*): quest'ultima è mostrata in figura sovrapposta al segnale originale a puntini, al fine di poter valutare le proporzioni relative.



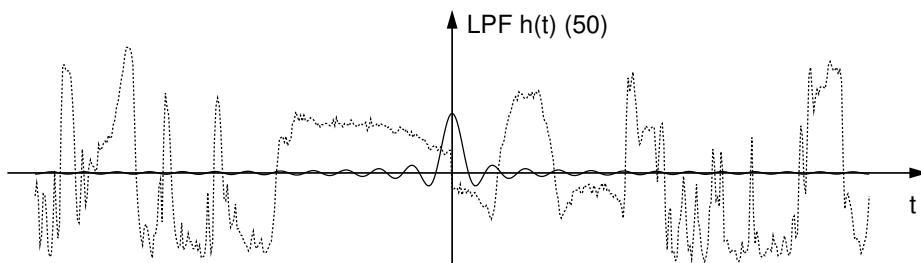
Applicando un filtro con banda inferiore, pari alla metà della banda precedente, l'effetto viene incrementato. Il modulo filtrato diventa



ed il segnale ricostruito è



Stringendo il filtro, ossia eliminando una maggiore quantità di alte frequenze, l'effetto di ammorbidente aumenta. Questo effetto si può anche interpretare nel tempo: se il filtro è più stretto l'uscita si ottiene convolvendo il segnale con una *sinc* più *larga* del caso precedente, come mostrato in figura

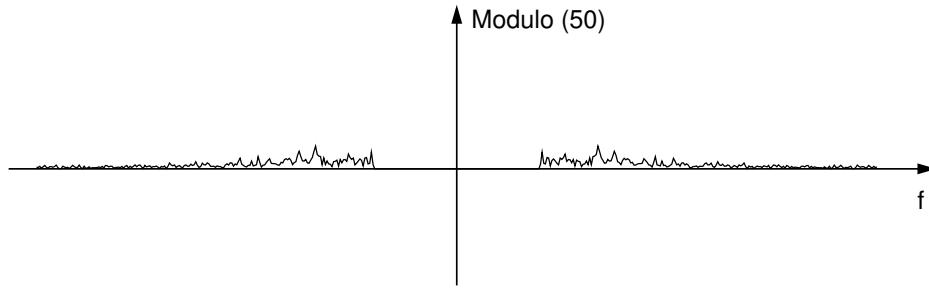


Attenzione: le due *sinc* mostrate nelle figure precedenti sono in scala fra di loro, ma NON con il segnale originale (per vedere qualcosa le ho moltiplicate entrambe per 512).

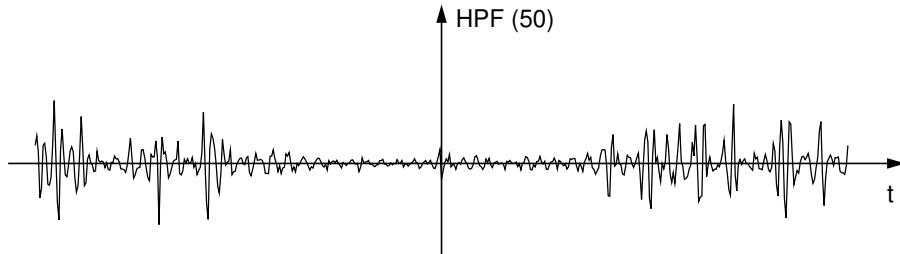
Esercizio 1.5 (da fare da soli) Perchè la *sinc* del secondo esempio è più bassa di quella del primo? Calcolate le antitrasformate dei due filtri, supponendo che il primo tagli a f_0 ed il secondo a $f_0/2$. Dimostrate che le due *sinc* hanno ampiezza in zero diversa ma stessa area. Cercate di capire perchè devono avere stessa area e perchè deve venire proprio quel valore...

1.13.3 Filtraggio Passa Alto (HPF)

Consideriamo ora un passa-alto ideale. Utilizziamo il duale del primo LPF visto: quello che lascia passare le frequenze che l'altro tagliava (e viceversa). Il suo effetto, ad esempio sul modulo della FT, è



L'antitrasformata del segnale filtrato è in figura: sono rimaste solo le alte frequenze, quindi si è perso l'andamento generale del segnale. In questo caso il segnale non è più morbido dell'originale, ma ha variazioni di rapidità comparabile. Se sommo l'uscita di questo filtro con quella del LPF duale ritrovo *esattamente* il segnale di partenza (come ovvio!).



L'antitrasformata del filtro si ottiene pensando il filtro HP come uno meno il filtro ideale LP corrispondente: quindi è un delta di Dirac nell'origine meno la *sinc* già vista.

Sinc e convoluzione*. Già che siamo in argomento ne aprofitto per farvi notare una proprietà della *sinc* poco nota. Consideriamo un filtro LP ideale, alto uno per $|f| \leq f_0$ e zero altrove: la sua antitrasformata vale $f_0 sinc(f_0 t)$. Consideriamo ora il segnale $s(t) = f_0 sinc(f_0 t)$, ovvero la sinc di prima, e filtriамola proprio col filtro LP visto. In frequenza significa moltiplicare per il rettangolo alto uno in $|f| \leq f_0$, cosa che lascia ovviamente invariata la FT di $s(t)$. Quindi anche nel tempo $s(t)$ non deve variare. Moltiplicare in frequenza per un rettangolo provoca tuttavia la convoluzione del segnale nel tempo con una *sinc*: l'unico modo per non avere contraddizione è notare come la convoluzione di una *sinc* (centrata in zero e di altezza opportuna) con se stessa non la modifichi. Ovvero $s(t) * s(t) = s(t)$! Questo vale solo per *sinc* ad area unitaria, in caso contrario compare anche un fattore di scala (una costante moltiplicativa). Se non sono centrate in zero il risultato è ancora la stessa *sinc*, ma traslata.

1.14 Sistemi senza memoria

I sistemi lineari LTI visti in precedenza sono, come detto, una sottoclasse importante di tutti i sistemi possibili. Esiste almeno un'altra classe di sistemi comunemente usati: i sistemi senza memoria. Considerando i sistemi LTI ci si può rendere conto che il valore dell'uscita in un certo istante dipende dai valori della funzione di ingresso in infiniti istanti. Se il sistema è causale la dipendenza riguarda solo i valori passati (oltre a quello presente), se il sistema non è causale l'uscita dipende anche dai valori dell'ingresso futuri.

I sistemi LPF definiti, ad esempio, in sezione 1.13.2 hanno risposta all'impulso di tipo **sinc**, ovvero infinita (sia verso il passato che verso il futuro). Il valore dell'uscita in un certo istante dipende quindi da tutti i valori, passati e futuri, della funzione in ingresso: l'uscita in un certo istante t_0 vale infatti (vedi eq. 1.22)

$$y(t_0) = \int_{-\infty}^{\infty} x(\tau) h(t_0 - \tau) d\tau$$

e se la risposta all'impulso $h(t)$ ha durata infinita (come ha una sinc) per calcolare $y(t_0)$ bisogna conoscere $x(t)$ per tutti i t . Se $h(t)$ ha una durata limitata l'uscita in un certo istante dipende dai valori dell'ingresso in un intervallo di pari durata. Quindi i sistemi lineari hanno memoria, visto che l'uscita dipende dalla "storia" dell'ingresso. I soli sistemi lineari senza memoria sono quelli che hanno risposta all'impulso di durata nulla, ossia del tipo $h(t) = a \delta(t)$: per tali sistemi la relazione ingresso-uscita è $y(t) = a x(t)$, ovvero l'uscita in un certo istante è uguale all'ingresso nello stesso istante moltiplicato per il valore a . Tali sistemi sono detti amplificatori (o attenuatori) ideali.

Un generico sistema senza memoria, lineare o non lineare, è caratterizzato dalla relazione ingresso-uscita: in realtà tutti i sistemi possono essere caratterizzati da tale relazione, per quelli senza memoria essa assume una forma particolare. Se il sistema è tempo invariante la relazione ingresso-uscita si può scrivere come

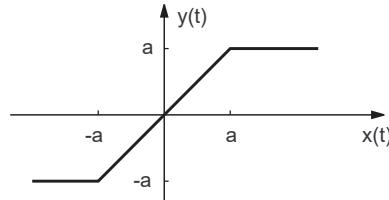
$$y(t) = g[x(t)] \quad (1.110)$$

essendo g una qualsiasi funzione reale. Se, ad esempio, $g(\alpha) = \alpha^2$ il corrispondente sistema è un quadratore e l'uscita si ottiene mediante $y(t) = x^2(t)$.

Se il sistema non è tempo invariante la 1.110 vale ancora, ma la funzione g è diversa in ogni istante (tali sistemi sono di scarsa utilità).

Da notare che nell'equazione 1.110 il tempo gioca il ruolo di un parametro, non della variabile indipendente: l'argomento della funzione g non è il tempo, sono i valori che assume la funzione $x(t)$. Infatti in un fissato istante t_0 la funzione $x(t_0)$ assume un qualche valore numerico x_0 : l'uscita del sistema si ottiene trasformando tale valore numerico mediante la funzione g , $y(t_0) = g(x_0)$.

Per questo motivo la relazione ingresso-uscita si può indicare mediante un grafico che abbia sulle ordinate $y(t)$ e sulle ascisse $x(t)$. Ad esempio un sistema non lineare senza memoria può essere definito da



Questo è un esempio di *soft-limiter*: per valori dell'ingresso fra $-a$ ed a l'uscita è uguale all'ingresso, per valori fuori intervallo l'uscita viene limitata. Il soft-limiter garantisce che $|y(t)| \leq a$.

Sapreste dimostrare che non è lineare?

Capitolo 2

Campionamento e quantizzazione

Nel capitolo precedente abbiamo trattato segnali continui sia nel tempo che in ampiezza, ovvero segnali analogici. Sia nelle telecomunicazioni che nell'elaborazione, analisi e sintesi dei segnali è di fondamentale importanza poter trattare segnali che siano a tempo discreto ed ampiezza discreta, ovvero segnali numerici (o digitali). I principali vantaggi della rappresentazione numerica sono

- la possibilità di eseguire le elaborazioni mediante algoritmi numerici per mezzo di elaboratori elettronici, ossia computer, sia di uso generale che dedicati (come i DSP, *digital signal processor*);
- poter raggiungere una migliore resistenza ai disturbi nelle trasmissioni, sia per la intrinseca maggior robustezza dei segnali numerici che per la possibilità di utilizzare codici per la protezione dagli errori;
- la possibilità di trasmettere molti segnali numerici, anche di tipo diverso (voce, immagini, dati, filmati), sullo stesso canale mediante tecniche di multiplazione.

Tutti i segnali del mondo reale sono peraltro analogici, serve quindi un modo per poterli convertire in formato numerico. Idealmente si vorrebbe poter trasformare un segnale analogico in forma numerica avendo la garanzia che non vi sia perdita di informazione, ovvero che dal segnale numerico si possa ritornare *esattamente* a quello analogico. La trasformazione da analogico a numerico avviene in due passi: la discretizzazione nel tempo (detta anche *campionamento*) e la discretizzazione in ampiezza (*quantizzazione*). Vedremo che, sotto certe condizioni, la prima trasformazione si può effettivamente realizzare senza alcuna perdita di informazione, seguendo le indicazioni del teorema del campionamento. Il passaggio alle ampiezze discrete porta invece inevitabilmente ad una perdita di informazione: tale perdita può essere resa piccola quanto si vuole ed in molti casi è irrilevante. Si pensi, ad esempio, al confronto fra i vecchi dischi in vinile (analogici) ed i nuovi CD, nei quali il suono è memorizzato in forma numerica senza che vi sia perdita di qualità percepita (per la maggior parte degli utilizzatori la qualità è migliorata, anche se vi sono alcuni dissidenti). La rappresentazione numerica dei segnali ha anche portato ad una ridefinizione del concetto di copia: un segnale numerico può essere copiato, dalla copia si può ottenere una nuova copia e così via per un numero illimitato di volte e tutte le copie sono *esattamente* uguali all'originale, mentre un segnale analogico ha una degradazione ad ogni copia. Le implicazioni di questo sono notevoli, si pensi ad esempio ai problemi di copyright, ma non riguardano questo corso.

2.1 Campionamento

Sia $x(t)$ un segnale analogico continuo definito su tutto l'asse reale. L'operazione del campionamento permette di ottenere un segnale $x[n]$ a tempo discreto, o sequenza di numeri, definito da

$$x[n] = x(nT_c) \quad (2.1)$$

Dove n è ovviamente intero e assume valori da meno infinito a più infinito. Il segnale discreto $x[n]$ assume quindi i valori del segnale $x(t)$ negli istanti nT_c , detti istanti di campionamento. I valori (numeri) $x(nT_c)$ vengono chiamati *campioni* del segnale $x(t)$, mentre T_c è il *periodo di campionamento* e $f_c = 1/T_c$ è la *frequenza di campionamento*.

La domanda che ci si pone è la seguente: è possibile ritornare a $x(t)$ (cioè al segnale analogico) partendo dalla sola conoscenza dei suoi campioni? La risposta è sì, ma solo per una certa classe di segnali, quelli limitati in banda, e se il periodo di campionamento è sufficientemente piccolo.

2.1.1 Teorema del campionamento

Il teorema del campionamento è noto anche come **teorema di Nyquist-Shannon**.

Per dimostrare quanto asserito nella sezione precedente vi sono diverse strade. Vediamo un modo basato sull'interpretazione spettrale del campionamento ideale: più avanti, dopo le necessarie premesse, analizzeremo diverse e più rigorose dimostrazioni.

A partire dai campioni di $x(t)$ costruiamo il segnale

$$x_c(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_c) = \sum_{n=-\infty}^{\infty} x(nT_c) \delta(t - nT_c) \quad (2.2)$$

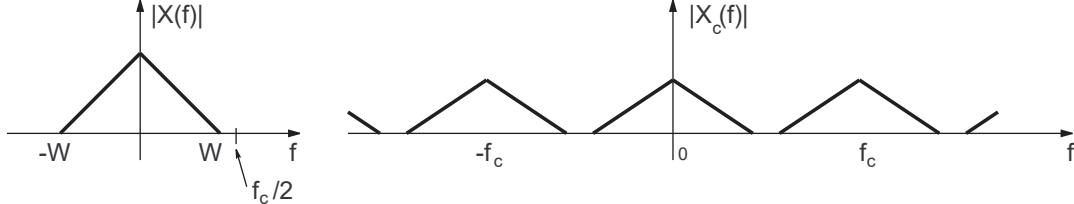
Il segnale $x_c(t)$ è detto il *campionamento ideale* di $x(t)$: ovviamente $x_c(t)$ è definito sull'intero asse t reale ed è analogico, anche se usa le funzioni delta di Dirac e quindi irrealizzabile in pratica. Il concetto fondamentale è che $x_c(t)$ contiene solo l'informazione dei campioni di $x(t)$.

La trasformata di Fourier di $x_c(t)$ si ricava facilmente per mezzo di quanto visto nel capitolo 1:

$$F[x_c(t)] = F \left[x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_c) \right] = X(f) * \frac{1}{T_c} \sum_{k=-\infty}^{\infty} \delta \left(f - \frac{k}{T_c} \right) = \frac{1}{T_c} \sum_{k=-\infty}^{\infty} X \left(f - \frac{k}{T_c} \right) \quad (2.3)$$

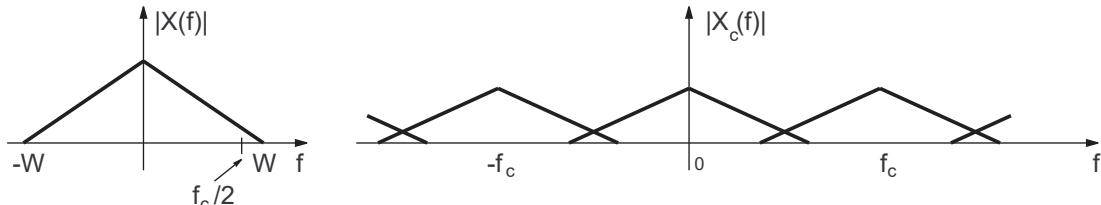
dove $X(f)$ è la trasformata di Fourier di $x(t)$ (che abbiamo assunto esista) e si sono usate le proprietà definite dalle eq. 1.58 e 1.92.

La trasformata di Fourier di $x_c(t)$ è quindi composta da infinite repliche di $X(f)$ equispaziate in frequenza a passo $f_c = 1/T_c$ e scalate di $1/T_c$: vediamo l'interpretazione grafica di questo risultato. Supponiamo che $x(t)$ sia un segnale in banda base con larghezza di banda W , ovvero che $X(f) = 0$ per $|f| \geq W$, e supponiamo inoltre che $f_c/2 > W$: la situazione risultante è



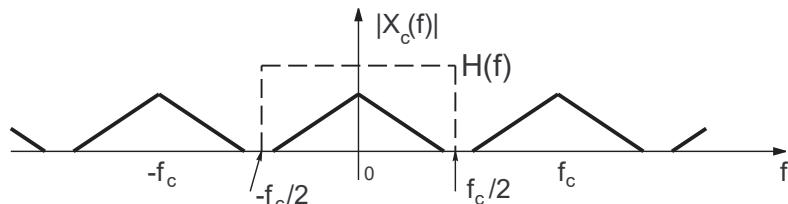
Come si vede, le repliche di $X(f)$ non sono sovrapposte: posso ritornare a $x(t)$ "isolando" la replica in banda base (cioè quella centrata sull'origine delle frequenze) e moltiplicando per T_c .

Se invece il segnale $x(t)$ non ha banda limitata, oppure se $f_c/2 < W$, la situazione è



Le repliche di $X(f)$ sono sovrapposte: negli intervalli in f dove c'è sovrapposizione si sommano le code delle diverse repliche. In tali condizioni non esiste la possibilità di isolare una sola replica e la ricostruzione non è possibile. La sovrapposizione delle repliche si chiama **aliasing**: questo fenomeno si verifica quando la frequenza di campionamento non è sufficiente e consiste nel "riplegamento" delle alte frequenze verso le basse. Vedremo più avanti esempi.

Se $x(t)$ è limitato in banda, cioè se $X(f) = 0$ per $|f| \geq W$, possiamo ricostruirlo a partire da $x_c(t)$ filtrando quest'ultimo con un filtro in frequenza $H_{LP}(f)$ che sia alto T_c per $|f| \leq f_c/2$ e zero altrove, come schematizzato in figura, purché valga la relazione $f_c \geq 2W$.

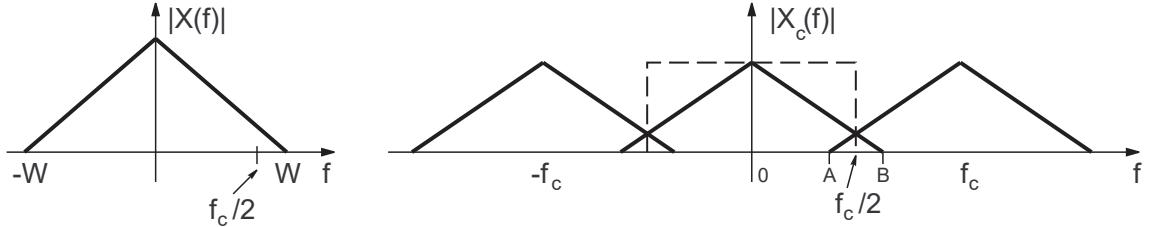


L'enunciato precedente è il **teorema del campionamento** o teorema di *Nyquist-Shannon*, e la condizione $f_c \geq 2W$ è nota come *condizione di Nyquist*.

2.1.2 Filtro anti-aliasing

I segnali del mondo reale possono avere durata breve o lunga, ma mai infinita. In base a quanto visto essi non sono quindi limitati in frequenza, ovvero hanno banda teorica infinita. Di fatto tutti i segnali si attenuano in frequenza, diventando sempre più piccoli al crescere di f : a partire da un certo valore di f si possono considerare trascurabili, dove ovviamente la definizione di cosa sia trascurabile dipende dai parametri di progetto.

Vediamo un esempio: il segnale vocale, teoricamente infinito, non ha praticamente energia per $|f| > 20$ KHz. Se lo ascoltate sulle radio in FM (quelle che trasmettono fra 88 e 108 MHz) sembra perfetto, anche se è stato filtrato a 15 KHz (che è la banda per tali trasmissioni). Se ascoltate le radio in AM (intorno a 1000 KHz), che hanno banda di 9 KHz, la qualità del segnale vocale è solo lievemente inferiore (bisogna farci caso per accorgersene). Se ascoltate una telefonata la qualità della voce risulta nettamente inferiore, anche se ancora del tutto comprensibile. Per le trasmissioni telefoniche la banda del segnale vocale viene infatti limitata all'intervallo fra 300 e 3400 Hz, intervallo che fu deciso dopo numerose prove come compromesso fra banda occupata e qualità avendo come obiettivo non la perfetta riproduzione (che infatti non c'è) ma la perfetta comprensibilità. Nel caso della telefonia numerica viene usata una frequenza di campionamento di 8 KHz, ovvero abbondante rispetto alla banda per lasciare margine ai filtri. Il problema è che se si campiona il segnale vocale a 8 KHz si ottiene una valanga di aliasing, a meno che non lo si filtri PRIMA del campionamento con un filtro passa-banda fra 300 e 3400 Hz. Chiariamo meglio il concetto: nella figura che segue, a sinistra, c'è un segnale con banda W maggiore della metà della frequenza di campionamento.



Se la ricostruzione avviene con il filtro indicato con linea tratteggiata nella parte destra della figura, ovvero mediante il filtro teorico di ricostruzione, ho sovrapposizione delle repliche nell'intervallo $[A, B]$. Il segnale ricostruito, che ha banda solo fino a $f_c/2$, è diverso dall'originale per due motivi: **1)** mancano le componenti fra $f_c/2$ e $B = W$ e **2)** la parte fra A e $f_c/2$ viene modificata dalla somma con la parte iniziale della replica successiva. Il primo effetto è ineliminabile, visto che campionando a f_c non posso avere frequenze maggiori di $f_c/2$, mentre il secondo effetto può essere eliminato se si filtra il segnale di partenza, prima del campionamento, mediante un filtro passa-basso di banda $f_c/2$. Questo filtro, sempre presente nei sistemi che campionano, si chiama filtro anti-aliasing.

2.1.3 Ricostruzione

Il teorema del campionamento non solo ci assicura che è possibile ricostruire esattamente un segnale campionato correttamente, ma ci indica anche come tale ricostruzione deve avvenire: filtrando il segnale campionato idealmente $x_c(t)$ (definito in 2.2) col filtro di ricostruzione

$$H_R(f) = T_c \prod \left(\frac{f}{f_c} \right) \quad (2.4)$$

La corrispondente risposta all'impulso è

$$h_R(t) = \text{sinc}(f_c t) = \text{sinc}(t/T_c) = \frac{\sin(\pi f_c t)}{\pi f_c t} \quad (2.5)$$

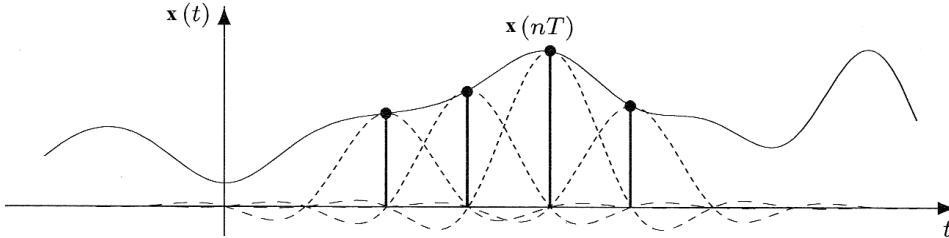
Il segnale $x(t)$ si può quindi ricostruire dai suoi campioni, nel tempo o in frequenza, mediante

$$X(f) = X_c(f) \cdot H_R(f) \quad x(t) = x_c(t) * h_R(t) \quad (2.6)$$

Espandiamo la ricostruzione nel tempo

$$x(t) = \text{sinc}(t/T_c) * \sum_{n=-\infty}^{\infty} x(nT_c) \delta(t - nT_c) = \sum_{n=-\infty}^{\infty} x(nT_c) \text{sinc}\left(\frac{t - nT_c}{T_c}\right) \quad (2.7)$$

L'ultima sommatoria dell'equazione precedente indica che il segnale $x(t)$ si può ottenere come somma di infinite *sinc* traslate a passi del periodo di campionamento T_c e scalate del valore del corrispondente campione. Ciascuna funzione $x(nT_c) \text{sinc}\left(\frac{t-nT_c}{T_c}\right)$ vale $x(nT_c)$ in $t = nT_c$ e vale zero in tutti gli altri istanti di campionamento $t = kT_c$ con $k \neq n$. In questo modo è evidente che i valori del segnale ricostruito negli istanti di campionamento sono proprio i valori dei campioni. I valori fra gli istanti di campionamento si ottengono dalla somma di *infinite sinc*, come schematizzato in figura.

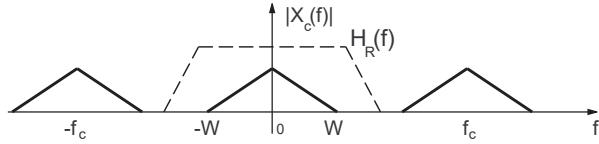


Si tratta, come evidente, di una operazione di *interpolazione* fra i valori dei campioni: per questo motivo $H_r(f)$ è detto filtro interpolatore e la sinc è detta funzione interpolante.

Da notare che quanto esposto è teorico: il filtro $H_R(f)$ è non causale e la relativa risposta all'impulso ha durata infinita. La realizzabilità pratica della ricostruzione passa dall'accettazione di alcuni compromessi...

Banda di guardia. Il filtro di ricostruzione introdotto nella 2.4 è ideale e quindi irrealizzabile esattamente. Un aiuto per realizzare la ricostruzione dai campioni si ottiene se si accetta di campionare ad una frequenza superiore alla frequenza limite di Nyquist.

In questo caso, schematizzato in figura, il filtro di ricostruzione deve essere piatto fra $-W$ e W ma ha più spazio per andare a zero: invece dei fronti ripidi di un rettangolo può assumere un andamento più morbido.



La corrispondente risposta all'impulso è ancora oscillatoria ed infinita, ma le code si smorzano più rapidamente (dopo poche oscillazioni diventano trascurabili). Troncando tale risposta e traslando l'uscita nel tempo (cosa che introduce un ritardo) il filtro può essere reso causale e quindi realizzabile. Lo spazio che si crea fra le repliche aumentando la frequenza di campionamento si chiama *banda di guardia*.

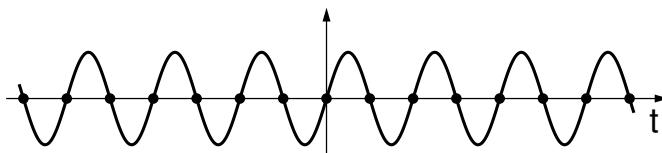
Come esempio possiamo citare ancora il campionamento utilizzato nei CD musicali. A fronte di una massima frequenza teorica di 20 KHz si utilizza una frequenza di campionamento di 44 KHz, lasciando quindi 4 KHz di banda di guardia.

2.1.4 Esempio: campionare una sinusoida

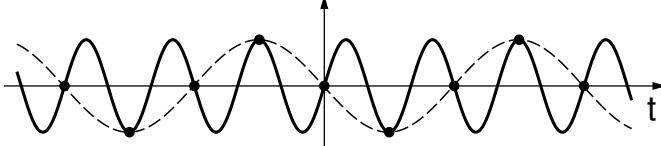
Il campionamento di un segnale sinusoidale (un tono puro) a frequenza f_0 fissata richiede un minimo di attenzione. Una sinusode generale è definita da $s(t) = \cos(2\pi f_0 t + \theta)$. La sua FT, come visto in eq. 1.83, vale

$$\mathcal{F}[\cos(2\pi f_0 t + \theta)] = \frac{1}{2} e^{j\theta} \delta(f - f_0) + \frac{1}{2} e^{-j\theta} \delta(f + f_0)$$

ovvero è composta da due impulsi centrati in $\pm f_0$. Se diciamo che la sua banda è f_0 possiamo campionarla con frequenza di campionamento $f_c = 2f_0$. Sembra ragionevole, ma può non funzionare. Nella figura sotto un seno viene campionato ogni $T_c = T_0/2$, essendo $T_0 = 1/f_0$, ovvero con $f_c = 2f_0$: come evidente tutti i campioni sono nulli, quindi il segnale ricostruito è identicamente zero! Cosa è successo? Il fatto è che la definizione di banda W di un segnale in banda base è $|X(f)| = 0$ per $|f| \geq W$, quindi un seno, che ha un impulso in f_0 , ha banda *maggior*e di f_0 . La banda del seno è $W = f_0 + \epsilon$, con ϵ positivo piccolo a piacere, ma non nullo. Questo fatto è da tenere in considerazione tutte le volte che bisogna campionare un segnale che ha un impulso in frequenza al limite della banda. Interessante anche l'interpretazione frequenziale: la periodizzazione della FT del seno a multipli di $2f_0$ causa la cancellazione degli impulsi (siete invitati a verificarlo da soli, partendo dall'eq. 1.82).



Aliasing. Lo studio del campionamento di una sinusoide si presta bene per un approfondimento del concetto di aliasing. Supponiamo di avere il segnale $x(t) = \sin(2\pi f_0 t)$, ovvero un seno a frequenza f_0 . Secondo il teorema del campionamento la frequenza minima di campionamento deve essere appena maggiore di $2f_0$. Vediamo cosa succede sottocampionando a $f_c = \frac{4}{3}f_0$, ovvero usando un periodo di campionamento $T_c = \frac{3}{4f_0} = \frac{3}{4}T_0$. La figura seguente mostra dove cadono i campioni: il segnale ricostruito (tratteggiato) è una sinusoide a frequenza $f_0/3$.



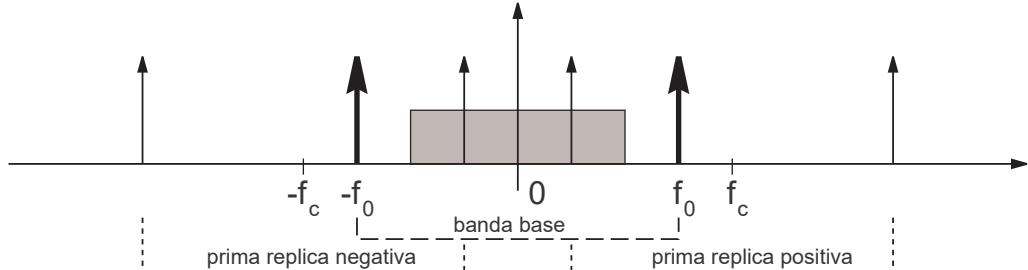
Analizziamo cosa è successo nel dominio delle frequenze: la FT del seno è

$$\mathcal{F}[\sin(2\pi f_0 t)] = \frac{1}{2j} [\delta(f - f_0) - \delta(f + f_0)]$$

I due impulsi che formano il modulo della FT sono indicati in figura con tratto più marcato. La prima replica positiva si ottiene traslando tali impulsi in $f_c = (4/3)f_0$, il risultato sono due impulsi in $f_c - f_0 = (1/3)f_0$ e $f_c + f_0 = (7/3)f_0$. In figura è mostrata anche la prima replica negativa. Il filtro di ricostruzione lascia passare le frequenze fra $-f_c/2$ e $f_c/2$: in figura è il rettangolo grigio. Come evidente sopravvivono solo due impulsi centrati in $\pm f_0/3$, che rappresentano appunto una sinusoide a frequenza un terzo dell'originale. La fase di tale sinusoide (come evidente dalla figura nel tempo è un seno sfasato di 180°) si può ricavare considerando anche la fase degli impulsi (e non solo il modulo): l'impulso in $f_0/3$ è moltiplicato per $-1/(2j)$, visto che deriva dall'impulso a frequenze negative, mentre quello in $-f_0/3$ è moltiplicato per $1/(2j)$. Il segnale ricostruito in frequenza è allora

$$X_r(f) = \frac{-1}{2j} \left[\delta(f - \frac{f_0}{3}) - \delta(f + \frac{f_0}{3}) \right]$$

per cui $x_r(t) = -\sin(\frac{f_0}{3}t)$.



Quanto visto dimostra l'effetto dell'aliasing, che "ribalta" le alte frequenze attorno a $f_c/2$, portandole sulle basse frequenze.

Questo si può interpretare anche come segue: campionando un segnale con frequenza $f_c = 1/T_c$ si ottengono, ovviamente, dei campioni. Vi sono infinite funzioni nel tempo che passano dai campioni trovati. La ricostruzione "alla Nyquist" sceglie, fra queste infinite funzioni, quella che ha banda minima. Tale banda può anche essere inferiore a $f_c/2$, ma è comunque garantito che la banda del segnale ricostruito sia minore o uguale a $f_c/2$. Se ho campionato una funzione con banda inferiore o uguale a $f_c/2$ la ricostruzione dai campioni permette di ottenere nuovamente e senza errori tale funzione. Se ho campionato non rispettando tale condizione, il segnale ricostruito sarà diverso da quello campionato, ovvero sarà un'altra funzione. Il nuovo segnale ricostruito sarà, in generale, diverso da quello originale anche nella banda di frequenza fra zero e $f_c/2$, visto che troncando il segnale originale in frequenza ottengo qualcosa che non passa più dai campioni del segnale non troncato.

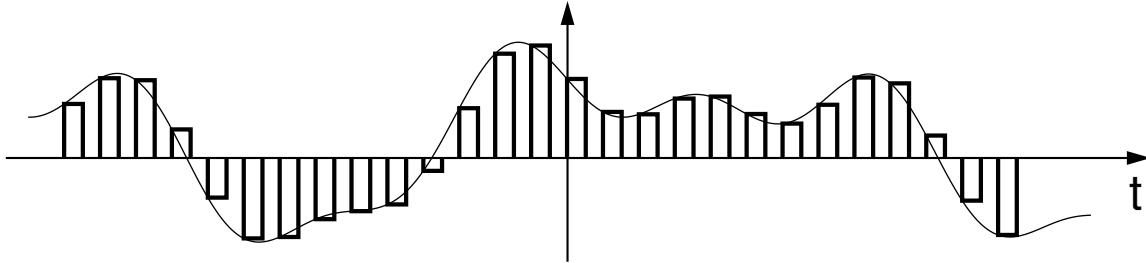
2.1.5 Sample & hold

Finora abbiamo visto il campionamento ideale. Quello reale differisce in tre ovvi aspetti:

1. il segnale campionato consiste di impulsi aventi durata ed ampiezza finite, anziché essere delta di Dirac a durata infinitesima ed ampiezza infinita;
2. i filtri di ricostruzione reali non sono filtri ideali;

3. i segnali da campionare hanno durata finita e quindi banda teoricamente infinita.

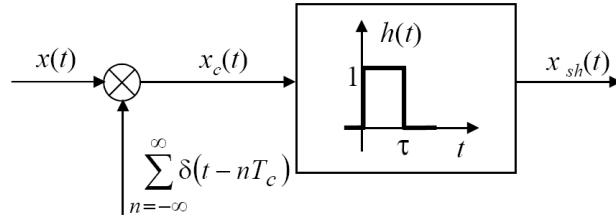
Uno dei sistemi di campionamento reali è il *sample & hold*, o campionamento a tenuta. Il valore campionato $x(nT_c)$ viene mantenuto costante per un tempo $\tau \leq T_c$, poi torna a zero (vedi figura).



Esistono circuiti che realizzano direttamente questo tipo di campionamento, ma per comodità di analisi si può pensare come ottenuto da un campionamento ideale (eq. 2.2) seguito dalla convoluzione con la funzione rettangolare

$$x_{sh}(t) = x_c(t) * h(t) = \left[x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_c) \right] * \prod \left(\frac{t - \tau/2}{\tau} \right) \quad (2.8)$$

ovvero secondo lo schema seguente (da notare il ritardo nel rettangolo, inserito affinché il sistema sia causale)



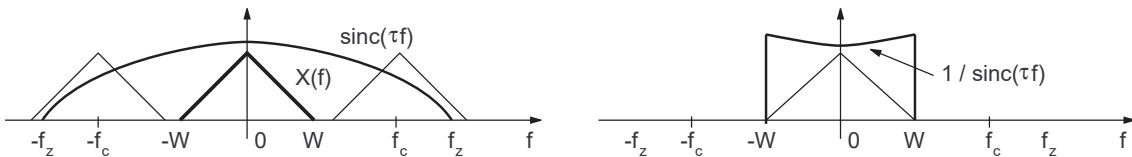
La FT di $x_{sh}(t)$ vale

$$X_{sh}(f) = X_c(f) \cdot H(f) = \frac{\tau}{T_c} \left[\sum_{k=-\infty}^{\infty} X(f - \frac{k}{T_c}) \right] \text{sinc}(\tau f) e^{-j\pi f \tau} \quad (2.9)$$

Questa FT è simile a quella di $X_c(f)$ (eq. 2.3) con le seguenti differenze:

- i) un fattore di scala τ ;
- ii) un ritardo temporale pari a $\tau/2$;
- iii) la moltiplicazione per la *sinc*.

Le prime due sono evidentemente irrilevanti ai fini della perfetta ricostruzione (purchè τ sia maggiore di zero). Il prodotto con la *sinc* è invece potenzialmente pericoloso, dato che provoca una distorsione dello spettro. Il primo zero della *sinc* è in $\pm f_z = 1/\tau$ e poichè $\tau \leq T_c$ segue che $f_z \geq 1/T_c = f_c$. Il segnale, se campionato correttamente, ha banda $W \leq f_c/2$, quindi lo zero è oltre la fine del segnale, come mostrato in figura (a sinistra)



La distorsione dello spettro può allora essere corretta mediante un filtro inverso nella banda del segnale, come in figura (a destra), ovvero usando, come filtro di ricostruzione

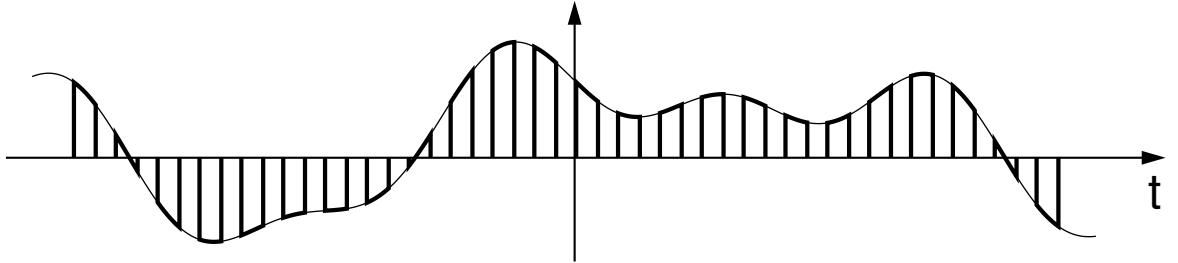
$$H_r(f) = \frac{T_c}{\tau \text{sinc}(\tau f)} \quad \text{in} \quad |f| \leq \frac{f_c}{2} \quad (2.10)$$

e zero altrove. Se vi fossero zeri nella banda del segnale non sarebbe possibile definire il filtro inverso, visto che comparirebbero divisioni per zero.

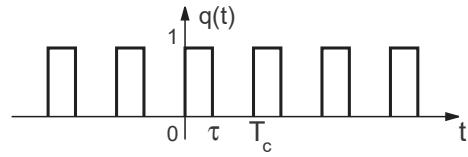
Il segnale ricostruito è $X_{SH}(f) \cdot H_r(f) = X(f) e^{-j\pi f \tau}$ e, nel tempo, $x(t - \tau/2)$, cioè il segnale originale ritardato di $\tau/2$. Da notare che al diminuire di τ , la *sinc* si allarga (in frequenza) ed il filtro di ricostruzione tende a diventare piatto. Per $\tau = T_c$ la *sinc* in frequenza è tale che il suo primo zero cade in $f_c = 1/T_c$.

2.1.6 Chopper sampling

Un diverso sistema reale di campionamento è il *chopper sampling*. Questo prevede, come mostrato nella figura seguente, di ottenere un segnale campionato $x_{ch}(t)$ che sia uguale al segnale originale $x(t)$ per un periodo $\tau < T_c$ in ogni intervallo T_c .



Per comodità di analisi si può pensare che il segnale $x_{ch}(t)$ sia ottenuto moltiplicando il segnale originale per una serie di rettangolini alti uno, ovvero per il segnale $q(t)$ in figura



Il segnale $q(t)$ si può scrivere come la convoluzione di un rettangolo con un treno di impulsi

$$q(t) = \prod \left(\frac{t - \tau/2}{\tau} \right) * \left[\sum_{n=-\infty}^{\infty} \delta(t - nT_c) \right] \quad (2.11)$$

e quindi

$$x_{ch}(t) = x(t) \cdot q(t) = x(t) \left[\prod \left(\frac{t - \tau/2}{\tau} \right) * \sum_{n=-\infty}^{\infty} \delta(t - nT_c) \right] \quad (2.12)$$

in frequenza si ottiene

$$\begin{aligned} X_{ch}(f) &= X(f) * Q(f) = X(f) * \left[\frac{\tau}{T_c} \text{sinc}(\tau f) e^{-j\pi f \tau} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T_c}\right) \right] = \\ &= X(f) * \left[\frac{\tau}{T_c} \sum_{k=-\infty}^{\infty} \text{sinc}\left(\frac{k\tau}{T_c}\right) e^{-j\pi \frac{k}{T_c} \tau} \delta\left(f - \frac{k}{T_c}\right) \right] = \\ &= \frac{\tau}{T_c} \sum_{k=-\infty}^{\infty} \text{sinc}\left(\frac{k\tau}{T_c}\right) e^{-j\pi \frac{k}{T_c} \tau} X\left(f - \frac{k}{T_c}\right) = \frac{\tau}{T_c} \sum_{k=-\infty}^{\infty} S_k X\left(f - \frac{k}{T_c}\right) \end{aligned} \quad (2.13)$$

Il coefficiente S_k , a τ fissato, è un numero complesso (non dipende da f). Quindi la FT di $x_{ch}(t)$ contiene infinite repliche di $X(f)$ centrate, al solito, in multipli di $1/T_c$ e scalate ciascuna di un valore S_k diverso.

La ricostruzione è allora possibile utilizzando un filtro che vale T_c/τ per $|f| < f_c/2$ e zero altrove, ovviamente se f_c rispetta la condizione di Nyquist.

Come esercizio lo studente è invitato a verificare che se $\tau = T_c$ si ottiene $X_{ch}(f) = X(f)$.

2.1.7 Teorema della dimensionalità

Una versione più generale del teorema del campionamento è il cosiddetto teorema della dimensionalità. Esso afferma che, quando il prodotto BT_0 è grande, un segnale reale può essere completamente specificato da

$$N = 2BT_0 \quad (2.14)$$

informazioni indipendenti. Il numero N è il numero delle dimensioni richieste per specificare il segnale (solitamente numeri reali), T_0 è la durata del segnale in secondi (l'intervallo temporale su cui è osservato) e B è la banda del segnale in Hz.

Da questo teorema discende il teorema del campionamento: se abbiamo un segnale di banda B osservato per T_0 secondi e vogliamo salvare i suoi campioni su di un calcolatore per poterlo analizzare o elaborare bisogna salvare almeno $2BT_0$ campioni, ovvero campionare ad una frequenza $f_c = N/T_0 \geq 2B$.

Le implicazioni del teorema della dimensionalità sono tuttavia più ampie: ad esempio ci dice che se un canale di trasmissione ha banda limitata, l'informazione che può trasportare è proporzionale al prodotto fra banda del sistema e tempo impiegato per la trasmissione.

2.2 Quantizzazione

La quantizzazione di un segnale analogico $s(t)$ è un'operazione non lineare irreversibile che consiste nel discretizzare le ampiezze di $s(t)$ ad un numero L prefissato e finito di valori possibili. Molto spesso si sceglie L fra le potenze del due, ovvero $L = 2^m$, per ragioni di efficienza quando si usano numeri binari. Supponiamo che $s(t)$ possa assumere valori nell'intervallo $[-V, V]$ in modo continuo e che si voglia quantizzare usando L livelli, ovvero usando l'insieme finito di valori, o alfabeto di quantizzazione $\mathcal{Q} = \{Q_0, Q_1, \dots, Q_{L-1}\}$. A ciascun livello di quantizzazione Q_i , $i = 0, \dots, L - 1$, corrisponde un intervallo \mathcal{I}_i di ampiezze di $s(t)$: se $s(t)$ ha ampiezza entro \mathcal{I}_i il segnale quantizzato $s_q(t)$ avrà ampiezza Q_i .

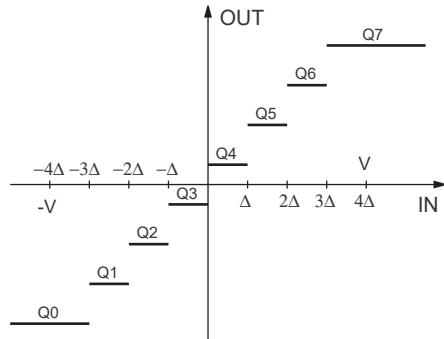
Il tipo più semplice di quantizzazione è la *quantizzazione uniforme*: l'intervallo di ampiezze di $s(t)$, $[-V, V]$, viene diviso in L parti (intervalli) uguali, ciascuno di ampiezza

$$\Delta = \frac{2V}{L}$$

L'intervallo \mathcal{I}_i ha estremi $-V + i\Delta$ e $-V + (i + 1)\Delta$. Il livello di quantizzazione Q_i va scelto in modo da minimizzare l'errore: in assenza di maggiori informazioni sull'andamento di $s(t)$ sembra ragionevole sceglierlo a metà intervallo, ovvero $Q_i = -V + (i + 0.5)\Delta$. In questo modo il massimo errore di quantizzazione sulle ampiezze vale $\Delta/2$, ovvero la metà del passo di quantizzazione Δ .

La regola di quantizzazione è allora la seguente: se l'ampiezza in un certo istante cade nell'intervallo \mathcal{I}_i il valore viene quantizzato a Q_i .

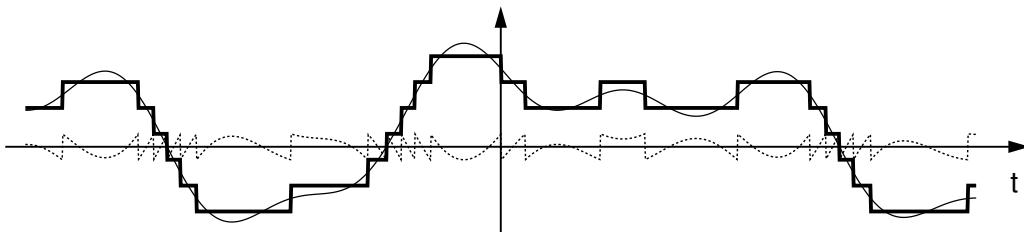
Vediamo un esempio: un quantizzatore uniforme a $L = 2^3 = 8$ livelli. Il valore da quantizzare (IN in figura) viene portato al livello corrispondente (OUT) all'intervalle nel quale cade. Con 8 livelli ho che $\Delta = V/4$. Se ad esempio il valore da quantizzare cade nell'intervallo $[0, \Delta]$ il valore in uscita è Q_4 , ossia $\Delta/2$, se cade in $[3\Delta, 4\Delta]$ diventa $Q_7 = 3.5\Delta$. Da notare che conviene estendere i due livelli estremi per gestire al meglio anche eventuali valori al di fuori dell'intervallo $[-V, V]$.



L'uscita di tale quantizzatore è in figura: il segnale $s(t)$ è in linea continua sottile, quello quantizzato $s_q(t)$ a 8 livelli in linea continua spessa. Viene anche visualizzato, con linea tratteggiata, l'errore di quantizzazione, definito da

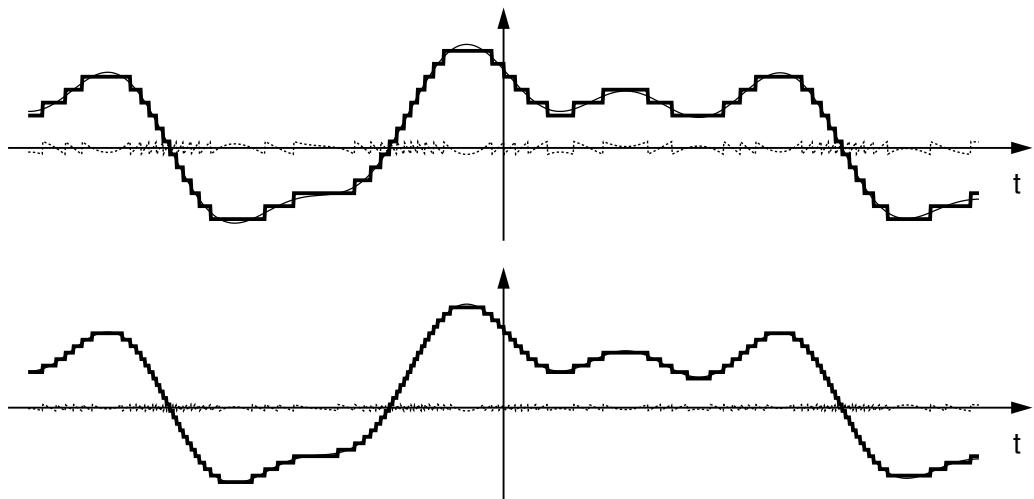
$$e_q(t) = s_q(t) - s(t) \quad (2.15)$$

come visto, $|e_q(t)| \leq \Delta/2 = V/L$.



Come evidente, la quantizzazione a 8 livelli è molto grossolana. Per avere una quantizzazione fine occorre aumentare molto il numero di livelli: in pratica si usano quantizzatori con almeno $L = 2^8 = 256$ livelli (la quantizzazione della musica sui CD audio usa $2^{16} = 65536$ livelli).

Le figure seguenti mostrano il miglioramento che si ottiene usando $L = 2^4 = 16$ e $L = 2^5 = 32$ livelli.



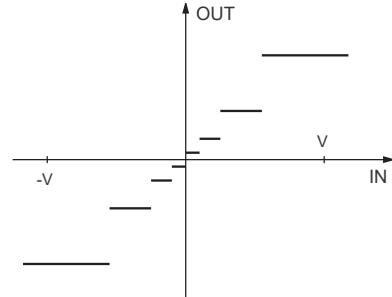
L'errore di quantizzazione, che causa una degradazione irreversibile dell'informazione, può essere analizzato in dettaglio. Per fare questo è però necessaria la teoria dei processi aleatori, che vedremo nel seguito del corso.

2.2.1 Quantizzazione non uniforme

Usare la quantizzazione uniforme, ovvero equispaziare i livelli, non sempre risulta essere la scelta migliore. Esistono segnali aventi una distribuzione delle ampiezze non uniforme: spesso i valori di ampiezza piccoli sono più probabili di quelli grandi o di picco, come ad esempio accade per i segnali vocali.

In questi casi è opportuno usare un *quantizzatore non uniforme* che abbia livelli più fitti a bassi valori di ampiezza e più rari a grandi valori. In questo modo l'errore di quantizzazione diminuisce per piccoli livelli (più probabili) ed aumenta per i grandi (meno probabili): in totale l'errore medio diminuisce.

Questo tipo di quantizzazione non uniforme è anche il più adatto quando il segnale è destinato ad essere ascoltato da un umano.



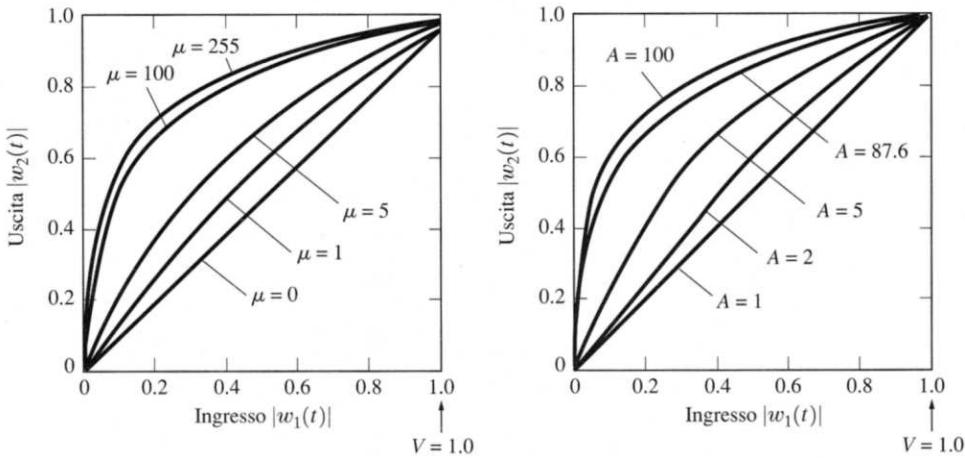
Il nostro sistema uditivo risponde infatti alle ampiezze del suono in maniera logaritmica, ovvero percepisce come uguali variazioni percentuali uguali: la differenza fra l'ampiezza 0.1 e 0.12 viene percepita dello stesso grado della variazione fra 1.0 e 1.2. Il fatto che l'errore di quantizzazione sia minore dove il segnale è piccolo causa quindi una migliore qualità percepita.

Da notare che il miglioramento dato dalla quantizzazione non uniforme diventa meno significativo all'aumentare del numero dei livelli. Il campo di maggior applicazione è nella quantizzazione del segnale telefonico: esso ha le caratteristiche viste, in più viene quantizzato a soli 256 livelli. Vediamo in dettaglio come si è deciso di trattarlo.

Segnale telefonico: leggi μ e A . * Lo stesso risultato della quantizzazione non uniforme si può ottenere elaborando prima il segnale analogico mediante un compressore, ovvero un dispositivo non lineare con amplificazione decrescente al crescere dell'ampiezza del segnale, e poi quantizzando mediante un quantizzatore uniforme. Questa tecnica, più semplice rispetto alla realizzazione di un quantizzatore non uniforme, è stata applicata negli Stati Uniti d'America a partire dagli anni '60 mediante la legge μ , definita da

$$w_2(t) = \text{sign}(w_1(t)) \frac{\ln(1 + \mu|w_1(t)|)}{\ln(1 + \mu)} \quad (2.16)$$

dove $w_1(t)$ è il segnale telefonico in ingresso, che deve essere normalizzato fra -1 e 1, ovvero $|w_1(t)| \leq 1$, $w_2(t)$ è il segnale compresso e μ è un parametro positivo che determina la quantità di compressione. La curva di conversione è in figura a sinistra (da [Cou02]) per alcuni valori di μ : si nota che $\mu = 0$ corrisponde ad una caratteristica lineare (e quindi ad un campionamento uniforme) e che al crescere di μ il grado di compressione aumenta.



Il valore $\mu = 255$ è tipico della telefonia nordamericana e giapponese.

In Europa si usa una legge lievemente differente, la cosiddetta legge A , definita da

$$|w_2(t)| = \begin{cases} \text{sign}(w_1(t)) \frac{A|w_1(t)|}{1+\ln(A)} & 0 \leq |w_1(t)| \leq \frac{1}{A} \\ \text{sign}(w_1(t)) \frac{1+\ln(A|w_1(t)|)}{1+\ln(A)} & \frac{1}{A} \leq |w_1(t)| \leq 1 \end{cases} \quad (2.17)$$

e mostrata in figura (a destra).

Naturalmente, quando in codifica si applica una compressione, il ricevitore deve effettuare l'operazione inversa, ossia una *espansione* con una legge reciproca per ripristinare i corretti valori del segnale. L'operazione di compressione/espansione viene indicata dal termine inglese *companding* (*compressing/expanding*) e, come già indicato, ha lo scopo di migliorare la qualità del segnale ricevuto.

2.3 Pulse Code Modulation (PCM)

La modulazione con codice ad impulsi (PCM, Pulse-Code Modulation) è un tipo di conversione analogico-numerica (ADC, analog to digital conversion) nella quale l'informazione contenuta in ciascun campione di un segnale analogico viene rappresentata attraverso "parole di codice" organizzate in un flusso di dati binari.

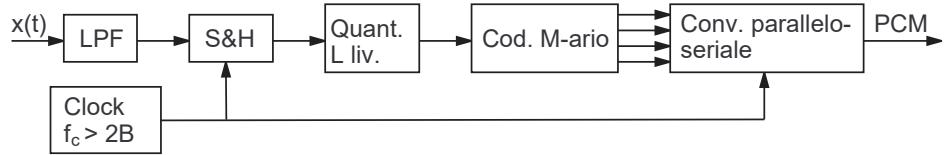
Una parola di codice è una sequenza di cifre (solitamente binarie, o *bit*): se ciascuna parola di codice ha n bit esistono $2^n = M$ parole di codice diverse e ciascuna rappresenta un diverso livello di ampiezza del segnale.

La codifica PCM viene utilizzata in diverse applicazioni, dal CD audio ai sistemi di trasmissione numerici. Le ragioni di tale diffusione sono molte, fra esse:

- la circuiteria che realizza tale tecnica è essenzialmente digitale e a basso costo;
- segnali PCM derivanti da qualsiasi tipo di sorgente analogica (audio, video, voce ecc.) possono essere multiplati con segnali dati (trasferimento file, comunicazioni fra calcolatori) e possono essere trasmessi su di una unica rete numerica. Questo tipo di coesistenza può essere realizzato attraverso tecniche di multiplazione a suddivisione di tempo, come vedremo più avanti;
- nei sistemi di telefonia numerica a lunga distanza con ripetitori, all'uscita di ogni ripetitore è possibile *rigenerare* un segnale PCM "pulito" dai disturbi presenti sul segnale PCM di ingresso. I disturbi possono tuttavia generare occasionalmente *errori* sui valori dei bit trasmessi.
- le prestazioni di un sistema numerico nei confronti del rumore possono essere superiori a quelle di un sistema analogico e possono ulteriormente essere migliorate con l'uso di specifiche tecniche di codifica.

Questi vantaggi compensano abbondantemente quello che è lo svantaggio principale del PCM, cioè la necessità di una banda sensibilmente più alta di quella del corrispondente segnale analogico.

Lo schema generale di un codificatore PCM è



Il segnale analogico in banda base $x(t)$, di banda B , viene filtrato dal filtro passa-basso (antialiasing) LPF di banda B e quindi campionato dal blocco Sample & Hold (S&H) che produce in uscita i campioni $x(nT_c)$ del segnale. La frequenza di campionamento deve rispettare la condizione di Nyquist, quindi $f_c \geq 2B$. Il blocco successivo è un quantizzatore a L livelli: in uscita a tale blocco si hanno i campioni quantizzati $x_q(nT_c)$, ovvero il segnale in ingresso reso discreto nel tempo (dal campionamento) e discreto in ampiezza (dal quantizzatore).

Teoricamente sarebbe possibile invertire l'ordine di campionatore e quantizzatore: in pratica non è tuttavia possibile, in quanto i quantizzatori reali hanno bisogno di un po' di tempo per produrre l'uscita. Intanto che il quantizzatore calcola il livello di uscita l'ingresso deve avere ampiezza costante: il campionatore sample and hold si occupa appunto di questo.

Il blocco seguente è il *codificatore M-ario* che trasforma i campioni quantizzati in parole di codice. Il codificatore usa un alfabeto M-ario e produce, per ogni campione, una parola di codice di ν cifre (*digit*) in parallelo. Poiché vi sono M^ν possibili parole di codice M-ario con ν cifre, per codificare in modo univoco L livelli di quantizzazione si richiede che $M^\nu \geq L$. Per evitare sprechi i parametri vanno scelti in modo da soddisfare l'uguaglianza, ossia

$$L = M^\nu \quad \nu = \log_M L$$

Solitamente si usa un alfabeto *binario*, ossia $M = 2$, e quindi ci si vincola ad un numero di livelli che sia potenza del due: $L = 2^\nu$. Valori tipici per ν sono 8, 12, 16, che corrispondono a 256, 4096 e 65536 livelli. Le cifre che compongono ciascuna parola di codice escono in parallelo dal codificatore: per un codificatore binario su hanno quindi ν bit in parallelo per ciascun campione. Il passo successivo è quello di serializzare le cifre (i bit), operazione compiuta dal convertitore parallelo/seriale (*shift register*).

Il codificatore PCM agisce quindi come un ADC, *Analog to Digital Converter*. Le cifre (i bit) seriali in uscita vengono emesse ogni T_c/ν secondi, visto che si producono ν cifre ogni campione ed i campioni sono separati da T_c secondi. Se $f_c = 1/T_c$ è la frequenza di campionamento si definisce

$$r = \nu f_c \geq 2\nu B$$

la frequenza di simboli (cifre) al secondo emesse. In inglese tale frequenza si chiama *symbol rate* o *signaling rate* e, nel caso binario, *bit-rate*. Come indicato nello schema del codificatore PCM, è necessario un circuito di temporizzazione (clock) che sincronizzi il campionamento con la conversione parallelo/seriale.

Occupazione di banda del segnale PCM. La stima precisa dell'occupazione di banda del segnale PCM non è semplicissima. Ci limitiamo ad una approssimazione: se B è la banda del segnale di ingresso $x(t)$ deve valere la condizione di Nyquist, ossia $f_c \geq 2B$. Quindi in numero di cifre al secondo è $r = \nu f_c \geq \nu 2B$. Secondo il teorema della dimensionalità la banda del segnale PCM è allora tale che

$$B_{PCM} \geq \frac{1}{2} r \geq \nu B \tag{2.18}$$

Questa stima è valida se si associa a ciascuna cifra un impulso di tipo *sinc*. Se, come spesso avviene, gli impulsi associati sono dei rettangolini la banda è maggiore (in teoria infinita). Se (sotto)stimiamo la banda dei rettangolini al primo zero in frequenza (il primo zero della trasformata, che è una *sinc* in frequenza), si ottiene la cosiddetta *banda al primo nullo*, che vale

$$B_{PCM} = r \geq 2\nu B \quad \text{banda al primo nullo} \tag{2.19}$$

Infatti la trasformata del rettangolino di durata $1/r$ è

$$\prod \left(\frac{t}{1/r} \right) \leftrightarrow \frac{1}{r} \operatorname{sinc} \left(\frac{f}{r} \right) \tag{2.20}$$

il cui primo zero è ovviamente in $f/r = 1$ (si è trascurata la traslazione nel tempo, che in frequenza è un termine di fase ininfluente per la ricerca degli zeri).

La banda al primo nullo è una stima più ragionevole dell'occupazione del segnale PCM e ci mostra come il PCM seriale abbia una banda notevolmente maggiore di quella del segnale analogico corrispondente.

Esempio 2.1 Codifica PCM di un segnale telefonico. Come visto, un segnale telefonico $s(t)$ occupa una banda convenzionale da 300 a 3400 Hz. La minima frequenza di campionamento è allora 6800 Hz. Per poter usare filtri antialiasing e di ricostruzione di costo ragionevole si inserisce una banda di guardia campionando ad una frequenza $f_c = 8000$ Hz, che è la frequenza di campionamento standard nei sistemi telefonici digitali in Europa e Stati Uniti d'America. Utilizzando un quantizzatore a $256 = 2^8$ livelli ed una codifica binaria otteniamo una bit-rate di $r_b = 8 \cdot 8000 = 64000$ bit/secondo, che si può anche indicare come $r_b = 64$ kbit/s o 64 kbps (*bps, bit per second*).

La stima della banda occupata al primo nullo è allora 64 KHz, da confrontare con la banda del segnale originale $s(t)$ che, per i sistemi di trasmissione reali (con bande di guardia) vale 4 KHz. In questo esempio la banda occupata è quindi ben 16 volte quella originale!

Esempio 2.2 Codifica PCM di un segnale audio (qualità CD). Il segnale audio a qualità CD è lo standard per la codifica di segnali ad alta fedeltà. La banda stimata per tali segnali è 20 KHz, la frequenza di campionamento, al solito un po' aumentata, è stata fissata a $f_c = 44.1$ KHz. Il quantizzatore usa $L = 65536 = 2^{16}$ livelli che vengono codificati in binario mediante 16 bit. Poiché il segnale è stereo (2 canali per destra e sinistra), la *bit-rate* risulta

$$r_{CD} = 2 \cdot 16 \cdot 44.1 \cdot 10^3 = 1.4112 \text{ Mbit/secondo}$$

La stima di banda al primo zero è $B_{CD} = 1.4112$ MHz, da confrontare con $2 \cdot 20$ KHz.

Quando questo segnale PCM viene scritto su CD si aggiunge una notevole quantità di bit per protezione dei dati, tanto che il flusso di bit effettivamente registrato è pari a 4.32 Mbit/s.

Schema riassuntivo del PCM.

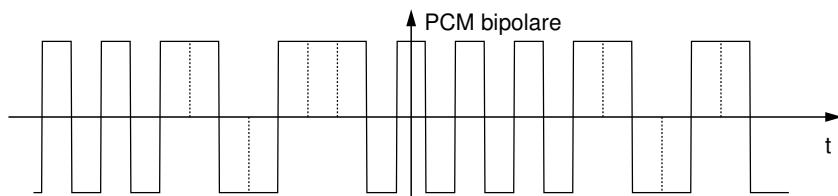
- $x(t)$ segnale passabasso di banda B ;
- campionamento a frequenza $f_c \geq 2B$, intervallo fra i campioni $T_c = 1/f_c$;
- quantizzazione a $L = 2^\nu$ livelli e codifica binaria, ovvero ν bit per campione;
- *bit-rate* $r = \nu f_c \geq 2\nu B$;
- banda minima $B_{PCM\ min} = \frac{1}{2}r \geq \nu B$;
- banda al primo zero $B_{PCM\ primo\ zero} = r \geq 2\nu B$.

2.3.1 Esempio: stima grafica della banda del PCM

Approfondiamo il concetto di banda occupata. Supponiamo di considerare il PCM binario (che è il più diffuso): uno dei modi più semplici per trasmetterlo è quello di associare a ciascun bit un rettangolino di durata pari al tempo di cifra $1/r$. Una possibile scelta è di trasmettere un rettangolino alto A se il bit da trasmettere vale 1 e un rettangolino alto zero (cioè niente) se il bit vale 0. Una tale sequenza di rettangolini si definisce *unipolare* ed ha lo svantaggio di non avere valor medio nullo, ovvero di avere una componente continua che non porta informazione ma "consuma" potenza in trasmissione. Ciò è ininfluente per certe applicazioni (il bus dati dei calcolatori elettronici usa effettivamente questa codifica) mentre diventa significativo per altre, tipicamente quando c'è trasmissione a distanza. In questi casi è meglio una codifica *bipolare*, ovvero tale da associare al valore 1 il livello $+A/2$ e al livello zero il livello $-A/2$, ottenendo un valor medio nullo (almeno se zeri e uni sono più o meno in pari numero).

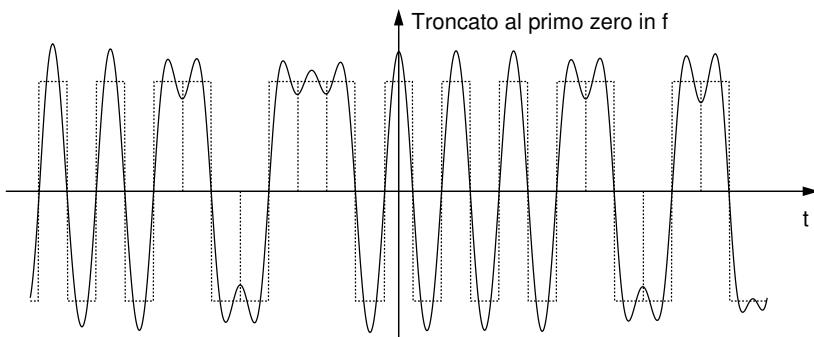
Come vedremo fra poco, associare rettangolini al segnale PCM corrisponde ad una trasmissione PAM (Pulse Amplitude Modulation), ovvero una modulazione ad ampiezza di impulso, visto che l'informazione viene portata dall'ampiezza di forme d'onda (impulsi) tutte uguali.

Un esempio di codifica PAM bipolare del segnale PCM è in figura: le linee tratteggiate verticali sono aggiunte per chiarezza, ma non appartengono al segnale. Ciascun rettangolo, di durata $1/r$, codifica il valore di un bit: il valore per $t = 0$ è 1, poi si ha la sequenza 01010110011...



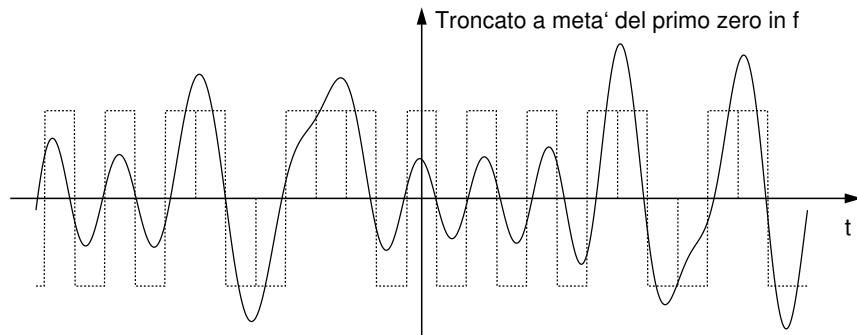
Il ricevitore del segnale PAM può essere realizzato campionando il valore del segnale ricevuto a metà intervallo (al centro di ogni rettangolino). Se il valore campionato è maggiore di zero si decide che è stato trasmesso 1, se è minore di zero si decide 0. Se vale esattamente zero non si sa cosa decidere, quindi si sceglie a caso (comunque capita raramente...).

La banda del segnale in figura è ovviamente infinita, visto che ha fronti di salita e discesa verticali. Se tronchiamo la banda al primo zero, tenendo quindi solo il primo lobo della trasformata del rettangolino, ovvero troncando a $|f| = r$, ciascun rettangolino diventa una funzione oscillante (vagamente simile ad una *sinc*), ed il segnale si trasforma in quello seguente (non sono rappresentati eventuali ritardi nella trasmissione)

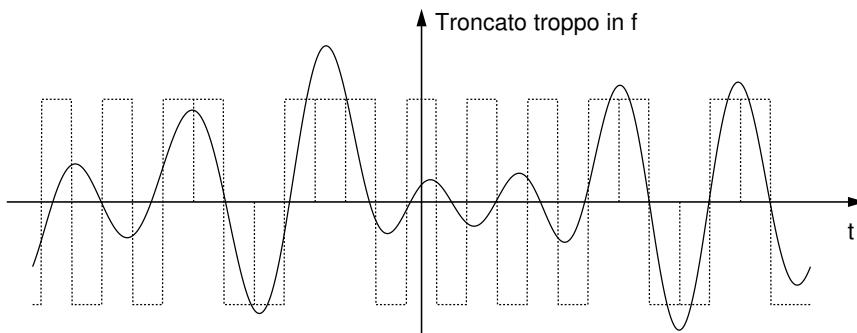


I rettangolini si sono ammorbidi, ma la strategia di ricezione descritta in precedenza continua a funzionare, lasciando anche ampio margine per contrastare gli eventuali disturbi.

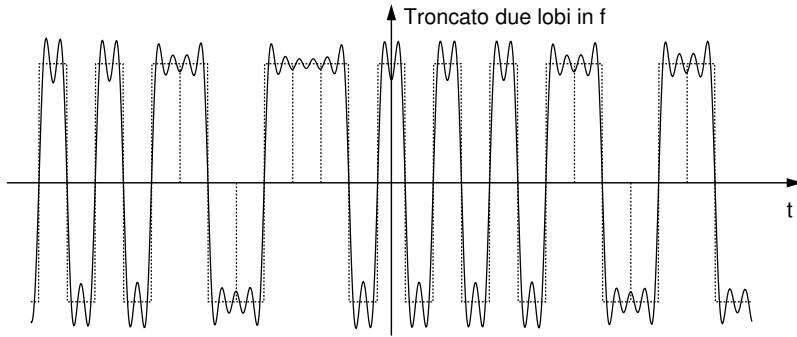
Se tronco a metà del primo lobo (dimezzo la banda rispetto a prima), ovvero al limite teorico, la ricostruzione è ancora possibile, ma il margine per contrastare i disturbi diventa minimo (in realtà nullo, per certe sequenze di bit), come mostrato in figura



Se esagero, troncando a frequenze ancora inferiori, la ricostruzione senza errori non è più possibile. Nella figura sotto il segnale è troncato a 2/3 della metà del primo zero ed alcuni dei bit sono ricostruiti sbagliati.

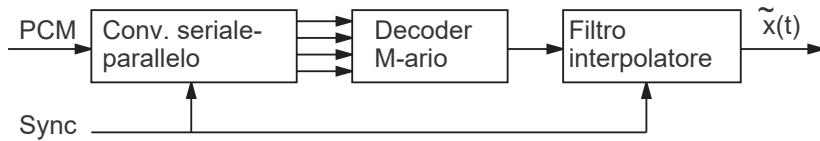


Come ultimo esempio vediamo come diventa il segnale PAM se tronco "abbondante", ovvero tenendo due lobi della sinc (al secondo zero):



2.3.2 Ricevitore PCM

Dal segnale PCM è possibile ritornare al segnale analogico di partenza (a parte gli effetti dell'errore di quantizzazione), mediante il ricevitore PCM, il cui schema è



Il primo blocco ricostruisce la parola di codice dalla sua versione serializzata, il secondo (il decodificatore M-ario) ricostruisce i valori dei campioni quantizzati $x_q(nT_c)$ dai quali, mediante un opportuno filtro interpolatore, si ottiene il segnale ricostruito $\tilde{x}(t)$, diverso dall'originale (in assenza di disturbi) a causa degli errori introdotti dalla quantizzazione. Un modo comune di realizzare il filtro interpolatore consiste in un blocco di tenuta, che mantiene il livello $x_q(nT_c)$ per un tempo $\tau \leq T_c$, seguito da un filtro passa-basso di larghezza di banda pari a quella del segnale di partenza $x(t)$, possibilmente con la forma vista per il campionamento *sample&hold* (funzione di τ , tempo di tenuta). Sia il convertitore seriale/parallelo che il blocco di tenuta hanno bisogno di un segnale di sincronizzazione per poter agire negli istanti dovuti.

I due problemi che restano da affrontare sono la valutazione quantitativa del rumore di quantizzazione e la qualità della ricostruzione in presenza di errori (bit errati) sul flusso dei dati. Per ora non abbiamo gli strumenti per poterlo fare, torneremo sull'argomento più avanti nel corso.

Se il segnale PCM deve essere trasmesso a distanza su di un canale di trasmissione (come avviene ad esempio per il segnale telefonico numerico e anche, ma con diversa interpretazione, per quello audio CD) è necessario progettare gli apparati necessari a trasferire i *bit*. Vi sono molti modi per farlo: uno, la PAM, lo vediamo ora, altri li vedremo in seguito.

2.4 Pulse Amplitude Modulation (PAM)

La PAM (*Pulse Amplitude Modulation*) è una tecnica di trasmissione dell'informazione che prevede l'uso di una sequenza di funzioni tutte uguali fra di loro (*Pulse*) equispaziate nel tempo. Ciascuna viene moltiplicata per un numero, la variazione di ampiezza che ne risulta giustifica il nome di *Amplitude Modulation*. Il segnale PAM può quindi essere rappresentato come

$$s_{PAM}(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kD) \quad (2.21)$$

L'informazione è contenuta nei valori a_k che possono assumere, a seconda delle applicazioni, sia valori discreti (ciascun a_k è uno fra un insieme finito e noto a priori di valori) sia valori continui (in questo caso i possibili valori di a_k sono infiniti). Le funzioni $p(t)$ possono, a loro volta, essere finite nel tempo (come nell'esempio già visto dei rettangolini di durata D) oppure di durata infinita.

Abbiamo già incontrato la PAM due volte. La prima in sezione 2.1.5, il segnale campionato Sample & Hold è evidentemente una PAM ad ampiezza continua (infiniti valori di a_k) con impulsi rettangolari di durata $\tau \leq D$. La seconda volta parlando di occupazione di banda del segnale PCM, con impulsi ancora rettangolari cadenzati di $1/r$ ad ampiezza discreta (solo due valori possibili).

Come regola generale la banda occupata non varia in modo significativo al variare del numero di livelli. Sembra quindi opportuno utilizzare PAM a molti (o infiniti) livelli, visto che più sono i livelli più informazione si può trasmettere. Infatti, se usare due livelli permette di associare un bit di informazione a ciascun impulso, utilizzarne, ad esempio, sedici permette di trasmettere 4 bit di informazione per impulso. In generale, se si usano $L = 2^m$ livelli si possono associare $m = \log_2 L$ bit a ciascun impulso.

Il problema nasce dal fatto che, a potenza in trasmissione fissata (come è in tutti i sistemi di trasmissione), più livelli si usano minore sarà la distanza fra di essi e quindi maggiore il rischio che, a causa del rumore sul canale, si possa sbagliare a decidere, al ricevitore, quale livello è stato trasmesso. Rimandiamo l'analisi del comportamento della PAM con rumore a quando avremo gli strumenti adatti. Per ora consideriamo la trasmissione come se il rumore non esistesse.

Una prima distinzione importante riguarda il canale sul quale gli impulsi viaggeranno: se è a banda infinita (o comunque larga) si usa la PAM in un certo modo, se è a banda stretta bisogna trasmettere in modo più raffinato.

2.4.1 PAM su canale a banda illimitata

Un canale è definito a banda illimitata se la sua larghezza di banda è tale da non distorcere in maniera significativa impulsi di durata D . Questo avviene se la banda B_c del canale è molto maggiore di $1/D$, come abbiamo già visto in sezione 2.3. Poiché stiamo considerando trasmissioni PAM in banda base, la banda del canale è intesa come passa basso: il canale ha banda B_c se la sua risposta in frequenza è piatta in $|f| < B_c$ e nulla (o comunque piccola) altrove. Per stabilire un limite, si considera a "banda illimitata" un canale che abbia

$$B_c \geq \frac{1}{D}$$

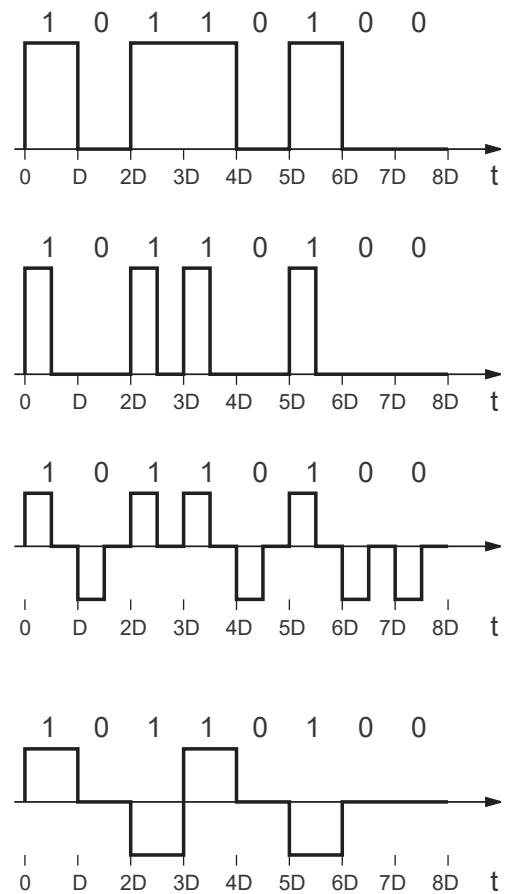
Se si verifica tale condizione si possono utilizzare impulsi di breve durata, ovvero di durata paragonabile a D , con la garanzia che questi non verranno distorti in modo significativo dal canale. Per descrivere i vari formati comunemente utilizzati si supponga di utilizzare una PAM binaria per trasmettere il messaggio binario 10110100.

Unipolare NRZ. I due livelli sono A e 0, quindi il segnale è unipolare, e gli impulsi sono rettangolini di durata D . Il termine NRZ significa *nonreturn-to-zero*, cioè che non vi è ritorno al valore zero prima dell'inizio dell'impulso successivo. Questo formato richiede che sia presente un segnale di sincronizzazione, visto che i rettangoli non sono separati fra di loro.

Unipolare RZ. I due livelli sono ancora A e 0, ma i rettangoli hanno durata $D/2$, poi il segnale torna al valore zero (da qui il nome RZ, *return-to-zero*). Il vantaggio è che non è richiesto un segnale di sincronizzazione, visto che sono presenti transizioni in posizione fissa (almeno per i bit a uno) che aiutano il ricevitore a mantenere la corretta sincronizzazione.

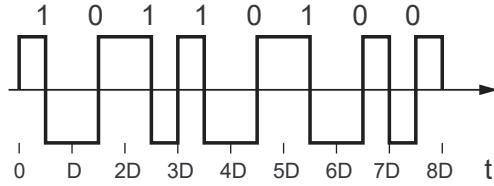
Polare RZ. Entrambi i formati sopra hanno versione polare, nella quale i due livelli usati sono $\pm A/2$. Il formato polare RZ è in figura, quello NRZ dovete immaginarlo (ma lo abbiamo già visto). Il vantaggio dei formati polari è l'assenza di componente continua, almeno se i bit 0 e 1 sono in egual proporzione.

Bipolare NRZ. Secondo questo formato il bit 0 si associa all'ampiezza zero, mentre il bit 1 assume alternativamente i valori $+A$ e $-A$. Questo formato non ha componente continua, neanche se zeri e uni sono in proporzione diseguale. Il suo vantaggio principale è che può essere decodificato anche in presenza di *inversione di segno*, cosa che si può verificare su linee telefoniche commutate (è l'unico formato di quelli qui presentati che gode di tale proprietà).



Split-Phase Manchester. Il bit 1 viene rappresentato da un impulso positivo, $+A/2$, nella prima metà dell'intervallo e negativo, $-A/2$, nella seconda metà. Il bit 0 ha segni opposti. Il segnale risultante ha assenza di componente continua qualunque sia la proporzione di zeri e uni. Inoltre è il segnale più facilmente sincronizzabile al ricevitore (non richiede clock separato), in quanto è sempre presente una transizione a metà intervallo.

Da notare che i formati RZ ed il Manchester utilizzano impulsi di durata $D/2$, ovvero mezzo intervallo di segnalazione: per questi il limite di "banda illimitata" diventa quindi $B_c \geq 2/D$.



PAM multilivello. Gli esempi visti utilizzano solo due valori possibili di ampiezza degli impulsi e quindi possono trasmettere un bit di informazione per volta (PAM binaria). È possibile utilizzare più livelli associando a ciascuno più bit di informazione. Come già visto se ho $L = 2^m$ livelli posso associare a ciascuno $m = \log_2 L$ bit.

Ad esempio nella PAM quaternaria si possono utilizzare i quattro livelli equispaziati $+3/2 A$, $A/2$, $-A/2$ e $-3/2 A$. A ciascun livello posso allora associare due bit di informazione, ad esempio secondo la tabella a lato. Questa associazione si chiama codice naturale, perché segue l'ordine della numerazione binaria.

livello	valore
$+(3/2) A$	11
$+ A/2$	10
$- A/2$	01
$-(3/2) A$	00

livello	valore
$+(3/2) A$	10
$+ A/2$	11
$- A/2$	01
$-(3/2) A$	00

Esiste una associazione migliore, il cosiddetto **codice di Gray**. Il difetto del codice naturale risiede nel fatto che livelli adiacenti rappresentano sequenze di bit che differiscono fra loro per un numero di cifre variabile: ad esempio fra 01 e 10 cambiano due cifre, fra 00 e 01 solo una. Il codice di Gray garantisce che vi sia una e una sola cifra diversa fra livelli adiacenti. Siccome gli errori di trasmissione consistono, in larga maggioranza, nel confondere un livello con uno adiacente, il codice di Gray minimizza il numero di bit errati.

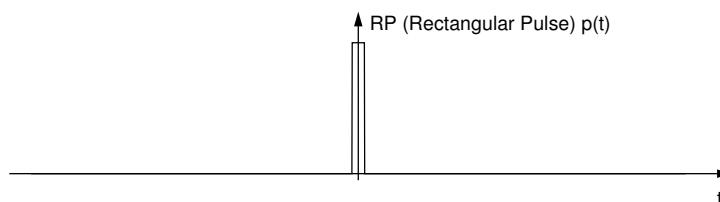
Utilizzando 8 livelli posso associare $\log_2 8 = 3$ bit a ciascun livello. La codifica naturale e quella di Gray sono mostrate in tabella. Da notare come la codifica naturale abbia sempre una transizione fra livelli adiacenti nella quale *tutti* i bit cambiano valore (nell'esempio la transizione fra 3 e 4, ovvero fra 011 e 100), mentre quella di Gray garantisce un solo cambio. Gray è anche ciclico (il passaggio dal valore massimo a zero prevede un solo cambio), anche se per la PAM tale proprietà è irrilevante.

livello	codice naturale	codice di Gray
$+(7/2) A$	111	100
$+(5/2) A$	110	101
$+(3/2) A$	101	111
$+ A/2$	100	110
$- A/2$	011	010
$-(3/2) A$	010	011
$-(5/2) A$	001	001
$-(7/2) A$	000	000

2.4.2 PAM su canale a banda stretta

Se la banda del canale non può essere considerata illimitata, ovvero non è maggiore di $1/D$, è necessario approfondire l'analisi. In un canale a banda stretta gli impulsi relativi ad un intervallo di segnalazione non possono essere confinati in una durata D , ma si allargano, andando ad interferire con i campioni vicini. Questa contaminazione può causare **ISI**, o *interferenza intersimbolica*: per fare in modo che non degradi le prestazioni della trasmissione è necessario applicare adeguate contromisure (vedi il seguito per una definizione precisa di cosa sia l'ISI).

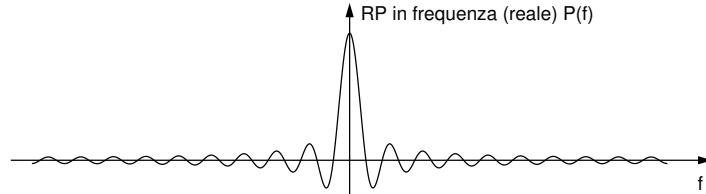
Vediamo in dettaglio cosa succede: consideriamo un singolo impulso rettangolare $p(t)$ di durata D ed ampiezza 1



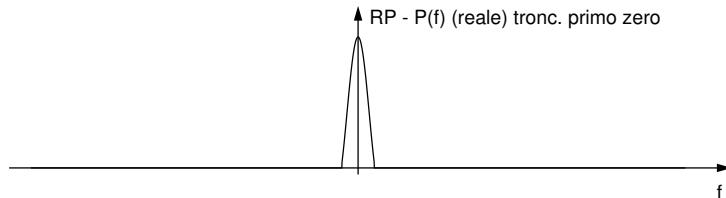
La sua FT è una *sinc*, e precisamente vale

$$F\left[\prod\left(\frac{t}{D}\right)\right] = F[p(t)] = P(f) = D \operatorname{sinc}(fD)$$

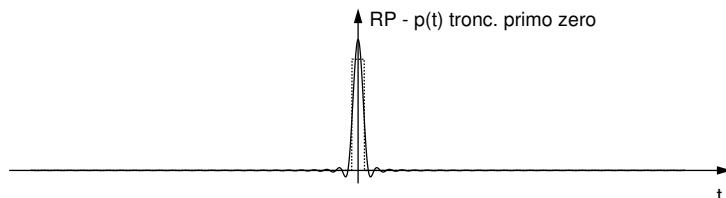
La parte reale di $P(f)$ è in figura (la parte immaginaria è nulla, visto che $p(t)$ è reale e pari)



Troncando $P(f)$ al primo zero (che è in $f = \pm 1/D$), tenendo quindi solo il primo lobo della *sinc*, si ottiene (in frequenza)

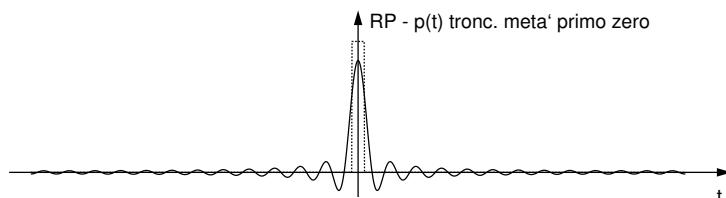


Antitrasformando si ottiene il rettangolino di partenza troncato in banda, ovvero (nel tempo)



Quindi $p(t)$ si è allargato nel tempo e non è più vero che è nullo al di fuori della sua durata nominale D . La funzione risultante è oscillante ed infinita (attenzione, NON è una *sinc*, visto che in frequenza non è un rettangolo) ed invade gli intervalli di tempo dedicati agli altri rettangolini, interferendo con questi ultimi. L'intereferenza è piccola e, a questo livello, può anche essere considerata trascurabile (infatti $1/D$ è la soglia per poter considerare il canale a banda illimitata).

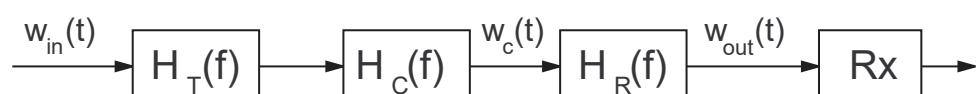
Troncando $P(f)$ a metà del primo lobo, ossia per $f = \pm 1/(2D)$ il rettangolino diventa (nel tempo)



Come evidente (e come ci si deve aspettare) $p(t)$ si è ancora allargato nel tempo (visto che si è stretto in frequenza), è sempre illimitato ed oscillante ed assomiglia di più ad una *sinc* (anche se NON è una *sinc*). L'interferenza con gli altri intervalli è evidentemente aumentata e non può più essere considerata trascurabile.

L'effetto del troncamento in frequenza su di una sequenza di rettangoli è già stato mostrato in modo qualitativo in sezione 2.3: è ora di entrare nel dettaglio.

Consideriamo il segnale PAM ed un generico sistema di trasmissione:



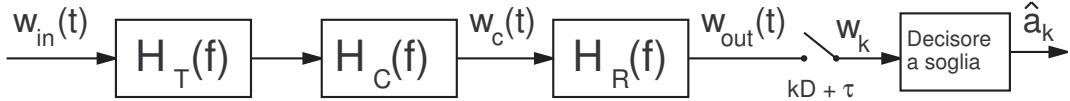
Il segnale in ingresso è (vedi eq. 2.21)

$$w_{in}(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kD) \quad (2.22)$$

Il trasmettitore è composto dal filtro in trasmissione avente risposta in frequenza $H_T(f)$, il canale viene modellato come un filtro a risposta $H_c(f)$ ed il ricevitore è composto dal filtro $H_R(f)$ seguito

dal blocco “ricevitore”, indicato come Rx. La più diffusa (ed efficiente) strategia per realizzare Rx consiste nel campionare il segnale ricevuto e filtrato $w_{out}(t)$ ad intervalli D , ovvero negli istanti $kD + \tau$, essendo τ un ritardo utile a compensare il tempo di transito del segnale sul canale (e gli eventuali ritardi dei filtri) e per spostare il campionamento al centro dell’intervallo. L’uscita del campionatore, che è ovviamente un *numero*, entra quindi in un decisore che sceglie uno fra i possibili valori di a_k (quello che è più probabile che sia stato trasmesso, solitamente il più vicino al valore campionato).

Quindi un generico sistema di trasmissione PAM si può schematizzare come in figura



Poichè il sistema è lineare, possiamo descrivere il segnale $w_{out}(t)$ come

$$w_{out}(t) = \sum_{n=-\infty}^{\infty} a_n h_e(t - nD - \tau)$$

dove $h_e(t)$ (ricavato nel seguito) è il risultato della trasformazione che $p(t)$ subisce passando nel sistema e τ è il ritardo introdotto. Campionando in $t = kD + \tau$ otteniamo

$$w_k = \sum_{n=-\infty}^{\infty} a_n h_e(kD + \tau - nD - \tau) = \sum_{n=-\infty}^{\infty} a_n h_e[(k - n)D]$$

Quello che si vuole ottenere è che w_k sia a_k , a meno di una eventuale costante, ovvero vogliamo che $w_k = C a_k$. Affinchè questo accada è necessario che $h_e(0) = C$ e che $h_e(mD) = 0$ per qualsiasi m intero diverso da zero. Se ciò non accade allora w_k conterrà anche altri termini, che dipendono dagli a_n per $n \neq k$: i termini non voluti sono quelli che danno origine all’interferenza intersimbolica o ISI. Da notare che per evitare l’ISI non è necessario che $h_e(t)$ sia limitato (diverso da zero) nell’intervallo di lunghezza D , basta che sia zero negli (infiniti) istanti di campionamento $h_e(mD) = 0$, $m \neq 0$. Da notare anche come w_k sia un *simbolo*, ovvero un numero (non una funzione del tempo) e che vogliamo che dipenda solo dal simbolo a_k e non dagli infiniti altri simboli a_n , $n \neq k$: è per questo che si parla di interferenza intersimbolica.

Vediamolo in dettaglio: gli impulsi $p(t)$ che formano il segnale in ingresso (se vi fa piacere potete pensarli come rettangolini, ma la forma può essere diversa) possono essere descritti anche per mezzo della convoluzione, ovvero posso scrivere che

$$w_{in}(t) = \sum_{k=-\infty}^{\infty} a_k [p(t) * \delta(t - kD)] = \left[\sum_{k=-\infty}^{\infty} a_k \delta(t - kD) \right] * p(t) \quad (2.23)$$

Il segnale che esce dal filtro in ricezione è allora

$$w_{out}(t) = \left[\sum_{k=-\infty}^{\infty} a_k \delta(t - kD) \right] * h_e(t) \quad (2.24)$$

essendo $h_e(t)$ la risposta impulsiva equivalente del sistema, cioè

$$h_e(t) = p(t) * h_T(t) * h_C(t) * h_R(t) \quad (2.25)$$

La funzione di trasferimento equivalente del sistema è

$$H_e(f) = P(f) H_T(f) H_C(f) H_R(f) \quad (2.26)$$

dove ovviamente $P(f)$ è la FT dell’impulso $p(t)$. Quindi è come avere un unico filtro $H_e(f)$ nel quale entra un treno di impulsi spaziati temporalmente di D , ciascuno con area a_k . Il problema è definire come deve essere $H_e(f)$ affinchè non vi sia interferenza intersimbolica (ISI). Normalmente il canale è dato, per cui si tratta di progettare correttamente la forma degli impulsi $p(t)$ ed i filtri in ricezione e trasmissione ($H_T(f)$ e $H_R(f)$).

Il primo che ha studiato il problema è stato Harry Nyquist (quello del teorema del campionamento) nel 1928. I suoi studi portarono alla scoperta di tre differenti metodi per scegliere $H_e(f)$ in modo da eliminare l’ISI. Noi vedremo in dettaglio il primo (che è il più importante) e descriveremo sommariamente gli altri due.

Primo criterio di Nyquist Il segnale in uscita al filtro di ricezione può essere scritto come in eq. 2.24. Se il blocco RX campiona $w_{out}(t)$ negli istanti $kD + \tau$ la condizione perchè non vi sia ISI è che

$$h_e(kD + \tau) = \begin{cases} C & \text{per } k = 0 \\ 0 & \text{per } k \neq 0 \end{cases} \quad (2.27)$$

Dove C è una qualsiasi costante reale diversa da zero. Questa condizione significa che si vuole che al k -esimo istante di campionamento solo una delle forme d'onda dia un contributo e tutte le altre siano zero. Se inviassimo un solo impulso $ap(t)$ all'istante $t = 0$ il segnale ricevuto sarebbe $w_{out}(t) = ah_e(t)$: se vale la condizione sopra allora il valore di w_{out} in $t = \tau$ è $w_{out}(\tau) = ah_e(\tau) = aC$, mentre il valore negli altri istanti di campionamento è zero ($w_{out}(kD + \tau) = ah_e(kD + \tau) = 0 \forall k \neq 0$).

Tanto per fare un esempio, supponiamo che

$$h_e(t) = C \operatorname{sinc}\left(\frac{t}{D}\right) \quad (2.28)$$

con C costante reale diversa da zero.

Se poniamo $\tau = 0$ notiamo subito come tale funzione soddisfi la condizione di Nyquist, visto che si annulla in tutti gli istanti $t = kD$, con $k \neq 0$.

Quindi se il sistema è progettato in modo che

$$H_e(f) = CD \prod\left(\frac{f}{1/D}\right) \quad (2.29)$$

non vi sarà ISI e la banda richiesta al canale per la trasmissione sarà $B_c = 1/(2D)$.

Questa soluzione alle condizioni in 2.27 ha il difetto di essere di problematica realizzabilità per due diversi motivi:

1. una risposta rettangolare in frequenza è di estrema difficoltà realizzativa: la risposta nel tempo è una *sinc* di durata infinita che va a zero lentamente (come $1/t$);
2. la sincronizzazione temporale del campionatore al ricevitore deve essere molto accurata, visto che le *sinc* degli intervalli adiacenti sono nulle solo se l'istante di campionamento è *esattamente* quello dovuto. Una sincronizzazione non accurata provoca l'insorgere di forte ISI.

Cerchiamo allora se esistono altre risposte che soddisfino la 2.27. Supponiamo, per semplicità, $\tau = 0$ (è solo un ritardo, possiamo sempre inserirlo dopo): se una $h_e(t)$ deve rispettare la 2.27, allora evidentemente vale che

$$C \delta(t) - \sum_{k=-\infty}^{\infty} h_e(kD) \delta(t - kD) = 0 \quad (2.30)$$

questa può anche essere scritta come

$$C \delta(t) - h_e(t) \cdot \sum_{k=-\infty}^{\infty} \delta(t - kD) = 0 \quad (2.31)$$

trasformando nel dominio delle frequenze (e usando l'eq. 1.92) si ottiene

$$C - H_e(f) * \frac{1}{D} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{D}) = 0 \quad (2.32)$$

ovvero

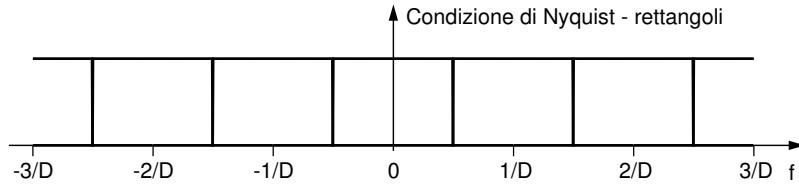
$$H_e(f) * \frac{1}{D} \sum_{k=-\infty}^{\infty} \delta(f - \frac{k}{D}) = C \quad (2.33)$$

che si può scrivere come

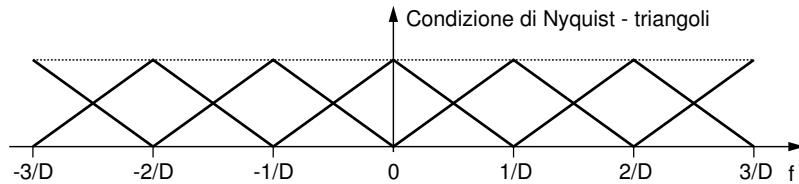
$$\sum_{k=-\infty}^{\infty} H_e(f - \frac{k}{D}) = CD = C' \quad (2.34)$$

Quindi una risposta che soddisfi la condizione nel tempo 2.27 deve rispettare la condizione in frequenza 2.33. Ovvero le repliche di $H_e(f)$ spaziate in frequenza di $f_0 = 1/D$ devono sommare ad una

costante reale. Vi sono infinite funzioni che soddisfano questa condizione: una la abbiamo già vista in eq. 2.29, ovvero un rettangolo in frequenza o una sinc nel tempo. La sua periodizzazione secondo la 2.33 è in figura



Se le repliche sono spaziate di $f_0 = 1/D$ è evidente che nessuna forma d'onda che sia limitata in frequenza ad un valore *inferiore* a $f_0/2 = 1/(2D)$ potrà mai soddisfare la 2.33. La soluzione a banda minima è allora proprio quella visualizzata. Esistono infinite soluzioni con banda maggiore di $1/(2D)$: una è nella figura seguente

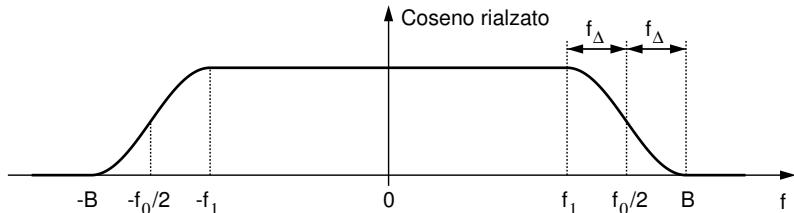


In frequenza sono triangoli di larghezza $1/D$ e, nel tempo, sinc^2 . Come evidente le repliche di $H_e(f)$ si incrociano a frequenza $f_0/2$ e relativi multipli. Benché questa risposta complessiva sia più facilmente realizzabile della *sinc* e abbia minori problemi per la sincronizzazione (le code della sinc^2 si smorzano più rapidamente) ha comunque il difetto di richiedere il *doppio* della banda minima.

Una soluzione ancora diversa prevede risposte di tipo **a coseno rialzato**, ovvero tali che

$$H_e(f) = \begin{cases} 1 & |f| < f_1 \\ \frac{1}{2} \left\{ 1 + \cos \left[\frac{\pi(|f|-f_1)}{2f_\Delta} \right] \right\} & f_1 \leq |f| \leq B \\ 0 & |f| > B \end{cases} \quad (2.35)$$

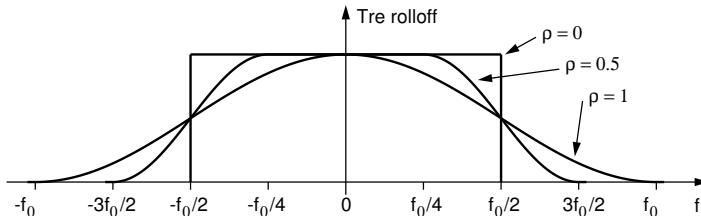
dove B è la banda di $H_e(f)$, con valori $f_0/2 < B < f_0$ (al solito $f_0 = 1/D$), f_1 è in posizione simmetrica a B rispetto ad $f_0/2$, ovvero $f_0/2 - f_1 = B - f_0/2$ e quindi $f_1 = f_0 - B$. Infine $f_\Delta = B - f_0/2$ rappresenta l'eccesso di banda rispetto alla minima assoluta. In figura è mostrato l'andamento del coseno rialzato



Il *fattore di decadimento* o *rolloff* è definito come

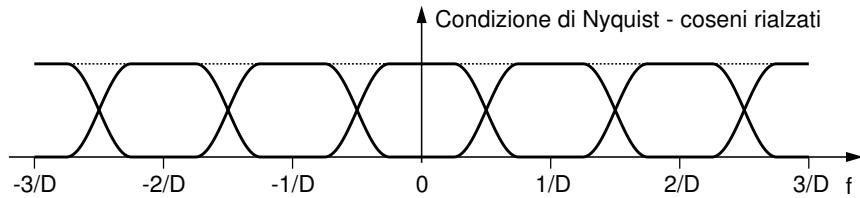
$$\rho = \frac{f_\Delta}{f_0/2}$$

i suoi limiti sono $0 \leq \rho \leq 1$: se vale zero il coseno rialzato diventa un rettangolo, se vale uno sparisce la parte piatta e $B = f_0$. Nella figura sopra $\rho = 1/4$, in quella sotto $\rho = 0, 0.5, 1$.

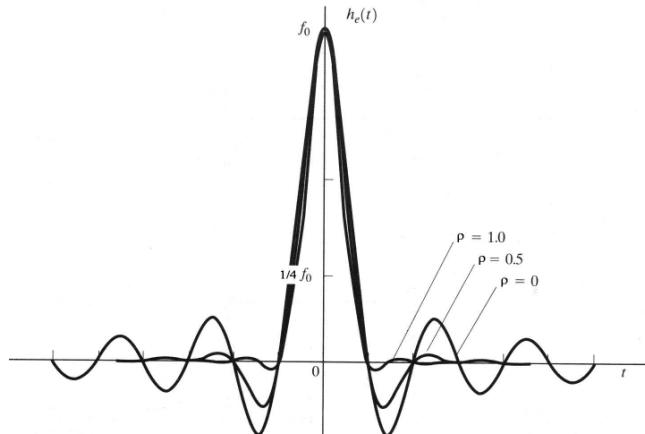


Come evidente, al crescere del *rolloff* aumenta la banda occupata, secondo la relazione $B = (1 + \rho)f_0/2$. Per qualsiasi valore di ρ fra zero e uno la risposta a coseno rialzato soddisfa la condizione 2.29

e quindi anche la 2.27. Nella figura seguente si vede la somma unitaria delle repliche in frequenza per $\rho = 1/2$.



Nel tempo i coseni rialzati si devono annullare in multipli di D : come mostrato nella figura che segue (adattata da [Cou02]) hanno andamento oscillante e, al crescere del *rolloff*, le code si smorzano più rapidamente.



Un filtro con risposta in frequenza tale da soddisfare la condizione 2.33 si chiama *filtro di Nyquist*. Al momento sembra che la sola condizione da rispettare per la progettazione sia il vincolo sulla risposta complessiva $H_e(f)$. Questo lascia diversi gradi di libertà, visto che posso decidere sia la forma dell'impulso trasmesso $p(t)$ che i filtri di ricezione e di trasmissione. Riprenderemo l'argomento più avanti nel corso e vedremo che vi sono altri vincoli da rispettare affinchè si possa minimizzare l'effetto del rumore sul canale.

Secondo e terzo criterio di Nyquist In base al **secondo criterio** di Nyquist è possibile eliminare gli effetti della interferenza intersimbolica (ISI) introducendo in modo controllato una quantità prefissata di ISI che possa essere esattamente cancellata al ricevitore. Seguendo questo criterio il ricevitore decide sempre correttamente, perlomeno in assenza di rumore. Una interessante applicazione, che non abbiamo il tempo di approfondire, è la cosiddetta PAM duobinaria: essa fornisce un modo realizzabile per trasmettere al limite teorico di banda $1/(2D)$ (anche la *sinc* arriva allo stesso limite, ma non è realizzabile).

Il **terzo criterio** di Nyquist per il controllo dell'ISI afferma che l'effetto di quest'ultimo è eliminato scegliendo una risposta impulsiva complessiva del sistema $h_e(t)$ in modo che l'integrale di $h_e(t)$ sul suo intervallo di segnalazione di durata D sia non nullo (una costante), mentre quello eseguito sugli altri intervalli sia nullo. Il ricevitore opera allora NON campionando $h_e(t)$, ma integrandolo sull'intervallo di segnalazione. Sono note diverse soluzioni che soddisfano il terzo criterio, ma le loro prestazioni in presenza di rumore sono inferiori rispetto a quelle che si trovano per i primi due criteri.

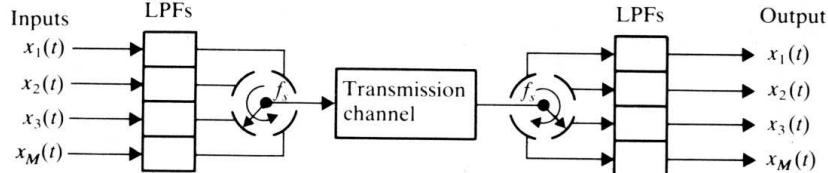
2.5 Multiplexing TDM e FDM

La multiplazione o *multiplexing* è una tecnica che permette a più segnali, omogenei o disomogenei, di essere trasmessi insieme su di un unico canale di comunicazione. Vi sono due strategie base:

1. Multiplazione a suddivisione in frequenza (**FDM**) o *Frequency Division Multiplexing*. Ciascun segnale occupa una diversa banda di frequenza, non sovrapposta a quella di alcun altro segnale.
2. Multiplazione a suddivisione nel tempo (**TDM**) o *Time Division Multiplexing*. Ciascun segnale usa il canale per una parte del tempo. L'asse temporale viene diviso in parti (di solito uguali), dette "slot", e ciascun segnale ne usa una ogni numero prefissato (ad esempio una slot ogni 16).

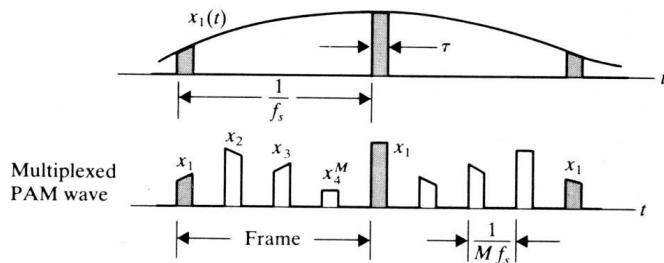
2.5.1 TDM analogica

La TDM consiste nell'intercalare temporalmente i campioni di diversi segnali in modo tale che essi possano essere trasmessi serialmente sullo stesso canale di comunicazione. Esiste sia in versione analogica che numerica. La versione analogica è quella più semplice ed antica, ormai praticamente inusata. Lo schema di massima (da [Car86]) è



Esiste un interruttore che ruota alla frequenza f_s , campionando quindi *chopper* a turno gli M segnali di ingresso $x_i(t)$, $i = 1, M$. Tutti i segnali $x_i(t)$ hanno identica banda W , e deve valere la relazione $f_s \geq 2W$. I filtri LPF di ingresso sono i filtri antialiasing, quindi anch'essi di banda W .

I campioni degli M segnali sono quindi inviati sul canale, ad una rate $r = Mf_s \geq 2MW$ campioni al secondo, come schematizzato in figura.



Al ricevitore esiste un interruttore rotante del tutto analogo a quello in trasmissione, che ha lo scopo di separare i campioni ed inviarli ciascuno al suo filtro di ricostruzione per riottenere il segnale $x_i(t)$ corrispondente. L'interruttore rotante in trasmissione si chiama *commutatore*, mentre quello in ricezione *decommutatore*. Naturalmente non si usano interruttori meccanici rotanti, ma più pratici componenti elettronici.

Indipendentemente da come viene realizzato il TDM, è necessaria una accurata sincronizzazione fra commutatore e decommutatore, pena il mischiare campioni di segnali diversi con esiti disastrosi.

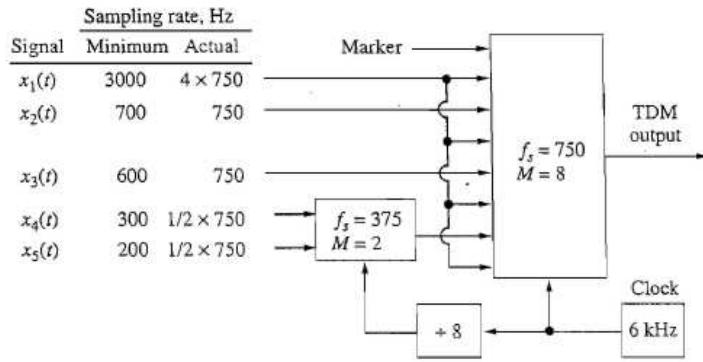
Una tecnica popolare di sincronizzazione prevede di dedicare una delle *slot* temporali ad un impulso di sincronizzazione che sia facilmente riconoscibile al ricevitore. Questo sistema funziona, ma ha il difetto di diminuire di uno il numero di segnali che possono essere inviati. Esistono altri metodi di sincronizzazione che non hanno tale difetto, più complessi, ma non li tratteremo.

L'occupazione di banda può essere stimata come visto in sezione 2.3: considerando che vengono inviati impulsi uno ogni $1/r$ secondi la banda al primo zero è $r = 2MW$. Questo corrisponde ad utilizzare il doppio della banda che si ottiene sommando le bande di ciascun segnale. Il limite teorico di occupazione di banda è la metà, ovvero MW , ma può essere avvicinato solo mediante una accurata progettazione.

Finora si è assunto che tutti i segnali abbiano la stessa banda, ma ciò non è necessario. Se i segnali hanno banda diversa ma simile si può campionare per il segnale di massima banda: gli altri saranno ovviamente sovraccampionati ma il sistema funziona ugualmente. Una strategia migliore quando le bande sono notevolmente diverse (come accade nei sistemi di telemetria analogica) è quella di usare una gerarchia di multipliatori per unire fra loro i segnali a banda inferiore e poi moltiplicarli a quelli di banda maggiore, come illustrato nell'esempio seguente.

Esempio 2.3 Supponiamo di dover trasmettere cinque segnali aventi minima frequenza di campionamento pari a 3000, 700, 600, 300 e 200 Hz. Se usassimo un multiplexer a cinque canali con frequenza di campionamento $f_s = 3000$ Hz per tutti i canali, la *rate* del TDM sarebbe $r = 5 \times 3000 = 15$ KHz, senza considerare gli impulsi di sincronizzazione.

Uno schema più efficiente prevede un multiplexer principale a 8 canali e $f_s = 750$ Hz e un multiplexer secondario a due canali con $f_s = 375$ Hz, come mostrato in figura.



I due segnali a banda inferiore vengono quindi combinati in un solo canale, il che significa che gli impulsi di tali segnali si trovano alternati nel segnale di uscita. Il segnale a banda maggiore viene applicato a quattro ingressi del multiplexer principale, e appare in quattro degli otto *slot* temporali in uscita, per una frequenza di campionamento equivalente pari a $4 \times 750 = 3000$ Hz. La rate totale in uscita, che comprende un marker per la sincronizzazione, vale $r = 8 \times 750 = 6$ KHz.

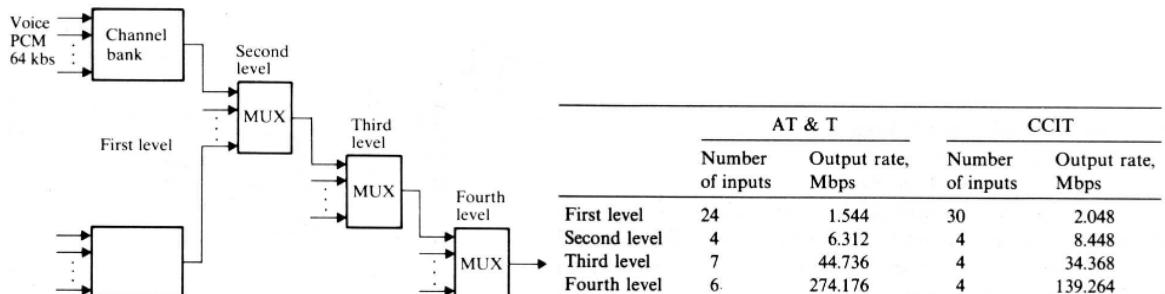
2.5.2 TDM numerica

Da un punto di vista teorico non vi sono grossi problemi: idealmente si può pensare che ciascuno dei segnali da multiplare sia convertito in PCM da un suo convertitore e che i bit in uscita da ciascuno di essi siano intercalati ed inviati sul canale di trasmissione. Ovviamente la velocità di trasmissione sul canale dei bit deve essere sufficiente a trasmettere i bit che arrivano da tutti i modulatori PCM. I problemi che nascono sono di tipo tecnologico.

Vediamo un esempio: supponiamo di voler trasmettere N canali telefonici utilizzando una TDM numerica. Come visto nell'esempio 2.1, ciascun segnale telefonico viene convertito PCM utilizzando una frequenza di campionamento $f_c = 8000$ Hz ed un quantizzatore ad 256 livelli, ossia 8 bit per campione. Il flusso di bit per ciascun canale telefonico è allora di 64 Kbit/secondo. Per mischiare i diversi segnali si possono utilizzare due diverse strategie: una prevede di prendere un bit per volta da ciascun segnale, l'altra di prendere 8 bit per volta, ovvero un campione. Il flusso di bit totale è, in ogni caso, pari a $N \cdot 64$ Kbit/secondo, più gli eventuali bit di sincronizzazione.

L'occupazione di banda minima teorica è $(1/2) \cdot N \cdot 64$ KHz e quella al primo zero il doppio: come già illustrato si ottiene quindi una occupazione di banda molto maggiore della somma delle bande dei segnali analogici di partenza ($N \cdot 4$ KHz).

La trasmissione TDM numerica è molto diffusa, ed esistono due principali standard per essi: uno della AT&T per nord America e Giappone e uno, detto CCIT, per l'Europa. Entrambi si basano su gerarchie di multiplatore, come schematizzato in figura (da [Car86])



Ad esempio il primo livello CCIT accetta in ingresso 30 segnali telefonici, ciascuno da 64 Kbit/sec, e produce in uscita un flusso a 2048 Kbit/sec. Siccome $64 \cdot 30 = 1920 < 2048$ è evidente che vi sono bit in più, dedicati alla sincronizzazione. Per entrambi i sistemi la banda richiesta è alta (quella che servirebbe per la trasmissione analogica è circa il 12 % di quella effettivamente occupata), ma viene compensata dalla flessibilità del sistema. Se, ad esempio, volete collegare due aziende con una rete dati, la Telecom vi affitta un canale da 2 Mbit/sec (o multipli): capite perchè?

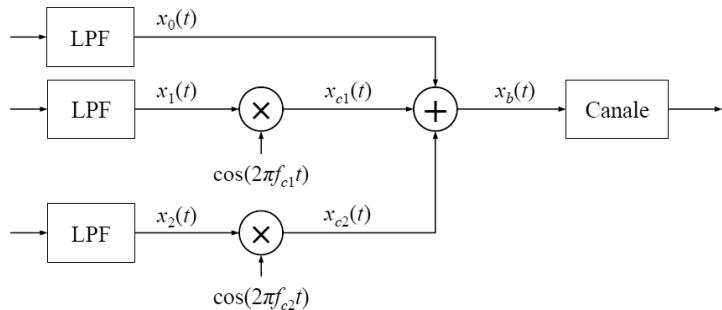
Analizziamo l'efficienza di banda, ovvero quanta banda deve essere allocata per la trasmissione PCM numerica in confronto con la banda originaria dei segnali analogici. L'*efficienza di banda* è definita come il rapporto fra la banda "analogica" e quella numerica. Nel sistema AT&T i quattro livelli di multiplazione possono trasmettere $24 \cdot 4 \cdot 7 \cdot 6 = 4032$ canali vocali PCM utilizzando una banda $B_T \geq r_b/2 \approx 137$ MHz (r_b , come da tabella, è circa 274 Mbit/sec.).

L'efficienza di banda è allora $(4032 \cdot 4\text{KHz})/137\text{MHz} \approx 12\%$.

Il PCM analogico può arrivare ad una efficienza di banda ben superiore, circa il 96 %. Nonostante questo la versione analogica è quasi inusata, a causa dei vantaggi già descritti della versione numerica. Un ulteriore vantaggio di quest'ultima è la possibilità di utilizzare dispositivi logici (ad esempio, TTL) dal costo molto contenuto.

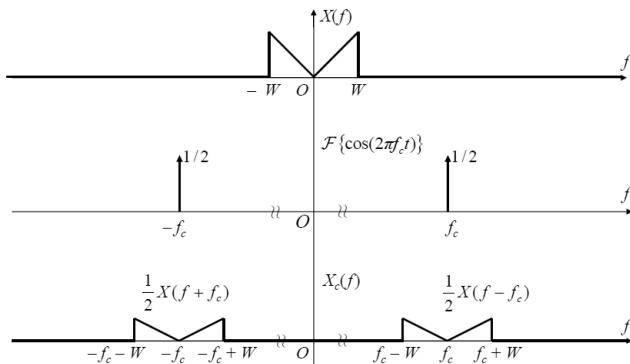
2.5.3 FDM

Un diverso modo per trasmettere più segnali indipendenti sullo stesso canale di trasmissione è quello di **traslare in frequenza** gli spettri dei segnali in modo che non interferiscano e quindi sommarli fra di loro. Un semplice schema a blocchi di un sistema FDM che opera su di un canale in banda base è



I segnali $x_i(t)$, $i = 0, 1, 2$, sono di tipo passabasso di banda W . I filtri passabasso LPF, anch'essi di banda W , assicurano il rispetto della banda. Come visto, moltiplicare per un coseno equivale a traslare in frequenza: se $x_c(t) = x(t) \cos(2\pi f_c t)$ allora

$$X_c(f) = X(f) * \frac{1}{2} [\delta(f - f_c) + \delta(f + f_c)] = \frac{1}{2} [X(f - f_c) + X(f + f_c)]$$



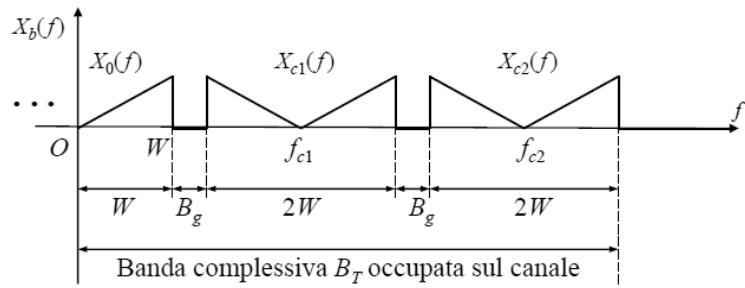
La moltiplicazione per un coseno con relativa traslazione in frequenza si chiama **modulazione DSB** (Double SideBand). La frequenza f_c si chiama frequenza **portante** (carrier), il segnale passabanda ottenuto si chiama **segnale modulato** mentre il segnale originale passabasso è il segnale **modulante**. Il nome *double sideband* indica che l'occupazione in banda del segnale viene raddoppiata rispetto alla banda del segnale modulante, da W a $2W$.

Nell'FDM si usa spesso una diversa modulazione, detta **SSB** (*single sideband*), che ha banda pari a quella del segnale modulante: vedremo tale modulazione, insieme ad altre che pure vengono usate per realizzare la trasmissione FDM, nel secondo semestre.

Nell'esempio dei tre segnali in banda base le due portanti f_{c1} e f_{c2} vanno scelte in modo tale che gli spettri dei tre segnali non si sovrappongano:

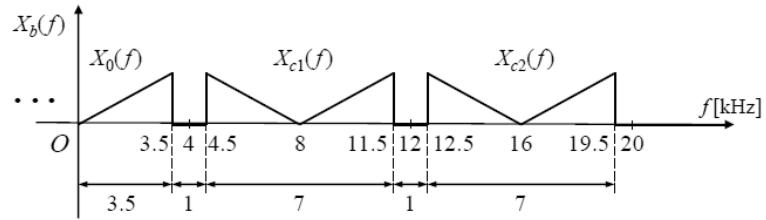
$$f_{c1} \geq 2W \quad f_{c2} \geq f_{c1} + 2W$$

Per garantire la realizzabilità dei filtri si lascia un po' di spazio fra gli spettri dei segnali: tale spazio, indicato come B_g nella figura seguente, si chiama **banda di guardia**.

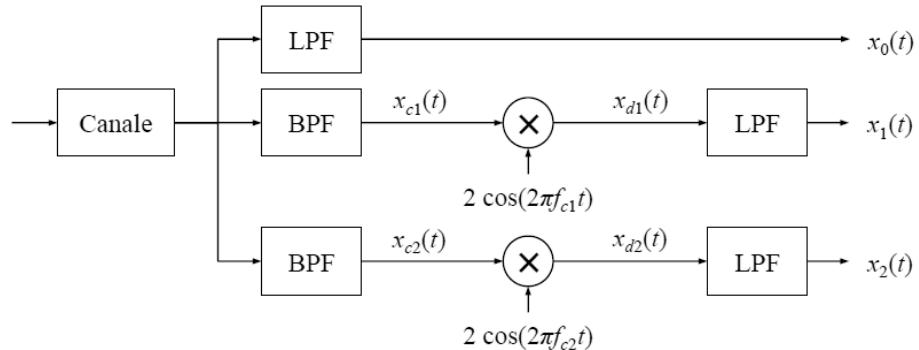


Ad esempio, se i segnali di ingresso sono segnali fonici di banda 3.5 KHz e lasciando una banda di guardia di 1 KHz fra gli spettri, si ottiene una occupazione effettiva sul canale

$$B_t = 3.5 + 1 + 7 + 1 + 7 = 19.5 \text{ KHz}$$

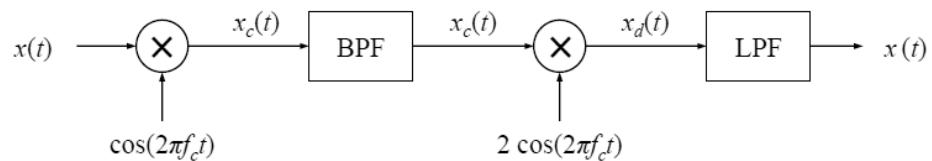


Al ricevitore i segnali possono essere ricostruiti utilizzando lo schema seguente:



Il segnale $x_0(t)$ viene ricostruito semplicemente mediante un filtro passabasso che abbia risposta unitaria in $f = [0, W]$ e che vada a zero entro la banda di guardia.

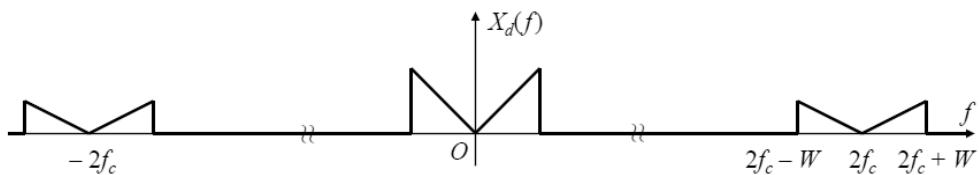
I segnali modulati DSB vengono isolati da un filtro passabanda con risposta unitaria in $f = [f_c - W, f_c + W]$ e quindi demodulati DSB mediante moltiplicazione per un coseno della stessa frequenza e fase del coseno usato per la modulazione. Tale demodulatore si chiama *demodulatore sincrono* e funziona secondo lo schema seguente:



$$\begin{aligned} x_d(t) &= x_c(t) 2 \cos(2\pi f_c t) = x(t) \cos(2\pi f_c t) 2 \cos(2\pi f_c t) = \\ &= 2x(t) \cos^2(2\pi f_c t) = 2x(t) \frac{1 + \cos(4\pi f_c t)}{2} = \\ &= x(t) + x(t) \cos(4\pi f_c t) \end{aligned}$$

$x_d(t)$ in frequenza vale

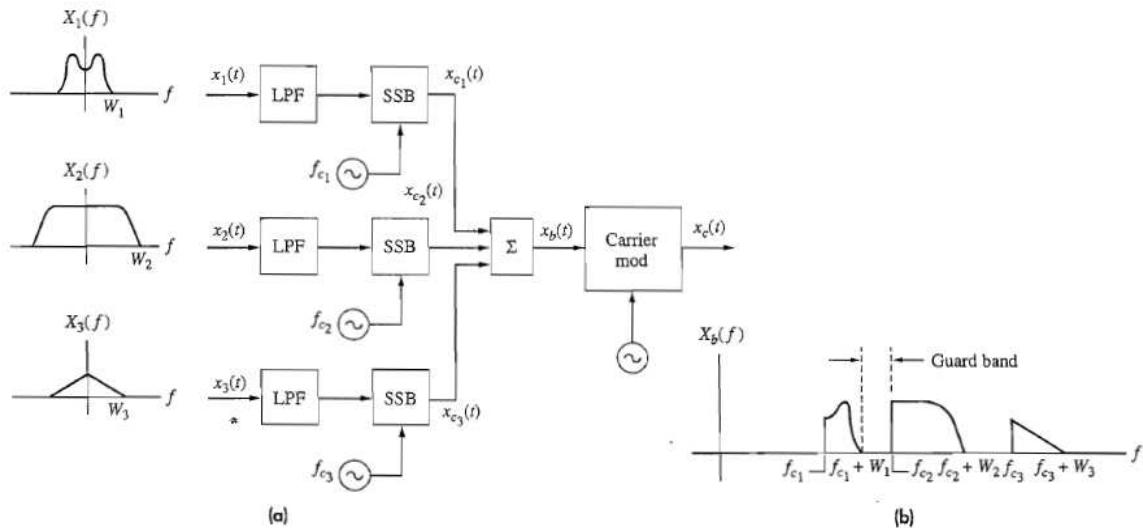
$$X_d(f) = X(f) + \frac{1}{2} [X(f - 2f_c) + X(f + 2f_c)]$$



e quindi il filtro passabasso in uscita ricostruisce $x(t)$.

Nell'esempio visto il canale di trasmissione è di tipo passabasso: questo consente di avere uno dei segnali da trasmettere in banda base. Nel caso di canale passabanda è necessario modulare in trasmissione e demodulare in ricezione tutti i segnali di ingresso.

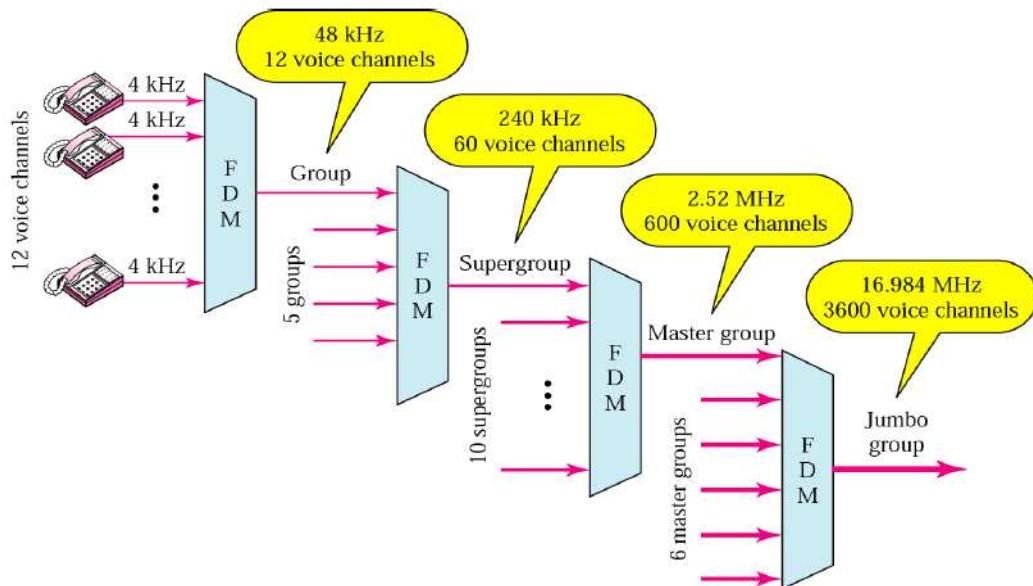
La moltiplicazione FDM può essere applicata ad un insieme di segnali analogici arbitrari, anche con bande differenti. Un possibile esempio è in figura.



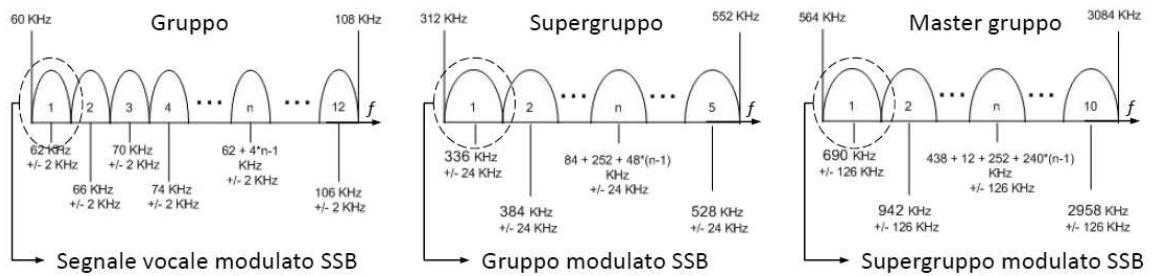
Gerarchie FDM

L'idea di FDM è relativamente semplice, ma i sistemi che ne fanno uso possono essere anche molto complessi e prevedere una struttura gerarchica. Come esempio vediamo come la compagnia statunitense AT&T ha sviluppato la FDM per applicazioni telefoniche.

La figura che segue mostra come 3600 segnali vocali, ciascuno di banda 4 KHz, vengono multiplexati su di un unico cavo coassiale:



I multiplexer ai vari livelli usano modulazioni SSB, e la struttura generale è



Number of voice channels	Bandwidth	Spectrum	AT&T
12	48 kHz	60–108 kHz	Group
60	240 kHz	312–552 kHz	Supergroup
300	1.232 MHz	812–2044 kHz	
600	2.52 MHz	564–3084 kHz	Mastergroup
900	3.872 MHz	8.516–12.388 MHz	
$N \times 600$			Mastergroup multiplex
3,600	16.984 MHz	0.564–17.548 MHz	Jumbogroup
10,800	57.442 MHz	3.124–60.566 MHz	Jumbogroup multiplex

Capitolo 3

Rappresentazione dei segnali

In questo capitolo si torna all'analisi dei segnali. Lo scopo è quello di trovarne una rappresentazione vettoriale. Essa risulterà molto utile, come vedremo, per il progetto ed il calcolo delle prestazioni di sistemi di trasmissione numerici. Nel corso della trattazione approfondiremo i concetti di spazio di funzioni e di serie di funzioni, che hanno applicazioni anche al di fuori delle comunicazioni elettriche.

3.1 Spazi di funzioni

Nel capitolo 1 sono stati già definiti i segnali ed alcune delle misure di interesse per la loro descrizione, ovvero media, energia e potenza. Definiamo ora altri descrittori, utili nel seguito.

3.1.1 Prodotto interno, ortogonalità e norma

Si definisce prodotto interno di due segnali $x(t)$ e $y(t)$ l'integrale

$$(x, y) = \int_a^b x(t)y^*(t)dt \quad (3.1)$$

I segnali $x(t)$ e $y(t)$ sono definiti sull'intervallo (a, b) e, in generale, complessi: ovviamente se $y(t)$ è reale vale che $y(t) = y^*(t)$. Se $(x, y) = 0$, i segnali $x(t)$ e $y(t)$ sono detti ortogonali.

Si definisce norma di un segnale $x(t)$ la radice quadrata di (x, x) :

$$\|x\| = \sqrt{(x, x)} \quad \|x\|^2 = (x, x) = \int_a^b |x(t)|^2 dt \quad (3.2)$$

Se $\|x\| \neq 0$ si dice che $x(t)$ è *normalizzabile*. Per normalizzare $x(t)$ lo si divide per la sua norma $\|x\|$. Se $\|x\| = 0$ allora $x(t) = 0$ per ogni $t \in (a, b)$ escluso eventualmente un insieme di punti di misura nulla. Un sottoinsieme di (a, b) ha misura nulla se può ricoprirsi con un numero finito (o infinito ma contabile) di intervalli di lunghezza complessiva minore di ogni preassegnato $\varepsilon > 0$.

Confrontando le equazioni 1.2 e 3.2 è evidente che $\|x\|^2$ è pari all'energia del segnale $x(t)$ nell'intervallo (a, b) . Si definisce infine distanza fra $x(t)$ e $y(t)$ la norma di $x(t) - y(t)$

$$\|x - y\| = \sqrt{\int_a^b |x(t) - y(t)|^2 dt} \quad (3.3)$$

Diseguaglianza di Schwarz Per ogni $x(t)$ e $y(t)$ definiti sullo stesso intervallo vale che il modulo del prodotto interno è minore o uguale al prodotto delle norme, ovvero

$$|(x, y)| \leq \|x\| \cdot \|y\| \quad (3.4)$$

dove il segno di uguale vale solo se $x(t) = \lambda y(t)$ per un qualche λ .

3.1.2 Funzioni linearmente indipendenti

Date n funzioni $x_i(t)$, $i = 1, 2, \dots, n$ definite su di un intervallo (a, b) , si dice che esse sono **linearmente indipendenti** se ogni relazione della forma

$$\sum_{i=1}^n c_i x_i(t) \equiv 0 \quad (3.5)$$

implica che $c_i = 0$ per $i = 1, 2, \dots, n$, ovvero che i coefficienti c_i siano tutti nulli.

Se ciò non è vero (ovvero se esiste un qualche insieme dei c_i non tutti nulli per cui la 3.5 vale) allora le n funzioni $x_i(t)$ sono linearmente dipendenti.

Esercizio 3.1 Dimostrare che le tre funzioni $1, t$ e t^2 sono linearmente indipendenti su ogni intervallo (a, b) .

Soluzione: supponiamo che esistano tre numeri c_1, c_2 e c_3 tali che $c_1 + c_2t + c_3t^2 = 0$. Deriviamo due volte....

NOTE

1. Se n funzioni $x_i(t), i = 1, 2, \dots, n$ sono linearmente indipendenti, nessuna di esse è identicamente nulla. Infatti, se $x_k(t) \equiv 0$ allora $\sum_{i=0}^n c_i x_i(t)$ con $c_i = 0$ per $i \neq k$ e $c_k \neq 0$ sarebbe identicamente nulla, quindi le funzioni sarebbero linearmente dipendenti.
2. Se n funzioni $x_i(t)$ sono linearmente indipendenti allora lo sono anche le $n - 1$ funzioni che rimangono dopo averne eliminata una qualsiasi. La dimostrazione è ovvia: se la sommatoria di $n - 1$ funzioni è nulla per $n - 1$ c_i non tutti nulli basta inserire la funzione tolta con coefficiente nullo per contraddirre l'ipotesi.
3. Se n funzioni $x_i(t)$ sono linearmente indipendenti non si può esprimere una di esse come combinazione lineare delle altre. Se infatti fosse vero che

$$x_k(t) = \sum_{i \neq k} c_i x_i(t)$$

si avrebbe

$$x_k(t) - \sum_{i \neq k} c_i x_i(t) \equiv 0$$

e le funzioni sarebbero linearmente dipendenti. Una definizione alternativa, e forse più chiara, di funzioni linearmente indipendenti è quindi che nessuna di esse si può esprimere come combinazione lineare delle altre.

Esercizio 3.2 Dimostrare che su ogni intervallo (a, b) le funzioni $\sin(\omega t)$ e $\cos(\omega t)$ sono linearmente indipendenti.

Soluzione: ammettiamo che esistano due numeri c_1 e c_2 tali che, $\forall t$,

$$c_1 \sin(\omega t) + c_2 \cos(\omega t) = 0$$

derivando si ottiene

$$c_1 \cos(\omega t) - c_2 \sin(\omega t) = 0$$

Il sistema formato dalle due equazioni sopra, nelle incognite c_1 e c_2 , ha come unica soluzione $c_1 = c_2 = 0$.

Esercizio 3.3 Dimostrare che su ogni intervallo (a, b) le funzioni $\sin(\omega t)$, $\cos(\omega t)$ e $\cos(\omega t - \varphi)$, con φ arbitrario, sono linearmente dipendenti.

Soluzione: $\cos(\omega t - \varphi) = \cos(\varphi) \cos(\omega t) + \sin(\varphi) \sin(\omega t)$

3.1.3 Ortogonalità ed indipendenza lineare

Date n funzioni $x_i(t), i = 1, 2, \dots, n$, ortogonali a due a due e tali per cui $\|x_i\| \neq 0 \forall i$, esse sono linearmente indipendenti.

Dimostrazione: neghiamo la tesi e dimostriamo che questo porta a negare l'ipotesi. Se le funzioni fossero linearmente dipendenti allora si potrebbe scrivere, per un qualche insieme dei coefficienti $\{c_i\}$ non tutti nulli,

$$x_k(t) = \sum_{i \neq k} c_i x_i(t)$$

Moltiplichiamo entrambi i membri per $x_k^*(t)$

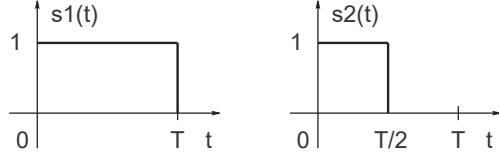
$$x_k(t)x_k^*(t) = \sum_{i \neq k} c_i x_i(t)x_k^*(t)$$

ed integriamo sull'intervallo di definizione (a, b)

$$\int_a^b x_k(t)x_k^*(t)dt = \int_a^b \sum_{i \neq k} c_i x_i(t)x_k^*(t)dt$$

$$\|x_k\|^2 = \sum_{i \neq k} c_i \int_a^b x_i(t)x_k^*(t)dt$$

Gli integrali al secondo membro sono tutti nulli, essendo le funzioni ortogonali fra loro. Quindi $\|x_k\|^2 = 0$, che contraddice l'ipotesi. In conclusione l'ortogonalità implica l'indipendenza lineare, mentre il contrario, in generale, NON E' VERO: l'indipendenza lineare non implica l'ortogonalità. Come esempio consideriamo le due funzioni, definite in $(0, T)$:



E' immediato riconoscere che esse sono linearmente indipendenti ma non ortogonali (se per voi non è così immediato vi conviene riguardare le definizioni).

3.1.4 Rappresentazione vettoriale delle funzioni

Date n funzioni $x_i(t)$, $i = 1, 2, \dots, n$ definite su di un intervallo (a, b) , si definisce sottospazio \mathcal{S} di $L_2(a, b)$ generato dalle n funzioni l'insieme di tutte le combinazioni lineari

$$\sum_{i=1}^n c_i x_i(t)$$

Dato un sottospazio \mathcal{S} , le funzioni $\varphi_j(t)$, $j = 1, 2, \dots, m$, formano una base di \mathcal{S} se:

- a) appartengono a \mathcal{S} ;
- b) sono linearmente indipendenti;
- c) ogni funzione di \mathcal{S} si può esprimere come combinazione lineare delle $\varphi_j(t)$;

Osservazione: poichè ogni $\varphi_j(t) \in \mathcal{S}$, ciascuna può esprimersi come combinazione lineare delle funzioni $x_i(t)$ che hanno generato \mathcal{S} :

$$\varphi_j(t) = \sum_{i=1}^n c_{ji} x_i(t) \quad j = 1, 2, \dots, m$$

Si vede immediatamente che ogni combinazione lineare delle m funzioni $\varphi_j(t)$ appartiene a \mathcal{S} .

TEOREMA: se m funzioni $\varphi_j(t)$ formano una base \mathcal{B} di \mathcal{S} , ogni $x(t) \in \mathcal{S}$ ha una unica rappresentazione del tipo

$$x(t) = \sum_{i=1}^m x_i \varphi_i(t) \tag{3.6}$$

I numeri x_i prendono il nome di componenti di $x(t)$ rispetto alla base \mathcal{B} .

Dimostrazione: per definizione di base, ogni $x(t) \in \mathcal{S}$ ha una rappresentazione del tipo 3.6. Per dimostrare che questa è unica ammettiamo che esista una seconda rappresentazione

$$x(t) = \sum_{i=1}^m x'_i \varphi_i(t)$$

Mediante sottrazione si ottiene

$$x(t) - x(t) = 0 = \sum_{i=1}^m (x_i - x'_i) \varphi_i(t)$$

da cui si deduce, per l'indipendenza lineare delle $\varphi_j(t)$, che $x_i = x'_i \forall i$. C.V.D.

Per il teorema precedente, data una base $\mathcal{B} = \{\varphi_j(t), j = 1, \dots, m\}$ di \mathcal{S} esiste una corrispondenza biunivoca fra le funzioni di \mathcal{S} ed i punti di uno spazio \mathcal{R}^m o \mathcal{C}^m a m dimensioni.

Ad ogni funzione $x(t) \in \mathcal{S}$ corrisponde uno e un solo punto $\bar{x} = (x_1, x_2, \dots, x_m)$ detto immagine di $x(t)$. Viceversa ad ogni punto \bar{x} dello spazio \mathcal{R}^m o \mathcal{C}^m corrisponde una ed una sola funzione di \mathcal{S} .

Una volta specificata una base, le operazioni lineari sulle funzioni di \mathcal{S} (somma e sottrazione fra funzioni e moltiplicazione di una funzione per una costante) si traducono in operazioni lineari sulle loro immagini: l'immagine di $x(t) \pm x'(t)$ è $\bar{x} \pm \bar{x}'$ e l'immagine di $\lambda x(t)$ è $\lambda \bar{x}$.

Si può dimostrare, ma non lo faremo, che uno spazio \mathcal{S} ha più di una base. Quindi esistono più modi di associare le funzioni di \mathcal{S} con i punti di uno spazio euclideo. Tuttavia lo spazio delle immagini di \mathcal{S} ha sempre le stesse dimensioni: *il numero delle funzioni che formano la base non cambia*.

Il numero delle funzioni base (m) è detto dimensione del sottospazio \mathcal{S} : si indica \mathcal{S}^m .

3.1.5 Basi ortonormali e calcolo dell'immagine di una funzione

Una volta fissata una base \mathcal{B} del sottospazio \mathcal{S}^m è semplice calcolare la funzione che corrisponde ad una certa immagine (punto dello spazio delle funzioni base). Calcolare invece le coordinate dell'immagine di una qualsiasi funzione $x(t) \in \mathcal{S}$ può non essere semplice. Lo diventa tuttavia se le funzioni di \mathcal{B} hanno tutte norma unitaria e sono a due a due ortogonali.

Base ortonormale: una base $\mathcal{B} = \{\varphi_j(t), j = 1, \dots, m\}$ si dice ortonormale se:

$$(\varphi_i, \varphi_j) = \delta_{ij} \quad i, j = 1, 2, \dots, m \quad (3.7)$$

essendo $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$.

Se la base \mathcal{B} è ortonormale il calcolo delle coordinate dell'immagine di $x(t) \in \mathcal{S}$ si ottiene come segue. Partiamo da

$$x(t) = \sum_{i=1}^m x_i \varphi_i(t)$$

dove le x_i sono ovviamente le incognite. Moltiplicando per $\varphi_k^*(t)$ ed integrando sull'intervallo (a, b) di definizione delle funzioni si ottiene

$$\int_a^b x(t) \varphi_k^*(t) dt = \sum_{i=1}^m x_i \int_a^b \varphi_i(t) \varphi_k^*(t) dt$$

ovvero

$$x_k = \int_a^b x(t) \varphi_k^*(t) dt = (x, \varphi_k) \quad (3.8)$$

avendo tenuto conto che, per definizione,

$$\int_a^b \varphi_i(t) \varphi_k^*(t) dt = \delta_{ik}$$

Proprietà

Energia Se $\mathcal{B} = \{\varphi_j(t), j = 1, \dots, m\}$ è una base ortonormale di \mathcal{S}^m e $x(t) \in \mathcal{S}^m$ allora, per un qualche insieme x_i , si ha

$$x(t) = \sum_{i=1}^m x_i \varphi_i(t)$$

L'energia di $x(t)$ si può allora esprimere come

$$\begin{aligned} \int_a^b |x(t)|^2 dt &= \int_a^b x(t) x^*(t) dt = \int_a^b \left[\sum_{i=1}^m x_i \varphi_i(t) \right] \left[\sum_{j=1}^m x_j^* \varphi_j^*(t) \right] dt = \\ &= \sum_{i=1}^m \sum_{j=1}^m x_i x_j^* \int_a^b \varphi_i(t) \varphi_j^*(t) dt = \sum_{i=1}^m |x_i|^2 \end{aligned} \quad (3.9)$$

Quindi l'energia di $x(t)$ è pari al quadrato della distanza della sua immagine dall'origine degli assi (teorema di Parseval): $|\bar{x}| = \|x\|$.

Analogamente la distanza dell'immagine di $x(t)$ dall'origine è pari alla norma $\|x\|$ di $x(t)$.

Distanza In modo analogo si vede che se anche $y(t) \in \mathcal{S}^m$, e quindi

$$y(t) = \sum_{i=1}^m y_i \varphi_i(t)$$

allora

$$\int_a^b |x(t) - y(t)|^2 dt = \int_a^b \left| \sum_{i=1}^m (x_i - y_i) \varphi_i(t) \right|^2 dt = \sum_{i=1}^m |x_i - y_i|^2 \quad (3.10)$$

da cui si deduce che la distanza fra le funzioni $x(t)$ e $y(t)$, definita nell'eq. 3.3 come $\|x - y\|$ è pari alla distanza euclidea (geometrica) delle loro immagini.

Prodotto interno Il prodotto interno delle funzioni $x(t)$ e $y(t) \in \mathcal{S}^m$ vale

$$(x, y) = \int_a^b x(t) y^*(t) dt = \sum_{i=1}^m x_i y_i^* = \bar{x} \cdot \bar{y} \quad (3.11)$$

ovvero è pari al prodotto scalare dei vettori immagine delle due funzioni.

Disegualanza di Schwarz Dalla geometria analitica è noto che il prodotto scalare dei vettori \bar{x} e \bar{y} vale $\bar{x} \cdot \bar{y} = |\bar{x}| |\bar{y}| \cos \theta$ essendo θ l'angolo compreso tra i due vettori.

Dalle relazioni viste sappiamo che

$$\bar{x} \cdot \bar{y} = (x, y) \quad \text{e} \quad |\bar{x}| = \|x\|$$

Sostituendo si ottiene

$$(x, y) = \|x\| \|y\| \cos \theta$$

da cui

$$|\cos \theta| = \frac{|(x, y)|}{\|x\| \|y\|} \leq 1$$

che è un modo per dimostrare la disegualanza di Schwarz (vedi eq. 3.4).

3.1.6 Cambiamento di Base

Siano: \mathcal{S}^m un sottospazio, $\mathcal{B}' = \{\varphi_j(t), j = 1, \dots, m\}$ e $\mathcal{B}'' = \{\psi_j(t), j = 1, \dots, m\}$ due sue basi ortonormali, \bar{x}' e \bar{x}'' le immagini di $x(t) \in \mathcal{S}^m$ corrispondenti rispettivamente a \mathcal{B}' e \mathcal{B}'' .

Quale è la relazione fra \bar{x}' e \bar{x}'' al variare di $x(t)$ in \mathcal{S}^m ?

Intanto possiamo notare che la distanza fra \bar{x}' e l'origine deve essere uguale alla distanza fra \bar{x}'' e l'origine, essendo tale distanza pari alla radice quadrata dell'energia di $x(t)$. Quindi \bar{x}' e \bar{x}'' sono entrambi sulla circonferenza (o sfera, o ipersfera...) di raggio $\sqrt{E_x}$ centrata sull'origine.

Consideriamo ora N funzioni $y_i(t) \in \mathcal{S}^m$, $i = 1, 2, \dots, m$ e le loro immagini \bar{y}'_i e \bar{y}''_i nelle basi \mathcal{B}' e \mathcal{B}'' : la distanza fra \bar{y}'_j e \bar{y}'_k deve essere uguale alla distanza fra \bar{y}''_j e \bar{y}''_k , per qualsiasi coppia $j, k = 1, 2, \dots, m$. Le relazioni spaziali reciproche fra le immagini non devono cioè cambiare qualsiasi sia la base ortonormale considerata.

Questo porta a concludere, pur senza una dimostrazione rigorosa, che la relazione cercata è:

Al variare di $x(t)$ in \mathcal{S}^m le immagini \bar{x}'' possono ottenersi dalle \bar{x}' mediante una rotazione rigida attorno all'origine degli assi.

3.1.7 Procedimento di Gram-Schmidt

E' un metodo che consente di trovare una base ortonormale di un sottospazio \mathcal{S} generato da un insieme noto di funzioni $\mathcal{X} = \{x_i(t), i = 1, \dots, n\}$ in generale né ortogonali né linearmente indipendenti.

Si tratta di determinare un insieme di funzioni ortonormali $\mathcal{B} = \{\varphi_i(t), i = 1, \dots, m\}$ tale che:

1. ogni funzione di \mathcal{B} appartenga ad \mathcal{S} ;
2. ogni $x(t) \in \mathcal{S}$ possa esprimersi come combinazione lineare delle $\varphi_i(t)$.

Notiamo due fatti. Il primo è che se $\|x_k\| = 0$ allora $x_k(t)$ è nulla su tutto l'intervallo di definizione (a, b) (a parte il solito eventuale insieme di punti a misura nulla). Il secondo è che se in \mathcal{X} vi sono funzioni a norma nulla e chiamiamo \mathcal{X}' l'insieme ottenuto da \mathcal{X} eliminando queste funzioni, allora i sottospazi generati da \mathcal{X} e \mathcal{X}' coincidono. Nella discussione che segue ammetteremo quindi che $\|x_i\| \neq 0$, $i = 1, 2, \dots, n$.

Descrizione del procedimento di Gram-Schmidt

Il primo passo consiste nel formare la prima funzione di \mathcal{B} partendo dalla prima funzione di \mathcal{X} :

$$\varphi_1(t) = \frac{x_1(t)}{\|x_1\|}$$

Ovviamente $\varphi_1(t)$ è normalizzata e $x_1(t) = \|x_1\|\varphi_1(t)$, ovvero $x_1(t)$ è combinazione lineare delle funzioni in \mathcal{B} .

Come secondo passo consideriamo la funzione ausiliaria

$$\theta_2(t) = x_2(t) - (x_2, \varphi_1)\varphi_1(t)$$

che rappresenta la "parte" della funzione $x_2(t)$ che non è rappresentabile per mezzo di $\varphi_1(t)$. Distinguiamo due casi:

- se $\|\theta_2\| = 0$ allora anche $x_2(t)$ è combinazione lineare delle funzioni in \mathcal{B} (che al momento sono una sola, cioè $\varphi_1(t)$):

$$x_2(t) = (x_2, \varphi_1)\varphi_1(t)$$

- se $\|\theta_2\| \neq 0$ si pone

$$\varphi_2(t) = \frac{\theta_2(t)}{\|\theta_2\|}$$

Evidentemente $\|\varphi_2\| = 1$. Inoltre è facile dimostrare che $\varphi_2(t)$ è ortogonale a $\varphi_1(t)$ e che

$$x_2(t) = \|\theta_2\|\varphi_2(t) + (x_2, \varphi_1)\varphi_1(t) = (x_2, \varphi_1)\varphi_1(t) + (x_2, \varphi_2)\varphi_2(t)$$

ossia che anche in questo caso $x_2(t)$ è combinazione lineare delle funzioni in \mathcal{B} (che sono due, cioè $\varphi_1(t)$ e $\varphi_2(t)$):

nota 3.1

Dimostriamo che $\varphi_2(t)$ è ortogonale a $\varphi_1(t)$.

$$(\theta_2, \varphi_1) = (x_2, \varphi_1) - (x_2, \varphi_1)(\varphi_1, \varphi_1) = (x_2, \varphi_1) - (x_2, \varphi_1) = 0$$

Quindi $\theta_2(t)$ è ortogonale a $\varphi_1(t)$ ma siccome $\theta_2(t) = \|\theta_2\|\varphi_2(t)$ anche $\varphi_2(t)$ lo è.

I passi successivi per ottenere le rimanenti funzioni in \mathcal{B} si possono riassumere come segue: quando si considera la funzione $x_k(t)$ si sono già trovate $l \leq k - 1$ funzioni ortonormali $\varphi_i(t)$. Costruendo la funzione ausiliaria

$$\theta_k(t) = x_k(t) - \sum_{i=1}^l (x_k, \varphi_i)\varphi_i(t)$$

possiamo nuovamente avere due casi:

- se $\|\theta_k\| = 0$ allora $x_k(t)$ è combinazione lineare delle funzioni in \mathcal{B} (che sono l).
- se $\|\theta_k\| \neq 0$ si pone

$$\varphi_{l+1}(t) = \frac{\theta_k(t)}{\|\theta_k\|}$$

La nuova funzione base $\varphi_{l+1}(t)$ è normalizzata ed ortogonale rispetto alle precedenti, ovvero

$$(\varphi_{l+1}, \varphi_i) = 0 \quad i = 1, \dots, l$$

Inoltre

$$x_k(t) = \|\theta_k\|\varphi_{l+1}(t) + \sum_{i=1}^l (x_k, \varphi_i)\varphi_i(t) = \sum_{i=1}^{l+1} (x_k, \varphi_i)\varphi_i(t)$$

Procedendo in questo modo fino a $k = n$ si ottengono tutte le funzioni di \mathcal{B} .

Le funzioni in \mathcal{B} sono evidentemente in numero $m \leq n$ e si può dimostrare il seguente teorema:

Condizione necessaria e sufficiente affinchè sia $m = n$ è che le funzioni in \mathcal{X} siano linearmente indipendenti.

NOTA permutando l'ordine delle $x_i(t)$ in \mathcal{X} il procedimento di Gram-Schmidt genera insiemi di funzioni base in generale differenti. Tali insiemi sono comunque formati tutti dallo stesso numero m di funzioni base.

Esempio 3.4 Date le funzioni

$$x_i(t) = \cos\left[\frac{2\pi}{T}t + (i-1)\frac{\pi}{4}\right] \quad i = 1, \dots, 8$$

definite nell'intervallo $(0, T)$, trovare una base ortonormale utilizzando la procedura di Gram-Schmidt.

La prima funzione base si ottiene mediante

$$\varphi_1(t) = \frac{x_1(t)}{\|x_1\|} = \frac{\cos\left(\frac{2\pi}{T}t\right)}{\sqrt{\int_0^T [\cos\left(\frac{2\pi}{T}t\right)]^2 dt}} = \sqrt{\frac{2}{T}} \cos\left(\frac{2\pi}{T}t\right)$$

Costruiamo la funzione ausiliaria

$$\begin{aligned} \theta_2(t) &= x_2(t) - (x_2, \varphi_1)\varphi_1(t) = \\ &= \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) - \int_0^T \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) \sqrt{\frac{2}{T}} \cos\left(\frac{2\pi}{T}t\right) dt \sqrt{\frac{2}{T}} \cos\left(\frac{2\pi}{T}t\right) = \\ &= \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) - \frac{2}{T} \int_0^T \left[\cos\left(\frac{2\pi}{T}t\right) \cos\frac{\pi}{4} - \sin\left(\frac{2\pi}{T}t\right) \sin\frac{\pi}{4} \right] \cos\left(\frac{2\pi}{T}t\right) dt \cos\left(\frac{2\pi}{T}t\right) = \\ &= \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) - \frac{2}{T} \int_0^T \left[\cos^2\left(\frac{2\pi}{T}t\right) \cos\frac{\pi}{4} - \sin\left(\frac{2\pi}{T}t\right) \sin\frac{\pi}{4} \cos\left(\frac{2\pi}{T}t\right) \right] dt \cos\left(\frac{2\pi}{T}t\right) = \end{aligned}$$

Usando le identità $\cos^2 \alpha = (1/2)[1 + \cos(2\alpha)]$ e $\sin \alpha \cos \alpha = (1/2)\sin(2\alpha)$ e considerando che gli integrali calcolati fra 0 e T di $\cos^2(4\pi t/T)$ e $\sin^2(4\pi t/T)$ valgono zero otteniamo

$$\begin{aligned} &= \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) - \frac{2}{T} \int_0^T \frac{1}{2} \cos\frac{\pi}{4} dt \cos\left(\frac{2\pi}{T}t\right) = \\ &= \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) - \frac{2}{T} \cos\frac{\pi}{4} \frac{1}{2} T \cos\left(\frac{2\pi}{T}t\right) = \\ &= \cos\left(\frac{2\pi}{T}t + \frac{\pi}{4}\right) - \cos\frac{\pi}{4} \cos\left(\frac{2\pi}{T}t\right) = \\ &= \cos\left(\frac{2\pi}{T}t\right) \cos\frac{\pi}{4} - \sin\left(\frac{2\pi}{T}t\right) \sin\frac{\pi}{4} - \cos\frac{\pi}{4} \cos\left(\frac{2\pi}{T}t\right) = -\sin\frac{\pi}{4} \sin\left(\frac{2\pi}{T}t\right) \end{aligned}$$

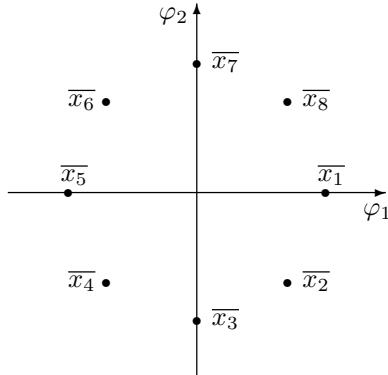
Da cui

$$\varphi_2(t) = \frac{\theta_2(t)}{\|\theta_2\|} = -\sqrt{\frac{2}{T}} \sin\left(\frac{2\pi}{T}t\right)$$

Tutte le rimanenti $\theta_i(t)$, $i = 3, \dots, 7$ sono nulle, cosicché le due funzioni $\varphi_1(t)$ e $\varphi_2(t)$ formano la base ortonormale richiesta. Le immagini o rappresentazioni delle $x_i(t)$ nello spazio delle funzioni base $\{\varphi_1(t), \varphi_2(t)\}$ sono:

$$\begin{aligned} \overline{x_1} &= \left(\sqrt{\frac{T}{2}}, 0\right) \\ \overline{x_2} &= \left(\sqrt{\frac{T}{2}} \cos\frac{\pi}{4}, -\sqrt{\frac{T}{2}} \sin\frac{\pi}{4}\right) \end{aligned}$$

ed analogamente per le altre (vedi figura).

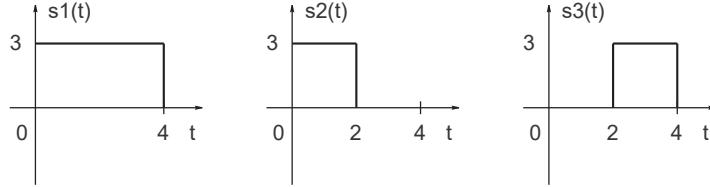


Lo stesso risultato si può ottenere notando che

$$\cos\left[\frac{2\pi}{T}t + (i-1)\frac{\pi}{4}\right] = \cos\left(\frac{2\pi}{T}t\right) \cos\left[(i-1)\frac{\pi}{4}\right] - \sin\left(\frac{2\pi}{T}t\right) \sin\left[(i-1)\frac{\pi}{4}\right]$$

e che le funzioni $\cos\left(\frac{2\pi}{T}t\right)$ e $\sin\left(\frac{2\pi}{T}t\right)$ sono ortogonali in $[0, T]$, anche se non sono normalizzate.

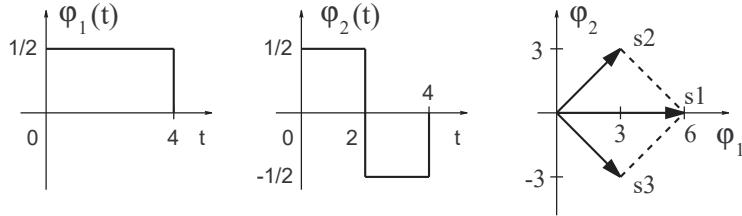
Esempio 3.5 Date le funzioni in figura $s_1(t) = 3$ per $0 < t < 4$ e zero altrove, $s_2(t) = 3$ per $0 < t < 2$ e zero altrove e $s_3(t) = 3$ per $2 < t < 4$ e zero altrove, definite nell'intervallo $(0, 4)$, trovare una base ortonormale utilizzando la procedura di Gram-Schmidt.



Poichè $\|s_1\|^2 = E_1 = 36$ possiamo ricavare subito $\varphi_1 = s_1(t)/6$. La componente di $s_1(t)$ lungo φ_1 , l'unica non nulla, vale quindi 6. Costruiamo la funzione ausiliaria

$$\theta_2(t) = s_2(t) - \int_{-\infty}^{\infty} s_2(t) \varphi_1(t) dt \varphi_1(t) = s_2(t) - 3\varphi_1(t)$$

Dato che $\|\theta_2\|^2 = 9$ si ottiene $\varphi_2(t) = \theta_2(t)/3$. La terza funzione base non c'è, dato che $s_3(t) = s_1(t) - s_2(t)$.



Le immagini delle funzioni sono: $\bar{s}_1 = (6, 0)$, $\bar{s}_2 = (3, 3)$ e $\bar{s}_3 = (3, -3)$, come mostrato in figura.

Da notare due proprietà:

- la somma dei vettori \bar{s}_2 e \bar{s}_3 è pari al vettore \bar{s}_1 , visto che $s_1(t) = s_2(t) + s_3(t)$;
- i vettori \bar{s}_2 e \bar{s}_3 formano un angolo retto (di $\pi/2$ radianti) dato che i segnali $s_2(t)$ ed $s_3(t)$ sono ortogonali (vedi sezione 3.1.1).

Lo studente è invitato a ripetere l'esercizio per i tre segnali: $x_1(t) = s_2(t)$, $x_2(t) = s_3(t)$ e $x_3(t) = s_1(t)$ verificando: a) che le funzioni base cambiano; b) che la rappresentazione nel piano delle funzioni base ruota; c) che le proprietà esposte sopra sono ancora valide.

3.1.8 Basi notevoli: Legendre e Hermite*

Polinomi di Legendre Le funzioni t^n , $n = 0, 1, \dots, N$ definite su $(-1, 1)$ generano un sottospazio che ha per base ortonormale le seguenti funzioni:

$$\begin{aligned} \varphi_0(t) &= \sqrt{1/2} \\ \varphi_1(t) &= \sqrt{3/2} t \\ \varphi_2(t) &= \sqrt{5/2} (3t^2 - 1)/2 \\ &\dots \\ \varphi_n(t) &= \sqrt{(2n+1)/2} P_n(t) \end{aligned}$$

Dove $P_n(t)$ sono i polinomi di Legendre, dati da

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} (t^2 - 1)^n \quad (3.12)$$

Polinomi di Hermite Le funzioni $t^n e^{-t^2/2}$, $n = 0, 1, \dots, N$ definite su $(-\infty, \infty)$ generano un sottospazio che ha per base ortonormale le funzioni di Hermite:

$$\varphi_n(t) = \sqrt{2^n n! \sqrt{\pi}} H_n(t) e^{-t^2/2}$$

Dove $H_n(t)$ sono i polinomi di Hermite, dati da

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n}(e^{t^2}) \quad (3.13)$$

3.2 Espansione in funzioni ortogonali

Consideriamo il problema di descrivere una funzione del tempo $g(t)$ su di un intervallo $[-T/2, T/2]$ per mezzo di un insieme discreto di coefficienti.

Usiamo una espansione in serie della forma

$$\hat{g}(t) = \sum_{k=1}^N g_k \psi_k(t) \quad -\frac{T}{2} \leq t \leq \frac{T}{2} \quad (3.14)$$

nella quale gli N coefficienti g_k dipendono dalla funzione $g(t)$ ma non dal tempo e le N funzioni $\psi_k(t)$ sono specificate indipendentemente da $g(t)$.

Il simbolo " $\hat{\cdot}$ " indica che $\hat{g}(t)$ è da considerarsi una approssimazione di $g(t)$. Se $g(t)$ si comporta "sufficientemente bene" (vedremo tra poco cosa vuol dire: comunque tutti i segnali di interesse per le telecomunicazioni si comportano bene) la differenza fra $g(t)$ e $\hat{g}(t)$ tende a zero al tendere di N all'infinito. Ovviamente anche le $\psi_k(t)$ devono essere adatte allo scopo. In alcune applicazioni si usa di proposito un numero finito di termini, ad esempio per "pulire" un segnale dal rumore di acquisizione (si usa nella stima della traiettoria di un oggetto a partire da osservazioni radar rumorose).

Come misura della differenza fra $g(t)$ e $\hat{g}(t)$ usiamo

$$\varepsilon = \|g - \hat{g}\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} |g(t) - \hat{g}(t)|^2 dt$$

ovvero l'energia del segnale differenza. Ovviamente l'errore ε è maggiore o uguale a zero. Se $\varepsilon = 0$ allora $g(t) = \hat{g}(t)$ per ogni $-T/2 \leq t \leq T/2$ (salvo il solito possibile insieme di punti a misura nulla). Data una generica funzione $g(t)$ affrontiamo ora il problema di scegliere i coefficienti g_k della sua espansione 3.14 in modo da minimizzare ε . Assumiamo che le funzioni $\psi_k(t)$ siano ortonormali.

Da notare che se noi sapessimo che, per un certo M , anche infinito, vale l'uguaglianza

$$g(t) = \sum_{k=1}^M g_k \psi_k(t) \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

allora i valori di g_k sarebbero facilmente ricavabili, in analogia con quanto già visto ed utilizzando l'ortonormalità delle $\psi_k(t)$, mediante

$$g_k = (g, \psi_k) = \int_{-\frac{T}{2}}^{\frac{T}{2}} g(t) \psi_k^*(t) dt$$

Il problema che stiamo considerando è diverso: data una serie che solo approssima $g(t)$ quali sono i valori di g_k , $k = 1, \dots, N$ che minimizzano l'errore di approssimazione? Il calcolo necessario a trovare tali g_k non è particolarmente difficile, ma non viene qui riportato (lo trovate su [Sak68] pag. 11).

La soluzione è che i g_k che minimizzano ε sono proprio

$$g_k = (g, \psi_k) = \int_{-\frac{T}{2}}^{\frac{T}{2}} g(t) \psi_k^*(t) dt$$

indipendentemente da N , numero di termini usati nella rappresentazione e qualsiasi siano le funzioni $\psi_k(t)$, purchè ortonormali. L'errore minimo di approssimazione, dipendente da N (e dalle $\psi_k(t)$) vale

$$\varepsilon_{min} = \|g\|^2 - \sum_{k=1}^N |g_k|^2 \quad (3.15)$$

Assumiamo ora di usare un numero infinito di funzioni $\varphi_k(t)$. La domanda è: sotto quali condizioni vale l'uguaglianza

$$g(t) = \sum_{k=1}^{\infty} g_k \psi_k(t) \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

ovvero, sotto quali condizioni l'errore

$$\varepsilon_N = \|g\|^2 - \sum_{k=1}^N |g_k|^2 \quad (3.16)$$

tende a zero per N che tende all'infinito? Se l'insieme delle funzioni $\psi_k(t)$ è tale che

$$\lim_{N \rightarrow \infty} \varepsilon_N = 0 \quad (3.17)$$

per ogni funzione $g(t)$ definita su $[-T/2, T/2]$ che soddisfa

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} |g(t)|^2 dt < \infty$$

si dice che l'insieme delle $\psi_k(t)$ è completo su $[-T/2, T/2]$, ed ogni funzione ad energia finita può essere rappresentata senza errori in termini delle $\psi_k(t)$.

Da notare che quando le $\psi_k(t)$ formano un insieme completo si ha l'uguaglianza fra $g(t)$ e

$$\hat{g}(t) = \sum_{k=1}^{\infty} g_k \psi_k(t)$$

nel senso che non vi è energia nella funzione errore $g(t) - \hat{g}(t)$. Possono esserci tuttavia dei punti t_i per i quali $g(t_i) \neq \hat{g}(t_i)$: ciò accade nei punti di discontinuità di $g(t)$.

Determinare se un insieme di funzioni ortonormali $\psi_k(t)$ è o non è completo è un complesso problema di analisi matematica che non tratteremo. Useremo invece insiemi completi di funzioni ortonormali, come ad esempio quelle che formano la serie di Fourier, asserendone la completezza senza provarla.

3.2.1 Teorema di Parseval

Se l'insieme delle funzioni $\psi_k(t)$, $k = 1, 2, \dots, \infty$, è completo ed ortonormale allora, per le equazioni 3.16 e 3.17, si ha

$$\|g\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} |g(t)|^2 dt = \sum_{k=1}^{\infty} |g_k|^2$$

Si osservi che

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} |g_k \psi_k(t)|^2 dt = |g_k|^2 \int_{-\frac{T}{2}}^{\frac{T}{2}} |\psi_k(t)|^2 dt = |g_k|^2$$

ovvero che $|g_k|^2$ rappresenta l'energia di $g_k \psi_k(t)$ nell'intervallo specificato.

Quindi il teorema di Parseval afferma che l'energia della funzione $g(t)$ è uguale alla somma delle energie delle funzioni $g_k \psi_k(t)$.

Il teorema di Parseval è una diretta conseguenza della completezza e della ortonormalità delle $\psi_k(t)$ e non è vero se le $\psi_k(t)$ non sono ortonormali. Da notare che un insieme di funzioni ortogonali è facilmente trasformabile in un insieme ortonormale: basta normalizzarle, dividendo ciascuna per la propria energia.

3.2.2 Serie di Fourier

Riprendiamo la serie di Fourier, già introdotta in sezione 1.7, per analizzarne i legami con quanto appena visto. Nella sua versione originale era definita mediante un particolare insieme di funzioni $\psi_k(t)$, la ben nota serie di seni e coseni di Fourier:

$$\psi_k(t) = \begin{cases} \cos\left(\frac{k-1}{2}\Omega t\right) & \text{per } k \text{ dispari} \\ \sin\left(\frac{k}{2}\Omega t\right) & \text{per } k \text{ pari} \end{cases} \quad k = 1, 2, \dots, \infty$$

dove $\Omega = 2\pi/T$. L'espansione in serie

$$g(t) = \sum_{k=1}^{\infty} g_k \psi_k(t) \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

può anche essere scritta

$$g(t) = \sum_{k=0}^{\infty} a_k \cos(k\Omega t) + \sum_{k=1}^{\infty} b_k \sin(k\Omega t) \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

E' facile verificare che questo insieme di $\psi_k(t)$ è ortogonale, anche se non è normalizzato. E' anche noto che è un insieme completo, pur se non è semplice dimostrarlo. Utilizzando le formule di Eulero (vedi 1.5), la serie di Fourier può essere rappresentata, in modo equivalente, da

$$g(t) = \sum_{k=-\infty}^{\infty} c_k e^{jk\Omega t} \quad -\frac{T}{2} \leq t \leq \frac{T}{2} \quad \Omega = \frac{2\pi}{T} \quad (3.18)$$

che è la forma ad esponenziali complessi vista in sezione 1.7. Per convenienza di notazione l'indice k varia da $-\infty$ a $+\infty$: l'alternativa è riarrangiare le funzioni in modo che k vari da 1 a $+\infty$, ma è più scomodo.

Anche l'insieme $\psi_k(t) = e^{jk\Omega t}$ è ovviamente completo ("ovviamente" se si parte dall'insieme completo di seni e coseni). Verifichiamo l'ortogonalità di tale insieme:

$$(e^{jk\Omega t}, e^{jm\Omega t}) = \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{jk\Omega t} (e^{jm\Omega t})^* dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{j(k-m)\Omega t} dt = \delta_{km} T$$

essendo al solito $\delta_{km} = 1$ se $k = m$ e 0 altrimenti. Le funzioni esponenziali non sono quindi normalizzate; dividendo ciascuna funzione per \sqrt{T} l'insieme risultante

$$\frac{1}{\sqrt{T}} e^{jk\Omega t} \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

è ortonormale.

Se espandiamo $g(t)$ nella forma

$$g(t) = \sum_{-\infty}^{\infty} g_k \frac{1}{\sqrt{T}} e^{jk\Omega t} \quad -\frac{T}{2} \leq t \leq \frac{T}{2}$$

con

$$g_k = \left(g, \frac{1}{\sqrt{T}} e^{jk\Omega t} \right) = \frac{1}{\sqrt{T}} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(t) e^{-jk\Omega t} dt$$

il legame con i coefficienti c_k dell'espansione 3.18, facilmente ricavabile, è

$$c_k = \frac{1}{\sqrt{T}} g_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(t) e^{-jk\Omega t} dt \quad (3.19)$$

E' immediato riconoscere che se $g(t)$ è una funzione reale vale

$$g_{-k} = (g_k)^* \quad \text{e} \quad c_{-k} = (c_k)^*$$

Il teorema di Parseval per la serie esponenziale di Fourier assume la forma

$$\|g\|^2 = \sum_{-\infty}^{\infty} |g_k|^2 = T \sum_{-\infty}^{\infty} |c_k|^2$$

3.2.3 Serie di Fourier per sequenze periodiche

Finora abbiamo ipotizzato che la funzione $g(t)$ sia definita solo sull'intervallo temporale $[-T/2, T/2]$ o, in modo equivalente, che il nostro interesse sia limitato a tale intervallo.

Tuttavia le funzioni $e^{jk\Omega t}$ sono periodiche, $\forall k$, con intervallo minimo comune T , dato che $\Omega = 2\pi/T$. Quindi la funzione

$$\frac{1}{\sqrt{T}} \sum_{k=-\infty}^{\infty} g_k e^{jk\Omega t}$$

è periodica di periodo T . Se perciò la funzione $g(t)$ è periodica con periodo T , la rappresentazione in serie di Fourier è valida per qualsiasi valore di t . In questo caso è anche evidente come i coefficienti g_k (o c_k) possano essere calcolati su di un qualsiasi intervallo di durata T .

$$g_k = \frac{1}{\sqrt{T}} \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} g(t) e^{-jk\Omega t} dt \quad \text{per } \tau \text{ arbitrario}$$

La potenza media di $g(t)$ è

$$P_g = \lim_{\Delta t \rightarrow \infty} \frac{1}{\Delta t} \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} |g(t)|^2 dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |g(t)|^2 dt = \frac{1}{T} \sum_{k=-\infty}^{\infty} |g_k|^2 = \sum_{k=-\infty}^{\infty} |c_k|^2 \quad (3.20)$$

che rappresenta un'altra forma del teorema di Parseval.

Le condizioni affinché una funzione $g(t)$ periodica o limitata in un intervallo in t abbia espansione in serie di Fourier e le proprietà della serie di Fourier sono argomento di altri corsi e non vengono qui ripetute (ma lo studente è tenuto a conoscerle). L'unica eccezione è il fenomeno di Gibbs.

3.2.4 Fenomeno di Gibbs

Consideriamo l'espansione parziale della funzione $g(t)$ in serie di Fourier

$$\widehat{g}_N(t) = \sum_{n=-N}^N c_n e^{jn\Omega t}$$

Supponiamo che $g(t)$ sia periodica di periodo T e che soddisfi una delle condizioni di convergenza (per esempio che $|g(t)|^2$ abbia integrale finito sul periodo oppure che $g(t)$ soddisfi le condizioni di Dirichlet, ovvero un numero finito di minimi, massimi e discontinuità per periodo e che sia assolutamente integrabile sul periodo). Se questo è vero vale che

$$\lim_{N \rightarrow \infty} \int_{-\frac{T}{2}}^{\frac{T}{2}} |g(t) - \widehat{g}_N(t)|^2 dt = 0$$

Supponiamo ora che $g(t)$ abbia una discontinuità a gradino in $t = t_0$: la somma parziale $\widehat{g}_N(t)$ (che è ovviamente una funzione continua) converge nel punto medio della discontinuità, e questo appare del tutto ragionevole. Tuttavia ai due lati della discontinuità $\widehat{g}_N(t)$ assume un comportamento oscillatorio smorzato, con periodo $\frac{T}{2N}$ e valore di picco pari a circa il 9% dell'ampiezza della discontinuità (per ogni lato), indipendentemente da N .

Quindi, al tendere di N all'infinito, le oscillazioni diventano di durata nulla nel tempo, ma restano di ampiezza costante, dando origine a due "spilli" (in inglese *spikes*) sopra e sotto la discontinuità.

Vediamo cosa succede. Per meglio renderci conto di cosa vuol dire rappresentare una funzione mediante espansione in serie e di quali sono i problemi di convergenza vediamo l'espansione in serie di Fourier della funzione periodica $v(t)$ costituita da un treno di impulsi rettangolari di durata τ , di ampiezza A e di periodo $T > \tau$. La funzione $v(t)$ è discontinua (a gradino) nei punti $t = \pm\tau/2$ e relativi periodici, per cui il valore in tali punti è *indefinito*. I segnali del mondo reale non hanno discontinuità, ma se il tempo di transizione è molto piccolo rispetto alla durata dell'impulso il modello ideale rappresenta con buona approssimazione quanto succede (le oscillazioni ci sono anche nei casi reali!).

Per calcolare i coefficienti di Fourier dell'espansione in serie di $v(t)$ utilizziamo l'equazione 3.19:

$$c_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} v(t) e^{-jk\Omega t} dt = \frac{1}{T} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} A e^{-jk\Omega t} dt = \frac{A}{T} \frac{\sin(k\Omega\tau)/2}{(k\Omega)/2}$$

o anche, essendo $\Omega = 2\pi/T$ e detta $f_0 = 1/T$,

$$c_k = Af_0\tau \operatorname{sinc}(kf_0\tau)$$

Da notare che $c_{-k} = c_k$. Inoltre, essendo $v(t)$ reale e pari, i c_k sono reali e l'espansione in serie utilizza i soli coseni. Nelle figure da 3.1 a 3.5 sono mostrate diverse espansioni in serie di Fourier di $v(t)$ che utilizzano un numero crescente di armoniche: è evidente il fenomeno di Gibbs. I valori usati per generare le figure sono: $A = 40$, $\tau = 20$ e $T = 80$.

Da notare che nei punti di discontinuità la serie passa per il valore $A/2$ e che, nei pressi di tali punti, il segnale approssimante presenta delle oscillazioni (*ripple*) attorno all'andamento del segnale $v(t)$. Nei pressi della discontinuità (salto) di ampiezza A si ottiene un picco che ha un valore massimo di circa $1.09 A$. Il valore minimo è della stessa ampiezza. Come evidente i due picchi non diminuiscono di ampiezza al crescere del numero di armoniche, solo diventano più brevi nel tempo. La presenza di tale fenomeno fa intuire che la serie di Fourier non converge uniformemente al segnale $v(t)$.

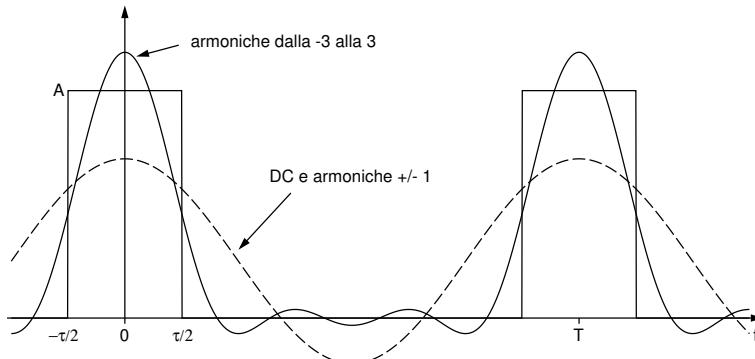


Figura 3.1: Serie di Fourier di un treno di impulsi rettangolari. Armoniche $[-1, 1]$ e $[-3, 3]$.

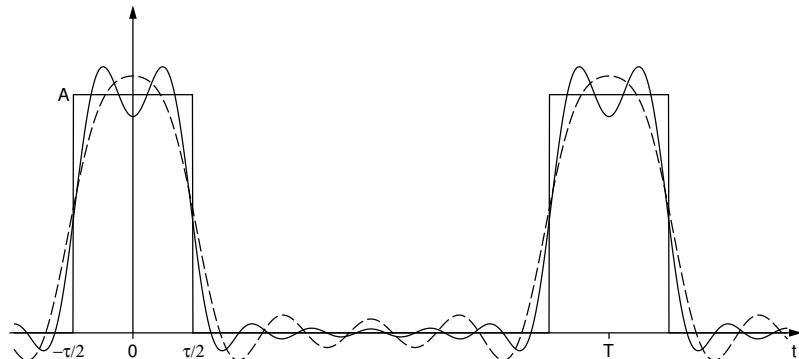


Figura 3.2: Serie di Fourier di un treno di impulsi rettangolari. Armoniche $[-5, 5]$ e $[-7, 7]$.

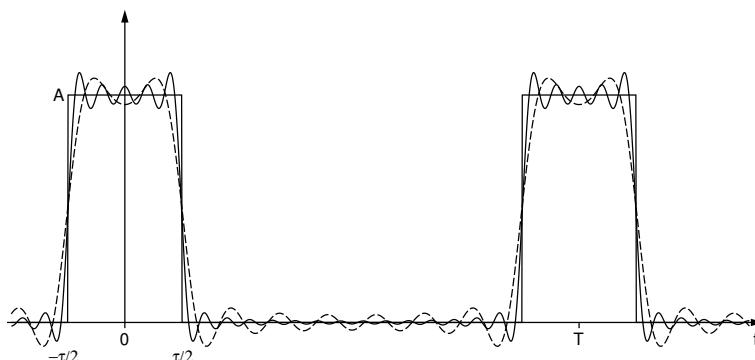
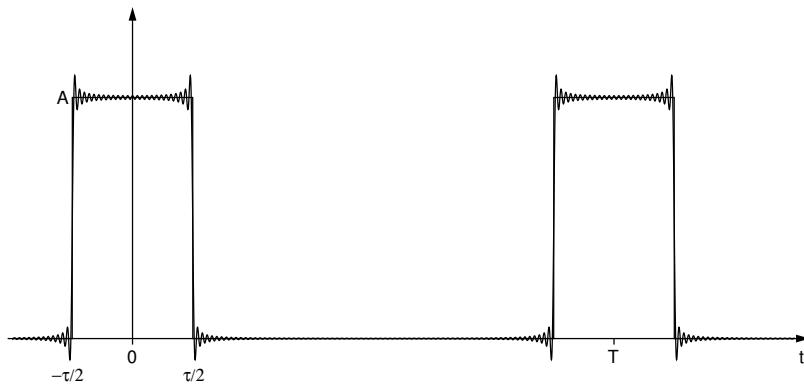
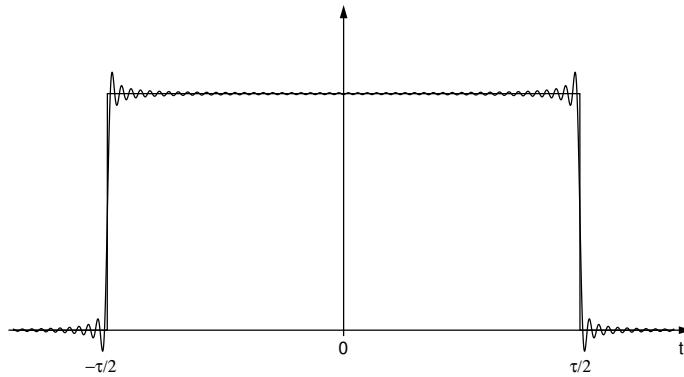


Figura 3.3: Serie di Fourier di un treno di impulsi rettangolari. Armoniche $[-9, 9]$ e $[-19, 19]$.

Figura 3.4: Espansione in serie di Fourier di un treno di impulsi rettangolari. Armoniche $[-100, 100]$.Figura 3.5: Serie di Fourier di un treno di impulsi rettangolari. Armoniche $[-200, 200]$, ingrandita.

3.2.5 Basi complete

Le abbiamo già viste, ma è utile riassumere. Si tratta di sviluppare una qualsiasi $x(t) \in L_2(a, b)$ mediante una base ortonormale $\mathcal{B} = \{\varphi_i(t)\}$ avente infinite funzioni. In generale non basta che \mathcal{B} abbia infinite funzioni per garantire che l'errore di approssimazione tenda a zero al crescere del numero delle funzioni base. Tuttavia esistono basi per cui è vero che

$$\lim_{n \rightarrow \infty} \|x(t) - \sum_{i=1}^n (x, \varphi_i) \varphi_i(t)\|^2 = 0 \quad (3.21)$$

Queste basi si dicono *complete*. Più precisamente si dice che la base $\mathcal{B} = \{\varphi_i(t)\}$ è completa se la 3.21 vale *per ogni* $x(t) \in L_2(a, b)$ (salvo le $x(t)$ a energia infinita o simili).

Nel seguito la 3.21 sarà scritta nel modo seguente

$$x(t) = \sum_{i=1}^{\infty} (x, \varphi_i) \varphi_i(t) \quad (3.22)$$

e si dirà che la serie a secondo membro converge a $x(t)$ in media quadratica. Si noti che questo non significa che la serie converge ad $x(t)$ in ogni punto, né che converge uniformemente su (a, b) .

Per inciso sono basi complete, oltre a quella di Fourier già vista, anche quelle di Legendre e di Hermite viste in sezione 3.1.8

3.2.6 Teorema del campionamento*

Il teorema el campionamento è stato dimostrato in sezione 2.1.1. Come visto, ogni funzione limitata in banda può essere rappresentata senza errori per mezzo dei suoi campioni, purchè la frequenza di campionamento sia pari ad almeno il doppio della banda della funzione. Vediamo di darne una interpretazione in termini di espansione in funzioni base ortogonali.

Un segnale $x(t)$ con spettro limitato nella banda W può essere rappresentato per $-\infty < t < \infty$ da

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \frac{\sin\{\pi f_s[t - (n/f_s)]\}}{\pi f_s[t - (n/f_s)]} \quad (3.23)$$

dove

$$a_n = f_s \int_{-\infty}^{\infty} x(t) \frac{\sin\{\pi f_s[t - (n/f_s)]\}}{\pi f_s[t - (n/f_s)]} dt \quad (3.24)$$

sono i campioni del segnale $x(t)$ presi agli istanti temporali equispaziati $t_n = n/f_s$ e dove la quantità f_s , detta frequenza di campionamento, deve rispettare la condizione $f_s \geq 2W$.

La dimostrazione si articola nelle seguenti fasi: dimostrazione che le funzioni base (le sinc) sono ortogonali, dimostrazione che gli a_n sono proprio i campioni di $x(t)$ e dimostrazione che le sinc sono una base completa sull'intero asse, ma solo per funzioni limitate in banda.

Ortogonalità delle sinc Vogliamo dimostrare che le funzioni

$$\varphi_n(t) = \frac{\sin\{\pi f_s[t - (n/f_s)]\}}{\pi f_s[t - (n/f_s)]} \quad (3.25)$$

per $-N < n < N$, definite su $L_2(-\infty, \infty)$, costituiscono un insieme ortogonale.

Si deve cioè dimostrare che

$$\int_{-\infty}^{\infty} \varphi_n(t) \varphi_m^*(t) dt = K_{nm} \delta_{nm}$$

Applicando il teorema di Parseval il primo membro diventa

$$\int_{-\infty}^{\infty} \varphi_n(t) \varphi_m^*(t) dt = \int_{-\infty}^{\infty} \Phi_n(f) \Phi_m^*(f) df$$

dove

$$\varphi_n(t) = \varphi_0(t - \frac{n}{f_s})$$

e

$$\Phi_n(f) = \mathcal{F}[\varphi_n(t)] = \Phi_0(f) e^{-j2\pi f \frac{n}{f_s}} = \frac{1}{f_s} \prod \left(\frac{f}{f_s} \right) e^{-j2\pi(nf/f_s)} \quad (3.26)$$

La trasformata di Fourier [Men79] di $\varphi_n(t)$, ovvero $\Phi_n(f)$, è un quindi rettangolo di ampiezza $1/f_s$ per $-f_s/2 < f < f_s/2$ e zero altrove. Quindi

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi_n(t) \varphi_m^*(t) dt &= \frac{1}{f_s^2} \int_{-\infty}^{\infty} \left[\prod \left(\frac{f}{f_s} \right) \right]^2 e^{-j2\pi(n-m)f/f_s} df \\ &= \frac{1}{f_s^2} \int_{-f_s/2}^{f_s/2} e^{-j2\pi(n-m)f/f_s} df = \frac{1}{f_s} \delta_{nm} \end{aligned}$$

Le funzioni $\varphi_n(t)$ sono quindi ortogonali con $K_{nm} = 1/f_s$, ovvero potrebbero essere rese ortonormali moltiplicando ciascuna per $\sqrt{f_s}$.

I coefficienti sono i campioni Visto che le sinc sono ortogonali, possiamo calcolare i coefficienti della scomposizione di $x(t)$ in tale base, ovvero gli a_n dell'eq. 3.24. Vogliamo dimostrare che sono proprio i campioni di $x(t)$. Utilizziamo nuovamente il teorema di Parseval

$$a_n = f_s \int_{-\infty}^{\infty} x(t) \varphi_n^*(t) dt = f_s \int_{-\infty}^{\infty} X(f) \Phi_n^*(f) df$$

utilizzando l'eq. 3.26 si ha

$$a_n = \int_{-f_s/2}^{f_s/2} X(f) e^{-j2\pi f(n/f_s)} df \quad (3.27)$$

Poichè $X(f)$ è zero, per ipotesi, per $f > |W|$ e $W \leq f_s/2$ gli estremi di integrazione possono essere estesi a $(-\infty, \infty)$ senza cambiare il valore dell'integrale. In questo caso l'integrale è l'antitrasformata di Fourier di $X(f)$ calcolata in $t = n/f_s$. Quindi $a_n = x(n/f_s)$. Da notare che questo risultato è vero solo se $x(t)$ è limitato in banda a W e $f_s \geq 2W$.

Le sinc formano una base completa solo per funzioni limitate in banda Sia $x(t)$ una funzione di $L_2(-\infty, \infty)$ che vogliamo espandere utilizzando come insieme di funzioni base ortonormali $2N + 1$ sinc definite in 3.25. Definiamo errore di approssimazione

$$e_N(t) = x(t) - \sum_{n=-N}^N (x, \varphi_n) \varphi_n(t) \quad (3.28)$$

Ci interessa trovare il limite di $\|e_N\|^2$ per $N \rightarrow \infty$.

La dimostrazione prosegue esprimendo $\|e_N\|^2$ in frequenza utilizzando il teorema di Parseval:

$$\|e_N\|^2 = \int_{-\infty}^{\infty} |E_N(f)|^2 df \quad (3.29)$$

essendo, dalla 3.28

$$E_N(f) = X(f) - \sum_{n=-N}^N (x, \varphi_n) \Phi_n(f) \quad (3.30)$$

Ancora per Parseval

$$(x, \varphi_n) = \int_{-\infty}^{\infty} X(f) \Phi_n^*(f) df = \int_{-\infty}^{\infty} X(\alpha) \Phi_n^*(\alpha) d\alpha \quad (3.31)$$

per cui

$$E_N(f) = X(f) - \sum_{n=-N}^N \int_{-\infty}^{\infty} X(\alpha) \Phi_n(f) \Phi_n^*(\alpha) d\alpha \quad (3.32)$$

Inseriamo l'espressione delle $\Phi_n(f)$ vista in 3.26

$$E_N(f) = X(f) - \Phi_0(f) \sum_{n=-N}^N \int_{-\infty}^{\infty} X(\alpha) \Phi_0^*(\alpha) e^{j2\pi n(\alpha-f)/f_s} d\alpha \quad (3.33)$$

facciamo tendere N all'infinito

$$E_\infty(f) = X(f) - \Phi_0(f) \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} X(\alpha) \Phi_0^*(\alpha) e^{j2\pi n(\alpha-f)/f_s} d\alpha \quad (3.34)$$

chiamiamo $\beta = \alpha - f$

$$E_\infty(f) = X(f) - \Phi_0(f) \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} X(\beta + f) \Phi_0^*(\beta + f) e^{j2\pi n\beta/f_s} d\beta \quad (3.35)$$

che può scriversi, utilizzando la formula di Poisson (vedi dopo)

$$E_\infty(f) = X(f) - f_s \sum_{n=-\infty}^{\infty} X(f + n f_s) \Phi_0(f) \Phi_0^*(f + n f_s) \quad (3.36)$$

poichè

$$\Phi_0(f) \Phi_0^*(f + n f_s) = \begin{cases} |\Phi_0(f)|^2 & n = 0 \\ 0 & n \neq 0 \end{cases}$$

e

$$|\Phi_0(f)|^2 = \begin{cases} 1/f_s & |f| < f_s/2 \\ 0 & |f| > f_s/2 \end{cases}$$

dalla 3.36 segue che

$$E_\infty(f) = \begin{cases} 0 & |f| < f_s/2 \\ X(f) & |f| > f_s/2 \end{cases} \quad (3.37)$$

In definitiva

$$\|e_\infty\|^2 = \int_{|f|>f_s/2} |X(f)|^2 df \quad (3.38)$$

ovvero $\|e_\infty\|^2$ è pari all'energia di $x(t)$ esterna alla banda $(-f_s/2, f_s/2)$. Se $x(t)$ è a banda limitata minore di $f_s/2$, cioè $X(f) = 0$ per $|f| > f_s/2$, $\|e_N\|^2$ tende a zero per N che tende all'infinito.

C.V.D.

Formula di Poisson Resta da giustificare il passaggio che porta dalla 3.35 alla 3.36: questo si ottiene mediante la formula di Poisson, ricavata, ad esempio, in [Men79] pag. 49 e 50.

$$T \sum_{k=-\infty}^{\infty} q(t - kT) = \sum_{k=-\infty}^{\infty} Q\left(\frac{k}{T}\right) e^{j2\pi kt/T} \quad (3.39)$$

essendo $Q()$ la trasformata di Fourier di $q()$.

Vale anche la duale rispetto ai domini tempo-frequenza:

$$\sum_{k=-\infty}^{\infty} Q\left(f - \frac{k}{T}\right) = T \sum_{k=-\infty}^{\infty} q(kT) e^{-j2\pi kfT} \quad (3.40)$$

dalla quale, ponendo $t = 0$ e $T = f_s$ e ricordando che

$$q(t) = \int_{-\infty}^{\infty} Q(\beta) e^{j2\pi\beta t} d\beta$$

si ottiene il passaggio cercato.

3.2.7 Espansioni che mantengono la convoluzione*

Come visto, una funzione a tempo continuo $x_a(t)$ può, in generale, essere espressa in termini di un insieme di funzioni base come

$$x_a(t) = \sum_{k=-\infty}^{\infty} c_k \varphi_k(t) \quad (3.41)$$

dove le funzioni $\{\varphi_k(t)\}$ sono un opportuno insieme di funzioni linearmente indipendenti. I coefficienti c_k possono allora essere pensati come una sequenza (segnaletico a tempo discreto) che rappresenta il segnale a tempo continuo. Se $x_a(t)$ è a banda limitata allora, per il teorema del campionamento, può essere espressa come

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{f_s}\right) \frac{\sin\{\pi f_s[t - (n/f_s)]\}}{\pi f_s[t - (n/f_s)]} \quad (3.42)$$

Questa espansione gode di una proprietà importante: mantiene la convoluzione. Cioè se:

$$x_1(t) = \sum_{k=-\infty}^{\infty} c_{1k} \varphi_k(t) \quad x_2(t) = \sum_{k=-\infty}^{\infty} c_{2k} \varphi_k(t) \quad x_3(t) = \sum_{k=-\infty}^{\infty} c_{3k} \varphi_k(t)$$

e

$$x_3(t) = \int_{-\infty}^{\infty} x_1(\tau) x_2(t - \tau) d\tau = x_1(t) * x_2(t)$$

allora

$$c_{3k} = \sum_{n=-\infty}^{\infty} c_{1n} c_{2(k-n)} = c_{1k} * c_{2k}$$

Un sistema lineare tempo invariante a tempo continuo può essere allora rappresentato da un sistema lineare a tempo discreto invariante alla traslazione.

Questa importante proprietà non appartiene al solo insieme di funzioni base visto (le sinc).

Si può dimostrare che gli insiemi di funzioni base $\{\varphi_k(t)\}$ che hanno questa proprietà sono tutti e solo quelli per cui, se $\Phi_k(j\Omega)$ è la trasformata di Fourier di $\varphi_k(t)$, vale che, per una qualche funzione $G()$,

$$\Phi_k(j\Omega) = [G(j\Omega)]^{-k}$$

3.2.8 Esempio: sulla natura di $\delta(t)^*$

Abbiamo visto in eq. 1.76 che

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega$$

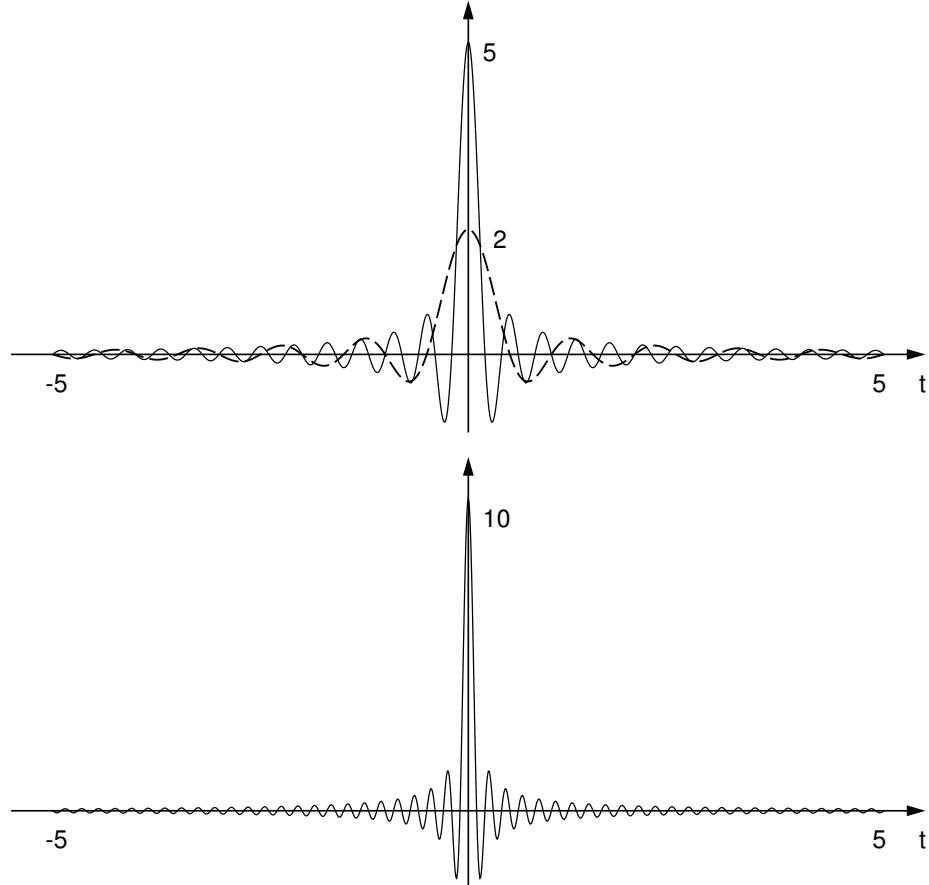
Per cercare di capire cosa succede scriviamo $\delta(t)$ nel modo equivalente

$$\delta(t) = \lim_{a \rightarrow \infty} \frac{1}{2\pi} \int_{-\pi a}^{\pi a} e^{j\omega t} d\omega \quad (3.43)$$

ed utilizziamo il cambio di variabile $z = j\omega t$, da cui $\omega = z/(jt)$, $dz = jtd\omega$, $d\omega = \frac{1}{jt}dz$.

$$\begin{aligned}\delta(t) &= \lim_{a \rightarrow \infty} \frac{1}{2\pi} \frac{1}{jt} \int_{-j\pi at}^{j\pi at} e^z dz = \lim_{a \rightarrow \infty} \frac{1}{2j\pi t} [e^{j\pi at} - e^{-j\pi at}] = \\ &= \lim_{a \rightarrow \infty} \frac{1}{\pi t} \sin(\pi at) = \lim_{a \rightarrow \infty} a \frac{\sin(\pi at)}{\pi at}\end{aligned}\quad (3.44)$$

Generando le funzioni per $a = 2, 5, 10$ possiamo renderci conto dell'andamento che porta a $\delta(t)$.



Il primo picco negativo (per $t > 0$) si ha per $\pi at = \frac{3}{2}\pi$, cioè $t = \frac{3}{2a}$, ed ha ampiezza pari a $-\frac{2a}{3\pi}$. I primi minimi e massimi vicino all'origine sono in tabella. Al crescere di a tali picchi aumentano il loro valore assoluto, tendendo all'infinito per $a \rightarrow \infty$, e si avvicinano all'origine. Quindi, al limite, il picco centrale è circondato da moltissime (infinite) oscillazioni di periodo tendente a zero ed ampiezza tendente a \pm infinito. Un vero mostro!

D'altra parte è anche possibile definire l'impulso come un rettangolo di durata in t da $-\varepsilon$ a ε ed altezza $\frac{1}{2\varepsilon}$ per $\varepsilon \rightarrow 0$.

Sembra che le due descrizioni non siano molto compatibili. Quale è quella giusta?

Il fatto è che $\delta(t)$ non è una funzione. Qualsiasi "oggetto" abbia la proprietà che, per ogni funzione $x(t)$ continua in 0,

$$\int_{-\infty}^{\infty} x(t)\delta(t)dt = x(0) \quad (3.45)$$

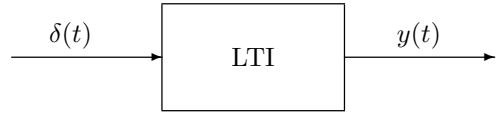
va bene. Ovvvero non ha senso chiedersi quale sia il "vero" $\delta(t)$, nè quale sia il suo valore per uno specificato istante t_0 . L'impulso di Dirac $\delta(t)$ assume significato solo se compare all'interno di un integrale, nel qual caso, con le dovute precauzioni, produce risultati privi di ambiguità.

Altre utili proprietà di $\delta(t)$, oltre alla 3.45, sono:

$$\int_{-\infty}^{\infty} x(t)\delta(t-a)dt = x(a) \quad x(t) \text{ continua in } a \quad (3.46)$$

$$\delta(t) = \delta(-t) \quad \text{e} \quad \delta(at) = \frac{1}{|a|}\delta(t) \quad (3.47)$$

Detto questo, succede di trovare l'impulso, in testi di comunicazioni elettriche, al di fuori del segno di integrale (anche in queste dispense, nel prossimo capitolo!). È ad esempio comune vedere figure del tipo:



Esse sono legittime (per degli ingegneri) considerando che il blocco esegue un integrale (di convoluzione) del segnale di ingresso: la funzione in uscita è quindi definita correttamente, anche se l'ingresso non lo è. In altre parole si accetta di usare temporaneamente $\delta(t)$ come se fosse una funzione, aspettando di incontrare un integrale che sistemi le cose...

Viene anche utilizzata la seguente relazione (vedi [Sak68] pag. 23):

$$\int_{-\infty}^{\infty} \delta(t)\delta(t+\tau)dt = \delta(\tau) \quad (3.48)$$

3.3 Sintesi ed analisi dei segnali

Abbiamo visto come un segnale $x(t)$, definito nell'intervallo (a, b) , appartenente ad uno spazio dei segnali S^m con base ortonormale $\mathcal{B} = \{\varphi_j(t), j = 1, \dots, m\}$ possa essere rappresentato da un vettore di coefficienti (numeri), detto immagine di $x(t)$, con m componenti: $\bar{x} = (x_1, x_2, \dots, x_m)$. Quindi

$$x(t) = \sum_{i=1}^m x_i \varphi_i(t) \quad (3.49)$$

Abbiamo anche visto come i coefficienti x_i si possano ricavare mediante prodotto interno:

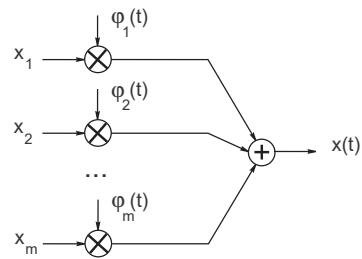
$$x_i = \int_a^b x(t) \varphi_i^*(t) dt = (x, \varphi_i) \quad (3.50)$$

In questa sezione vedremo come si possano realizzare le operazioni descritte dalle eq. 3.49 e 3.50 mediante opportuni schemi circuituali.

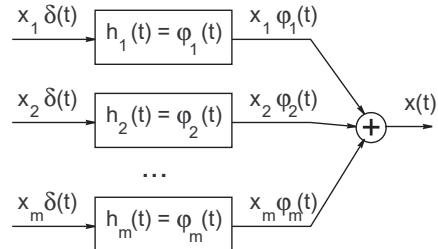
3.3.1 Sintesi

La sintesi di un segnale, descritta dall'eq. 3.49, si può realizzare in due modi.

Il primo è la trasposizione semplice dell'eq. 3.49, secondo lo schema in figura. Questo schema richiede la disponibilità di opportuni generatori in grado di produrre le funzioni $\varphi_i(t)$ all'istante desiderato, cosa di non sempre facile realizzazione.

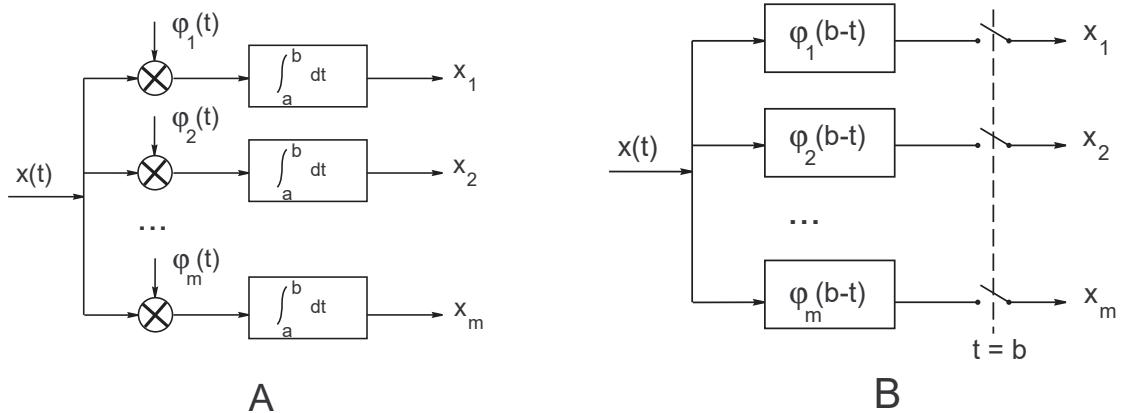


Un metodo alternativo, spesso più interessante sotto l'aspetto pratico, è quello che utilizza un banco di filtri. Se ciascun filtro ha risposta all'impulso pari ad una delle funzioni base $\varphi_i(t)$ e l'ingresso è proprio un impulso pesato mediante il relativo coefficiente x_i , la somma delle uscite dei filtri è il segnale $x(t)$.



3.3.2 Analisi

L'analisi di un segnale, descritta dall'eq. 3.50, si può anch'essa realizzare mediante due diversi schemi circuituali.



Lo schema **A** è la trasposizione semplice dell'eq. 3.50 e richiede, come per il corrispondente schema di sintesi, la disponibilità di opportuni generatori in grado di produrre le funzioni $\varphi_i(t)$ all'istante desiderato.

Lo schema **B**, noto come schema a filtri adattati, è di più semplice realizzazione. Vediamo come funziona. Il segnale $x(t)$ entra in un filtro avente risposta all'impulso $h_i(t) = \varphi_i(b - t)$: la relativa uscita vale

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h_i(t - \tau) d\tau = \int_{-\infty}^{\infty} x(\tau) \varphi_i(b - t + \tau) d\tau$$

campionare nell'istante $t = b$ produce

$$y(b) = \int_{-\infty}^{\infty} x(\tau) \varphi_i(b - b + \tau) d\tau = \int_{-\infty}^{\infty} x(\tau) \varphi_i(\tau) d\tau = x_i$$

ovvero il risultato desiderato.

3.4 Riassunto sugli spazi di funzioni

Prodotto interno di due segnali $x(t)$ e $y(t)$ definiti su (a, b)

$$(x, y) = \int_a^b x(t)y^*(t)dt$$

Se $(x, y) = 0$, i segnali sono detti **ortogonali**.

Norma

$$\|x\| = \sqrt{(x, x)} \quad \|x\|^2 = (x, x) = \int_a^b |x(t)|^2 dt$$

Se $\|x\| \neq 0$ si dice che $x(t)$ è *normalizzabile*. Per normalizzare $x(t)$ lo si divide per la sua norma $\|x\|$.

Se $\|x\| = 0 \Rightarrow x(t) = 0$ per ogni $t \in (a, b)$...

$\|x\|^2$ è l'**energia** del segnale $x(t)$ nell'intervallo (a, b) .

Distanza fra $x(t)$ e $y(t)$

$$\|x - y\| = \sqrt{\int_a^b |x(t) - y(t)|^2 dt}$$

Diseguaglianza di Schwarz

$$|(x, y)| \leq \|x\| \cdot \|y\|$$

dove il segno di uguale vale solo se $x(t) = \lambda y(t)$ per un qualche λ .

Funzioni linearmente indipendenti Date n funzioni $x_i(t)$, $i = 1, 2, \dots, n$ definite in (a, b) , sono linearmente indipendenti se ogni relazione della forma

$$\sum_{i=1}^n c_i x_i(t) \equiv 0$$

implica che $c_i = 0$ per $i = 1, 2, \dots, n$.

NOTE Se n funzioni $x_i(t)$, $i = 1, 2, \dots, n$ sono linearmente indipendenti, nessuna di esse è identicamente nulla. Se n funzioni $x_i(t)$ sono linearmente indipendenti non si può esprimere una di esse come combinazione lineare delle altre.

$$x_k(t) \neq \sum_{i \neq k} c_i x_i(t)$$

Ortogonalità ed indipendenza lineare Date n funzioni $x_i(t)$, $i = 1, 2, \dots, n$, ortogonali a due a due e tali per cui $\|x_i\| \neq 0 \forall i$, esse sono linearmente indipendenti.

Rappresentazione vettoriale delle funzioni Date n funzioni $x_i(t)$, $i = 1, 2, \dots, n$ definite su di un intervallo (a, b) , si definisce **sottospazio** \mathcal{S} di $L_2(a, b)$ generato dalle n funzioni l'insieme di tutte le combinazioni lineari

$$x(t) = \sum_{i=1}^n c_i x_i(t)$$

Problema: ogni funzione $x(t) \in \mathcal{S}$ si può descrivere in molti modi (la rappresentazione non è unica).

Soluzione: lo stesso spazio \mathcal{S} si può ottenere mediante un insieme di funzioni linearmente indipendenti $\varphi_j(t)$, $j = 1, 2, \dots, m$, dette base \mathcal{B} di \mathcal{S} .

Ci sono molte possibili basi per \mathcal{S} , ma tutte della stessa dimensione $m \leq n$. Esprimendo le funzioni in \mathcal{S}^\ddagger come combinazione lineare delle φ_j

$$x(t) = \sum_{i=1}^m x_i \varphi_i(t)$$

la rappresentazione è **unica**. I numeri x_i prendono il nome di componenti di $x(t)$ rispetto alla base \mathcal{B} . Ad ogni funzione $x(t) \in \mathcal{S}$ corrisponde uno e un solo punto $\bar{x} = (x_1, x_2, \dots, x_m)$ detto **immagine** di $x(t)$. Viceversa ad ogni punto \mathbf{x} dello spazio \mathbb{R}^m o \mathcal{C}^m corrisponde una ed una sola funzione di \mathcal{S} .

Problema: ottenere $x(t)$ da \bar{x} è facile, ottenere \bar{x} da $x(t)$ no.

Soluzione: usare una base **ortonormale**.

Una base $\mathcal{B} = \{\varphi_j(t), j = 1, \dots, m\}$ si dice ortonormale se:

$$(\varphi_i, \varphi_j) = \delta_{ij} \quad i, j = 1, 2, \dots, m$$

essendo $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$.

Se la base \mathcal{B} è ortonormale il calcolo delle coordinate dell'immagine di $x(t) \in \mathcal{S}$ si ottiene facilmente

$$x_k = \int_a^b x(t) \varphi_k^*(t) dt = (x, \varphi_k)$$

Proprietà

Energia L'energia di $x(t)$ è pari al quadrato della distanza della sua immagine dall'origine degli assi (teorema di Parseval): $|\bar{x}| = \|x\|$. Analogamente la distanza dell'immagine di $x(t)$ dall'origine è pari alla norma $\|x\|$ di $x(t)$.

Distanza La distanza fra le funzioni $x(t)$ e $y(t)$, ovvero $\|x - y\|$ è pari alla distanza euclidea (geometrica) delle loro immagini.

Prodotto interno Il prodotto interno delle funzioni $x(t)$ e $y(t)$, $\bar{x} \cdot \bar{y}$, è pari al prodotto scalare dei vettori immagine delle due funzioni.

Procedimento di Gram-Schmidt E' un metodo che consente di trovare una base ortonormale di un sottospazio \mathcal{S} generato da un insieme noto di funzioni $\mathcal{X} = \{x_i(t), i = 1, \dots, n\}$ in generale né ortogonali né linearmente indipendenti.

Sintesi ed analisi dei segnali

$$x(t) = \sum_{i=1}^m x_i \varphi_i(t) \quad x_i = \int_a^b x(t) \varphi_i^*(t) dt = (x, \varphi_i)$$

Capitolo 4

Segnali e sistemi a tempo discreto

Un **segnale a tempo discreto** è un segnale che esiste solo in un insieme discreto di istanti temporali. Ad esempio, un segnale analogico $x(t)$ campionato con frequenza di campionamento $f_c = 1/T_c$ origina il segnale (campionato) a tempo discreto $x(nT_c)$, che esiste solo negli infiniti istanti nT_c , con n intero. Un segnale a tempo discreto può tuttavia non essere il campionamento di un segnale analogico: ad esempio può essere generato a calcolatore a partire da un algoritmo. Per avere la massima generalità indicheremo quindi un segnale a tempo discreto con la notazione $x(n)$: in generale $x(n)$ è una sequenza di infiniti numeri che possono o meno essere quantizzati. La sequenza $x(n)$ di numeri reali o complessi è definita per ogni numero intero n : l'indice n è detto anche tempo discreto. Ovviamente non sempre n è un indice temporale, tuttavia lo indicheremo comunque, per semplicità, come "tempo discreto". Se i valori possibili dei numeri $x(n)$ sono discreti, ovvero se la sequenza è quantizzata in ampiezza, si parla di *segnali numerici* o digitali.

Un **sistema a tempo discreto** è definito come una regola per trasformare un segnale (in ingresso) a tempo discreto $x(n)$ in un segnale a tempo discreto $y(n)$ in uscita.

4.1 Segnali a tempo discreto

Definiamo due sequenze di uso comune:

$$\delta(n) = \begin{cases} 0 & n \neq 0 \\ 1 & n = 0 \end{cases} \quad u(n) = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases}$$

$\delta(n)$ è la *sequenza campione unitario* o impulso a tempo discreto (da notare che l'impulso a tempo discreto non presenta le complicazioni matematiche dell'impulso a tempo continuo).

$u(n)$ è la *sequenza gradino unitario*. Le due sequenze sono legate dalle relazioni

$$u(n) = \sum_{k=-\infty}^n \delta(k) \quad \delta(n) = u(n) - u(n-1)$$

Poichè vale

$$\delta(n-k) = \begin{cases} 0 & n \neq k \\ 1 & n = k \end{cases} \tag{4.1}$$

una arbitraria sequenza $f(n)$ può essere scritta come somma pesata di sequenze impulso

$$f(n) = \sum_{k=-\infty}^{\infty} f(k)\delta(n-k) \tag{4.2}$$

Definizioni

Una *sequenza esponenziale reale* è una sequenza i cui valori sono della forma a^n con a reale.

Una sequenza $x(n)$ è *periodica* di periodo N se $x(n) = x(n+N)$ per ogni n . Il periodo N è intero ed è il più piccolo intero per cui vale la relazione precedente.

Una *sequenza sinusoidale* ha valori della forma $A \cos(\omega_0 n + \phi)$ ed una *sequenza esponenziale complessa* ha valori $Ae^{j\omega_0 n}$.

Le sequenze sinusoidale ed esponenziale complessa sono periodiche di periodo $2\pi/\omega_0$ se questo numero è intero. Sono periodiche di periodo maggiore se $2\pi/\omega_0$ è un numero razionale e non sono periodiche

altrimenti. Chiameremo ω_0 la pulsazione (o frequenza in radienti) di tali sequenze siano o meno periodiche. ω_0 può essere scelta in modo arbitrario, ma non si perde in generalità considerandola vincolata ad essere in $[0, 2\pi]$ (o in $[-\pi, \pi]$), in quanto, ad esempio, $Ae^{j\omega_0 n} = Ae^{j(\omega_0 + k2\pi)n}$ per ogni k intero.

Si dice che la sequenza y è una versione *ritardata* o *traslata* della sequenza x se $y(n) = x(n - n_0)$, dove n_0 è un intero.

4.2 Trasformata di Fourier per segnali a tempo discreto

La trasformata di Fourier della generica sequenza $x(n)$ è definita come

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} \quad (4.3)$$

Dalla definizione è evidente che la funzione $X(e^{j\omega})$ è continua in ω e *periodica* di periodo pari a 2π e quindi può essere espansa in serie di Fourier. Ricordando la definizione di serie di Fourier (equazioni 1.85 e 1.86) si nota che la 4.3 è proprio la sua espansione in serie, e che quindi i coefficienti sono

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega \quad (4.4)$$

Vediamolo meglio: una generica funzione $s(\omega)$ periodica di periodo 2π ha espansione in serie di Fourier (dalla 1.85)

$$s(\omega) = \sum_{n=-\infty}^{\infty} c_n e^{-jn\omega_0\omega}$$

e poichè $\omega_0 = 2\pi/T_0 = 2\pi/2\pi = 1$ la 4.3 è proprio una serie di Fourier.

Le relazioni 4.3 e 4.4 valgono per qualsiasi sequenza, purchè la 4.3 converga, e sono le definizioni della trasformata di Fourier **diretta** e **inversa** per sequenze. Quindi la trasformata di Fourier per una sequenza non è altro che la serie di Fourier che si ottiene utilizzando i valori della sequenza come coefficienti e imponendo che il periodo sia 2π .

Come nel caso continuo, vi sono diversi criteri di convergenza: uno è quello della sommabilità in valore assoluto,

$$\sum_{n=-\infty}^{\infty} |x(n)| < \infty$$

Un altro criterio di esistenza è che la sequenza sia ad energia finita, ovvero che

$$\sum_{n=-\infty}^{\infty} |x(n)|^2 < \infty$$

Quest'ultima condizione non assicura la convergenza uniforme.

Fra le proprietà della trasformata di Fourier per sequenze vi sono le seguenti:

1. se $x(n)$ è reale, $X(e^{j\omega})$ ha simmetria hermitiana, ovvero $X(e^{j\omega}) = X^*(e^{-j\omega})$;
2. se $x(n)$ è reale e pari (ossia $x(n) = x(-n)$), allora $X(e^{j\omega})$ è pari e reale;
3. se $x(n)$ è reale e dispari (ossia $x(n) = -x(-n)$), allora $X(e^{j\omega})$ è dispari e immaginaria pura.

4.2.1 Esempi di trasformate di segnali a tempo discreto

Vediamo alcuni esempi di trasformate di Fourier di sequenze tipiche.

Sequenza rettangolare

Sia $x(n)$ la sequenza rettangolare alta uno di durata $2N$: $x(n) = 1$ per $n \in [-N, N]$ e zero altrove. La sua trasformata di Fourier vale allora (vedi 4.3);

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} = \sum_{n=-N}^{N} e^{-j\omega n}$$

Utilizzando il cambio di variabile $m = n + N$, da cui $n = m - N$, si ottiene

$$X(e^{j\omega}) = \sum_{n=-N}^N e^{-j\omega n} = \sum_{m=0}^{2N} e^{-j\omega(m-N)} = e^{j\omega N} \sum_{m=0}^{2N} e^{-j\omega m}$$

e ricordando la formula della serie geometrica di ragione α

$$\sum_{k=0}^{M-1} \alpha^k = \begin{cases} \frac{1-\alpha^M}{1-\alpha} & \forall \alpha \text{ complesso } \neq 1 \\ M & \text{se } \alpha = 1 \end{cases}$$

sostituendo $\alpha = e^{-j\omega}$ si ottiene

$$X(e^{j\omega}) = e^{j\omega N} \frac{1 - e^{-j\omega(2N+1)}}{1 - e^{-j\omega}} = \frac{e^{j\omega N} - e^{-j\omega(N+1)}}{e^{-j\omega/2} (e^{j\omega/2} - e^{-j\omega/2})}$$

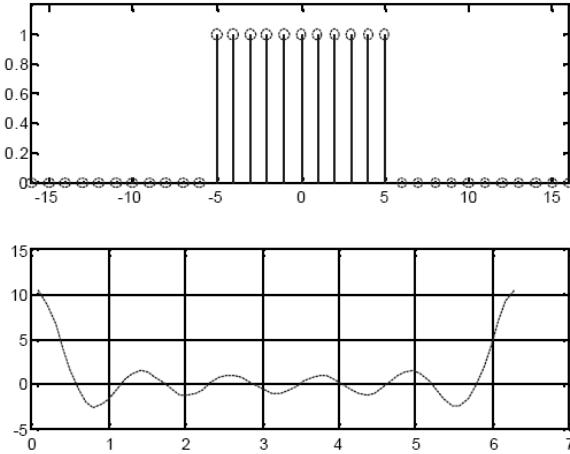
ovvero

$$X(e^{j\omega}) = e^{j\omega/2} \frac{e^{j\omega N} - e^{-j\omega(N+1)}}{e^{j\omega/2} - e^{-j\omega/2}} = \frac{e^{j\omega(N+1/2)} - e^{-j\omega(N+1/2)}}{e^{j\omega/2} - e^{-j\omega/2}}$$

e quindi

$$X(e^{j\omega}) = \frac{\sin[\omega(N+1/2)]}{\sin(\omega/2)}$$

La sequenza rettangolare e la sua trasformata per il caso $N = 5$ sono in figura: è visualizzato solo un periodo della trasformata (che è ovviamente periodica di periodo 2π).



Impulsi e costanti

La FT della sequenza impulso unitario $\delta(n)$ è immediata:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \delta(n) e^{-j\omega n} = e^{-j\omega 0} = 1 \quad (4.5)$$

e per un impulso traslato

$$\delta(n - K) \longleftrightarrow e^{-j\omega K} \quad (4.6)$$

che, per inciso, è la solita regola della traslazione. Meno diretta è la trasformata del segnale discreto $x(n) = 1 \forall n$, ovvero della sequenza costante di ampiezza unitaria: la definizione produce

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} 1 e^{-j\omega n} \quad (4.7)$$

che non è sommabile in senso ordinario. Per calcolarla consideriamo il segnale periodico di periodo T costituito dal treno di impulsi

$$d(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT)$$

Essendo $d(t)$ periodico si può espandere in serie di Fourier (vedi 1.85):

$$d(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad \omega_0 = \frac{2\pi}{T}$$

dove i coefficienti, dalla 1.86, sono

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} d(t) e^{-jn\omega_0 t} dt = \frac{1}{T}$$

Quindi

$$d(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT) = \sum_{n=-\infty}^{\infty} \frac{1}{T} e^{jn\omega_0 t} \quad (4.8)$$

Da questa si ricava una relazione fondamentale:

$$\sum_{k=-\infty}^{\infty} \delta(t - kT) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{j2\pi nt/T} \quad (4.9)$$

Possiamo ora trovare la FT della sequenza costante $x(n) = 1 \forall n$, che vale (dalla 4.7 sostituendo $\omega = 2\pi f$)

$$\sum_{n=-\infty}^{\infty} 1 e^{-j2\pi fn} = \sum_{k=-\infty}^{\infty} \delta(f - k) \quad (4.10)$$

che quindi esiste, nel senso delle distribuzioni, come un treno di impulsi in f (o ω , se preferite).

4.2.2 Trasformata di Fourier di un segnale campionato

Supponiamo di avere un segnale analogico $x_a(t)$ e di considerare il segnale a tempo discreto $x(n) = x_a(nT_c)$ che si ottiene dal campionamento di $x_a(t)$ con periodo T_c (o frequenza $f_c = 1/T_c$).

Che relazioni vi sono fra la trasformata di Fourier di $x_a(t)$ e quella della sequenza $x(n)$?

La trasformata di Fourier di $x_a(t)$ è definita da

$$X_a(\omega) = \int_{-\infty}^{\infty} x_a(t) e^{-j\omega t} dt \quad x_a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_a(\omega) e^{j\omega t} d\omega \quad (4.11)$$

Quindi si può scrivere, utilizzando un semplice cambio di variabile, che

$$\begin{aligned} x(n) &= x_a(nT_c) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_a(\omega) e^{j\omega nT_c} d\omega \\ &= \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \int_{(2l-1)\pi/T_c}^{(2l+1)\pi/T_c} X_a(\omega) e^{j\omega nT_c} d\omega \\ &= \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \int_{-\pi/T_c}^{\pi/T_c} X_a(\omega + \frac{2\pi l}{T_c}) e^{j\omega nT_c} e^{j2\pi ln} d\omega \end{aligned} \quad (4.12)$$

e poichè $e^{j2\pi ln} = 1$ si ottiene, invertendo l'ordine di sommatoria ed integrale,

$$x(n) = \frac{1}{2\pi} \int_{-\pi/T_c}^{\pi/T_c} \left[\sum_{l=-\infty}^{\infty} X_a(\omega + \frac{2\pi l}{T_c}) \right] e^{j\omega nT_c} d\omega \quad (4.13)$$

chiamando $\Omega = \omega T_c$

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{1}{T_c} \sum_{l=-\infty}^{\infty} X_a(\frac{\Omega}{T_c} + \frac{2\pi l}{T_c}) \right] e^{j\Omega n} d\Omega \quad (4.14)$$

Confrontando infine con la definizione di FT per sequenze inversa (eq. 4.4) si arriva a

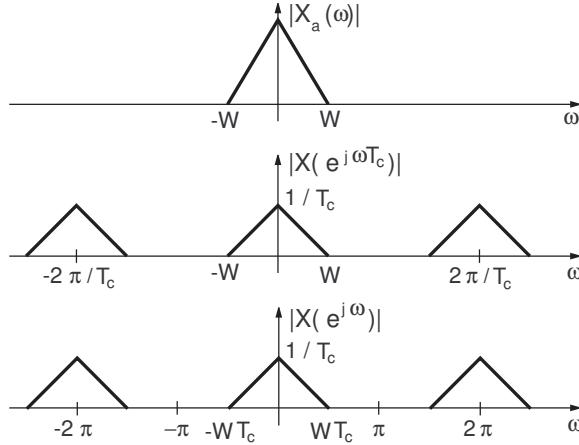
$$X(e^{j\omega}) = \frac{1}{T_c} \sum_{l=-\infty}^{\infty} X_a(\frac{\omega}{T_c} + \frac{2\pi l}{T_c}) \quad (4.15)$$

o anche

$$X(e^{j\omega T_c}) = \frac{1}{T_c} \sum_{l=-\infty}^{\infty} X_a(\omega + \frac{2\pi l}{T_c}) \quad (4.16)$$

Quindi la trasformata di Fourier del segnale a tempo discreto $x(n)$ ottenuto compionando con frequenza $f_c = 1/T_c$ il segnale analogico $x_a(t)$ è la somma di infinite repliche di $X_a(\omega)$ scalate e traslate di multipli di $2\pi/T_c = 2\pi f_c$.

Se $x_a(t)$ è limitata in banda, ossia se $X_a(\omega) = 0$ per $|\omega| \geq W$ e se $2\pi/T_c = 2\pi f_c \geq 2W$, cioè se $f_c \geq W/\pi$, le repliche non si sovrappongono ed il segnale $x_a(t)$ può essere esattamente ricostruito dai suoi campioni $x(n) = x_a(nT_c)$, come mostrato in figura.



Chiariamo meglio: $x(e^{j\omega})$ è periodica di periodo 2π . Consideriamo $\omega \in [-\pi, \pi]$, ossia un periodo di $x(e^{j\omega})$. In eq. 4.15 si vede che $X_a()$ dipende da ω/T_c e quindi un periodo in omega corrisponde all'intervallo

$$\frac{\omega}{T_c} \in \left[-\frac{\pi}{T_c}, \frac{\pi}{T_c} \right] = [-\pi f_c, \pi f_c] = [-2\pi \frac{f_c}{2}, 2\pi \frac{f_c}{2}]$$

Se $X_a(\omega) = 0$ per $|\omega| \geq W$ allora $X_a(\omega/T_c) = 0$ per $|\omega/T_c| \geq W$, cioè $|\omega| \geq WT_c$. Se vale la condizione di Nyquist, cioè se $W \leq 2\pi f_c/2 = \pi/T_c$ allora $WT_c \leq \pi$ e le repliche non si sovrappongono.

Per inciso, questa è un'altra **dimostrazione del teorema del campionamento**, probabilmente la migliore delle tre viste finora.

4.3 Sistemi a tempo discreto

Un sistema a tempo discreto è definito come una regola per trasformare una sequenza in ingresso $x(n)$ in una sequenza in uscita $y(n)$. Questo si può indicare come

$$y(n) = L[x(n)] \quad \xrightarrow{x(n)} [L[\quad]] \xrightarrow{y(n)}$$

Esempio 4.1 Vediamo alcuni semplici esempi di sistemi a tempo discreto:

- | | |
|----------------------------|---|
| $y(n) = x^2(n)$ | Sistema non lineare. Il valore di y in un particolare n_0 dipende solo dal valore di $x(n_0)$ (sistema senza memoria) |
| $y(n) = nx(n)$ | Sistema lineare, senza memoria, tempo variante |
| $y(n) = 2x(n) + 3x(n - 1)$ | Sistema lineare, invariante alla traslazione, a memoria finita |

Imponendo dei vincoli sulla trasformazione $L[\]$ si possono definire delle classi di sistemi a tempo discreto.

4.3.1 Linearità e tempo invarianza

Un sistema L è **lineare** se vale il principio della sovrapposizione degli effetti

$$L[ax_1(n) + bx_2(n)] = aL[x_1(n)] + bL[x_2(n)] = ay_1(n) + by_2(n)$$

per ogni $a, b, x_1(n)$ e $x_2(n)$.

Un sistema L è **invariante alla traslazione** (o tempo invariante) se

$$L[x(n - k)] = y(n - k)$$

per ogni k intero. In altre parole una traslazione della sequenza di ingresso provoca la stessa traslazione nella sequenza di uscita.

Esercizio 4.2 Dimostrare che il sistema $y(n) = |x(n)|$ è non lineare e invariante alla traslazione.

Dimostrare che il sistema $y(n) = nx(n)$ è lineare e non invariante alla traslazione.

Nel seguito del capitolo tratteremo solo sistemi lineari invarianti alla traslazione, o **LTI**, senza doverlo per forza ripetere ogni volta e a meno che non sia diversamente indicato.

4.3.2 Risposta all'impulso

Chiameremo $h(n)$ la risposta di un sistema alla sequenza impulso $\delta(n)$

$$h(n) = L[\delta(n)]$$

Notiamo che la sequenza $h(n)$ non è necessariamente nulla per $n < 0$. Se $h(n) = 0$ per $n < 0$ il sistema si dice **causale**.

Convoluzione discreta Se un sistema è lineare si può esprimere la sua uscita, ricordando la 4.2, come

$$y(n) = L[x(n)] = L \left[\sum_{k=-\infty}^{\infty} x(k) \delta(n-k) \right] = \sum_{k=-\infty}^{\infty} x(k) L[\delta(n-k)]$$

L'ultimo passaggio deriva dalla linearità, il penultimo è sempre vero (anche per non lineari). Se il sistema è invariante alla traslazione si ottiene

$$y(n) = \sum_{k=-\infty}^{\infty} x(k) h(n-k) = x(n) * h(n) \quad (4.17)$$

che è chiamata somma di convoluzione La 4.17 vale quindi *solo per sistemi LTI*.

La somma di convoluzione è commutativa: l'ordine con il quale la si esegue non conta, per cui $x(n) * h(n) = h(n) * x(n)$ e la 4.17 si può equivalentemente scrivere come

$$y(n) = \sum_{k=-\infty}^{\infty} x(n-k) h(k) \quad (4.18)$$

Se sia il sistema che il suo ingresso sono *causal*, ovvero se $h(n) = 0$ e $x(n) = 0$ per $n < 0$, l'uscita, causale anch'essa, si ottiene da

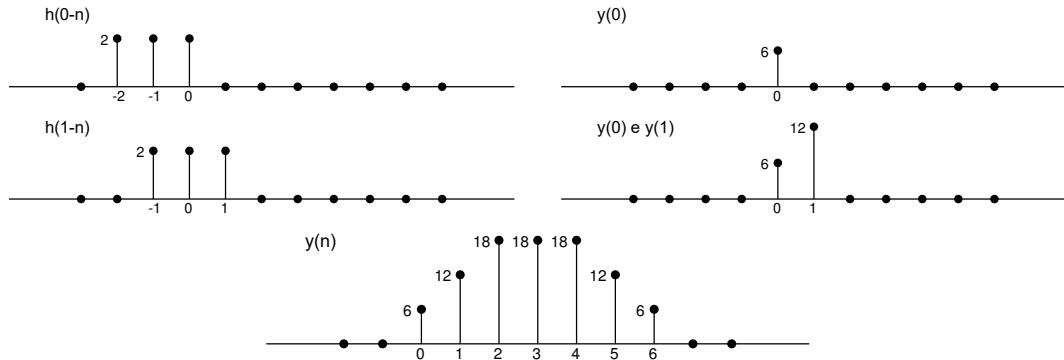
$$y(n) = \sum_{k=0}^n x(k) h(n-k) = \sum_{k=0}^n x(n-k) h(k)$$

Esempio 4.3 Calcolare la sequenza di uscita di un sistema sapendo che:

$$h(n) = 2 \cdot [u(n) - u(n-3)] \quad x(n) = 3 \cdot [u(n) - u(n-5)]$$

Le due sequenze di partenza ed il procedimento grafico per il calcolo della loro convoluzione sono illustrate nelle figure seguenti. Partendo dalla 4.17 si ribalta $h(n)$ e lo si trasla di zero ottenendo $h(0-n)$: moltiplicando per $x(n)$ rimane un solo valore in zero, e cioè $3 \cdot 2 = 6$, che è il valore della convoluzione in zero $\rightarrow y(0) = 6$. Quindi si ribalta $h(n)$ e lo si trasla di uno, ottenendo $h(1-n)$: moltiplicando per $x(n)$ rimangono solo due valori, in zero e uno, entrambi pari a 6. La somma dei due campioni è 12, e quindi $y(1) = 12$. Iterando il procedimento si trovano tutti i valori di $y(n)$.





La lunghezza (numero di valori non nulli) di $x(n)$ è 5, quella di $h(n)$ è 3 e quella della loro convoluzione $y(n)$ vale 7. Questa regola è generale: la lunghezza della convoluzione è la somma delle lunghezze meno uno. È anche evidente come la convoluzione discreta e quella continua si comportino in modo simile: nel caso continuo la convoluzione di due rettangoli è un trapezio, come nel caso discreto.

4.3.3 Stabilità

Abbiamo visto che i vincoli di linearità e di invarianza alla traslazione definiscono una classe di sistemi che sono rappresentati dalla somma di convoluzione. Gli ulteriori vincoli di stabilità e causalità definiscono una classe più ristretta di sistemi di importanza pratica: i cosiddetti sistemi realizzabili. Definiamo *sistema stabile* (BIBO) un sistema per cui ogni ingresso limitato in ampiezza provoca una uscita limitata in ampiezza. I sistemi discreti LTI sono stabili *se e solo se*

$$S = \sum_{n=-\infty}^{\infty} |h(n)| < \infty \quad (4.19)$$

Dimostriamolo. Se la 4.19 è vera e $x(n)$ è limitata, cioè $|x(n)| < M$ per ogni n , allora dalla 4.18 segue che

$$|y(n)| = \left| \sum_{k=-\infty}^{\infty} h(k)x(n-k) \right| < M \sum_{k=-\infty}^{\infty} |h(k)| < \infty$$

Perciò y è limitata. Si dimostra che la condizione 4.19 è anche necessaria facendo vedere che se $S = \infty$ allora si può trovare un ingresso limitato che dà luogo ad una uscita non limitata. Tale è la sequenza

$$x(n) = \begin{cases} \frac{h^*(-n)}{|h(-n)|} & h(-n) \neq 0 \\ 0 & h(-n) = 0 \end{cases}$$

Chiaramente $x(n)$ è limitata, visto che $|x(n)| = 1$ per $h(-n) \neq 0$ e zero altrimenti. Il valore dell'uscita per $n = 0$ è

$$y(0) = \sum_{k=-\infty}^{\infty} x(-k)h(k) = \sum_{k=-\infty}^{\infty} \frac{|h(k)|^2}{|h(k)|} = S$$

Perciò se $S = \infty$, la sequenza di uscita non è limitata.

Esercizio 4.4 Dire per quali a il sistema avente risposta all'impulso $h(n) = a^n u(n)$ è stabile.

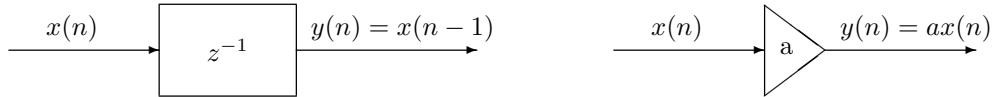
Soluzione: bisogna verificare per quali valori di a la somma

$$S = \sum_{n=-\infty}^{\infty} |h(n)| = \sum_{k=0}^{\infty} |a|^k$$

è limitata. Se $|a| < 1$ la somma converge a $1/(1 - |a|)$ ed il sistema è stabile. Se $|a| \geq 1$ la somma non converge ed il sistema non è stabile.

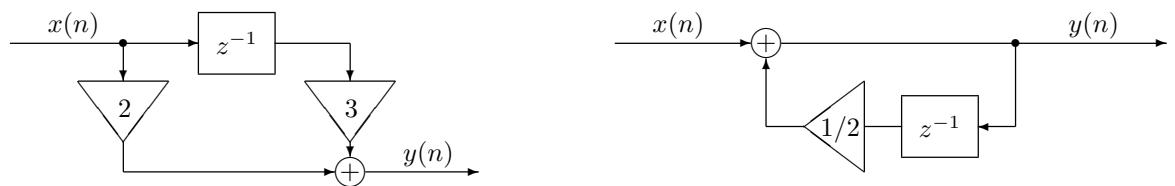
Sistemi discreti in scatola di montaggio

Vi sono due semplici sistemi di interesse particolare: l'elemento di ritardo ed il moltiplicatore. Il primo è definito dalla relazione $y(n) = x(n - 1)$, ovvero l'uscita è pari all'ingresso ritardato di uno. Il secondo è definito dalla relazione $y(n) = a x(n)$, ovvero l'uscita è pari all'ingresso moltiplicato per il numero (solitamente reale) a . Questi sistemi saranno rappresentati come in figura, la lettera a nel triangolo che rappresenta il moltiplicatore è il suo *guadagno*. Il significato di z^{-1} nel blocco di ritardo verrà chiarito in seguito.



Si può dimostrare che un arbitrario sistema lineare invariante alla traslazione e causale può essere realizzato combinando insieme blocchi di ritardo e moltiplicatori. Basta quindi avere a disposizione questi blocchi per costruire un qualsiasi sistema realizzabile. Per i sistemi non causali serve anche un "blocco anticipatore"...

Esercizio 4.5 Trovare la risposta $h(n)$ alla sequenza impulso dei due sistemi in figura.



Soluzione: $h_1(n) = 2\delta(n) + 3\delta(n - 1)$ e $h_2(n) = (1/2)^n u(n)$.

Esercizio 4.6 Costruire il sistema che realizza la risposta all'impulso dell' esercizio 4.3.

4.3.4 Funzione di trasferimento e trasformata z

Inviamo ad un sistema lineare invariante alla traslazione la sequenza esponenziale

$$x(n) = r^n$$

Possiamo calcolare l'uscita, conoscendo la risposta all'impulso $h(n)$, mediante la 4.18

$$y(n) = \sum_{k=-\infty}^{\infty} r^{n-k} h(k) = r^n \sum_{k=-\infty}^{\infty} h(k) r^{-k} \quad (4.20)$$

La sequenza di uscita è anch'essa esponenziale, moltiplicata per il valore $H(r)$ della **trasformata z**

$$H(z) = \sum_{n=-\infty}^{\infty} h(n) z^{-n} \quad (4.21)$$

della sequenza $h(n)$. Possiamo quindi concludere che

$$L[z^n] = H(z) z^n \quad (4.22)$$

per ogni z , reale o complessa, per cui la serie dell'equazione 4.21 converge. Chiameremo $H(z)$ la *funzione di trasferimento* in z del sistema (in inglese è anche detta *system function*).

La funzione di trasferimento dell'elemento di ritardo si trova come segue: se $x(n) = z^n$, allora $y(n) = x(n - 1) = z^{n-1} = z^{-1} z^n$. Quindi $H(z) = z^{-1}$.

Per il moltiplicatore un analogo ragionamento porta a $H(z) = a$.

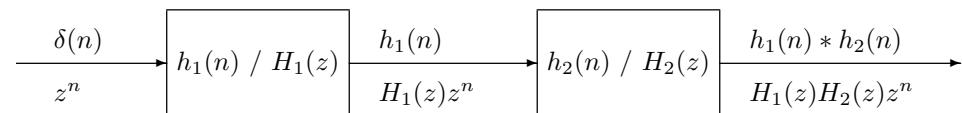


Figura 4.1: Sistemi discreti in cascata.

Sistemi in cascata Due sistemi lineari invarianti alla traslazione in cascata (vedi fig. 4.1) formano un sistema con risposta alla sequenza impulso pari alla convoluzione discreta delle due risposte. Dal momento che l'ordine in cui si esegue la convoluzione non conta, si ha che non conta l'ordine in cui sono posti i due sistemi in cascata. Chiamando $h(n)$ e $H(z)$ la risposta all'impulso e la funzione di trasferimento del sistema in cascata in fig. 4.1, si ha

$$h(n) = h_1(n) * h_2(n) \quad H(z) = H_1(z)H_2(z) \quad (4.23)$$

Convoluzione nel dominio della trasformata z Per definizione $H(z)$, $H_1(z)$ e $H_2(z)$ nella 4.23 sono le trasformate z di $h(n)$, $h_1(n)$ e $h_2(n)$ rispettivamente. Siccome poi le sequenze $h_1(n)$ e $h_2(n)$ sono arbitrarie, ne segue che la trasformata z della convoluzione di due sequenze è il prodotto delle loro trasformate z . Quindi, poichè

$$y(n) = x(n) * h(n) \quad \text{allora} \quad Y(z) = X(z)H(z) \quad (4.24)$$

Sistemi in parallelo Due sistemi lineari invarianti alla traslazione in parallelo formano un sistema con risposta alla sequenza impulso pari alla somma delle due risposte e funzione di trasferimento pari alla somma delle risposte:

$$h(n) = h_1(n) + h_2(n) \quad H(z) = H_1(z) + H_2(z) \quad (4.25)$$

4.3.5 Funzione di trasferimento e trasformata di Fourier

E' possibile ricavare la funzione di trasferimento di un sistema discreto anche nel dominio della trasformata di Fourier per sequenze. Ripetiamo quindi la trattazione svolta usando il nuovo dominio. Inviamo ad un sistema lineare invariante alla traslazione la sequenza esponenziale complessa con pulsazione ω

$$x(n) = e^{j\omega n} \quad \text{per} \quad n \in (-\infty, \infty)$$

Possiamo calcolare l'uscita, conoscendo la risposta all'impulso $h(n)$, mediante la 4.18

$$y(n) = \sum_{k=-\infty}^{\infty} e^{j\omega(n-k)} h(k) = e^{j\omega n} \sum_{k=-\infty}^{\infty} h(k) e^{-j\omega k} \quad (4.26)$$

La sequenza di uscita è anch'essa esponenziale complessa della stessa pulsazione ω , moltiplicata per il valore $H(e^{j\omega})$ della *trasformata di Fourier*

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n) e^{-j\omega n} \quad (4.27)$$

della sequenza $h(n)$. Possiamo quindi concludere che

$$L[e^{j\omega n}] = H(e^{j\omega}) e^{j\omega n} \quad (4.28)$$

Chiameremo anche $H(e^{j\omega})$ *funzione di trasferimento* (in omega) o *risposta in frequenza* del sistema. La trasformata z di una sequenza è fortemente legata alla trasformata di Fourier della stessa sequenza. I valori della trasformata z calcolata per $z = e^{j\omega}$ (ovvero sul circolo di raggio unitario centrato nell'origine del piano complesso) sono infatti i valori della trasformata di Fourier della stessa sequenza (ammesso ovviamente che siano soddisfatti i criteri di convergenza). La trasformata z è quindi interamente determinata dai suoi valori sul circolo unitario. Si può dire che, per i segnali a tempo discreto, la trasformata z sia una generalizzazione della trasformata di Fourier, così come, per i segnali a tempo continuo, la trasformata di Laplace è una generalizzazione della trasformata di Fourier.

Ovviamente anche nel dominio di Fourier la convoluzione si ottiene mediante prodotto delle trasformate, e la 4.24 diventa: se

$$y(n) = x(n) * h(n) \quad \text{allora} \quad Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega}) \quad (4.29)$$

4.4 Simulazione discreta di sistemi continui

Consideriamo un sistema continuo con funzione di trasferimento $H_c(\omega)$ avente in ingresso la funzione $x_c(t)$ ed uscita $y_c(t)$. Vogliamo trovare un sistema discreto avente funzione di trasferimento $H(e^{j\omega})$ tale che, se la sequenza in ingresso $x(n)$ è formata dai campioni $x_c(nT)$ della funzione $x_c(t)$, allora la sequenza in uscita $y(n)$ sia formata dai campioni $y_c(nT)$ di $y_c(t)$. Ovvero

$$\text{Se } x(n) = x_c(nT) \quad \text{allora} \quad y(n) = y_c(nT) \quad (4.30)$$

Se un tale sistema esiste, lo chiameremo un simulatore discreto di $H_c(\omega)$. Come vedremo, l'equazione 4.30 non può, in generale, essere vera per ogni $x_c(t)$. La simulazione è possibile solo se si restringe la classe delle funzioni di ingresso $x_c(t)$.

Ingresso esponenziale. Supponiamo che $x_c(t)$ sia un'esponenziale, come visto (eq. 1.31) anche $y_c(t)$ è esponenziale (Fig. 4.2)

$$x_c(t) = e^{j\omega_0 t} \quad y_c(t) = H_c(\omega_0) e^{j\omega_0 t} \quad (4.31)$$

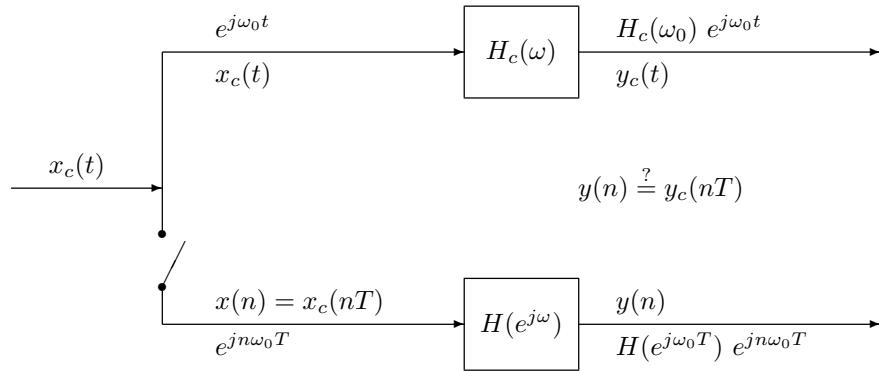


Figura 4.2: Simulazione discreta di un sistema continuo.

I valori campionati $x_c(nT)$ di $x_c(t)$ formano una sequenza esponenziale

$$x(n) = x_c(nT) = e^{jn\omega_0 T}$$

quindi la risposta del sistema discreto $H(e^{j\omega})$ è data da (vedi 4.28)

$$y(n) = H(e^{j\omega_0 T}) e^{jn\omega_0 T} \quad (4.32)$$

Poichè $y_c(nT) = H_c(\omega_0) e^{jn\omega_0 T}$, per soddisfare la 4.30 bisogna scegliere $H(e^{j\omega})$ tale che

$$H(e^{j\omega_0 T}) = H_c(\omega_0) \quad (4.33)$$

Ingresso arbitrario. Generalizziamo il risultato ottenuto. Sia ora $x_c(t)$ un segnale arbitrario con trasformata di Fourier $X_c(\omega)$. Ovviamente vale che:

$$\begin{aligned} x_c(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} X_c(\omega) e^{j\omega t} d\omega \\ Y_c(\omega) &= X_c(\omega) H_c(\omega) \quad y_c(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_c(\omega) H_c(\omega) e^{j\omega t} d\omega \end{aligned} \quad (4.34)$$

Calcolando $x_c(t)$ in $t = nT$ possiamo esprimere $x(n)$ come

$$x(n) = x_c(nT) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_c(\omega) e^{jn\omega T} d\omega$$

e calcolando $y_c(t)$ in $t = nT$ (dalla 4.34) possiamo esprimere $y_c(n)$ come

$$y_c(nT) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_c(\omega) H_c(\omega) e^{jn\omega T} d\omega \quad (4.35)$$

La risposta del ramo basso (quello discreto) si ricava da $Y(\omega) = X_c(\omega)H(e^{j\omega T})$, ovvero

$$y(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_c(\omega) H(e^{j\omega T}) e^{jn\omega T} d\omega \quad (4.36)$$

Quindi, affinché la 4.30 sia vera, gli integrali in 4.36 e 4.35 devono essere uguali per ogni n . Questo accade se

$$H(e^{j\omega T}) = H_c(\omega) \quad (4.37)$$

per ogni ω per cui $X_c(\omega) \neq 0$. Se l'equazione 4.30 deve valere per ogni $x_c(t)$, allora 4.37 deve valere per ogni ω . Questo non è tuttavia sempre possibile, essendo $H(e^{j\omega T})$ una funzione periodica di ω mentre, in generale, $H_c(\omega)$ non lo è. Per soddisfare la condizione 4.30 bisogna limitare la classe delle possibili funzioni in ingresso.

4.4.1 Teorema del simulatore

Se $x_c(t)$ è limitata in banda, cioè se

$$X_c(\omega) = 0 \quad \text{per} \quad |\omega| > \sigma \quad \sigma = \frac{\pi}{T} \quad (4.38)$$

e $H(e^{j\omega})$ è tale che

$$H(e^{j\omega T}) = H_c(\omega) \quad \text{per} \quad |\omega| < \sigma \quad (4.39)$$

allora la condizione 4.30 è soddisfatta.

Dimostrazione: Dalle equazioni 4.38 e 4.39 segue che

$$X_c(\omega)H(e^{j\omega T}) = X_c(\omega)H_c(\omega) \quad \text{per ogni } \omega$$

Quindi i due integrali in 4.36 e 4.35 sono uguali per ogni ω .

Risposta all'impulso del simulatore discreto Chiamando $h(n)$ la risposta alla sequenza impulso del sistema discreto si ha (vedi eq. 4.27) che

$$H(e^{j\omega T}) = \sum_{n=-\infty}^{\infty} h(n) e^{-jn\omega T} \quad (4.40)$$

Quindi i valori di $h(n)$ sono i coefficienti della espansione in serie di Fourier della funzione periodica $H(e^{j\omega T})$ (di periodo 2σ in ω , o 2π in ωT). I valori di $h(n)$ possono quindi essere valutati (vedi eq. 3.19) mediante

$$h(n) = \frac{1}{2\sigma} \int_{-\sigma}^{\sigma} H(e^{j\omega T}) e^{jn\omega T} d\omega = \frac{1}{2\sigma} \int_{-\sigma}^{\sigma} H_c(\omega) e^{jn\omega T} d\omega \quad (4.41)$$

Consideriamo la funzione di trasferimento $H_\sigma(\omega)$ ottenuta da $H_c(\omega)$ troncandola per $|\omega| > \sigma$, ovvero

$$H_\sigma(\omega) = \begin{cases} H_c(\omega) & |\omega| < \sigma \\ 0 & |\omega| > \sigma \end{cases}$$

la corrispondente risposta all'impulso, $h_\sigma(t)$, vale

$$h_\sigma(t) = \frac{1}{2\pi} \int_{-\sigma}^{\sigma} H_c(\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_\sigma(\omega) e^{j\omega t} d\omega \quad (4.42)$$

e, per confronto con la 4.41,

$$h(n) = T h_\sigma(nT) \quad (4.43)$$

Dalla discussione precedente segue che per trovare il simulatore discreto del sistema continuo $H_c(\omega)$ possiamo valutare $h(n)$ campionando la risposta all'impulso $h_\sigma(t)$ del sistema troncato $H_\sigma(\omega)$. In alternativa possiamo specificare la risposta in frequenza del sistema discreto $H(e^{j\omega})$ seguendo la 4.39.

Esempio 4.7 Vogliamo simulare un differenziatore, ossia un sistema continuo con risposta in frequenza

$$H_c(\omega) = j\omega$$

Poichè

$$h_\sigma(t) = \frac{1}{2\pi} \int_{-\sigma}^{\sigma} j\omega e^{j\omega t} d\omega = \sigma t \frac{\cos \sigma t - \sin \sigma t}{\pi t^2} \quad (4.44)$$

la 4.43 ci permette di scrivere, essendo $\sigma = \pi/T$,

$$h(n) = T h_\sigma(nT) = \frac{(-1)^n}{nT} \quad \text{per } n \neq 0 \quad (4.45)$$

e $h(0) = 0$. Quindi, se $X_c(\omega) = 0$ per $|\omega| > \sigma$, allora $y(nT) = x'_c(nT)$. Questo porta all'identità

$$x'_c(nT) = \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \frac{(-1)^k}{kT} x_c(nT - kT) \quad (4.46)$$

che esprime la derivata di una funzione limitata in banda in termini dei suoi valori campionati.

Esempio 4.8 Vogliamo simulare un filtro passa-basso ideale $H_c(\omega)$ con frequenza di taglio pari a $\omega_c < \sigma$, cioè

$$H_c(\omega) = \begin{cases} 1 & |\omega| < \omega_c \\ 0 & |\omega| > \omega_c \end{cases}$$

In questo caso $H_\sigma(\omega) = H_c(\omega)$, dato che $H_c(\omega) = 0$ per $|\omega| > \sigma > \omega_c$. Quindi:

$$h_\sigma(t) = h_c(t) = \frac{\sin \omega_c t}{\pi t} \quad h(n) = Th_\sigma(nT) = \frac{\sin \omega_c T n}{\pi n} \quad (4.47)$$

Se $\omega_c = \sigma$ il filtro diventa un passa-tutto, ovvero non modifica un ingresso limitato in banda a σ . Infatti, in questo caso,

$$h(n) = \frac{\sin \sigma T n}{\pi n} = \frac{\sin \pi n}{\pi n} = \delta(n)$$

4.5 Proprietà delle sommatorie

Analizziamo in questa sezione alcune proprietà delle sommatorie ad infiniti termini.

4.5.1 Manipolazione degli indici delle sommatorie infinite

Sia $s(n)$ una generica sequenza illimitata. Supponiamo di voler valutare la somma di tutti i suoi termini:

$$\sum_{n=-\infty}^{\infty} s(n) \quad (4.48)$$

Siccome l'operazione di somma è commutativa, posso sommare i termini nell'ordine che preferisco. Ad esempio posso eseguire la somma dei termini di indice pari, poi quella con indice dispari e infine sommare i due risultati. Indico questo modo di procedere mediante

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{n \text{ pari}} s(n) + \sum_{n \text{ dispari}} s(n) \quad (4.49)$$

La notazione sopra è precisa ma non operativa: posso esprimere lo stesso concetto scrivendo

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{r=-\infty}^{\infty} s(2r) + \sum_{k=-\infty}^{\infty} s(2k+1) \quad (4.50)$$

È evidente che mentre l'indice r assume tutti i possibili valori interi, $2r$ assumerà tutti i possibili valori pari (compreso lo zero). Allo stesso modo, mentre l'indice k assume tutti i possibili valori

interi, $(2k + 1)$ assumerà tutti i possibili valori dispari. Posso anche unire le due sommatorie e scrivere che

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{l=-\infty}^{\infty} [s(2l) + s(2l+1)] \quad (4.51)$$

o anche, se voglio,

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{l=-\infty}^{\infty} \sum_{h=0}^1 s(2l+h) \quad (4.52)$$

La doppia sommatoria sopra esprime il fatto che posso sommare gli infiniti valori di $s(n)$ sommando due termini consecutivi per volta e poi sommare tutti i risultati.

Nulla vieta di estendere questo risultato: ad esempio posso dividere $\sum_{n=-\infty}^{\infty} s(n)$ in quattro sommatorie (anzichè nelle due della eq. 4.49) considerando una sommatoria sugli indici esattamente divisibili per 4 ($0, 4, 8, 12, \dots$), la seconda sugli indici divisibili per quattro con resto uno ($1, 5, 9, \dots$), la terza sugli indici divisibili per quattro con resto due ($2, 6, 10, \dots$) e l'ultima sugli indici divisibili per quattro con resto tre ($3, 7, 11, \dots$). In tal caso la 4.50 diventa

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{r=-\infty}^{\infty} s(4r) + \sum_{r=-\infty}^{\infty} s(4r+1) + \sum_{r=-\infty}^{\infty} s(4r+2) + \sum_{r=-\infty}^{\infty} s(4r+3) \quad (4.53)$$

e la 4.52 diventa

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{m=-\infty}^{\infty} \sum_{n=0}^3 s(4m+n) \quad (4.54)$$

(gli indici delle sommatorie li posso chiamare col nome che preferisco....).

A questo punto posso generalizzare il risultato: per qualsiasi numero $N \geq 1$ vale che

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{m=-\infty}^{\infty} \sum_{n=0}^{N-1} s(mN+n) \quad (4.55)$$

e poichè m assume tutti i valori interi da $-\infty$ a ∞ , vale anche che

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{m=-\infty}^{\infty} \sum_{n=0}^{N-1} s(n-mN) \quad (4.56)$$

4.5.2 Rappresentazione delle sequenze periodiche mediante sommatorie

Sia $\tilde{x}(n)$ una sequenza periodica di periodo pari a N . Ciò significa che vale la relazione

$$\tilde{x}(n+kN) = \tilde{x}(n) \quad \forall k \text{ intero} \quad (4.57)$$

Sia ora $x(n)$ una sequenza (aperiodica) tale che

$$x(n) = \begin{cases} \tilde{x}(n) & \text{per } 0 \leq n \leq N-1 \\ 0 & \text{altrove} \end{cases} \quad (4.58)$$

ovvero uguale alla sequenza periodica $\tilde{x}(n)$ in un periodo e zero altrove.

Posso allora esprimere la sequenza periodica come combinazione di infinite repliche di $x(n)$ traslate, secondo lo schema seguente:

$$\tilde{x}(n) = \begin{cases} \dots & \dots \\ x(n+N) & \text{per } -N \leq n \leq -1 \\ x(n) & \text{per } 0 \leq n \leq N-1 \\ x(n-N) & \text{per } N \leq n \leq N+N-1 \\ x(n-2N) & \text{per } 2N \leq n \leq 2N+N-1 \\ x(n-3N) & \text{per } 3N \leq n \leq 3N+N-1 \\ \dots & \dots \end{cases} \quad (4.59)$$

Siccome $x(n)$ è nulla per $n \notin [0, N-1]$, posso ottenere $\tilde{x}(n)$ come somma di tutte le sequenze traslate, ovvero come

$$\tilde{x}(n) = \sum_{k=-\infty}^{\infty} x(n-kN) \quad (4.60)$$

Poichè k assume tutti i valori interi da $-\infty$ a ∞ , posso scrivere la somma precedente nel modo equivalente

$$\tilde{x}(n) = \sum_{k=-\infty}^{\infty} x(n+kN) \quad (4.61)$$

cosa che si può anche ottenere cambiando l'indice della sommatoria da k a $l = -k$ oppure osservando che, per *qualsiasi* sequenza $s(n)$, deve valere che

$$\sum_{n=-\infty}^{\infty} s(n) = \sum_{n=-\infty}^{\infty} s(-n)$$

visto che posso ottenere il risultato di una somma sommando i fattori in ordine qualsiasi (la somma è commutativa).

4.6 Trasformata Discreta di Fourier (DFT)

Finora abbiamo visto la trasformata di Fourier per funzioni analogiche, la serie di Fourier, anch'essa per funzioni analogiche, e la trasformata di Fourier per segnali a tempo discreto (sequenze). Nel caso delle sequenze, la FT è una funzione a tempo (omega) continuo, non discreto. Serve qualcosa che permetta di rappresentare i segnali a tempo discreto nel dominio della frequenza ottenendo un segnale a tempo (frequenza) discreto, in modo da poter eseguire l'analisi di Fourier per mezzo di un calcolatore numerico. La risposta a questa esigenza è la trasformata discreta di Fourier, o DTF, che trasforma una sequenza finita di numeri in una diversa sequenza finita di numeri.

Per arrivare alla DFT, partiamo dalla sequenza $x(n)$ di lunghezza limitata, ovvero definita per soli N valori, $n = 0, \dots, N-1$. Si può anche pensare come una sequenza illimitata ma tale che $x(n) = 0$ per $n \notin [0, N-1]$ (le due interpretazioni sono equivalenti). Per cercare di ottenere una rappresentazione discreta, usiamo un sistema già visto: consideriamo la sequenza periodica $\tilde{x}(n)$ che si ottiene replicando infinite volte $x(n)$

$$\tilde{x}(n) = \sum_{m=-\infty}^{\infty} x(n-mN) \quad (4.62)$$

Consideriamo la FT di $\tilde{x}(n)$, definita dalla 4.3, nella sua versione in f (che si ottiene da $\omega = 2\pi f$):

$$\tilde{X}(f) = \sum_{n=-\infty}^{\infty} \tilde{x}(n) e^{-j2\pi fn} \quad (4.63)$$

che può essere scritta nella forma equivalente (vedi eq. 4.56)

$$\tilde{X}(f) = \sum_{m=-\infty}^{\infty} \sum_{n=0}^{N-1} \tilde{x}(n-mN) e^{-j2\pi f(n-mN)} \quad (4.64)$$

Sfruttando la periodicità di $\tilde{x}(n)$, ovvero il fatto che $\tilde{x}(n-mN) = x(n)$, si ha

$$\tilde{X}(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi fn} \sum_{m=-\infty}^{\infty} e^{j2\pi fmN} \quad (4.65)$$

Ma, dalla 4.9, considerando $T = 1/N$, sappiamo che

$$\sum_{m=-\infty}^{\infty} e^{j2\pi fmN} = \frac{1}{N} \sum_{h=-\infty}^{\infty} \delta(f - \frac{h}{N})$$

e quindi

$$\tilde{X}(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi fn} \frac{1}{N} \sum_{h=-\infty}^{\infty} \delta(f - \frac{h}{N}) \quad (4.66)$$

Cambiando l'ordine delle sommatorie si ottiene

$$\tilde{X}(f) = \sum_{h=-\infty}^{\infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi hn/N} \delta(f - \frac{h}{N}) \quad (4.67)$$

Chiamando

$$S_h = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi hn/N} \quad (4.68)$$

si ha

$$\tilde{X}(f) = \sum_{h=-\infty}^{\infty} S_h \delta(f - \frac{h}{N}) \quad (4.69)$$

Ciascun valore S_h è un numero che dipende solo da h e dalla sequenza aperiodica $x(n)$. La FT $\tilde{X}(f)$ è allora un treno di impulsi equispaziati in frequenza con passo $1/N$, ciascuno di area S_h . Ma le aree S_h sono periodiche di periodo N , infatti

$$S_{h+kN} = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi(h+kN)n/N} = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi hn/N} e^{-j2\pi kn} = S_h \quad (4.70)$$

visto che $e^{-j2\pi kn} = 1$ per ogni k, n interi.

Allora la FT della sequenza periodica $\tilde{x}(n)$ è completamente specificata da N numeri, ovvero i valori di S_h in un periodo. Inoltre ciascun S_h può essere calcolato mediante una sommatoria di N termini (vedi 4.68). In totale si devono quindi calcolare N valori i S_h , ciascuno mediante una sommatoria finita di N addendi: questo può essere eseguito per mezzo di un calcolatore numerico senza alcuna approssimazione, a differenza di quanto avviene con le FT che richiedono sommatorie con infiniti termini o integrali.

Si definisce allora la *Trasformata Discreta di Fourier* diretta, o **DFT**, della sequenza $x(n)$, $n = 0, \dots, N-1$, come

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N} nk} \quad 0 \leq k \leq N-1 \quad (4.71)$$

E' evidente il legame con gli S_h definiti in precedenza. Da notare che la DFT trasforma una sequenza di N numeri (reali o complessi), in una nuova sequenza di N numeri, in generale complessi. Quindi la conoscenza degli N numeri $X(k)$, $k = 0, \dots, N-1$ è equivalente alla conoscenza della sequenza $x(n)$. La trasformazione inversa, o **IDFT**, si ottiene da

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi}{N} nk} \quad 0 \leq n \leq N-1 \quad (4.72)$$

Il fattore di normalizzazione $1/N$ compare normalmente nella IDFT, alcuni testi la inseriscono invece nella diretta (non cambia molto...).

L'uso della DFT è di fondamentale importanza nello studio e nella elaborazione dei segnali in forma numerica. La sua notevole diffusione è anche dovuta all'esistenza di algoritmi di calcolo veloce, detti Fast Fourier Transform, o FFT, che riducono di diversi ordini di grandezza il numero di operazioni necessarie al calcolo della DFT se N è grande: si passa da una complessità computazionale dell'ordine di N^2 per il calcolo diretto della 4.71 a $(N/2) \log_2 N$ per il calcolo veloce (se N è una potenza del due). Gli algoritmi di FFT li vedrete in un corso successivo e non verranno esposti in questa sede.

4.6.1 Proprietà della DFT

Per la DFT valgono le solite proprietà della trasformata di Fourier, con le modifiche del caso.

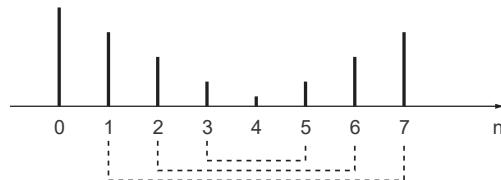
Linearità La trasformata discreta di Fourier è un operatore lineare, come evidente dalla definizione. Date due generiche sequenze $x(n)$ e $y(n)$ della stessa lunghezza N , vale quindi che

$$DFT[ax(n) + by(n)] = aX(k) + bY(k)$$

essendo ovviamente $X(k)$ la DFT di $x(n)$ e $Y(k)$ la DFT di $y(n)$.

Simmetria Se $x(n)$ è reale allora $X(k)$ ha simmetria hermitiana, ovvero parte reale (e modulo) pari e parte immaginaria (e fase) dispari. Se $x(n)$ è pari allora $X(k)$ è puramente reale, se $x(n)$ è dispari $X(k)$ è puramente immaginaria. Val anche la combinazione delle proprietà, per cui se $x(n)$ è reale e pari allora $X(k)$ è pari e reale e se $x(n)$ è reale e dispari $x(k)$ è dispari e puramente immaginaria.

Il problema che nasce è il seguente: se $x(n)$ è definita fra 0 e $N - 1$ cosa significa, ad esempio, dire che è pari? Ovviamente una sequenza è pari se $x(n) = x(-n) \forall n$, ma $x(n)$ NON è definita per n negativi! Il fatto è che bisogna tener conto di come la DFT è stata ricavata: vale per sequenze limitate (aperiodiche), ma è stata derivata periodizzando la sequenza di ingresso. Quando si dice $x(n)$ pari si intende quindi $\tilde{x}(n)$ pari, ovvero che sia pari la sequenza che si ottiene periodizzando $x(n)$. Quindi, ad esempio, $\tilde{x}(1) = \tilde{x}(-1)$ è vera se $x(1) = x(N - 1)$, $\tilde{x}(2) = \tilde{x}(-2)$ è vera se $x(2) = x(N - 2)$ e così via. Se $x(n)$ è lunga $N = 8$ campioni, è "pari" se gli elementi uniti dalle linee tratteggiate in figura sono uguali.



Traslazione Vi sono due modi per intendere la traslazione. Il primo è definire una sequenza $t(n) = x(n - A)$: se $x(n)$ ha valori in $n \in [0, N - 1]$ allora $t(n)$ ha valori in $n \in [A, A + N - 1]$. La DFT di $t(n)$ è

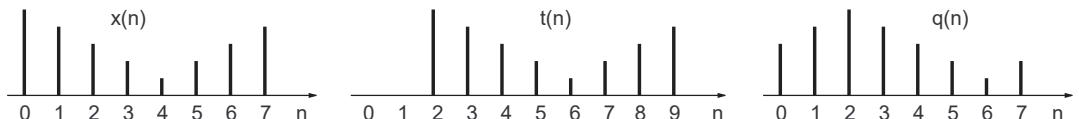
$$T(k) = \sum_{l=A}^{A+N-1} t(l) e^{-j\frac{2\pi}{N} lk}$$

e cambiando indice della sommatoria da l a $n = l - A$, ovvero sostituendo $l = n + A$,

$$T(k) = \sum_{n=0}^{N-1} t(n + A) e^{-j\frac{2\pi}{N}(n+A)k} = e^{-j\frac{2\pi}{N}Ak} \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk} = e^{-j\frac{2\pi}{N}Ak} X(k)$$

che è la solita regola di traslazione delle trasformate di Fourier.

Lo stesso risultato si ottiene definendo una traslazione *circolare* della sequenza $x(n)$: essa equivale a traslare la sequenza periodizzata, ovvero $q(n) = \tilde{x}(n - A)$ per $n \in [0, N - 1]$. In figura sono rappresentate le due traslazioni per $N = 8$ e $A = 2$. Nel caso di traslazione circolare i campioni che *escono* a destra a causa della traslazione *rientrano* a sinistra.



I due tipi di traslazione hanno lo stesso effetto sulla DFT in quanto la periodizzazione di $t(n)$ produce la stessa sequenza che si ottiene periodizzando $q(n)$. Gli algoritmi che eseguono la DFT assumono implicitamente che la sequenza sia definita in $[0, N - 1]$, per cui, se si vuole traslare la sequenza di ingresso, bisogna eseguire una traslazione circolare.

4.6.2 Uso della DFT per il calcolo di convoluzioni discrete*

La convoluzione di due sequenze corrisponde, nel dominio della DFT, al prodotto delle trasformate. Vi sono però differenze importanti rispetto alle analoghe proprietà già viste per le altre FT. La prima, evidente, è che posso moltiplicare fra loro due DFT solo se esse sono della stessa lunghezza, ovvero se hanno lo stesso numero di campioni. Se così non è devo allungare la sequenza più corta con abbastanza zeri da portarla alla lunghezza della più lunga.

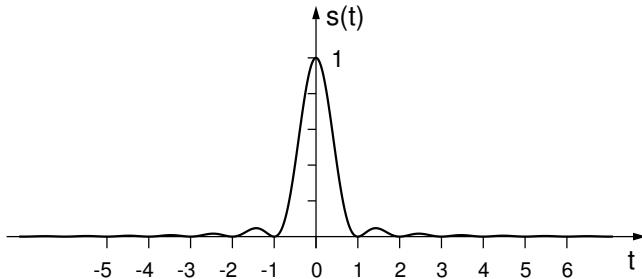
La seconda differenza la illustriamo con un esempio: siano $x(n)$ e $y(n)$ due sequenze entrambe di lunghezza N . Se eseguo la convoluzione $x(n) * y(n)$ secondo la definizione in 4.17 o 4.18 ottengo una sequenza lunga $2N - 1$ campioni, come visto nell'esempio 4.3.

Se calcolo le DFT di entrambe le sequenze, $X(k)$ e $Y(k)$, ottengo due sequenze lunghe N . Il loro prodotto è ancora della stessa lunghezza, così come l'antitrasformata del prodotto. È allora evidente che l'antitrasformata del prodotto delle DFT non può corrispondere alla convoluzione delle sequenze, se non altro perché le lunghezze sono diverse.

Il fatto è che la DFT assume la periodizzazione delle sequenze, per cui la traslazione, come visto nella sezione precedente, viene interpretata in modo circolare. Il modo per ottenere la corretta sequenza di convoluzione è quello di allungare le due sequenze di ingresso, prima della DFT diretta, con abbastanza zeri da portare entrambe le loro lunghezze alla somma delle lunghezze originali meno uno. I dettagli di tale tecnica, di fondamentale utilità nel calcolo numerico, verranno ripresi in un corso seguente.

4.6.3 Analisi dei segnali mediante DFT

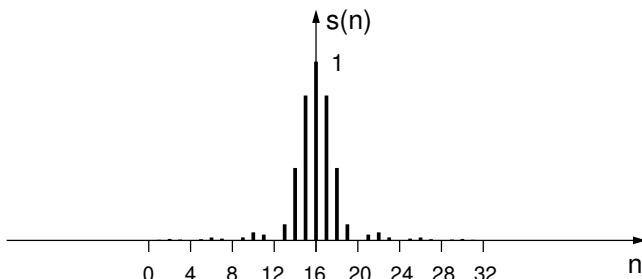
Sia $x_a(t)$ un segnale analogico passabasso a banda limitata, ovvero $X_a(f) = 0$ per $|f| \geq W$. Dalla teoria sappiamo che $x_a(t)$ non può essere limitato nel tempo, ma supponiamo, per il momento, che lo sia, ovvero che sia pari a zero per t al di fuori di un certo intervallo (a, b) . Torneremo in seguito su questa questione, per analizzarla correttamente (equivale a troncare nel tempo). Come esempio consideriamo il segnale $s(t) = \text{sinc}^2(t)$ in figura: è illimitato nel tempo, ma se considero un certo numero di lobi, ad esempio 4, posso supporre che il resto sia trascurabile e considerarla come se fosse limitata.



La funzione $s(t)$ ha massimo pari a 1 in $t = 0$ ed i suoi zeri sono a valori interi di t . La trasformata di Fourier $S(f)$ è un triangolo alto uno con base da $f = -1$ a $f = 1$ Hz, ovvero $S(f) = \Delta(-f)$ ma siccome il triangolo è pari, $S(f) = \Delta(f)$. Quindi $s(t)$ ha banda limitata pari a $W = 1$ Hz.

Un segnale a banda limitata può essere campionato con frequenza di campionamento f_c pari ad almeno il doppio di W . Siccome abbiamo supposto che $x_a(t)$ sia anche limitato nel tempo, la sequenza utile (cioè con valori non nulli) dei suoi campioni, $x(n) = x_a(n/f_c) = x_a(nT_c)$ ha un numero N limitato di campioni.

Nell'esempio, campioniamo $s(t)$ con frequenza $f_c = 4$ Hz, piuttosto abbondante (il limite è 2 Hz), e consideriamo solo i primi 4 lobi. La sequenza che risulta dal campionamento è la seguente



Siccome ci sono 4 campioni a lobo e limito il campionamento a 8 lobi ho che la sequenza campionata $s(n) = s(t = -4 + n/4)$ ha $N = 32$ campioni: in realtà sarebbero 33, se considerassi anche l'ultimo campione dell'ultimo intervallo, ma per convenzione non si usa (e poi è nullo).

Consideriamo ora la trasformata di Fourier per sequenze del segnale campionato $x(n)$ (eq. 4.3):

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} = \sum_{n=0}^{N-1} x(n) e^{-j\omega n} \quad (4.73)$$

Se continuiamo a trascurare il troncamento nel tempo ed abbiamo campionato seguendo la condizione di Nyquist, sappiamo (eq. 4.16) che

$$X(e^{j\omega T_c}) = \frac{1}{T_c} \sum_{l=-\infty}^{\infty} X_a(\omega + \frac{2\pi l}{T_c})$$

ovvero che

$$X(e^{j\omega T_c}) = \frac{1}{T_c} X_a(\omega) \quad \text{per } |\omega| \leq \frac{\pi}{T_c} \quad (4.74)$$

o anche

$$X(e^{j\omega}) = \frac{1}{T_c} X_a\left(\frac{\omega}{T_c}\right) \quad \text{per } |\omega| \leq \pi \quad (4.75)$$

ovvero che la replica centrata in $\omega = 0$ della funzione periodica $X(e^{j\omega T_c})$ è uguale alla trasformata di Fourier $X_a(\omega)$ del segnale analogico $x_a(t)$, a parte per un fattore di scalamento.

Consideriamo ora la DFT della sequenza campionata $x(n)$ (eq. 4.71)

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk} \quad 0 \leq k \leq N-1$$

e confrontiamola con la 4.73: evidentemente

$$X(k) = X(e^{j\frac{2\pi}{N}k}) \quad 0 \leq k \leq N-1$$

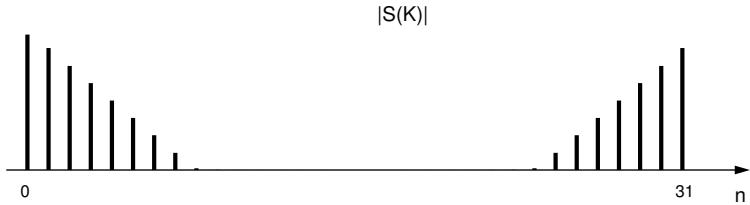
Quindi la DFT è il campionamento di $X(e^{j\omega})$ mediante N campioni equispaziati nell'intervallo $\omega \in [0, 2\pi]$ (0 compreso, 2π escluso) ed il legame fra i valori della DFT $X(k)$ e la trasformata di Fourier del segnale analogico $x_a(t)$ è, considerando la 4.75:

$$X(k) = X(e^{j\frac{2\pi}{N}k}) = \frac{1}{T_c} X_a\left(\frac{2\pi}{NT_c}k\right) \quad (4.76)$$

La relazione precedente vale solo per $k \leq N/2$, ovvero per la prima metà della DFT, che corrisponde alle frequenze positive di $X_a(\omega)$, visto che $X(k)$ è il campionamento di $X(e^{j\omega})$ per $\omega \in [0, 2\pi)$, mentre la 4.75 vale per $\omega \in (-\pi, \pi)$.

Poichè $X(e^{j\omega})$ è periodica, la seconda metà della DFT corrisponde alla prima metà della prima replica di $X(e^{j\omega})$, e quindi alle frequenze negative di $X_a(\omega)$.

Torniamo all'esempio: la DFT di $s(n)$ è

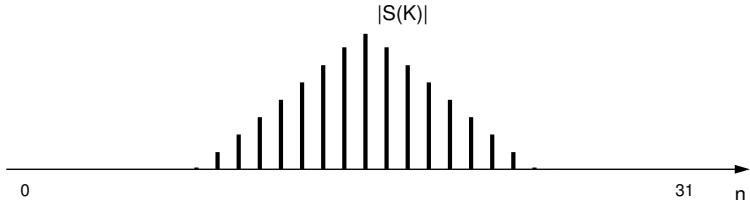


Come evidente prima compaiono i campioni della parte positiva del triangolo, poi quelli della parte negativa. Poichè si è campionato a 4 Hz, ovvero con tempo di campionamento $T_c = 1/4$ sec., i campioni sono equispaziati in ω di

$$\frac{2\pi}{NT_c} = \frac{2\pi}{32/4} = \frac{2\pi}{8}$$

La fine del triangolo corrisponde al campione numero 8 (contando da zero), e quindi a $\omega = 2\pi$, cioè a $f = 1$ Hz, che è la corretta banda del segnale $s(t)$.

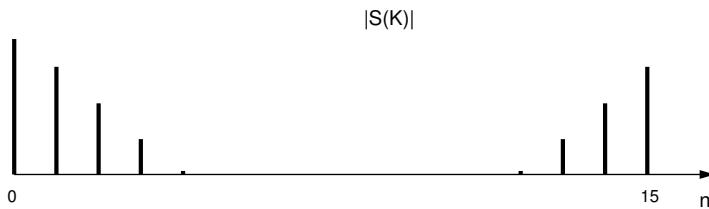
Se voglio posso visualizzare la DFT portando il campione 0 a metà sequenza mediante traslazione circolare, ottendo



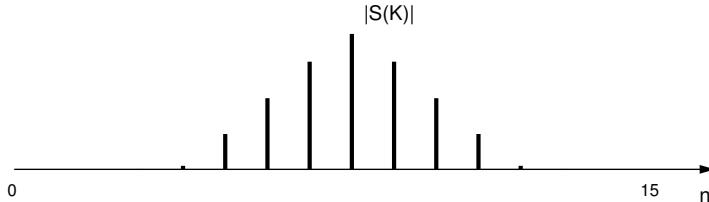
che appare una migliore rappresentazione della trasformata di Fourier (campionata) di $s(t) = \text{sinc}^2(t)$.

Il risultato ottenuto dalla DFT di $s(t)$ sembra confermare che la parte di $\text{sinc}^2(t)$ al di fuori dei primi quattro lobi (per parte) sia effettivamente trascurabile, visto che le differenze rispetto al risultato teorico sono (almeno visivamente) indistinguibili.

Proviamo a visualizzare il risultato che si ottiene troncando maggiormente la $\text{sinc}^2(t)$: consideriamo solo due lobi per parte. Campionando sempre alla frequenza di 4 Hz (il doppio del minimo necessario) ottengo 16 campioni (4 per lobo su quattro lobi, due per parte). La relativa DFT (modulo) è

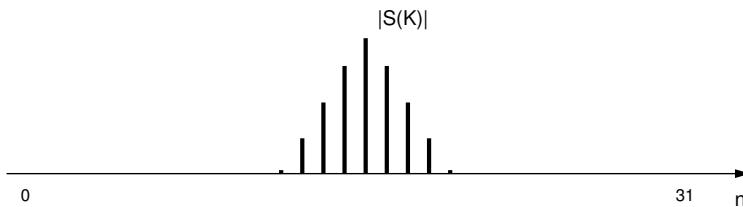


e la versione "centrata" (d'ora in poi visualizzeremo solo le versioni centrate, che sono più chiare)

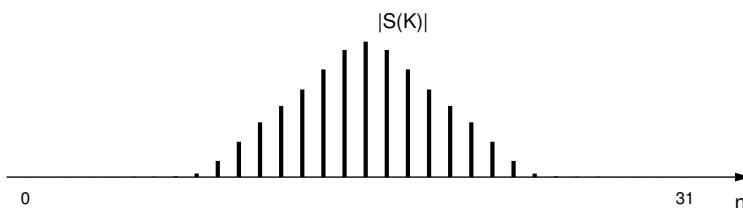


Anche in questo caso la DFT appare come un triangolo, ma i campioni sono pochi. C'è un modo per aumentare la risoluzione nel dominio della DFT, in modo da osservare con maggior dettaglio cosa accade?

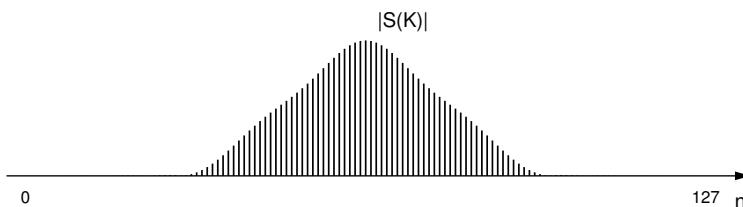
Proviamo ad aumentare la risoluzione aumentando la frequenza di campionamento: con $f_c = 8$ Hz ho 32 campioni (8 per lobo) e la DFT diventa



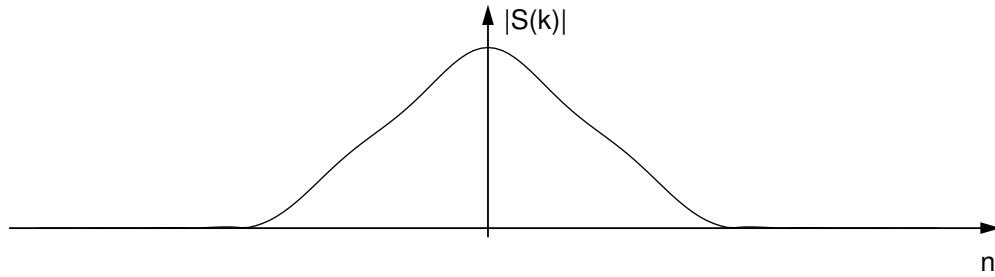
Ho il doppio dei campioni, ma quelli significativi sono sempre nello stesso numero. Questo è ovvio, con $f_c = 8$ Hz "esploro" il doppio dell'asse delle frequenze col doppio dei campioni, quindi la spaziatura in frequenza dei campioni non varia. Se voglio rendere i campioni più fitti in frequenza la strada è un'altra: torno al campionamento a 4 Hz, ma aggiungo in fondo alla sequenza campionata (di 16 campioni) altri 16 campioni tutti nulli. La DFT diventa



Come evidente ora la risoluzione in frequenza migliora! Aumentiamo ancora la risoluzione: 16 campioni originali più 112 zeri (totale 128 campioni):



Quando il numero di campioni aumenta conviene cambiare modo di visualizzazione: invece di disegnare i campioni come segmenti verticali disegno la linea che unisce la sommità degli stessi, come se fosse una funzione continua. La DFT precedente, visualizzata in questo modo, diventa



Ora è evidente che il troncamento nel tempo ha reso la trasformata diversa da un triangolo. Considerare solo i primi due lobi della funzione $s(t) = \text{sinc}^2(t)$ equivale a moltiplicare $s(t)$ per un rettangolo alto uno fra $t = -2$ e $t = 2$ e zero altrove, ovvero considerare la funzione troncata

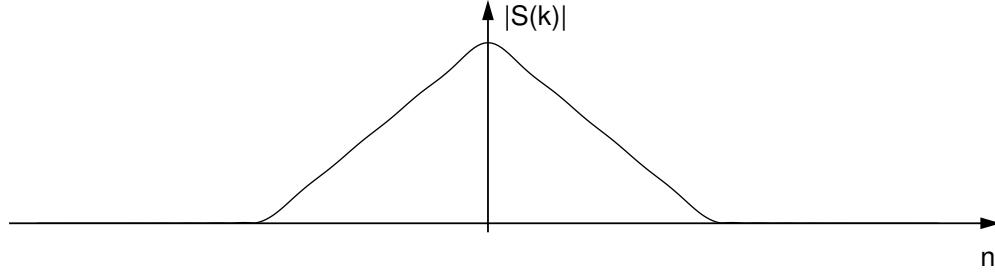
$$s_t(t) = \text{sinc}^2(t) \cdot \prod \left(\frac{t}{4} \right)$$

questa ha trasformata di Fourier che si ottiene per convoluzione fra la trasformata della sinc e la trasformata del rettangolo, ovvero

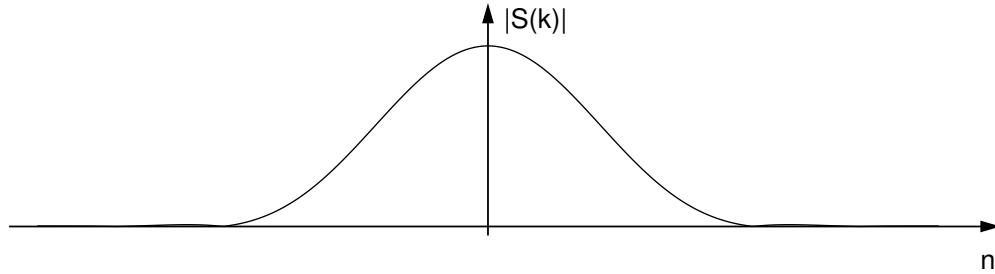
$$S_t(f) = \Lambda(f) * 4 \text{sinc}(4f)$$

cioè la convoluzione fra un triangolo di base $[-1, 1]$ ed una sinc alta 4 in $f = 0$ con zeri in posizione $f_z = h/4$ Hz, h intero diverso da zero. La convoluzione con la sinc deforma il triangolo, che assume l'andamento "oscillante" in figura.

Se tronco al quarto lobo, come nel caso visto in precedenza, tronco con un rettangolo alto uno fra -4 e 4 . La sua trasformata è ancora una sinc, ma più stretta (gli zeri distano $1/8$ di Hz) e quindi l'effetto è ridotto. Infatti la DFT, opportunamente infittita, è



Per finire ecco cosa succede se tronco al primo lobo, cioè fra -1 e 1



Quanto visto illustra l'uso della DFT per l'analisi in frequenza delle funzioni utilizzando un calcolatore numerico (e passando dal teorema del campionamento). Eseguendo i passaggi con la cura necessaria la DFT fornisce lo strumento adatto per la visualizzazione della trasformata di Fourier di segnali analogici: tutte le figure di trasformate di Fourier dirette o inverse che sono state presentate in queste dispense sono ottenute proprio mediante il suo uso.

Capitolo 5

Teoria della probabilità

La teoria della probabilità riguarda fenomeni multipli (sequenziali o simultanei) e lo studio delle loro medie. Esempi di tali fenomeni sono: l'emissione elettronica, il gioco del lotto, il transito di automobili su di una strada, il rumore (fruscio) di un amplificatore audio, gli errori su di un canale di trasmissione, il lancio di un dado e molti altri.

E' stato osservato che in questi ed altri fenomeni certe medie tendono ad un valore costante al crescere del numero di osservazioni e che il loro valore resta lo stesso quando vengono valutate usando una sottosequenza specificata prima che l'esperimento venga compiuto. Ad esempio, nel lancio della moneta, la percentuale di "testa" approssima 0,5 (o una qualche altra costante) al crescere del numero di tentativi e la stessa media si ottiene considerando solo un lancio ogni tre.

Lo scopo della teoria della probabilità è quello di descrivere e prevedere tali medie in termini di probabilità degli eventi.

La probabilità di un evento \mathcal{A} è un numero $P(\mathcal{A})$ assegnato all'evento. Tale numero può essere interpretato come segue:

se un esperimento viene ripetuto n volte e l'evento \mathcal{A} si verifica $n_{\mathcal{A}}$ volte allora,
con un alto grado di certezza, la frequenza relativa $n_{\mathcal{A}}/n$ del verificarsi di \mathcal{A} è vicino a
 $P(\mathcal{A})$, ovvero

$$P(\mathcal{A}) \approx \frac{n_{\mathcal{A}}}{n}$$

purchè n sia sufficientemente grande.

Questa interpretazione è ovviamente imprecisa, non essendo ben definiti i termini sottolineati. L'imprecisione non può essere evitata in quanto la teoria della probabilità è legata ai fenomeni fisici solo in termini inesatti, proprio come ogni altra teoria fisica. Tuttavia la teoria è esatta, è derivata logicamente da assiomi definiti chiaramente e, quando applicata a problemi reali, funziona!

Come confronto consideriamo un resistore: esso viene schematizzato come un componente a due terminali. La tensione ai suoi estremi dipende dalla corrente che lo attraversa secondo: $R = v(t)/i(t)$. Questa è solo un'astrazione: un resistore reale è un dispositivo complesso con induttanza e capacità distribuite e senza terminali chiaramente specificati. La relazione sopra è quindi valida con certi errori, in certi intervalli di frequenza, per certi limiti di corrente e di temperatura ecc. Tuttavia per sviluppare la teoria dei circuiti elettrici si ignorano queste imprecisioni: si definisce un resistore ideale (che non esiste) e si va avanti.

Per applicare la teoria della probabilità a problemi reali è necessario distinguere i seguenti passi.

PASSO 1: OSSERVAZIONE (fisico). Si determinano, in modo non esatto, le probabilità $P(\mathcal{A}_i)$ di certi eventi \mathcal{A}_i . Per trovare le $P(\mathcal{A}_i)$ si può usare la frequenza relativa in un "grande" numero di prove od il ragionamento utilizzando simmetrie fisiche (se un dado è non truccato l'uscita di ciascuna delle sue sei facce ha probabilità 1/6) o ancora il numero di eventi favorevoli diviso il numero di eventi possibili (se in una scatola ho 10 palline rosse e 20 palline bianche, la probabilità di pescare a caso una pallina bianca è 20/30).

PASSO 2: DEDUZIONE (concettuale). Assumiamo che le probabilità soddisfino certi assiomi e, per mezzo del ragionamento deduttivo, determiniamo, a partire dalle probabilità $P(\mathcal{A}_i)$ di certi eventi \mathcal{A}_i le probabilità $P(\mathcal{B}_i)$ di altri eventi \mathcal{B}_i . Ad esempio, nel lancio del dado, a partire da

$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$ deduciamo che la probabilità che esca un numero dispari è $1/2$.

PASSO 3: PREDIZIONE (fisico). Si realizza una predizione fisica basata sui numeri $P(\mathcal{B}_i)$ ottenuti al passo 2. Se, ad esempio, tiriamo il dado 1000 volte, la nostra predizione è che l'evento "esce un numero dispari" si verifichi circa 500 volte.

La maggior parte della teoria della probabilità (e di quanto vedremo) riguarda il passo 2: da certe probabilità ne deriviamo altre in modo deduttivo. La probabilità di un evento verrà interpretata come un numero assegnato all'evento, come la massa è assegnata ad un corpo od una resistenza ad un resistore. Nello sviluppo della teoria non ci occuperemo del "significato fisico" di tale numero, esattamente come viene fatto in analisi dei circuiti, in meccanica, in elettromagnetismo ed in ogni altra disciplina scientifica. Ovviamente tutte queste teorie sarebbero senza valore se non servissero a risolvere problemi reali. Dobbiamo assegnare una specifica, ancorché approssimata, resistenza ad un resistore ed una probabilità ad un evento (PASSO 1). Dobbiamo anche trovare un significato fisico alle conclusioni ricavate della teoria (PASSO 3), ma il legame fra concetti ed osservazioni deve essere separato dalla struttura logica della teoria (PASSO 2).

Un po' di storia. La nascita della teoria della probabilità viene generalmente fissata nel 1654 in Francia, quando Antoine Gombaud, chevalier de Méré, scrittore e giocatore d'azzardo, chiese al matematico Blaise Pascal di aiutarlo relativamente ad alcune questioni relative al gioco dei dadi (in Francia a quel tempo il gioco d'azzardo era legale). Pascal si mise al lavoro in collaborazione epistolare con Pierre Fermat, e la teoria della probabilità ebbe inizio. Il problema, interessante anche perché ancor oggi non molta gente è in grado di risolverlo, era il seguente: il cavalier De Mere osservò che lanciando tre dadi le combinazioni con somma 11 erano un po' più frequenti delle combinazioni con somma 12 e ciò gli sembrava strano poichè sia la somma 11 che la somma 12 si possono ottenere con tre dadi esattamente in sei modi diversi, cioè:

$$11 = \{(6, 4, 1), (6, 3, 2), (5, 5, 1), (5, 4, 2), (5, 3, 3), (4, 4, 3)\}$$

$$12 = \{(6, 5, 1), (6, 4, 2), (6, 3, 3), (5, 5, 2), (5, 4, 3), (4, 4, 4)\}$$

quindi egli considerava i due casi equiprobabili e perciò si aspettava che le due combinazioni dovessero avere la stessa frequenza. Pascal risolse il problema dimostrando che le combinazioni a somma 11 sono 27, mentre quelle a somma 12 sono solo 25, poichè conta anche l'ordine dei numeri sui dadi. Il metodo sviluppato (dopo aver lanciato dadi innumerevoli volte) si chiama metodo classico. Come vedremo fra poco si basa sulla seguente considerazione: supponiamo che un gioco abbia n possibili uscite, delle quali m sono quelle che ci interessano. La probabilità è allora m/n . Questo approccio richiede che un esperimento (gioco) sia divisibile in uscite equiprobabili: questo non sempre è possibile e a volte non è facile capire se le uscite siano o meno ugualmente probabili.

Un diverso metodo, noto come metodo basato sulla frequenza relativa, si basa sul ripetere un gioco (esperimento) molte volte. La probabilità di vincere è circa uguale al numero di uscite favorevoli diviso il numero di prove effettuate. Pascal e Fermat usarono questo metodo per verificare i risultati ottenuti col metodo classico.

Ci sono altre versioni della storia, nelle quali il questito posto da de Méré è diverso (ad esempio in [Dev08], libro di piacevole lettura se siete interessati alla storia della scienza). Ciò che sembra certo è che de Méré pose il famoso quesito dopo essersi quasi rovinato col gioco dei dadi: prima si arricchì scommettendo sulla favorevole probabilità di ottenere un 6 su quattro lanci di dadi e poi andò in bancarotta scommettendo sulla scarsa probabilità di ottenere un doppio sei in 24 lanci di due dadi (come commentò in seguito Pascal con ironia era piuttosto ignorante in materia di calcolo matematico).

Sebbene vi siano esempi di calcolo delle probabilità anche precedenti, il lavoro di Pascal e Fermat pose le basi della moderna concezione della teoria della probabilità e causò una rivoluzione intellettuale. Per la prima volta si avevano strumenti per l'analisi dei giochi d'azzardo e, più importante ancora, per la predizione di eventi futuri. In precedenza la sola idea di poter prevedere la probabilità di occorrenza di eventi futuri era inconcepibile.

Nel corso del diciottesimo secolo compaiono le prime applicazioni della teoria della probabilità al di fuori dei giochi, principalmente nei campi delle assicurazioni e medico. Purtroppo sia il metodo classico che quello basato sulla frequenza relativa hanno grossi problemi ad essere applicati, e nonostante gli sforzi compiuti, la teoria della probabilità non è diventata una disciplina rigorosa (ed

usabile) fino al 1933, quando Andrei Nikolaevich Kolmogorov sviluppò un approccio rigoroso, noto come metodo assiomatico.

Ai giorni nostri la teoria della probabilità ha numerosissime applicazioni nelle più varie discipline. Nonostante la sua importanza e diffusione soffre di uno strano problema: la maggior parte delle persone (compresi puttroppo molti degli studenti che escono da Ingegneria) non ne sa quasi nulla, tanto da sbagliare il calcolo delle probabilità anche in casi elementari. Per averne un esempio guardate quanti siti internet e quanti programmi televisivi offrono metodi per predire i numeri del lotto e quanta gente gioca ai giochi d'azzardo legali (statali) o clandestini per speranze di vincita non corrispondenti alle probabilità del gioco. In base alla teoria della probabilità ed alle quote pagate l'unica predizione possibile sul gioco del lotto è la seguente: più giocate più la probabilità di perdere tende alla certezza. Alla fine del capitolo vi spiegherò perché.

5.1 Definizioni di probabilità

5.1.1 Definizione classica

La probabilità $P(\mathcal{A})$ di un evento \mathcal{A} è determinata a priori come:

$$P(\mathcal{A}) = \frac{N_{\mathcal{A}}}{N} \quad (5.1)$$

dove N è il numero delle uscite possibili e $N_{\mathcal{A}}$ il numero di uscite favorevoli all'evento \mathcal{A} .

Il significato dei numeri N e $N_{\mathcal{A}}$ non è tuttavia sempre chiaro. Facciamo un esempio: trovare la probabilità che tirando due dadi esca il numero sette.

Soluzione 1: N è 11 (esce 2, 3, ... 12), $N_{\mathcal{A}} = 1$, per cui $P(7) = 1/11$. Questo risultato è errato.

Soluzione 2: $N = 21$, ovvero tutte le possibili coppie di uscite senza distinguere fra il primo ed il secondo dado. $N_{\mathcal{A}} = 3$, ovvero gli eventi (3, 4), (5, 2) e (6, 1) e $P(7) = 3/21$. Anche questo risultato è errato.

Soluzione 3: $N = 36$, ovvero tutte le possibili coppie di uscite distinguendo fra il primo ed il secondo dado. $N_{\mathcal{A}} = 6$, ovvero gli eventi (3, 4), (4, 3), (2, 5), (5, 2), (1, 6) e (6, 1), $P(7) = 6/36$. Questo risultato è giusto.

L'esempio mostra che la definizione deve essere migliorata: un modo è quello di richiedere che il numero totale di eventi possibili sia riferito ad uscite ugualmente probabili. In questo modo introduciamo però il concetto di probabilità nella definizione di probabilità (e non è logicamente coerente). Vi sono anche altri problemi: ad esempio vi sono casi nei quali la definizione non si può applicare. Se un dado è truccato e la probabilità del 4 vale 0.4 non è possibile ricavare questa probabilità mediante la definizione. Se poi il numero di uscite è infinito applicare la definizione classica può risultare problematico. Come esempio si può citare il paradosso di Bertrand, in [Pap91] pag. 9-10.

Benché la definizione classica si sia rivelata inadatta come fondamento della teoria della probabilità resta comunque un utile strumento per la determinazione di probabilità reali.

5.1.2 Definizione basata sulla frequenza relativa

In base a questa definizione, la probabilità $P(\mathcal{A})$ di un evento \mathcal{A} è il limite

$$P(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{n_{\mathcal{A}}}{n} \quad (5.2)$$

dove $n_{\mathcal{A}}$ è il numero di occorrenze di \mathcal{A} ed n è il numero di tentativi.

Questa definizione appare ragionevole. Tuttavia, sebbene il concetto di frequenza relativa sia fondamentale nelle applicazioni, il suo uso come base per la teoria è questionabile. In particolare, poichè in un esperimento reale il numero di tentativi può anche essere grande ma non può essere infinito, ne segue che $P(\mathcal{A})$ può solo essere nota in modo approssimativo, non esatto. A cavallo fra il 1800 ed il 1900 furono fatti tentativi per fondare la teoria della probabilità su questa definizione (in particolare da Richard von Mises nel 1919), ma i risultati ottenuti non vennero considerati soddisfacenti.

5.1.3 Definizione assiomatica

Si deve, come già detto, al matematico Kolmogorov, che formalizza l'approccio assiomatico in un articolo del 1933. Si suppone che il lettore sia a conoscenza dei concetti fondamentali della teoria degli insiemi. L'approccio assiomatico alla probabilità si basa sui tre postulati (o *assiomi*) seguenti e su null'altro:

Postulato 1) La probabilità $P(\mathcal{A})$ di un evento \mathcal{A} è un numero positivo assegnato all'evento

$$P(\mathcal{A}) \geq 0$$

Postulato 2) La probabilità dell'evento certo vale 1.

Postulato 3) Se gli eventi \mathcal{A} e \mathcal{B} sono mutuamente esclusivi, allora

$$P(\mathcal{A} + \mathcal{B}) = P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$$

Definizioni. L'evento certo \mathcal{S} è l'evento che si verifica ad ogni tentativo. L'unione $\mathcal{A} + \mathcal{B}$ o $\mathcal{A} \cup \mathcal{B}$ di due eventi \mathcal{A} e \mathcal{B} è l'evento che si verifica quando \mathcal{A} o \mathcal{B} o entrambi si verificano. Gli eventi \mathcal{A} e \mathcal{B} sono mutuamente esclusivi se l'occorrenza di uno essi esclude l'occorrenza dell'altro. L'intersezione $\mathcal{A}\mathcal{B}$ o $\mathcal{A} \cap \mathcal{B}$ di due eventi \mathcal{A} e \mathcal{B} è l'evento che si verifica quando si verificano entrambi.

L'insieme evento certo è anche detto spazio campione o spazio degli eventi: i suoi elementi sono tutte e solo le possibili uscite (risultati) dell'esperimento considerato. I suoi sottoinsiemi sono gli eventi. In Appendice C di questo capitolo (5.16) sono riportati alcuni dei simboli usati nella teoria degli insiemi.

Corollari agli assiomi della probabilità

Teorema 1: la probabilità dell'evento impossibile \emptyset è zero:

$$P(\emptyset) = 0$$

Infatti, detto \mathcal{A} un generico evento, l'intersezione fra \mathcal{A} e \emptyset è \emptyset , cioè $\mathcal{A}\emptyset = \mathcal{A} \cap \emptyset = \emptyset$, l'unione di \mathcal{A} con \emptyset vale $\mathcal{A} + \emptyset = \mathcal{A}$ e quindi, essendo eventi mutuamente esclusivi, si ottiene

$$P(\mathcal{A}) = P(\mathcal{A} + \emptyset) = P(\mathcal{A}) + P(\{\emptyset\})$$

Teorema 2: la probabilità che non si verifichi un evento \mathcal{A} è il complemento a uno della probabilità che si verifichi

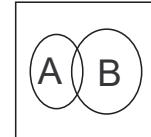
$$P(\overline{\mathcal{A}}) = 1 - P(\mathcal{A})$$

Infatti $\mathcal{A} + \overline{\mathcal{A}} = \mathcal{S}$ e $\mathcal{A}\overline{\mathcal{A}} = \{\emptyset\}$, quindi $1 = P(\mathcal{S}) = P(\mathcal{A} + \overline{\mathcal{A}}) = P(\mathcal{A}) + P(\overline{\mathcal{A}})$.

Teorema 3: se \mathcal{A} e \mathcal{B} sono due eventi

$$P(\mathcal{A} + \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A}\mathcal{B}) \leq P(\mathcal{A}) + P(\mathcal{B})$$

dove l'uguaglianza vale solo se gli eventi \mathcal{A} e \mathcal{B} sono mutuamente esclusivi (la semplice dimostrazione è omessa).



Il diagramma aiuta a capire le relazioni fra gli eventi: il rettangolo è lo spazio degli eventi o evento certo \mathcal{S} , gli eventi sono sottoinsiemi di \mathcal{S} . In figura gli eventi \mathcal{A} e \mathcal{B} sono rappresentati come parzialmente sovrapposti, quindi non sono mutuamente esclusivi (la zona comune rappresenta l'evento intersezione). L'area del rettangolo vale uno e le aree degli eventi rappresentano le relative probabilità.

Esercizio 5.1 Dato un mazzo di 52 carte, costituito, come noto, da 13x4 carte (le tredici sono dall'asso al 10 più le tre figure JQK, i quattro sono picche, denari, cuori e fiori) trovare le probabilità di:

- estrarre (a caso e senza guardare) l'otto di picche ($1/52$);
- estrarre una figura ($3/13$);
- estrarre un asso o una carta di fiori. Soluzione (vedi teorema 3):

$$P(\text{asso}) + P(\text{fiori}) - P(\text{asso e fiori}) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

- ottenere due assi estraendo contemporaneamente due carte.

Soluzione: ci sono $52 \cdot 51 = 2652$ coppie ordinate possibili, gli eventi favorevoli sono $3 \cdot 4 = 12$ (che sono tutte le coppie ordinate di assi), per cui la probabilità cercata è pari a $12/2652 \approx 0.0045$.

Lo stesso risultato si ottiene come

$$\binom{4}{2} / \binom{52}{2} = 12/2652$$

come vedremo nel seguito (sezione 5.9).

5.1.4 Interpretazione soggettiva

Secondo questa interpretazione la probabilità esprime il grado di fiducia di un singolo individuo nel verificarsi di un evento. La probabilità non è quindi una proprietà dell'evento, ma l'espressione di un giudizio soggettivo. Questa definizione riguarda quindi più la psicologia che la scienza e viene spesso usata in situazioni comuni per esprimere la fiducia che si ha sul verificarsi di un evento. In tale senso è applicabile anche ad eventi singoli (che non si ripeteranno) o ad eventi che, pur ripetibili, non si sono ancora verificati.

5.2 Definizione di esperimento

In teoria della probabilità è fondamentale definire con precisione i concetti di esperimento, dei suoi possibili risultati e degli eventi di cui si vuole studiare la probabilità. Dato un esperimento, l'insieme di tutti i suoi possibili risultati si chiama **spazio degli eventi** e si indica con \mathcal{S} . I punti di tale spazio sono i **risultati** (o uscite) dell'esperimento ed i sottoinsiemi di \mathcal{S} sono gli **eventi**. Lo spazio degli eventi \mathcal{S} è anche indicato come spazio campione, spazio dei campioni, spazio delle prove ed *evento certo*. L'insieme vuoto $\{\emptyset\}$ è detto *evento impossibile* e l'evento $\{\xi_i\}$ che contiene un singolo risultato ξ_i è detto *evento elementare*. Gli eventi elementari sono mutuamente esclusivi.

Nella applicazione della teoria della probabilità a problemi fisici, la definizione dei risultati di un esperimento non è sempre unica. Illustriamo questa ambiguità con un esempio, il lancio di un dado a sei facce: tre diversi giocatori, A, B e C, possono interpretarlo come segue:

- per il giocatore A i risultati dell'esperimento sono le sei facce del dado. lo spazio degli eventi è l'insieme $\mathcal{S} = \{f_1, f_2, f_3, f_4, f_5, f_6\}$ dei sei possibili risultati. Ciascun evento possibile è un sottoinsieme di \mathcal{S} e quindi abbiamo a disposizione $2^6 = 64$ eventi distinti. L'evento $\{\text{pari}\}$ consiste dei tre risultati f_2, f_4 e f_6 .
- per il giocatore B conta solo se il risultato è pari o dispari. Per lui l'esperimento ha due risultati. Questi formano lo spazio degli eventi $\mathcal{S} = \{\text{pari}, \text{dispari}\}$. Tale spazio ha due elementi, e quindi possiamo formare $2^2 = 4$ eventi distinti. L'evento $\{\text{pari}\}$ consiste di un solo risultato.
- per il giocatore C conta solo la posizione che il dado assume sul tavolo dopo il lancio. Il suo spazio degli eventi è quindi composto da tutti i punti del tavolo ed è formato da infiniti risultati. L'evento $\{\text{pari}\}$ non esiste ed esistono infiniti eventi distinti.

Tutte le possibili interpretazioni di un esperimento fisico sono legittime, ma è necessario rendersi conto che, in teoria della probabilità, ciascuna interpretazione corrisponde ad un diverso esperimento. Nel seguito quando introdurremo un esperimento, assumeremo sempre che i possibili risultati siano chiaramente identificati. Per inciso, d'ora in poi l'esperimento del lancio del dado avrà come spazio degli eventi $\mathcal{S} = \{f_1, f_2, f_3, f_4, f_5, f_6\}$.

Nella interpretazione in termini di frequenza relativa di vari risultati useremo la seguente terminologia: effettuare un esperimento una volta si chiama *prova*. Ad ogni prova osserviamo un singolo risultato ξ_i . Diremo che l'evento \mathcal{A} si è verificato durante la prova se il sottoinsieme \mathcal{A} contiene l'elemento ξ_i . L'evento certo si verifica ad ogni prova e l'evento impossibile non si verifica mai. L'evento $\mathcal{A} + \mathcal{B}$ si verifica se si verifica \mathcal{A} oppure \mathcal{B} oppure entrambi. L'evento \mathcal{AB} si verifica se entrambi gli eventi \mathcal{A} e \mathcal{B} si verificano. Se gli eventi \mathcal{A} e \mathcal{B} sono mutuamente esclusivi e si verifica \mathcal{A} , non può verificarsi \mathcal{B} . Se $\mathcal{A} \subset \mathcal{B}$ e si verifica \mathcal{A} , allora si verifica \mathcal{B} . Ad ogni prova o si verifica \mathcal{A} o si verifica $\overline{\mathcal{A}}$.

Classe degli eventi

Abbiamo visto che gli eventi sono sottoinsiemi dello spazio \mathcal{S} . In generale non è opportuno considerare sempre come eventi tutti i possibili sottoinsiemi di \mathcal{S} . Una ragione è quella di semplificarsi la vita. Ad esempio, nel lancio del dado, l'evento $\{f_2, f_5\}$ (ovvero l'evento che si verifica se il risultato è 2 oppure 5) in genere non interessa a nessuno. La ragione vera è però un'altra, ed è di natura matematica. In certi casi, che riguardano spazi degli eventi con infiniti risultati, risulta impossibile assegnare le probabilità a tutti i possibili sottoinsiemi in modo da soddisfare gli assiomi visti più uno che è la generalizzazione dell'assioma III, noto come assioma della infinita additività.

La necessità di questo assioma aggiuntivo nasce dalla seguente considerazione: se n eventi $\mathcal{A}_1 \dots \mathcal{A}_n$ sono mutuamente esclusivi, allora la ripetuta applicazione dell'assioma III porta a

$$P(\mathcal{A}_1 + \dots + \mathcal{A}_n) = P(\mathcal{A}_1) + \dots + P(\mathcal{A}_n) \quad (5.3)$$

L'estensione al caso di infiniti eventi non può tuttavia essere ottenuta in questo modo. Essa è una condizione aggiuntiva, nota appunto come **assioma della infinita additività**: se gli infiniti eventi $\mathcal{A}_1, \mathcal{A}_2, \dots$ sono mutuamente esclusivi, allora

$$P(\mathcal{A}_1 + \mathcal{A}_2 + \dots) = P((\mathcal{A}_1) + P((\mathcal{A}_2) + \dots \quad (5.4)$$

L'approfondimento di questi aspetti, che coinvolge i campi di Borel, non viene qui trattato (volendo, lo trovate in [Pap91], pag. 22 e seguenti).

Spazi finiti

Se lo spazio \mathcal{S} consiste di N elementi o risultati ed N è un numero finito, allora le probabilità di tutti gli eventi possono essere espresse in termini delle probabilità $P\{\xi_i\} = p_i$ degli eventi elementari $\{\xi_i\}$. Dagli assiomi segue che

$$p_i \geq 0 \quad p_1 + p_2 + \dots + p_N = 1$$

Supponiamo che \mathcal{A} sia un evento formato dagli r eventi elementari $\xi_{k_1}, \dots, \xi_{k_r}$. In questo caso \mathcal{A} può essere scritto come l'unione degli eventi elementari $\xi_{k_1}, \dots, \xi_{k_r}$, dall'assioma III, segue che

$$P(\mathcal{A}) = P\{\xi_{k_1}\} + \dots + P\{\xi_{k_r}\} = p_{k_1} + \dots + p_{k_r} \quad (5.5)$$

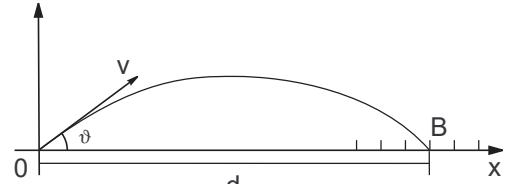
Attenzione a non confondere gli insiemi (o spazi) finiti con quelli **contabili**: un insieme si dice *contabile* se esso è *finito* o *numerabile*. Un insieme è numerabile se esso è equinumeroso con \mathbb{N} , ovvero con l'insieme dei numeri interi (che è infinito, ma numerabile). Ci sono insiemi infiniti non numerabili, ad esempio l'insieme dei numeri reali compresi fra zero e uno.

5.3 Causalità e casualità

A volte viene citata una controversia fra determinismo (causalità) e casualità. Il conflitto non ha ragione di essere se osserviamo che le teorie scientifiche non sono scoperte delle leggi di natura, ma invenzioni della mente umana. Le loro conseguenze vengono presentate in forma deterministica se siamo interessati ad eventi singoli, in forma probabilistica se siamo interessati alle regolarità presenti in numerosi eventi multipli. In entrambi i casi i risultati sono validi con limitazioni. Nel caso deterministico *entro un certo errore e se i parametri rilevanti sono in certi intervalli*. Nel caso probabilistico *con un alto grado di certezza se il numero degli eventi è grande a sufficienza*. Inoltre lo stesso evento può essere interpretato in entrambi i modi. Vediamo un esempio:

Un sasso viene lanciato con velocità iniziale v con una direzione che forma un angolo θ con l'orizzontale (vedi figura). Vogliamo trovare la distanza $d = OB$ fra l'origine ed il punto colpito B . Dalle leggi del moto di Newton segue che:

$$d = \frac{v^2}{g} \sin(2\theta)$$



Questo sembra un risultato esatto di una legge causale, ma non è così. Per trovare la soluzione abbiamo usato un modello idealizzato che trascura l'attrito dell'aria, le variazioni di g e le incertezze nei valori di v e θ (e in assenza di vento). Il risultato deve quindi essere accettato con limitazioni: vale con un certo errore ε se i fattori trascurati sono minori di δ . Possiamo dare una interpretazione probabilistica dello stesso esperimento. Supponiamo che l'area bersaglio sia formata da n buchi numerati vicini fra di loro e che vogliamo trovare in quale buco finirà il sasso. A causa delle incertezze viste non possiamo dare una risposta esatta. Possiamo tuttavia porci una differente domanda: se lanciamo molti sassi identici, tutti nominalmente con la stessa velocità e con lo stesso angolo, quanti di questi cadranno nel buco i -esimo? Questa è una tipica domanda probabilistica, cui possiamo rispondere in termini di probabilità.

Lo stesso problema fisico può quindi essere analizzato in forma deterministica o probabilistica. Si potrebbe obiettare che che il problema è deterministico, e che solo la nostra ignoranza sui parametri ci costringe a considerarlo probabilistico. Una possibile risposta è che i fisici non si occupano dei valori veri, ma solo di quelli osservati. Una diversa risposta è più "filosofica". Dopo la scoperta delle leggi di Newton l'universo veniva considerato come un grande meccanismo. In base a tali leggi se fossimo riusciti a stabilire con precisione la velocità e la posizione di tutte le particelle che lo compongono

(compito difficile ma non teoricamente impossibile) avremmo avuto accesso a tutta la storia, passata e futura, dell'universo. Questo crea grossi problemi alla nostra esistenza, in quanto, ad esempio, un tale universo non ha spazio per il libero arbitrio. Fortunatamente scoperte successive hanno dimostrato che esiste un livello di casualità ineliminabile a livello di particelle elementari (principio di indeterminazione, meccanica quantistica). Questo crea un limite teorico (e non solo pratico) alla precisione con la quale è possibile conoscere velocità e posizione di un oggetto e rende inevitabile il passaggio alla trattazione probabilistica quando le precisioni richieste sono sufficientemente alte. Ad esempio, qualcuno ha dimostrato che gli effetti quantistici diventano significativi, nel problema della traiettoria di una palla da biliardo, quando il numero di rimbalzi sulle sponde è almeno pari a dodici.

5.4 Probabilità condizionata

La probabilità condizionata di \mathcal{A} assumendo \mathcal{B} è la probabilità che si verifichi un certo evento \mathcal{A} sapendo che si è verificato un altro evento \mathcal{B} (dello stesso esperimento).

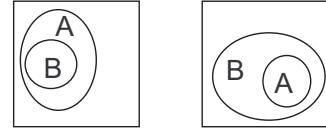
La probabilità di \mathcal{A} assumendo \mathcal{B} si indica $P(\mathcal{A}/\mathcal{B})$ e vale, per definizione

$$P(\mathcal{A}/\mathcal{B}) = \frac{P(\mathcal{A}\mathcal{B})}{P(\mathcal{B})} \quad \text{dove } P(\mathcal{B}) \neq 0 \quad (5.6)$$

Se $P(\mathcal{B}) = 0$ la formula non è applicabile: d'altra parte non ha significato cercare una probabilità condizionata ad un evento a probabilità nulla.

Proprietà:

- se $\mathcal{B} \subset \mathcal{A}$ allora $P(\mathcal{A}/\mathcal{B}) = 1$
- se $\mathcal{A} \subset \mathcal{B}$ allora $P(\mathcal{A}/\mathcal{B}) = P(\mathcal{A})/P(\mathcal{B}) \geq P(\mathcal{A})$



Interpretazione in termini di frequenza relativa. Siano $n_{\mathcal{A}}$, $n_{\mathcal{B}}$ e $n_{\mathcal{AB}}$ il numero di occorrenze degli eventi \mathcal{A} , \mathcal{B} e \mathcal{AB} in n tentativi.

$$\begin{aligned} P(\mathcal{A}) &\approx \frac{n_{\mathcal{A}}}{n} & P(\mathcal{B}) &\approx \frac{n_{\mathcal{B}}}{n} & P(\mathcal{AB}) &\approx \frac{n_{\mathcal{AB}}}{n} \\ P(\mathcal{A}/\mathcal{B}) &= \frac{P(\mathcal{AB})}{P(\mathcal{B})} \approx \frac{n_{\mathcal{AB}}/n}{n_{\mathcal{B}}/n} = \frac{n_{\mathcal{AB}}}{n_{\mathcal{B}}} \end{aligned}$$

che ha la seguente interpretazione: se scartiamo tutte le prove nelle quali l'evento \mathcal{B} non si verifica, la $P(\mathcal{A}/\mathcal{B})$ corrisponde alla frequenza relativa $n_{\mathcal{AB}}/n_{\mathcal{B}}$ dell'evento \mathcal{A} nelle prove non scartate.

Da notare che, a \mathcal{B} specificato, la probabilità condizionata $P(\mathcal{A}/\mathcal{B})$ è una probabilità, ovvero soddisfa i tre assiomi visti.

Esempio 5.2 Nel lancio del dado (non truccato) trovare la probabilità dell'evento $\mathcal{A} = \{\text{esce il numero } 2\}$ assumendo l'evento $\mathcal{B} = \{\text{esce un numero pari}\}$.

Chiaramente $P(\mathcal{A}) = 1/6$ e $P(\mathcal{B}) = 3/6$, inoltre $P(\mathcal{AB}) = P(\mathcal{A})$ poiché $\mathcal{AB} = \mathcal{A}$. Quindi

$$P(\mathcal{A}/\mathcal{B}) = P(\text{due/pari}) = \frac{P(\mathcal{A})}{P(\mathcal{B})} = \frac{1}{6} \frac{6}{3} = \frac{1}{3}$$

Esempio 5.3 Una scatola contiene tre palline bianche, b_1 , b_2 e b_3 , e due palline rosse, r_1 e r_2 . Estraiamo casualmente due palline in successione. Quanto vale la probabilità che la prima sia bianca e la seconda rossa?

Vediamo due soluzioni:

SOLUZIONE 1 Lo spazio degli eventi consiste in tutte le coppie ordinate che possiamo formare con cinque palline

$$(b_1, b_2) \ (b_1, b_3) \ (b_1, r_1) \ (b_1, r_2) \ \dots \ (r_2, b_1) \ (r_2, b_2) \ (r_2, b_3) \ (r_2, r_1)$$

Il numero delle coppie possibili è quindi $5 \cdot 4 = 20$. L'evento {prima bianca, seconda rossa} consiste delle 6 uscite

$$(b_1, r_1) \ (b_1, r_2) \ (b_2, r_1) \ (b_2, r_2) \ (b_3, r_1) \ (b_3, r_2)$$

La probabilità cercata è allora 6/20.

SOLUZIONE 2 Utilizziamo le probabilità condizionate.

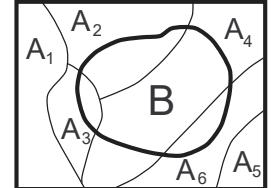
La probabilità dell'evento $\mathcal{B}_1 = \{\text{prima bianca}\}$ vale $P(\mathcal{B}_1) = 3/5$. Togliendo una pallina bianca ne rimangono due rosse e due bianche; la probabilità dell'evento $\mathcal{R}_2 = \{\text{seconda rossa}\}$ assumendo \mathcal{B}_1 vale quindi 2/4. Perciò:

$$P(\mathcal{B}_1 \mathcal{R}_2) = P(\mathcal{R}_2 / \mathcal{B}_1) P(\mathcal{B}_1) = \frac{2}{4} \cdot \frac{3}{5} = \frac{6}{20}$$

5.5 Teoremi della probabilità totale e di Bayes

Sia $\mathcal{U} = [\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n]$ una **partizione** di \mathcal{S} , ovvero una collezione di n sottoinsiemi mutuamente esclusivi A_i di \mathcal{S} tali che la loro unione sia pari a \mathcal{S} . Sia \mathcal{B} un evento arbitrario. Il teorema della probabilità totale sostiene che

$$P(\mathcal{B}) = P(\mathcal{B}/\mathcal{A}_1)P(\mathcal{A}_1) + \dots + P(\mathcal{B}/\mathcal{A}_n)P(\mathcal{A}_n) \quad (5.7)$$



Dimostrazione: chiaramente

$$\mathcal{B} = \mathcal{B}\mathcal{S} = \mathcal{B}(\mathcal{A}_1 + \dots + \mathcal{A}_n) = \mathcal{B}\mathcal{A}_1 + \dots + \mathcal{B}\mathcal{A}_n$$

Gli eventi $\mathcal{B}\mathcal{A}_i$ e $\mathcal{B}\mathcal{A}_j$, $i \neq j$, sono mutuamente esclusivi perché sono mutuamente esclusivi gli eventi \mathcal{A}_i e \mathcal{A}_j . Quindi

$$P(\mathcal{B}) = P(\mathcal{B}\mathcal{A}_1) + \dots + P(\mathcal{B}\mathcal{A}_n)$$

Poichè $P(\mathcal{B}\mathcal{A}_i) = P(\mathcal{B}/\mathcal{A}_i)P(\mathcal{A}_i)$ (dall'eq. 5.6) la tesi è dimostrata.

La probabilità $P(\mathcal{B}\mathcal{A}_i)$ si può scrivere in due modi differenti:

$$P(\mathcal{B}\mathcal{A}_i) = P(\mathcal{B}/\mathcal{A}_i)P(\mathcal{A}_i) = P(\mathcal{A}_i/\mathcal{B})P(\mathcal{B})$$

da cui

$$P(\mathcal{A}_i/\mathcal{B}) = \frac{P(\mathcal{B}/\mathcal{A}_i)P(\mathcal{A}_i)}{P(\mathcal{B})} \quad (5.8)$$

Inserendo ora l'espressione di $P(\mathcal{B})$ secondo il teorema della probabilità totale otteniamo il **teorema di Bayes**

$$P(\mathcal{A}_i/\mathcal{B}) = \frac{P(\mathcal{B}/\mathcal{A}_i)P(\mathcal{A}_i)}{P(\mathcal{B}/\mathcal{A}_1)P(\mathcal{A}_1) + \dots + P(\mathcal{B}/\mathcal{A}_n)P(\mathcal{A}_n)} \quad (5.9)$$

$P(\mathcal{A}_i)$ e $P(\mathcal{A}_i/\mathcal{B})$ sono spesso indicate come probabilità a priori e probabilità a posteriori, rispettivamente.

Esempio 5.4 In un paese, il 70% della popolazione ha i capelli scuri, il 20% biondi e il 10% rossi. Risulta poi che hanno gli occhi chiari l'8% degli scuri, il 50% dei biondi e il 25% dei rossi. Se sapete che una certa persona ha gli occhi chiari, che probabilità c'è che abbia i capelli scuri?

SOLUZIONE Le probabilità date, con ovvia notazione, sono:

$$P(S) = 0.7 \quad P(B) = 0.2 \quad P(R) = 0.1$$

$$P(C/S) = 0.08 \quad P(C/B) = 0.5 \quad P(C/R) = 0.25$$

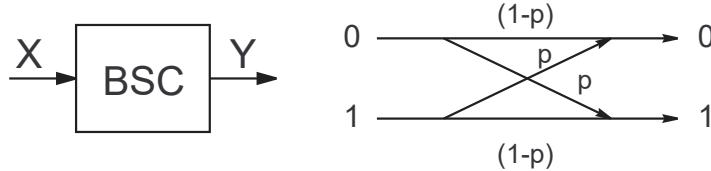
Per calcolare $P(S/C)$ utilizziamo il teorema di Bayes

$$P(S/C) = \frac{P(C/S)P(S)}{P(C)} = \frac{P(C/S)P(S)}{P(C/S)P(S) + P(C/B)P(B) + P(C/R)P(R)}$$

$$P(S/C) = \frac{0.08 \cdot 0.7}{0.08 \cdot 0.7 + 0.5 \cdot 0.2 + 0.25 \cdot 0.1} = \frac{0.056}{0.056 + 0.1 + 0.025} = \frac{0.056}{0.181} \approx 0.31$$

5.5.1 Esempio: BSC

Si consideri il canale binario simmetrico (BSC) in figura



dove: $X = \{0, 1\}$, $Y = \{0, 1\}$, $p = P(\{Y = 0\}/\{X = 1\}) = P(\{Y = 1\}/\{X = 0\})$

Chiamiamo inoltre $P_0 = P(\{X = 0\})$.

Il canale binario simmetrico BSC rappresenta un semplice modello di canale "rumoroso". In tale canale, a causa del rumore, non sempre il simbolo di uscita coincide col simbolo di ingresso. Il canale è detto simmetrico perché la probabilità di transizione (cioè $P(Y \neq X)$) non dipende dal simbolo trasmesso.

Considerando come esperimento il passaggio di un simbolo sul canale possiamo definire lo spazio degli eventi \mathcal{S} come l'insieme di tutte le possibili coppie (X, Y) : $\mathcal{S} = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

La probabilità di ciascun evento è facilmente calcolabile: ad esempio

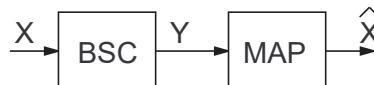
$$P(0, 0) = P(X = 0)P(\{Y = 0\}/\{X = 0\}) = P_0(1 - p).$$

La probabilità di transizione (o errore sul canale) si ottiene dalla legge della probabilità totale:

$$\begin{aligned} P(X \neq Y) &= P(X = 0)P(\{Y \neq X\}/\{X = 0\}) + P(X = 1)P(\{Y \neq X\}/\{X = 1\}) \\ &= P_0p + (1 - P_0)p = p \end{aligned}$$

Nei sistemi di trasmissione si usa distinguere fra probabilità di transizione (o errore sul canale), che è la probabilità che un simbolo arrivi errato, e probabilità di errore, che solitamente indica la probabilità che un simbolo arrivi errato considerando le tecniche di codifica (protezione dal rumore) presenti a monte e a valle del canale. Ovvero, dato un canale con una certa probabilità di transizione, posso applicare tecniche di codifica e/o di decisione ottima in modo tale che la probabilità di avere effettivamente un errore in ricezione sia minore della probabilità di transizione. In genere questo può essere fatto inserendo opportuni simboli di protezione, ovvero trasmettendo dati in più che mi permettano di "correggere" gli errori sul canale. Come questo si realizza in pratica è argomento di corsi successivi. Vediamo qui un esempio di cosa vuol dire "decidere" il simbolo trasmesso e come questo porti, in alcuni casi, ad avere una probabilità di errore minore della probabilità di transizione: questo esempio è solo un assaggio delle tecniche di decisione, senza applicazioni pratiche dirette, giusto per introdurre concetti di notevole importanza.

Supponiamo di aggiungere al nostro sistema un nuovo componente:



Il blocco indicato MAP ha il compito di generare una uscita $\hat{X} \in \{0, 1\}$ tale che sia minima la probabilità di errore definita come $P(e) = P(\hat{X} \neq X)$.

Il blocco MAP riceve in ingresso Y , cioè l'uscita del BSC, e conosce sia $P_0 = P(\{X = 0\})$ che p , probabilità di transizione sul canale. Il nostro compito è quello di progettare il blocco MAP minimizzando la $P(e)$. Vedremo che non sempre la scelta ovvia ($\hat{X} = Y$) è quella che minimizza $P(e)$. Chiamiamo X_0 il caso in cui sia stato trasmesso $X = 0$ e X_1 il caso in cui sia stato trasmesso $X = 1$. Analogamente Y_0 ed Y_1 indicheranno che l'uscita del BSC vale $Y = 0$ e $Y = 1$. Per trovare il criterio che minimizza $P(e)$ chiamiamo $P(X_i/Y_j)$, $i, j = 0, 1$, la probabilità condizionata che sia stato trasmesso X_i dato che l'uscita del BSC è Y_j . Se, dato che ho ricevuto Y_j , $P(X_0/Y_j) > P(X_1/Y_j)$, ovvero la probabilità che sia stato trasmesso $X = 0$ è maggiore della probabilità che sia stato trasmesso $X = 1$, allora la scelta che minimizza $P(e)$ è $\hat{X} = 0$. Se $P(X_1/Y_j) > P(X_0/Y_j)$ la scelta è $\hat{X} = 1$. Se sono uguali la scelta è indifferente. Le quantità $P(X_i/Y_j)$ sono note come probabilità a posteriori, ovvero dopo l'osservazione, ed un ricevitore che opera secondo il criterio esposto si chiama MAP (Massima probabilità a Posteriori, da Maximum A Posteriori).

Restano da calcolare le $P(X_i/Y_j)$: usando l'eq. 5.8 si ottiene

$$P(X_i/Y_j) = P(Y_j/X_i) \frac{P(X_i)}{P(Y_j)}$$

Supponiamo che sia stato ricevuto Y_0 , ovvero $Y = 0$; dobbiamo confrontare le due probabilità

$$P(X = 0/Y = 0) \stackrel{?}{<= >} P(X = 1/Y = 0)$$

cioè

$$P(Y = 0/X = 0) \frac{P(X = 0)}{P(Y = 0)} \stackrel{?}{<= >} P(Y = 0/X = 1) \frac{P(X = 1)}{P(Y = 0)}$$

Sostituendo i valori noti ed eliminando $P(Y = 0)$, che è una quantità positiva (il caso $P(Y = 0) = 0$ si tratta a parte...) otteniamo

$$(1 - p)P_0 \stackrel{?}{<= >} p(1 - P_0)$$

quindi, essendo sia p che P_0 non negative, devo confrontare

$$P_0 \stackrel{?}{<= >} p$$

Se $P_0 > p$ allora devo scegliere $\hat{X} = 0$ quando ricevo $Y = 0$, se $P_0 < p$ devo scegliere $\hat{X} = 1$. Se $P_0 = p$ posso scegliere sia 0 che 1, ottenendo la stessa $P(e/Y_0)$ (è indifferente).

Se ricevo Y_1 , ovvero $Y = 1$, un analogo calcolo mi porta scegliere $\hat{X} = 1$ se $(1 - P_0) > p$ e $\hat{X} = 0$ se $(1 - P_0) < p$.

Vediamo come funziona il blocco MAP in alcuni casi:

1. Se $P_0 = 1 - P_0 = 0.5$ e $p < 0.5$ scelgo $\hat{X} = 0$ per $Y = 0$ e $\hat{X} = 1$ per $Y = 1$. $P(e) = p$.
2. Se $P_0 = 1 - P_0 = 0.5$ e $p > 0.5$ scelgo $\hat{X} = 1$ per $Y = 0$ e $\hat{X} = 0$ per $Y = 1$. $P(e) = 1 - p < p$.
3. Se $P_0 = 1 - P_0 = 0.5$ e $p = 0.5$ la scelta è indifferente. $P(e) = 0.5$ ed il canale è inutile (ottengo lo stesso risultato se lancio una moneta e decido di conseguenza, ignorando l'uscita del canale).
4. Se $P_0 = 0.1$, $(1 - P_0) = 0.9$ e $p = 0.2$ scelgo $\hat{X} = 1$ per $Y = 0$ e $\hat{X} = 1$ per $Y = 1$. $P(e) = 0.1 < p$. Da notare che il risultato migliore si ottiene decidendo sempre per 1, ignorando l'uscita del canale.

I risultati ottenuti sono abbastanza banali, l'esempio riguarda il metodo, di grande utilità nei casi reali e più complicati.

Calcolo della $P(e)$

Trovare la probabilità di errore richiede, in generale, di esaminare tutti i casi possibili mediando poi i risultati in base alle probabilità. Questo si ottiene mediante il teorema della probabilità totale. Nell'esempio appena visto la $P(e)$ si calcola come:

$$P(e) = P(e/Y_0)P(Y_0) + P(e/Y_1)P(Y_1)$$

dove, sempre usando la probabilità totale,

$$P(Y_0) = P(Y_0/X_0)P(X_0) + P(Y_0/X_1)P(X_1) = (1 - p)P_0 + p(1 - P_0)$$

$$P(Y_1) = P(Y_1/X_0)P(X_0) + P(Y_1/X_1)P(X_1) = pP_0 + (1 - p)(1 - P_0)$$

5.6 Eventi indipendenti

Due eventi \mathcal{A} e \mathcal{B} si dicono **indipendenti** se vale la relazione

$$P(\mathcal{A}\mathcal{B}) = P(\mathcal{A}) P(\mathcal{B}) \tag{5.10}$$

Una formulazione equivalente è:

$$P(\mathcal{A}/\mathcal{B}) = \frac{P(\mathcal{A}\mathcal{B})}{P(\mathcal{B})} = P(\mathcal{A}) \tag{5.11}$$

Ovvero la probabilità di occorrenza di \mathcal{A} non è influenzata dal verificarsi o meno di \mathcal{B} . In questo caso si dice che l'evento \mathcal{A} è statisticamente indipendente da \mathcal{B} . Naturalmente se è vera l'eq. 5.11 è vera anche

$$P(\mathcal{B}/\mathcal{A}) = \frac{P(\mathcal{A}\mathcal{B})}{P(\mathcal{A})} = P(\mathcal{B})$$

Quindi l'indipendenza di \mathcal{A} da \mathcal{B} implica l'indipendenza di \mathcal{B} da \mathcal{A} .

Il concetto di indipendenza statistica è fondamentale. Se ci limitiamo allo sviluppo assiomatico delle probabilità, l'indipendenza è definita dall'eq. 5.10 e non vi sono problemi. Stabilire invece se due eventi del mondo reale siano o meno indipendenti può essere piuttosto complicato. Come caso limite vi presento il seguente: esistono religioni che considerano l'universo come un tutto, dove qualsiasi cosa accada è legata a tutte le altre in modo olistico. Se siete seguaci di tali religioni gli eventi indipendenti semplicemente non esistono.

Anche limitandoci al campo scientifico sorgono comunque delle difficoltà: spesso considerare indipendenti due eventi è solo una ipotesi che va verificata. Da notare che spesso il progresso scientifico ha avuto per base la scoperta che due eventi considerati indipendenti in realtà non lo sono (e viceversa): ad esempio la puntura di un certo tipo di zanzara e l'ammalarsi di malaria (che, come indica il nome, era considerata frutto dell'aria insalubre).

NOTA: attenzione a non fare confusione fra eventi indipendenti ed eventi mutuamente esclusivi. Poichè per due eventi mutuamente esclusivi vale che $P(\mathcal{A}\mathcal{B}) = 0$ essi NON sono indipendenti (infatti l'uscita di uno infilza l'uscita dell'altro).

Indipendenza di $\overline{\mathcal{A}}$ e $\overline{\mathcal{B}}$

Se gli eventi \mathcal{A} e \mathcal{B} sono indipendenti lo sono anche gli eventi $\overline{\mathcal{A}}$ e $\overline{\mathcal{B}}$? L'intuizione ci dice di sì, e possiamo dimostrarlo. Sappiamo che

$$P(\mathcal{A}\mathcal{B}) = P(\mathcal{A}) P(\mathcal{B})$$

e calcoliamo

$$\begin{aligned} P(\overline{\mathcal{A}}) P(\overline{\mathcal{B}}) &= [1 - P(\mathcal{A})] [1 - P(\mathcal{B})] \\ P(\overline{\mathcal{A}}\overline{\mathcal{B}}) &= P(\overline{\mathcal{A}}/\overline{\mathcal{B}}) P(\overline{\mathcal{B}}) = [1 - P(\mathcal{A}/\overline{\mathcal{B}})] [1 - P(\mathcal{B})] \end{aligned}$$

ma

$$P(\mathcal{A}/\overline{\mathcal{B}}) = \frac{P(\overline{\mathcal{B}}/\mathcal{A})P(\mathcal{A})}{P(\overline{\mathcal{B}})} = \frac{[1 - P(\mathcal{B}/\mathcal{A})]P(\mathcal{A})}{1 - P(\mathcal{B})} = \frac{[1 - P(\mathcal{B})]P(\mathcal{A})}{1 - P(\mathcal{B})}$$

l'ultimo passaggio, $P(\mathcal{B}/\mathcal{A}) = P(\mathcal{B})$, deriva direttamente dall'indipendenza fra \mathcal{A} e \mathcal{B} . Quindi

$$P(\overline{\mathcal{A}}\overline{\mathcal{B}}) = [1 - \frac{[1 - P(\mathcal{B})]P(\mathcal{A})}{1 - P(\mathcal{B})}] [1 - P(\mathcal{B})] = [1 - P(\mathcal{A})] [1 - P(\mathcal{B})]$$

Allora

$$P(\overline{\mathcal{A}}\overline{\mathcal{B}}) = P(\overline{\mathcal{A}}) P(\overline{\mathcal{B}})$$

e gli eventi $\overline{\mathcal{A}}$ e $\overline{\mathcal{B}}$ sono indipendenti.

5.6.1 Indipendenza di tre eventi

Tre eventi \mathcal{A}_1 , \mathcal{A}_2 e \mathcal{A}_3 sono detti (mutuamente) indipendenti se sono indipendenti a coppie

$$P(\mathcal{A}_i\mathcal{A}_j) = P(\mathcal{A}_i)P(\mathcal{A}_j) \quad i \neq j$$

e se

$$P(\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3) = P(\mathcal{A}_1)P(\mathcal{A}_2)P(\mathcal{A}_3)$$

Da notare quindi che tre eventi possono essere indipendenti a coppie ma non indipendenti.

Non è facile trovare un esempio di tre eventi indipendenti a coppie ma non indipendenti: uno (un po' cervellotico) è il seguente. Supponiamo di lanciare in successione tre monete. La prime due sono normali monete affidabili, la terza moneta è magica: esce testa se le prime due monete mostrano la stessa faccia, altrimenti esce croce. Chiamiamo \mathcal{A}_1 l'evento "esce testa sulla moneta 1", \mathcal{A}_2 l'evento "esce testa sulla moneta 2" e \mathcal{A}_3 l'evento "esce testa sulla moneta 3". Le relative probabilità sono: $P(\mathcal{A}_1) = P(\mathcal{A}_2) = P(\mathcal{A}_3) = 0.5$. E' ora facile vedere che i tre eventi sono indipendenti a coppie ma non indipendenti, infatti

$$P(\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3) = P(\mathcal{A}_1)P(\mathcal{A}_2) = \frac{1}{4} \neq P(\mathcal{A}_1)P(\mathcal{A}_2)P(\mathcal{A}_3)$$

La tabella che segue riporta le probabilità degli eventi possibili

Moneta 1	Moneta 2	Moneta 3	Probabilità
T	C	C	1/4
T	T	T	1/4
C	C	T	1/4
C	T	C	1/4

5.7 Esperimenti con infiniti risultati

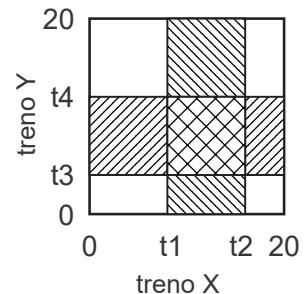
La teoria delle probabilità non è in grado di trattare esperimenti che abbiano infiniti risultati, se non in casi molto particolari. Un modo per rendersene conto è il seguente: supponiamo che un esperimento abbia N possibili risultati, con N finito. In questo caso se un risultato ha probabilità zero vuol dire che non si verifica mai, ovvero che è un risultato impossibile. Consideriamo ora un esperimento con infiniti risultati: come esempio usiamo il lancio di una freccetta su di un bersaglio quando i risultati sono le distanze fra il punto colpito dalla freccetta ed il centro del bersaglio. Come ovvio i possibili risultati sono infiniti, visto che supponiamo di avere infinita precisione nel determinare le distanze. Ciascun risultato ha allora probabilità zero, ma ogni volta che si lancia una freccetta un qualche punto viene colpito: quindi, in questo caso, i risultati a probabilità zero non sono impossibili, visto che se ne verifica uno tutte le volte che si esegue l'esperimento! Il problema vero è il seguente: se sono un lanciatore decente, la probabilità di colpire vicino al centro (diciamo entro due palmi) sarà maggiore della probabilità che il mio lancio colpisca molto lontano (a distanza maggiore di due metri). Con la teoria delle probabilità non c'è tuttavia modo di esprimere questo fatto, visto che comunque la probabilità di colpire un punto è sempre zero. Posso ovviamente approssimare l'analisi dividendo il bersaglio in N corone circolari concentriche: in questo modo l'esperimento ha un numero finito di risultati ed è trattabile, ma è solo una approssimazione.

Per trattare in modo esatto gli esperimenti con infiniti risultati è stata sviluppata la teoria delle variabili aleatorie, che vedremo nel prossimo capitolo. Con la teoria delle probabilità qualcosa si può tuttavia ottenere, almeno nel caso semplificato nel quale tutti gli infiniti risultati siano *equiprobabili*¹. L'esempio che segue illustra come.

5.7.1 Esempio: treni

In questo esempio vedremo come calcolare alcune probabilità sfruttando il concetto di indipendenza. Due treni, X e Y arrivano in stazione in maniera casuale fra le ore 8.00 e le 8.20 (otto e venti minuti) del mattino. Il treno X si ferma per 4 minuti, il treno Y per 5 minuti. Assumendo, per ipotesi, che i due treni arrivino in modo indipendente, vogliamo determinare alcune probabilità legate ai tempi x ed y di arrivo dei treni X e Y .

E' necessario specificare precisamente in cosa consiste l'esperimento: le sue uscite sono tutti i punti (x, y) del quadrato in figura. L'evento \mathcal{A} è definito come $\{\text{il treno } X \text{ arriva nell'intervallo } [t_1, t_2]\}$, ovvero $\mathcal{A} = \{t_1 \leq x \leq t_2\}$ (considerare o meno gli istanti terminali è indifferente). L'evento \mathcal{A} è rappresentato dalla striscia verticale in figura, la sua probabilità è pari a $(t_2 - t_1)/20$, se t_1 e t_2 sono espressi in minuti. Questo risultato è l'interpretazione dell'informazione che il treno arriva in maniera casuale, ovvero che nessun istante è privilegiato rispetto ad un altro (quando introdurremo le variabili aleatorie vedremo definizioni maggiormente rigorose, l'assunzione appena fatta corrisponde a dire che la variabile aleatoria x è uniformemente distribuita).



In modo analogo l'evento \mathcal{B} è definito come $\{\text{il treno } Y \text{ arriva nell'intervallo } [t_3, t_4]\}$, ovvero $\mathcal{B} = \{t_3 \leq y \leq t_4\}$: \mathcal{B} è rappresentato dalla striscia orizzontale e la sua probabilità è pari a $(t_4 - t_3)/20$. Interpretando l'indipendenza del tempo di arrivo come indipendenza degli eventi \mathcal{A} e \mathcal{B} , possiamo ottenere la probabilità dell'evento intersezione

$$P(\mathcal{AB}) = P(\mathcal{A})P(\mathcal{B}) = \frac{(t_2 - t_1)(t_4 - t_3)}{20 \cdot 20}$$

¹ Attenzione, qui il concetto di equiprobabilità è piuttosto confuso, dato che tutti gli eventi hanno probabilità zero... Di fatto si richiede che non vi siano eventi "privilegiati" rispetto ad altri, nel senso che la probabilità di capitare in un piccolo intervallo non varia al variare della posizione dell'intervallo.

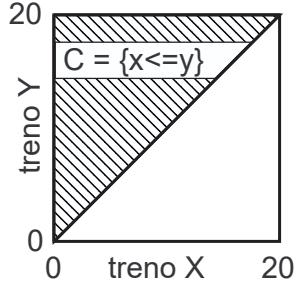
L'evento \mathcal{AB} è il rettangolo centrale: la sua probabilità è pari alla sua area diviso l'area totale, che è 400. In generale la probabilità che il punto (x, y) sia in una regione \mathcal{R} è pari all'area di \mathcal{R} diviso l'area totale. Vediamo ora di calcolare le probabilità di alcuni eventi.

- A) Trovare la probabilità che il treno X arrivi prima del treno Y . Questa è la probabilità dell'evento

$$\mathcal{C} = \{x \leq y\}$$

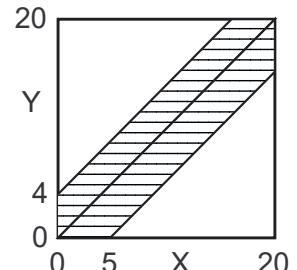
come mostrato in figura. La sua probabilità è

$$P(\mathcal{C}) = \frac{200}{400} = 0.5$$



- B) Trovare la probabilità che i treni si incontrino in stazione. Affinchè ciò avvenga x deve essere minore di $y + 5$ e y deve essere minore di $x + 4$. Questo è l'evento $\mathcal{D} = \{-4 \leq x - y \leq 5\}$, come mostrato in figura (area tratteggiata). La sua probabilità si ottiene mediante il calcolo delle aree e vale

$$P(\mathcal{D}) = \frac{159.5}{400}$$



- C) Assumendo che i treni si incontrino, trovare la probabilità che il treno X arrivi prima del treno Y . La risposta, da quanto visto sulle probabilità condizionate, vale

$$P(\mathcal{C}/\mathcal{D}) = \frac{P(\mathcal{CD})}{P(\mathcal{D})} = \frac{72}{159.5}$$

visto che

$$P(\mathcal{CD}) = \frac{1}{2} - \frac{16 \cdot 16}{2 \cdot 400} = \frac{72}{400}$$

Questo tipo di esercizio è comune negli scritti e nei compitini.

5.8 Esperimenti congiunti

Consideriamo due esperimenti: il primo è il lancio di un dado affidabile, definito da

$$\mathcal{S}_1 = \{f_1, \dots, f_6\} \quad P_1\{f_i\} = \frac{1}{6}$$

Il secondo esperimento è il lancio di una moneta affidabile, definito da

$$\mathcal{S}_2 = \{T, C\} \quad P_2\{T\} = P_2\{C\} = \frac{1}{2}$$

Eseguiamo entrambi gli esperimenti e cerchiamo di trovare la probabilità di ottenere due sul dado e testa sulla moneta. Se facciamo la ragionevole supposizione che i risultati del primo esperimento siano indipendenti da quelli del secondo, concludiamo che la probabilità è pari a $1/6 \cdot 1/2$.

La conclusione ottenuta è ragionevole: tuttavia la nozione di indipendenza usata per trovarla è in disaccordo con la definizione data in 5.10. In tale definizione, gli eventi \mathcal{A} e \mathcal{B} erano sottoinsiemi dello stesso spazio. Per usare l'indipendenza bisogna quindi costruire uno spazio \mathcal{S} che abbia come sottoinsiemi gli eventi "due" e "testa". Questo si realizza come segue: i due esperimenti sono visti come un singolo esperimento i risultati del quale sono le coppie $\xi_1 \xi_2$, dove ξ_1 è una delle sei facce del dado e ξ_2 è testa o croce. Lo spazio risultante consiste di dodici elementi

$$\mathcal{S} = \{f_1 T, \dots, f_6 T, f_1 C, \dots, f_6 C\}$$

In tale spazio $\{\text{due}\}$ non è un evento elementare ma un sottoinsieme costituito da due elementi: $\{\text{due}\} = \{f_2 T, f_2 C\}$. In modo analogo $\{\text{testa}\}$ è un evento con sei elementi: $\{\text{testa}\} = \{f_1 T, \dots, f_6 T\}$.

Per completare la definizione dell'esperimento dobbiamo assegnare una probabilità a tutti i sottoinsiemi di \mathcal{S} . Chiaramente l'evento {due} si verifica se il dado mostra "due" qualsiasi sia il valore mostrato dalla moneta. Ne concludiamo che la probabilità dell'evento {due} è pari a 1/6. Similmente concludiamo che la probabilità dell'evento {testa} è pari a 1/2. L'intersezione degli eventi {due} e {testa} è l'evento elementare $\{f_2, T\}$. Assumendo che gli eventi {due} e {testa} siano indipendenti nel senso della definizione 5.10, concludiamo che $P(\{f_2, T\}) = \frac{1}{6} \cdot \frac{1}{2}$, in accordo con la nostra prima conclusione.

5.8.1 Prodotto cartesiano di due insiemi

Dati due insiemi \mathcal{S}_1 \mathcal{S}_2 con elementi ξ_1 e ξ_2 rispettivamente, formiamo tutte le coppie *ordinate* $\xi_1\xi_2$ nelle quali ξ_1 è ogni elemento di \mathcal{S}_1 e ξ_2 è ogni elemento di \mathcal{S}_2 . Il prodotto cartesiano di \mathcal{S}_1 e \mathcal{S}_2 è l'insieme \mathcal{S} formato da tutte queste coppie:

$$\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2$$

Se \mathcal{A} è un sottoinsieme di \mathcal{S}_1 e \mathcal{B} è un sottoinsieme di \mathcal{S}_2 , allora l'insieme $\mathcal{C} = \mathcal{A} \times \mathcal{B}$ costituito da tutte le coppie $\xi_1\xi_2$ nelle quali $\xi_1 \in \mathcal{A}$ e $\xi_2 \in \mathcal{B}$ è un sottoinsieme di \mathcal{S} .

Se formiamo in modo analogo gli insiemi $\mathcal{A} \times \mathcal{S}_2$ e $\mathcal{S}_1 \times \mathcal{B}$, vediamo che la loro intersezione è l'insieme $\mathcal{A} \times \mathcal{B}$

$$\mathcal{A} \times \mathcal{B} = (\mathcal{A} \times \mathcal{S}_2) \cap (\mathcal{S}_1 \times \mathcal{B}) \quad (5.12)$$

5.8.2 Prodotto cartesiano di due esperimenti

Il prodotto cartesiano di due esperimenti \mathcal{S}_1 \mathcal{S}_2 è un nuovo esperimento $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2$ i cui eventi sono tutti i prodotti cartesiani

$$\mathcal{A} \times \mathcal{B}$$

dove \mathcal{A} è un evento di \mathcal{S}_1 e \mathcal{B} è un evento di \mathcal{S}_2 , più tutte le unioni ed intersezioni di tali prodotti cartesiani. In questo nuovo esperimento, le probabilità degli eventi $\mathcal{A} \times \mathcal{S}_2$ e $\mathcal{S}_1 \times \mathcal{B}$ sono tali per cui

$$P(\mathcal{A} \times \mathcal{S}_2) = P_1(\mathcal{A}) \quad P(\mathcal{S}_1 \times \mathcal{B}) = P_2(\mathcal{B}) \quad (5.13)$$

dove $P_1(\mathcal{A})$ è la probabilità dell'evento \mathcal{A} nell'esperimento \mathcal{S}_1 e $P_2(\mathcal{B})$ è la probabilità dell'evento \mathcal{B} nell'esperimento \mathcal{S}_2 . La 5.13 è motivata dalla interpretazione di \mathcal{S} come esperimento congiunto: l'evento $\mathcal{A} \times \mathcal{S}_2$ dell'esperimento \mathcal{S} si verifica se si verifica l'evento \mathcal{A} dell'esperimento \mathcal{S}_1 , qualsiasi sia il risultato dell'esperimento \mathcal{S}_2 . Lo stesso ragionamento vale per $\mathcal{S}_1 \times \mathcal{B}$. Questo giustifica le due equazioni 5.13.

Le probabilità degli eventi $\mathcal{A} \times \mathcal{B}$ e delle loro unioni ed intersezioni non possono tuttavia essere espresse, in generale, in termini di P_1 e P_2 : servono informazioni aggiuntive sugli esperimenti \mathcal{S}_1 e \mathcal{S}_2 .

Esperimenti congiunti indipendenti In molti casi, gli eventi $\mathcal{A} \times \mathcal{S}_2$ e $\mathcal{S}_1 \times \mathcal{B}$ dell'esperimento congiunto \mathcal{S} sono indipendenti per ogni \mathcal{A} e \mathcal{B} . In questo caso possiamo concludere (da 5.10 e 5.13) che

$$P(\mathcal{A} \times \mathcal{B}) = P(\mathcal{A} \times \mathcal{S}_2) P(\mathcal{S}_1 \times \mathcal{B}) = P_1(\mathcal{A}) P_2(\mathcal{B}) \quad (5.14)$$

Questo completa la definizione dell'esperimento congiunto \mathcal{S} , poiché tutti i suoi eventi sono unioni ed intersezioni di eventi della forma $\mathcal{A} \times \mathcal{B}$.

Notiamo in particolare che l'evento elementare di \mathcal{S} $\{\xi_1\xi_2\}$ può essere scritto come prodotto cartesiano $\{\xi_1\} \times \{\xi_2\}$ degli eventi elementari $\{\xi_1\}$ e $\{\xi_2\}$ degli esperimenti \mathcal{S}_1 e \mathcal{S}_2 e quindi

$$P(\{\xi_1\xi_2\}) = P_1(\{\xi_1\}) P_2(\{\xi_2\})$$

Esempio 5.5 La scatola B_1 contiene 10 palle bianche e 5 rosse. La scatola B_2 contiene 20 palle bianche e 20 rosse. Si estrae una palla da ciascuna scatola: quale è la probabilità che la palla da B_1 sia bianca e quella da B_2 sia rossa?

Quello descritto può essere considerato un esperimento congiunto. Assumendo che i due esperimenti che lo formano siano indipendenti la probabilità cercata vale $\frac{10}{15} \cdot \frac{20}{40} = \frac{1}{3}$.

Esempio 5.6 Consideriamo l'esperimento del lancio della moneta. Sia p la probabilità di ottenere "testa" e $q = 1 - p$ quella di ottenere "croce". Lanciamo la moneta due volte: otteniamo lo spazio degli eventi

$$\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \quad \mathcal{S}_1 = \mathcal{S}_2 = \{T, C\}$$

quindi $\mathcal{S} = \{TT, TC, CT, CC\}$. Assumendo che i due esperimenti \mathcal{S}_1 e \mathcal{S}_2 siano indipendenti, possiamo ottenere le probabilità di tutti gli eventi elementari di \mathcal{S} :

$$P(TT) = P_1(T)P_2(T) = p^2 \quad P(TC) = P_1(T)P_2(C) = pq$$

$$P(CC) = P_1(C)P_2(C) = q^2 \quad P(CT) = P_1(C)P_2(T) = qp$$

Da queste vogliamo ottenere la probabilità dell'evento $\mathcal{T}_1 = \{\text{testa al primo lancio}\} = \{TT, TC\}$. Poiché \mathcal{T}_1 consiste dei due risultati $\{TT\}$ e $\{TC\}$, dalla 5.5 si ha:

$$P(\mathcal{T}_1) = P(TT) + P(TC) = p^2 + pq = p$$

Lo stesso valore si trova ovviamente anche dalla 5.14 essendo $\mathcal{T}_1 = \{T\} \times \mathcal{S}_2$.

5.8.3 Prodotto cartesiano di N esperimenti

Dati n esperimenti $\mathcal{S}_1, \dots, \mathcal{S}_n$, definiamo come loro prodotto cartesiano

$$\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_n \tag{5.15}$$

l'esperimento che ha come elementi (eventi elementari) le *ennupple* $\xi_1 \dots \xi_n$ dove ξ_i è un elemento dell'insieme \mathcal{S}_i . Gli eventi di questo spazio \mathcal{S} sono tutti gli insiemi della forma

$$\mathcal{A}_1 \times \dots \times \mathcal{A}_n$$

con $\mathcal{A}_i \subset \mathcal{S}_i$, più tutte le loro unioni ed intersezioni. Se gli esperimenti sono indipendenti e $P_i(\mathcal{A}_i)$ è la probabilità dell'evento \mathcal{A}_i nell'esperimento \mathcal{S}_i , allora

$$P(\mathcal{A}_1 \times \dots \times \mathcal{A}_n) = P_1(\mathcal{A}_1)P_2(\mathcal{A}_2) \cdots P_n(\mathcal{A}_n) \tag{5.16}$$

Esempio 5.7 Lanciamo una moneta n volte. Sia p la probabilità di ottenere "testa" e $q = 1 - p$ quella di ottenere "croce" in ogni singolo lancio. Lo spazio degli eventi è $\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_n$, formato dai 2^n elementi $\xi_1 \dots \xi_n$ tali che $\xi_i = T \circ C$. Chiaramente:

$$P\{\xi_1 \dots \xi_n\} = P_1\{\xi_1\} \cdots P_n\{\xi_n\} \quad P_i\{\xi_i\} = \begin{cases} p & \text{se } \xi_i = T \\ q & \text{se } \xi_i = C \end{cases} \tag{5.17}$$

Se, in particolare, $p = q = 1/2$, allora $P\{\xi_1 \dots \xi_n\} = 1/2^n$.

Se l'evento elementare $\{\xi_1 \dots \xi_n\}$ consiste di k teste e $n - k$ croci in ordine specificato la sua probabilità vale (vedi eq. 5.17):

$$P\{\xi_1 \dots \xi_n\} = p^k q^{n-k}$$

L'evento $\mathcal{T}_1 = \{\text{testa al primo lancio}\}$ consiste dei 2^{n-1} eventi elementari $\{\xi_1 \dots \xi_n\}$ aventi $\xi_1 = T$ e $\xi_i = T \circ C$ per $i > 0$. L'evento \mathcal{T}_1 può allora essere scritto come prodotto cartesiano

$$\mathcal{T}_1 = \{T\} \times \mathcal{S}_2 \times \dots \times \mathcal{S}_n$$

avente probabilità (vedi eq. 5.16):

$$P(\mathcal{T}_1) = P_1(T)P_2(\mathcal{S}_2) \cdots P_n(\mathcal{S}_n) = P_1(T) = p$$

Analogamente si trovano le probabilità di ottenere testa all'i-esimo lancio (p) o croce al k-esimo lancio (q).

5.9 Prove ripetute

In teoria della probabilità le prove ripetute possono essere considerate in due modi diversi. Il primo è il legame approssimato fra probabilità $P(\mathcal{A})$ di un evento \mathcal{A} in un esperimento \mathcal{S} e la frequenza relativa di occorrenza di \mathcal{A}

$$P(\mathcal{A}) = \frac{n_{\mathcal{A}}}{n}$$

Il secondo è la creazione dell'esperimento $\mathcal{S} \times \cdots \times \mathcal{S}$.

Ad esempio, il lancio ripetuto di una moneta può essere interpretato nei seguenti due modi:

Interpretazione 1 (fisica) L'esperimento è il lancio *singolo* di una moneta (affidabile). Lo spazio degli eventi ha due elementi e la probabilità di ogni evento elementare è pari a $1/2$. Una prova è il lancio della moneta una volta.

Se noi lanciamo la moneta n volte e "testa" compare n_T volte, allora, quasi certamente, $n_T/n \simeq 1/2$ purchè n sia abbastanza grande. Quindi la prima interpretazione delle prove ripetute è la relazione imprecisa che lega probabilità e frequenze relative osservate.

Interpretazione 2 (concettuale) L'esperimento è ora il lancio della moneta n volte, dove n è un qualsiasi numero. Lo spazio degli eventi ha 2^n elementi e la probabilità di ciascun elemento (o evento elementare) è pari a $1/2^n$. Una prova è il lancio della moneta n volte. Tutti gli eventi hanno probabilità esatte definite a priori. E' possibile dare una interpretazione in termini di frequenza relativa a queste probabilità, ma per fare questo bisogna ripetere *il lancio della moneta n volte* per un grande numero di volte.

5.9.1 Prove di Bernoulli

E' noto che, se un insieme ha n elementi, allora il numero totale di sottoinsiemi ciascuno dei quali è composto da k elementi è pari a

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{1 \cdot 2 \cdots k} = \frac{n!}{k!(n-k)!} \quad (5.18)$$

Mediante tale relazione vogliamo trovare la probabilità che un evento \mathcal{A} si verifichi k volte in n prove indipendenti dello stesso esperimento \mathcal{S} . Se

$$P(\mathcal{A}) = p \quad P(\overline{\mathcal{A}}) = q \quad p + q = 1$$

la probabilità cercata, detta $p_n(k)$, vale

$$p_n(k) = P\{\mathcal{A} \text{ si verifica } k \text{ volte in qualsiasi ordine}\} = \binom{n}{k} p^k q^{n-k} \quad (5.19)$$

Questo risultato è noto come *teorema fondamentale*. La sua prova è semplice: l'evento $\{\mathcal{A} \text{ si verifica } k \text{ volte su } n \text{ tentativi in uno specifico ordine}\}$ è il prodotto cartesiano $\mathcal{B}_1 \times \cdots \times \mathcal{B}_n$ dove k degli eventi \mathcal{B}_i sono \mathcal{A} ed $(n-k)$ sono $\overline{\mathcal{A}}$. La probabilità di questo evento, vedi 5.16, vale

$$P(\mathcal{B}_1) \cdots P(\mathcal{B}_n) = p^k q^{n-k}$$

L'evento $\{\mathcal{A} \text{ si verifica } k \text{ volte su } n \text{ tentativi in un qualsiasi ordine}\}$ è l'unione degli $\binom{n}{k}$ eventi $\{\mathcal{A} \text{ si verifica } k \text{ volte su } n \text{ tentativi in uno specifico ordine}\}$ e siccome tali eventi sono mutuamente esclusivi concludiamo, in accordo con la 5.3, che è vera la 5.19.

Esempio 5.8 Un dado affidabile viene lanciato 5 volte. Quale è la probabilità $p_5(2)$ che il "sei" appaia due volte?

Nel lancio singolo del dado la probabilità di ottenere "sei" è pari a $1/6$. Quindi basta inserire nella 5.19 i valori: $p = 1/6$, $q = 5/6$, $n = 5$, $k = 2$ per ottenere la probabilità cercata:

$$p_5(2) = \binom{5}{2} p^2 q^3 = \frac{5!}{2!3!} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 \approx 0.161$$

Esempio 5.9 Una coppia di dadi affidabili viene lanciata quattro volte. Quale è la probabilità $p_4(0)$ che il "sette" non compaia mai?

Lo spazio degli eventi del lancio di due dadi è formato dai 36 elementi $f_i f_j$. L'evento $\mathcal{A} = \{\text{sette}\}$ è l'insieme formato dai 6 elementi

$$f_1 f_6, \quad f_2 f_5, \quad f_3 f_4, \quad f_4 f_3, \quad f_5 f_2, \quad f_6 f_1$$

Quindi $P(\mathcal{A}) = 6/36 = 1/6$, $P(\overline{\mathcal{A}}) = 5/6$. Con $n = 4$ e $k = 0$ la 5.19 porta a

$$p_4(0) = \binom{4}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^4 = \left(\frac{5}{6}\right)^4 \approx 0.482$$

Esempio 5.10 Piazziamo a caso n punti nell'intervallo $[0, T]$. Quale è la probabilità che k di questi punti siano nell'intervallo (t_1, t_2) ?

Questo può essere considerato un esempio di problema di prove ripetute. L'esperimento \mathcal{S} è il piazzamento di un singolo punto nell'intervallo $[0, T]$. In tale esperimento l'evento \mathcal{A} è l'evento che si verifica se il punto piazzato cade nell'intervallo (t_1, t_2) ; la sua probabilità è

$$P(\mathcal{A}) = p = \frac{t_2 - t_1}{T} \quad \text{e} \quad P(\overline{\mathcal{A}}) = q = 1 - p$$

Quindi

$$P\{k \text{ punti cadono in } (t_1, t_2)\} = \binom{n}{k} p^k q^{n-k} \quad (5.20)$$

5.9.2 Analisi di $p_n(k) = \binom{n}{k} p^k q^{n-k}$

Esaminiamo ora l'andamento di $p_n(k)$ come funzione di k per n fissato. Vediamo due esempi: in entrambi $n = 9$, nel primo $p = q = 1/2$, nel secondo $p = 1/3$ e $q = 2/3$. I valori che si ottengono sono mostrati in tabella e in fig. 5.1.

E' evidente che $p_9(k)$ cresce fino ad un massimo e quindi decresce. Questo andamento vale per qualsiasi scelta di n , p e q , anche se, in generale, non si ottengono valori simmetrici. Si può dimostrare (ma non lo faremo qui) che il massimo della generica funzione $p_n(k)$ si ottiene per

$$k = k_{max} = [(n+1)p]$$

dove la parentesi quadra indica il massimo intero che non supera $(n+1)p$. Se $(n+1)p$ è un intero, allora il massimo si ottiene per *due* valori di k : $k_1 = (n+1)p$ e $k_2 = k_1 - 1$ (vedi tabella).

k	$\binom{9}{k}$	$p_9(k) \ p = \frac{1}{2}$	$p_9(k) \ p = \frac{1}{3}$
0	1	0.001953	0.026012
1	9	0.017578	0.117055
2	36	0.070313	0.234111
3	84	0.164063	0.273129
4	126	0.246094	0.204847
5	126	0.246094	0.102423
6	84	0.164063	0.034141
7	36	0.070313	0.007316
8	9	0.017578	0.000914
9	1	0.001953	0.000051

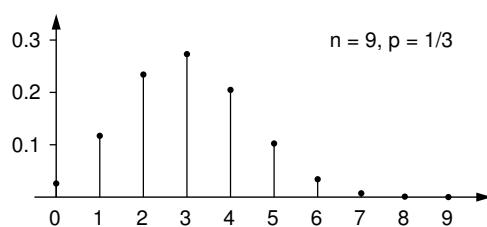
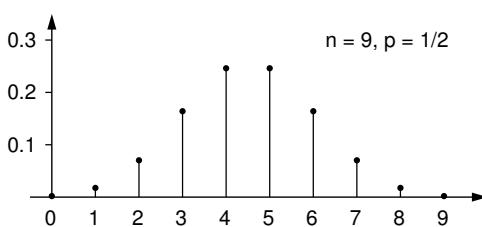


Figura 5.1: $p_n(k)$ con $n = 9$ e p come indicato.

NOTA: il primo esempio ($n = 9$, $p = q = 1/2$) corrisponde al lancio per nove volte di una moneta affidabile. Il numero di combinazioni possibili è $2^9 = 512$, e qualsiasi combinazione di teste e croci ha la stessa probabilità, ovvero $1/512 \approx 0.001953$. Perchè consideriamo "speciali" le combinazioni di 9 teste o di 9 croci? Perchè si verificano solo una volta (in media) su 512 lanci. Al contrario le combinazioni con 4 o 5 teste si verificano con probabilità di $(126 + 126)/512 = 252/512$, ovvero circa la metà delle volte.

k compreso fra k_1 e k_2

La probabilità che il numero k di occorrenze dell'evento \mathcal{A} in n prove sia compreso fra k_1 e k_2 è:

$$P\{k_1 \leq k \leq k_2\} = \sum_{k=k_1}^{k_2} p_n(k) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{(n-k)} \quad (5.21)$$

Questo risultato deriva dal fatto che gli eventi { \mathcal{A} si verifica k volte } per k che assume i valori da k_1 a k_2 sono mutuamente esclusivi e che la loro unione forma l'evento { $k_1 \leq k \leq k_2$ }.

Esempio 5.11 Supponiamo di aver acquistato 10^4 componenti. La probabilità che un componente sia difettoso è pari a $p = 0.1$. Quale è la probabilità che il numero totale di componenti difettosi sia inferiore o uguale a 1100?

$$P\{0 \leq k \leq 1100\} = \sum_{k=0}^{1100} \binom{10^4}{k} (0.1)^k (0.9)^{(10^4-k)}$$

Questo risultato è ovviamente corretto, ma non facile da calcolare. Considerate che $(10^4)! \approx 2.8 \cdot 10^{35659}$, un numero che scritto per esteso richiederebbe 35659 + 1 cifre, ovvero circa dodice pagine di queste dispense.² Considerando che è probabilmente sufficiente avere da due a quattro cifre significative sul risultato, sarebbe utile trovare un metodo approssimato per il calcolo...

5.10 Formula di approssimazione di Gauss

Le equazioni 5.19 e 5.21 permettono il calcolo delle probabilità degli eventi corrispondenti in modo esatto, ma non sono sempre pratiche da calcolare (vedi esempio 5.11). In questa sezione troveremo due semplici formule approssimate per valutare queste probabilità.

Gaussiana In teoria delle probabilità la funzione gaussiana (o normale), definita da

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (5.22)$$

gioca un ruolo molto importante come il suo integrale

$$G(x) = \int_{-\infty}^x g(y) dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy \quad (5.23)$$

Come noto

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

da cui segue che

$$G(\infty) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$$

Inoltre, poiché $g(-x) = g(x)$ segue che $G(-x) = 1 - G(x)$.

L'integrale in 5.23 non è risolvibile in forma chiusa. Introduciamo la funzione

$$Q(k) = 1 - G(k) = \frac{1}{\sqrt{2\pi}} \int_k^{\infty} e^{-\frac{\lambda^2}{2}} d\lambda \quad (5.24)$$

I valori di Q si trovano su numerosi testi, compreso il [Car86]. Per grandi valori di k la funzione Q può essere approssimata mediante

$$Q(k) \approx \frac{1}{k\sqrt{2\pi}} e^{-\frac{k^2}{2}} \quad (5.25)$$

che fornisce risultati accurati per $k > 3$.

I alcuni testi invece della funzione Q potete trovare la funzione erf, definita da

$$\text{erf}(k) = G(k) - \frac{1}{2} = \frac{1}{\sqrt{2\pi}} \int_0^k e^{-\frac{\lambda^2}{2}} d\lambda = \frac{1}{2} - Q(k) \quad (5.26)$$

²Sulla mia calcolatrice scientifica il massimo numero fattoriale calcolabile è $69! \approx 1.7112 \cdot 10^{98}$, da 70! in poi da errore.

In generale si può definire come gaussiana di media m e varianza σ^2 la funzione

$$g_{m\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

da cui

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_1} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = G\left(\frac{x_1-m}{\sigma}\right) = 1 - Q\left(\frac{x_1-m}{\sigma}\right)$$

5.10.1 Teorema di DeMoivre–Laplace

Si può dimostrare che, se $npq \gg 1$,

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}} \quad (5.27)$$

Questo risultato è noto come Teorema di DeMoivre–Laplace o formula di approssimazione di Gauss e può essere espresso come uguaglianza al limite: il rapporto dei due lati tende a 1 quando $n \rightarrow \infty$. La dimostrazione, che non viene qui riportata, è basata sulla *formula di Stirling*

$$n! \approx n^n e^{-n} \sqrt{2\pi n} \quad (5.28)$$

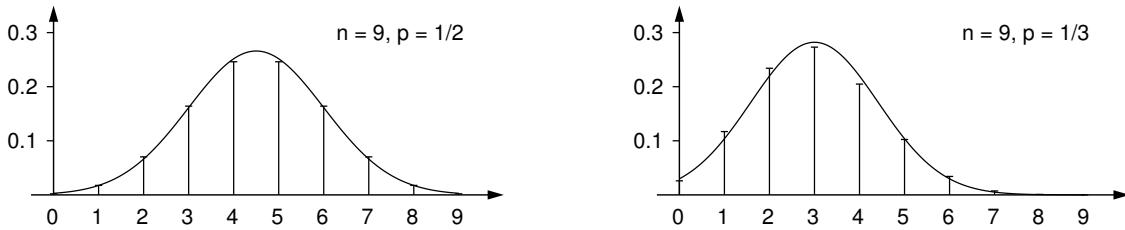


Figura 5.2: Formula di approssimazione di Gauss per i casi indicati.

In figura 5.2 si possono osservare i risultati della formula di approssimazione di Gauss nei due casi già trattati. L'approssimazione è visibile in quanto npq non è molto maggiore di uno: vale 2.25 per $p = 1/2$ e 2 per $p = 1/3$.

5.10.2 Valutazione approssimata di $P\{k_1 \leq k \leq k_2\}$

Usando l'approssimazione 5.27 si può dimostrare che

$$P\{k_1 \leq k \leq k_2\} = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{n-k} \approx G\left(\frac{k_2-np}{\sqrt{npq}}\right) - G\left(\frac{k_1-np}{\sqrt{npq}}\right) \quad (5.29)$$

L'approssimazione è accurata se $npq \gg 1$ e le differenze $k_1 - np$ e $k_2 - np$ sono dell'ordine di \sqrt{npq}

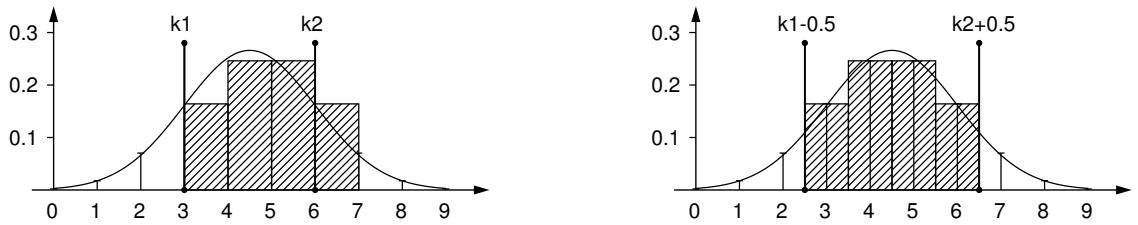
Dimostrazione: inserendo la 5.27 nella 5.21 e chiamando $\sigma^2 = npq$ otteniamo

$$\sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=k_1}^{k_2} e^{-(k-np)^2/2\sigma^2} \quad (5.30)$$

Le gaussiane all'interno della sommatoria sono quasi costanti in ogni intervallo di lunghezza 1 in quanto, per assunzione, $\sigma^2 = npq \gg 1$; quindi l'area in tali intervalli è circa uguale alla sua ordinata (vedi fig. 5.3, a sinistra). Da ciò segue che la parte destra della 5.30 può essere approssimata dall'integrale della gaussiana nell'intervallo (k_1, k_2)

$$\frac{1}{\sigma\sqrt{2\pi}} \sum_{k=k_1}^{k_2} e^{-(k-np)^2/2\sigma^2} \approx \frac{1}{\sigma\sqrt{2\pi}} \int_{k_1}^{k_2} e^{-(x-np)^2/2\sigma^2} dx \quad (5.31)$$

da cui segue la 5.29.

Figura 5.3: Integrali per la valutazione approssimata di $P\{k_1 \leq k \leq k_2\}$.

Correzione dell'errore La sommatoria nella parte a sinistra della 5.30 consiste di $k_2 - k_1 + 1$ termini, mentre l'integrale in 5.31 approssima solo $k_2 - k_1$ rettangoli. Se $k_2 - k_1 \gg 1$ l'errore risultante può essere trascurato. In caso contrario è meglio usare una approssimazione più accurata, allargando i limiti di integrazione di 0.5 per parte (vedi fig. 5.3, a destra), ottenendo la formula di approssimazione

$$P\{k_1 \leq k \leq k_2\} = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{(n-k)} \approx \mathbf{G}\left(\frac{k_2 + 0.5 - np}{\sqrt{npq}}\right) - \mathbf{G}\left(\frac{k_1 - 0.5 - np}{\sqrt{npq}}\right) \quad (5.32)$$

Esercizio 5.12 Una moneta affidabile viene lanciate 10000 volte. Quale è la probabilità che il numero di teste sia fra 4900 e 5100 ?

Soluzione: per ipotesi $n = 10000$, $p = q = 0.5$, $np = 5000$, $npq = 2500$, $k_1 = 4900$, $k_2 = 5100$. Quindi, usando la 5.29 ne concludiamo che la probabilità cercata è circa

$$\mathbf{G}(2) - \mathbf{G}(-2) = 2\mathbf{G}(2) - 1 \approx 0.9545$$

In appendice A (di questo capitolo) trovate quanto serve per valutare **G**, **erf** e **Q**. Gli studenti sono tenuti a saper usare tabelle e grafici per il calcolo delle probabilità e a saper effettuare l'interpolazione lineare per i valori intermedi.

Calcolo del risultato dell'esempio 5.11 Si tratta di calcolare

$$P\{0 \leq k \leq 1100\} = \sum_{k=0}^{1100} \binom{10^4}{k} (0.1)^k (0.9)^{(10^4-k)}$$

Poichè $npq = 10^4 \cdot 0.1 \cdot 0.9 = 900 \gg 1$ possiamo usare la formula di approssimazione di Gauss

$$P\{0 \leq k \leq 1100\} \approx \mathbf{G}\left(\frac{1100 - 1000}{30}\right) - \mathbf{G}\left(\frac{0 - 1000}{30}\right) = \mathbf{G}\left(\frac{100}{30}\right) - \mathbf{G}\left(\frac{-1000}{30}\right)$$

e poichè $\mathbf{G}(-x) = 1 - \mathbf{G}(x)$

$$P\{0 \leq k \leq 1100\} \approx \mathbf{G}(3.33333) - 1 + \mathbf{G}(33.33333)$$

Entrambi i valori sono maggiori di 3, quindi si approssimano secondo

$$\mathbf{G}(x) \approx 1 - \frac{1}{x} g(x)$$

$$\mathbf{G}(3.33333) \approx 1 - \frac{1}{3.33333} \frac{1}{\sqrt{2\pi}} e^{-\frac{3.33333^2}{2}} \approx 0.99954$$

$$\mathbf{G}(33.33333) \approx 1 - \frac{1}{33.33333} \frac{1}{\sqrt{2\pi}} e^{-\frac{33.33333^2}{2}} \approx 1$$

e

$$P\{0 \leq k \leq 1100\} \approx 0.99954$$

Esempio 5.13 Come interpolare linearmente i valori di $\text{erf}()$ in tabella. Supponiamo di dover calcolare

$$\mathbf{G}(0.33333) = \frac{1}{2} + \text{erf}(0.33333)$$

Dalla tabella in Appendice sappiamo che

$$\text{erf}(0.30) = 0.11791 \quad \text{erf}(0.35) = 0.13683$$

Il valore è compreso fra i due e si può approssimare mediante interpolazione lineare:

$$\text{erf}(0.33333) \approx 0.11791 + \frac{(0.33333 - 0.3) \cdot (0.13683 - 0.11791)}{0.35 - 0.3} \approx 0.13052$$

quindi

$$\mathbf{G}(0.33333) = \frac{1}{2} + \text{erf}(0.33333) \approx 0.63052$$

5.11 La legge dei grandi numeri

In accordo con l'interpretazione delle probabilità in termini di frequenze relative, se un evento \mathcal{A} con $P(\mathcal{A}) = p$ si verifica k volte in n prove, allora $k \approx np$. E' ora di riformulare questa affermazione euristica in termini rigorosi (teorema al limite).

Osserviamo intanto che $k \approx np$ non significa che k sarà vicino a np . Infatti (vedi 5.27):

$$P\{k = np\} = \binom{n}{np} p^{np} q^{n-np} \approx \frac{1}{\sqrt{2\pi npq}} \rightarrow 0 \quad \text{per } n \rightarrow \infty \quad (5.33)$$

L'approssimazione $k \approx np$ significa che il rapporto k/n è vicino a p . In termini rigorosi si enuncia come segue:

per ogni $\varepsilon > 0$ la probabilità che $\left| \frac{k}{n} - p \right| < \varepsilon$ tende a 1 per $n \rightarrow \infty$.

La semplice dimostrazione di questo teorema, noto come legge dei grandi numeri, è omessa.

5.12 Teorema di Poisson

Abbiamo visto in (5.19) che la probabilità che l'evento \mathcal{A} si verifichi k volte su n tentativi in qualsiasi ordine vale

$$\binom{n}{k} p^k q^{n-k} \quad (5.34)$$

Forniamo una formula di approssimazione per il calcolo di questa probabilità valida se $p \ll 1$. Come visto, se n è così grande che $np \approx npq \gg 1$ possiamo usare l'approssimazione di DeMoivre-Laplace (5.27). Se invece np è dell'ordine di 1, l'eq. 5.27 non è più applicabile. In questi casi si può usare la seguente approssimazione: per k nell'ordine di np ,

$$\binom{n}{k} p^k q^{n-k} \approx e^{-np} \frac{(np)^k}{k!} \quad (5.35)$$

Infatti, se k è dell'ordine di np , allora $k \ll n$ e $kp \ll 1$. Quindi

$$\begin{aligned} n(n-1) \cdots (n-k+1) &\approx n \cdot n \cdots n = n^k \\ q &= 1-p \approx e^{-p} \quad q^{n-k} \approx e^{-(n-k)p} \approx e^{-np} \end{aligned}$$

e segue la 5.35. Questa approssimazione, portata al limite, diventa il **teorema di Poisson**: se

$$n \rightarrow \infty \quad p \rightarrow 0 \quad np \rightarrow a$$

allora

$$\lim_{n \rightarrow \infty} \frac{n!}{k!(n-k)!} p^k q^{n-k} = e^{-a} \frac{a^k}{k!} \quad (5.36)$$

Applicando infine la 5.35 è possibile ottenere l'approssimazione

$$P\{k_1 \leq k \leq k_2\} \approx e^{-np} \sum_{k=k_1}^{k_2} \frac{(np)^k}{k!} \quad (5.37)$$

5.13 Punti casuali di Poisson

Una importante applicazione del teorema di Poisson è la valutazione approssimata della 5.20 quando T ed n tendono all'infinito. Ripetiamo il problema. Piazziamo a caso n punti nell'intervallo $(-T/2, T/2)$ e chiamiamo $P\{k \text{ in } t_a\}$ la probabilità che k di questi punti giacciono in un intervallo (t_1, t_2) di lunghezza $t_2 - t_1 = t_a$. Come abbiamo visto in 5.20

$$P\{k \text{ in } t_a\} = \binom{n}{k} p^k q^{n-k} \quad \text{dove} \quad p = \frac{t_a}{T} \quad (5.38)$$

Assumiamo ora che $n \gg 1$ e $t_a \ll T$. Applicando la 5.35 concludiamo che

$$P\{k \text{ in } t_a\} \approx e^{-nt_a/T} \frac{(nt_a/T)^k}{k!} \quad (5.39)$$

per k nell'ordine di nt_a/T . Supponiamo ora che n e T crescano senza limiti, ma che il loro rapporto

$$\lambda = \frac{n}{T}$$

resti costante (λ è ovviamente la densità di punti per unità di lunghezza). Il risultato è un insieme infinito di punti che coprono l'asse t . L'equazione 5.38 non è applicabile, ma la 5.39 lo è:

$$P\{k \text{ in } t_a\} = e^{-\lambda t_a} \frac{(\lambda t_a)^k}{k!} \quad (5.40)$$

5.14 APPENDICE A: valori di erf, G, Q

La funzione gaussiana $g(x)$ è definita da

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

I suoi integrali sono le funzioni **G**, **Q** e **erf**, definite da

$$\begin{aligned} \mathbf{G}(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy \\ \mathbf{Q}(x) &= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{y^2}{2}} dy \\ \mathbf{erf}(x) &= \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{y^2}{2}} dy \end{aligned}$$

Ovviamente

$$\mathbf{Q}(x) = 1 - \mathbf{G}(x) \quad \mathbf{erf}(x) = \mathbf{G}(x) - \frac{1}{2} \quad \mathbf{Q}(x) = \frac{1}{2} - \mathbf{erf}(x)$$

$$g(-x) = g(x) \quad \mathbf{G}(-x) = 1 - \mathbf{G}(x) \quad \mathbf{G}(x) = \frac{1}{2} + \mathbf{erf}(x)$$

I valori di $\text{erf}(x)$ per $0 < x \leq 3$ sono in tabella (copiati da [Pap91]). Per grandi valori di x vale l'approssimazione

$$\mathbf{G}(x) \approx 1 - \frac{1}{x} g(x)$$

x	erf(x)	x	erf(x)	x	erf(x)	x	erf(x)
0.05	0.01994	0.80	0.28814	1.55	0.43943	2.30	0.48928
0.10	0.03983	0.85	0.30234	1.60	0.44520	2.35	0.49061
0.15	0.05962	0.90	0.31594	1.65	0.45053	2.40	0.49180
0.20	0.07926	0.95	0.32894	1.70	0.45543	2.45	0.49286
0.25	0.09871	1.00	0.34134	1.75	0.45994	2.50	0.49379
0.30	0.11791	1.05	0.35314	1.80	0.46407	2.55	0.49461
0.35	0.13683	1.10	0.36433	1.85	0.46784	2.60	0.49534
0.40	0.15542	1.15	0.37493	1.90	0.47128	2.65	0.49597
0.45	0.17364	1.20	0.38493	1.95	0.47441	2.70	0.49653
0.50	0.19146	1.25	0.39435	2.00	0.47726	2.75	0.49702
0.55	0.20884	1.30	0.40320	2.05	0.47982	2.80	0.49744
0.60	0.22575	1.35	0.41149	2.10	0.48214	2.85	0.49781
0.65	0.24215	1.40	0.41924	2.15	0.48422	2.90	0.49813
0.70	0.25804	1.45	0.42647	2.20	0.48610	2.95	0.49841
0.75	0.27337	1.50	0.43319	2.25	0.48778	3.00	0.49865

Funzione Q

La funzione $Q(k)$ è definita come

$$Q(k) = 1 - G(k) = \frac{1}{\sqrt{2\pi}} \int_k^\infty e^{-\frac{\lambda^2}{2}} d\lambda$$

I suoi valori per $0 \leq k \leq 7$ sono in figura (presi da [Car86]) : per valori maggiori può essere approssimata da

$$Q(k) \approx \frac{1}{k\sqrt{2\pi}} e^{-\frac{k^2}{2}}$$

che è accurata per $k > 3$.

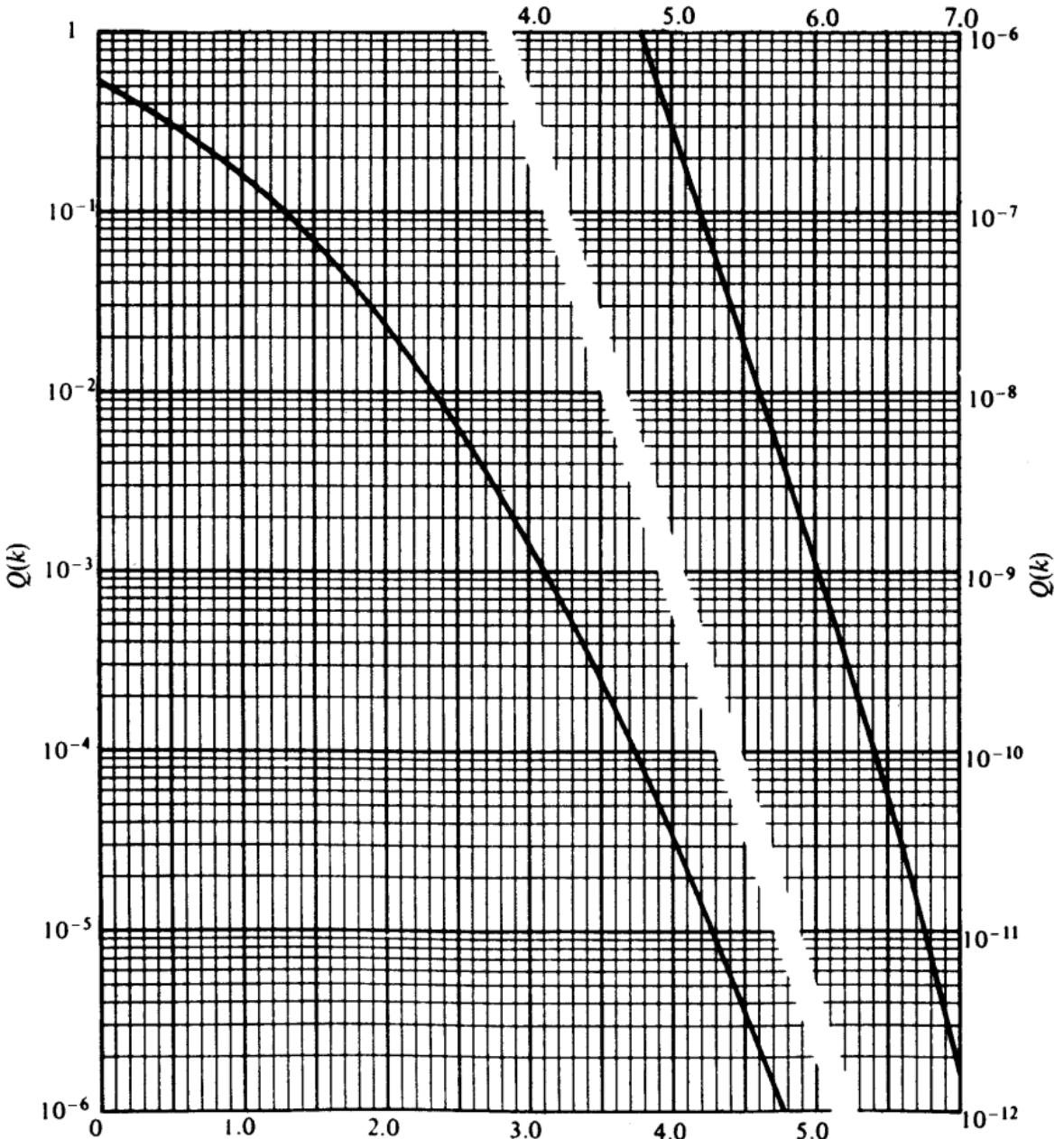


Figura 5.4: Valori della funzione $Q(k)$ per $0 \leq k \leq 7$.

5.15 APPENDICE B: il gioco del lotto

Dal sito www.lottomatica.it:

Nato a Genova nel 1500, il Lotto è un gioco di antica tradizione e grande popolarità. Il Lotto consiste nell'estrarre 5 numeri da un'urna che ne contiene 90 presso ognuna delle 10 ruote d'estrazione. Le estrazioni avvengono due volte alla settimana: il mercoledì e il sabato. Si possono giocare fino ad un massimo di 10 numeri su una o più ruote secondo le seguenti combinazioni: estratto, ambo, terno, quaterna, cinquina.

Le vincite sono così distribuite:

cinquina = 1.000.000 volte la posta

quaterna = 80.000 volte la posta

terno = 4.250 volte la posta

ambo = 250 volte la posta

estratto = 11.232 volte la posta

Un gioco viene considerato equo quando le quote pagate sono proporzionate alla frequenza di vincita. Se la combinazione favorevole esce in media una volta su dieci la vincita di dieci volte la posta è equa: in questo modo se giocate abbastanza a lungo andate in pari. Ovviamente nessun gioco d'azzardo organizzato paga quote eque, perché il margine di guadagno sarebbe nullo. Un esempio può essere il gioco della roulette: la pallina si può fermare in 37 diverse caselle, numerate da zero a trentasei. Se azzeccate il numero vi pagano 36 volte la posta. Quindi pagano 36 volte la posta per una probabilità di 1 a 37. Ciò è sufficiente ai casinò per guadagnare, e non è particolarmente iniquo.

Vediamo il gioco del lotto: la cinquina viene pagata 1.000.000 di volte la posta, ma la sua probabilità di uscita è pari a 1 volta su 44.000.000 circa ($1/43949268$). La probabilità di uscita di una quaterna è di $1/511.038$, quella di un terno $1/11.748$, quella di un ambo $1/400$ e quella di un estratto $1/18$ (confrontatele con le quote).

Nel 1999 il gioco del lotto ha reso circa 20.000 miliardi di vecchie lire alle casse statali: supponendo che i giocatori in Italia siano 20 milioni ciascuno di essi ha perso, in media, un milione di lire. I dati esatti, ricavati dal sito del Ministero dell'Economia e delle Finanze, agenzia delle entrate (www.agenziaentrate.it/documentazione/studi/entrate) danno le seguenti entrate, al netto delle vincite: anno 1996 6.267 Ml (miliardi di lire), 1997 8.784 Ml, 1998 12.309 Ml, 1999 19.577 Ml: dati più recenti non sono disponibili. Notizie stampa stimano tuttavia in circa 12 miliardi di euro le entrate tributarie dirette e indirette ottenute con lo sfruttamento statale del gioco d'azzardo per l'anno 2014. *Da "repubblica.it" del 6 aprile 2020: nel 2019 gli italiani hanno speso per lotto, slot-machine e Gratta & Vinci 110 miliardi di euro, 1.833 a testa, neonati compresi. Un passatempo occasionale per tanti, una mania border-line per il milione di persone catalogate dal Cnr a rischio moderato, un droga per le 400mila vittime secondo il Cnr di gioco problematico. Le vincite sono il 75% della somma spesa per slot & C.*

La discussione sui motivi per cui esistono giocatori del lotto non riguarda questo corso, mentre il calcolo delle probabilità di uscita delle varie combinazioni si.

Estratto. La probabilità che un numero venga estratto per primo è pari a $1/90$. La probabilità che venga estratto per secondo è $1/89$ per la probabilità che non sia già stato estratto come primo ($89/90$), ovvero ancora $(1/89)(89/90) = 1/90$. Con lo stesso ragionamento si trova che $1/90$ è la probabilità che venga estratto come terzo, quarto o quinto. Siccome gli eventi sono indipendenti la probabilità che un numero venga estratto è pari a $5(1/90) = 1/18$. Lo stesso risultato si ottiene considerando che vengono estratti 5 numeri su 90: la probabilità di indovinarne uno è $5/90$.

Ambo. Per il calcolo della probabilità conviene ragionare come segue: con 90 numeri non ripetuti si possono formare 4005 ambi diversi (vedi eq. 5.18), secondo

$$\binom{90}{2} = \frac{90!}{2!(90-2)!} = \frac{90 \cdot 89 \cdot 88!}{2 \cdot 88!} = \frac{90 \cdot 89}{2} = 4005$$

Ogni volta si estraggono 5 numeri, che formano $\binom{5}{2} = 10$ ambi diversi. Essendo tutti eventi equiprobabili ne risulta che la probabilità di indovinare un ambo è $10/4005 = 1/400.5$.

Terno. Con lo stesso ragionamento visto per l'ambo otteniamo che vi sono

$$\binom{90}{3} = \frac{90!}{3!(90-3)!} = 117480$$

terni possibili. Ad ogni estrazione ne escono $\binom{5}{3} = 10$ e la probabilità di indovinarne uno è 1/11748.

I calcoli per **quaterna** e **cinquina** sono analoghi.

5.16 APPENDICE C: simboli della teoria degli insiemi

Simboli in teoria degli insiemi		
Simbolo	Nome	Significato / esempio
\mathcal{N}	Insieme dei numeri naturali	$\mathcal{N} = \{0, 1, 2, 3, \dots\}$
\mathcal{Z}	Insieme dei numeri interi relativi	$\mathcal{Z} = \{\dots - 3, -2, -1, 0, 1, 2, 3, \dots\}$
\mathcal{Q}	Insieme dei numeri razionali	$1/4$ e $3/17$ sono elementi dell'insieme \mathcal{Q}
\mathcal{R}	Insieme dei numeri reali	$\pi, 1, \sqrt{2}, 5/3$ sono elementi dell'insieme \mathcal{R}
\mathcal{C}	Insieme dei numeri complessi	$2 + j0, -\pi - j\sqrt{2}, 3$ sono elementi dell'insieme \mathcal{C}
\emptyset	Insieme vuoto	indica un insieme privo di elementi
\mathcal{S}	Insieme che contiene tutti i possibili risultati	Spazio degli eventi o Evento certo
\in	Appartiene	$2 \in \mathcal{N}$
\notin	Non appartiene	$0.5 \notin \mathcal{N}$
\cup o $+$	Unione	Indica l'unione fra insiemi
\cap	Intersezione	$A \cap B$ o AB indica l'intersezione fra insiemi.
\subset	Sottoinsieme proprio	$A \subset B$ indica che A è sottoinsieme di B
\subseteq	Sottoinsieme improprio	$A \subseteq B$ indica che A è sottoinsieme di B o coincidente
\supset	Contiene	$A \supset B$ indica che A contiene B
\supseteq	Contiene (anche coincidenti)	$A \supseteq B$ indica che A contiene B o coincidono
$\bar{\mathcal{A}}$	Complemento	$\mathcal{S} - \mathcal{A} = \bar{\mathcal{A}}$

Capitolo 6

Variabili aleatorie

6.1 Definizione di variabile aleatoria

Una variabile aleatoria (abbreviata nel seguito v.a.) è una *regola* che associa un numero $\mathbf{x}(\zeta)$, solitamente reale, ad ogni possibile uscita ζ di un esperimento. In inglese la v.a. si chiama *random variable o RV*.

L'assegnazione può essere arbitraria, ma spesso il numero associato ha un significato: la tensione in un certo istante ai capi di un resistore rumoroso, l'altezza di una persona, il costo di un componente, la portata di un fiume.

I vantaggi delle v.a. sono molteplici: uno dei più importanti è quello di poter sviluppare una teoria generale che prescinda dal particolare esperimento che si considera.

- Esempio 6.1** (a) Nel lancio del dado assegnamo ad ognuna delle sei possibili uscite f_i il numero $\mathbf{x}(f_i) = 10i$. Quindi: $\mathbf{x}(f_1) = 10, \dots, \mathbf{x}(f_6) = 60$.
(b) Nello stesso esperimento assegnamo il numero 1 ad ogni uscita dispari ed il numero 0 ad ogni uscita pari. Quindi: $\mathbf{x}(f_1) = \mathbf{x}(f_3) = \mathbf{x}(f_5) = 1$ e $\mathbf{x}(f_2) = \mathbf{x}(f_4) = \mathbf{x}(f_6) = 0$.

Una v.a. è quindi una funzione il cui dominio è l'insieme \mathcal{S} delle uscite di un esperimento e i cui valori sono un insieme di numeri, purchè soddisfi alcune condizioni che vedremo presto. Il simbolo $\mathbf{x}(\zeta)$ indicherà il numero assegnato allo specifico evento ζ , mentre il simbolo \mathbf{x} indicherà la regola di corrispondenza fra ogni specifica uscita ζ ed il numero assegnato ad essa.

Eventi generati da una v.a. Nello studio delle v.a. sorgono domande del tipo: quale è la probabilità che la v.a. \mathbf{x} sia minore di un certo numero x oppure quale è la probabilità che la v.a. \mathbf{x} sia compresa fra due numeri x_1 e x_2 ? Per rispondere a queste domande dobbiamo essere in grado di esprimere le condizioni su \mathbf{x} in termini di eventi (che sono i soli per cui abbiamo definito il concetto di probabilità). Ad esempio, la notazione

$$\{\mathbf{x} \leq x\}$$

rappresenta il sottoinsieme di \mathcal{S} consistente nelle uscite ζ tali che $\mathbf{x}(\zeta) \leq x$. Quindi $\{\mathbf{x} \leq x\}$ non è un insieme di numeri, bensì un insieme di uscite dell'esperimento. Analogamente:

$\{x_1 \leq \mathbf{x} \leq x_2\}$ rappresenta il sottoinsieme di \mathcal{S} composto da tutte le uscite ζ tali che $x_1 \leq \mathbf{x}(\zeta) \leq x_2$ dove x_1 e x_2 sono due numeri dati;

$\{\mathbf{x} = x\}$ rappresenta il sottoinsieme di \mathcal{S} composto da tutte le uscite ζ tali che $\mathbf{x}(\zeta) = x$;
 $\{\mathbf{x} \in A\}$ dove A è un sottoinsieme di numeri sull'asse reale, rappresenta il sottoinsieme di \mathcal{S} composto da tutte le uscite ζ tali che $\mathbf{x}(\zeta) \in A$.

- Esempio 6.2** Consideriamo la v.a. definita nell'esempio 6.1 (a):

- l'insieme $\{\mathbf{x} \leq 35\}$ consiste degli elementi f_1, f_2 e f_3 ;
- l'insieme $\{\mathbf{x} \leq 5\}$ è vuoto;
- l'insieme $\{20 \leq \mathbf{x} \leq 35\}$ consiste degli elementi f_2 e f_3 ;
- l'insieme $\{\mathbf{x} = 40\}$ consiste dell'elemento f_4 ;
- l'insieme $\{\mathbf{x} = 35\}$ è vuoto.

Definizione formale Una v.a. è una regola che associa un numero $\mathbf{x}(\zeta)$ ad ogni uscita ζ di un esperimento. Per essere garantiti che le condizioni su \mathbf{x} siano eventi bisogna che tale assegnazione soddisfi le due condizioni seguenti:

I L'insieme $\{\mathbf{x} \leq x\}$ è un evento per ogni x .

II Le probabilità degli eventi $\{\mathbf{x} = \infty\}$ e $\{\mathbf{x} = -\infty\}$ valgono 0:

$$P\{\mathbf{x} = \infty\} = 0 \quad P\{\mathbf{x} = -\infty\} = 0$$

La seconda condizione dice che, nonostante sia permesso che \mathbf{x} valga ∞ o $-\infty$ per certe uscite, chiediamo che queste uscite formino un insieme con probabilità nulla.

v.a. complessa Una v.a. complessa \mathbf{z} è la somma

$$\mathbf{z} = \mathbf{x} + j\mathbf{y}$$

di due v.a. \mathbf{x} e \mathbf{y} reali. A meno che non sia specificato diversamente assumeremo che tutte le v.a. nel seguito siano reali.

6.2 Funzione di distribuzione di probabilità di una v.a.

Gli elementi dell'insieme \mathcal{S} che sono contenuti nell'evento $\{\mathbf{x} \leq x\}$ cambiano al cambiare del numero x . La probabilità $P\{\mathbf{x} \leq x\}$ dell'evento $\{\mathbf{x} \leq x\}$ è quindi un numero che dipende da x . Questo numero è $F_x(x)$ ed è chiamato la funzione distribuzione (cumulativa) di probabilità della v.a. \mathbf{x} .

Definizione: La funzione di distribuzione della v.a. \mathbf{x} è la funzione

$$F_x(x) = P\{\mathbf{x} \leq x\} \quad (6.1)$$

definita per ogni x da $-\infty$ a ∞ . In questa notazione la variabile x può essere sostituita da qualsiasi simbolo, per cui la 6.1 può anche essere scritta, ad esempio, come $F_x(w) = P\{\mathbf{x} \leq w\}$. Ciò non toglie che, in assenza di ambiguità, si possa anche usare la notazione semplificata $F(x)$.

Esempio 6.3 Nell'esperimento del lancio della moneta la probabilità di *testa* vale p e quella di *croce* vale q . Definiamo la v.a. \mathbf{x} tale che

$$\mathbf{x}(\text{testa}) = 1 \quad \mathbf{x}(\text{croce}) = 0$$

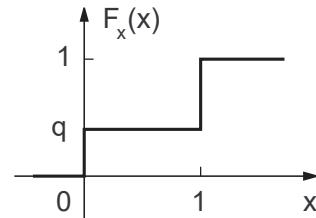
e vediamo l'andamento della funzione di distribuzione $F(x)$ per ogni x da $-\infty$ a ∞ .

Se $x \geq 1$ allora $F(x) = P\{\mathbf{x} \leq x\} = P\{\text{testa, croce}\} = 1$.

Se $0 \leq x < 1$ allora $F(x) = P\{\mathbf{x} \leq x\} = P\{\text{croce}\} = q$.

Se $x < 0$ allora $F(x) = P\{\mathbf{x} \leq x\} = P\{\emptyset\} = 0$.

La distribuzione $F_x(x)$ risultante è in figura.



Statistiche

Diremo che le statistiche di una v.a. \mathbf{x} sono note se possiamo determinare la probabilità $P\{\mathbf{x} \in R\}$ che \mathbf{x} sia in un insieme R dell'asse x costituito dall'unione od intersezione di un numero contabile di intervalli. Dalla 6.1 e dagli assiomi segue che le statistiche di \mathbf{x} sono note se è nota la sua funzione di distribuzione $F_x(x)$.

6.2.1 Proprietà della funzione di distribuzione di probabilità

Definiamo le due notazioni $F_x(x^+)$ e $F_x(x^-)$ come i limiti

$$F_x(x^+) = \lim_{\varepsilon \rightarrow 0} F_x(x + \varepsilon) \quad F_x(x^-) = \lim_{\varepsilon \rightarrow 0} F_x(x - \varepsilon) \quad \varepsilon > 0$$

$F_x(x^+)$ è detto limite da destra, mentre $F_x(x^-)$ è detto limite da sinistra.

La funzione di distribuzione $F_x(x)$ ha le seguenti proprietà:

1. $F_x(+\infty) = 1$ e $F_x(-\infty) = 0$

2. è una funzione nondecrescente di x :

$$\text{se } x_1 < x_2 \text{ allora } F_x(x_1) \leq F_x(x_2)$$

3. se $F_x(x_0) = 0$ allora $F_x(x) = 0$ per ogni $x \leq x_0$

4. $P\{\mathbf{x} > x\} = 1 - F_x(x)$

5. la funzione di distribuzione $F_x(x)$ è continua da destra: $F_x(x^+) = F_x(x)$, mentre può accadere che $F_x(x^-) \neq F_x(x)$

6. $P\{x_1 < \mathbf{x} \leq x_2\} = F_x(x_2) - F_x(x_1)$

7. $P\{\mathbf{x} = x\} = F_x(x) - F_x(x^-) = F_x(x^+) - F_x(x^-)$

La proprietà 5. richiede qualche approfondimento. Se una generica funzione $s(t)$ ha una discontinuità a gradino (un salto) nel punto $t = t_0$ allora, nell'interpretazione più comune, si dice che il suo valore in t_0 è indefinito, cioè $s(t_0)$ non si sa quanto vale. Consideriamo ora la funzione $F_x(x)$ dell'esempio 6.3: nel punto $x = 1$ ha una discontinuità a gradino, ovvero un salto fra il valore q ed il valore 1. Tuttavia $F_x(1)$ non è indefinita. Infatti $F_x(1) = P\{\mathbf{x} \leq 1\}$, ovvero è la probabilità che la v.a. \mathbf{x} sia minore o uguale a 1: poiché \mathbf{x} può assumere solo i due valori zero e uno è evidente che $F_x(1) = 1$. Quindi la funzione di distribuzione di probabilità assume, in presenza di una discontinuità a gradino, il valore più alto ovvero, essendo una funzione non decrescente, il valore a destra. Questa proprietà è una diretta conseguenza della definizione 6.1, in particolare dell'uguale nel \leq (se fosse definita solo minore sarebbe continua da sinistra, e tutta la teoria delle v.a. sarebbe diversa).

Nella proprietà 6. ponete attenzione a dov'è l'uguale: le possibili varianti sono:

$$\begin{aligned} P\{x_1 \leq \mathbf{x} \leq x_2\} &= P\{x = x_1\} + P\{x_1 < \mathbf{x} \leq x_2\} = F_x(x_2) - F_x(x_1^-) \\ P\{x_1 \leq \mathbf{x} < x_2\} &= P\{x = x_1\} + P\{x_1 < \mathbf{x} \leq x_2\} - P\{x = x_2\} = F_x(x_2^-) - F_x(x_1^-) \\ P\{x_1 < \mathbf{x} < x_2\} &= P\{x_1 < \mathbf{x} \leq x_2\} - P\{x = x_2\} = F_x(x_2^-) - F_x(x_1) \end{aligned} \quad (6.2)$$

v.a. di tipo continuo, discreto e misto Una v.a. \mathbf{x} è di tipo *continuo* se la sua funzione di distribuzione $F_x(x)$ è continua. In questo caso $F_x(x^-) = F_x(x)$ da cui segue che $P\{\mathbf{x} = x\} = 0$ per ogni x .

Una v.a. \mathbf{x} è di tipo *discreto* se la sua funzione di distribuzione $F_x(x)$ è a scala, ovvero costante ovunque tranne che in un certo numero di punti x_i di salto (discontinuità), dove vale che

$$F_x(x_i) - F_x(x_i^-) = P\{\mathbf{x} = x_i\} = p_i \quad (6.3)$$

La funzione distribuzione di probabilità dell'esempio 6.3 è di tipo discreto.

Una v.a. \mathbf{x} è di tipo *misto* se la sua funzione di distribuzione $F(x)$ ha delle discontinuità ma non è a scala (ovvero contiene anche tratti continui non costanti).

Da notare che se un insieme \mathcal{S} ha un numero finito di elementi, allora ogni v.a. definita su di esso è di tipo discreto. Tuttavia una v.a. può essere di tipo discreto anche se \mathcal{S} ha un numero infinito di elementi.

6.3 Funzione di densità di probabilità di una v.a.

La derivata

$$f_x(x) = \frac{dF_x(x)}{dx} \quad (6.4)$$

di $F(x)$ è chiamata funzione di densità di probabilità della v.a. \mathbf{x} .

Se la v.a. \mathbf{x} è di tipo discreto ed assume i valori x_i con probabilità p_i (come in 6.3), allora

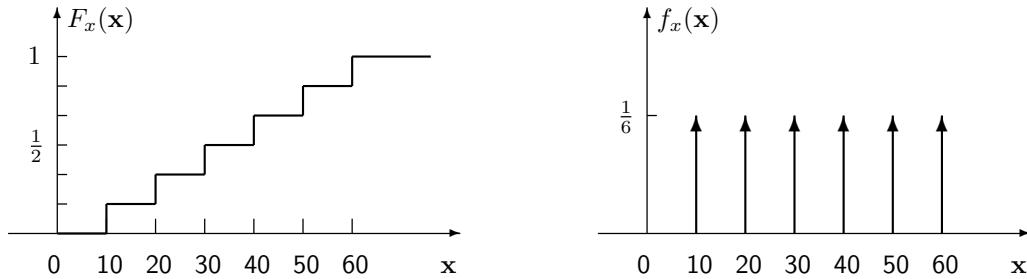
$$f(x) = \sum_i p_i \delta(x - x_i) \quad p_i = P\{\mathbf{x} = x_i\} \quad (6.5)$$

dove $\delta(t)$ è la funzione impulso. I termini $p_i \delta(x - x_i)$ sono normalmente disegnati come frecce verticali in $x = x_i$ di altezza p_i (vedi esempio 6.4).

Esempio 6.4 Consideriamo la v.a. definita nell'esempio 6.1 (a). La v.a. \mathbf{x} è di tipo discreto ed assume i sei valori $x_1 = 10, \dots, x_6 = 60$ con probabilità $1/6$. Quindi

$$f_x(x) = \frac{1}{6} [\delta(x - 10) + \delta(x - 20) + \dots + \delta(x - 60)]$$

Le funzioni distribuzione e densità di probabilità di \mathbf{x} sono visualizzate in figura.



6.3.1 Proprietà della funzione di densità di probabilità

Poichè $F_x(x)$ è non decrescente, segue che

$$f_x(x) \geq 0 \quad (6.6)$$

Integrando l'equazione 6.4 ed usando il fatto che $F(-\infty) = 0$ otteniamo

$$F_x(x) = \int_{-\infty}^x f_x(\xi) d\xi \quad (6.7)$$

Poichè $F(\infty) = 1$

$$\int_{-\infty}^{\infty} f_x(x) dx = 1 \quad (6.8)$$

Dalla 6.7 segue che

$$F_x(x_2) - F_x(x_1) = \int_{x_1}^{x_2} f_x(x) dx \quad (6.9)$$

Quindi (occhio agli estremi !)

$$P\{x_1 < \mathbf{x} \leq x_2\} = \int_{x_1}^{x_2} f_x(x) dx \quad (6.10)$$

Se la v.a. \mathbf{x} è di tipo continuo i segni di uguale in $P\{\dots\}$ possono essere messi come si vuole, ma se \mathbf{x} è di tipo discreto e vi sono discontinuità di $F_x(x)$ in x_1 e/o x_2 bisogna mettere gli estremi corretti.
Dalla 6.10 segue che, se la v.a. \mathbf{x} è di tipo continuo,

$$P\{x < \mathbf{x} \leq x + \Delta x\} = P\{x \leq \mathbf{x} \leq x + \Delta x\} = \int_x^{x+\Delta x} f_x(x) dx \approx f_x(x) \Delta x \quad (6.11)$$

purchè Δx sia abbastanza piccolo. Da ciò deriva che $f_x(x)$ può essere definita come il limite

$$f_x(x) = \lim_{\Delta x \rightarrow 0} \frac{P\{x \leq \mathbf{x} \leq x + \Delta x\}}{\Delta x} \quad (6.12)$$

Quindi la probabilità che \mathbf{x} sia in un piccolo intervallo di lunghezza Δx specificata è proporzionale a $f_x(x)$, ed è massima se l'intervallo contiene il punto x_m dove $f_x(x)$ è massima. Tale punto è detto *moda* o *valore più probabile* di \mathbf{x} . Una v.a. è detta unimodale se ha un singolo punto x_m di massimo. La $P\{\mathbf{x} = x\}$ è diversa da zero solo per quei punti x per i quali $f_x(x)$ ha un impulso.

Per v.a. continue vale allora che, per ogni x , $P\{\mathbf{x} = x\} = 0$: la differenza fra i valori di x possibili e quelli impossibili è che se $f_x(x) > 0$ è possibile che \mathbf{x} assuma il valore x , se $f_x(x) = 0$ no. Inoltre, sempre per v.a. continue, se $f_x(x_1) > f_x(x_2)$ la probabilità che \mathbf{x} assuma valori in un piccolo intorno Δx di x_1 è maggiore della probabilità che assuma valori in un intorno Δx di x_2 . La definizione di densità di probabilità permette di trattare in modo rigoroso esperimenti ad infinite uscite, cosa impossibile con la teoria della probabilità classica.

6.4 Media e varianza

Come visto, la conoscenza della funzione di distribuzione di probabilità di una v.a. definisce completamente le sue caratteristiche statistiche, ovvero non vi è altro che serva conoscere. Anche la densità di probabilità definisce completamente le statistiche di una v.a., cosa ovvia, visto che $F_x(x)$ e $f_x(x)$ possono essere ricavate l'una dall'altra.

A partire da tale conoscenza è spesso utile ricavare informazioni più limitate, che descrivano in maniera incompleta ma sintetica certe caratteristiche della v.a. in esame. Se, ad esempio, definisco una v.a. che rappresenta il guadagno mensile di una persona, posso essere interessato a quanto essa guadagni in media e magari ad un indice che rappresenti la variabilità di tali guadagni: per un dipendente a stipendio fisso la variabilità è bassa (zero, a parte aumenti, inflazione e tredicesima), per un libero professionista può essere anche molto alta.

Le due informazioni di sintesi più importanti sono la media e la varianza.

Media di una v.a. Il *valore atteso* o **media** di una qualsiasi v.a. \mathbf{x} è, per definizione, l'integrale

$$E\{\mathbf{x}\} = \bar{x} = \int_{-\infty}^{\infty} x f_x(x) dx \quad (6.13)$$

Questo numero verrà anche indicato come η_x o η .

Linearità dell'operatore di media Dalla definizione 6.13 segue immediatamente che l'operatore di media è lineare, per cui

$$E\{a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_3 \mathbf{x}_3\} = a_1 E\{\mathbf{x}_1\} + a_2 E\{\mathbf{x}_2\} + \dots + a_3 E\{\mathbf{x}_3\}$$

Se $z = x + jy$ è una **v.a. complessa** la sua media si esprime allora come

$$E\{z\} = E\{x\} + jE\{y\}$$

Da notare anche che, detta a una arbitraria costante, la sua media statistica vale $E\{a\} = a$.

Media di una v.a. discreta. Per una v.a. discreta, l'integrale in 6.13 può essere scritto come una sommatoria. Supponiamo che la v.a. \mathbf{x} assuma i valori x_i con probabilità p_i . In questo caso

$$f(x) = \sum_i p_i \delta(x - x_i)$$

Inserendola nella 6.13 ed usando l'identità

$$\int_{-\infty}^{\infty} x \delta(x - x_i) dx = x_i$$

otteniamo

$$E\{\mathbf{x}\} = \bar{x} = \sum_i p_i x_i \quad p_i = P\{\mathbf{x} = x_i\} \quad (6.14)$$

Varianza di una v.a. La **varianza** di una qualsiasi v.a. \mathbf{x} è, per definizione,

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx \quad (6.15)$$

La costante positiva σ_x è detta **deviazione standard**.

Dalla 6.15 segue che σ_x^2 è la media di $(\mathbf{x} - \bar{x})^2$, ovvero la media delle differenze al quadrato della v.a. dal suo valor medio. Quindi si può scrivere anche come

$$\sigma_x^2 = E\{(\mathbf{x} - \bar{x})^2\} = E\{\mathbf{x}^2 - 2\mathbf{x}\bar{x} + \bar{x}^2\} = E\{\mathbf{x}^2\} - E^2\{\mathbf{x}\} \quad (6.16)$$

cioè come il *valore quadratico medio* $E\{\mathbf{x}^2\}$ meno il valor medio al quadrato: ovviamente

$$E\{\mathbf{x}^2\} = \int_{-\infty}^{\infty} x^2 f(x) dx$$

Dalla linearità dell'operatore di media segue che

$$E\{\mathbf{x}^2 - 2\mathbf{x}\bar{x} + \bar{x}^2\} = E\{\mathbf{x}^2\} - 2E\{\mathbf{x}\bar{x}\} + E\{\bar{x}^2\} = E\{\mathbf{x}^2\} - 2\bar{x}^2 + \bar{x}^2 = E\{\mathbf{x}^2\} - E^2\{\mathbf{x}\}$$

La varianza σ_x^2 è un indice di quanto la v.a. \mathbf{x} sia concentrata attorno al suo valor medio, ovvero della sua variabilità. Ad esempio se una v.a. assume un solo valore (banale ma possibile) la sua media è tale valore e la sua varianza è zero (infatti non varia mai). Riprenderemo l'argomento in sezione 6.6.

Varianza di una v.a. discreta. Per una v.a. discreta, l'integrale in 6.15 può essere scritto come una sommatoria. Supponiamo che la v.a. \mathbf{x} assuma i valori x_i con probabilità p_i . In questo caso

$$f(x) = \sum_i p_i \delta(x - x_i)$$

Inserendola nella 6.15 otteniamo

$$\sigma^2 = \sum_i p_i (x_i - \bar{x})^2 \quad p_i = P\{x = x_i\}$$

NOTA: se la funzione $f_x(x)$ viene interpretata come una densità di massa sull'asse x , allora $\eta_x = E\{\mathbf{x}\}$ equivale al baricentro (centro di gravità), il valore quadratico medio $E\{\mathbf{x}^2\}$ è il momento di inerzia rispetto all'origine e la varianza σ_x^2 è il momento centrale di inerzia.

6.4.1 Momenti*

Oltre a media e varianza è possibile definire altri (infiniti) numeri che rappresentino una qualche informazione sulla v.a. in esame. Quelli più diffusi sono i **momenti**.

Si definisce momento ennesimo:

$$m_n = E\{\mathbf{x}^n\} = \int_{-\infty}^{\infty} x^n f(x) dx \quad (6.17)$$

Si definisce momento centrale ennesimo:

$$\mu_n = E\{(\mathbf{x} - \eta)^n\} = \int_{-\infty}^{\infty} (x - \eta)^n f(x) dx \quad (6.18)$$

I momenti centrali sono legati ai momenti da

$$\mu_n = E\{(\mathbf{x} - \eta)^n\} = E\left\{\sum_{k=0}^n \binom{n}{k} \mathbf{x}^k (-\eta)^{n-k}\right\} = \sum_{k=0}^n \binom{n}{k} m_k (-\eta)^{n-k}$$

Esistono anche i momenti assoluti normali e centrali, definiti come

$$E\{|\mathbf{x}|^n\} \quad E\{|(\mathbf{x} - \eta)|^n\} \quad (6.19)$$

ed i momenti generalizzati (normali ed assoluti) riferiti ad una qualsiasi costante a :

$$E\{(\mathbf{x} - a)^n\} \quad E\{|(\mathbf{x} - a)|^n\} \quad (6.20)$$

I numeri η_x e σ_x^2 danno una caratterizzazione limitata di $f(x)$. I momenti aggiungono informazione, caratterizzando meglio $f(x)$. Sotto certe ipotesi, conoscere m_n per ogni n permette di determinare $f(x)$ in modo univoco (ciò è noto agli studiosi del settore come *problema dei momenti*).

6.5 V.a. a distribuzione o densità notevoli

Abbiamo visto esempi di v.a. definite a partire da esperimenti. Ora vogliamo considerare v.a. aventi specifiche densità o distribuzioni di probabilità senza riferimento ad un particolare esperimento. Per fare questo dobbiamo verificare sotto quali condizioni questo è possibile.

6.5.1 Teorema di esistenza

Data una funzione $f(x)$ o il suo integrale

$$F(x) = \int_{-\infty}^x f(\xi) d\xi \quad (6.21)$$

possiamo costruire una v.a. \mathbf{x} con distribuzione $F(x)$ o densità $f(x)$ se valgono le proprietà seguenti:

la funzione $f(x)$ deve essere non negativa e la sua area deve essere pari a 1. La funzione $F(x)$ deve essere continua a destra e, al crescere di x da $-\infty$ a ∞ , deve crescere monotonicamente da 0 a 1.

Il fatto di poter definire una v.a. solo attraverso la sua distribuzione o la sua densità di probabilità è un altro dei vantaggi della teoria delle v.a. rispetto alla teoria della probabilità, in quanto permette di ricavare relazioni e teoremi senza dover specificare ogni volta l'esperimento fisico collegato, come è necessario fare in teoria della probabilità.

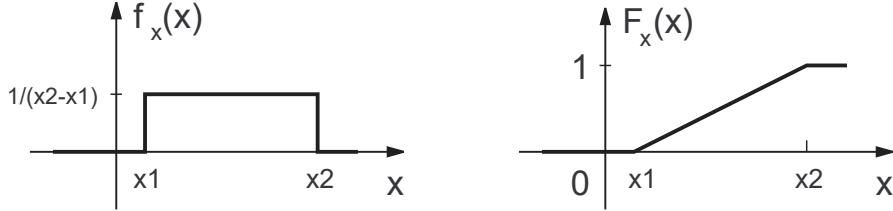
Nel seguito vediamo alcune densità notevoli.

6.5.2 Uniforme

Una v.a. \mathbf{x} è detta *uniforme* fra x_1 e x_2 se la sua densità è costante nell'intervallo (x_1, x_2) e nulla altrove.

$$f_x(x) = \begin{cases} \frac{1}{x_2 - x_1} & x_1 \leq x \leq x_2 \\ 0 & \text{altrove} \end{cases} \quad (6.22)$$

La corrispondente distribuzione $F_x(x)$ è una rampa. $F_x(x)$ è nulla prima di x_1 , vale 1 dopo x_2 e vale $(x - x_1)/(x_2 - x_1)$ per $x_1 \leq x \leq x_2$.



Useremo la notazione $U(x_1; x_2)$ per indicare che una v.a. è uniforme fra x_1 e x_2 .

Esempio 6.5 Un insieme di resistenze possono essere rappresentate da una v.a. \mathbf{r} uniforme fra 900 e 1100 Ω . Quale è la probabilità che \mathbf{r} sia fra 950 e 1050?

Poichè $f_r(r) = 1/200$ fra 900 e 1100 e zero altrove, si ottiene

$$P\{950 \leq \mathbf{r} \leq 1050\} = \frac{1}{200} \int_{950}^{1050} dr = 0.5$$

Se \mathbf{x} è uniforme nell'intervallo (x_1, x_2) , allora $f_x(x) = 1/(x_2 - x_1)$ in tale intervallo e zero altrove. Quindi la sua **media** è

$$E\{\mathbf{x}\} = \bar{x} = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} x dx = \frac{x_1 + x_2}{2}$$

Una **proprietà generale** di *qualsiasi* v.a. \mathbf{x} è che se la linea verticale $x = a$ è l'asse di simmetria di $f_x(x)$ allora $E\{\mathbf{x}\} = a$: in particolare se $f_x(-x) = f_x(x)$ allora $E\{\mathbf{x}\} = 0$.

Per una v.a. \mathbf{x} uniforme in (x_1, x_2) la densità $f_x(x)$ è infatti simmetrica rispetto alla linea $x = (x_1 + x_2)/2$.

La **varianza** di una v.a. \mathbf{x} uniforme in (x_1, x_2) è

$$\sigma_x^2 = E\{\mathbf{x}^2\} - E^2\{\mathbf{x}\} = \frac{x_2^2 + x_1^2 + x_1 x_2}{3} - \left(\frac{x_1 + x_2}{2}\right)^2 = \frac{(x_2 - x_1)^2}{12}$$

Provate a calcolarlo da soli, ricordando che $a^3 - b^3 = (a - b)(a^2 + b^2 + ab)$.

Esempio 6.6 Se \mathbf{x} è uniforme nell'intervallo $(-c, c)$, allora $\bar{x} = 0$ e

$$\sigma_x^2 = E\{\mathbf{x}^2\} = \frac{1}{2c} \int_{-c}^c x^2 dx = \frac{c^2}{3}$$

6.5.3 Normale o Gaussiana

Una v.a. è detta *normale* o *gaussiana* se la sua densità è la curva gaussiana $\mathbf{g}(x)$ definita in 5.22, traslata e scalata

$$f(x) = \frac{1}{\sigma} \mathbf{g}\left(\frac{x - \eta}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\eta)^2}{2\sigma^2}} \quad (6.23)$$

Questa è una curva a campana, simmetrica rispetto alla linea $x = \eta$ e la sua area è pari a 1. La corrispondente distribuzione di probabilità è data da (vedi 5.23)

$$F(x) = \mathbf{G}\left(\frac{x - \eta}{\sigma}\right) \quad (6.24)$$

Nelle equazioni precedenti η è la media e σ^2 la varianza. Useremo la notazione $N(\eta; \sigma)$ per indicare che una v.a. è normale o gaussiana con i parametri η e σ .

Esempio 6.7 Una v.a. \mathbf{x} è $N(1000; 50)$. Vogliamo trovare la probabilità che \mathbf{x} sia fra 900 e 1050. Chiaramente

$$P\{900 \leq \mathbf{x} \leq 1050\} = F(1050) - F(900) = \mathbf{G}(1) - \mathbf{G}(-2) = \mathbf{G}(1) + \mathbf{G}(2) - 1$$

avendo usato la proprietà $\mathbf{G}(-y) = 1 - \mathbf{G}(y)$. Dai valori tabulati di \mathbf{G} (vedi Appendice) la probabilità cercata vale circa 0.819.

Esercizio 6.8 Data una v.a. \mathbf{x} di tipo $N(\eta; \sigma)$ trovare la probabilità che \mathbf{x} sia fra $\eta - \sigma$ e $\eta + \sigma$, ovvero in un intorno $\pm\sigma$ di η . [RISPOSTA: 0.682]

Esercizio 6.9 Data una v.a. \mathbf{x} di tipo $N(\eta; \sigma)$ trovare la probabilità che \mathbf{x} sia fra $\eta - 3\sigma$ e $\eta + 3\sigma$, ovvero in un intorno $\pm 3\sigma$ di η . [RISPOSTA: 0.9973]

Nella equazione 6.23 η è la media e σ^2 la varianza. Dimostriamolo.

1) Ovviamente si può calcolare la media mediante l'eq. 6.13, ma è più semplice notare che $f(x)$ è simmetrica rispetto a $x = \eta$: quindi $E\{\mathbf{x}\} = \bar{x} = \eta$.

2) L'area di $f(x)$ è pari a 1, quindi

$$\int_{-\infty}^{\infty} e^{-(x-\eta)^2/2\sigma^2} dx = \sigma\sqrt{2\pi}$$

differenziamo rispetto a σ

$$\int_{-\infty}^{\infty} \frac{(x-\eta)^2}{\sigma^3} e^{-(x-\eta)^2/2\sigma^2} dx = \sqrt{2\pi}$$

moltiplichiamo ambo i membri per $\sigma^2/\sqrt{2\pi}$

$$\int_{-\infty}^{\infty} (x-\eta)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\eta)^2/2\sigma^2} dx = \sigma^2$$

e concludiamo che $E\{(\mathbf{x} - \eta)^2\} = \sigma^2$.

6.5.4 Binomiale

Una v.a. ha una distribuzione *binomiale* di ordine n se assume i valori $0, 1, \dots, n$ con probabilità

$$P\{\mathbf{x} = k\} = \binom{n}{k} p^k q^{n-k} \quad p + q = 1 \quad (6.25)$$

Quindi \mathbf{x} è di tipo discreto e la sua densità vale

$$f(x) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \delta(x - k) \quad (6.26)$$

La corrispondente distribuzione è una funzione a scala e nell'intervallo $(0, n)$ vale

$$F(x) = \sum_{k=0}^m \binom{n}{k} p^k q^{n-k} \quad m \leq x < m+1 \quad (6.27)$$

Come già visto nella sezione 5.10.2, se n è grande $F(x)$ è simile alla distribuzione di una v.a. gaussiana $N(np, \sqrt{npq})$, ovvero

$$F(x) \simeq \mathbf{G}\left(\frac{x - np}{\sqrt{npq}}\right) \quad (6.28)$$

Attenzione: se n è grande la distribuzione tende alla distribuzione gaussiana, ma la densità NON tende alla densità gaussiana (una è una sequenza di impulsi, l'altra è continua!).

6.5.5 Poisson

Una v.a. ha una distribuzione di *Poisson* con parametro a se assume i valori $0, 1, \dots, n, \dots$ con probabilità

$$P\{\mathbf{x} = k\} = e^{-a} \frac{a^k}{k!} \quad k = 0, 1, \dots \quad (6.29)$$

Quindi \mathbf{x} è di tipo discreto e la sua densità vale

$$f(x) = e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} \delta(x - k) \quad (6.30)$$

La corrispondente distribuzione è una funzione a scala.

Chiamando $p_k = P\{\mathbf{x} = k\}$, dalla 6.29 segue che

$$\frac{p_{k-1}}{p_k} = \frac{e^{-a} \frac{a^{k-1}}{(k-1)!}}{e^{-a} \frac{a^k}{k!}} = \frac{k}{a}$$

Se il rapporto k/a è minore di uno, cioè se $k < a$, allora $p_{k-1} < p_k$. Questo mostra che, al crescere di k , p_k cresce, raggiungendo il massimo per $k = [a]$ (al solito $[a]$ significa il massimo intero che non supera a). Più precisamente:

se $a < 1$ allora p_k è massima per $k = 0$;

se $a > 1$ ma non è un intero il massimo è per $k = [a]$;

se a è un intero allora p_k è massima per $k = a - 1$ e per $k = a$.

La **media**, il valore quadratico medio e la **varianza** della v.a. definita in 6.30 sono:

$$E\{\mathbf{x}\} = a \quad E\{\mathbf{x}^2\} = a^2 + a \quad \sigma^2 = a \quad (6.31)$$

La dimostrazione (per i curiosi) è sul [Pap91] pag. 108, ma non siete tenuti a saperla.

Solitamente una v.a. di Poisson è il numero di eventi che occorrono in un certo intervallo di tempo (o di spazio) se: (i) la lunghezza del periodo di osservazione è fissato in precedenza, (ii) gli eventi accadono con densità media costante, (iii) il numero di eventi che accadono in intervalli disgiunti è statisticamente indipendente. Possibili esempi sono il numero di automobili che transitano in un punto fisso in 5 minuti o il numero di telefonate ricevute da un call-center in un'ora, supponendo che la densità media non vari.

La distribuzione di Poisson può essere usata per approssimare la Binomiale quando n è grande e p è piccolo: in questo caso $a = np$. L'utilità consiste nel fatto che i conti per il calcolo della binomiale sono assai più complessi di quelli per Poisson (vedi sezione 5.13).

Esiste anche la versione *continua*: detto $\lambda > 0$ il parametro di tale distribuzione, la sua densità è definita da

$$f_x(x) = f_x(x; \lambda) = \lambda e^{-\lambda x} \quad \text{per } x \geq 0 \quad (6.32)$$

e zero altrove. La corrispondente distribuzione vale

$$F_x(x) = F_x(x; \lambda) = 1 - e^{-\lambda x} \quad \text{per } x \geq 0 \quad (6.33)$$

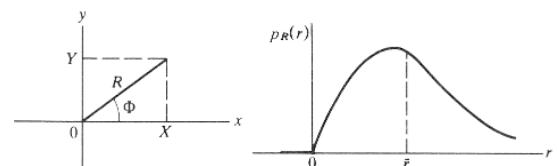
e zero per $x < 0$.

La media vale $\bar{x} = 1/\lambda$ e la varianza $\sigma_x^2 = 1/\lambda^2$.

6.5.6 Rayleigh

Il modello di Rayleigh descrive una v.a. continua prodotta da due v.a. gaussiane mediante la trasformazione in figura. Siano \mathbf{x} e \mathbf{y} due v.a. *indipendenti* gaussiane con media e varianza

$$\eta_1 = \eta_2 = 0 \quad \sigma_1 = \sigma_2 = \sigma$$



La v.a. $\mathbf{r} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2}$ ha la densità di probabilità di Rayleigh (vedi figura)

$$f_r(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} u(r) \quad (6.34)$$

Poichè la v.a. \mathbf{r} non può essere negativa (per definizione), deve avere un valor medio positivo anche se \mathbf{x} e \mathbf{y} sono a media nulla. La media ed il valore quadratico medio di \mathbf{r} sono

$$\bar{r} = \sqrt{\frac{\pi}{2}} \sigma \quad \overline{r^2} = 2\sigma^2 \quad (6.35)$$

dove σ^2 è la varianza delle due gaussiane.

La distribuzione di \mathbf{r} si ottiene facilmente integrando la densità in 6.34

$$F_r(r) = P\{\mathbf{r} \leq r\} = \left(1 - e^{-r^2/2\sigma^2}\right) u(r) \quad (6.36)$$

6.6 Diseguaglianza di Tchebycheff

Una misura della concentrazione di una v.a. attorno al suo valor medio η è la sua varianza σ^2 . Infatti la probabilità che la v.a. \mathbf{x} sia fuori da un arbitrario intervallo $(\eta - \varepsilon, \eta + \varepsilon)$ è trascurabile se il rapporto σ/ε è abbastanza piccolo. Più precisamente Tchebycheff ha dimostrato che

$$P\{|\mathbf{x} - \eta| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \quad \forall \varepsilon > 0 \quad (6.37)$$

La dimostrazione è basata sul fatto che

$$P\{|\mathbf{x} - \eta| \geq \varepsilon\} = \int_{|\mathbf{x} - \eta| \geq \varepsilon} f(x) dx$$

quindi

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \eta)^2 f(x) dx \geq \int_{|\mathbf{x} - \eta| \geq \varepsilon} (x - \eta)^2 f(x) dx \geq \varepsilon^2 \int_{|\mathbf{x} - \eta| \geq \varepsilon} f(x) dx$$

e la 6.37 segue.

Da notare che il limite imposto dalla diseguaglianza di Tchebycheff può essere troppo alto per densità note. Ad esempio se \mathbf{x} è una v.a. gaussiana ed assumiamo come valore di $\varepsilon = 3\sigma$ si ottiene che

$$P\{|\mathbf{x} - \eta| \geq 3\sigma\} = 2 - 2\mathbf{G}(3) \approx 0.0027$$

mentre la diseguaglianza di Tchebycheff fornisce il valore

$$P\{|\mathbf{x} - \eta| \geq 3\sigma\} \leq \frac{1}{9} \approx 0.1111$$

L'importanza della diseguaglianza di Tchebycheff consiste nel fatto che vale per ogni $f(x)$ e può quindi essere usata anche se $f(x)$ non è nota.

6.7 Distribuzioni e densità di probabilità condizionate

Come visto in sezione 5.4, la probabilità di un evento \mathcal{A} assumendo \mathcal{B} è data da

$$P(\mathcal{A}/\mathcal{B}) = \frac{P(\mathcal{A}\mathcal{B})}{P(\mathcal{B})} \quad \text{con} \quad P(\mathcal{B}) \neq 0$$

La *distribuzione condizionata* $F(x/\mathcal{B})$ di una v.a. \mathbf{x} assumendo \mathcal{B} è definita come la probabilità condizionata dell'evento $\{\mathbf{x} \leq x\}$:

$$F(x/\mathcal{B}) = P\{\mathbf{x} \leq x/\mathcal{B}\} = \frac{P\{\mathbf{x} \leq x, \mathcal{B}\}}{P\{\mathcal{B}\}} \quad (6.38)$$

Nell'equazione precedente, $\{\mathbf{x} \leq x, \mathcal{B}\}$ è l'intersezione degli eventi $\{\mathbf{x} \leq x\}$ e \mathcal{B} , ovvero l'evento che consiste di tutti i risultati ξ tali che $\mathbf{x}(\xi) \leq x$ e $\xi \in \mathcal{B}$.

La definizione di $F(x/\mathcal{B})$ è quindi analoga alla definizione 6.1 di $F(x)$ sostituendo tutte le probabilità con probabilità condizionate. Da ciò segue che $F(x/\mathcal{B})$ ha le stesse proprietà di $F(x)$. In particolare:

$$F(\infty/\mathcal{B}) = 1 \quad F(-\infty/\mathcal{B}) = 0 \quad (6.39)$$

$$P\{x_1 < \mathbf{x} \leq x_2/\mathcal{B}\} = F(x_2/\mathcal{B}) - F(x_1/\mathcal{B}) = \frac{P\{x_1 < \mathbf{x} \leq x_2, \mathcal{B}\}}{P(\mathcal{B})} \quad (6.40)$$

La *densità condizionata* $f(x/\mathcal{B})$ è la derivata di $F(x/\mathcal{B})$:

$$f(x/\mathcal{B}) = \frac{dF(x/\mathcal{B})}{dx} = \lim_{\Delta x \rightarrow 0} \frac{P\{\mathbf{x} \leq \mathbf{x} \leq x + \Delta x / \mathcal{B}\}}{\Delta x} \quad (6.41)$$

è non negativa e la sua area vale 1.

Per trovare $F(x/\mathcal{B})$ dobbiamo, in generale, conoscere l'esperimento sottostante. Tuttavia, se \mathcal{B} è un evento che può essere espresso in termini della v.a. \mathbf{x} allora, per la determinazione di $F(x/\mathcal{B})$, è sufficiente la conoscenza di $F(x)$. I due casi che seguono chiariscono il concetto.

I. Vogliamo trovare la distribuzione condizionata di una v.a. \mathbf{x} assumendo che $\mathbf{x} \leq a$, dove a è un numero tale che $F(a) \neq 0$. Questo è un caso speciale della 6.38 dove $\mathcal{B} = \{\mathbf{x} \leq a\}$. Si tratta, nota $F(x)$, di trovare

$$F(x/\mathbf{x} \leq a) = P\{\mathbf{x} \leq x / \mathbf{x} \leq a\} = \frac{P\{\mathbf{x} \leq x, \mathbf{x} \leq a\}}{P\{\mathbf{x} \leq a\}}$$

Se $x \geq a$, allora $P\{\mathbf{x} \leq x, \mathbf{x} \leq a\} = P\{\mathbf{x} \leq a\}$ e quindi

$$F(x/\mathbf{x} \leq a) = \frac{P\{\mathbf{x} \leq a\}}{P\{\mathbf{x} \leq a\}} = 1 \quad x \geq a$$

Se $x < a$, allora $P\{\mathbf{x} \leq x, \mathbf{x} \leq a\} = P\{\mathbf{x} \leq x\}$ e quindi

$$F(x/\mathbf{x} \leq a) = \frac{P\{\mathbf{x} \leq x\}}{P\{\mathbf{x} \leq a\}} = \frac{F(x)}{F(a)} \quad x < a$$

Differenziando $F(x/\mathbf{x} \leq a)$ rispetto ad x otteniamo la densità corrispondente

$$f(x/\mathbf{x} \leq a) = \frac{f(x)}{F(a)} = \frac{f(x)}{\int_{-\infty}^a f(x) dx} \quad \text{per} \quad x < a \quad (6.42)$$

e 0 per $x \geq a$. È facile dimostrare che l'area di $f(x/\mathbf{x} \leq a)$ è 1, infatti, dalla 6.42

$$\int_{-\infty}^{\infty} f(x/\mathbf{x} \leq a) dx = \frac{1}{F(a)} \int_{-\infty}^a f(x) dx = \frac{F(a)}{F(a)} = 1$$

II. Vogliamo trovare la distribuzione condizionata di una v.a. \mathbf{x} assumendo che $b < \mathbf{x} \leq a$. In questo caso la 6.38 porta a

$$F(x/b < \mathbf{x} \leq a) = P\{\mathbf{x} \leq x / b < \mathbf{x} \leq a\} = \frac{P\{\mathbf{x} \leq x, b < \mathbf{x} \leq a\}}{P\{b < \mathbf{x} \leq a\}}$$

Se $x \geq a$, allora $P\{\mathbf{x} \leq x, b < \mathbf{x} \leq a\} = P\{b < \mathbf{x} \leq a\}$ e quindi

$$F(x/b < \mathbf{x} \leq a) = \frac{F(a) - F(b)}{F(a) - F(b)} = 1 \quad x \geq a$$

Se $b \leq x < a$, allora $P\{\mathbf{x} \leq x, b < \mathbf{x} \leq a\} = P\{b < \mathbf{x} \leq x\}$ e quindi

$$F(x/b < \mathbf{x} \leq a) = \frac{F(x) - F(b)}{F(a) - F(b)} \quad b \leq x < a$$

Infine se $x < b$, allora $P\{\mathbf{x} \leq x, b < \mathbf{x} \leq a\} = P\{\emptyset\}$ e quindi

$$F(x/b < \mathbf{x} \leq a) = 0 \quad x < b$$

La corrispondente densità è data da

$$f(x/b < \mathbf{x} \leq a) = \frac{f(x)}{F(a) - F(b)} \quad \text{per} \quad b \leq x < a \quad (6.43)$$

e 0 altrove.

Esempio 6.10 Vogliamo trovare la densità condizionata $f(x/|\mathbf{x} - \eta| \leq k\sigma)$ di una v.a. \mathbf{x} gaussiana $N(\eta; \sigma)$. Poiché

$$P\{|\mathbf{x} - \eta| \leq k\sigma\} = P\{\eta - k\sigma \leq \mathbf{x} \leq \eta + k\sigma\} = \mathbf{G}(k) - \mathbf{G}(-k) = 2\mathbf{G}(k) - 1$$

dalla 6.43 concludiamo che

$$f(x/|\mathbf{x} - \eta| \leq k\sigma) = \frac{1}{2\mathbf{G}(k) - 1} \frac{e^{-(x-\eta)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$

per \mathbf{x} fra $\eta - k\sigma$ e $\eta + k\sigma$ e 0 altrove. Questa densità è detta *normale (o gaussiana) troncata*.

Media condizionata. La media condizionata di una v.a. \mathbf{x} assumendo \mathcal{B} si ottiene dalle equazioni 6.13 e 6.14 inserendo, rispettivamente, densità e probabilità condizionate:

$$E\{\mathbf{x}/\mathcal{B}\} = \int_{-\infty}^{\infty} x f_x(x/\mathcal{B}) dx \quad \mathbf{x} \text{ v.a. continua} \quad (6.44)$$

$$E\{\mathbf{x}/\mathcal{B}\} = \sum_i x_i P\{\mathbf{x} = x_i/\mathcal{B}\} \quad \mathbf{x} \text{ v.a. discreta} \quad (6.45)$$

I valori condizionati della varianza, del valore quadratico medio e dei momenti si ottengono in modo analogo utilizzando nelle rispettive definizioni la densità condizionata al posto di quella non condizionata.

6.8 Probabilità totale e teorema di Bayes

Il teorema di Bayes è stato definito in sezione 5.5 per le probabilità: ovviamente si può estendere alle variabili aleatorie.

1. Ponendo $\mathcal{B} = \{\mathbf{x} \leq x\}$ nella equazione 5.7 otteniamo

$$P\{\mathbf{x} \leq x\} = P\{\mathbf{x} \leq x/\mathcal{A}_1\}P(\mathcal{A}_1) + \cdots + P\{\mathbf{x} \leq x/\mathcal{A}_n\}P(\mathcal{A}_n)$$

Quindi (vedi 6.38 e 6.41)

$$F(x) = F(x/\mathcal{A}_1)P(\mathcal{A}_1) + \cdots + F(x/\mathcal{A}_n)P(\mathcal{A}_n) \quad (6.46)$$

$$f(x) = f(x/\mathcal{A}_1)P(\mathcal{A}_1) + \cdots + f(x/\mathcal{A}_n)P(\mathcal{A}_n) \quad (6.47)$$

Dove gli eventi $\mathcal{A}_1, \dots, \mathcal{A}_n$ formano una partizione dello spazio degli eventi \mathcal{S} .

2. Dalla identità

$$P(\mathcal{A}/\mathcal{B}) = \frac{P(\mathcal{B}/\mathcal{A})P(\mathcal{A})}{P(\mathcal{B})} \quad (6.48)$$

(vedi eq. 5.8) segue che

$$P(\mathcal{A}/\mathbf{x} \leq x) = \frac{P\{\mathbf{x} \leq x/\mathcal{A}\}}{P\{\mathbf{x} \leq x\}} P(\mathcal{A}) = \frac{F(x/\mathcal{A})}{F(x)} P(\mathcal{A}) \quad (6.49)$$

3. Inserendo $\mathcal{B} = \{x_1 < \mathbf{x} \leq x_2\}$ nella 6.48 concludiamo che (vedi 6.40)

$$\begin{aligned} P(\mathcal{A}/x_1 < \mathbf{x} \leq x_2) &= \frac{P\{x_1 < \mathbf{x} \leq x_2/\mathcal{A}\}}{P\{x_1 < \mathbf{x} \leq x_2\}} P(\mathcal{A}) = \\ &= \frac{F(x_2/\mathcal{A}) - F(x_1/\mathcal{A})}{F(x_2) - F(x_1)} P(\mathcal{A}) \end{aligned} \quad (6.50)$$

4. La probabilità condizionata $P(\mathcal{A}/\mathbf{x} = x)$ dell'evento \mathcal{A} assumendo $\mathbf{x} = x$ non può essere definita come in 5.6 poichè, in generale, $P\{\mathbf{x} = x\} = 0$. La definiamo come limite

$$P(\mathcal{A}/\mathbf{x} = x) = \lim_{\Delta x \rightarrow 0} P\{\mathcal{A}/x < \mathbf{x} \leq x + \Delta x\} \quad (6.51)$$

usando l'eq. 6.50 con le sostituzioni $x_1 = x$ e $x_2 = x + \Delta x$ e portando al limite si ottiene

$$P(\mathcal{A}/\mathbf{x} = x) = \frac{f(x/\mathcal{A})}{f(x)} P(\mathcal{A}) \quad (6.52)$$

Quindi $P(\mathcal{A}/\mathbf{x} = x)$ esiste anche se $P\{\mathbf{x} = x\} = 0$, purchè $f(x) \neq 0$, ovvero se $\{\mathbf{x} = x\}$ è un evento a probabilità zero ma non impossibile.

Teorema della probabilità totale Come visto (eq. 6.39)

$$F(\infty/\mathcal{A}) = \int_{-\infty}^{\infty} f(x/\mathcal{A}) dx = 1$$

Moltiplicando la 6.52 per $f(x)$, integrando e tenendo conto dell'equazione sopra otteniamo

$$\int_{-\infty}^{\infty} P(\mathcal{A}/\mathbf{x} = x) f(x) dx = P(\mathcal{A}) \quad (6.53)$$

che è la versione continua del teorema della probabilità totale dell'eq. 5.7.

Teorema di Bayes Dalle equazioni 6.52 e 6.53 segue che

$$f(x/\mathcal{A}) = \frac{P(\mathcal{A}/\mathbf{x} = x)}{P(\mathcal{A})} f(x) = \frac{P(\mathcal{A}/\mathbf{x} = x) f(x)}{\int_{-\infty}^{\infty} P(\mathcal{A}/\mathbf{x} = x) f(x) dx} \quad (6.54)$$

che è la versione continua del teorema di Bayes dell'eq. 5.9.

6.9 Funzioni di una variabile aleatoria

Supponiamo che \mathbf{x} sia una variabile aleatoria e $g(x)$ una funzione della variabile reale x . L'espressione

$$\mathbf{y} = g(\mathbf{x}) \quad (6.55)$$

è una nuova v.a. definita come segue: per un dato ξ , $\mathbf{x}(\xi)$ è un numero e $g[\mathbf{x}(\xi)]$ è un altro numero specificato in termini di $\mathbf{x}(\xi)$ e $g(x)$. Questo numero è il valore $\mathbf{y}(\xi) = g[\mathbf{x}(\xi)]$ assegnato alla v.a. \mathbf{y} . La funzione distribuzione di probabilità $F_y(y)$ della v.a. così formata è la probabilità dell'evento $\{\mathbf{y} \leq y\}$ consistente di tutti i risultati ξ tali che $\mathbf{y}(\xi) = g[\mathbf{x}(\xi)] \leq y$:

$$F_y(y) = P\{\mathbf{y} \leq y\} = P\{g(\mathbf{x}) \leq y\} \quad (6.56)$$

Per uno specifico y , i valori di x tali che $g(x) \leq y$ formano un insieme sull'asse x denotato da R_y . Chiaramente $g[\mathbf{x}(\xi)] \leq y$ se $\mathbf{x}(\xi)$ è un numero nell'insieme R_y . Quindi

$$F_y(y) = P\{\mathbf{x} \in R_y\} \quad (6.57)$$

L'equazione precedente porta alla conclusione che, affinché $g(\mathbf{x})$ sia una v.a., la funzione $g(x)$ deve avere le seguenti proprietà:

1. il suo dominio deve includere l'intervallo di valori della v.a. \mathbf{x} .
2. Deve essere una funzione di *Baire*, cioè: per ogni y l'insieme R_y tale che $g(x) \leq y$ deve consistere della unione ed intersezione di un numero contabile di intervalli. Solo in questo modo $\{\mathbf{y} \leq y\}$ è un evento.
3. Gli eventi $\{g(\mathbf{x}) = \pm\infty\}$ devono avere probabilità zero.

Quindi la 6.55 definisce una nuova v.a. \mathbf{y} solo se le tre condizioni precedenti sono soddisfatte. Nel seguito della sezione si affronta il problema di definire le statistiche di $\mathbf{y} = g(\mathbf{x})$ a partire dalla conoscenza delle statistiche di \mathbf{x} e della funzione $g()$.

6.9.1 Determinazione della distribuzione di $\mathbf{y} = g(\mathbf{x})$

Vedremo qui come esprimere la funzione di distribuzione $F_y(y)$ della v.a. $\mathbf{y} = g(\mathbf{x})$ in termini della funzione di distribuzione $F_x(x)$ della v.a. \mathbf{x} e della funzione $g(x)$. Per fare questo dobbiamo determinare l'insieme R_y dell'asse x tale che $g(x) \leq y$ e la probabilità che \mathbf{x} sia in questo intervallo. Questo metodo verrà illustrato per mezzo di 5 casi diversi. Se non diversamente specificato assumeremo che $F_x(x)$ sia continua.

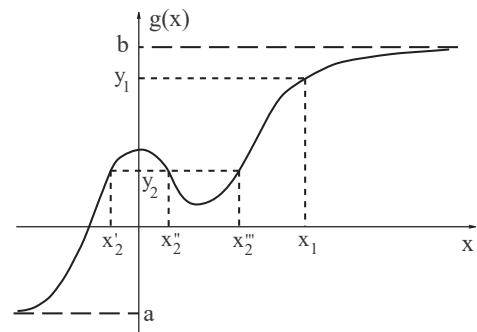
Caso 1. (g continua generale) Iniziamo con la funzione $g(x)$ in figura. Come si può notare $g(x)$ è fra a e b per ogni x .

Questo porta alla conclusione che se $y \geq b$, allora $g(x) \leq y$ per ogni x : quindi $P(\mathbf{y} \leq y) = 1$. Se $y < a$, non vi sono valori di x tali che $g(x) \leq y$ e quindi $P(\mathbf{y} \leq y) = 0$. Ovvero

$$F_y(y) = \begin{cases} 1 & y \geq b \\ 0 & y < a \end{cases}$$

Per i valori x_1 ed y_1 come in figura, osserviamo che $g(x) \leq y_1$ per $x \leq x_1$, per cui

$$F_y(y_1) = P\{\mathbf{x} \leq x_1\} = F_x(x_1)$$



Infine notiamo che

$$g(x) \leq y_2 \quad \text{se} \quad x \leq x'_2 \quad \text{oppure se} \quad x''_2 \leq x \leq x'''_2$$

e quindi

$$F_y(y_2) = P\{\mathbf{x} \leq x'_2\} + P\{x''_2 \leq \mathbf{x} \leq x'''_2\} = F_x(x'_2) + F_x(x'''_2) - F_x(x''_2)$$

essendo gli eventi $P\{\mathbf{x} \leq x'_2\}$ e $P\{x''_2 \leq \mathbf{x} \leq x'''_2\}$ mutuamente esclusivi.

Esempio 6.11

$$\mathbf{y} = a\mathbf{x} + b$$

Per trovare $F_y(y)$ dobbiamo trovare i valori di x tali che $a\mathbf{x} + b \leq y$. Quindi

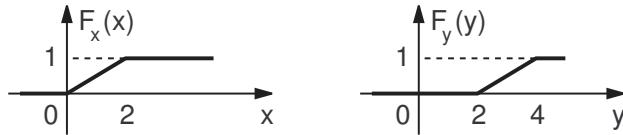
$$F_y(y) = P\left\{\mathbf{x} \leq \frac{y-b}{a}\right\} = F_x\left(\frac{y-b}{a}\right) \quad a > 0$$

$$F_y(y) = P\left\{\mathbf{x} \geq \frac{y-b}{a}\right\} = 1 - F_x\left(\frac{y-b}{a}\right) \quad a < 0$$

Caso speciale 1: Sia $y = x + 2$, quindi $a = 1 > 0$ e $b = 2$. Applicando il risultato precedente

$$F_y(y) = F_x(y-2)$$

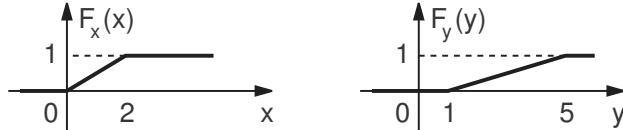
e se, ad esempio, la v.a. \mathbf{x} è uniforme in $[0, 2]$ ne risulta che \mathbf{y} è uniforme in $[2, 4]$.



Caso speciale 2: Sia $y = 2x + 1$, quindi $a = 2 > 0$ e $b = 1$. Applicando il risultato precedente

$$F_y(y) = F_x\left(\frac{y-1}{2}\right)$$

e se, ad esempio, la v.a. \mathbf{x} è uniforme in $[0, 2]$ ne risulta che \mathbf{y} è uniforme in $[1, 5]$.



Esempio 6.12

$$\mathbf{y} = \mathbf{x}^2$$

Se $y \geq 0$ allora $x^2 \leq y$ per $-\sqrt{y} \leq x \leq \sqrt{y}$ e quindi

$$F_y(y) = P\{-\sqrt{y} \leq x \leq \sqrt{y}\} = F_x(\sqrt{y}) - F_x(-\sqrt{y}) \quad y > 0$$

Se $y < 0$ non ci sono valori di x tali che $x^2 < y$ e

$$F_y(y) = P\{\emptyset\} = 0 \quad y < 0$$

Caso speciale: se \mathbf{x} è uniforme nell'intervallo $(-1, 1)$ (vedi Fig. 6.1), allora

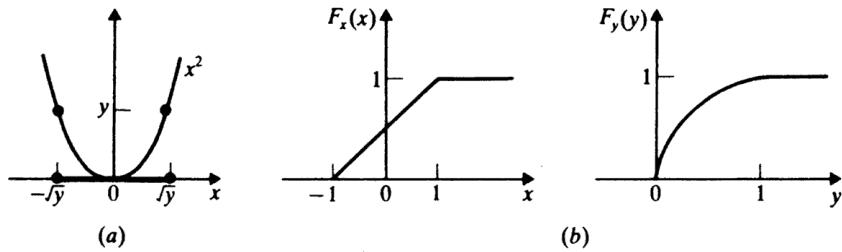
$$F_x(x) = \frac{1}{2} + \frac{x}{2} \quad |x| < 1$$

e quindi

$$F_y(y) = \sqrt{y} \quad \text{per} \quad 0 \leq y \leq 1 \quad \text{e} \quad F_y(y) = \begin{cases} 1 & y > 1 \\ 0 & y < 0 \end{cases}$$

Questo risultato si ottiene da:

$$F_y(y) = F_x(\sqrt{y}) - F_x(-\sqrt{y}) = \frac{1}{2} + \frac{\sqrt{y}}{2} - \left(\frac{1}{2} - \frac{\sqrt{y}}{2}\right) = \sqrt{y}$$

Figura 6.1: $y = x^2$ per x uniforme in $(-1, 1)$.**Esempio 6.13**

$$y = |x|$$

Se $y < 0$ l'equazione non ha soluzioni, e quindi $F_y(y) = 0$.

Se $y = 0$ si ha che $F_y(0) = P\{y \leq 0\} = P\{y = 0\} = P\{x = 0\} = F_x(0^+) - F_x(0^-)$.

Se $y > 0$ vale che $F_y(y) = P\{y \leq y\} = P\{-y \leq x \leq y\} = F_x(y) - F_x(-y^-)$.

Caso speciale: se x è uniforme nell'intervallo $[-2, 2]$ allora $F_x(x) = (1/2) + (x/4)$ e

$$F_y(y) = \frac{1}{2} + \frac{y}{4} - \left(\frac{1}{2} - \frac{y}{4} \right) = \frac{y}{2} \quad 0 \leq y \leq 2$$

e, ovviamente, $F_y(y) = 0$ per $y < 0$ e $F_y(y) = 1$ per $y > 2$. Quindi la v.a. y è uniforme in $[0, 2]$.

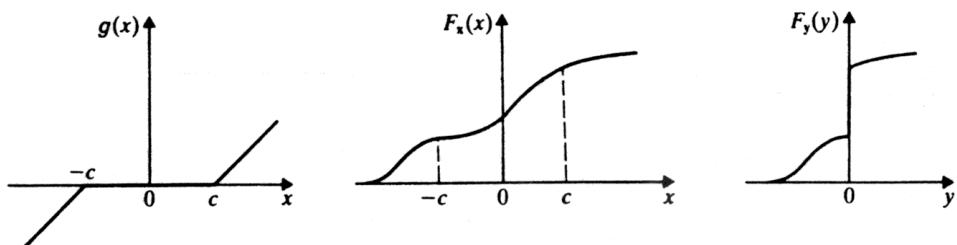
Caso 2. Se la funzione $g(x)$ è **costante in un intervallo** (x_0, x_1)

$$g(x) = y_1 \quad x_0 < x \leq x_1$$

allora

$$P\{y = y_1\} = P\{x_0 < x \leq x_1\} = F_x(x_1) - F_x(x_0) \quad (6.58)$$

e quindi $F_y(y)$ è discontinua in $y = y_1$ e presenta un salto di ampiezza $F_x(x_1) - F_x(x_0)$.

Figura 6.2: $g(x)$ costante per $-c \leq x \leq c$.

Esempio 6.14 Nella figura 6.2 è mostrata la funzione $g(x)$ tale che

$$g(x) = 0 \quad \text{per} \quad -c \leq x \leq c \quad \text{e} \quad g(x) = \begin{cases} x - c & x > c \\ x + c & x < -c \end{cases}$$

In questo caso $F_y(y)$ è discontinua per $y = 0$ e presenta un salto di ampiezza $F_x(c) - F_x(-c)$. Inoltre

$$\text{Se } y \geq 0 \quad \text{allora} \quad P\{y \leq y\} = P\{x \leq y + c\} = F_x(y + c)$$

$$\text{Se } y < 0 \quad \text{allora} \quad P\{y \leq y\} = P\{x \leq y - c\} = F_x(y - c)$$

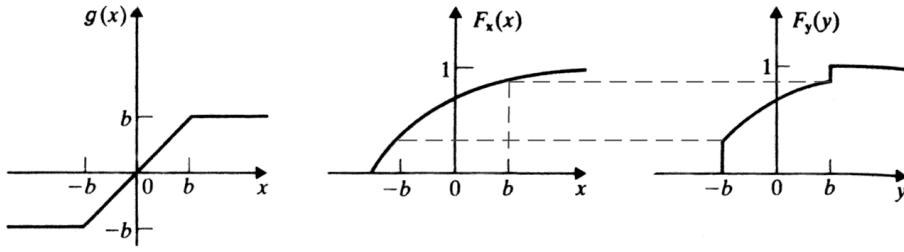


Figura 6.3: Limiter.

Esempio 6.15 Limiter La curva $g(x)$ di Fig. 6.3 è costante per $x \leq -b$ e $x \geq b$ e nell'intervallo $(-b, b)$ è una linea retta inclinata di 45 gradi. Quindi $g(x) = x$ per $-b \leq x \leq b$. Ponendo $\mathbf{y} = g(\mathbf{x})$ segue che $F_y(y)$ è discontinua per $y = g(-b) = -b$ e $y = g(b) = b$.

$$\text{Se } y \geq b \quad \text{allora } F_y(y) = 1$$

$$\text{Se } -b \leq y < b \quad \text{allora } F_y(y) = F_x(y)$$

$$\text{Se } y < -b \quad \text{allora } F_y(y) = 0$$

Caso 3. Supponiamo che $g(x)$ sia una **funzione a gradini** (costante a tratti):

$$g(x) = g(x_i) = y_i \quad x_{i-1} < x \leq x_i$$

In questo caso la v.a. $\mathbf{y} = g(\mathbf{x})$ è di tipo discreto ed assume i valori y_i con

$$P\{\mathbf{y} = y_i\} = P\{x_{i-1} < \mathbf{x} \leq x_i\} = F_x(x_i) - F_x(x_{i-1})$$

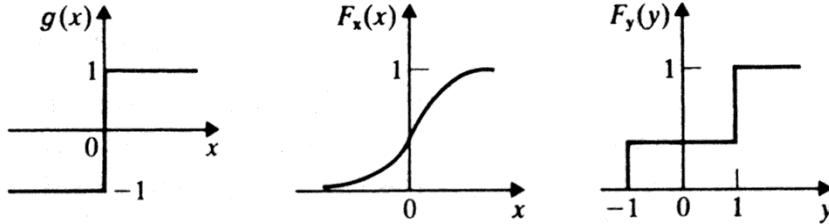


Figura 6.4: Hard Limiter.

Esempio 6.16 Hard Limiter Se

$$g(x) = \begin{cases} +1 & x > 0 \\ -1 & x \leq 0 \end{cases}$$

allora \mathbf{y} assume i valori ± 1 con

$$P\{\mathbf{y} = -1\} = P\{\mathbf{x} \leq 0\} = F_x(0)$$

$$P\{\mathbf{y} = +1\} = P\{\mathbf{x} > 0\} = 1 - F_x(0)$$

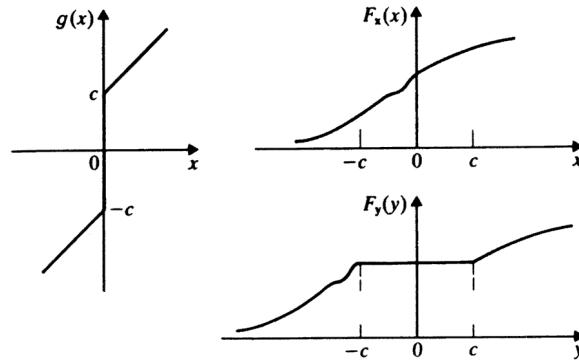
e quindi $F_y(y)$ è una funzione a gradini come in Fig. 6.4

Caso 4. Supponiamo che $g(x)$ sia **discontinua** in $x = x_0$ e tale che

$$g(x) < g(x_0^-) \quad \text{per } x < x_0 \quad \text{e} \quad g(x) > g(x_0^+) \quad \text{per } x > x_0$$

In questo caso, se y è fra $g(x_0^-)$ e $g(x_0^+)$, allora $g(x) < y$ per $x < x_0$. Quindi

$$F_y(y) = P\{\mathbf{x} \leq x_0\} = F_x(x_0) \quad \text{per } g(x_0^-) \leq y \leq g(x_0^+)$$

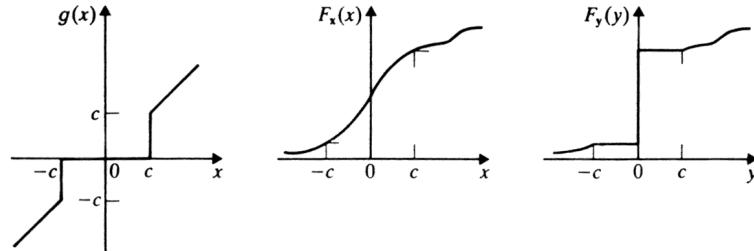
Figura 6.5: $g(x)$ discontinua in zero.

Esempio 6.17 Se $g(x)$ è discontinua in zero e tale che

$$g(x) = \begin{cases} x + c & x \geq 0 \\ x - c & x < 0 \end{cases}$$

(vedi Fig. 6.5) si ha che $g(x_0^-) = -c$ e $g(x_0^+) = c$ e $F_y(y) = F_x(0)$ per $|y| \geq c$. Inoltre

$$\begin{aligned} \text{Se } y \geq c & \text{ allora } g(x) \leq y \text{ per } x \leq y - c \text{ e quindi } F_y(y) = F_x(y - c) \\ \text{Se } -c \leq y \leq c & \text{ allora } g(x) \leq y \text{ per } x \leq 0 \text{ e quindi } F_y(y) = F_x(0) \\ \text{Se } y \leq -c & \text{ allora } g(x) \leq y \text{ per } x \leq y + c \text{ e quindi } F_y(y) = F_x(y + c) \end{aligned}$$

Figura 6.6: $g(x)$ discontinua in $\pm c$.

Esempio 6.18 La funzione $g(x)$ in Fig. 6.6 vale zero nell'intervallo $(-c, c)$ ed è discontinua per $x = \pm c$ con valori $g(c^+) = c$, $g(c^-) = 0$, $g(-c^-) = -c$ e, $g(-c^+) = 0$. Quindi $F_y(y)$ è discontinua in $y = 0$ ed è costante per $0 < y \leq c$ e $-c \leq y < 0$:

$$\begin{aligned} \text{Se } y \geq c & \text{ allora } g(x) \leq y \text{ per } x \leq y \text{ e quindi } F_y(y) = F_x(y) \\ \text{Se } 0 \leq y < c & \text{ allora } g(x) \leq y \text{ per } x < c \text{ e quindi } F_y(y) = F_x(c) \\ \text{Se } -c \leq y < 0 & \text{ allora } g(x) \leq y \text{ per } x \leq -c \text{ e quindi } F_y(y) = F_x(-c) \\ \text{Se } y < -c & \text{ allora } g(x) \leq y \text{ per } x \leq y \text{ e quindi } F_y(y) = F_x(y) \end{aligned}$$

Caso 5. Se la v.a. \mathbf{x} è di tipo discreto ed assume i valori x_k con probabilità p_k allora la v.a. $\mathbf{y} = g(\mathbf{x})$ è anch'essa di tipo discreto ed assume i valori $y_k = g(x_k)$.

Se $y_k = g(x)$ per un solo $x = x_k$ allora $P\{\mathbf{y} = y_k\} = P\{\mathbf{x} = x_k\} = p_k$. Se invece $y_k = g(x)$ per più valori x_j di x allora $P\{\mathbf{y} = y_k\}$ si ottiene sommando le probabilità delle x_j .

Esempio 6.19

$$\mathbf{y} = \mathbf{x}^2$$

- a) se \mathbf{x} assume i valori $1, 2, \dots, 6$ con probabilità $1/6$ allora \mathbf{y} assume i valori $1^2, 2^2, \dots, 6^2$ con probabilità $1/6$.
- b) se \mathbf{x} assume i valori $-2, -1, 0, 1, 2, 3$ con probabilità $1/6$ allora \mathbf{y} assume i valori $0, 1, 4, 9$ con probabilità rispettivamente $1/6, 1/3, 1/3, 1/6$.

6.9.2 Determinazione della densità di $\mathbf{y} = g(\mathbf{x})$

Vogliamo trovare la densità di probabilità di $\mathbf{y} = g(\mathbf{x})$ in termini della densità $f_x(x)$ di \mathbf{x} e di $g(x)$. Per prima cosa consideriamo che se l'insieme R dell'asse y non è nei valori assunti dalla funzione $g(x)$ qualsiasi sia x , allora la probabilità che $g(\mathbf{x})$ sia in R è zero e quindi $f_y(y) = 0$ per $y \in R$. E' allora sufficiente considerare i valori di y tali che, per qualche x , $g(x) = y$.

TEOREMA FONDAMENTALE Per trovare $f_y(y)$ per uno specifico y bisogna risolvere l'equazione $y = g(x)$. Chiamando x_n le sue radici reali

$$y = g(x_1) = \dots = g(x_n) = \dots \quad (6.59)$$

vedremo che

$$f_y(y) = \frac{f_x(x_1)}{|g'(x_1)|} + \dots + \frac{f_x(x_n)}{|g'(x_n)|} + \dots \quad (6.60)$$

dove $g'(x_k)$ è la derivata di $g(x)$ calcolata in x_k .

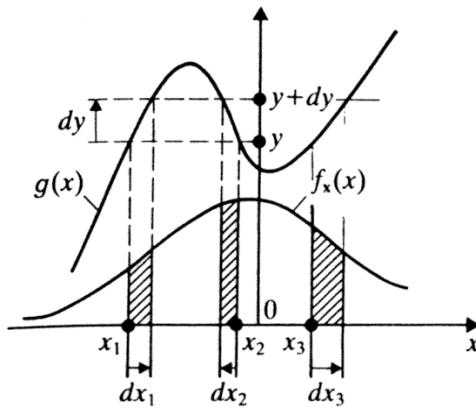


Figura 6.7: Illustrazione del teorema fondamentale.

Dimostrazione: assumiamo che, come in Fig. 6.7, l'equazione $y = g(x)$ abbia tre soluzioni. Come visto (eq. 6.11)

$$f_y(y) dy = P\{y < \mathbf{y} \leq y + dy\}$$

Bisogna quindi trovare l'insieme dei valori x tali che $y < g(x) \leq y + dy$ e la probabilità che \mathbf{x} sia in tale insieme. Come si vede in figura, l'insieme consiste dei tre intervalli

$$x_1 < x < x_1 + dx_1$$

$$x_2 + dx_2 < x < x_2$$

$$x_3 < x < x_3 + dx_3$$

dove $dx_1 > 0$ e $dx_3 > 0$, ma $dx_2 < 0$. Da ciò segue che

$$P\{y < \mathbf{y} < y + dy\} = P\{x_1 < \mathbf{x} < x_1 + dx_1\} + P\{x_2 + dx_2 < \mathbf{x} < x_2\} + P\{x_3 < \mathbf{x} < x_3 + dx_3\}$$

La parte destra è uguale all'area tratteggiata in figura 6.7. Poichè

$$\begin{aligned} P\{x_1 < \mathbf{x} < x_1 + dx_1\} &= f_x(x_1)dx_1 & dx_1 &= dy/g'(x_1) \\ P\{x_2 + dx_2 < \mathbf{x} < x_2\} &= f_x(x_2)dx_2 & dx_2 &= dy/g'(x_2) \\ P\{x_3 < \mathbf{x} < x_3 + dx_3\} &= f_x(x_3)dx_3 & dx_3 &= dy/g'(x_3) \end{aligned}$$

$$\tan \alpha = g'(x_1) = \frac{dy}{dx_1}$$

concludiamo che

$$f_y(y) dy = \frac{f_x(x_1)}{|g'(x_1)|} dy + \frac{f_x(x_2)}{|g'(x_2)|} dy + \frac{f_x(x_3)}{|g'(x_3)|} dy$$

e la 6.60 segue.

Notiamo infine che se $g(x) = y_1$ = costante per ogni x nell'intervallo (x_0, x_1) allora (vedi 6.58) $F_y(y)$ è discontinua per $y = y_1$ e perciò $f_y(y)$ contiene un impulso $\delta(y - y_1)$ di area $F_x(x_1) - F_x(x_0)$.

ESEMPI

Vediamo nel seguito alcune applicazioni delle equazioni 6.57 e 6.60

$$1. \quad \mathbf{y} = a\mathbf{x} + b \quad g'(x) = a$$

L'equazione $y = ax + b$ ha una sola soluzione $x = (y - b)/a$ per ogni y . Quindi

$$f_y(y) = \frac{1}{|a|} f_x\left(\frac{y-b}{a}\right) \quad (6.61)$$

Caso particolare: se \mathbf{x} è uniforme nell'intervallo (x_1, x_2) allora \mathbf{y} è uniforme in $(ax_1 + b, ax_2 + b)$.

$$2. \quad \mathbf{y} = \frac{1}{\mathbf{x}} \quad g'(x) = -(1/x^2) = -y^2$$

L'equazione $y = 1/x$ ha una sola soluzione $x = 1/y$ per ogni y . Quindi

$$f_y(y) = \frac{1}{y^2} f_x\left(\frac{1}{y}\right) \quad (6.62)$$

$$3. \quad \mathbf{y} = a\mathbf{x}^2 \quad a > 0 \quad g'(x) = 2ax \quad |g'(x)| = 2a\sqrt{y/a}$$

Se $y < 0$ l'equazione $y = ax^2$ non ha soluzioni reali e $f_y(y) = 0$. Se $y > 0$ ha due soluzioni

$$x_1 = \sqrt{\frac{y}{a}} \quad x_2 = -\sqrt{\frac{y}{a}}$$

e

$$f_y(y) = \frac{1}{2a\sqrt{y/a}} \left[f_x\left(\sqrt{\frac{y}{a}}\right) + f_x\left(-\sqrt{\frac{y}{a}}\right) \right] \quad y > 0 \quad (6.63)$$

Notiamo che $F_y(y) = 0$ per $y < 0$ e

$$F_y(y) = P\left\{-\sqrt{\frac{y}{a}} \leq \mathbf{x} \leq \sqrt{\frac{y}{a}}\right\} = F_x\left(\sqrt{\frac{y}{a}}\right) - F_x\left(-\sqrt{\frac{y}{a}}\right) \quad \text{se } y > 0 \quad (6.64)$$

$$4. \quad \mathbf{y} = a \sin(\mathbf{x} + \theta) \quad a > 0$$

Se $|y| > a$ l'equazione $y = a \sin(x + \theta)$ non ha soluzione e $f_y(y) = 0$.

Se $|y| \leq a$ l'equazione $y = a \sin(x + \theta)$ ha infinite soluzioni

$$x_n = \arcsin\left(\frac{y}{a}\right) - \theta \quad n = \dots, -1, 0, 1, \dots$$

La derivata di $g(x)$ calcolata nei punti soluzione vale $g'(x_n) = a \cos(x_n + \theta)$. Per ottenerla in funzione di y notiamo che $\cos^2(\alpha) + \sin^2(\alpha) = 1$, cioè $\cos(\alpha) = \sqrt{1 - \sin^2(\alpha)}$. Quindi

$$g'(x_n) = a \cos(x_n + \theta) = \sqrt{a^2 - y^2}$$

e

$$f_y(y) = \frac{1}{\sqrt{a^2 - y^2}} \sum_{n=-\infty}^{\infty} f_x(x_n) \quad |y| < a \quad (6.65)$$

Esempio 6.20 Un sasso lascia l'origine degli assi sotto l'influenza della forza di gravità e la sua velocità iniziale forma un angolo di φ con l'asse orizzontale. Come noto, il sasso colpisce il suolo a distanza

$$\mathbf{d} = \frac{v^2}{g} \sin(2\varphi)$$

dall'origine. Assumendo che φ sia una v.a. uniforme fra 0 e $\pi/2$ trovare la densità di \mathbf{d} .

Chiaramente $\mathbf{d} = a \sin \mathbf{x}$ con la v.a. $\mathbf{x} = 2\varphi$ uniforme fra 0 e π ed $a = v^2/g$. Poiché $0 < d < a$ l'equazione $d = a \sin x$ ha esattamente due soluzioni nell'intervallo $(0, \pi)$, $x_1 \in [0, \pi/2]$ e $x_2 \in [\pi/2, \pi]$, quindi

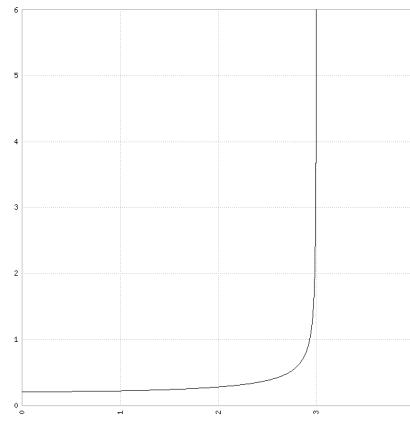
$$f_d(d) = \frac{1}{\sqrt{a^2 - d^2}} f_x(x_1) + \frac{1}{\sqrt{a^2 - d^2}} f_x(x_2)$$

e dato che $f_x(x) = 1/\pi$ nell'intervallo $[0, \pi]$

$$f_d(d) = \frac{1}{\sqrt{a^2 - d^2}} \left[\frac{1}{\pi} + \frac{1}{\pi} \right] = \frac{2}{\pi \sqrt{a^2 - d^2}}$$

per $0 < d < a$ e zero altrove.

In figura il caso $a = 3$.



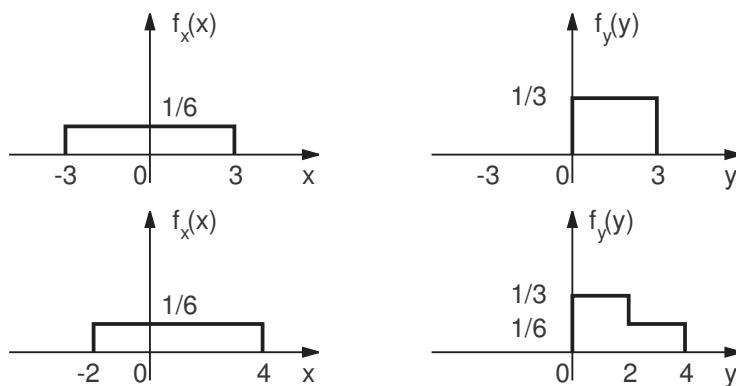
5. $y = |\mathbf{x}|$

Se $y < 0$ l'equazione $y = |\mathbf{x}|$ non ha soluzione e $f_y(y) = 0$. Per $y > 0$ ci sono due soluzioni, $x_1 = -y$ e $x_2 = y$. Inoltre $|g'(x)| = 1$, a parte per $x = 0$ dove non è definita. Quindi

$$f_y(y) = f_x(y) + f_x(-y) \quad y > 0$$

E, a essere precisi, bisogna aggiungere $f_y(0) = f_x(0)$ per $y = 0$, ovvero $f_y(0) = f_x(0)$.

Questo corrisponde a "ribaltare" la parte per $x < 0$ di $f_x(x)$ ed aggiungerla a quella esistente: nella figura che segue sono rappresentati graficamente due esempi di tale tecnica.



Come compito a casa potete provare a risolvere $y = a|\mathbf{x}|$ e $y = a|\mathbf{x}| + b$.

6.9.3 Problema inverso

Abbiamo visto come, a partire da una v.a. \mathbf{x} nota e da una funzione $g(x)$, si possa trovare la distribuzione e la densità della v.a. $\mathbf{y} = g(\mathbf{x})$. Consideriamo il problema inverso: data \mathbf{x} vogliamo trovare una funzione $g(x)$ tale che la distribuzione della v.a. $\mathbf{y} = g(\mathbf{x})$ sia uguale ad una specificata $F_y(y)$.

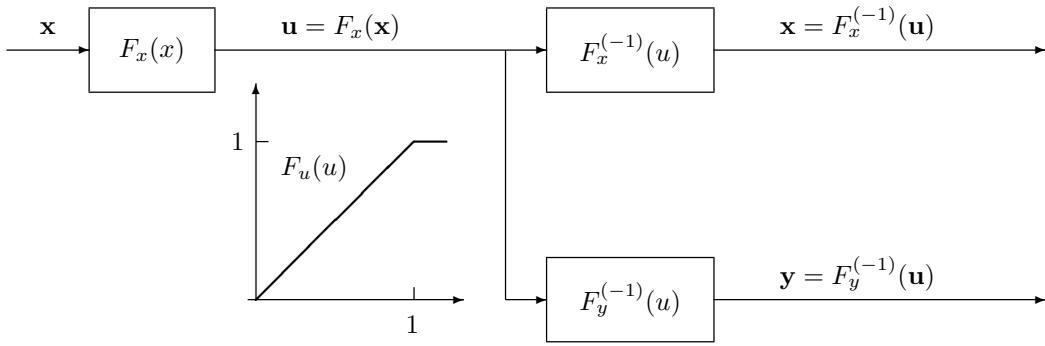
Da $F_x(x)$ ad una distribuzione uniforme Data una v.a. con distribuzione $F_x(x)$, vogliamo trovare una funzione $g(x)$ tale che la v.a. $\mathbf{u} = g(\mathbf{x})$ sia uniformemente distribuita nell'intervallo $(0, 1)$. La soluzione è che se

$$\mathbf{u} = F_x(\mathbf{x}) \quad \text{allora} \quad F_u(u) = u \quad \text{per} \quad 0 \leq u \leq 1 \quad (6.66)$$

Dimostrazione: supponiamo che x sia un numero arbitrario e che $u = F_x(x)$. Dalla monotonicità di $F_x(x)$ segue che $\mathbf{u} \leq u$ se $\mathbf{x} \leq x$. Quindi

$$F_u(u) = P\{\mathbf{u} \leq u\} = P\{\mathbf{x} \leq x\} = F_x(x) = u$$

La v.a. \mathbf{u} può essere considerata come l'uscita di un sistema nonlineare senza memoria (vedi figura) con ingresso \mathbf{x} e uscita $F_x(\mathbf{x})$. Quindi, se inviamo \mathbf{u} in ingresso ad un altro sistema che abbia risposta inversa otteniamo di nuovo $\mathbf{x} = F_x^{(-1)}(\mathbf{u})$.



Da una distribuzione uniforme a $F_y(y)$ Data una v.a. \mathbf{u} uniforme in $(0, 1)$ vogliamo trovare una funzione $g(u)$ tale che la distribuzione della v.a. $\mathbf{y} = g(\mathbf{u})$ sia una specificata funzione $F_y(y)$. Dalla figura precedente, notando come $F_x(x)$ sia del tutto arbitraria, ricaviamo che la funzione $g(u)$ cercata è l'inversa di $F_y(y)$

$$\text{Se } \mathbf{y} = F_y^{(-1)}(\mathbf{u}) \quad \text{allora} \quad P\{\mathbf{y} \leq y\} = F_y(y)$$

Da $F_x(x)$ a $F_y(y)$ Consideriamo infine il caso generale. Data $F_x(x)$ e $F_y(y)$, trovare $g(x)$ tale che la distribuzione di $\mathbf{y} = g(\mathbf{x})$ sia $F_y(y)$. Dalla figura precedente e da quanto visto finora è immediato concludere che

$$\text{Se } \mathbf{y} = F_y^{(-1)}(F_x(\mathbf{x})) \quad \text{allora} \quad P\{\mathbf{y} \leq y\} = F_y(y)$$

6.9.4 Media di $\mathbf{y} = g(\mathbf{x})$.

Data una v.a. \mathbf{x} ed una funzione $g(x)$, formiamo la v.a. $\mathbf{y} = g(\mathbf{x})$. Come visto, la media di \mathbf{y} è data da (vedi 6.13):

$$E\{\mathbf{y}\} = \bar{y} = \int_{-\infty}^{\infty} y f_y(y) dy \quad (6.67)$$

Sembra quindi che per trovare la media di \mathbf{y} sia necessario trovare la sua densità $f_y(y)$. Questo tuttavia non è necessario, infatti

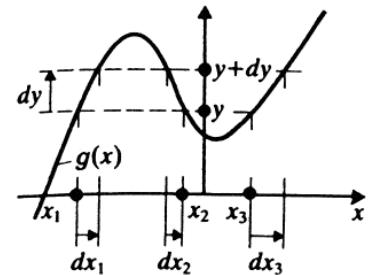
$$E\{\mathbf{y}\} = E\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(x) f_x(x) dx \quad (6.68)$$

Dimostrazione*:

Usiamo la curva in figura: poichè $y = g(x_1) = g(x_2) = g(x_3)$ si ottiene

$$f_y(y) dy = f_x(x_1) dx_1 + f_x(x_2) dx_2 + f_x(x_3) dx_3$$

e moltiplicando per y



$$y f_y(y) dy = g(x_1) f_x(x_1) dx_1 + g(x_2) f_x(x_2) dx_2 + g(x_3) f_x(x_3) dx_3$$

Quindi ad ogni differenziale in eq. 6.67 corrispondono uno o più differenziali nell'eq. 6.68. Mentre dy copre l'asse y i corrispondenti dx non si sovrappongono e coprono l'asse x . Quindi gli integrali 6.67 e 6.68 sono uguali.

Se \mathbf{x} è di tipo discreto si ottiene

$$E\{\mathbf{y}\} = E\{g(\mathbf{x})\} = \sum_i g(x_i) P\{\mathbf{x} = x_i\} \quad (6.69)$$

Analoghe relazioni si ottengono per valore quadratico medio, varianza e momenti.

6.10 Due variabili aleatorie

Siano date due v.a. \mathbf{x} e \mathbf{y} , definite sullo stesso esperimento. Il problema che si affronta in questa sezione è la determinazione delle loro relazioni: conoscere le statistiche di ciascuna di esse non è sufficiente, visto che manca l'informazione di quali sono i legami fra di esse. Spesso sono proprio i legami fra diverse v.a. ad essere interessanti: ad esempio se \mathbf{x} rappresenta il reddito di una persona e \mathbf{y} la sua età.

Per definire le relazioni fra due v.a. è necessario determinare le loro statistiche *congiunte*, ovvero la probabilità che il punto (x, y) sia in una specificata regione del piano x, y . Per evitare confusione le statistiche di ciascuna v.a. sono dette statistiche *marginali* e, come già detto, non sono sufficienti a determinare le statistiche congiunte.

Nel seguito vedremo che le statistiche congiunte delle v.a. \mathbf{x} e \mathbf{y} sono completamente determinate se la probabilità dell'evento

$$\{\mathbf{x} \leq x\} \cap \{\mathbf{y} \leq y\} = \{\mathbf{x} \leq x, \mathbf{y} \leq y\}$$

è nota per ogni x e y .

6.10.1 Distribuzione e densità congiunte

La distribuzione congiunta $F_{xy}(x, y)$ (o $F(x, y)$) di due v.a. \mathbf{x} e \mathbf{y} è definita come la probabilità dell'evento $\{\mathbf{x} \leq x, \mathbf{y} \leq y\}$, ovvero

$$F_{xy}(x, y) = P\{\mathbf{x} \leq x, \mathbf{y} \leq y\} \quad (6.70)$$

Valgono le seguenti proprietà:

$$F(-\infty, y) = 0 \quad F(x, -\infty) = 0 \quad F(\infty, \infty) = 1$$

Dalla definizione 6.70 seguono inoltre le seguenti identità:

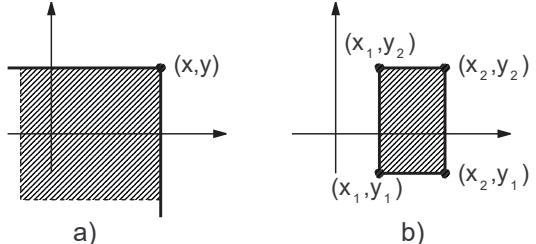
$$P\{x_1 < \mathbf{x} \leq x_2, \mathbf{y} \leq y\} = F(x_2, y) - F(x_1, y)$$

$$P\{\mathbf{x} \leq x, y_1 < \mathbf{y} \leq y_2\} = F(x, y_2) - F(x, y_1)$$

$$P\{x_1 < \mathbf{x} \leq x_2, y_1 < \mathbf{y} \leq y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)$$

Per capirne il significato si può fare riferimento alla figura: in a) $F(x, y)$ corrisponde alla probabilità che \mathbf{x} ed \mathbf{y} siano nell'area infinita, linee estreme comprese, tratteggiata. In b) si è evidenziato l'evento

$$\{x_1 < \mathbf{x} \leq x_2, y_1 < \mathbf{y} \leq y_2\}$$



La densità di probabilità congiunta di due v.a. \mathbf{x} e \mathbf{y} è, per definizione,

$$f_{xy}(x, y) = \frac{\delta^2 F_{xy}(x, y)}{\delta x \delta y} \quad (6.71)$$

Quindi

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(\alpha, \beta) d\alpha d\beta \quad (6.72)$$

Data una arbitraria regione D del piano x, y vale che

$$P\{(x, y) \in D\} = \iint_D f(x, y) dx dy \quad (6.73)$$

inoltre

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1 \quad \text{e} \quad f(x, y) \geq 0 \quad \forall x, y$$

Statistiche marginali

Nello studio di più v.a. le statistiche di ciascuna sono dette marginali. Le statistiche marginali di \mathbf{x} e \mathbf{y} possono essere espresse in termini delle statistiche congiunte $F(x, y)$ e $f(x, y)$:

$$F_x(x) = F(x, \infty) \quad F_y(y) = F(\infty, y) \quad (6.74)$$

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad f_y(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad (6.75)$$

V.a. congiuntamente gaussiane

Due v.a. \mathbf{x} e \mathbf{y} sono congiuntamente gaussiane (o normali) se

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)} \left[\frac{(x-\eta_1)^2}{\sigma_1^2} - 2r\frac{(x-\eta_1)(y-\eta_2)}{\sigma_1\sigma_2} + \frac{(y-\eta_2)^2}{\sigma_2^2} \right]} \quad (6.76)$$

con $|r| \leq 1$. La funzione $f(x, y)$ sopra verrà indicata $N(\eta_1, \eta_2; \sigma_1, \sigma_2; r)$: η_1 e η_2 sono i valori medi di \mathbf{x} e \mathbf{y} , mentre σ_1 e σ_2 sono le rispettive deviazioni standard. Il coefficiente di correlazione r verrà spiegato nel seguito.

Le statistiche marginali sono gaussiane e non dipendono da r :

$$f_x(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-\eta_1)^2}{2\sigma_1^2}} \quad f_y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(y-\eta_2)^2}{2\sigma_2^2}} \quad (6.77)$$

La dimostrazione, non richiesta, è su [Pap91] pag. 128. Quindi se due v.a. sono congiuntamente gaussiane lo sono anche marginalmente, mentre l'inverso non è, in generale, vero.

Indipendenza

Due v.a. \mathbf{x} e \mathbf{y} sono dette (statisticamente) indipendenti se gli eventi $\{\mathbf{x} \in \mathcal{A}\}$ e $\{\mathbf{y} \in \mathcal{B}\}$ sono indipendenti, ovvero, come visto, se

$$P(\{\mathbf{x} \in \mathcal{A}, \mathbf{y} \in \mathcal{B}\}) = P(\{\mathbf{x} \in \mathcal{A}\})P(\{\mathbf{y} \in \mathcal{B}\})$$

Se ciò è vero allora

$$F_{xy}(x, y) = F_x(x)F_y(y) \quad \text{e} \quad f_{xy}(x, y) = f_x(x)f_y(y) \quad (6.78)$$

Entrambe le uguaglianze sono necessarie e sufficienti per l'indipendenza.

Quindi se due v.a. sono indipendenti allora la conoscenza delle statistiche marginali è sufficiente per determinare le statistiche congiunte.

Teorema: Se le v.a. \mathbf{x} e \mathbf{y} sono indipendenti, allora anche le v.a. $\mathbf{z} = g(\mathbf{x})$ e $\mathbf{u} = h(\mathbf{y})$ lo sono.

Simmetria circolare

La densità congiunta di due v.a. \mathbf{x} e \mathbf{y} è circolarmente simmetrica se dipende solo dalla distanza dall'origine:

$$f_{xy}(x, y) = g(\rho) \quad \rho = \sqrt{x^2 + y^2} \quad (6.79)$$

Si può dimostrare che, se due v.a. sono a densità circolarmente simmetrica ed indipendenti, allora esse sono gaussiane con media nulla ed uguale varianza.

6.10.2 Funzioni di due variabili aleatorie

Date due v.a. \mathbf{x} e \mathbf{y} ed una funzione $g(x, y)$ formiamo la v.a.

$$\mathbf{z} = g(\mathbf{x}, \mathbf{y}) \quad (6.80)$$

Vogliamo ora esprimere le statistiche di \mathbf{z} in termini della funzione $g()$ e delle statistiche congiunte di \mathbf{x} e \mathbf{y} .

Se z è un numero dato, chiameremo D_z la regione del piano x, y tale che $g(\mathbf{x}, \mathbf{y}) \leq z$ (D_z può non essere semplicemente connessa). Chiaramente:

$$\{\mathbf{z} \leq z\} = \{g(\mathbf{x}, \mathbf{y}) \leq z\} = \{(\mathbf{x}, \mathbf{y}) \in D_z\}$$

quindi, vedi eq. 6.73,

$$F_z(z) = P(\{\mathbf{z} \leq z\}) = P\{(x, y) \in D_z\} = \iint_{D_z} f(x, y) dx dy \quad (6.81)$$

Per trovare $F_z(z)$ bisogna quindi trovare la regione D_z per ogni z e valutare l'integrale nell'eq. 6.81. La densità $f_z(z)$ si valuta in modo analogo

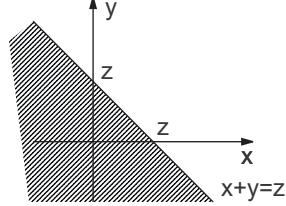
$$f_z(z) dz = P(\{z < \mathbf{z} \leq z + dz\}) = \iint_{\Delta D_z} f(x, y) dx dy \quad (6.82)$$

o per mezzo di $f_z(z) = dF_z(z)/dz$.

Esempio 6.21

Sia $\mathbf{z} = \mathbf{x} + \mathbf{y}$. La regione D_z del piano x, y tale che $x + y \leq z$ è tratteggiata in figura, alla sinistra della linea $x + y = z$. Quindi

$$F_z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-\alpha} f(x, \alpha) dx d\alpha$$



Possiamo ricavare $f_z(z)$ sia differenziando $F_z(z)$ che direttamente dall'eq. 6.82, notando che la regione ΔD_z tale che $z < x + y \leq z + dz$ è la striscia diagonale infinita limitata dalle linee $x + y = z$ e $x + y = z + dz$. Le coordinate di un punto in questa regione sono $(z - y, y)$ e l'area differenziale vale $dy dz$. Quindi

$$f_z(z) dz = \int_{-\infty}^{\infty} f_{xy}(z - y, y) dy dz \quad (6.83)$$

e

$$f_z(z) = \int_{-\infty}^{\infty} f_{xy}(z - y, y) dy \quad (6.84)$$

Indipendenza e convoluzione Se due v.a. \mathbf{x} e \mathbf{y} sono indipendenti allora, come visto, $f(x, y) = f_x(x)f_y(y)$. Utilizzando l'esempio sopra possiamo quindi trovare che, se $\mathbf{z} = \mathbf{x} + \mathbf{y}$,

$$f_z(z) = \int_{-\infty}^{\infty} f_x(z - y) f_y(y) dy = f_x(z) * f_y(z) \quad (6.85)$$

In conclusione:

se due v.a. \mathbf{x} e \mathbf{y} sono indipendenti la densità della loro somma è pari alla convoluzione delle loro densità.

Media di $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$.

La media della v.a. $\mathbf{z} = g(\mathbf{x}, \mathbf{y})$ si può ovviamente ottenere mediante

$$E\{\mathbf{z}\} = \int_{-\infty}^{\infty} z f_z(z) dz \quad (6.86)$$

E' anche possibile esprimere la media di \mathbf{z} in funzione di $g()$ e delle statistiche congiunte di \mathbf{x} e \mathbf{y} : analogamente a quanto visto nel caso delle funzioni di una v.a. si può dimostrare che

$$E\{\mathbf{z}\} = E\{g(\mathbf{x}, \mathbf{y})\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{xy}(x, y) dx dy \quad (6.87)$$

da questa segue che

$$E \left\{ \sum_{k=1}^n a_k g_k(\mathbf{x}, \mathbf{y}) \right\} = \sum_{k=1}^n a_k E \{g_k(\mathbf{x}, \mathbf{y})\} \quad (6.88)$$

in particolare vale quindi

$$E\{\mathbf{x} + \mathbf{y}\} = E\{\mathbf{x}\} + E\{\mathbf{y}\} \quad (6.89)$$

mentre, in generale, $E\{\mathbf{x}\mathbf{y}\} \neq E\{\mathbf{x}\}E\{\mathbf{y}\}$.

6.10.3 Covarianza e correlazione

Le statistiche congiunte contengono l'informazione completa dei legami fra due v.a. A volte è utile avere informazioni più sintetiche, anche se incomplete, di tali legami. Questo si ottiene mediante due numeri, il coefficiente di correlazione e la covarianza. Essi sono del tutto equivalenti, nel senso che noto uno l'altro si calcola immediatamente (sono lo stesso numero, a parte per un fattore di scala positivo). Entrambi i numeri possono essere negativi, nulli o positivi e ovviamente hanno lo stesso segno. Se sono positivi si dice che le due v.a. sono correlate positivamente, se sono nulli si dice che sono scorrelate e se sono negativi le v.a. sono correlate negativamente. Ad esempio, se la v.a. \mathbf{x} rappresenta l'altezza degli individui di una popolazione e la v.a. \mathbf{y} il peso, ci si può aspettare una correlazione positiva; se \mathbf{x} rappresenta l'altezza e \mathbf{y} il reddito verosimilmente la covarianza è nulla e se \mathbf{x} rappresenta la percentuale di tempo passata davanti alla TV e \mathbf{y} il quoziente intellettuale probabilmente esiste correlazione negativa (ipotesi personale dell'autore).

La covarianza C o C_{xy} di due v.a. \mathbf{x} e \mathbf{y} è, per definizione

$$C_{xy} = E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})\} \quad (6.90)$$

che può anche essere scritta, in modo equivalente,

$$C_{xy} = E\{\mathbf{xy}\} - E\{\mathbf{x}\}E\{\mathbf{y}\} \quad (6.91)$$

Dalla 6.90 si deduce che se \mathbf{x} e \mathbf{y} tendono ad essere entrambe dallo stesso lato rispetto al valore medio (se $\mathbf{x} > \bar{\mathbf{x}}$ ho probabilità maggiore di 0.5 che $\mathbf{y} > \bar{\mathbf{y}}$) allora la covarianza è positiva.

Normalizzando la covarianza si ottiene il coefficiente di correlazione r o r_{xy} (o ρ) delle v.a. \mathbf{x} e \mathbf{y} : esso è, per definizione

$$r_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \quad (6.92)$$

E' sempre vero che $|r| \leq 1$ e $|C| \leq \sigma_x \sigma_y$.

Se $r_{xy} = 1$ ho quindi la massima correlazione positiva, se vale -1 la massima negativa e se vale zero sono scorrelate.

Due v.a. sono dette scorrelate se la loro covarianza è nulla. Le seguenti sono definizioni equivalenti di scorrelazione di due v.a.:

$$C_{xy} = 0 \quad r_{xy} = 0 \quad E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$$

Attenzione che se due v.a. sono scorrelate questo non significa che non siano legate.

Esempio 6.22 Sia \mathbf{x} una v.a. uniforme in $[-1, 1]$ e sia $\mathbf{y} = \mathbf{x}^2$.

\mathbf{x} e \mathbf{y} sono scorrelate, nonostante abbiano dipendenza funzionale (che è la forma più forte di dipendenza). La dimostrazione è semplice:

$$E\{\mathbf{x}\} = 0 \quad E\{\mathbf{y}\} = E\{\mathbf{x}^2\} = \int_{-1}^1 x^2 \frac{1}{2} dx = \frac{1}{3} \quad E\{\mathbf{xy}\} = E\{\mathbf{x}^3\} = \int_{-1}^1 x^3 \frac{1}{2} dx = 0$$

quindi $E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$.

Da notare che le v.a. \mathbf{x} , \mathbf{y} hanno la stessa covarianza e lo stesso coefficiente di correlazione delle v.a. $\mathbf{x} - \bar{\mathbf{x}}$, $\mathbf{y} - \bar{\mathbf{y}}$.

Due v.a. sono dette ortogonal se $E\{\mathbf{xy}\} = 0$ e si indica $\mathbf{x} \perp \mathbf{y}$.

a) Se \mathbf{x} e \mathbf{y} sono scorrelate allora $\mathbf{x} - \bar{\mathbf{x}} \perp \mathbf{y} - \bar{\mathbf{y}}$.

b) Se \mathbf{x} e \mathbf{y} sono scorrelate e $\bar{\mathbf{x}} = 0$ oppure $\bar{\mathbf{y}} = 0$ allora $\mathbf{x} \perp \mathbf{y}$.

Da quanto visto finora ci si può aspettare che due v.a. indipendenti siano anche scorrelate: ciò è corretto ed è l'enunciato del teorema seguente:

Teorema: se due v.a. sono indipendenti, cioè se $f(x, y) = f_x(x)f_y(y)$, allora sono scorrelate.

La dimostrazione si ottiene dimostrando che $E\{\mathbf{xy}\} = E\{\mathbf{x}\}E\{\mathbf{y}\}$, mediante

$$E\{\mathbf{xy}\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_x(x) f_y(y) dx dy = \int_{-\infty}^{\infty} xf_x(x) dx \int_{-\infty}^{\infty} yf_y(y) dy = E\{\mathbf{x}\}E\{\mathbf{y}\}$$

Ovviamente la scorrelazione non implica l'indipendenza (vedi esempio 6.22).

NOTA: se due v.a. sono scorrelate esse non sono, in generale, indipendenti. Se tuttavia due v.a. congiuntamente gaussiane sono scorrelate allora sono anche indipendenti. Per dimostrarlo basta considerare l'eq. 6.76 imponendo $r = 0$ per ottenere facilmente che $f(x, y) = f_x(x)f_y(y)$.

Varianza della somma di due v.a.

Se $\mathbf{z} = \mathbf{x} + \mathbf{y}$ allora, vedi eq. 6.89, $\bar{z} = E\{\mathbf{z}\} = E\{\mathbf{x}\} + E\{\mathbf{y}\} = \bar{x} + \bar{y}$. Quindi

$$\sigma_z^2 = E\{(\mathbf{z} - \bar{z})^2\} = E\{[(\mathbf{x} - \bar{x}) + (\mathbf{y} - \bar{y})]^2\}$$

espandendo il quadrato e usando la linearità dell'operatore di media si ottiene

$$\sigma_z^2 = E\{(\mathbf{x} - \bar{x})^2\} + 2E\{(\mathbf{x} - \bar{x})(\mathbf{y} - \bar{y})\} + E\{(\mathbf{y} - \bar{y})^2\}$$

e, dalle definizioni di covarianza e coefficiente di correlazione,

$$\sigma_z^2 = \sigma_x^2 + 2C_{xy} + \sigma_y^2 = \sigma_x^2 + 2r\sigma_x\sigma_y + \sigma_y^2 \quad (6.93)$$

Quindi se due v.a. \mathbf{x} e \mathbf{y} sono scorrelate ($r = 0$) vale che

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 \quad (6.94)$$

Riassumendo: la media della somma di due v.a. è sempre la somma delle medie mentre la varianza è la somma delle varianze solo se le v.a. sono scorrelate (o, a maggior ragione, indipendenti).

6.10.4 Distribuzioni condizionate

In questa sezione ci occupiamo di definire distribuzione e densità condizionate di una v.a. quando la condizione riguarda una seconda v.a. e sono note le statistiche congiunte delle due v.a. Da tener presente quanto già visto sulle distribuzioni condizionate in sez. 6.7.

Le distribuzioni condizionate possono essere espresse in termini di probabilità condizionate:

$$F_z(z/\mathcal{U}) = P\{\mathbf{z} \leq z/\mathcal{U}\} = \frac{P\{\mathbf{z} \leq z, \mathcal{U}\}}{P(\mathcal{U})} \quad (6.95)$$

$$F_{zw}(z, w/\mathcal{U}) = P\{\mathbf{z} \leq z, \mathbf{w} \leq w/\mathcal{U}\} = \frac{P\{\mathbf{z} \leq z, \mathbf{w} \leq w, \mathcal{U}\}}{P(\mathcal{U})} \quad (6.96)$$

Le densità corrispondenti si ottengono per differenziazione.

Esempio 6.23 Vogliamo trovare la distribuzione condizionata $F_y(y/\mathbf{x} \leq x)$ e la densità condizionata $f_y(y/\mathbf{x} \leq x)$. Ponendo $\mathcal{U} = \{\mathbf{x} \leq x\}$ nell'eq. 6.95 si ottiene

$$F_y(y/\mathbf{x} \leq x) = \frac{P\{\mathbf{x} \leq x, \mathbf{y} \leq y\}}{P\{\mathbf{x} \leq x\}} = \frac{F(x, y)}{F_x(x)} \quad (6.97)$$

$$f_y(y/\mathbf{x} \leq x) = \frac{\delta F_y(y/\mathbf{x} \leq x)}{\delta y} = \frac{\delta F(x, y)/\delta y}{F_x(x)} \quad (6.98)$$

Esempio 6.24 Vogliamo trovare la distribuzione condizionata $F_{xy}(x, y/\mathcal{U})$ nel caso in cui $\mathcal{U} = \{x_1 < \mathbf{x} \leq x_2\}$. Essa vale

$$F(x, y/x_1 < \mathbf{x} \leq x_2) = \frac{P\{\mathbf{x} \leq x, \mathbf{y} \leq y, x_1 < \mathbf{x} \leq x_2\}}{P\{x_1 < \mathbf{x} \leq x_2\}} = \\ = \begin{cases} 0 & x < x_1 \\ \frac{F(x_2, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)} & x > x_2 \\ \frac{F(x, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)} & x_1 < x \leq x_2 \end{cases} \quad (6.99)$$

Poichè $f = \delta^2 F / \delta x \delta y$ si ottiene

$$f(x, y/x_1 < \mathbf{x} \leq x_2) = \frac{f(x, y)}{F_x(x_2) - F_x(x_1)} \quad x_1 < \mathbf{x} \leq x_2 \quad (6.100)$$

e zero altrove.

Esempio 6.25 Determinare la densità condizionata di \mathbf{y} assumendo $\mathbf{x} = x$ è un po' più difficile, visto che non è possibile usare direttamente l'eq. 6.95 dato che, in generale, l'evento $\{\mathbf{x} = x\}$ ha probabilità zero. Il risultato si ottiene mediante passaggio al limite. Definiamo l'evento $\mathcal{U} = \{x_1 < \mathbf{x} \leq x_2\}$ ed applichiamo l'eq. 6.95:

$$F_y(y/x_1 < \mathbf{x} \leq x_2) = \frac{P\{\mathbf{y} \leq y, x_1 < \mathbf{x} \leq x_2\}}{P\{x_1 < \mathbf{x} \leq x_2\}} = \frac{F(x_2, y) - F(x_1, y)}{F_x(x_2) - F_x(x_1)} \quad (6.101)$$

Differenziando rispetto a y otteniamo

$$f_y(y/x_1 < \mathbf{x} \leq x_2) = \frac{\int_{x_1}^{x_2} f(x, y) dx}{F_x(x_2) - F_x(x_1)} \quad (6.102)$$

Abbiamo qui applicato l'eq. 6.72, specializzata in

$$\frac{\delta F(x, y)}{\delta y} = \int_{-\infty}^x f(\alpha, y) d\alpha$$

Per trovare $f_y(y/\mathbf{x} = x)$ poniamo $x_1 = x$ e $x_2 = x + \Delta x$ nell'eq. 6.102

$$f_y(y/x < \mathbf{x} \leq x + \Delta x) = \frac{\int_x^{x+\Delta x} f(\alpha, y) d\alpha}{F_x(x + \Delta x) - F_x(x)} \approx \frac{f(x, y)\Delta x}{f_x(x)\Delta x} \quad (6.103)$$

Quindi

$$f_y(y/\mathbf{x} = x) = \lim_{\Delta x \rightarrow 0} f_y(y/x < \mathbf{x} \leq x + \Delta x) = \frac{f(x, y)}{f_x(x)} \quad (6.104)$$

Se non c'è rischio di ambiguità la funzione $f_y(y/\mathbf{x} = x)$ si può scrivere nella forma $f(y/x)$. Definendo in modo simile $f(x/y)$ possiamo quindi concludere che

$$f(y/x) = \frac{f(x, y)}{f_x(x)} \quad f(x/y) = \frac{f(x, y)}{f_y(y)} \quad (6.105)$$

Se le v.a. \mathbf{x} e \mathbf{y} sono indipendenti vale ovviamente che

$$f(x, y) = f(x)f(y) \quad f(y/x) = f(y) \quad f(x/y) = f(x) \quad (6.106)$$

NOTA: per uno specifico x , la funzione $f(y/x)$ è un *profilo* di $f(x, y)$, essendo l'intersezione della superficie $f(x, y)$ col piano $x = \text{costante}$. La densità condizionata $f(y/x)$ è l'equazione di tale curva intersezione, normalizzata per mezzo del fattore $1/f(x)$ in modo da avere area unitaria. Naturalmente è necessario che $f_x(x) \neq 0$.

Teorema di Bayes e della probabilità totale Dall'eq. 6.105 segue che

$$f(x/y) = \frac{f(y/x)f(x)}{f(y)} \quad (6.107)$$

Poichè

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad \text{e} \quad f(x, y) = f(y/x)f(x)$$

possiamo concludere che (probabilità totale)

$$f(y) = \int_{-\infty}^{\infty} f(y/x)f(x) dx \quad (6.108)$$

Inserendo quest'ultimo risultato nell'eq. 6.107 otteniamo il Teorema di Bayes per densità

$$f(x/y) = \frac{f(y/x)f(x)}{\int_{-\infty}^{\infty} f(y/x)f(x) dx} \quad (6.109)$$

6.10.5 Medie condizionate

Applicando l'eq. 6.68 alle densità condizionate otteniamo la media condizionata di $g(\mathbf{y})$:

$$E\{g(\mathbf{y})/\mathcal{U}\} = \int_{-\infty}^{\infty} g(y) f(y/\mathcal{U}) dy \quad (6.110)$$

Usando il passaggio ai limiti possiamo anche esprimere la media condizionata $E\{g(\mathbf{y})/x\}$ all'evento $\mathcal{U} = \{\mathbf{x} = x\}$. In particolare

$$\eta_{y/x} = E\{\mathbf{y}/x\} = \int_{-\infty}^{\infty} y f(y/x) dy = \int_{-\infty}^{\infty} y \frac{f(x,y)}{f(x)} dy = \frac{1}{f_x(x)} \int_{-\infty}^{\infty} y f(x,y) dy \quad (6.111)$$

è la media condizionata di \mathbf{y} assumendo $\mathbf{x} = x$ e

$$\sigma_{y/x}^2 = E\{(\mathbf{y} - \eta_{y/x})^2/x\} = \int_{-\infty}^{\infty} (y - \eta_{y/x})^2 f(y/x) dy \quad (6.112)$$

è la varianza condizionata.

Per un dato x , il primo integrale nell'eq. 6.111 è il baricentro (centro di gravità) della striscia verticale $(x, x + \delta x)$. Il luogo di tali punti, al variare di x da $-\infty$ a ∞ , è la funzione

$$\varphi(x) = \int_{-\infty}^{\infty} y f(y/x) dy = E\{\mathbf{y}/x\} \quad (6.113)$$

detta linea di regressione.

NOTA: se due v.a. sono legate funzionalmente, ovvero se $\mathbf{y} = g(\mathbf{x})$, allora

$$\varphi(x) = E\{\mathbf{y}/x\} = E\{g(x)/x\} = g(x)$$

Legge di GALTON

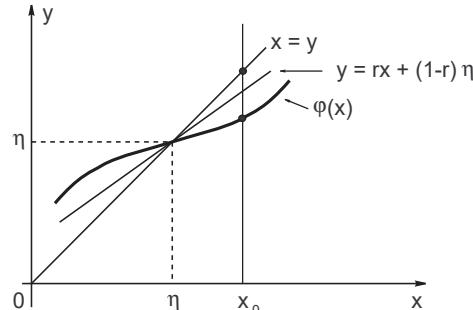
Il termine *regressione* si deve alla seguente osservazione di Sir Francis Galton: "I valori estremi di una popolazione regrediscono verso il loro valor medio".

Questo significa, ad esempio, che i figli di genitori alti sono (quando cresciuti) mediamente più bassi dei genitori, anche se mediamente più alti della media, ed i figli di genitori bassi sono più alti dei genitori. E' evidente che se fosse vero il contrario la popolazione si dividerebbe dopo un certo tempo in giganti e nani.

In termini statistici si esprime così: supponiamo che le v.a. \mathbf{x} e \mathbf{y} modellino rispettivamente l'altezza dei genitori e quella dei relativi figli.

Queste v.a. hanno la stessa media e varianza e sono positivamente correlate:

$$\eta_x = \eta_y = \eta \quad \sigma_x = \sigma_y = \sigma \quad r > 0$$



In base alla legge di Galton se $x > \eta$ allora $y < x$ e se $x < \eta$ allora $y > x$, ovvero:

$$E\{\mathbf{y}/x\} = \varphi(x) = \begin{cases} < x & \text{se } x > \eta \\ > x & \text{se } x < \eta \end{cases}$$

La linea di regressione $\varphi(x)$ è sotto la retta $y = x$ per $x > \eta$ e sopra se $x < \eta$.

Se le v.a. \mathbf{x} e \mathbf{y} sono congiuntamente gaussiane la linea di regressione è retta e vale

$$E\{\mathbf{y}/x\} = \varphi(x) = \eta_y + r \frac{\sigma_y}{\sigma_x} (x - \eta_x) \quad (6.114)$$

e se $\sigma_y = \sigma_x$ e $\eta_y = \eta_x$ si ottiene $\varphi(x) = (1 - r)\eta + rx$.

Per v.a. arbitrarie la legge di Galton non è ovviamente vera, ma il termine regressione è usato lo stesso per identificare ogni media condizionata.

6.10.6 Stima di v.a.

Il problema della stima è fondamentale nelle applicazioni della teoria della probabilità e richiederebbe uno spazio adeguato. Ci limiteremo qui ad introdurre l'argomento e le idee principali.

Stima mediante costante Come noto, conoscere la funzione di distribuzione di probabilità $F_y(y)$ (o la densità $f_y(y)$) significa determinare completamente le statistiche della v.a. y . Questo non significa, ovviamente, che conoscere $F_y(y)$ ci permette di predire il valore $y(\xi)$ di y in un esperimento ancora da compiere.

Supponiamo tuttavia di voler stimare l'incognita $y(\xi)$ per mezzo della costante \mathcal{C} . La conoscenza di $F_y(y)$ ci può guidare a scegliere il miglior valore di \mathcal{C} . Stimando y con \mathcal{C} , ad ogni prova si verificherà un errore di stima $y(\xi) - \mathcal{C}$: si tratta di scegliere \mathcal{C} in modo da minimizzare, in un qualche senso, l'errore medio. Un possibile criterio per scegliere \mathcal{C} è quello di chiedere che l'errore sia "vicino" a zero in una lunga serie di prove:

$$\frac{y(\xi_1) - \mathcal{C} + y(\xi_2) - \mathcal{C} + \dots + y(\xi_n) - \mathcal{C}}{n} \approx 0$$

Questo risultato si ottiene se \mathcal{C} è il valore medio di y .

Un diverso criterio per la scelta di \mathcal{C} è quello di minimizzare il valore medio di $|y(\xi) - \mathcal{C}|$. In questo caso il valore di \mathcal{C} risulta essere la mediana di y , ovvero il più piccolo numero m tale che $F_y(m) = 0.5$.

Minimo errore quadratico medio (MS) Un diverso criterio è quello di rendere minimo, in media, $|y(\xi) - \mathcal{C}|^2$. La stima al minimo errore quadratico medio (MS nel seguito, da *Mean Square error*) è di notevole importanza nella teoria della stima, fondamentalmente perchè, oltre a portare ad utili risultati, permette una certa semplificazione dei calcoli. Nel seguito ci occuperemo solo di stima MS. Come vedremo (eq. 6.117), il valore migliore per \mathcal{C} secondo la stima MS è ancora il valor medio di y .

Stima di una v.a. in termini di un'altra v.a. Supponiamo ora che, ad ogni prova, noi osserviamo il valore $x(\xi)$ della v.a. x . In base a questa osservazione vogliamo stimare y non con un valore costante unico \mathcal{C} , ma per mezzo di un numero che, ad ogni prova, dipende dal valore osservato $x(\xi)$. Cerchiamo quindi una stima di $y(\xi)$ per mezzo di una funzione $c(x)$ della v.a. x . Il problema che ne deriva è la scelta ottima della funzione $c()$. Si potrebbe pensare che conoscere $x(\xi)$ permetta di conoscere ξ e quindi $y(\xi)$. Questo non è in generale vero in quanto lo stesso valore $x(\xi) = x$ può essere assunto dalla v.a. x per più uscite dell'esperimento. Lo stesso numero $x(\xi)$ viene osservato per tutti gli ξ nell'insieme $\{x(\xi) = x\}$. I valori che y assume per gli ξ in $\{x(\xi) = x\}$ sono, in generale, diversi ed il miglior criterio nel senso MS è di scegliere $c(x)$ che sia il valor medio di y per quei valori. Ovvero $c(x)$ deve essere pari alla media condizionata di y assumendo che $x = x$: vediamo un esempio.

Esempio 6.26 Sia \mathcal{S} lo spazio di tutti i bambini di una certa comunità e sia y la v.a. che rappresenta l'altezza di ogni bambino. Una particolare uscita ξ è un bambino e $y(\xi)$ la sua altezza. Se vogliamo stimare y con un numero \mathcal{C} dobbiamo scegliere (criterio MS) la media di y . Supponiamo ora di pesare i bambini. Se di un bambino ξ conosciamo il peso (sia $x(\xi)$ la v.a. che rappresenta il peso di ξ) possiamo eseguire una stima più accurata della sua altezza, prendendo come valore di $c(x)$ l'altezza media dei bambini che pesano x . La stima ottima MS è quindi la media condizionata $E\{y/x\}$, essendo x il peso del bambino.

In termini di teoria delle probabilità, la stima MS di una v.a. y per mezzo di una costante \mathcal{C} si esprime in questo modo: trovare \mathcal{C} tale che il secondo momento (errore MS)

$$e = E\{(y - \mathcal{C})^2\} = \int_{-\infty}^{\infty} (y - \mathcal{C})^2 f(y) dy \quad (6.115)$$

della differenza (errore) $(y - \mathcal{C})$ sia minimo.

Ovviamente e dipende da \mathcal{C} ed è minimo se

$$\frac{de}{d\mathcal{C}} = - \int_{-\infty}^{\infty} 2(y - \mathcal{C}) f(y) dy = 0 \quad (6.116)$$

da cui

$$-2 \int_{-\infty}^{\infty} y f(y) dy + 2\mathcal{C} \int_{-\infty}^{\infty} f(y) dy = -2 \int_{-\infty}^{\infty} y f(y) dy + 2\mathcal{C} = 0$$

che ha per soluzione

$$\mathcal{C} = \int_{-\infty}^{\infty} y f(y) dy = E\{\mathbf{y}\} \quad (6.117)$$

che è il risultato che avevamo anticipato.

Stima MS non lineare

Vogliamo stimare \mathbf{y} con una funzione $c(\mathbf{x})$ della v.a. \mathbf{x} . Si tratta di trovare $c()$ che minimizza l'errore MS medio

$$e = E\{[\mathbf{y} - c(\mathbf{x})]^2\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [y - c(x)]^2 f(x, y) dx dy \quad (6.118)$$

La soluzione (che non dimostriamo) è

$$c(x) = E\{\mathbf{y}/x\} = \int_{-\infty}^{\infty} y f(y/x) dy = \varphi(x) \quad (6.119)$$

dove $\varphi(x)$ è la linea di regressione vista in sezione 6.10.5.

Se le v.a. \mathbf{y} e \mathbf{x} sono legate funzionalmente, ovvero se $\mathbf{y} = g(\mathbf{x})$, allora $E\{\mathbf{y}/x\} = g(x)$, $c(x) = g(x)$ e l'errore MS risultante è zero. Questo risultato è ovvio: se osservo \mathbf{x} e $\mathbf{y} = g(\mathbf{x})$ allora \mathbf{y} è nota esattamente.

Se le v.a. \mathbf{y} e \mathbf{x} sono indipendenti, allora $E\{\mathbf{y}/x\} = E\{\mathbf{y}\}$ = costante. Conoscere \mathbf{x} non ha effetto sulla stima di \mathbf{y} e $g(x)$ non dipende da \mathbf{x} , essendo costante.

Stima MS lineare

La soluzione del problema della stima MS non lineare è basata sulla conoscenza della linea di regressione $\varphi(x)$. Un problema più semplice è la stima MS lineare di \mathbf{y} in termini di \mathbf{x} : non è altrettanto buona, ma la soluzione è più facile da trovare.

La stima lineare MS della v.a. \mathbf{y} in termini di \mathbf{x} consiste nello stimare \mathbf{y} per mezzo della funzione lineare $c(x) = Ax + B$. Il problema è quello di trovare le costanti A e B in modo da minimizzare l'errore MS.

$$e = E\{[\mathbf{y} - (Ax + B)]^2\} \quad (6.120)$$

L'errore $e = e_m$ minimo si ottiene se (non lo dimostriamo):

$$A = \frac{\mu_{11}}{\mu_{20}} = \frac{r\sigma_y}{\sigma_x} \quad B = \bar{y} - A\bar{x} \quad (6.121)$$

Dove $\mu_{ik} = E\{(\mathbf{x} - \bar{\mathbf{x}})^k (\mathbf{y} - \bar{\mathbf{y}})^i\}$ sono i momenti centrali ed r è il coefficiente di correlazione.

L'errore MS minimo vale

$$e_m = \mu_{02} - \frac{\mu_{11}^2}{\mu_{20}} = \sigma_y^2(1 - r^2) \quad (6.122)$$

Se le v.a. \mathbf{x} e \mathbf{y} sono scorrelate, ovvero se $r = 0$, si ottiene $A = 0$ e $B = \bar{y}$, ovvero la stima non dipende da \mathbf{x} e $e_m = \sigma_y^2$.

Se le v.a. \mathbf{x} e \mathbf{y} sono tali che $r = 1$ si ottiene $e_m = 0$ e la stima è ottima.

Terminologia: $A\mathbf{x} + B$ è la stima lineare *non omogenea* di \mathbf{y} in termini di \mathbf{x} .

Se la v.a. \mathbf{y} viene stimata per mezzo di una linea retta $a\mathbf{x}$ passante per l'origine la stima è detta *omogenea*.

La v.a. \mathbf{x} è il *dato* della stima. la v.a. $\varepsilon = \mathbf{y} - (A\mathbf{x} + B)$ è l'*errore* della stima ed il numero $e = E\{\varepsilon^2\}$ è l'errore quadratico medio.

Nota: in generale, la stima non lineare $\varphi(x) = E\{\mathbf{y}/x\}$ non è una linea retta e l'errore MS medio risultante $E\{[\mathbf{y} - \varphi(x)]^2\}$ è minore dell'errore e_m della stima lineare $A\mathbf{x} + B$.

Tuttavia, se le v.a. \mathbf{x} e \mathbf{y} sono congiuntamente gaussiane allora (vedi eq. 6.114) la linea di regressione

$$\varphi(x) = \frac{r\sigma_y x}{\sigma_x} + \bar{y} - \frac{r\sigma_y \bar{x}}{\sigma_x} \quad (6.123)$$

è una retta e le due stime coincidono.

Il principio di ortogonalità*

Inserendo i valori di 6.121 nell'eq. 6.120 si può facilmente ottenere

$$E\{[\mathbf{y} - (\mathbf{Ax} + \mathbf{B})] \mathbf{x}\} = 0 \quad (6.124)$$

Un altro modo per ricavare il risultato precedente si ottiene notando che l'errore MS dell'eq. 6.120 è una funzione di A e B ed è minimo se $\delta e/\delta A = 0$ e $\delta e/\delta B = 0$. La prima equazione porta a

$$\frac{\delta e}{\delta A} = E\{[\mathbf{y} - (\mathbf{Ax} + \mathbf{B})](-\mathbf{x})\} = 0$$

da cui deriva nuovamente l'eq. 6.124. Questa equazione ci dice che la stima MS lineare ottima $\mathbf{Ax} + \mathbf{B}$ della v.a. \mathbf{y} è tale che l'errore di stima $[\mathbf{y} - (\mathbf{Ax} + \mathbf{B})]$ è ortogonale al dato \mathbf{x} (vedi sezione 6.10.3). Questo risultato è noto come principio di ortogonalità ed è fondamentale nella stima MS.

6.11 Teorema del limite centrale

Date n v.a. indipendenti \mathbf{x}_i formiamo la loro somma $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \dots + \mathbf{x}_n$.

La v.a. \mathbf{x} ha media $\bar{x} = \bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_n$ e varianza $\sigma_x^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \dots + \sigma_{x_n}^2$. Il teorema del limite centrale asserisce che, sotto certe condizioni generali, la distribuzione $F_x(x)$ di \mathbf{x} approssima, al crescere di n , una distribuzione gaussiana con pari media e varianza, ovvero che

$$F_x(x) \approx \mathbf{G}\left(\frac{x - \bar{x}}{\sigma_x}\right)$$

Inoltre, se le v.a. \mathbf{x}_i sono di tipo continuo, allora la densità $f_x(x)$ di \mathbf{x} approssima, al crescere di n , una densità gaussiana

$$f_x(x) \approx \frac{1}{\sigma_x \sqrt{2\pi}} e^{-(x-\bar{x})^2/2\sigma_x^2} \quad (6.125)$$

Questo importante teorema può essere espresso come limite: se $\mathbf{z} = (\mathbf{x} - \bar{\mathbf{x}})/\sigma_x$ allora

$$\lim_{n \rightarrow \infty} F_z(z) = \mathbf{G}(z) \quad \lim_{n \rightarrow \infty} f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

rispettivamente per i casi generale e continuo. La dimostrazione è omessa.

Il teorema del limite centrale deriva dalle proprietà della convoluzione, dato che $f_x(x) = f_{x_1}(x_1) * f_{x_2}(x_2) * \dots * f_{x_n}(x_n)$ e che la convoluzione di un grande numero di funzioni positive tende ad una funzione gaussiana (accettatelo senza dimostrazione).

Questo teorema è di fondamentale importanza nelle comunicazioni elettriche (e non solo). Il rumore presente nei sistemi di telecomunicazione deriva, in generale, dalla somma di tanti contributi che si possono considerare indipendenti (ad esempio il rumore termico dei vari componenti). La somma di tutti i contributi si può quindi, in base al teorema del limite centrale, considerare come un rumore gaussiano. Ciò giustifica l'ampio spazio dato alla trattazione del rumore gaussiano nella teoria delle comunicazioni elettriche.

Quanto l'approssimazione sia buona per un certo n dipende dalla forma delle densità delle v.a. \mathbf{x}_i . Se le \mathbf{x}_i sono i.i.d. (indipendenti ed identicamente distribuite) l'approssimazione è generalmente buona per $n \approx 30$. Se le densità sono inoltre "morbide", si ottengono buoni risultati già per n pari a 5 o 6.

Esempio 6.27 Le n v.a. \mathbf{x}_i sono i.i.d. ed uniformemente distribuite nell'intervallo $[0, T]$: la densità di probabilità $f_x(x_i)$ di ciascuna di esse è in fig. 6.8 [a]. Confronderemo la densità $f_x(x)$ della loro somma \mathbf{x} con l'approssimazione gaussiana dell'eq. 6.125 per i casi $n = 2$ e $n = 3$.

In questo problema i dati sono

$$\bar{x}_i = \frac{T}{2} \quad \sigma_i^2 = \frac{T^2}{12} \quad \bar{x} = n\bar{x}_i = n\frac{T}{2} \quad \sigma^2 = n\frac{T^2}{12}$$

Per $n = 2$, $f_x(x)$ è il triangolo ottenuto dalla convoluzione di due rettangoli, come mostrato in fig. 6.8 [b] (linea continua).

$$\bar{x} = T \quad \sigma^2 = \frac{T^2}{6} \quad f(x) \approx \frac{1}{T} \sqrt{\frac{3}{\pi}} e^{-3(x-T)^2/T^2}$$

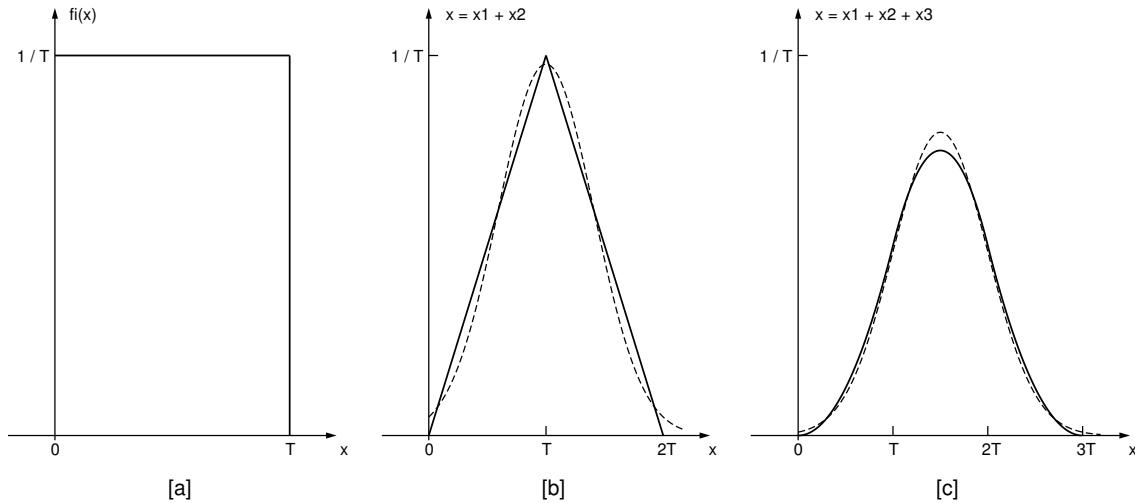


Figura 6.8: Esempio di somma di v.a. i.i.d. uniformi.

L'approssimazione gaussiana è in [b] linea tratteggiata.

Per $n = 3$, $f_x(x)$ consiste di tre tratti di parabola ottenuti convolvendo un triangolo con un rettangolo, come mostrato in fig. 6.8 [c] (linea continua).

$$\bar{x} = \frac{3T}{2} \quad \sigma^2 = \frac{T^2}{4} \quad f(x) \approx \frac{1}{T} \sqrt{\frac{2}{\pi}} e^{-2(x-1.5T)^2/T^2}$$

L'approssimazione gaussiana è in [c] linea tratteggiata.

Come facile osservare, l'approssimazione è buona persino per questi piccoli valori di n .

Capitolo 7

Processi aleatori

7.1 Definizioni

Abbiamo visto come una v.a. sia una regola per assegnare ad ogni risultato ξ di un esperimento \mathcal{S} un numero $\mathbf{x}(\xi)$.

Un processo aleatorio (p.a.) $x(t)$ è una regola per assegnare ad ogni ξ una funzione $x(t, \xi)$.

Quindi un p.a. è una famiglia di funzioni del tempo (o di altra variabile indipendente) che dipendono dal parametro ξ o, in modo equivalente, una funzione di t e ξ . Il dominio di ξ è l'insieme di tutte le uscite dell'esperimento ed il dominio di t è un insieme \mathcal{R} di numeri reali. Se \mathcal{R} è l'asse reale, allora $x(t)$ è un processo a tempo continuo. Se \mathcal{R} è l'insieme dei numeri interi, o un insieme che può essere messo in corrispondenza biunivoca con esso, allora $x(t)$ è un processo a tempo discreto. Inoltre possiamo distinguere fra processi aleatori a valori continui e a valori discreti.

I p.a. sono una ulteriore estensione della teoria delle probabilità, dopo le v.a., e hanno il fondamentale compito di permettere l'analisi di quei fenomeni aleatori che variano nel tempo. Se, ad esempio, si vuole progettare un sistema di trasmissione è necessario conoscere che segnali viaggeranno su di esso. Ovviamente essi sono ignoti (non posso sapere cosa trasmetterò in futuro) ma è di fondamentale importanza conoscere le loro caratteristiche medie. Allo stesso modo è necessario modellare il rumore che andrà a disturbare le trasmissioni, per mettere in atto le corrette procedure di difesa.

Useremo la notazione $x(t)$ per rappresentare un p.a. omettendo, come per le v.a., la sua dipendenza da ξ . Il p.a. $x(t)$ ha le seguenti interpretazioni:

1. è una famiglia (o insieme) di funzioni $x(t, \xi)$. In questa interpretazione t e ξ sono variabili.
2. è una singola funzione del tempo (detta *realizzazione* o *campione* del processo). In questo caso t è una variabile e ξ è fissata.
3. se t è fissato e ξ è variabile, allora $x(t)$ è una v.a. pari allo stato del processo all'istante t .
4. se t e ξ sono fissati allora $x(t)$ è un numero.

Un esempio fisico di un p.a. è il moto di particelle microscopiche in collisione con le molecole di un fluido (moto browniano). Il p.a. risultante $x(t)$ consiste nel moto di tutte le particelle. Una singola realizzazione (o campione) $x(t, \xi)$ di questo p.a. è il moto di una specifica particella.

Un altro esempio è la tensione $x(t) = r \cos(\omega t + \varphi)$ in uscita da un generatore con ampiezza r e fase φ casuali. In questo caso il p.a. $x(t)$ consiste di una famiglia di onde sinusoidali: una singola realizzazione è la funzione (a ξ fissato)

$$x(t, \xi) = r(\xi) \cos[\omega t + \varphi(\xi)]$$

Entrambi gli esempi sono p.a., ma esiste una fondamentale differenza fra di loro. Il primo esempio consiste in una famiglia di funzioni che non possono essere descritte in termini di un numero fissato di parametri. Inoltre il futuro di un campione $x(t, \xi)$ di $x(t)$ non può essere determinato conoscendone il passato. Infine, sotto certe condizioni, le sue statistiche possono essere determinate conoscendo una singola realizzazione (campione del p.a.). Il secondo esempio consiste di una famiglia di onde sinusoidali pure ed il p.a. risultante è completamente specificato dalle v.a. r e φ . Inoltre, se $x(t, \xi)$ è noto per $t \leq t_0$ è noto anche per $t > t_0$. Infine un singolo campione $x(t, \xi)$ di $x(t)$ non può specificare le proprietà dell'intero p.a. Questi concetti verranno approfonditi nel seguito del capitolo.

Uguaglianza

Due p.a. $x(t)$ e $y(t)$ sono uguali se i loro rispettivi campioni $x(t, \xi)$ e $y(t, \xi)$ sono identici per ogni ξ e t . In modo analogo l'uguaglianza $z(t) = x(t) + y(t)$ significa che $z(t, \xi) = x(t, \xi) + y(t, \xi)$ per ogni ξ e t . Derivate, integrali ed ogni altra operazione sui processi sono definite similmente in termini delle corrispondenti operazioni per ogni campione del p.a.

Le definizioni precedenti possono essere rilassate sostituendo l'uguaglianza in senso stretto con uguaglianze nel senso dell'errore quadratico medio (MSE).

Ad esempio due p.a. sono uguali nel senso MS se $E\{[x(t) - y(t)]^2\} = 0$ per ogni t .

Note storiche

Un p.a. in teoria dei fenomeni aleatori è un oggetto matematico solitamente definito come una collezione di variabili aleatorie. Storicamente le v.a. sono state associate a (o indicizzate mediante) un insieme di numeri, di solito visti come punti nel tempo, fornendo l'interpretazione di un p.a. che rappresenta valori numerici di un qualche sistema casualmente variante nel tempo, come la crescita di una popolazione di batteri, una corrente elettrica che fluttua a causa del rumore termico o il movimento di una molecola di gas.

Due esempi di p.a. sono il processo di Wiener, o moto Browniano, usato ad esempio da L. Bachelier¹ per studiare le variazioni di prezzo nella Borsa di Parigi, e il processo di Poisson, usato ad esempio da A.K. Erlang² per studiare il numero di chiamate telefoniche effettuate in un determinato periodo di tempo. Questi due p.a. sono considerati i più importanti e centrali nella teoria dei p.a. e sono stati scoperti ripetutamente ed in modo indipendente, sia prima che dopo i due esempi citati.

I p.a. che vedremo sono formati da v.a. indicizzate da numeri interi o intervalli di numeri reali. Se le v.a. sono indicizzate nel piano cartesiano o in qualche spazio Euclideo di maggiori dimensioni, la collezione di v.a. è di solito chiamata *random field* anzichè *random process* o *stochastic process*.

7.2 Statistiche dei processi aleatori

In questa sezione si defiscono le informazioni necessarie per una conoscenza completa delle statistiche di un p.a. e le informazioni parziali che spesso si utilizzano quando l'informazione completa non è disponibile o non è necessaria.

7.2.1 Distribuzioni e densità

Un p.a. è una infinità non contabile di v.a., una per ogni t . Fissando t , $x(t)$ è una v.a. con distribuzione

$$F(x, t) = P\{x(t) \leq x\} \quad (7.1)$$

Questa funzione dipende da t ed è la probabilità dell'evento $\{x(t) \leq x\}$ consistente di tutti i risultati ξ tali che, al tempo t , i campioni $x(t, \xi)$ del p.a. $x(t)$ non eccedano il numero x .

La funzione $F(x, t)$ è la **distribuzione del primo ordine** del p.a. $x(t)$. La sua derivata rispetto ad x ,

$$f(x, t) = \frac{\delta F(x, t)}{\delta x} \quad (7.2)$$

è la **densità del primo ordine** di $x(t)$.

La conoscenza delle statistiche del primo ordine equivale alla conoscenza delle infinite statistiche *marginali* delle v.a. che si ottengono fissando il tempo in $x(t)$. Come visto in sezione 6.10, la conoscenza delle statistiche marginali non è sufficiente a descrivere le relazioni fra più v.a.: servono anche le statistiche congiunte.

Se si considerano le infinite coppie di v.a. che si ottengono mediante tutti i possibili valori degli istanti temporali t_1 e t_2 , si possono definire le statistiche congiunte *a coppie* di tali v.a.

La **distribuzione del secondo ordine** del p.a. $x(t)$ è proprio la distribuzione congiunta

$$F(x_1, x_2; t_1, t_2) = P\{x(t_1) \leq x_1, x(t_2) \leq x_2\} \quad (7.3)$$

delle v.a. $x(t_1)$ e $x(t_2)$. La **densità del secondo ordine** corrispondente è

$$f(x_1, x_2; t_1, t_2) = \frac{\delta^2 F(x_1, x_2; t_1, t_2)}{\delta x_1 \delta x_2} \quad (7.4)$$

¹1870-1946, considerato il pioniere della matematica finanziaria. Nella sua tesi di dottorato del 1900 è stato il primo ad introdurre il p.a. ora chiamato moto Browniano

²1878-1929, pioniere della teoria delle code e della *traffic engineering*.

Si noti che deve valere:

$$\begin{aligned} F(x_1; t_1) &= F(x_1, \infty; t_1, t_2) \\ f(x_1; t_1) &= \int_{-\infty}^{\infty} f(x_1, x_2; t_1, t_2) dx_2 \end{aligned}$$

Come accadeva per le v.a., le statistiche (congiunte) del secondo ordine comprendono le statistiche (marginali) del primo ordine.

Considerando le v.a. a gruppi di tre, si può notare come le statistiche congiunte a coppie non siano sufficienti a descrivere in modo completo le relazioni fra di esse: servono anche le statistiche congiunte del terzo ordine. Proseguendo ci si accorge che, per definire completamente un p.a., servono le statistiche congiunte di *ogni ordine* (fino all'infinito): la **distribuzione di ordine** n di $x(t)$ è la distribuzione congiunta $F(x_1, \dots, x_n; t_1, \dots, t_n)$ delle v.a. $x(t_1), \dots, x(t_n)$.

Per la determinazione delle proprietà statistiche di un processo aleatorio è necessario conoscere la funzione $F(x_1, \dots, x_n; t_1, \dots, t_n)$ per ogni x_i, t_i ed n .

Ovviamente le **densità di ordine** n $f(x_1, \dots, x_n; t_1, \dots, t_n)$ sono un modo equivalente di descrivere l'informazione richiesta.

La quantità di informazione necessaria per descrivere completamente un generico p.a. è quindi di dimensioni formidabili: fortunatamente vi sono molte applicazioni per le quali è sufficiente una conoscenza parziale delle statistiche di $x(t)$. In particolare sono spesso utili i valori delle medie di $x(t)$ e $x^2(t)$: tali medie possono essere espresse in termini delle statistiche del primo ordine di $x(t)$.

7.2.2 Media

La media $\eta(t)$ del p.a. $x(t)$ è il valore atteso della v.a. $x(t)$:

$$\eta(t) = E\{x(t)\} = \int_{-\infty}^{\infty} xf(x, t) dx \quad (7.5)$$

In generale la media di un p.a. è quindi una funzione del tempo, ovvero vi è un valore medio per ciascun istante t . È inoltre evidente che la media dipende dalle sole statistiche del primo ordine.

7.2.3 Autocorrelazione e Autocovarianza

L'**autocorrelazione** $R(t_1, t_2)$ del p.a. $x(t)$ è il valore atteso del prodotto $x_1(t)x_2(t)$:

$$R(t_1, t_2) = E\{x(t_1)x(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (7.6)$$

Il valore di $R(t_1, t_2)$ sulla diagonale $t_1 = t_2 = t$ è il valore quadratico medio della v.a. $x(t)$ (a t fissato) ed è definito essere la **potenza media** del p.a. $x(t)$:

$$E\{x^2(t)\} = \int_{-\infty}^{\infty} x^2 f(x, t) dx = R(t, t) \geq 0 \quad (7.7)$$

Anche la potenza media è, in generale, una funzione del tempo, mentre l'autocorrelazione è una funzione delle due variabili temporali t_1 e t_2 .

L'autocorrelazione dipende dalle statistiche del secondo ordine, mentre la potenza media da quelle del primo.

L'**autocovarianza** $C(t_1, t_2)$ del p.a. $x(t)$ è la covarianza delle v.a. $x(t_1)$ e $x(t_2)$:

$$C(t_1, t_2) = R(t_1, t_2) - \eta(t_1)\eta(t_2) \quad (7.8)$$

ed il suo valore $C(t, t)$ sulla diagonale $t_1 = t_2 = t$ è la varianza della v.a. $x(t)$.

Autocorrelazione e autocovarianza possono essere indicate anche come

$$R(t_1, t_2) = R_x(t_1, t_2) = R_{xx}(t_1, t_2) \quad C(t_1, t_2) = C_x(t_1, t_2) = C_{xx}(t_1, t_2)$$

per evitare ambiguità in particolare con la correlazione incrociata o la covarianza incrociata definite nel seguito.

Il **coefficiente di correlazione** del p.a. $x(t)$ è il coefficiente di correlazione delle v.a. $x(t_1)$ e $x(t_2)$ al variare di t_1 e t_2 (vedi sezione 6.10.3):

$$r(t_1, t_2) = \frac{C(t_1, t_2)}{\sqrt{C(t_1, t_1) C(t_2, t_2)}} \quad (7.9)$$

ovviamente

$$|r(t_1, t_2)| \leq 1 \quad \text{e} \quad r(t, t) = 1 \quad (7.10)$$

Nell'analisi dei sistemi di telecomunicazione riveste grande importanza il calcolo delle potenze: ad esempio, il principale indice di qualità delle trasmissioni analogiche è il rapporto fra la potenza di segnale e quella di rumore a destinazione (entrambi modellati come p.a.).

Benchè la potenza media di un p.a. dipenda dalle sole statistiche del primo ordine, per il calcolo delle potenze in una catena di sistemi le statistiche del primo ordine non sono sufficienti. Ad esempio, se $x(t)$ è il processo in ingresso ad un sistema lineare e $y(t)$ la corrispondente uscita vedremo che, mentre la media di $y(t)$ dipende dalla sola media di $x(t)$, la potenza media di $y(t)$ può essere calcolata solo conoscendo $R_x(t_1, t_2)$ (non basta conoscere la potenza media $E\{x^2(t)\}$ di $x(t)$). Per un'ampia classe di applicazioni la conoscenza della funzione di autocorrelazione e del valor medio sono sufficienti per la corretta progettazione.

7.3 Esempi

In questa sezione si presentano esempi di p.a. notevoli e si calcolano alcune delle loro statistiche.

7.3.1 Processo sinusoidale

Il processo sinusoidale è descritto da una sinusoide a frequenza fissata: fra le varie realizzazioni variano l'ampiezza e la fase, che vengono modellate mediante opportune v.a.

Un p.a. sinusoidale è quindi definito da

$$x(t) = \mathbf{r} \cos(\omega t + \varphi) \quad (7.11)$$

dove le v.a. \mathbf{r} e φ sono assunte indipendenti. Solitamente si impone che φ sia uniformemente distribuita nell'intervallo $(-\pi, \pi)$, ovvero che la fase sia casuale: in questo modo si tiene conto della arbitrarietà nella definizione dell'istante $t = 0$. Le statistiche della v.a. \mathbf{r} non vengono per ora specificate, varieranno a seconda dell'applicazione (anche se \mathbf{r} fosse un numero, anzichè una v.a., $x(t)$ sarebbe comunque un p.a.).

Per calcolare il **valor medio** del p.a. si parte dalla definizione:

$$E\{x(t)\} = \int_{-\infty}^{\infty} x f(x, t) dx$$

A t fissato il p.a. è costituito dal prodotto di due v.a., \mathbf{r} e $\mathbf{z} = g(\varphi) = \cos(\omega t + \varphi)$:

$$E\{x(t)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r z f(r, z; t) dr dz = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r \cos(\omega t + \varphi) f(r, \varphi; t) dr d\varphi$$

Dalla indipendenza di \mathbf{r} e φ segue l'indipendenza di \mathbf{r} e \mathbf{z} (vedi sezione 6.10.1). Quindi vale che $f(r, z; t) = f(r)f(z; t)$ e $f(r, \varphi; t) = f(r)f(\varphi)$. Le densità $f(r)$ e $f(\varphi)$ non dipendono dal tempo, e neanche quella di \mathbf{z} (anche se il legame fra φ e \mathbf{z} contiene il tempo).

$$E\{x(t)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r \cos(\omega t + \varphi) f(r) f(\varphi) dr d\varphi = \int_{-\infty}^{\infty} r f(r) dr \int_{-\infty}^{\infty} \cos(\omega t + \varphi) f(\varphi) d\varphi$$

Il primo integrale è il valor medio di \mathbf{r} , il secondo può essere calcolato ricordando che φ è uniforme in $(-\pi, \pi)$:

$$E\{x(t)\} = \bar{r} \int_{-\pi}^{\pi} \cos(\omega t + \varphi) \frac{1}{2\pi} d\varphi = 0$$

L'**autocorrelazione** $R_x(t_1, t_2)$ del processo è (dalla definizione)

$$R_x(t_1, t_2) = E\{\mathbf{r} \cos(\omega t_1 + \varphi) \mathbf{r} \cos(\omega t_2 + \varphi)\}$$

usando la relazione trigonometrica

$$\cos(\alpha) \cos(\beta) = \frac{1}{2} [\cos(\alpha + \beta) + \cos(\alpha - \beta)]$$

si arriva a

$$R_x(t_1, t_2) = \frac{1}{2} E\{\mathbf{r}^2\} E\{\cos[\omega(t_1 - t_2)] + \cos(\omega t_1 + \omega t_2 + 2\varphi)\}$$

e poichè

$$E\{\cos(\omega t_1 + \omega t_2 + 2\varphi)\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(\omega t_1 + \omega t_2 + 2\varphi) d\varphi = 0$$

si ottiene che

$$R(t_1, t_2) = \frac{1}{2} E\{\mathbf{r}^2\} \cos[\omega(t_1 - t_2)] \quad (7.12)$$

Per ottenere questo risultato si sono usate normali formule trigonometriche, la linearità dell'operatore di media e l'ovvio fatto che la media di una funzione NON aleatoria è la funzione stessa, ovvero che $E\{\cos[\omega(t_1 - t_2)]\} = \cos[\omega(t_1 - t_2)]$.

La **potenza media** del processo sinusoidale si ottiene da

$$E\{x^2(t)\} = R(t, t) = \frac{1}{2} E\{\mathbf{r}^2\} \cos[\omega(t - t)] = \frac{1}{2} E\{\mathbf{r}^2\} \quad (7.13)$$

7.3.2 Processo di Poisson

In sezione 5.13 sono stati definiti i punti casuali di Poisson, aventi le seguenti proprietà:

- il numero $n(t_1, t_2)$ dei punti che cadono nell'intervallo (t_1, t_2) di lunghezza $t_a = t_2 - t_1$ è una v.a. di Poisson con parametro λt_a (vedi eq. 5.40):

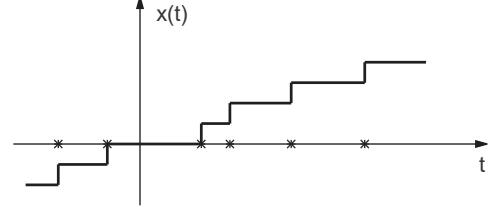
$$P\{n(t_1, t_2) = k\} = e^{-\lambda t_a} \frac{(\lambda t_a)^k}{k!} \quad (7.14)$$

- se gli intervalli (t_1, t_2) e (t_3, t_4) non sono sovrapposti allora le v.a. $n(t_1, t_2)$ e $n(t_3, t_4)$ sono indipendenti.

Usando i punti t_i formiamo il p.a.

$$x(t) = n(0, t)$$

in figura. Il p.a. $x(t)$ è un processo a stati discreti che consiste in una famiglia di funzioni crescenti "a scalini" discontinue nei punti t_i .



Fissato t , $x(t)$ è una v.a. di Poisson con parametro λt e media

$$E\{x(t)\} = \eta(t) = \lambda t \quad (7.15)$$

La sua autocorrelazione vale:

$$R(t_1, t_2) = \begin{cases} \lambda t_2 + \lambda^2 t_1 t_2 & \text{se } t_1 \geq t_2 \\ \lambda t_1 + \lambda^2 t_1 t_2 & \text{se } t_1 \leq t_2 \end{cases} \quad (7.16)$$

Questo risultato si ottiene come segue: se $t_1 = t_2$ allora (vedi eq. 6.31)

$$E\{x^2(t)\} = \lambda t + \lambda^2 t^2 \quad (7.17)$$

e la 7.16 è vera.

Poichè $R(t_1, t_2) = R(t_2, t_1)$ è sufficiente provare la 7.16 per $t_1 < t_2$. Le v.a. $x(t_1)$ e $[x(t_2) - x(t_1)]$ sono indipendenti dato che gli intervalli $(0, t_1)$ e (t_1, t_2) non sono sovrapposti. Inoltre esse sono v.a. di Poisson con parametro λt_1 e $\lambda(t_2 - t_1)$ rispettivamente. Quindi

$$E\{x(t_1)[x(t_2) - x(t_1)]\} = E\{x(t_1)\}E\{x(t_2) - x(t_1)\} = \lambda t_1 \lambda(t_2 - t_1)$$

Usando l'identità

$$x(t_1)x(t_2) = x(t_1)[x(t_1) + x(t_2) - x(t_1)]$$

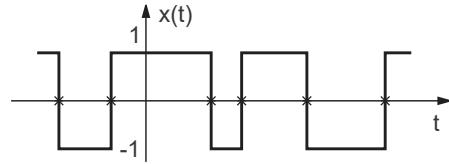
e la 7.17 concludiamo che

$$R(t_1, t_2) = E\{x(t_1)x(t_2)\} = \lambda t_1 + \lambda^2 t_1^2 + \lambda t_1 \lambda(t_2 - t_1)$$

da cui segue la 7.16.

7.3.3 Processo telegrafico casuale

Usando ancora i punti di Poisson t_i formiamo un processo $x(t)$ tale che $x(t) = 1$ se il numero di punti nell'intervallo $(0, t)$ è pari e $x(t) = -1$ se è dispari. Chiamando $P(k)$ la probabilità che il numero di punti nell'intervallo $(0, t)$ sia k concludiamo che



$$P\{x(t) = 1\} = P(0) + P(2) + \dots = e^{-\lambda t} \left[1 + \frac{(\lambda t)^2}{2!} + \dots \right] = e^{-\lambda t} \cosh(\lambda t)$$

e

$$P\{x(t) = -1\} = P(1) + P(3) + \dots = e^{-\lambda t} \left[\lambda t + \frac{(\lambda t)^3}{3!} + \dots \right] = e^{-\lambda t} \sinh(\lambda t)$$

Quindi

$$E\{x(t)\} = e^{-\lambda t} [\cosh(\lambda t) - \sinh(\lambda t)] = e^{-2\lambda t} \quad (7.18)$$

risultato che ha significato per $t > 0$. Inoltre

$$R(t_1, t_2) = e^{-2\lambda|t_1 - t_2|} \quad (7.19)$$

Il p.a. illustrato si chiama processo telegrafico semi-casuale perchè il suo valore in $t = 0$ non è casuale, ma vale $x(0) = 1$. Per togliere questa certezza formiamo il prodotto $y(t) = \mathbf{a}x(t)$ con \mathbf{a} v.a. che assume i valori ± 1 con uguale probabilità. La v.a. \mathbf{a} è inoltre indipendente da $x(t)$. Il p.a. $y(t)$ è detto processo telegrafico casuale.

Poichè $E\{\mathbf{a}\} = 0$ e $E\{\mathbf{a}^2\} = 1$ si ha che la media del p.a. $y(t)$ vale

$$E\{y(t)\} = E\{\mathbf{a}\}E\{x(t)\} = 0 \quad (7.20)$$

e la sua autocorrelazione è

$$E\{y(t_1)y(t_2)\} = E\{\mathbf{a}^2\}E\{x(t_1)x(t_2)\} = e^{-2\lambda|t_1 - t_2|} \quad (7.21)$$

Da notare che, al crescere di $t \rightarrow \infty$, i p.a. $x(t)$ e $y(t)$ hanno statistiche asintoticamente uguali.

Supponiamo ora di voler trovare valor medio ed autocorrelazione di un processo telegrafico casuale definito fra 0 ed A anzichè fra -1 e 1 . Tale p.a. si può ovviamente definire, a partire dal p.a. $y(t)$, come

$$z(t) = \frac{A}{2} [y(t) + 1] = \frac{A}{2} y(t) + \frac{A}{2}$$

Essendo A una costante si ottiene immediatamente che

$$E\{z(t)\} = E\{\frac{A}{2}y(t)\} + E\{\frac{A}{2}\} = 0 + \frac{A}{2} = \frac{A}{2}$$

L'autocorrelazione è

$$\begin{aligned} E\{z(t_1)z(t_2)\} &= E\left\{\left(\frac{A}{2}y(t_1) + \frac{A}{2}\right)\left(\frac{A}{2}y(t_2) + \frac{A}{2}\right)\right\} = \\ &= E\left\{\frac{A^2}{4}y(t_1)y(t_2) + \frac{A^2}{4}y(t_1) + \frac{A^2}{4}y(t_2) + \frac{A^2}{4}\right\} = \\ &= \frac{A^2}{4}E\{y(t_1)y(t_2)\} + \frac{A^2}{4} = \frac{A^2}{4} \left(e^{-2\lambda|\tau|} + 1\right) = R_z(\tau) \end{aligned}$$

avendo posto $\tau = t_2 - t_1$. Per τ che tende all'infinito $R_z(\tau)$ vale $A^2/4$, che è il valor medio al quadrato. La potenza media vale infine

$$E\{z^2(t)\} = R_z(0) = \frac{A^2}{2}$$

7.3.4 Rumore bianco

Il rumore bianco è un p.a. $v(t)$ tale che i suoi valori (v.a.) $v(t_1)$ e $v(t_2)$ sono scorrelati per ogni t_1 e t_2 ($t_1 \neq t_2$), ovvero tale che

$$C(t_i, t_j) = 0 \quad \forall t_i \neq t_j \quad (7.22)$$

Si può dimostrare che l'autocovarianza di un p.a. rumore bianco deve essere nella forma

$$C(t_1, t_2) = q(t_1) \delta(t_1 - t_2) \quad q(t) \geq 0 \quad (7.23)$$

Se le v.a. $v(t_1)$ e $v(t_2)$ sono anche indipendenti (oltre che scorrelate) il p.a. $v(t)$ si dice rumore strettamente bianco.

Il rumore bianco modella gran parte dei disturbi presenti nei canali di trasmissione: riprenderemo l'argomento più avanti.

7.3.5 Processo gaussiano

Un p.a. $x(t)$ è detto gaussiano se le v.a. $x(t_1), \dots, x(t_n)$ sono congiuntamente gaussiane per ogni n e t_1, \dots, t_n . Le statistiche di un p.a. gaussiano sono completamente determinate conoscendo la sua media $\eta(t)$ e la sua autocovarianza $C(t_1, t_2)$. Infatti, siccome

$$E\{x(t)\} = \eta(t) \quad \text{e} \quad \sigma_x^2(t) = C(t, t)$$

allora la densità del primo ordine $f(x, t)$ di $x(t)$ è la densità gaussiana $N[\eta(t); \sqrt{C(t, t)}]$.

7.4 Due processi aleatori

Le statistiche congiunte di due p.a. reali sono determinate in termini delle distribuzioni congiunte delle v.a.

$$x(t_1), \dots, x(t_n), y(t'_1), \dots, y(t'_n)$$

Il p.a. complesso $z(t) = x(t) + jy(t)$ è specificato in termini delle statistiche congiunte dei p.a. reali $x(t)$ e $y(t)$.

Per un p.a. complesso le definizioni vanno lievemente modificate, inserendo in alcuni punti il complesso coniugato. Le definizioni che seguono tengono conto di ciò e ovviamente sono valide anche per p.a. reali (in tal caso il complesso coniugato non fa nulla).

L'autocorrelazione di un p.a. $x(t)$, reale o complesso, è, per definizione,

$$R(t_1, t_2) = R_x(t_1, t_2) = R_{xx}(t_1, t_2) = E\{x(t_1)x^*(t_2)\} \quad (7.24)$$

Da cui segue che

$$R(t_2, t_1) = E\{x(t_2)x^*(t_1)\} = R^*(t_1, t_2) \quad (7.25)$$

Ovviamente

$$R(t, t) = E\{|x(t)|^2\} \geq 0 \quad (7.26)$$

è sempre reale.

L'autocovarianza $C(t_1, t_2)$ di un p.a. $x(t)$ è la covarianza delle v.a. $x(t_1)$ e $x(t_2)$:

$$C(t_1, t_2) = R(t_1, t_2) - \eta(t_1)\eta^*(t_2) \quad (7.27)$$

essendo $\eta(t) = E\{x(t)\}$.

La correlazione incrociata di due p.a. $x(t)$ e $y(t)$ è la funzione

$$R_{xy}(t_1, t_2) = E\{x(t_1)y^*(t_2)\} = R_{yx}^*(t_2, t_1) \quad (7.28)$$

Similmente, la covarianza incrociata è

$$C_{xy}(t_1, t_2) = R_{xy}(t_1, t_2) - \eta_x(t_1)\eta_y^*(t_2) \quad (7.29)$$

Due p.a. $x(t)$ e $y(t)$ sono detti (mutuamente) ortogonal se

$$R_{xy}(t_1, t_2) = 0 \quad \forall t_1, t_2 \quad (7.30)$$

7.5 Processi stazionari

Una importante classe di p.a. consiste di quelli con statistiche invarianti rispetto alla traslazione nel tempo.

Un p.a. $x(t)$ è detto **stazionario in senso stretto** (o **SSS**) se le sue proprietà statistiche sono invarianti rispetto ad una qualsiasi traslazione dell'origine. Questo significa che i processi $x(t)$ e $x(t+c)$ hanno le stesse statistiche per ogni c .

Due processi $x(t)$ ed $y(t)$ sono detti *congiuntamente stazionari* se le statistiche congiunte di $x(t)$ e $y(t)$ sono le stesse delle statistiche congiunte di $x(t+c)$ e $y(t+c)$ per ogni c .

Un processo complesso $z(t) = x(t) + jy(t)$ è stazionario se i processi $x(t)$ e $y(t)$ sono congiuntamente stazionari.

Dalla definizione segue che la densità di ordine n di un p.a. SSS deve essere tale che

$$f(x_1, \dots, x_n; t_1, \dots, t_n) = f(x_1, \dots, x_n; t_1 + c, \dots, t_n + c) \quad (7.31)$$

per ogni c . Quindi, in particolare, deve valere che $f(x; t) = f(x; t+c)$ per ogni c . Allora la densità del primo ordine di $x(t)$ deve essere indipendente dal tempo t :

$$f(x; t) = f(x) \quad (7.32)$$

Ne segue che il valor medio di un p.a. SSS non dipende dal tempo, ovvero è un numero. Infatti

$$E\{x(t)\} = \int_{-\infty}^{\infty} x f_x(x, t) dx = \int_{-\infty}^{\infty} x f_x(x) dx = \eta_x \quad (7.33)$$

Analogamente $f(x_1, x_2; t_1, t_2) = f(x_1, x_2; t_1 + c, t_2 + c)$ deve essere indipendente da c e quindi

$$f(x_1, x_2; t_1, t_2) = f(x_1, x_2; \tau) \quad \tau = t_1 - t_2 \quad (7.34)$$

ovvero le statistiche del secondo ordine dipendono solo dalla differenza fra t_1 e t_2 .

Da ciò segue che la funzione di autocorrelazione dipende solo dalla differenza degli istanti temporali:

$$R(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; \tau) dx_1 dx_2 = R(\tau) \quad \tau = t_1 - t_2 \quad (7.35)$$

Un p.a. è detto **stazionario in senso lato** (o **SSL**) se la sua media è costante (non dipende dal tempo)

$$E\{x(t)\} = \eta_x \quad (7.36)$$

e la sua autocorrelazione dipende solo da $\tau = t_1 - t_2$:

$$E\{x(t+\tau)x^*(t)\} = R_x(\tau) \quad (7.37)$$

Poichè τ è la distanza da t a $t+\tau$ la funzione $R_x(\tau)$ può essere scritta anche nella forma simmetrica

$$R_x(\tau) = E\{x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})\} \quad (7.38)$$

Da notare che la **potenza media** di un p.a. stazionario (SSS e SSL) è indipendente dal tempo t e vale $R_x(0)$:

$$R_x(0) = E\{x(t)x^*(t)\} = E\{|x(t)|^2\} \quad (7.39)$$

Dalla 7.37 segue che l'autocovarianza di un p.a. $x(t)$ SSL dipende solo da $\tau = t_1 - t_2$:

$$C_x(\tau) = R_x(\tau) - |\eta_x|^2 \quad (7.40)$$

ed il coefficiente di correlazione vale

$$r_x(\tau) = \frac{C_x(\tau)}{C_x(0)} \quad (7.41)$$

Quindi $C_x(\tau)$ è la covarianza delle v.a. $x(t+\tau)$ e $x(t)$ e $r_x(\tau)$ il loro coefficiente di correlazione.

Da notare che, ovviamente, se un p.a. è SSS è anche SSL. L'inverso non è, in generale, vero. Se tuttavia un p.a. SSL è gaussiano allora è anche SSS.

7.5.1 Processi ergodici

Una importante sottoclasse dei processi stazionari è quella dei processi **ergodici**. Questi ultimi sono caratterizzati dal fatto che le statistiche temporali calcolate su di un qualsiasi campione coincidono con le medie di insieme dell'intero p.a.

Definizione: se un p.a. è ergodico allora la media temporale di ciascuna delle funzioni campione $x_k(t)$ coincide con il valore atteso del processo. Ovvero, per ogni campione $x_k(t)$ vale che

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_k(t) dt \equiv E\{x(t)\} \equiv \int_{-\infty}^{\infty} x f_x(x) dx \quad (7.42)$$

Esempio 7.1 Consideriamo il p.a. sinusoidale

$$x(t) = A \cos(\omega_0 t + \theta)$$

dove A e ω_0 sono costanti e θ è una v.a. uniformemente distribuita in $(0, 2\pi)$. Valutiamo media e potenza del p.a. $x(t)$:

$$E\{x(t)\} = \int_{-\infty}^{\infty} x(t, \theta) f_{\theta}(\theta) d\theta = \int_0^{2\pi} A \cos(\omega_0 t + \theta) \frac{1}{2\pi} d\theta = 0$$

$$E\{x^2(t)\} = \int_0^{2\pi} [A \cos(\omega_0 t + \theta)]^2 \frac{1}{2\pi} d\theta = \int_0^{2\pi} \frac{A^2}{2} [1 + \cos(2\omega_0 t + 2\theta)] \frac{1}{2\pi} d\theta = \frac{A^2}{2} \quad (7.43)$$

Entrambi i valori non dipendono da t , essendo il processo stazionario (affermazione non dimostrata).

Calcoliamo ora le medie temporali: si ipotizza di utilizzare un campione del processo, quindi nelle formule che seguono il valore di θ è fissato (è un numero, non una v.a.).

$$\begin{aligned} & \frac{1}{T_0} \int_0^{T_0} A \cos(\omega_0 t + \theta) dt = 0 \\ \text{e} \quad & \frac{1}{T_0} \int_0^{T_0} [A \cos(\omega_0 t + \theta)]^2 dt = \frac{A^2}{2} \end{aligned} \quad (7.44)$$

dove $T_0 = 1/f_0$ e $\omega_0 = 2\pi f_0$ e si è utilizzato l'operatore di media temporale per funzioni periodiche.

Il risultato ottenuto è un indicatore che il processo potrebbe essere ergodico. In effetti lo è, ma dimostrarlo non è semplice (bisogna verificare *tutte* le statistiche temporali e di insieme). In generale è piuttosto difficile dimostrare l'ergodicità di un processo: in casi reali spesso si ipotizza che un p.a. sia ergodico quando alcune delle sue medie temporali coincidono con quelle d'insieme.

Esempio 7.2 Consideriamo ancora il p.a.

$$x(t) = A \cos(\omega_0 t + \theta)$$

dove ω_0 è costante, θ è una v.a. uniformemente distribuita in $(0, 2\pi)$ ma A non è più una costante: ad esempio A potrebbe essere una v.a. uniformemente distribuita in $(-1, 1)$.

Il nuovo p.a. non è ergodico (provate ad immaginare cosa succede se si analizza il campione di $x(t)$ che si ottiene quando $A = 0$).

7.5.2 Rumore bianco stazionario

Un p.a. $n(t)$ rumore bianco stazionario almeno SSL avrà autocovarianza (vedi equazione 7.23)

$$C(\tau) = q \delta(\tau) \quad q \geq 0 \quad (7.45)$$

Detto η il valor medio di $n(t)$ (essendo SSL η non dipende dal tempo) l'autocorrelazione è

$$R(\tau) = C(\tau) + \eta^2 = q \delta(\tau) + \eta^2 \quad q \geq 0 \quad (7.46)$$

La potenza media del p.a. non è esprimibile, in quanto il valore $R(0)$ non esiste (è il valore di $\delta(t)$ in $t = 0$). La potenza media del p.a. rumore bianco è infatti infinita.

7.6 Sistemi con ingresso aleatorio

Dato un p.a. $x(t)$ assegniamo, in accordo con una qualche regola, a ciascuno dei suoi campioni $x(t, \xi_i)$ una funzione $y(t, \xi_i)$. Abbiamo quindi creato un nuovo p.a.

$$y(t) = T[x(t)]$$

i cui campioni sono le funzioni $y(t, \xi_i)$. Il nuovo p.a. $y(t)$ può essere considerato come l'uscita di un sistema che ha per ingresso il p.a. $x(t)$. Il sistema è completamente specificato in termini dell'operatore $T[\cdot]$.

Il sistema è deterministico se opera sulla sola variabile t considerando ξ_i un parametro. Ovvero, se due campioni $x(t, \xi_1)$ e $x(t, \xi_2)$ sono identici in t lo sono anche le rispettive uscite $y(t, \xi_1)$ e $y(t, \xi_2)$. Se questo non accade il sistema è detto aleatorio. Noi considereremo solo sistemi di tipo deterministico. L'uscita di un generico sistema può essere espressa in termini di $T[\cdot]$ e delle statistiche dell'ingresso. Non è tuttavia semplice farlo. Vedremo due importanti casi speciali: i sistemi senza memoria ed i sistemi lineari.

7.6.1 Sistemi senza memoria

Un sistema è detto senza memoria se la sua uscita è

$$y(t) = g[x(t)]$$

dove $g(x)$ è una funzione di x . Ad uno specificato istante t_1 l'uscita $y(t_1)$ dipende quindi solo da $x(t_1)$ e non dai valori passati e futuri di $x(t)$. Di conseguenza la densità del primo ordine $f_y(y; t)$ di $y(t)$ può essere espressa in termini della corrispondente densità $f_x(x; t)$ di $x(t)$, in analogia a quanto visto per le v.a. in sezione 6.9 (a t fissato un p.a. è una v.a.). Il valor medio dell'uscita si può esprimere come

$$E\{y(t)\} = \int_{-\infty}^{\infty} g(x) f_x(x; t) dx \quad (7.47)$$

Poichè $y(t_1) = g[x(t_1)]$ e $y(t_2) = g[x(t_2)]$, la densità del secondo ordine $f_y(y_1, y_2; t_1, t_2)$ di $y(t)$ è determinata in termini della corrispondente densità $f_x(x_1, x_2; t_1, t_2)$ di $x(t)$ (questo vale per le densità di ordine n qualsiasi). Inoltre

$$E\{y(t_1)y(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1) g(x_2) f_x(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (7.48)$$

Stazionarietà: se l'ingresso ad un sistema senza memoria è SSS anche la sua uscita lo è. Se l'ingresso è SSL l'uscita può non essere stazionaria in nessun senso.

Esempio 7.3 sia dato il sistema senza memoria

$$y(t) = x^2(t)$$

Troviamo la densità del primo ordine dell'uscita in funzione di quella dell'ingresso utilizzando quanto visto in sezione 6.9.2.

Se $y > 0$ l'equazione $y = x^2$ ha le due soluzioni $\pm x = \pm\sqrt{y}$ e $y'(x) = \pm 2\sqrt{y}$. Quindi

$$f_y(y; t) = \frac{1}{2\sqrt{y}} [f_x(\sqrt{y}; t) + f_x(-\sqrt{y}; t)] \quad y > 0$$

Esempio 7.4 Hard Limiter.

Consideriamo un sistema senza memoria tale che

$$g(x) = \begin{cases} 1 & \text{se } x > 0 \\ -1 & \text{se } x < 0 \end{cases}$$

Assumiamo che il p.a. in ingresso $x(t)$ sia stazionario (almeno SSL): l'uscita $y(t)$ assume i valori ± 1 con probabilità

$$\begin{aligned} P\{y(t) = 1\} &= P\{x(t) > 0\} = 1 - F_x(0; t) = 1 - F_x(0) \\ P\{y(t) = -1\} &= P\{x(t) < 0\} = F_x(0; t) = F_x(0) \end{aligned}$$

il valor medio di $y(t)$ vale quindi

$$E\{y(t)\} = 1 \cdot P\{y(t) = 1\} - 1 \cdot P\{y(t) = -1\} = 1 - 2F_x(0; t) = 1 - 2F_x(0)$$

Calcoliamo l'autocorrelazione di $y(t)$: il prodotto $y(t + \tau)y(t)$ vale 1 se $x(t + \tau)x(t) > 0$ e vale -1 se $x(t + \tau)x(t) < 0$. Quindi

$$R_y(\tau) = P\{x(t + \tau)x(t) > 0\} - P\{x(t + \tau)x(t) < 0\} \quad (7.49)$$

In questo caso è quindi evidente che se $x(t)$ è SSL anche $y(t)$ è SSL.

NOTA: lo studente è invitato a riflettere sul fatto che il sistema non ha uscita definita se l'ingresso è nullo. Che proprietà deve avere $x(t)$ affinché i risultati visti sopra siano corretti?

7.6.2 Sistemi lineari (LTI)

La notazione $y(t) = L[x(t)]$ indica che $y(t)$ è l'uscita di un sistema lineare con ingresso $x(t)$. Questo significa che

$$L[a_1x_1(t) + a_2x_2(t)] = a_1L[x_1(t)] + a_2L[x_2(t)] \quad (7.50)$$

per ogni $a_1, a_2, x_1(t)$ e $x_2(t)$. La 7.50 è vera anche se i coefficienti a_1 e a_2 sono v.a., essendo il sistema deterministico. L'eq. 7.50 è nota come *principio di sovrapposizione degli effetti*.

D'ora in avanti assumeremo che il sistema sia anche tempo invariante, cioè che

$$y(t + c) = L[x(t + c)] \quad \forall c$$

L'uscita si può quindi esprimere mediante convoluzione dell'ingresso con la risposta all'impulso $h(t) = L[\delta(t)]$ del sistema

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(t - \alpha) h(\alpha) d\alpha \quad (7.51)$$

Dalle definizioni seguono alcune importanti conseguenze (valide per sistemi LTI):

1. se $x(t)$ è un p.a. gaussiano anche $y(t)$ è un p.a. gaussiano;
2. se $x(t)$ è SSS anche $y(t)$ è SSS;
3. se $x(t)$ è SSL allora i p.a. $x(t)$ e $y(t)$ sono congiuntamente SSL.

Media: per ogni sistema lineare (non solo LTI) vale che

$$\eta_y(t) = E\{y(t)\} = E\{L[x(t)]\} = L[E\{x(t)\}] \quad (7.52)$$

ovvero la media dell'uscita $\eta_y(t)$ è la risposta alla media $\eta_x(t)$ dell'ingresso

$$\eta_y(t) = L[\eta_x(t)] \quad (7.53)$$

Se il sistema è LTI

$$E\{y(t)\} = \eta_y(t) = \int_{-\infty}^{\infty} E\{x(t - \alpha)\} h(\alpha) d\alpha = \eta_x(t) * h(t) \quad (7.54)$$

Questo risultato è valido se $E\{x(t)\} < \infty \forall t$ e se il sistema LTI è stabile, cioè se

$$\int_{-\infty}^{\infty} |h(\tau)| d\tau < \infty$$

Se $x(t)$ è almeno SSL il suo valor medio non dipende dal tempo e

$$\eta_y = \eta_x \int_{-\infty}^{\infty} h(\alpha) d\alpha = \eta_x H(0) \quad (7.55)$$

Autocorrelazione dell'uscita di un sistema LTI:

$$\begin{aligned} R_{yy}(t_1, t_2) &= E\{y(t_1)y(t_2)\} = \\ &= E \left\{ \int_{-\infty}^{\infty} x(t_1 - \alpha) h(\alpha) d\alpha \int_{-\infty}^{\infty} x(t_2 - \beta) h(\beta) d\beta \right\} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha) h(\beta) E\{x(t_1 - \alpha)x(t_2 - \beta)\} d\alpha d\beta = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{xx}(t_1 - \alpha, t_2 - \beta) h(\alpha) h(\beta) d\alpha d\beta \end{aligned} \quad (7.56)$$

Se il p.a. $x(t)$ è stazionario (almeno SSL) allora

$$R_{yy}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{xx}(\tau - \alpha + \beta) h(\alpha) h(\beta) d\alpha d\beta = R_{xx}(\tau) * h(\tau) * h^*(-\tau) \quad (7.57)$$

Questo risultato è di notevole importanza, in quanto permette di calcolare l'autocorrelazione dell'uscita conoscendo l'autocorrelazione dell'ingresso e la risposta all'impulso. La potenza media di $y(t)$ si ottiene da $R_{yy}(0)$, ovvero da

$$E\{y^2(t)\} = R_{yy}(0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{xx}(-\alpha + \beta) h(\alpha) h(\beta) d\alpha d\beta \quad (7.58)$$

Come evidente non è sufficiente conoscere la potenza dell'ingresso $E\{x^2(t)\} = R_{xx}(0)$ per poter calcolare la potenza dell'uscita.

Esempio 7.5 Rumore bianco stazionario.

Se l'ingresso ad un sistema LTI è un rumore bianco SSL a media nulla con autocorrelazione

$$R_{xx}(\tau) = q \delta(\tau)$$

allora l'energia dell'uscita vale

$$E\{y^2(t)\} = q \int_{-\infty}^{\infty} |h(t)|^2 dt \quad (7.59)$$

dove l'integrale è l'energia della risposta all'impulso, detta anche energia del filtro.

7.7 Densità spettrale di potenza

La **densità spettrale di potenza** di un p.a. SSL (o SSS) $x(t)$, reale o complesso, è, per definizione, la trasformata di Fourier della sua autocorrelazione $R_x(\tau) = E\{x(t + \tau)x(t)\}$:

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau \quad (7.60)$$

Poichè $R(-\tau) = R^*(\tau)$ ne segue che S_x è una funzione reale di ω . S_x è anche detta spettro di potenza ed è positiva o nulla $\forall \omega$ (vedi sezione 7.7.3). Dalla 7.60 segue che

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau} d\omega \quad (7.61)$$

Se $x(t)$ è un p.a. reale allora $R_{xx}(\tau)$ è reale e pari, per cui $S_x(\omega)$ è pari e reale e le eq. 7.60 e 7.61 si possono esprimere usando solo $\cos(\omega\tau)$.

Lo spettro di potenza di p.a. non stazionari non ha (quasi) utilità pratica e non viene qui trattato.

Esempio 7.6 il p.a. telegrafico casuale di sezione 7.3.3 ha autocorrelazione $R(\tau) = e^{-2\lambda|\tau|}$ ed il suo spettro di potenza vale

$$S(\omega) = \frac{4\lambda}{4\lambda^2 + \omega^2} \quad (7.62)$$

È possibile esprimere lo spettro di potenza dell'uscita di un sistema LTI in funzione dello spettro dell'ingresso e della risposta all'impulso. Infatti, dall'eq. 7.57 si ricava che

$$S_y(\omega) = S_x(\omega)H(\omega)H^*(\omega) = S_x(\omega)|H(\omega)|^2 \quad (7.63)$$

Vediamo di capire perchè la densità spettrale di potenza (o spettro di potenza) ha questo nome: la potenza media di ogni p.a. SSL $x(t)$ vale

$$E\{|x(t)|^2\} = R_{xx}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) d\omega \quad (7.64)$$

Quindi la potenza media di un p.a. si può ottenere integrando su tutto l'asse frequenziale $S_x(\omega)$, che è quindi la densità (di potenza).

Se l'integrale copre solo una certa banda di frequenze (doppia e simmetrica rispetto all'origine) si ottiene la potenza media del p.a. in quella banda. Ciò si vede bene considerando un filtro LTI passa basso o passa banda ideale (modulo unitario e fase nulla) ed usando l'eq. 7.63 per trovare lo spettro di potenza dell'uscita, come verrà mostrato in sezione 7.7.3.

Si definisce **densità spettrale incrociata** di due p.a. $x(t)$ ed $y(t)$ congiuntamente stazionari la trasformata di Fourier della loro correlazione incrociata:

$$S_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j\omega\tau} d\tau \quad (7.65)$$

Poichè $R_{xy}(\tau) = R_{yx}(-\tau)$, vale che

$$S_{xy}(\omega) = S_{yx}^*(\omega) \quad (7.66)$$

Da notare che $S_{xy}(\omega)$ non è necessariamente pari, reale o positiva.

Per un sistema LTI valgono le seguenti uguaglianze:

$$\begin{aligned} R_{xy}(\tau) &= R_{xx}(\tau) * h^*(-\tau) & R_{yy}(\tau) &= R_{xy}(\tau) * h(\tau) \\ S_{xy}(\omega) &= S_{xx}(\omega) H^*(\omega) & S_{yy}(\omega) &= S_{xy}(\omega) H(\omega) \end{aligned} \quad (7.67)$$

Dato un sistema LTI avente in ingresso un qualsiasi p.a. $x(t)$ almeno SSL e ricordando l'eq. 7.63 segue che

$$E\{|y(t)|^2\} = R_{yy}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) |H(\omega)|^2 d\omega \geq 0 \quad (7.68)$$

7.7.1 Rumore bianco

Se $n(t)$ è un p.a. rumore bianco SSL a valor medio nullo, allora

$$\begin{aligned} R_{xx}(\tau) &= q \delta(\tau) & S_{xx}(\omega) &= q \\ R_{yy}(\tau) &= q [h(\tau) * h^*(-\tau)] & S_{yy}(\omega) &= q |H(\omega)|^2 \end{aligned} \quad (7.69)$$

Da notare, in particolare, che lo spettro di densità di potenza è pari ad una costante e quindi che la potenza media del p.a. rumore bianco è *infinita* (eq. 7.64).

Il fatto che lo spettro di densità di potenza sia piatto giustifica il fatto che tale processo sia chiamato *bianco*: infatti la luce bianca è quella che contiene in uguale proporzione tutti i colori (anche lo spettro della luce bianca è piatto). Seguendo tale analogia, quando un processo bianco passa attraverso un filtro (che varia il suo spettro), si chiama *colorato*.

7.7.2 Trasformata di Hilbert

Un sistema LTI con risposta in frequenza

$$H(\omega) = -j \operatorname{sgn}(\omega) = \begin{cases} -j & \omega > 0 \\ j & \omega < 0 \\ 0 & \omega = 0 \end{cases} \quad (7.70)$$

è detto filtro in quadratura. La corrispondente risposta all'impulso è $h(t) = \frac{1}{\pi t}$. Evidentemente $H(\omega)$ in eq. 7.70 è un filtro passa-tutto che sfasa di -90° : ad esempio se l'ingresso è $\cos(\omega t)$ l'uscita vale $\cos(\omega t - 90^\circ) = \sin(\omega t)$ e se l'ingresso è $\sin(\omega t)$ l'uscita vale $\sin(\omega t - 90^\circ) = -\cos(\omega t)$.

La risposta di questo filtro ad un p.a. reale $x(t)$ [o a un segnale reale $x(t)$] è detta la trasformata di Hilbert di $x(t)$ e viene indicata con $\check{x}(t)$.

$$\check{x}(t) = x(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\alpha)}{t - \alpha} d\alpha \quad (7.71)$$

Dalle eq. 7.66 e 7.67 segue che

$$\begin{aligned} S_{x\ddot{x}}(\omega) &= j S_{xx}(\omega) \operatorname{sgn}(\omega) = -S_{\dot{x}x}(\omega) \\ S_{\dot{x}\ddot{x}}(\omega) &= S_{xx}(\omega) \end{aligned} \quad (7.72)$$

La trasformata di Hilbert di $\dot{x}(t)$ produce $-x(t)$ togliendo la componente continua:

$$\begin{aligned} \check{X}(\omega) &= -j \operatorname{sgn}(\omega) X(\omega) \\ \check{\check{X}}(\omega) &= (-j)^2 [\operatorname{sgn}(\omega)]^2 X(\omega) = \begin{cases} -X(\omega) & \omega \neq 0 \\ 0 & \omega = 0 \end{cases} \end{aligned} \quad (7.73)$$

Il p.a. complesso $z(t) = x(t) + j\check{x}(t)$ è detto il segnale analitico associato ad $x(t)$. Chiaramente $z(t)$ è la risposta del sistema avente risposta in frequenza

$$H(\omega) = 1 + j[-j \operatorname{sgn}(\omega)] = 2 U(\omega)$$

all'ingresso $x(t)$ e (vedi eq. 7.63)

$$\begin{aligned} S_{zz}(\omega) &= 4 S_{xx}(\omega) U(\omega) = 2 S_{xx}(\omega) + 2j S_{\dot{x}x}(\omega) \\ R_{zz}(\tau) &= 2 R_{xx}(\tau) + 2j R_{\dot{x}x}(\tau) \end{aligned} \quad (7.74)$$

7.7.3 Teorema di Wiener - Khinchin

La densità spettrale di potenza è stata definita in 7.60 come la trasformata di Fourier della funzione di autocorrelazione. Tuttavia non è ancora chiaro il legame con lo spettro di densità di potenza per segnali deterministici visto nella prima parte del corso.

Sia $x(t) = x(t, \xi)$ un p.a. ed indichiamo con $x(t, \xi_i)$ una sua realizzazione. Sia poi $x_T(t, \xi_i)$ la versione troncata di tale realizzazione:

$$x_T(t, \xi_i) = \begin{cases} x(t, \xi_i) & \text{se } |t| < \frac{1}{2}T \\ 0 & \text{altrimenti} \end{cases} \quad (7.75)$$

La corrispondente trasformata di Fourier vale

$$X_T(f, \xi_i) = \int_{-\infty}^{\infty} x_T(t, \xi_i) e^{-j2\pi f t} dt = \int_{-T/2}^{T/2} x(t, \xi_i) e^{-j2\pi f t} dt \quad (7.76)$$

A ξ_i fissato $x_T(t, \xi_i)$ è un segnale deterministico che dipende solo da t , e $X_T(f, \xi_i)$ la sua trasformata di Fourier. L'energia di $x_T(t, \xi_i)$ si può quindi calcolare, utilizzando il teorema di Parseval, come

$$E_T(\xi_i) = \int_{-\infty}^{\infty} x_T^2(t, \xi_i) dt = \int_{-\infty}^{\infty} |X_T(f, \xi_i)|^2 df \quad (7.77)$$

Se non si fissa ξ le formule precedenti restano valide: ovviamente $x_T(t)$ è un p.a. (pari a $x(t)$ troncato), $X_T(f)$ è anche lui un p.a. (complesso) nella variabile f e E_T è una v.a.

L'energia media del p.a. troncato si ottiene allora calcolando la media di E_T :

$$\overline{E_T} = \int_{-T/2}^{T/2} \overline{x^2(t)} dt = \int_{-\infty}^{\infty} \overline{x_T^2(t)} dt = \int_{-\infty}^{\infty} \overline{|X_T(f)|^2} df \quad (7.78)$$

dove l'operatore di media è il solito: ad esempio

$$\int_{-T/2}^{T/2} \overline{x^2(t)} dt = \int_{-T/2}^{T/2} \left[\int_{-\infty}^{\infty} x^2 f_x(x, t) dx \right] dt \quad (7.79)$$

La potenza media è l'energia media per unità di tempo, ovvero

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \overline{x^2(t)} dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} \overline{x_T^2(t)} dt \quad (7.80)$$

ciò conferma l'intuizione: la potenza media (temporale) di un p.a. si ottiene mediando (in senso statistico) le potenze medie (temporali) delle singole realizzazioni del p.a.

La stessa relazione si può esprimere in frequenza utilizzando la 7.78:

$$P = \int_{-\infty}^{\infty} \left[\lim_{T \rightarrow \infty} \frac{1}{T} \overline{|X_T(f)|^2} \right] df \quad (7.81)$$

dove la funzione entro le parentesi quadre è ovviamente la densità spettrale di potenza del p.a.

$$S_x(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \overline{|X_T(f)|^2} \quad (7.82)$$

essendo, come visto,

$$X_T(f) = \int_{-T/2}^{T/2} x(t) e^{-j2\pi ft} dt \quad (7.83)$$

(la 7.81 ha senso purchè integrale e limite esistano).

La procedura vista è anche, per inciso, il metodo comunemente utilizzato per stimare la densità spettrale di potenza di un p.a. ignoto disponendo di alcune realizzazioni in un intervallo finito.

Resta da stabilire il legame fra lo spettro di densità di potenza appena visto e quello definito in precedenza. Di questo si occupa il

Teorema di Wiener-Khinchin *La densità spettrale di potenza di un p.a. stazionario (almeno in senso lato) è uguale alla trasformata di Fourier della funzione di autocorrelazione del p.a.*

$$S_x(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \overline{|X_T(f)|^2} = \mathcal{F}[R_x(\tau)] = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau \quad (7.84)$$

purchè $R_x(\tau)$ sia sufficientemente piccola al crescere di τ , in modo che

$$\int_{-\infty}^{\infty} |\tau R_x(\tau)| d\tau < \infty \quad (7.85)$$

Questa condizione è molto stretta: ad esempio esclude i p.a. a valor medio non nullo e il p.a. sinusoidale. Di fatto il teorema di Wiener-Khinchin è valido (almeno per le comunicazioni elettriche) purchè esista la trasformata di Fourier dell'autocorrelazione. Di solito si accetta l'esistenza nella versione "estesa", ovvero ammettendo le trasformate delle funzioni generalizzate (con la definizione normale le trasformate, ad esempio, della costante, della sinc e delle sinusoidi non esistono).

La dimostrazione del teorema di Wiener-Khinchin non viene qui riportata (si trova, ad esempio, in [Cou02] pag. 400 e seguenti).

Dalla definizione di densità spettrale di potenza e dal teorema di Wiener-Khinchin seguono le seguenti **proprietà**:

1. $S_x(f)$ è reale;
2. $S_x(f) \geq 0 \forall f$;
3. se $x(t)$ è reale $S_x(f)$ è pari, ovvero $S_x(-f) = S_x(f)$;
4. la potenza del processo vale $P = \int_{-\infty}^{\infty} S_x(f) df$
se $x(t)$ è stazionario, almeno in senso lato, $P = \int_{-\infty}^{\infty} S_x(f) df = \overline{x^2(t)} = R_x(0)$
5. $S_x(0) = \int_{-\infty}^{\infty} R_x(\tau) d\tau$

La proprietà 2. merita un approfondimento. Come visto, dato un p.a. $x(t)$,

$$E\{x^2(t)\} = R_x(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) d\omega \geq 0$$

che dimostra come l'area totale dello spettro di potenza sia positiva o nulla. Si può dimostrare una condizione più forte su S_x , ovvero che:

$$S_x(\omega) \geq 0 \quad \forall \omega \quad (7.86)$$

Dimostrazione: formiamo un sistema ideale passa banda con funzione di trasferimento

$$H(\omega) = \begin{cases} 1 & \omega_1 < \omega < \omega_2 \\ 0 & \text{altrove} \end{cases}$$

Sia $y(t)$ l'uscita di tale sistema quando l'ingresso vale $x(t)$: come visto (eq. 7.63):

$$S_{yy}(\omega) = \begin{cases} S_x(\omega) & \omega_1 < \omega < \omega_2 \\ 0 & \text{altrove} \end{cases}$$

ma

$$0 \leq E\{y^2(t)\} = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} S_{yy}(\omega) d\omega = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} S_x(\omega) d\omega$$

Quindi l'area di $S_x(\omega)$ è positiva in qualsiasi intervallo. Ciò è possibile solo se $S_x(\omega)$ è una funzione positiva. CVD

Inoltre, se $\omega_1 = \omega_0 - \epsilon$ e $\omega_2 = \omega_0 + \epsilon$ con ϵ abbastanza piccolo, vale che

$$E\{y^2(t)\} \approx \frac{\epsilon}{\pi} S_x(\omega_0) \quad (7.87)$$

7.7.4 Proprietà

Se una funzione $R(\tau)$ è l'autocorrelazione di un processo $x(t)$ SSL allora la sua trasformata di Fourier $S_x(\omega)$ è positiva, come visto. Se $R(\tau)$ è una funzione avente trasformata di Fourier positiva allora possiamo trovare un p.a. $x(t)$ che abbia autocorrelazione $R(\tau)$. Quindi condizione necessaria e sufficiente affinché una funzione $R(\tau)$ sia una autocorrelazione è la positività della sua trasformata di Fourier. Questa condizione (la positività della trasformata di Fourier) può anche essere espressa nel dominio del tempo, dove diventa

$$\sum_{i,j} a_i a_j^* R(\tau_i - \tau_j) \geq 0 \quad \forall a_i a_j \tau_i \tau_j \quad (7.88)$$

(non chiesta all'orale).

Dalla eq. 7.61 mediante operazione di modulo si ottiene che

$$|R(\tau)| = \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau} d\omega \right| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) |e^{j\omega\tau}| d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) d\omega = R(0) \quad (7.89)$$

il che implica che la funzione di autocorrelazione $R(\tau)$ è massima nell'origine: $R(\tau) \leq R(0) \forall \tau$. Se $R(\tau)$ non è periodica il suo massimo è solo nell'origine. Altrimenti vale il seguente

TEOREMA (non dimostrato): se $R(\tau_1) = R(0)$ per un qualche $\tau_1 \neq 0$ allora $R(\tau)$ è periodica di periodo τ_1 :

$$R(\tau + \tau_1) = R(\tau) \quad \forall \tau$$

COROLLARIO: se $R(\tau_1) = R(\tau_2) = R(0)$ per $\tau_1 \neq 0$ e $\tau_2 \neq 0$ e τ_1/τ_2 è irrazionale allora $R(\tau)$ è costante.

Correlazione incrociata: dati due p.a. SSL $x(t)$ e $y(t)$ la loro correlazione incrociata ha la seguente proprietà:

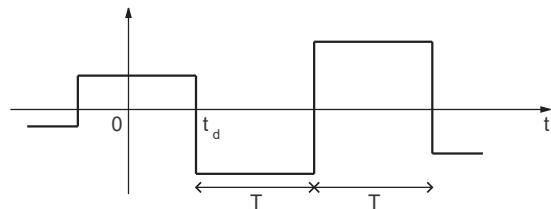
$$R_{xy}^2(\tau) \leq R_{xx}(0) R_{yy}(0) \quad (7.90)$$

7.8 Processo digitale casuale

Il p.a. digitale casuale (senza memoria) è definito come un insieme di impulsi rettangolari di durata T ed ampiezza a_k . Detto $p(t)$ il rettangolo alto 1 in $[0, T]$ e zero altrove, il p.a. digitale casuale è

$$x(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT - t_d)$$

L'ampiezza a_k è una v.a. discreta con media nulla e varianza σ_a^2 . La v.a. t_d , uniformemente distribuita fra 0 e T , rende il processo stazionario. In figura è mostrato un esempio (campione) del p.a. *Senza memoria* significa che le ampiezze nei diversi intervalli sono statisticamente indipendenti: $E\{a_j a_k\} = 0$ per $j \neq k$.



Se la v.a. a_k può assumere solo due diversi valori il p.a. si chiama **binario casuale**. Poichè $x(t) = a_k$ se t cade nel k -esimo intervallo, vale che:

$$\begin{aligned} E\{x(t)\} &= E\{a_k\} = 0 \\ E\{x^2(t)\} &= E\{a_k^2\} = \sigma_a^2 = \bar{a}_k^2 \end{aligned} \quad (7.91)$$

Per calcolare la funzione autocorrelazione di $x(t)$, $R_x(t_1, t_2)$, distinguiamo i casi in cui t_1 e t_2 cadono nello stesso intervallo da quelli in cui cadono in intervalli diversi. Se $|t_2 - t_1| > T$ allora t_1 e t_2 sono certamente in intervalli diversi e

$$E\{x(t_1)x(t_2)\} = E\{a_j a_k\} = 0 \quad (7.92)$$

Questo risultato vale anche se $|t_2 - t_1| < T$ ma t_1 e t_2 sono in intervalli diversi (adiacenti). Se t_1 e t_2 sono nello stesso intervallo, allora

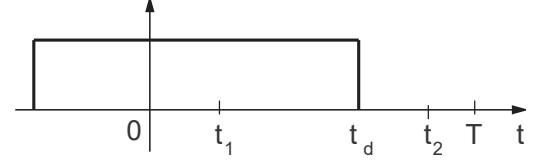
$$E\{x(t_1)x(t_2)\} = E\{a_k^2\}$$

Sempre nel caso $|t_2 - t_1| < T$ chiamiamo \mathcal{A} l'evento { t_1 e t_2 sono in intervalli adiacenti}: chiaramente

$$E\{x(t_1)x(t_2)\} = E\{a_j a_k\} P(\mathcal{A}) + E\{a_k^2\} [1 - P(\mathcal{A})] = E\{a_k^2\} [1 - P(\mathcal{A})]$$

Per calcolare $P(\mathcal{A})$ basta considerare l'intervalllo $0 < t < T$ in figura, dato che $P(\mathcal{A})$ ha lo stesso valore in qualsiasi intervallo di lunghezza T .

Se $t_1 < t_2$, allora t_1 e t_2 sono in intervalli adiacenti se $t_1 < t_d < t_2$ e



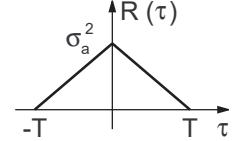
$$P(\mathcal{A}) = P\{t_1 < t_d < t_2\} = \int_{t_1}^{t_2} \frac{1}{T} dt_d = \frac{t_2 - t_1}{T}$$

Ovviamente se $t_2 < t_1$ si ottiene

$$P(\mathcal{A}) = \frac{t_1 - t_2}{T}$$

Quindi

$$E\{x(t_1)x(t_2)\} = \sigma_a^2 \left[1 - \frac{|t_1 - t_2|}{T} \right] \quad |t_1 - t_2| < T \quad (7.93)$$



e zero altrove. Attenzione: l'altezza del triangolo vale $E\{a_k^2\}$ che, nel caso in esame, è uguale a σ_a^2 essendo gli a_k a media nulla.

L'autocorrelazione $R_x(\tau)$ dipende solo dalla differenza $\tau = t_1 - t_2$ ed è in figura.

Esercizio 7.7 Trovare $R_x(\tau)$ e η_x del p.a. binario casuale $x(t)$ sapendo che a_k assume i valori $\pm V$ con uguale probabilità ed in modo indipendente in ogni intervallo.

Esercizio 7.8 Ripetere l'esercizio precedente se a_k assume i valori 0 e A con uguale probabilità ed in modo indipendente in ogni intervallo. NOTA: questo esercizio è più complicato del precedente, si risolve o direttamente o utilizzando la sovrapposizione (vedi prossimo argomento).

7.9 P.a. combinazione di altri p.a.

Alcuni p.a. possono essere visti come combinazione di altri processi. Vedremo i casi di p.a. ottenuto da somma di due altri p.a. e di p.a. ottenuto tramite prodotto.

7.9.1 Sovrapposizione

Supponiamo che il p.a. $z(t)$ si ottenga per somma o differenza di due p.a.

$$z(t) = x(t) \pm w(t)$$

Assumiamo che i p.a. $x(t)$ e $w(t)$ siano congiuntamente stazionari, oltre che individualmente stazionari, in modo tale che

$$R_{xw}(t_1, t_2) = R_{xw}(t_2 - t_1)$$

Allora

$$R_z(\tau) = R_x(\tau) + R_w(\tau) \pm [R_{xw}(\tau) + R_{wx}(\tau)]$$

e

$$S_z(f) = S_x(f) + S_w(f) \pm [S_{xw}(f) + S_{wx}(f)]$$

dove $S_{xw}(f) = FT\{R_{xw}(\tau)\}$ è la densità spettrale incrociata.

Se $x(t)$ e $w(t)$ sono scorrelati e $\eta_x \eta_w = 0$ allora

$$R_{xw}(\tau) = R_{wx}(\tau) = 0 \quad (7.94)$$

Sotto queste condizioni

$$\begin{aligned} R_z(\tau) &= R_x(\tau) + R_w(\tau) \\ S_z(f) &= S_x(f) + S_w(f) \\ E\{z^2(t)\} &= E\{x^2(t)\} + E\{w^2(t)\} \end{aligned} \quad (7.95)$$

Ovvero si ha sovrapposizione delle autocorrelazioni, delle densità spettrali e delle potenze medie. Quando valgono le eq. 7.94 e 7.95 i segnali casuali (p.a.) $x(t)$ e $w(t)$ sono detti **INCOERENTI**. I segnali in uscita da due sorgenti indipendenti sono usualmente incoerenti e la sovrapposizione delle potenze medie è un effetto fisico comune. Ad esempio, se due musicisti suonano all'unisono ma senza una sincronizzazione perfetta, la potenza media acustica raddoppia. Se fossero perfettamente sincronizzati la potenza media sarebbe quadruplicata.

7.9.2 Modulazione

È definita dal prodotto

$$z(t) = x(t) \cos(\omega_c t + \Phi) \quad (7.96)$$

dove $x(t)$ è un p.a. stazionario e Φ un angolo casuale (v.a.) indipendente da $x(t)$ ed uniformemente distribuito su di un intervallo lungo 2π . Se non inserissimo Φ il processo $z(t)$ non sarebbe stazionario: includere Φ tiene semplicemente in conto la scelta arbitraria dell'origine dei tempi quando $x(t)$ e $\cos(\omega_c t)$ provengono da sorgenti indipendenti.

Essendo la modulazione tempo variante dobbiamo trovare $R_z(t_1, t_2) = E\{z(t_1)z(t_2)\}$ considerando il valore atteso rispetto sia ad $x(t)$ che a Φ . Dopo aver espanso il coseno otteniamo:

$$R_z(t_1, t_2) = \frac{1}{2}E\{x(t_1)x(t_2)\}\{\cos[\omega_c(t_1-t_2)]+E[\cos(\omega_c t_1+\omega_c t_2+2\Phi)]\} = \frac{1}{2}R_x(t_1, t_2)\cos[\omega_c(t_1-t_2)]$$

e quindi, se $\tau = t_1 - t_2$,

$$R_z(\tau) = \frac{1}{2}R_x(\tau)\cos(\omega_c\tau) \quad (7.97)$$

La sua trasformata di Fourier vale (se $\omega_c = 2\pi f_c$)

$$S_z(f) = \frac{1}{4}[S_x(f-f_c) + S_x(f+f_c)] \quad (7.98)$$

che è la traslazione dello spettro di $x(t)$ attorno a $\pm f_c$.

La modulazione è un caso particolare della combinazione di due p.a. mediante prodotto:

$$z(t) = x(t)y(t) \quad (7.99)$$

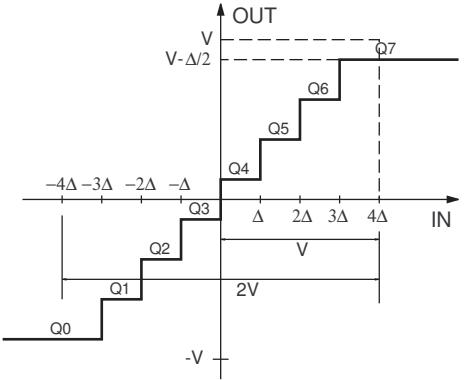
Se i processi $x(t)$ e $y(t)$ sono indipendenti e stazionari (congiuntamente ed individualmente) vale che:

$$\begin{aligned} R_z(\tau) &= R_x(\tau) R_y(\tau) \\ S_z(f) &= S_x(f) * S_y(f) \end{aligned} \quad (7.100)$$

7.10 Rumore di quantizzazione

La quantizzazione di un segnale campionato introduce un errore sulle ampiezze, come visto in sez. 2.2.

Si consideri il quantizzatore uniforme a $L = 2^3 = 8$ livelli in figura. Il segnale da campionare $x(t)$ ha ampiezza compresa fra $-V$ e V . Il valore da quantizzare (IN) viene portato al livello corrispondente (OUT) all'intervalle nel quale cade. Con 8 livelli ho che $\Delta = V/4$. Il livello di ricostruzione è a metà intervallo: se il valore da quantizzare cade nell'intervallo $[0, \Delta]$ il valore in uscita è $Q_4 = \Delta/2$, se cade in $[3\Delta, 4\Delta]$ diventa $Q_7 = 3.5\Delta$. Da notare che conviene estendere i due livelli estremi per gestire al meglio anche eventuali valori al di fuori dell'intervallo $[-V, V]$.



L'esempio sopra riguarda un classico quantizzatore uniforme (anche se solitamente hanno più livelli). Sono possibili altre scelte, come ad esempio spostare il livello di ricostruzione ad uno degli estremi dell'intervallo, ma poiché sono meno efficienti non verranno trattati.

Per ogni valore di ingresso x_k si otterrà quindi un valore di uscita

$$y_k = x_k + n_q$$

essendo n_q l'errore introdotto dal quantizzatore, di ampiezza ovviamente compresa fra $-\Delta/2$ e $\Delta/2$. Nell'ipotesi (ragionevole) che il segnale abbia una distribuzione simmetrica e che il numero di livelli di quantizzazione sia sufficientemente elevato, il rumore di quantizzazione ha, con un'ottima approssimazione, una distribuzione uniforme nell'intervallo $[-\Delta/2, \Delta/2]$. Quindi

$$\overline{n_q^2} = \int_{-\infty}^{\infty} n_q^2 f_n(n_q) dn_q = \int_{-\Delta/2}^{\Delta/2} n_q^2 \frac{1}{\Delta} dn_q = \frac{\Delta^2}{12} = \frac{V^2}{3L^2}$$

l'ultimo passaggio deriva dal fatto che $\Delta = 2V/L$. Ovvimente $\overline{n_q^2} = N_q$, ovvero è la potenza media del rumore di quantizzazione, che ha valor medio nullo (nelle ipotesi in oggetto).

Esempio 7.9 I CD musicali contengono i campioni del segnale audio: essi sono ottenuti, per ciascuno dei due canali stereo, utilizzando una frequenza di campionamento di 44 KHz ed un quantizzatore uniforme a 16 bit, ovvero con un numero di livelli $L = 2^{16} = 65536$. La corretta valutazione del rumore di quantizzazione si ottiene calcolando il rapporto fra la potenza di segnale e quella di rumore. Supponendo che i campioni del segnale possano essere modellati mediante una v.a. x uniformemente distribuita fra $-V$ e V , la potenza di segnale vale

$$S = E\{\mathbf{x}^2\} = \int_{-V}^V x^2 \frac{1}{2V} dx = \frac{V^2}{3}$$

Utilizzando il valore della potenza di rumore appena calcolato possiamo allora scrivere che

$$\left(\frac{S}{N_q}\right) = \frac{V^2}{3} \frac{3L^2}{V^2} = L^2 = 4294967296 \approx 4.3 \cdot 10^9$$

Ovvero che la potenza di segnale è più di 4 miliardi di volte quella del rumore. Esprimendo il risultato in dB (che verranno definiti nel prossimo capitolo) si ottiene un valore di 96,3 dB.

Un risultato lievemente diverso si può ottenere ipotizzando che il segnale sia modellabile mediante una sinusoide di ampiezza massima V , ovvero tramite il processo aleatorio sinusoidale $x(t) = V \cos(\omega_c t + \theta)$, con V e ω_c fissati e θ v.a. uniforme in $[0, 2\pi]$. Come visto, la potenza di segnale diventa

$$S = E\{x^2(t)\} = \frac{V^2}{2}$$

e il nuovo rapporto segnale/rumore è

$$\left(\frac{S}{N_q}\right) = \frac{V^2}{2} \frac{3L^2}{V^2} = \frac{3L^2}{2} = 6442450944 \approx 6.4 \cdot 10^9 \approx 98 \text{ dB}$$

Esiste una formula approssimata secondo la quale il rapporto segnale/rumore dovuto alla quantizzazione si può esprimere in modo approssimato in dB in funzione del numero n di bit del quantizzatore uniforme mediante

$$\left(\frac{S}{N_q} \right) \approx 6n + 1.76 \quad \text{dB}$$

Applicata al caso che stiamo analizzando produce il valore di 97.76 dB.

I risultati appena calcolati sono in realtà ottimistici, nel senso che la potenza media di un segnale audio di ampiezza massima V è normalmente più bassa (dalle parti di $S = V^2/100$). Questo non ha molta importanza, visto che un segnale ad alta fedeltà richiede che il rapporto segnale/rumore sia almeno nell'ordine del milione.

Capitolo 8

Rumore e Modulazioni

8.1 Guadagno, attenuazione e decibel

Dato un sistema, possiamo definire il suo guadagno in potenza come il rapporto fra la potenza in uscita e quella in ingresso:

$$g \triangleq \frac{P_{out}}{P_{in}} \quad (8.1)$$

La definizione è spesso applicata a sistemi senza distorsioni, ovvero quelli nei quali il segnale in uscita è una copia esatta, a meno di una costante moltiplicativa, del segnale in ingresso. La definizione tuttavia è valida anche per sistemi distorcenti.

Poichè i valori di g possono essere molto grandi è conveniente esprimere in scala logaritmica usando i decibel:

$$g_{dB} \triangleq 10 \log_{10} g = 10 \log_{10} \frac{P_{out}}{P_{in}} \quad (8.2)$$

dove la "B" in dB si scrive maiuscola in onore di Alexander Graham Bell (il primo che ha utilizzato la misura delle potenze in scala logaritmica).

Il guadagno di potenza è ovviamente sempre positivo o nullo, ma il suo valore espresso in dB può essere negativo: questo accade quando $g < 1$. Il valore di 0 dB corrisponde ad un guadagno unitario. L'eq. 8.2 si può ovviamente invertire:

$$g = 10^{(g_{dB}/10)} \quad (8.3)$$

I decibel rappresentano rapporti di potenza, ma anche la potenza di un segnale può essere espressa in dB rapportandola a un watt o a un milliwatt:

$$P_{dBW} = 10 \log_{10} \frac{P}{1W} \quad P_{dBm} = 10 \log_{10} \frac{P}{1mW} \quad (8.4)$$

Usando i dBm possiamo allora esprimere le relazioni di potenza ingresso uscita come

$$P_{out\ dBm} = g_{dB} + P_{in\ dBm} \quad (8.5)$$

In tabella sono riportati alcuni valori tipici.

Da notare che un raddoppio del rapporto P_{out}/P_{in} causa un incremento di 3 dB. Ovvero, se

$$10 \log_{10} \frac{P_{out}}{P_{in}} = \alpha$$

allora

$$10 \log_{10} \frac{2P_{out}}{P_{in}} = \alpha + 3$$

Valore in dB	Valore corrispondente
3 dB	$\rightarrow 10^{0.3} = 1.995 \dots \approx 2$
6 dB	$\rightarrow 10^{0.6} = 3.981 \dots \approx 4$
10 dB	$\rightarrow 10^1 = 10$
20 dB	$\rightarrow 10^2 = 100$
30 dB	$\rightarrow 10^3 = 1000$
60 dB	$\rightarrow 10^6 = 1000000$
90 dB	$\rightarrow 10^9 = 1000000000$

Attenzione: essendo legate al rapporto di due grandezze, le misure in dB sono **adimensionali**.

In generale si assume che i rapporti in dB siano sempre relativi a energie o potenze, anche partendo da altre grandezze da cui energia e potenza dipendono non linearmente, come tensioni o correnti. Questo introduce nei calcoli un fattore 20 che può creare confusione.

È ciò che accade quando si devono trattare rapporti in dB fra due grandezze indicanti tensioni o correnti elettriche, per esprimere una amplificazione di tensione o di corrente. Infatti in questi

casi non si intende il rapporto fra le grandezze stesse, ma fra le potenze che le tensioni o le correnti svilupperebbero se applicate a una medesima impedenza. Quindi, essendo la potenza P proporzionale al quadrato della tensione V o della corrente I , sfruttando le proprietà dei logaritmi si ricavano ed utilizzano le formule seguenti:

$$\text{Rapporto di potenze}_{dB} = 10 \log_{10} \frac{P_{out}}{P_{in}} = 10 \log_{10} \left(\frac{V_{out}}{V_{in}} \right)^2 = 20 \log_{10} \frac{V_{out}}{V_{in}}$$

$$\text{Rapporto di potenze}_{dB} = 10 \log_{10} \frac{P_{out}}{P_{in}} = 10 \log_{10} \left(\frac{I_{out}}{I_{in}} \right)^2 = 20 \log_{10} \frac{I_{out}}{I_{in}}$$

Come conseguenza un blocco di amplificazione può avere il guadagno espresso in dB senza che sia necessario specificare se tale guadagno sia in ampiezza o in potenza.

Facciamo un esempio: supponiamo che un amplificatore guadagni 10 in tensione, ovvero che $V_{out} = 10 V_{in}$. Il guadagno in potenza vale ovviamente 100, ovvero $P_{out} = 100 P_{in}$. Quindi devo specificare se il guadagno che indico è in tensione o in potenza. Se però dico che il guadagno vale 20 dB non è necessario che specifichi altro.

Attenuazione

Ogni mezzo trasmittivo passivo "consuma" la potenza del segnale che transita in esso. È allora preferibile parlare di attenuazione anziché di guadagno: l'attenuazione L è definita da

$$L \triangleq \frac{1}{g} = \frac{P_{in}}{P_{out}} \quad (8.6)$$

$$L_{dB} = -g_{dB} = 10 \log_{10} \frac{P_{in}}{P_{out}} \quad (8.7)$$

Nel caso di mezzi trasmittivi quali linee di trasmissione, cavi coassiali e guide d'onda la potenza in uscita diminuisce *esponenzialmente* con la distanza, secondo una relazione del tipo:

$$P_{out} = 10^{-(\alpha l/10)} P_{in}$$

dove l è la distanza fra sorgente e destinazione ed α è il *coefficiente di attenuazione* espresso in dB per unità di lunghezza, in generale dipendente dalla frequenza. Quindi

$$L = 10^{(\alpha l/10)} \quad L_{dB} = \alpha l \quad (8.8)$$

ovvero l'attenuazione espressa in dB è proporzionale alla distanza (lunghezza).

Nella tabella che segue sono riportati alcuni coefficienti di attenuazione tipici.

Mezzo trasmittivo	Frequenza	Attenuazione in dB / Km
Doppino o piattina, diametro 0.3 cm	1 KHz	0.05
Doppino ritorto	10 KHz	2
	100 KHz	3
	300 KHz	6
Cavo coassiale, diametro 1 cm	100 KHz	1
	1 MHz	2
	3 MHz	4
Cavo coassiale, diametro 15 cm	100 MHz	1.5
Guida d'onda rettangolare, 2 x 2.5 cm	10 GHz	5
Fibra ottica	4×10^{14} Hz	10

Esprimere l'attenuazione in dB nasconde, ad un primo sguardo, l'effettiva attenuazione del segnale. Ad esempio, supponiamo di avere un cavo lungo 30 Km con coefficiente di attenuazione pari a $\alpha = 3$ dB / Km. L'attenuazione totale è quindi $L_{dB} = 3 \cdot 30 = 90$ dB, ovvero $P_{out} = 10^{-9} P_{in}$. Se raddoppiamo la lunghezza otteniamo $L_{dB} = 180$ dB e $P_{out} = 10^{-18} P_{in}$. Questi valori vogliono dire che se uso una potenza in trasmissione di un megawatt (10^6 W) ottengo una potenza in ricezione di un picowatt (10^{-12} W). (*Naturalmente non è possibile inviare una simile potenza in un cavo!*)

Nel caso di propagazione libera (onde radio), considerando la propagazione che avviene in linea retta fra due antenne, l'attenuazione è dovuta alla dispersione sferica delle onde radio ed è proporzionale al quadrato della distanza, secondo

$$L = \left(\frac{4\pi l}{\lambda} \right)^2 = \left(\frac{4\pi f l}{c} \right)^2 \quad (8.9)$$

dove λ è la lunghezza d'onda, f la frequenza del segnale c la velocità della luce. Esprimendo c in km/sec ($3 \cdot 10^5$ km/sec), l deve essere in km e f in Hz.

Esprimendo l in kilometri e f in GHz (10^9 Hz) si ottiene

$$L_{dB} = 92.4 + 20 \log_{10} f + 20 \log_{10} l \quad (8.10)$$

Come evidente l'attenuazione (in dB) cresce ora come il logaritmo della distanza: raddoppiare la distanza provoca "solo" un aumento di 6 dB dell'attenuazione.

8.2 Rumore

Segnali elettrici indesiderati sono inevitabili in qualsiasi sistema di comunicazione e provengono da diversi tipi di sorgenti. Queste ultime possono essere classificate in due distinti gruppi: interferenze dovute ad attività umane e rumore di origine naturale. Le interferenze di origine umana provengono da altri sistemi di comunicazione, da motori a scoppio o elettrici, dall'apertura e chiusura di interruttori, dalla rete elettrica a 50 Hz, dai neon, dai sistemi di calcolo e da altre simili attività. Il rumore di origine naturale comprende i disturbi atmosferici (lampi), la radiazione extraterrestre, il rumore che si forma nei circuiti (rumore termico). Gli effetti dei segnali di disturbo possono essere ridotti o anche quasi eliminati per mezzo di una accurata progettazione, ma rimane sempre una certa quantità ineliminabile di rumore (segnali casuali) che rappresenta il limite fondamentale alle prestazioni dei sistemi di comunicazione. Questo significa che dato un certo canale (inevitabilmente rumoroso), è possibile trasmettere informazione su di esso in modo da rendere trascurabili gli effetti del rumore solo se la quantità di informazione che viaggia nell'unità di tempo è inferiore ad una certa soglia teoricamente determinabile, detta capacità del canale (vedi la *Teoria dell'Informazione*).

8.2.1 Rumore Termico

Il rumore termico è dovuto al movimento caotico di particelle elettricamente cariche (di solito elettroni) all'interno di mezzi conduttori. Dato un elemento metallico conduttore di resistenza R (misurata in Ω , cioè ohm) alla temperatura assoluta di T gradi Kelvin¹, il moto degli elettroni produce un rumore $v(t)$ (misurato in Volt) ai suoi estremi. Dalla meccanica quantistica ed in accordo con il teorema del limite centrale è noto che $v(t)$ è un processo aleatorio gaussiano a media nulla e varianza pari a

$$\sigma_v^2 = \frac{2(\pi k T)^2}{3h} R \quad (8.11)$$

dove k è la costante di Boltzmann, $k = 1.3807 \cdot 10^{-23}$ J/K (joule / gradi kelvin), e h è la costante di Planck, $h = 6.626 \cdot 10^{-34}$ Js (joule per secondo).

La densità spettrale di potenza vale

$$S_v(f) = 2R \left[\frac{h|f|}{2} + \frac{h|f|}{e^{h|f|/(kT)} - 1} \right] \quad (8.12)$$

A temperatura ambiente e per frequenze inferiori a 1000 GHz si ha che $[h|f|/(kT)] < 1/5$ e si può usare l'approssimazione $e^x \approx 1 + x$. La densità spettrale del rumore termico può allora essere approssimata da

$$S_v(f) = 2RkT \quad (8.13)$$

ovvero un rumore gaussiano bianco. L'unico problema di questa approssimazione è che il suo integrale sull'intero asse f vale ∞ (ovvero che $v(t)$ ha potenza infinita), il che è fisicamente impossibile. Il rumore termico viene comunque sempre filtrato dai sistemi reali, che hanno banda limitata.

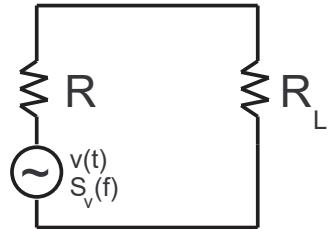
Nel seguito verrà utilizzata esclusivamente la versione semplificata 8.13, più che accurata per i casi di interesse pratico di frequenza al di sotto dei 100 GHz e temperature abbastanza lontane dallo zero assoluto.

¹I gradi Kelvin misurano il moto delle particelle: zero gradi Kelvin sono allo zero assoluto, dove le particelle sono ferme, e corrispondono a -273.15 gradi centigradi. Non esiste nell'universo una temperatura inferiore.

Possiamo ora caratterizzare una sorgente di rumore mediante la massima potenza di rumore che può essere trasferita ad un carico.

In figura è visualizzato il circuito equivalente di Thevenin di un resistore rumoroso R con carico R_L . La massima densità spettrale di potenza disponibile sul carico R_L si ottiene quando il carico è adattato ($R_L = R^*$) e vale

$$S_a(f) = \frac{S_v(f)}{4R} = \frac{1}{2} k\mathcal{T} \quad \text{W / Hz}$$



Quindi la potenza disponibile per una sorgente di rumore termico non dipende dal valore della resistenza R , ma solo dalla sua temperatura (la tensione a circuito aperto dipende invece da R).

La potenza disponibile in una certa banda di frequenza vale allora

$$P_a = \int_{-B}^B S_v(f) df = \int_{-B}^B \frac{1}{2} k\mathcal{T} df = k\mathcal{T} B \quad (8.14)$$

Vista la formula sopra è possibile caratterizzare una qualsiasi sorgente di rumore mediante la sua **temperatura equivalente di rumore**, definita da

$$\mathcal{T} = \frac{P_a}{kB} \quad (8.15)$$

essendo P_a la potenza disponibile della sorgente in una banda pari a B .

Da notare che \mathcal{T} rappresenta la temperatura di una sorgente di rumore termico, mentre per sorgenti di rumore generali la temperatura equivalente di rumore non ha alcun legame con la temperatura fisica del dispositivo.

Da notare infine come il rumore di origine termica sia tipicamente molto piccolo: ad esempio, una resistenza a temperatura ambiente produce una potenza disponibile di rumore su banda di 1 MHz pari a $P_a \approx 4 \cdot 10^{-15}$ W. Questo numero si ottiene dall'eq. 8.14 considerando che la **temperatura ambiente** è definita $T_0 = 290$ K, un po' freddina (circa 17° C) ma tale che $kT_0 \approx 4 \cdot 10^{-21}$.

In condizioni normali il rumore termico è quindi tipicamente trascurabile. Vi sono tuttavia dei casi in cui il segnale è così debole da essere paragonabile al rumore termico (ad esempio al ricevitore di sistemi di comunicazione a grande distanza).

8.2.2 Rumore bianco e rumore filtrato

Oltre alle resistenze vi sono molti altri tipi di sorgente di rumore che sono di tipo gaussiano ed hanno densità spettrale di potenza praticamente piatta (costante) su di un ampio intervallo di frequenze. Questo tipo di rumore ha componenti in uguale proporzione per tutte le frequenze ed è chiamato rumore bianco, in analogia con la luce bianca, che è formata da una uguale proporzione di frequenze luminose (colori).

La densità spettrale del rumore bianco è quindi

$$G(f) = \frac{\eta}{2} \quad (8.16)$$

dove il fattore 1/2 viene inserito per indicare che metà della potenza è associata alle frequenze positive e metà a quelle negative: per questo motivo si chiama densità spettrale di potenza *bilatera*. In alternativa si può definire la densità di potenza monolatera come $G(f) = \eta$, intendendo che vengono considerate le sole frequenze positive. Nel seguito considereremo solo densità bilatera.

Dato che $G(f)$ è nota è immediato calcolare la funzione di autocorrelazione, che è l'antitrasformata di Fourier della densità spettrale di potenza:

$$R(\tau) = \int_{-\infty}^{\infty} \frac{\eta}{2} e^{j f \tau} df = \frac{\eta}{2} \delta(\tau) \quad (8.17)$$

Poichè $R(\tau \neq 0) = 0$ due differenti campioni di rumore bianco (non importa quanto vicini) sono scorrelati e quindi (dato che il processo è gaussiano) statisticamente indipendenti. Da un punto di vista matematico il rumore bianco è allora il massimo esempio di segnale variabile e casuale. Se però ascoltiamo il rumore bianco (inviandolo ad un altoparlante) il risultato è un fruscio continuo e monotono: alcuni lo usano per conciliare il sonno. Infatti il cervello umano si è evoluto per distinguere le strutture presenti nel mondo circostante ed il rumore bianco semplicemente non ne ha.

Consideriamo ora un rumore gaussiano bianco $x(t)$ di densità spettrale $G_x(f) = \eta/2$ applicato all'ingresso di un filtro LTI avente funzione di trasferimento $H(f)$. La corrispondente uscita $y(t)$ è, come visto, un p.a. gaussiano descritto da:

$$G_y(f) = \frac{\eta}{2} |H(f)|^2 \quad (8.18)$$

$$R_y(\tau) = \frac{\eta}{2} \mathcal{F}^{-1}[|H(f)|^2] \quad (8.19)$$

$$\overline{y^2} = \frac{\eta}{2} \int_{-\infty}^{\infty} |H(f)|^2 df \quad (8.20)$$

L'eq. 8.18 mostra che la densità spettrale del rumore bianco filtrato assume la forma di $|H(f)|^2$: per questo si dice che il rumore in uscita è rumore *colorato* ed ha contenuto in frequenza nella banda passante del filtro. Come conseguenza la funzione di autocorrelazione perde il suo aspetto impulsivo e campioni diversi possono essere correlati fra loro.

Esempio 8.1 Sia $H(f)$ un filtro ideale passabasso di banda B (ovvero di altezza unitaria per $-B < f < B$ e zero altrove). Se l'ingresso è un rumore bianco con densità di potenza bilatera $G(f) = \eta/2$, la sua uscita è un rumore gaussiano colorato avente densità spettrale di potenza ed autocorrelazione

$$G_y(f) = \frac{\eta}{2} \prod \left(\frac{f}{2B} \right) \quad R_y(\tau) = \eta B \operatorname{sinc}(2B\tau)$$

Data la forma della funzione di autocorrelazione i campioni del rumore di uscita sono scorrelati solo se ottenuti a distanza $n/(2B)$ con n intero diverso da zero e correlati altrimenti.

Banda equivalente di rumore

Il rumore bianco filtrato ha solitamente potenza finita: per evidenziare questo fatto chiamiamo *potenza media di rumore* il valore $N = \overline{y^2}$ e scriviamo l'eq. 8.20 nella forma

$$N = \frac{\eta}{2} \int_{-\infty}^{\infty} |H(f)|^2 df = \eta \int_0^{\infty} |H(f)|^2 df$$

L'integrale dipende solo dalla risposta in frequenza del filtro, quindi possiamo semplificare la discussione sulla potenza di rumore definendo la banda equivalente di rumore B_N di un filtro come

$$B_N \triangleq \frac{1}{g} \int_0^{\infty} |H(f)|^2 df \quad (8.21)$$

dove $\sqrt{g} = |H(f)|_{max}$ è il guadagno del filtro, spesso l'ampiezza del filtro alla frequenza centrale. La potenza del rumore filtrato si può allora esprimere come

$$N = g\eta B_N \quad (8.22)$$

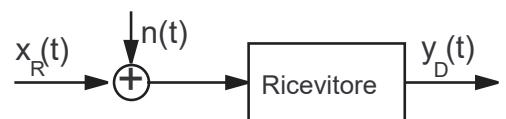
L'effetto del filtro viene quindi separato in due parti: la selettività in frequenza, descritta da B_N , e il guadagno in potenza (o attenuazione) rappresentato da g . Ovviamente B_N rappresenta la banda passante di un filtro ideale (a rettangolo di altezza g) che farebbe passare la stessa quantità di rumore del filtro reale. Per i filtri normali, B_N è un po' più grande della banda passante a 3 dB. Ad esempio, per un filtro passa basso del primo ordine B_N è circa il 50 % in più. La differenza diminuisce per filtri maggiormente selettivi, cosicché spesso si può approssimare B_N con la banda passante a 3 dB. Riassumendo, se $y(t)$ è un rumore gaussiano bianco a media nulla filtrato, allora

$$\overline{y^2} = \sigma_y^2 = N = g\eta B_N \quad \sigma_y = \sqrt{N} = \sqrt{g\eta B_N} \quad (8.23)$$

8.2.3 Rumore additivo e $(S/N)_D$

Il rumore che agisce nei sistemi di trasmissione ha solitamente un comportamento *additivo*, nel senso che il rumore si aggiunge al segnale che porta informazione in diversi punti del tragitto fra sorgente e destinazione.

Per semplificare l'analisi tutto il rumore viene concentrato in un unico segnale di rumore e viene aggiunto in un solo punto dello schema: all'ingresso del ricevitore.



Questo modello tiene conto del fatto che il punto più vulnerabile al rumore in un sistema di trasmissione è proprio l'ingresso del ricevitore, dove il segnale è più debole. Inoltre non è difficile trasformare sorgenti di rumore in altri punti del sistema in modo da poterle trattare come se fossero nel punto indicato.

Poichè il ricevitore è un sistema lineare, la somma di segnale e rumore all'ingresso produce in uscita la somma di segnale voluto e rumore:

$$y_D(t) = x_D(t) + n_D(t) \quad (8.24)$$

dove $x_D(t)$ e $n_D(t)$ significano rispettivamente segnale e rumore a destinazione. La potenza in uscita si ottiene come media di

$$E\{y_D^2(t)\} = E\{x_D^2(t) + 2x_D(t)n_D(t) + n_D^2(t)\} \quad (8.25)$$

Per calcolarla si tratta il segnale come il campione di un processo ergodico noto e si fanno le seguenti ipotesi sul rumore (del tutto ragionevoli):

1. il rumore proviene da una sorgente ergodica ed ha media nulla e densità spettrale di potenza nota $G_n(f)$;
2. il rumore è fisicamente indipendente dal segnale e quindi scorrelato.

Sotto queste condizioni la media del prodotto incrociato $x_D(t)n_D(t)$ vale zero, essendo $x_D(t)$ e $n_D(t)$ *incoerenti*. Quindi

$$\overline{y_D^2} = E\{y_D^2(t)\} = E\{x_D^2(t) + n_D^2(t)\} = \overline{x_D^2} + \overline{n_D^2} \quad (8.26)$$

ovvero abbiamo la sovrapposizione delle potenze di segnale e di rumore a destinazione.

Definendo ora $S_D \triangleq \overline{x_D^2}$ e $N_D \triangleq \overline{n_D^2}$ otteniamo che il rapporto segnale/rumore a destinazione è

$$\left(\frac{S}{N}\right)_D = \frac{S_D}{N_D} = \overline{x_D^2} / \overline{n_D^2} \quad (8.27)$$

Tale valore può essere misurato sperimentalmente in questo modo: si spegne il segnale e si misura la potenza del rumore a destinazione N_D , si riaccende il segnale e (visto che il rumore non si può spegnere!) si misura $S_D + N_D$. Il rapporto delle due misure fornisce $(S_D + N_D)/N_D = (S/N)_D + 1$. Molto spesso la densità spettrale di rumore si assume costante e pari a $G_n(f) = \eta/2$, ovvero si assume che il rumore sia bianco. Se il ricevitore ha banda equivalente di rumore B_N e guadagno g_R , la potenza del rumore a destinazione vale allora

$$N_D = g_R \eta B_N \quad (8.28)$$

La densità di rumore η può anche essere espressa in termini di temperatura equivalente di rumore \mathcal{T}_N all'ingresso del ricevitore, ovvero

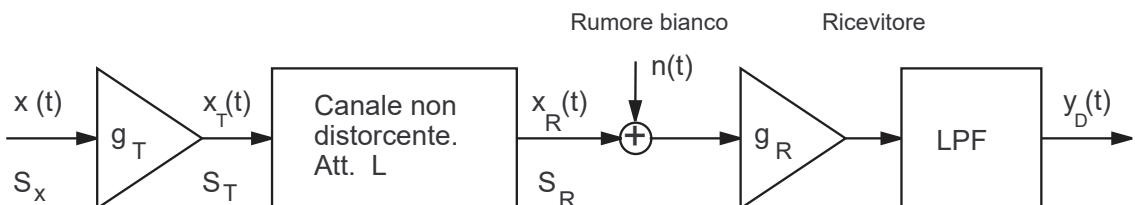
$$\eta = k \mathcal{T}_N = k \mathcal{T}_0 (\mathcal{T}_N/\mathcal{T}_0) \approx 4 \cdot 10^{-21} (\mathcal{T}_N/\mathcal{T}_0) \text{ W/Hz} \quad (8.29)$$

essendo $\mathcal{T}_0 = 290$ K. Valori tipici di \mathcal{T}_N variano da circa $0.2\mathcal{T}_0$ (60 K) per sistemi a basso rumore a $10\mathcal{T}_0$ o più per sistemi rumorosi.

Se il rumore bianco è anche gaussiano si indica sinteticamente come AWGN (*additive white gaussian noise*).

8.2.4 Trasmissione analogica in banda base

Un semplice sistema per la trasmissione analogica in banda base può essere schematizzato come in figura:



I due blocchi triangolari sono ovviamente amplificatori di guadagno g_T e g_R : in questo schema i guadagni sono in potenza, per cui, ad esempio, $S_T = g_T S_x$. Il guadagno in tensione è allora dato dalla radice quadrata del guadagno in potenza, $x_T(t) = \sqrt{g_T} x(t)$. Se il guadagno è espresso in dB non è necessario specificare se si intende in potenza o in tensione (perchè è uguale).

L'ingresso $x(t)$ viene generato da una sorgente di informazione e lo scopo è di ottenere una copia dello stesso in uscita. La sorgente viene modellata come un processo aleatorio ergodico avente *larghezza di banda* W , ovvero tale che ogni campione $x(t)$ ha contenuto spettrale trascurabile per $|f| > W$.

Il canale è assunto essere senza distorsione sulla banda del segnale, cosicchè la sua uscita è pari all'ingresso ritardato e attenuato di un fattore L :

$$x_R(t) = \sqrt{\frac{g_T}{L}} x(t - t_d)$$

Il rumore è assunto essere AWGN con densità spettrale bilatera $G_n(f) = \eta/2$. Il ricevitore è composto da un amplificatore di guadagno g_R seguito da un filtro passabasso: la sua uscita è la funzione (processo aleatorio) $y_D(t) = x_D(t) + n_D(t)$. Ci interessa calcolare il rapporto segnale/rumore a destinazione. Le potenze medie in trasmissione e ricezione sono legate da:

$$\begin{aligned} S_x &\triangleq \overline{x^2} & S_T &= g_T \overline{x^2} = g_T S_x \\ S_R &= \overline{x_R^2} = S_T/L & S_D &= \overline{x_D^2} = g_R S_R \end{aligned} \quad (8.30)$$

Il valore di S_D assume che il filtro finale passabasso abbia risposta unitaria nella banda del segnale. Il filtro ha il compito fondamentale di eliminare le componenti di rumore al di fuori della banda del segnale: in condizioni ideali è quindi un filtro LPF di forma rettangolare (in frequenza) con banda $B = W$ e altezza unitaria. La potenza di rumore a destinazione vale allora:

$$N_D = g_R \eta W$$

ed il rapporto segnale/rumore a destinazione è

$$\left(\frac{S}{N}\right)_D = \frac{S_R}{\eta W} = \gamma \quad (8.31)$$

Il risultato appena ottenuto mostra che il rapporto segnale/rumore a destinazione dipende da tre parametri fondamentali: la potenza del segnale ricevuto, la densità del rumore AWGN e la banda del segnale W . Questo risultato è considerato il riferimento per la qualità delle trasmissioni, ed è di uso così comune che gli è stato assegnato un nome, ovvero γ (gamma).

Da notare che ovviamente il valore di g_R non varia il risultato, poichè l'amplificatore finale agisce in modo uguale sia sul segnale che sul rumore. Inoltre, poichè il rumore dipende dalla banda su cui è calcolato, un segnale a banda larga è più degradato di uno a banda stretta (a pari η).

Essendo un rapporto di potenze, il valore di $(S/N)_D$ è solitamente espresso in dB. Esprimendo la potenza di segnale in milliwatt (o dBm) ed il rumore in termini di temperatura equivalente di rumore \mathcal{T}_N vale che:

$$\left(\frac{S}{N}\right)_{DdB} = 10 \log_{10} \left(\frac{S_R}{k\mathcal{T}_N W} \right) = S_{RdBm} + 174 - 10 \log_{10} \left(\frac{\mathcal{T}_N}{\mathcal{T}_0} W \right) \quad (8.32)$$

dove il valore 174 deriva dalla conversione dell'eq. 8.29 in milliwatt.

Infatti

$$\begin{aligned} 10 \log_{10} \left(\frac{S_R}{k\mathcal{T}_N W} \right) &= 10 \left[\log_{10} S_R - \log_{10} \left(k\mathcal{T}_0 \frac{\mathcal{T}_N}{\mathcal{T}_0} W \right) \right] \\ 10 \log_{10} \left(\frac{S_R}{k\mathcal{T}_N W} \right) &= 10 \left[\log_{10} S_R - \log_{10} (k\mathcal{T}_0) - \log_{10} \left(\frac{\mathcal{T}_N}{\mathcal{T}_0} W \right) \right] \end{aligned}$$

dove

$$\log_{10} (k\mathcal{T}_0) \approx \log_{10} (4 \cdot 10^{-21}) \approx -20.4$$

e poichè

$$S_{RdBm} = 10 \log_{10} \left(\frac{S_R}{10^{-3}} \right) = 10 \log_{10}(S_R) + 30$$

si ottiene

$$\left(\frac{S}{N}\right)_{DdB} = S_{RdBm} - 30 + 204 - 10 \log_{10} \left(\frac{\mathcal{T}_N}{\mathcal{T}_0} W \right)$$

e la 8.32 segue.

La tabella che segue, presa da [Car86], indica i valori del rapporto segnale/rumore a destinazione e della banda richiesti per alcune tipiche trasmissioni analogiche:

Tipo di segnale	Intervallo di frequenza	$(S/N)_D$ in dB
Segnale vocale appena intellegibile	500 Hz - 2 KHz	5 - 10
Segnale vocale di qualità telefonica	200 Hz - 3.2 KHz	25 - 35
Audio a qualità radio (AM)	100 Hz - 5 KHz	40 - 50
Audio ad alta fedeltà	20 Hz - 20 KHz	55 - 65
Segnale televisivo (analogico)	50 Hz - 5 MHz	45 - 55

In termini di potenza trasmessa il rapporto $(S/N)_D$ si può anche esprimere come:

$$\left(\frac{S}{N}\right)_D = \frac{S_T}{L\eta W} \quad (8.33)$$

Da notare infine che i risultati esposti sono basati sulle ipotesi di rumore AWGN, di trasmissione senza distorsioni e di filtri ideali e sono quindi da intendersi come limiti teorici, anche se tipicamente i valori raggiungibili non sono molto lontani.

8.2.5 Ripetitori analogici

Quando i parametri nell'eq. 8.33 sono fissati e la potenza in trasmissione richiesta risulta troppo alta, una possibile soluzione è quella di usare ripetitori. Per farlo si divide il percorso del segnale in m tratte inserendo un sistema ricevitore/trasmettitore alla fine di ciascuna tratta. Se supponiamo di dividere in m parti uguali e che l'attenuazione di ciascuna tratta sia L_1 avremo un rapporto segnale/rumore alla fine di ciascun tratto pari a

$$\left(\frac{S}{N}\right)_1 = \frac{S_T}{L_1\eta W} \quad (8.34)$$

I ripetitori sono solitamente progettati in modo che il guadagno del trasmettitore sia uguale alla perdita di tratta: in queste condizioni si può dimostrare che il rapporto segnale/rumore a destinazione vale circa (vedi [Car86] appendice A)

$$\left(\frac{S}{N}\right)_D \approx \frac{1}{m} \left(\frac{S}{N}\right)_1 = \frac{L}{mL_1} \left(\frac{S_T}{L\eta W}\right) \quad (8.35)$$

Il vantaggio è ovviamente che l'attenuazione cresce molto più che linearmente con la distanza, cosicché $L/(mL_1) \gg 1$.

Esempio 8.2 Consideriamo un sistema di trasmissione via cavo con attenuazione totale di $L = 140$ dB = 10^{14} e $\mathcal{T}_N = 5\mathcal{T}_0$. Supponiamo di voler trasmettere un segnale audio ad alta fedeltà di banda $W = 20$ kHz e di volere un rapporto segnale/rumore a destinazione $(S/N)_D \geq 60$ dB.

La potenza di segnale al ricevitore si calcola usando l'eq. 8.32

$$S_{RdBm} + 174 - 10 \log_{10}(5 \cdot 20 \cdot 10^3) \geq 60 \text{ dB}$$

cioè $S_R \geq -64$ dBm $\approx 4 \cdot 10^{-7}$ mW e $S_T = LS_R \geq 4 \cdot 10^7$ mW = 40000 W. Anche senza essere particolarmente esperti in sistemi di trasmissione è evidente che non è possibile usare una simile potenza in trasmissione in un cavo!

Supponiamo che sia possibile inserire un ripetitore a metà strada. L'attenuazione di ciascuna tratta risulta essere $L_1 = 70$ dB = 10^7 , visto che l'attenuazione in un cavo espressa in dB è direttamente proporzionale alla sua lunghezza. Il miglioramento risultante (eq. 8.35) è

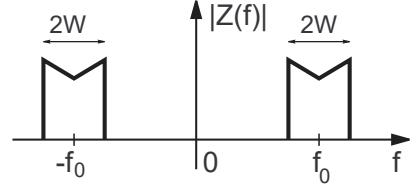
$$\frac{L}{mL_1} = \frac{10^{14}}{2 \cdot 10^7} = 5 \cdot 10^6$$

che riduce la potenza in trasmissione richiesta a $S_T \geq 8$ mW, che è un valore molto più ragionevole (S_T fra 10 e 20 mW fornisce un corretto margine di sicurezza).

8.3 Rappresentazione di segnali passa-banda

Consideriamo un segnale deterministico $z(t)$ con trasformata di Fourier $Z(f)$. Diremo che $z(t)$ è un segnale *passa-banda* se $Z(f)$ è non trascurabile solo in una banda di frequenza di ampiezza $2W$ centrata intorno alla frequenza $\pm f_0$ (vedi figura). La frequenza $f_0 > 0$ è la frequenza centrale dell'intervallo (o frequenza della portante, come vedremo fra poco).

La quantità W è detta la *larghezza di banda* (monolatera) del segnale; se $f_0 \gg W$ il segnale è detto a banda stretta. Se invece $Z(f)$ ha componenti non trascurabili intorno alla frequenza zero il segnale è detto *passa-basso*. Molti dei segnali passa-banda di interesse per le comunicazioni sono segnali ottenuti modulando una portante con un segnale passa-basso che contiene l'informazione.



Noi vogliamo trovare una relazione fra il segnale modulato passa-banda ed il segnale passa-basso che si ottiene dopo la demodulazione. Per fare questo ricordiamo la trasformata di Hilbert vista in sezione 7.7.2. Sia $\check{z}(t)$ la trasformata di Hilbert di $z(t)$: come visto, trasformare due volte $z(t)$ produce ancora $z(t)$ cambiato di segno (e con la componente continua rimossa).

Definiamo *pre-inviluppo* del segnale $z(t)$ il segnale

$$z_+(t) = z(t) + j\check{z}(t) \quad (8.36)$$

Da notare che $z_+(t)$ è un segnale complesso, poiché sia $z(t)$ che $\check{z}(t)$ sono reali. La caratteristica importante di $z_+(t)$ si vede bene in frequenza: sia

$$Z_+(f) = \mathcal{F}[z_+(t)] \quad (8.37)$$

la trasformata di Fourier di $z_+(t)$. Allora

$$Z_+(f) = Z(f) + j(-j\text{sgn}f)Z(f) = \begin{cases} 2Z(f) & f > 0 \\ Z(0) & f = 0 \\ 0 & f < 0 \end{cases} \quad (8.38)$$

ovvero il pre-inviluppo di un segnale non ha contenuto in frequenza (la sua trasformata di Fourier è nulla) per frequenza negativa.

Consideriamo il significato fisico del pre-inviluppo. Se $z(t)$ è un segnale passa-banda centrato su $\pm f_0$, può essere conveniente esprimere nella forma

$$z(t) = x(t) \cos(2\pi f_0 t) - y(t) \sin(2\pi f_0 t) \quad (8.39)$$

o anche

$$z(t) = a(t) \cos[2\pi f_0 t + \theta(t)] \quad (8.40)$$

essendo i segnali $x(t)$, $y(t)$, $a(t)$ e $\theta(t)$ tutti passa-basso. I segnali $x(t)$ e $y(t)$ sono le componenti in fase e in quadratura del segnale passa-banda $z(t)$, mentre $a(t)$ e $\theta(t)$ sono l'ampiezza e la fase. Le relazioni fra questi segnali sono:

$$a(t) = \sqrt{x^2(t) + y^2(t)} \quad (8.41)$$

$$\theta(t) = \sin^{-1} \left[\frac{y(t)}{a(t)} \right] = \cos^{-1} \left[\frac{x(t)}{a(t)} \right] \quad (8.42)$$

Da notare che possiamo esprimere $z(t)$ mediante le equazioni 8.39 o 8.40 per *ogni* valore della frequenza f_0 . Se f_0 è proprio la frequenza centrale (che a volte coincide con la portante) allora $a(t)$ è l'uscita di un demodulatore ad inviluppo. Vedremo che la funzione $a(t)$ è indipendente dal valore scelto per la frequenza f_0 : per questo motivo $a(t)$ è detta inviluppo di $z(t)$. Se f_0 è uguale alla frequenza centrale, e solo per questo valore, la fase $\theta(t)$ non contiene termini che crescono linearmente con il tempo.

Le funzioni $x(t)$ e $y(t)$ possono essere ricavate da $z(t)$ per qualsiasi scelta di f_0 per mezzo della demodulazione sincrona. Vediamo come: assumiamo che

$$z(t) \leftrightarrow Z(f)$$

sia tale che $Z(f)$ sia diversa da zero solo per $|f \pm f_0| < f_0$. Allora

$$z_+(t) \leftrightarrow Z_+(f)$$

è diversa da zero solo per $|f - f_0| < f_0$, avendo $Z_+(f)$ valori non nulli solo per frequenze positive. Traslando in frequenza di f_0 si ottiene un nuovo segnale

$$v(t) = z_+(t) e^{-j2\pi f_0 t} \quad (8.43)$$

la cui trasformata ha valori non nulli solo per $|f| < f_0$: $v(t)$ è quindi un segnale passa-basso. Consideriamo ora $z(t)$ espressa in termini di $v(t)$:

$$\begin{aligned} z(t) &= \operatorname{Re}\{z(t) + j\dot{z}(t)\} = \operatorname{Re}\{z_+(t)\} \\ &= \operatorname{Re}\{v(t) e^{j2\pi f_0 t}\} \\ &= \operatorname{Re}\{v(t)\} \cos 2\pi f_0 t - \operatorname{Im}\{v(t)\} \sin 2\pi f_0 t \end{aligned} \quad (8.44)$$

Essendo $v(t)$ un segnale passa-basso, anche $\operatorname{Re}\{v(t)\}$ e $\operatorname{Im}\{v(t)\}$ lo sono; dal confronto con l'eq. 8.39 si vede che

$$x(t) = \operatorname{Re}\{v(t)\} = \operatorname{Re}\{e^{-j2\pi f_0 t} z_+(t)\} \quad (8.45)$$

$$y(t) = \operatorname{Im}\{v(t)\} = \operatorname{Im}\{e^{-j2\pi f_0 t} z_+(t)\} \quad (8.46)$$

e quindi

$$a(t) = \sqrt{x^2(t) + y^2(t)} = |v(t)| = |e^{-j2\pi f_0 t} z_+(t)| = |z_+(t)| \quad (8.47)$$

Il segnale $a(t)$, che corrisponde all'uscita di un demodulatore ad inviluppo, si chiama inviluppo del segnale $z(t)$: siccome $a(t) = |z_+(t)|$ il segnale $z_+(t)$ è detto pre-inviluppo di $z(t)$. Dato che $|e^{-j2\pi f_0 t}| \equiv 1$, l'inviluppo è indipendente dal valore scelto per f_0 .

Da quanto visto si ha anche che

$$\theta(t) = \tan^{-1} \frac{\operatorname{Im}[e^{-j2\pi f_0 t} z_+(t)]}{\operatorname{Re}[e^{-j2\pi f_0 t} z_+(t)]} = \angle z_+(t) e^{-j2\pi f_0 t} = \angle v(t) \quad (8.48)$$

La rappresentazione per mezzo delle componenti in quadratura di $z(t)$ può quindi essere ottenuta da $z_+(t)$ in modo univoco. È immediato vedere che vale anche l'opposto: le componenti in quadratura determinano in modo univoco $z_+(t)$.

SINTESI

Quanto visto si può sintetizzare nel modo seguente: dato un *qualsiasi* segnale passa-banda *reale* $z(t)$ di larghezza di banda (monolatera) W ed una frequenza f_0 , tipicamente scelta entro la banda di $z(t)$, è sempre possibile trovare un segnale (complesso) $v(t)$, *passa-basso* di banda W , tale che

$$\begin{aligned} z(t) &= \operatorname{Re}\{v(t) e^{j2\pi f_0 t}\} \\ &= \operatorname{Re}\{v(t)\} \cos 2\pi f_0 t - \operatorname{Im}\{v(t)\} \sin 2\pi f_0 t \end{aligned} \quad (8.49)$$

dove il segnale $v(t)$, che varia in funzione della frequenza f_0 , è detto *equivalente passa-basso* di $z(t)$ e vale $v(t) = z_+(t) e^{-j2\pi f_0 t}$.

8.3.1 Scomposizione di Rice del rumore gaussiano passa banda

Sia $n(t)$ un processo stazionario gaussiano di tipo passa banda caratterizzato da uno spettro di densità di potenza $G_n(f)$: per ogni frequenza f_c interna alla banda del processo si possono determinare due processi passa basso gaussiani, congiuntamente stazionari, $n_i(t)$ e $n_q(t)$ per cui:

$$n(t) = n_i(t) \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t)$$

Se il processo passa banda è un rumore gaussiano a media nulla, vi sono due casi particolari di estremo interesse.

Caso 1. Se $G_n(f)$ è piatto all'interno della banda B_t ($G_n(f) = \eta/2$ in banda) e f_c è la frequenza centrale di B_t , allora $n_i(t)$ e $n_q(t)$ sono processi gaussiani passa basso a media nulla i cui spettri di densità di potenza risulteranno piatti nella banda $[0, B_t/2]$ e di ampiezza pari a η , come mostrato in figura 8.1.

Pertanto sarà:

$$N_r = \overline{n^2} = \overline{n_i^2} = \overline{n_q^2}$$

inoltre $n_i(t)$ e $n_q(t)$, per qualunque istante di tempo t fissato, sono due variabili aleatorie gaussiane scorrelate (e quindi indipendenti), ovvero

$$\overline{n_i(t) n_q(t)} = \overline{n_i(t)} \overline{n_q(t)} = 0 \quad \forall t$$

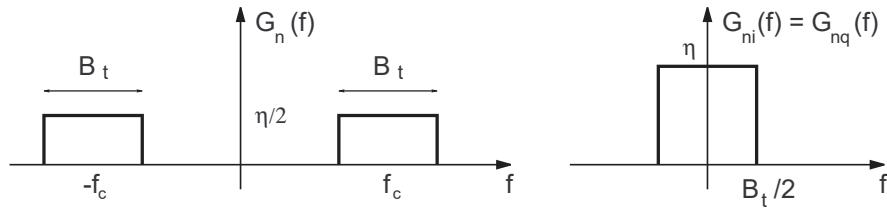


Figura 8.1: Scomposizione di Rice: caso 1.

Caso 2. Se $G_n(f)$ è piatto e pari a $\eta/2$ all'interno della banda B_t e f_c è posto ad uno degli estremi della banda B_t , allora $n_i(t)$ e $n_q(t)$ sono processi gaussiani passa basso a media nulla i cui spettri di densità di potenza risulteranno piatti nella banda $[0, B_t]$ e pari a $\eta/2$, come mostrato in figura 8.2.

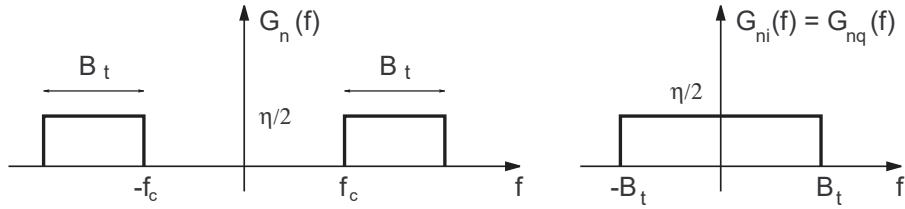


Figura 8.2: Scomposizione di Rice: caso 2.

Come già osservato per il caso 1 precedente, sarà:

$$N_r = \overline{n^2} = \overline{n_i^2} = \overline{n_q^2}$$

inoltre $n_i(t)$ e $n_q(t)$, per qualunque istante di tempo t fissato, sono due variabili aleatorie gaussiane scorrelate (e quindi indipendenti), ovvero

$$\overline{n_i(t) n_q(t)} = \overline{n_i(t)} \overline{n_q(t)} = 0 \quad \forall t$$

8.4 Modulazioni analogiche

Non sempre i segnali sono adatti ad essere inviati direttamente al canale trasmissivo, in quanto, ad esempio, le caratteristiche in frequenza del canale potrebbero essere incompatibili con lo spettro del segnale che si vuole trasmettere.

Definizione Chiamasi modulazione quel processo in cui un segnale periodico, detto **portante**, viene sistematicamente alterato sulla base di un segnale, detto **modulante**, che reca l'informazione, al fine di poter adattare le caratteristiche di quest'ultimo a quelle del canale di comunicazione. Il segnale uscita di tale processo è detto segnale **modulato**. Il processo di modulazione deve essere reversibile, ovvero deve esistere un secondo processo, detto demodulazione, che, in base alla conoscenza (con più o meno dettaglio) del segnale portante, sia in grado di ottenere dal segnale modulato il segnale modulante, recante l'informazione che la sorgente vuole inviare alla destinazione.

Le modulazioni si dividono in analogiche o numeriche, a seconda del tipo di segnale modulante. Vedremo solo le modulazioni analogiche, che comprendono le seguenti sottoclassi principali:

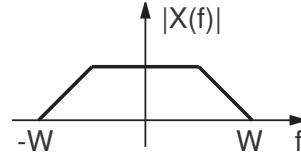
- Modulazioni lineari o di ampiezza
 - Modulazione di ampiezza (AM o AM standard)
 - Modulazione a doppia banda laterale (DSB)
 - Modulazione a singola banda laterale (SSB)
 - Modulazione a banda vestigiale (VSB)

²La parte relativa alla scomposizione di Rice e alle modulazioni è stata scritta in collaborazione col Prof. Sandro Zappatore, utilizzando in larga parte le sue dispense

- Modulazioni angolari o esponenziali
 - Modulazione di fase (PM)
 - Modulazione di frequenza (FM)

Segnale modulante.

Definiamo $x(t)$ il segnale analogico reale modulante. Per tutte le modulazioni che vedremo il segnale $x(t)$ deve rispettare le seguenti specifiche: deve essere normalizzato (ovvero l'ampiezza del suo modulo deve essere inferiore o uguale a uno, cioè $|x(t)| \leq 1$), di tipo passa-basso e con larghezza di banda W in Hz ($|X(f)| = 0$ per $|f| > W$). In figura un esempio di come potrebbe apparire il modulo della trasformata di Fourier di un tipico segnale modulante (nel seguito le figure nel dominio delle frequenze assumeranno l'andamento visualizzato).



Modulo della trasformata di Fourier del segnale modulante (esempio).

Canale e rumore. Per valutare l'effetto del rumore introdotto dal canale sulla qualità del segnale demodulato assumeremo che il canale sia ideale, cioè non introduca distorsioni (lineari o non) per le quali sarebbe necessario prevedere qualche forma di equalizzazione (ove possibile). In altri termini, si supporrà che il rumore sia l'unico responsabile del degrado della qualità del segnale demodulato. Il canale verrà quindi modellato nello stesso modo visto per la trasmissione in banda base (sezione 8.2.4), ovvero senza distorsione sulla banda del segnale, cosicché la sua uscita è pari all'ingresso ritardato e attenuato di un fattore L , e il rumore è assunto essere AWGN con densità spettrale bilatera $G_n(f) = \eta/2$.

Misura della qualità del segnale. Un modo oggettivo per esprimere la qualità del segnale demodulato consiste nel valutare il **rapporto segnale/disturbo** del segnale demodulato, rapporto comunemente indicato come $(S/N)_D$. $(S/N)_D$ è anche detto rapporto segnale/disturbo di post demodulazione o rapporto segnale/disturbo a destinazione. Se $x(t)$ e $x_D(t)$ sono, rispettivamente, il segnale modulante trasmesso e il segnale demodulato, allora $(S/N)_D$ è definito come segue:

$$\left(\frac{S}{N}\right)_D = \frac{\overline{x(t)^2}}{\overline{(x(t) - x_D(t))^2}} = \frac{S_x}{N_D}$$

$$N_D = \overline{(x(t) - x_D(t))^2}$$

Osservazione Prove sperimentali hanno evidenziato che l'intellegibilità di un messaggio in banda fonica richiede un rapporto segnale disturbo a destinazione di almeno 10 - 15 dB. Tale limite non è di tipo tecnologico, essendo dovuto al sistema uditivo umano.

8.5 Modulazioni Lineari

Iniziamo riassumendo le caratteristiche principali delle varie modulazioni lineari. Ciascuna di esse verrà poi analizzata con maggior dettaglio. Le formule che seguono sono quindi una sintesi di quanto verrà esposto nel seguito e mi aspetto che sappiate scriverle durante l'esame orale.

Le modulazioni lineari che verranno descritte sono:

DSB, modulazione a doppia banda laterale o *Double Side Band*,

AM, modulazione di ampiezza o *Amplitude Modulation*,

SSB, modulazione a banda laterale unica o *Single Side Band*, nelle due versioni a banda laterale unica superiore (Upper SSB o USSB) e inferiore (Lower SSB o LSSB).

La modulazione a banda laterale vestigiale **VSB** o *Vestigial Side Band* sarà descritta sommariamente.

Le equazioni che portano dal segnale modulante $x(t)$ a quello modulato $y(t)$ sono:

$$\begin{aligned} \text{DSB} \quad y(t) &= A_c x(t) \cos(\omega_c t + \theta) = A_c x(t) \cos(2\pi f_c t + \theta) \\ \text{AM} \quad y(t) &= A_c (1 + \mu x(t)) \cos(\omega_c t + \theta) \\ \text{USSB} \quad y(t) &= \frac{1}{2} A_c [x(t) \cos(\omega_c t) - \dot{x}(t) \sin(\omega_c t)] \\ \text{LSSB} \quad y(t) &= \frac{1}{2} A_c [x(t) \cos(\omega_c t) + \dot{x}(t) \sin(\omega_c t)] \end{aligned} \tag{8.50}$$

dove μ (indice di modulazione della AM) è compreso fra 0 e 1 e $\check{x}(t)$ è la trasformata di Hilbert di $x(t)$.

La larghezza di banda delle modulazioni lineari è $2W$ per AM e DSB e W per USSB e LSSB.

La potenza del segnale $x(t)$ vale, per definizione, $S_x = \overline{x^2(t)} \leq 1$. Le potenze medie in trasmissione $S_t = y^2(t)$ sono:

$$\begin{aligned} \text{DSB} \quad S_t &= \frac{1}{2} A_c^2 S_x \\ \text{AM} \quad S_t &= \frac{1}{2} A_c^2 [1 + \mu^2 S_x] \\ \text{SSB} \quad S_t &= \frac{1}{4} A_c^2 S_x \end{aligned} \quad (8.51)$$

Nota: nell'AM la potenza in trasmissione si divide fra una parte dedicata alla portante, pari a $P_c = \frac{1}{2} A_c^2$, ed una parte sulle bande laterali, pari a $2P_{bs} = \frac{1}{2} A_c^2 \mu^2 S_x$. Quindi $S_t = P_c + 2P_{bs}$ e $P_c \geq 2P_{bs}$. Quindi almeno il 50 % della potenza in trasmissione è concentrata nella portante e non dipende da $x(t)$.

Le trasformate di Fourier delle modulazioni DSB ed AM valgono:

$$\begin{aligned} \text{DSB} \quad Y(f) &= \frac{1}{2} A_c [X(f - f_c) + X(f + f_c)] \\ \text{AM} \quad Y(f) &= \frac{1}{2} A_c \{\mu[X(f - f_c) + X(f + f_c)] + [\delta(f - f_c) + \delta(f + f_c)]\} \end{aligned} \quad (8.52)$$

Le prestazioni delle modulazioni lineari in presenza di rumore, espresse mediante il rapporto segnale/rumore a destinazione in riferimento al valore γ per trasmissione in banda base, ovvero

$$\gamma = \left(\frac{S}{N} \right)_{D \text{ banda base}} = \frac{S_R}{\eta W}$$

sono:

$$\begin{aligned} \text{DSB} \quad \left(\frac{S}{N} \right)_D &= \gamma \\ \text{AM demod. sincrono} \quad \left(\frac{S}{N} \right)_D &= \frac{\mu^2 S_x}{1 + \mu^2 S_x} \gamma \\ \text{SSB} \quad \left(\frac{S}{N} \right)_D &= \gamma \end{aligned} \quad (8.53)$$

Vediamo ora in dettaglio ciascuna modulazione.

8.5.1 Modulazione DSB

Come visto il segnale modulato DSB si ottiene dal modulante mediante

$$y(t) = A_c x(t) \cos(\omega_c t + \theta) = A_c x(t) \cos(2\pi f_c t + \theta) \quad (8.54)$$

Nel processo di modulazione DSB, il segnale modulante $x(t)$ altera, istante per istante, l'ampiezza di una portante cosinusoidale. La figura 8.3 mostra un esempio di segnale modulante e del corrispondente segnale modulato: da notare che normalmente la frequenza della portante è molto maggiore della massima frequenza del segnale modulante, mentre in figura, per chiarezza di disegno, tale condizione non è del tutto rispettata.

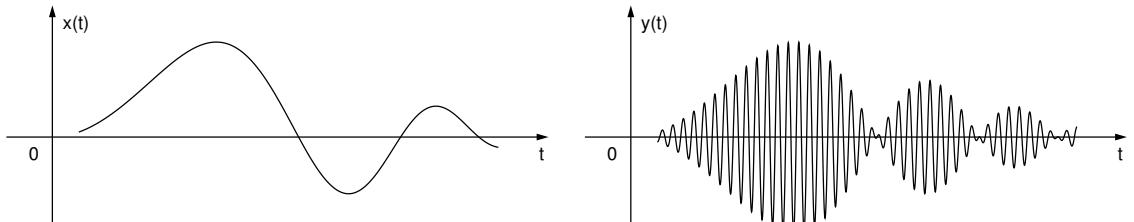


Figura 8.3: Esempio di segnale modulante $x(t)$ e del corrispondente $y(t)$ modulato DSB.

Da osservare come la portante subisca una inversione di fase pari a 180 gradi in tutti i punti per i quali il segnale modulante attraversa lo zero (vedi ingrandimento a lato). Il modulo del segnale modulante prende il nome di *inviluppo* del segnale modulato (vedi figura 8.4). Poiché $x(t)$ può assumere valori negativi, l'inviluppo del segnale modulato differisce dal segnale modulante. La ragione del nome è da attribuirsi al fatto che entro i suoi limiti è contenuto il segnale portante, come facilmente verificabile graficamente.

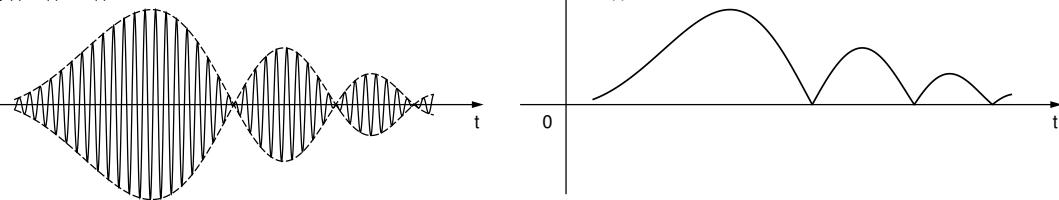
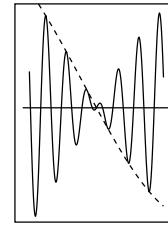


Figura 8.4: Segnale modulato DSB con sovrapposto inviluppo (a sinistra) e inviluppo.

Analizziamo ora l'effetto della modulazione DSB nel dominio di Fourier. Se nell'equazione 8.54 poniamo, senza perdita di generalità, $\theta = 0$, trasformando secondo Fourier e ricordando le proprietà della convoluzione e dell'impulso di Dirac otteniamo

$$Y(f) = 0.5A_c \cdot X(f) * [\delta(f - f_c) + \delta(f + f_c)] \quad (8.55)$$

poichè

$$X(f) * \delta(f - f_c) = \int_{-\infty}^{+\infty} X(f - \lambda)\delta(f - f_c)d\lambda = X(f - f_c) \quad (8.56)$$

si ha

$$Y(f) = 0.5A_c[X(f - f_c) + X(f + f_c)] \quad (8.57)$$

Graficamente la trasformata $Y(f)$ risulta essere quella mostrata in figura 8.5.

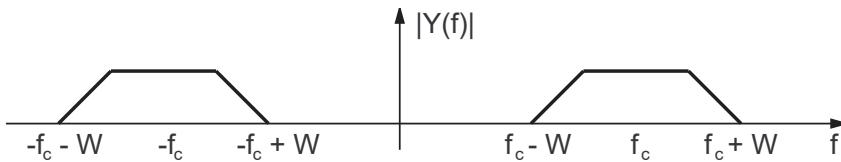


Figura 8.5: Modulo della trasformata di Fourier di un segnale modulato DSB.

Moltiplicare, nel dominio tempo, $x(t)$ per un coseno a frequenza f_c , corrisponde, nel dominio delle frequenze, a traslarne rigidamente lo spettro attorno alla frequenza f_c . È evidente che:

- un segnale modulato DSB è di tipo passa banda;
- la larghezza di banda del segnale modulato è doppia rispetto a quella del segnale modulante: $B_t = 2W$

Lo schema a blocchi del modulatore è illustrato in figura 8.6:

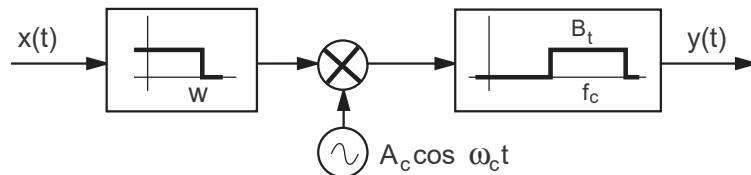


Figura 8.6: Schema a blocchi del modulatore DSB.

Il primo filtro rende il segnale all'ingresso del mixer a banda limitata; il filtro passa banda, posto dopo il mixer, evita che eventuali componenti di frequenza non essenziali vengano immesse sul canale.

Da un punto di vista puramente teorico i due filtri sono superflui, tuttavia diventano indispensabili a causa del fatto che i componenti elettrici usati per realizzare il modulatore presentano caratteristiche (dovute a non linearità) che possono far discostare il reale funzionamento del sistema da quello teorico. Calcoliamo ora la potenza media in trasmissione, questo può essere fatto in due modi:

1. Utilizzando quanto visto per i processi aleatori (metodo rigoroso).

Il segnale $y(t) = A_c x(t) \cos(\omega_c t + \theta)$ è il prodotto di due processi aleatori (vedi sezione 7.9.2).

Poichè i due p.a. $x(t)$ e $A_c \cos(\omega_c t + \theta)$ sono generati da apparati differenti è corretto assumere i) che θ sia una v.a. uniforme in $[0, 2\pi]$ (cosa che modella la scelta arbitraria dell'origine dei tempi quando due p.a. provengono da sorgenti indipendenti) e ii) che i due p.a. siano indipendenti. Siamo quindi nelle condizioni di applicabilità dell'equazione 7.100 e la potenza si ottiene mediante prodotto delle potenze, ovvero

$$S_t = \overline{y^2(t)} = \frac{1}{2} A_c^2 S_x$$

(la potenza di $A_c \cos(\omega_c t + \theta)$ è stata calcolata in sezione 7.5.1).

2. In modo empirico (ignorando i processi aleatori). Se ignoriamo il termine di fase θ

$$S_t = \overline{y^2(t)} = A_c^2 \overline{(x(t) \cos \omega_c t)^2} = 0.5 A_c^2 \overline{x^2(t)(1 + \cos 2\omega_c t)}$$

Se il calcolo della potenza è effettuato su un multiplo del periodo di portante, allora $\overline{\cos 2\omega_c t} = 0$.

Se inoltre la larghezza di banda di $x(t)$ è piccola rispetto ad f_c , $x(t) \cos 2\omega_c t \approx 0$.

Poichè $S_x = \overline{x^2(t)}$ è la potenza media del segnale modulante, la potenza media in trasmissione di una DSB vale:

$$S_t = \frac{1}{2} A_c^2 S_x$$

Demodulazione DSB. La demodulazione è il processo inverso della modulazione e consiste nel riportare in banda base l'informazione contenuta nel segnale modulato ricevuto. Lo schema a blocchi del demodulatore DSB è illustrato in figura 8.7.

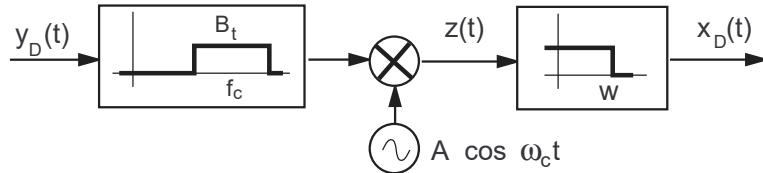


Figura 8.7: Schema a blocchi del demodulatore DSB.

Sia:

$$y_D(t) = A_r x(t) \cos(\omega_c t)$$

il segnale DSB ricevuto, avente larghezza di banda B_t . Il filtro passa banda posto all'ingresso del ricevitore lascia passare inalterato il segnale ricevuto: il suo compito è quello di impedire che possano giungere al mixer altri segnali eventualmente presenti sul canale e di limitare la potenza di rumore. Dopo il mixer si ha:

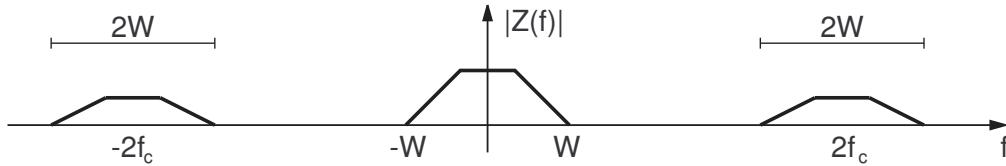
$$z(t) = A_r x(t) \cos(\omega_c t) \cdot A \cos(\omega_c t) = A_r A x(t) \cos^2(\omega_c t) = \frac{1}{2} A_r A x(t) [1 + \cos(2\omega_c t)]$$

Il segnale $z(t)$, la cui trasformata

$$Z(f) = \frac{1}{2} A_r A \left[X(f) + \frac{1}{2} X(f - 2f_c) + \frac{1}{2} X(f + 2f_c) \right]$$

è riportata in figura 8.8, è quindi applicato al filtro passa basso, avente larghezza di banda W , che provvede a eliminare tutte le componenti di $Z(f)$ che sono al di fuori della banda (-w, +w). Pertanto

$$x_D(t) = \frac{1}{2} A A_r x(t) \quad (8.58)$$

Figura 8.8: Modulo della trasformata di Fourier di $z(t)$.

Il segnale in uscita dal filtro passa basso è una copia esatta (a meno di un fattore di scala) del segnale modulante. Il demodulatore presentato prevede l'impiego di un oscillatore locale (quello presente nel ricevitore) perfettamente sincrono con quello usato nel trasmettitore: la sincronicità richiede che la frequenza e la fase dei due oscillatori siano perfettamente uguali. Per tale ragione il demodulatore DSB è detto **sincrono** o coerente. Per chiarire quanto sia importante il sincronismo, supponiamo che l'oscillatore locale sia solo isocrono (cioè abbia la stessa frequenza, ma non la stessa fase) con quello impiegato in trasmissione, ovvero

$$o_D(t) = A \cos(\omega_c t + \theta)$$

con θ angolo arbitrario tra 0 e 2π .

$$z(t) = A_r x(t) \cos(\omega_c t) A \cos(\omega_c t + \theta) = \frac{1}{2} A_r A x(t) [\cos(\theta) + \cos(2\omega_c t + \theta)]$$

Pertanto, dopo il filtro passa basso, si ha

$$x_D(t) = \frac{1}{2} A A_r x(t) \cos(\theta) \quad (8.59)$$

La potenza del segnale ricevuto risulta ridotta di un fattore $\cos^2 \theta$; ci sono perciò dei casi (ad esempio per $\theta = 0.5\pi$) per cui il segnale ricevuto è identicamente nullo. In altri termini, un ricevitore isocrono può causare una diminuzione della potenza del segnale demodulato. Problemi maggiori si hanno in caso di completa asincronia tra l'oscillatore locale e quello presente nel trasmettitore.

Prestazione della DSB in presenza di rumore. Se a destinazione si usa un demodulatore sincrono si ha che (vedi figura 8.9):

$$\left(\frac{S}{N}\right)_R = \frac{S_R}{\eta B_t} = \frac{S_R}{2\eta W} \quad (8.60)$$

All'ingresso del ricevitore si ha il segnale somma del rumore introdotto dal canale più $y_D(t) = A_r x(t) \cos(\omega_c t)$, ovvero il segnale trasmesso e attenuato dal canale (si trascurano fase e ritardo, che sono ininfluenti).

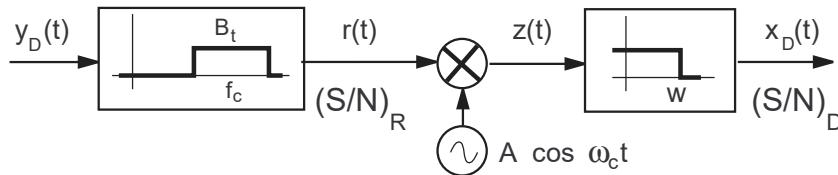


Figura 8.9: Schema del demodulatore sincrono.

La potenza ricevuta è quindi, per la parte del segnale

$$S_R = \frac{1}{2} A_r^2 S_x$$

La potenza del rumore non si può esprimere all'ingresso del ricevitore: essendo il rumore AWGN con densità spettrale bilatera $G_n(f) = \eta/2$ la sua potenza è infatti infinita. Si può tuttavia esprimere dopo il filtro passabanda iniziale (ideale, ovvero di ampiezza unitaria): esso lascia invariata la parte

di segnale, cioè $y_D(t)$, ma limita il rumore alla sola banda B_t . Chiamando $r(t)$ il segnale dopo il filtro passabanda si ha che

$$r(t) = y_D(t) + n(t) = A_r x(t) \cos(\omega_c t) + n(t)$$

La potenza di $n(t)$ è evidentemente $N_R = \eta B_t = 2\eta W$ e ne risulta il rapporto segnale rumore in ricezione in eq. 8.60. Per definizione il rapporto segnale/disturbo di prerivelazione ($S/N)_R$ è il rapporto tra la potenza del segnale ricevuto, misurato all'ingresso del ricevitore, e la potenza di rumore dovuta al canale **relativamente alla banda di trasmissione**. Nell'ipotesi di rumore additivo gaussiano bianco (AGWN), supposto il filtro di ingresso del ricevitore (filtro di "front-end") passa banda ideale, piatto nella banda di trasmissione B_t , allora $(S/N)_R$ è il rapporto tra la potenza di segnale e di rumore all'uscita del filtro di ingresso (ovvero quello che abbiamo calcolato).

Continuando l'analisi separatamente per segnale e rumore (ipotesi corretta visto che il rumore è additivo ed indipendente dal segnale e che il demodulatore è un sistema lineare) e ricordando che il segnale a destinazione è quello in eq. 8.58 ne deriva che la potenza di segnale a destinazione vale

$$S_D = \overline{x_D^2(t)} = \frac{1}{4} A^2 A_r^2 S_x = \frac{1}{2} A^2 S_R \quad (8.61)$$

Occupiamoci ora del rumore. Scomponendo il rumore secondo Rice (caso 1)

$$\begin{aligned} r(t) &= A_r x(t) \cos(2\pi f_c t) + n_i(t) \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t) = \\ &= [A_r x(t) + n_i(t)] \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t) \end{aligned}$$

Dopo il mixer si ha

$$\begin{aligned} z(t) &= A [A_r x(t) + n_i(t)] \cos^2(2\pi f_c t) - A n_q(t) \sin(2\pi f_c t) \cos(2\pi f_c t) = \\ &= \frac{1}{2} A [A_r x(t) + n_i(t)] (1 - \cos(4\pi f_c t)) - \frac{1}{2} A n_q(t) \sin(4\pi f_c t) \end{aligned}$$

Il filtro passa basso posto dopo il mixer provvede a eliminare le componenti centrate alla frequenza doppia della portante. Perciò in uscita si avrà:

$$\frac{1}{2} A [A_r x(t) + n_i(t)]$$

La potenza di segnale S_D è quindi quella vista, mentre la potenza di rumore a destinazione vale

$$N_D = \frac{1}{4} A^2 \overline{n_i^2} = \frac{1}{4} A^2 \eta B_t = \frac{1}{2} A^2 \eta W$$

Il rapporto segnale/rumore a destinazione vale allora

$$\left(\frac{S}{N}\right)_D = \frac{0.5 A^2 S_R}{0.5 A^2 \eta W} = \frac{S_R}{\eta W} = 2 \left(\frac{S}{N}\right)_R = \gamma$$

Le prestazioni di una DSB, misurate in termini di rapporto segnale/disturbo a destinazione, sono le stesse ottenibili con una trasmissione equivalente in banda base. In altri termini, il processo di modulazione e il successivo processo di demodulazione non peggiorano (né migliorano) la qualità del segnale informativo a destinazione rispetto al caso di una trasmissione "diretta" (in banda base) equivalente (cioè utilizzando la stessa potenza media in trasmissione e a parità di rumore) del segnale modulante.

8.5.2 Modulazione AM

Un segnale $y(t)$ si dice modulato in ampiezza (o semplicemente modulato AM) da un segnale modulante $x(t)$ (normalizzato, passabasso di banda W) se

$$y(t) = A_c (1 + \mu x(t)) \cos(\omega_c t + \theta) \quad (8.62)$$

dove la costante μ è detta **indice di modulazione** ed è compresa tra 0 e 1. L'AM è molto simile alla DSB: vi sono due modi di pensare l'AM derivata da una DSB

- l'AM è una DSB con segnale modulante uguale a $(1 + \mu x(t))$; esiste cioè nel segnale modulante una componente continua finita. Lo schema a blocchi corrispondente è riportato in figura 8.10.

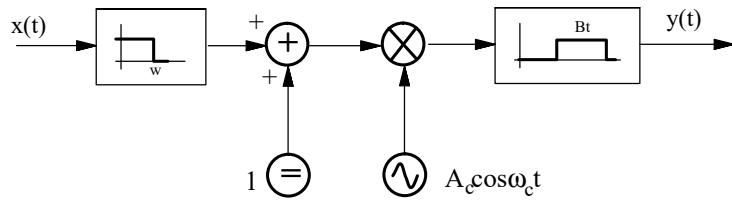


Figura 8.10: Schema a blocchi del modulatore AM.

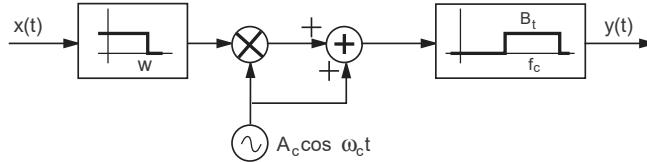


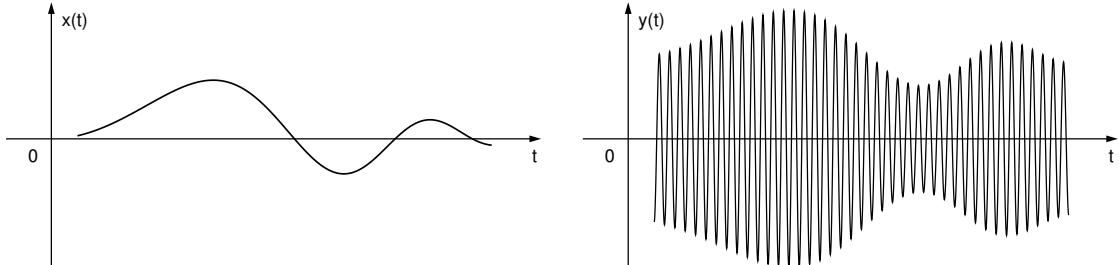
Figura 8.11: Schema a blocchi alternativo del modulatore AM.

2. l'AM è una DSB, con modulante $\mu x(t)$ a cui viene reinserita la portante che altrimenti risulterebbe soppressa. In altri termini l'AM può essere ottenuta da una DSB aggiungendovi il segnale portante. Per tale ragione spesso la DSB viene detta AM a portante soppressa. Lo schema corrispondente è in questo caso quello di figura 8.11.

I due schemi a blocchi sono del tutto equivalenti, anche se corrispondono a due possibili scelte operative differenti con cui realizzare un modulatore AM. Se il segnale modulante $x(t)$ è quello in figura 8.12 a sinistra, allora il corrispondente segnale modulato AM è visibile in figura 8.12 a destra. Non vi sono, a differenza di quanto accadeva per una DSB, cambiamenti di fase della portante dovuti al segnale modulante: è infatti facile provare che:

$$1 + \mu x(t) \geq 0 \quad \forall t \quad (8.63)$$

essendo, per ipotesi, $0 < \mu < 1$ e $|x(t)| \leq 1$.

Figura 8.12: Esempio di segnale modulante $x(t)$ e del corrispondente $y(t)$ modulato AM.

L'inviluppo di un segnale AM è rappresentato in figura 8.13. L'inviluppo di un segnale AM, a meno di una costante additiva e moltiplicativa, riproduce il segnale modulante. Estrarre l'inviluppo di una AM equivale perciò a demodulare il segnale (purchè ovviamente valga la condizione 8.63).

Ponendo, senza perdita di generalità, $\theta = 0$ nell'eq. 8.62 e seguendo il procedimento seguito per la DSB, nel dominio delle frequenze si ha:

$$Y(f) = \frac{1}{2} A_c \mu X(f) * [\delta(f - f_c) + \delta(f + f_c)] + \frac{1}{2} A_c [\delta(f - f_c) + \delta(f + f_c)]$$

quindi

$$Y(f) = \frac{1}{2} A_c \{ \mu [X(f - f_c) + X(f + f_c)] + [\delta(f - f_c) + \delta(f + f_c)] \}$$

Lo spettro dell'AM è uguale a quello della DSB salvo per la presenza di due impulsi di Dirac in corrispondenza della frequenza di portante e della attenuazione di $X(f)$ del fattore μ (si veda figura 8.14). La larghezza di banda dell'AM è quindi pari al doppio della banda del segnale modulante: $B_t = 2W$.

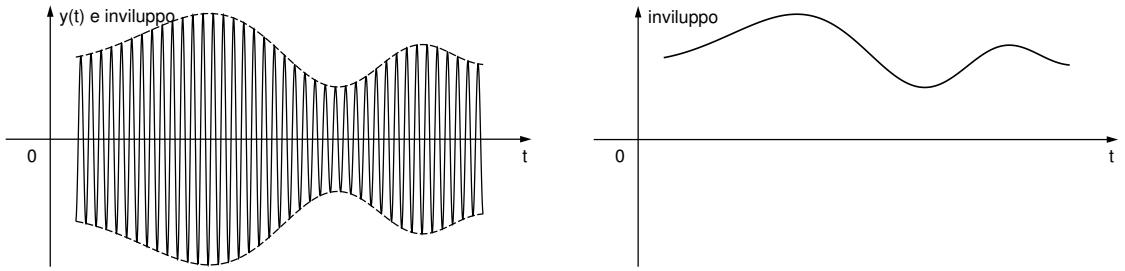


Figura 8.13: Segnale modulato AM con sovrapposto inviluppo (a sinistra) e inviluppo.

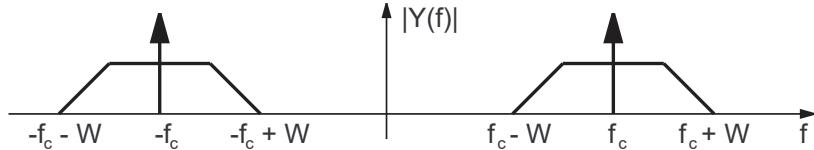


Figura 8.14: Modulo della trasformata di Fourier di un segnale modulato AM.

Calcoliamo ora la potenza media in trasmissione: il calcolo, analogo a quello per la DSB, porta a

$$S_t = \overline{y^2(t)} = A_c^2 \overline{[(1 + \mu x(t)) \cos \omega_c t]^2} = \frac{1}{2} A_c^2 [1 + \mu^2 S_x]$$

L'espressione della potenza dell'AM evidenzia come una parte dell'energia sia dovuta alla presenza della portante mentre l'altra sia da imputarsi alle due bande laterali e, quindi, all'informazione contenuta nel segnale modulante. Allora

$$S_t = P_c + 2P_{bs}$$

S_t è la potenza dell'AM, $P_c = \frac{1}{2} A_c^2$ è la potenza dovuta alla portante e $P_{bs} = \frac{1}{4} A_c^2 \mu^2 S_x$ è la potenza di ciascuna banda laterale. Poiché $\mu^2 S_x \leq 1$ vale che $2P_{bs} \leq P_c$: quindi

$$P_c \geq \frac{1}{2} S_t \quad \text{e} \quad P_{bs} \leq \frac{1}{4} S_t$$

Pertanto almeno il 50% della potenza trasmessa è concentrata nella portante e al massimo il 25% risiede in ciascuna banda laterale.

Demodulazione AM. Poiché l'AM può essere pensata a tutti gli effetti come una DSB in cui il segnale modulante presenta una componente continua (tale da impedire che, nel complesso, il modulante possa assumere valori negativi), allora un ricevitore DSB demodulerà anche un segnale AM. In figura 8.15 è presentato il demodulatore AM sincrono: come evidente è identico al demodulatore DSB (figura 8.7) a parte per l'aggiunta del condensatore finale che elimina la componente continua.

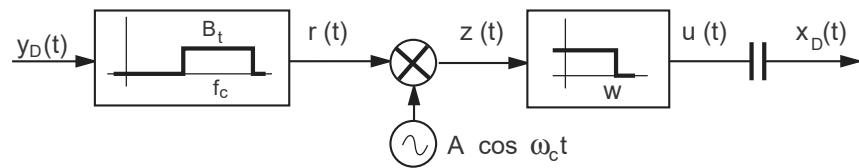


Figura 8.15: Schema a blocchi del demodulatore sincrono per l'AM.

L'AM presenta un grosso vantaggio rispetto alla DSB: l'inviluppo di un segnale AM è una copia (a meno di un fattore di scala e di traslazione) del segnale modulante. Ciò permette una realizzazione alternativa del ricevitore, detto **ricevitore ad inviluppo**. Lo schema relativo è mostrato in figura 8.16.

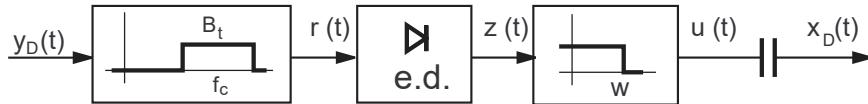
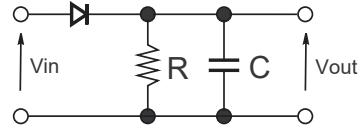
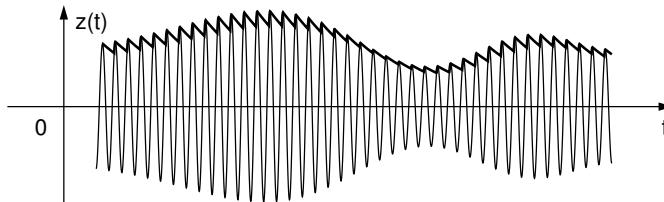


Figura 8.16: Schema a blocchi del demodulatore a inviluppo.

Il principio di funzionamento è il seguente: il segnale $r(t)$ è la copia attenuata del segnale modulato $y(t)$ più il rumore filtrato. Trascurando il rumore si può estrarre l'inviluppo mediante un semplice blocco *envelope detector* (vedi figura).



Supponendo il diodo ideale, la tensione in uscita segue quella di ingresso se è maggiore di quella presente sul condensatore. Se è minore vince la tensione di scarica del condensatore. Se, come usuale, la banda del segnale è molto più bassa di quella della portante, ovvero se $W \ll f_c$, il segnale estratto è molto simile all'inviluppo (vedi figura 8.17).

Figura 8.17: Uscita del blocco e.d. (*envelope detector*) in assenza di rumore.

Vantaggi della ricezione ad inviluppo:

- a) non richiede nessun oscillatore locale, pertanto non vi sono problemi di sincronizzazione;
- b) economico: il ricevitore è estremamente semplice e i componenti impiegati sono di tipo comune;
- c) si può dimostrare che, almeno per tutte le condizioni di funzionamento significative, le prestazioni di un ricevitore ad inviluppo sono le stesse del ricevitore sincrono.

Svantaggi:

- la ricezione ad inviluppo è attuabile solo per l'AM. L'AM richiede, a parità di potenza del segnale "informativo", una maggior potenza complessiva del segnale modulato, comprendendo quest'ultimo la componente di portante che tuttavia non reca alcuna informazione.

L'AM è scarsamente economica per collegamenti punto-punto, ma ha trovato larghissimo impiego per trasmissioni "broadcast", poiché la ricezione ad inviluppo permette la realizzazione di ricevitori economici e quindi alla portata di molti utenti.

Prestazione della AM in presenza di rumore con demodulatore sincrono. Lo schema del ricevitore sincrono è in fig. 8.15. Il calcolo delle prestazioni dell'AM demodulata coerentemente segue il procedimento seguito per lo studio della DSB. Il rapporto segnale/rumore in ricezione vale, come visto

$$\left(\frac{S}{N} \right)_R = \frac{S_R}{\eta B_t} = \frac{S_R}{2\eta W}$$

dove $S_R = (1/2)A_r^2(1 + \mu^2 S_x)$. Dopo il filtro di ingresso si ha

$$r(t) = y_D(t) + n(t) = A_r(1 + \mu x(t)) \cos \omega_c t + n(t)$$

Utilizzando la scomposizione di Rice per il rumore come visto (caso 1), si ha:

$$\begin{aligned} r(t) &= A_r(1 + \mu x(t)) \cos(2\pi f_c t) + n_i(t) \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t) = \\ &= [A_r(1 + \mu x(t)) + n_i(t)] \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t) \end{aligned}$$

Dopo il mixer si ha

$$\begin{aligned} z(t) &= A[A_r(1 + \mu x(t)) + n_i(t)] \cos^2(2\pi f_c t) - A n_q(t) \sin(2\pi f_c t) \cos(2\pi f_c t) = \\ &= \frac{1}{2} A[A_r(1 + \mu x(t)) + n_i(t)] (1 + \cos(4\pi f_c t)) - \frac{1}{2} A n_q(t) \sin(4\pi f_c t) \end{aligned}$$

Il filtro passa basso posto dopo il mixer provvede a eliminare le componenti centrate alla frequenza doppia della portante. Quindi

$$u(t) = \frac{1}{2} A[A_r(1 + \mu x(t)) + n_i(t)]$$

Il condensatore elimina infine la componente continua:

$$x_D(t) = \frac{1}{2} A[A_r \mu x(t) + n_i(t)]$$

La potenza media di segnale a destinazione è

$$S_D = \frac{1}{4} A^2 A_r^2 \mu^2 S_x = \frac{1}{4} A^2 \frac{\mu^2 S_x}{1 + \mu^2 S_x} (1 + \mu^2 S_x) A_r^2 = \frac{1}{4} A^2 \frac{\mu^2 S_x}{1 + \mu^2 S_x} 2 S_R$$

La potenza media di rumore a destinazione (uguale a quella della DSB) è

$$N_D = \frac{1}{4} A^2 \overline{n_i^2} = \frac{1}{4} A^2 \eta B_t = \frac{1}{2} A^2 \eta W$$

Il rapporto segnale/rumore a destinazione vale allora

$$\left(\frac{S}{N}\right)_D = \frac{0.5 A^2 \frac{\mu^2 S_x}{1 + \mu^2 S_x} S_R}{0.5 A^2 \eta W} = \frac{\mu^2 S_x}{1 + \mu^2 S_x} \frac{S_R}{\eta W} = \frac{\mu^2 S_x}{1 + \mu^2 S_x} \gamma$$

o anche

$$\left(\frac{S}{N}\right)_D = 2 \frac{\mu^2 S_x}{1 + \mu^2 S_x} \left(\frac{S}{N}\right)_R$$

Poichè μ e S_x sono entrambi minori o uguali a uno, vale che

$$\frac{\mu^2 S_x}{1 + \mu^2 S_x} \leq \frac{1}{2}$$

Le prestazioni di una AM valutate impiegando un demodulatore coerente, misurate in termini di rapporto segnale/disturbo a destinazione, sono quindi sempre inferiori a quelle di una trasmissione equivalente in banda base. Ciò è da imputarsi al fatto che almeno il 50% della potenza trasmessa complessiva è impiegata per trasmettere la portante, che non porta alcun contributo informativo (non facendo parte dello spettro del segnale modulante).

Prestazione della AM in presenza di rumore con demodulatore a inviluppo. Lo schema del ricevitore ad inviluppo è in fig. 8.16. Come già visto per il demodulatore AM coerente si ha:

$$\left(\frac{S}{N}\right)_R = \frac{S_R}{\eta B_t} = \frac{S_R}{2\eta W}$$

dove $S_R = (1/2)A_r^2(1 + \mu^2 S_x)$. Dopo il filtro di ingresso si ha

$$r(t) = y_D(t) + n(t) = A_r(1 + \mu x(t)) \cos \omega_c t + n(t)$$

Utilizziamo la scomposizione di Rice per il rumore come visto (caso 1):

$$\begin{aligned} r(t) &= A_r(1 + \mu x(t)) \cos(2\pi f_c t) + n_i(t) \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t) = \\ &= [A_r(1 + \mu x(t)) + n_i(t)] \cos(2\pi f_c t) - n_q(t) \sin(2\pi f_c t) \end{aligned}$$

Nell'ipotesi di impiegare un estrattore di inviluppo ideale (blocco e.d., envelope detector), si ottiene (vedi eq. 8.41)

$$A(t) = +\sqrt{[A_r(1 + \mu x(t)) + n_i(t)]^2 + n_q^2(t)}$$

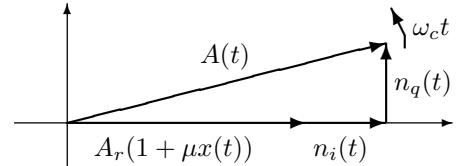
Dall'espressione di $A(t)$ emerge chiaramente che, in generale, il rumore non è più additivo, compiendo quale fattore moltiplicativo del segnale nello sviluppo del quadrato della somma. Tuttavia, se $A_r^2 \gg \sigma_n^2$ (σ_n^2 è la varianza del rumore, ovvero, essendo a media nulla, la sua potenza N), cioè se si è nell'ipotesi di "grandi rapporti segnale/disturbo" di prorivelazione, allora il primo termine sotto radice predomina decisamente sul secondo e

$$A(t) \approx [A_r(1 + \mu x(t)) + n_i(t)]$$

Che tale approssimazione sia ragionevole quando $A_r^2 \gg \sigma_n^2$ emerge chiaramente tracciando il diagramma fasoriale (su piano rotante a velocità ω_c) su cui siano riportati il segnale e le componenti del rumore (il diagramma è relativo a $r(t)$, l'inviluppo è il modulo del vettore indicato con $A(t)$):

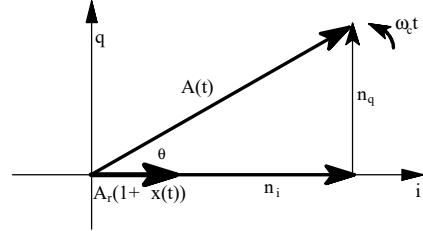
In condizioni di grandi rapporti segnale/disturbo, n_q è molto piccolo e quindi $A(t)$ può essere approssimato dalla sua componente in fase

$$A(t) \approx A_r(1 + \mu x(t)) + n_i(t)$$



che è, a meno di un fattore di scala, ciò che si aveva in uscita dal ricevitore AM sincrono. Pertanto, per i casi in cui $A_r^2 \gg \sigma^2$, il rumore si può ragionevolmente ritenere ancora additivo. **Nell'ipotesi di grandi rapporti segnali/disturbo le prestazioni di una AM demodulata ad inviluppo coincidono con quelle ottenibili con un ricevitore coerente.**

Viceversa se $A_r^2 \ll \sigma^2$, il diagramma fasoriale è quello in figura. Poiché $A_r(1 + \mu x(t))$ è per la maggior parte del tempo molto più piccolo di $n_i(t)$ e/o $n_q(t)$, allora $A(t) \approx n(t)$; di fatto il ricevitore demodula solo rumore.



Affinchè si possa ritenere $A_r^2 \gg \sigma^2$, se $A_n(t) = +\sqrt{n_i^2(t) + n_q^2(t)}$ è l'inviluppo del rumore, deve essere $A_n(t) \ll A_r$ per la maggior parte del tempo, ad esempio per il 99% del tempo. La condizione di cui sopra diviene $Pr(A_n > A_r) = 0.01$: essendo $A_n(t)$ distribuito secondo Rayleigh

$$\int_{A_r}^{\infty} \frac{A_n}{\sigma^2} e^{-\frac{A_n^2}{2\sigma^2}} dA_n = e^{-\frac{A_r^2}{2\sigma^2}} \Big|_{A_r}^{\infty} = e^{-\frac{A_r^2}{2\sigma^2}} = 0.01$$

e quindi

$$-\frac{A_r^2}{2\sigma^2} = \ln 0.01 \quad \frac{A_r^2}{\sigma^2} \approx 9.2$$

perciò la condizione dei grandi rapporti S/N, per $S_x = 1$, è soddisfatta se

$$\left(\frac{S}{N}\right)_R \geq 10dB$$

8.5.3 Modulazione SSB

Consideriamo la trasformata di Fourier di un generico segnale reale modulato DSB, rappresentata in figura 8.18.

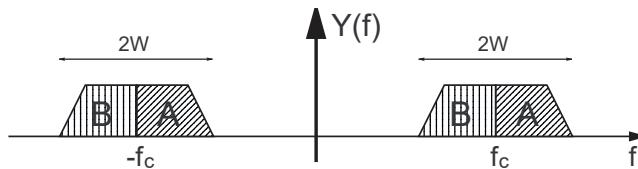


Figura 8.18: Trasformata di Fourier di un segnale modulato DSB.

È evidente come le due bande laterali "A" e "B" siano presenti due volte, una sulle frequenze positive ed una in quelle negative. Non vi è quindi perdita di informazione eliminando tali duplicati.

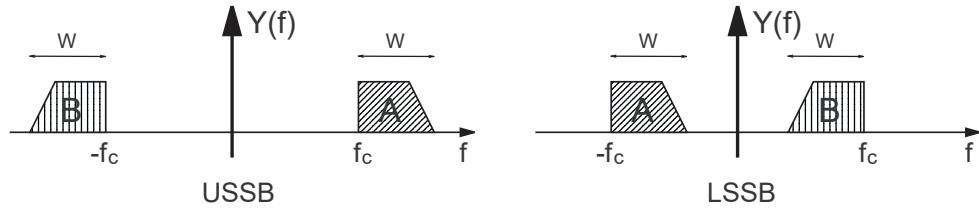


Figura 8.19: Modulo della trasformata di Fourier delle modulazioni USSB e LSSB.

Un segnale SSB può quindi essere pensato come ottenuto da una DSB a cui è stata cancellata una delle due bande laterali. A seconda di quale banda sia stata rimossa si parla di USSB (Upper SSB) o di LSSB (Lower SSB). Gli spettri sono raffigurati in figura 8.19.

La SSB può quindi essere ottenuta da una DSB mediante l'applicazione di filtro passa-alto o passa-basso ideale centrato sulla frequenza di portante f_c . La larghezza di banda di un segnale modulato SSB è $B_t = W$.

Poichè è presente una sola banda laterale, la potenza in trasmissione di una SSB è pari alla metà di quella della DSB, quindi

$$S_t = \frac{1}{4} S_x A_c^2$$

Sebbene sia molto semplice rappresentare la SSB nel dominio delle frequenze, non lo è altrettanto nel dominio del tempo. Si può dimostrare che l'espressione nel tempo di un segnale modulato SSB è

$$y(t) = \frac{1}{2} A_c [x(t) \cos(\omega_c t) \pm \check{x}(t) \sin(\omega_c t)]$$

ove $\check{x}(t)$ è la trasformata di Hilbert (vedi sezione 7.7.2) del segnale modulante. Se nella formula precedente si usa il segno "+" si ottiene una LSSB, viceversa con il segno "-" il processo di modulazione fornisce una USSB.

Vi sono due principali schemi di modulatori SSB: il modulatore con filtro a forza bruta (figura 8.20) e il modulatore in quadratura (figura 8.21): entrambe le figure sono riferite ad una USSB.

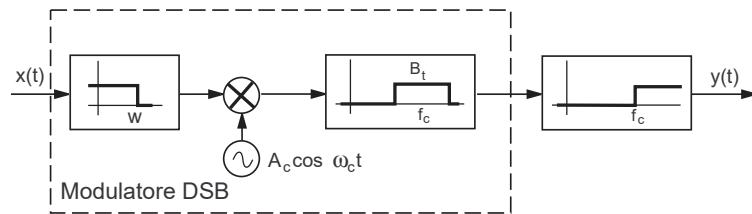


Figura 8.20: Modulatore con filtro a forza bruta.

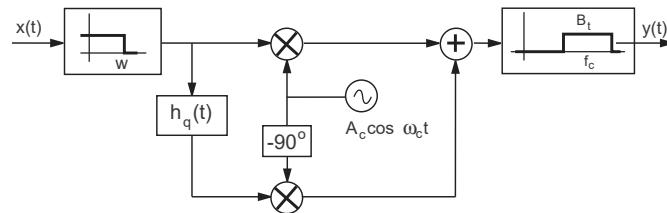


Figura 8.21: Modulatori in quadratura (h_q è il filtro di Hilbert).

Demodulazione SSB Il demodulatore sincrono (vedi figura 8.9) è in grado di demodulare la SSB, infatti all'uscita del mixer si avrà

$$\begin{aligned}
 z(t) &= \frac{1}{2} A_r [x(t) \cos \omega_c t \pm \check{x}(t) \sin \omega_c t] A \cos \omega_c t = \\
 &= \frac{1}{4} A A_r [x(t) (1 + \cos 2\omega_c t) + \check{x}(t) \sin \omega_c t \cos \omega_c t] = \\
 &= \frac{1}{4} A A_r [x(t) (1 + \cos 2\omega_c t) + \check{x}(t) \sin 2\omega_c t]
 \end{aligned}$$

Il segnale $z(t)$ è quindi applicato al filtro passa basso, di larghezza di banda W , che provvede ad eliminare tutte le componenti centrate attorno al doppio della frequenza portante. Perciò

$$x_D(t) = \frac{1}{4} A A_r x(t)$$

SSB in pratica Si è visto come la SSB sia ottenibile o impiegando filtri (a forza bruta) passa basso o passa alto, posti in cascata ad un modulatore DSB, oppure mediante due modulatori DSB in quadratura, in cui il ramo in quadratura è modulato dalla trasformata di Hilbert del segnale modulante. È da notare che:

1. i filtri passa basso e/o alto ideali non sono realizzabili. I filtri reali non possono presentare fronti a 90 gradi: la "salita" del fronte deve essere effettuata in un intervallo di frequenza finito. Una regola empirica dice che, detto β tale intervallo, allora $200\beta \geq f_c$.
2. il filtro di Hilbert è realizzabile solo per un intervallo di frequenza limitato. Inoltre, esistono segnali (ad esempio un impulso rettangolare) che se trasformati richiederebbero energia impulsiva infinita. Tali segnali non sono perciò trasformabili.

I vincoli di cui sopra **impediscono** di realizzare una "vera" SSB. La SSB risulta pertanto realizzabile solo per una limitata classe di segnali modulanti, quelli che non presentano energia in un intervallo di frequenze nell'intorno dell'origine. A questa classe appartengono, ad esempio, i segnali vocali (vedi figura 8.22).

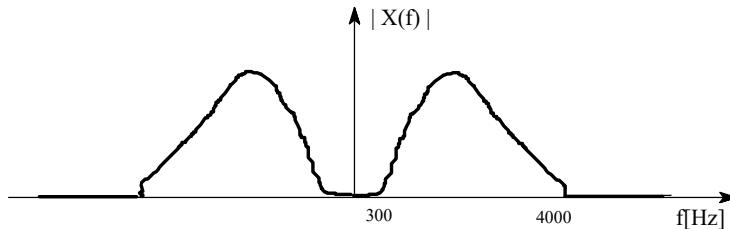


Figura 8.22: Spettro di un tipico segnale vocale.

Tra le due bande laterali del segnale modulato DSB è presente un "vuoto": è appunto questo "vuoto" che permette di realizzare una SSB usando filtri passa alto (o basso) reali, come illustrato in figura 8.23.

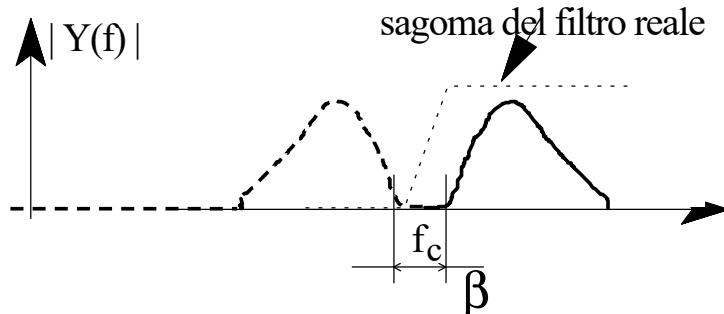


Figura 8.23: Filtraggio con filtro reale di una DSB modulata da un segnale vocale.

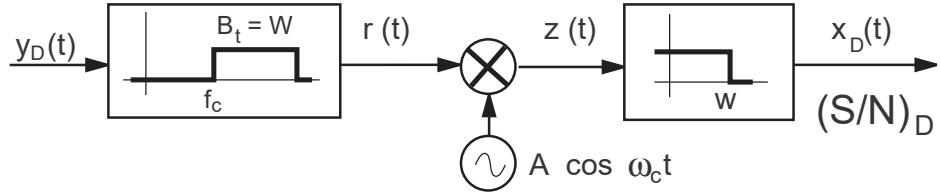


Figura 8.24: Schema del ricevitore sincrono per una USSB.

Prestazioni della SSB. Lo schema del ricevitore sincrono per una USSB è in figura 8.24. Il ricevitore per LSSB è ovviamente del tutto simile salvo che il filtro di front-end sarà un passa banda ideale con frequenze di taglio pari a $f_c - B_t$ e f_c . Inoltre tale ricevitore sincrono è anche in grado di demodulare una VSB purchè con banda vestigiale piccola rispetto a W . Nel caso di una USSB si ha:

$$\left(\frac{S}{N}\right)_R = \frac{S_R}{\eta B_t} = \frac{S_R}{\eta W}$$

dove $S_R = 0.25A_r^2 S_x$ e $r(t) = 0.5A_r(x(t) \cos \omega_c t - \dot{x}(t) \sin \omega_c t) + n(t)$.

Utilizzando la scomposizione di Rice per il rumore (caso 2)

$$\begin{aligned} r(t)) &= \frac{1}{2} A_r x(t) \cos 2\pi f_c t + n_i(t) \cos 2\pi f_c t - \frac{1}{2} A_r \dot{x}(t) \sin 2\pi f_c t - n_q(t) \sin 2\pi f_c t = \\ &= [\frac{1}{2} A_r x(t) + n_i(t)] \cos 2\pi f_c t - [\frac{1}{2} A_r \dot{x}(t) + n_q(t)] \sin 2\pi f_c t \end{aligned}$$

Dopo il mixer si ha

$$\begin{aligned} z(t)) &= A[\frac{1}{2} A_r x(t) + n_i(t)] \cos^2 2\pi f_c t - [\frac{1}{2} A_r \dot{x}(t) + n_q(t)] \sin 2\pi f_c t \cos 2\pi f_c t = \\ &= \frac{1}{2} A[\frac{1}{2} A_r x(t) + n_i(t)] (1 - \cos 4\pi f_c t) - \frac{1}{2} A[\frac{1}{2} A_r \dot{x}(t) + n_q(t)] \sin 4\pi f_c t \end{aligned}$$

Il filtro passa basso posto dopo il mixer provvede ad eliminare le componenti centrate alla frequenza doppia della portante. Perciò in uscita si avrà:

$$x_D(t) = \frac{1}{2} A[\frac{1}{2} A_r x(t) + n_i(t)]$$

La potenza di segnale a destinazione è quindi

$$S_D = \frac{1}{4} A^2 \frac{1}{4} A_r^2 S_x = \frac{1}{4} A^2 S_R$$

e la potenza di rumore a destinazione, ricordando che $B_t = W$, è

$$N_D = \frac{1}{4} A^2 \overline{n_i^2} = \frac{1}{4} A^2 \eta B_t = \frac{1}{4} A^2 \eta W$$

Quindi

$$\left(\frac{S}{N}\right)_D = \frac{0.25A^2 S_R}{0.25A^2 \eta W} = \frac{S_R}{\eta W} = \gamma = \left(\frac{S}{N}\right)_R$$

Pertanto, come per la DSB, anche le prestazioni della SSB, misurate in termini di rapporto segnale/disturbo a destinazione, sono le stesse ottenibili con una trasmissione equivalente in banda base.

8.5.4 Modulazione VSB

La SSB è di difficile realizzabilità a causa della presenza del filtro passa alto o passa basso posto in cascata al modulatore DSB. Filtri reali richiedono che lo spettro del segnale modulante sia caratterizzato da un "vuoto" in corrispondenza delle basse frequenze. Allora, se l' SSB è in qualche modo utilizzabile per la trasmissione di segnali fonici, poco si presta per la trasmissione di segnali televisivi, che, a differenza dei primi, presentano componenti in frequenza non trascurabili anche in prossimità della frequenza zero.

La VSB risulta essere una soluzione di compromesso, presentando caratteristiche intermedie tra la DSB e la SSB. La VSB è ottenuta applicando un *filtro con simmetria vestigiale* in cascata ad un trasmettitore DSB. Tale filtro fa sì che una banda laterale della DSB passi quasi tutta a meno di una piccola porzione (vestigia), porzione che viene inclusa a partire dall'altra banda laterale che risulta, pertanto, quasi completamente rimossa. Il filtro vestigiale $H_v(f)$ deve essere tale che

$$H_v(f - f_c) + H_v(f + f_c) = C, |f| \leq B$$

essendo C una costante e B la banda del segnale modulante.

La risposta in frequenza di un filtro vestigiale è riportata in figura 8.25.

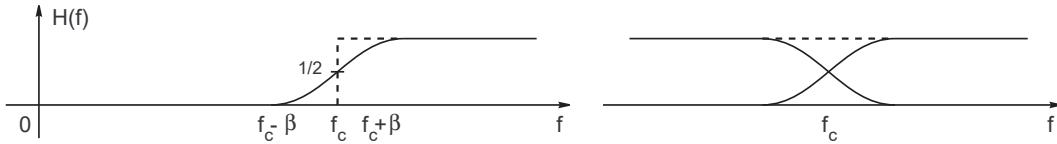


Figura 8.25: Filtro a simmetria vestigiale

La forma esatta del filtro non è così importante come la sua simmetria (che è quella che permette la ricostruzione di $x(t)$ senza distorsioni).

Si può dimostrare che l'espressione nel tempo è

$$y(t) = \frac{1}{2} A_c [x(t) \cos(\omega_c t) \pm x_q(t) \sin(\omega_c t)]$$

essendo $x_q(t)$ un segnale ottenuto dalla convoluzione del segnale modulante con una funzione che dipende dal filtro vestigiale usato. Poiché la VSB è una soluzione di compromesso tra DSB e SSB allora risulta:

per quanto riguarda la banda di trasmissione

$$B_t = W + \beta \Rightarrow B_{tSSB} = W < B_t < 2W = B_{tDSB}$$

per quanto riguarda la potenza

$$S_{tSSB} = \frac{1}{4} A_c^2 S_x < S_t < \frac{1}{2} A_c^2 S_x = S_{tDSB}$$

Un ricevitore sincrono è in grado di compiere la demodulazione.

Infine, esiste una versione di VSB con portante (VSB-C) che può essere ottenuta applicando il filtro vestigiale ad un segnale AM. Se $\beta \geq 0.3 W$, la VSB-C può essere demodulata, a fronte di un degrado di qualità accettabile, ad inviluppo con un notevoli vantaggi in termini di complessità e di costo del ricevitore. Le trasmissioni televisive analogiche sono effettuate con una VSB-C.

8.6 Modulazioni Angolari

In questo tipo di modulazioni il segnale modulante altera il contenuto di fase della portante, mentre l'inviluppo del segnale modulato è costante. Sono anche note come modulazioni esponenziali.

Sia $x(t)$ il segnale reale modulante normalizzato ($|x(t)| < 1$) di tipo passa-basso con larghezza di banda W in Hz. Le modulazioni angolari che trattiamo sono: PM (modulazione di fase) e FM (modulazione di frequenza). Per entrambe il segnale modulato è

$$y(t) = A_c \cos(\omega_c t + \theta(t)) \quad (8.64)$$

dove $\theta(t)$ è funzione del segnale modulante $x(t)$.

La potenza in trasmissione è quella del p.a. sinusoidale casuale, e vale

$$S_T = \frac{1}{2} A_c^2$$

Infatti per il calcolo della potenza media è corretto considerare l'angolo di fase $\theta(t)$ come una v.a. uniforme in $[0, 2\pi]$.

Si definisce *angolo totale di fase* il segnale $\phi(t) = \omega_c t + \theta(t)$.

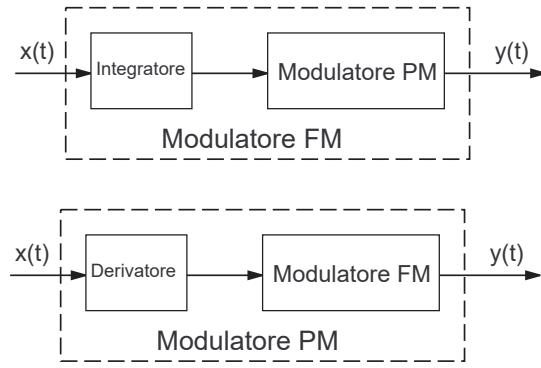


Figura 8.26: Modulatore FM ottenuto da un modulatore PM e modulatore PM ottenuto da un modulatore FM.

Si definisce *frequenza istantanea* $f_i(t)$ la derivata rispetto al tempo dell'angolo totale di fase:

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi}{dt} = \frac{\omega_c}{2\pi} + \frac{1}{2\pi} \frac{d\theta}{dt} = f_c + \frac{1}{2\pi} \frac{d\theta}{dt}$$

Le equazioni che legano $x(t)$ a $\theta(t)$ sono, a seconda della modulazione:

$$\begin{array}{ll} \text{PM} & \theta(t) = \theta_\Delta x(t) \\ \text{FM} & \theta(t) = 2\pi f_\Delta \int_{t_0}^t x(\tau) d\tau + \theta(t_0) \end{array} \quad t \geq t_0 \quad (8.65)$$

Il fatto che θ_Δ possa variare solo fra i limiti indicati deriva ovviamente dal fatto che $x(t)$ è normalizzato e che $\theta(t)$ è un angolo: se varia per più di 2π il suo valore non può essere determinato senza ambiguità.

Se t_0 è scelto in modo tale che $\theta(t_0) = 0$ si può usare la notazione

$$\theta(t) = 2\pi f_\Delta \int^t x(\tau) d\tau$$

Nella PM la fase della portante è direttamente proporzionale al segnale modulante, mentre nella FM la frequenza istantanea è direttamente proporzionale al segnale modulante.

Le costanti θ_Δ e f_Δ sono dette rispettivamente **deviazione** (massima) **di fase e di frequenza**. Dato un segnale modulato in frequenza la fase è

$$\theta(t) = 2\pi f_\Delta \int^t x(\tau) d\tau$$

Allora la FM può essere pensata come una PM modulata dall'integrale del segnale modulante. Viceversa, dato un segnale modulato in fase, la frequenza istantanea è

$$f_i(t) = f_c + \frac{1}{2\pi} \theta_\Delta \frac{dx(t)}{dt}$$

Allora la PM può essere pensata come una FM modulata dalla derivata del segnale modulante. Vi è quindi una strettissima parentela tra PM e FM: un modulatore FM può essere ottenuto da uno PM previa aggiunta di un integratore per il segnale modulante; un modulatore PM può essere realizzato impiegando un trasmettitore FM modulato dalla derivata del segnale modulante (vedi figura 8.26).

L'evoluzione nel tempo di un segnale modulato PM e FM è esemplificata in Figura 8.27, ove PM e FM vengono anche confrontate con un segnale modulato AM.

8.6.1 PM a banda stretta - PMNB

Si definisce modulazione PM a banda stretta (PM Narrow Band) una modulazione di fase caratterizzata da $\theta_\Delta \ll 1$ radiano.

Sotto tale ipotesi è semplice calcolare la trasformata del segnale modulato a partire dalla definizione:

$$y(t) = A_c \cos(\omega_c t + \theta(t)) = A_c [\cos(\omega_c t) \cos(\theta(t)) - \sin(\omega_c t) \sin(\theta(t))]$$

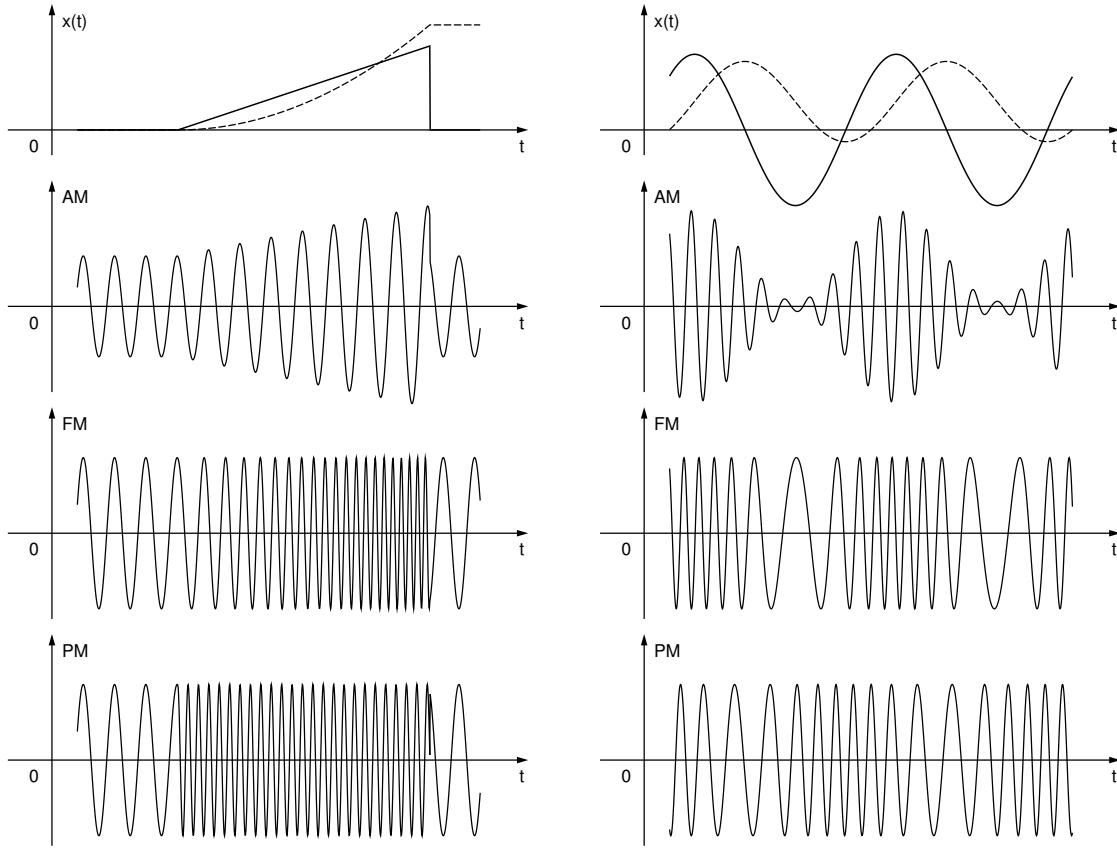


Figura 8.27: Esempio di andamenti nel tempo di segnali AM, FM e PM. Tratteggiato l'integrale di $x(t)$ (non in scala). Tutte le modulazioni hanno la stessa frequenza della portante.

poichè $\theta_\Delta \ll 1$ si ha che $\theta(t) = \theta_\Delta x(t) \ll 1$ e valgono le approssimazioni

$$\cos \theta(t) \approx 1 \quad \sin \theta(t) \approx \theta(t)$$

per cui

$$y(t) \approx A_c [\cos(\omega_c t) - \theta(t) \sin(\omega_c t)] = A_c [\cos(\omega_c t) - \theta_\Delta x(t) \sin(\omega_c t)]$$

la cui trasformata di Fourier vale

$$Y(f) = \frac{1}{2} A_c \{ [\delta(f - f_c) + \delta(f + f_c)] + j \theta_\Delta [-X(f - f_c) + X(f + f_c)] \} \quad (8.66)$$

La larghezza di banda è allora $2W$. Da notare la somiglianza con lo spettro della modulazione AM, anche se la portante è qui in quadratura rispetto a quella alterata dal modulante e la potenza associata alle bande laterali è molto piccola rispetto a quella della portante.

Per i motivi anzidetti, un ricevitore sincrono DSB (con oscillatore locale $\sin(\omega_c t)$) è in grado di demodulare (con qualità più che buona) una NBPM.

Per le modulazioni angolari a banda larga è non banale calcolare la trasformata di Fourier del segnale modulato, salvo quando il modulante è un tono sinusoidale (modulazione tonale).

8.6.2 Spettro della PM/FM tonale*

Consideriamo una FM con segnale modulante tonale dato da

$$x(t) = A_m \cos(2\pi f_m t)$$

la fase risulta

$$\theta(t) = \frac{A_m f_\Delta}{f_m} \sin 2\pi f_m t$$

Consideriamo adesso una PM con segnale modulante dato da

$$x(t) = A_m \sin(2\pi f_m t)$$

La fase è

$$\theta(t) = \theta_\Delta A_m \sin 2\pi f_m t$$

Dal confronto dalle due fasi è evidente come sia possibile condurre, ai fini della determinazione dello spettro di una modulazione angolare tonale, un'analisi congiunta. Infatti, definendo β come:

$$\begin{aligned} \beta &= \theta_\Delta A_m && \text{per la PM} \\ \beta &= \frac{f_\Delta A_m}{f_m} && \text{per la FM} \end{aligned} \quad (8.67)$$

l'angolo di fase si può scrivere per entrambe le modulazioni come $\theta(t) = \beta \sin(2\pi f_m t)$: β è detto **indice di modulazione** della PM e della FM.

Il segnale modulato vale quindi, per entrambe le modulazioni,

$$y(t) = A_c \cos(\omega_c t + \theta(t)) = A_c \cos(\omega_c t) \cos(\beta \sin 2\pi f_m t) - A_c \sin(\omega_c t) \sin(\beta \sin 2\pi f_m t)$$

Sebbene $y(t)$ non sia necessariamente periodica, i termini

$$\cos(\beta \sin 2\pi f_m t) \quad \text{e} \quad \sin(\beta \sin 2\pi f_m t)$$

sono periodici di periodo $1/f_m$ e pertanto possono essere espansi utilizzando le serie di Fourier con frequenza fondamentale pari a f_m . Esistono quindi dei coefficienti tali che:

$$\begin{aligned} \cos(\beta \sin 2\pi f_m t) &= J_0(\beta) + \sum_{n \text{ pari}}^{+\infty} 2J_n(\beta) \cos 2\pi n f_m t \\ \sin(\beta \sin 2\pi f_m t) &= \sum_{n \text{ dispari}}^{+\infty} 2J_n(\beta) \sin 2\pi n f_m t \end{aligned}$$

$J_n(\beta)$ sono i coefficienti della serie di Fourier

$$J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{j(\beta \sin \lambda - n\lambda)} d\lambda$$

Le funzioni $J_n(\beta)$ sono dette *funzioni di Bessel del primo tipo di ordine n*: le funzioni di Bessel godono, tra le altre, di queste due proprietà

$$\begin{aligned} J_{-n}(\beta) &= (-1)^n J_n(\beta) \\ \sum_{n=-\infty}^{+\infty} [J_n(\beta)]^2 &= 1 \quad (\text{Teorema di Parseval}) \end{aligned}$$

Il segnale PM/FM tonale modulato si può quindi scrivere come

$$\begin{aligned} y(t) &= A_c J_0(\beta) \cos(2\pi f_c t) + \\ &+ A_c \sum_{n \text{ dispari}}^{+\infty} J_n(\beta) [\cos 2\pi(f_c + n f_m)t - \cos 2\pi(f_c - n f_m)t] + \\ &+ A_c \sum_{n \text{ pari}}^{+\infty} J_n(\beta) [\cos 2\pi(f_c + n f_m)t + \cos 2\pi(f_c - n f_m)t] \end{aligned}$$

Usando le proprietà delle funzioni di Bessel si ha che:

$$\begin{aligned} y(t) &= A_c \left[\sum_{n \text{ dispari}}^{+\infty} J_n(\beta) \cos 2\pi(f_c + n f_m)t + \sum_{n \text{ pari}}^{+\infty} J_n(\beta) \cos 2\pi(f_c + n f_m)t \right] = \\ &= A_c \sum_{n=-\infty}^{+\infty} J_n(\beta) \cos 2\pi(f_c + n f_m)t \end{aligned}$$

Pertanto, lo *spettro di una FM e/o PM tonale consiste di una riga alla frequenza di portante e di una serie infinita di righe laterali*, distanziate l'una dall'altra della frequenza del tono modulante. In figura 8.28 sono riportati gli andamenti delle funzioni di Bessel relativamente ad $n=0, 1, 5$ e 10 . E' da notare come:

- le oscillazioni tendano a smorzarsi al crescere di β
- $|J_n(\beta)|$ sia monotonicamente decrescente al crescere di n/β
- $|J_n(\beta)| \ll 1$ se $|n/\beta| \gg 1$

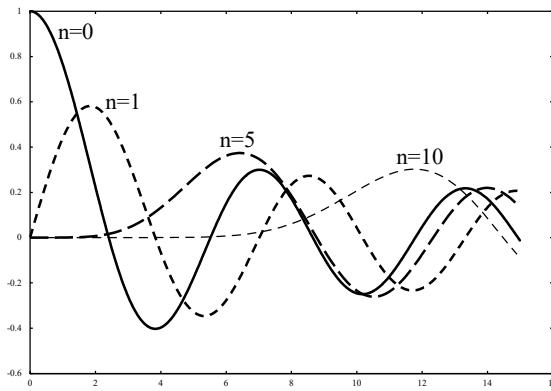


Figura 8.28: Andamento delle funzioni di Bessel per diversi valori di n .

8.6.3 Osservazioni sullo spettro della PM/FM tonale

Tenendo conto delle proprietà delle funzioni di Bessel, per FM e PM con modulante tonale valgono le seguenti osservazioni:

- a) le righe dello spettro sono infinite, quindi, da un punto di vista teorico, la larghezza di banda è infinita;
- b) la potenza associata alle righe caratterizzate da elevati valori di n è trascurabile rispetto a quella associabile alle righe relative a $n = 0, 1, 2, \dots$;
- c) elevati indici di modulazione comportano larghezze di banda crescenti;
- d) l'ampiezza della portante è funzione dell' indice di modulazione, pertanto anche la riga di portante porta "infomazione".

Per FM e PM si rende necessario riformulare la definizione di larghezza di banda. Sarà perciò data una definizione di larghezza di banda operativa, cioè si definirà la larghezza di banda come quell'intervallo di frequenza che contiene una certa percentuale (valori tipici sono 95%, 99%, 99,9%) della potenza totale del segnale modulato. Tale definizione è ovviamente caratterizzata da un certo grado di arbitrarietà legata al massimo grado di distorsione accettabile. Esempi di come si distribuiscono le righe spettrali al variare di β per una PM o FM tonale sono riportati in figura 8.29 (a f_m fissa).

8.6.4 Larghezza di banda per FM e PM

A parte la modulazioni angolari a banda stretta, per le quali $B_t \approx 2W$, in genere $\beta > 1$. Poichè

$$|J_n(\beta)| \ll 1 \quad \text{se} \quad \left| \frac{n}{\beta} \right| \gg 1$$

una possibile approssimazione della banda potrebbe consistere nel considerare significative solo quelle righe per cui

$$\left| \frac{n}{\beta} \right| \leq 1 \quad \text{cioè} \quad |n| \leq \beta$$

Usando tale approssimazione la larghezza di banda di una modulazione tonale (con tono a frequenza f_m) è

$$B_t \approx 2\beta f_m = \begin{cases} 2\phi_\Delta A_m f_m & \text{per una PM} \\ 2f_\Delta A_m & \text{per una FM} \end{cases}$$

Tale modo di procedere fornisce tuttavia solo un limite inferiore della larghezza di banda, nel senso che trascurare le componenti in frequenza al di fuori della banda così trovata comporta, in genere, distorsioni non trascurabili per quel che riguarda il segnale demodulabile. Un modo più corretto di procedere è quello di fissare una certa soglia ε (0.1 - 0.01 corrispondenti al 99% e al 99.99% della potenza totale) e quindi stabilire il numero M di righe dello spettro per cui risulta

$$|J_M(\beta)| > \varepsilon \quad \text{e} \quad |J_{M+1}(\beta)| < \varepsilon$$

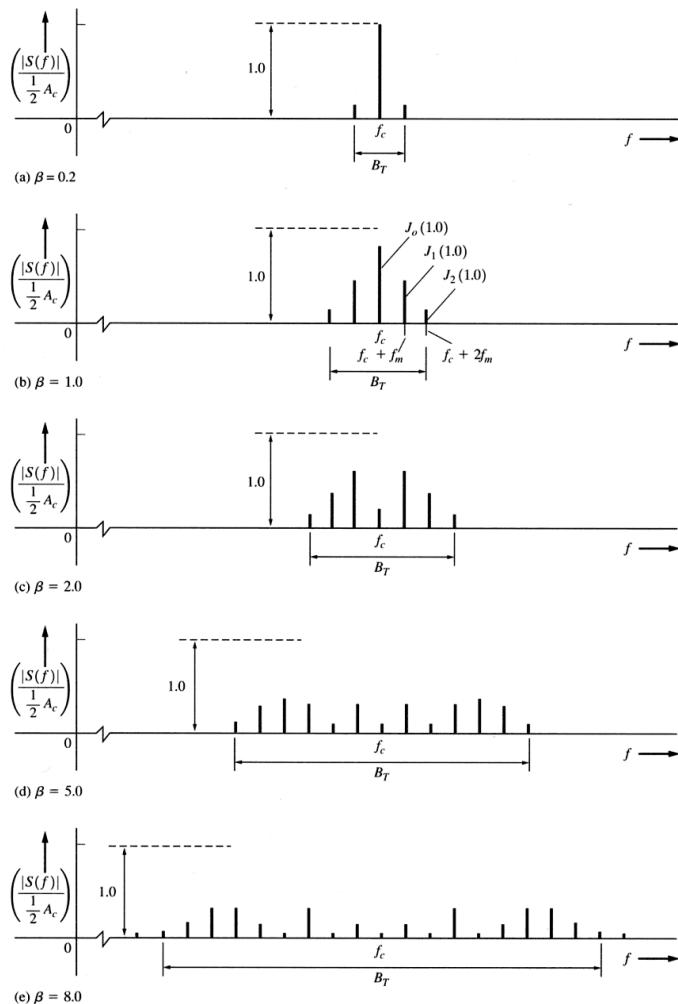


Figura 8.29: Distribuzione delle righe spettrali al variare dell'indice di modulazione β per una PM e FM con modulante tonale.

Ovviamente M è funzione di β e di ε . Osservando le curve ottenibili per $\varepsilon = 0.1$ e $\varepsilon = 0.01$ si può notare come la retta $M(\beta) \approx \beta + 2$ sia in "mezzo" alle due curve. Allora

$$B_t \approx 2M(\beta)f_m \approx 2(\beta + 2)f_m$$

Per FM con *modulante pieno* ($A_m = 1$) allora

$$B_t \approx 2(f_\Delta + 2f_m)$$

In alternativa è possibile utilizzare una formula analoga che, in generale, sottostima la banda di trasmissione. Tale formula è comunemente detta *regola del Carson*

$$B_t \approx 2(f_\Delta + f_m)$$

Le stime della larghezza di banda viste sono riferite ad un segnale modulante tonale. Nel caso di segnale modulante generico con larghezza di banda W le cose si complicano, in quanto per le *modulazioni angolari non vale il principio di sovrapposizione degli effetti*.

Definendo D , detto **indice di deviazione**, come

$$D = \frac{f_\Delta}{W}$$

per una **FM** si considera comunque accettabile una delle due seguenti approssimazioni

$B_t \approx 2(D + 2)W = 2(f_\Delta + 2W) \quad 2 \leq D \leq 10 \quad (8.68)$
--

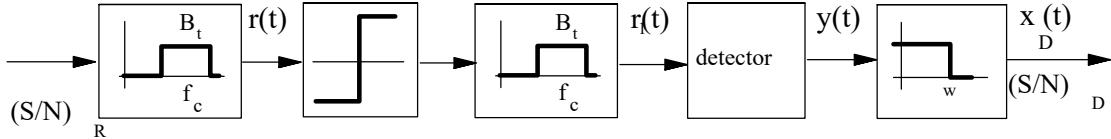
$B_t \approx 2(D + 1)W = 2(f_\Delta + W) \quad D > 10 \quad (8.69)$

Per una modulazione **PM** la banda può essere stimata come

$$B_t \approx 2(\theta_\Delta + 1)W \quad (8.70)$$

8.6.5 Prestazioni dei sistemi di modulazione angolare

Lo schema di massima del demodulatore angolare è:



Il rapporto segnale/rumore in ricezione vale

$$\left(\frac{S}{N}\right)_R = \frac{S_R}{\eta B_t}$$

dove B_t è la larghezza di banda della FM o della PM.

Dopo il filtro di ricezione, se $n(t)$ è il rumore filtrato, si ha il segnale

$$r(t) = A_c \cos(\omega_c t + \theta(t)) + n(t)$$

ovvero, esprimendo il rumore in termini di inviluppo e fase,

$$r(t) = A_c \cos(\omega_c t + \theta(t)) + A_n(t) \cos(\omega_c t + \phi_n(t)) = A_v(t) \cos(\omega_c t + \theta_v(t))$$

ove $A_v(t)$ e $\theta_v(t)$ sono, rispettivamente, l'inviluppo e la fase di $r(t)$.

Per determinare $A_v(t)$ e $\theta_v(t)$ conviene fare ricorso a un diagramma fasoriale dei segnali su un piano rotante a velocità ω_c (vedi Fig. 8.30).

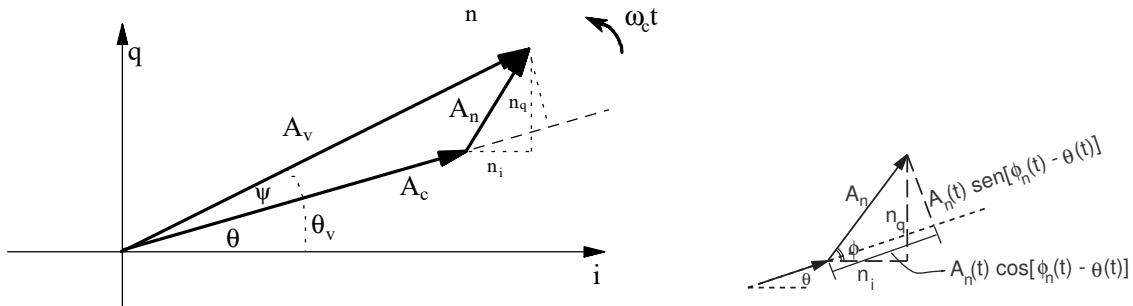


Figura 8.30: Diagramma fasoriale dei segnali su un piano rotante.

$$\theta_v(t) = \theta(t) + \psi(t)$$

$$\psi(t) = \arctan \frac{A_n(t) \sin[\phi_n(t) - \theta(t)]}{A_c + A_n(t) \cos[\phi_n(t) - \theta(t)]}$$

$r(t)$ attraversa il blocco di "hard-limiter" e il secondo filtro passa banda. Si può dimostrare che se f_c è sufficientemente più grande di B_t (come di fatto è in tutti i casi di interesse pratico), allora all'uscita del secondo BPF è presente un segnale modulato angolarmente, con l'angolo di fase corrotto dal rumore, ma ad inviluppo costante. Perciò

$$r_i(t) = A \cos(\omega_c t + \theta_v(t))$$

Come si può notare dall'espressione di $\theta_v(t)$, il rumore non è, in generale, additivo, essendo $\psi(t)$ combinazione non lineare di segnale e rumore. Nell'ipotesi di grandi rapporti segnale/disturbo, ovvero se $A_c^2 \gg \sigma^2$

$$\psi(t) \approx \arctan \left(\frac{A_n(t) \sin[\phi_n - \theta(t)]}{A_c} \right) \approx \frac{A_n(t) \sin[\phi_n - \theta(t)]}{A_c}$$

L'approssimazione è motivata dal fatto che, per la maggior parte del tempo $A_c \gg A_n(t)$. Inoltre, poichè $\phi_n(t)$ è una variabile aleatoria uniformemente distribuita in $[0, 2\pi]$, almeno per quanto riguarda le statistiche, $[\phi_n(t) - \theta(t)]$ differisce da $\phi_n(t)$ per il valor medio. Per quanto concerne lo studio degli effetti del rumore, allora la differenza $[\phi_n(t) - \theta(t)]$ può essere sostituita da $\phi_n(t)$. Quindi

$$\psi(t) \approx \frac{A_n(t) \sin(\phi_n(t))}{A_c} = \frac{n_q(t)}{A_c} = \frac{n_q(t)}{\sqrt{2 S_R}}$$

Dato che $\psi(t)$ dipende adesso solo dalla componente in quadratura del rumore, il rumore in $\theta_v(t)$ è di nuovo additivo; in altri termini, *sotto l'ipotesi di grandi rapporti segnale/disturbo di prorivelazione, il rumore di fase è additivo*. Inoltre, come evidente, $\psi(t)$ diminuisce al crescere di S_R .

Va anche osservato che l'espressione semplificata ottenuta per $\psi(t)$ è proprio il rumore di fase, nell'ipotesi di grandi rapporti segnale/disturbo, quando il segnale modulante è nullo e al sistema ricevente giunge solo portante non modulata e rumore.

Poichè $n_q(t)$ è piatto in $[-B_t/2, B_t/2]$ e con densità spettrale pari a η , nella stessa banda anche $\psi(t)$ risulterà piatto e con densità spettrale pari a $\eta/(2 S_R)$. Le modulazioni PM ed FM hanno rumore con banda $B_t \gg W$, quindi il filtro che taglia a W è assolutamente indispensabile.

Nel caso di una PM, il detector è un dispositivo in grado di estrarre la fase del segnale ricevuto.

Nel caso di una FM, il detector provvederà ad estrarre le variazioni di frequenza subite dal segnale ricevuto

8.6.6 Prestazioni della PM

All'uscita del detector si ha

$$y(t) = \theta_v(t) = \theta(t) + \psi(t) = \theta_\Delta x(t) + \psi(t)$$

Il filtro passa basso non altera il segnale, ma limita la banda del rumore, pertanto

$$x_D(t) = \theta_\Delta x(t) + \xi(t)$$

con $\xi(t)$ rumore ottenuto dal filtraggio di $\psi(t)$. Lo spettro di $\xi(t)$ è piatto nella banda $[-W, W]$ e presenta la stessa densità spettrale di $\psi(t)$, pari a $\eta/(2 S_R)$.

Allora la potenza media del segnale demodulato e del rumore a destinazione saranno

$$S_D = \theta_\Delta^2 S_x \quad N_D = \frac{\eta W}{S_R}$$

Pertanto

$$\left(\frac{S}{N} \right)_D = \frac{\theta_\Delta^2 S_x}{\frac{\eta W}{S_R}} = \theta_\Delta^2 S_x \frac{S_R}{\eta W} = \theta_\Delta^2 S_x \gamma$$

Le prestazioni ottenibili con una PM sono quindi migliori di quelle di una trasmissione equivalente in banda base per un fattore pari a $\theta_\Delta^2 S_x$.

Osservazioni:

- a parità di qualità del segnale demodulato, una PM consente di utilizzare meno potenza rispetto a una trasmissione in banda base e, quindi, anche rispetto a una DSB, SSB e AM;
- i risultati ottenuti sono validi entro i limiti per cui valgono le approssimazioni compiute, cioè se $A_c^2 \gg \sigma^2$. Con motivazioni analoghe a quelle usate per l'AM demodulata ad inviluppo, un limite inferiore per $(S/N)_R$, affinchè possa valere l'ipotesi di grandi rapporti segnale/disturbo, è individuabile in 10 dB. 10 dB rappresenta la soglia al di sotto della quale il rumore non può più ritenersi additivo e il detector demodula, praticamente, solo rumore;
- tale risparmio in termini di potenza di trasmissione è pagato al prezzo di una maggior larghezza di banda di trasmissione: si parla perciò di **"scambio banda-potenza"**;
- dovendo essere $\theta_D < \pi$ al fine di evitare ambiguità in sede di demodulazione, il massimo guadagno che una PM permette di conseguire nei confronti di una trasmissione equivalente in banda base è pari a circa 10 dB.

8.6.7 Prestazioni della FM

All'uscita del detector si ha

$$y(t) = \frac{1}{2\pi} \frac{d\theta_v(t)}{dt} = f_\Delta x(t) + \frac{1}{2\pi} \frac{d\psi(t)}{dt} = f_\Delta x(t) + \frac{1}{2\pi} \frac{dn_q(t)/dt}{\sqrt{2S_R}} = f_\Delta x(t) + \xi(t)$$

Ricordando che se la trasformata di Fourier di $s(t)$ vale $S(f)$ la trasformata della derivata vale

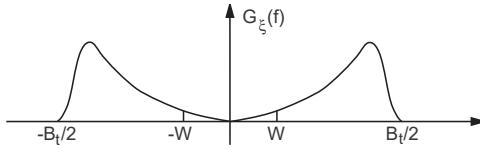
$$\frac{ds(t)}{dt} \leftrightarrow j2\pi f S(f)$$

lo spettro di densità di potenza del rumore $\xi(t)$ è,

$$G_\xi(f) = (2\pi f)^2 \left(\frac{1}{2\pi} \frac{1}{\sqrt{2S_R}} \right)^2 \eta = \frac{\eta f^2}{2S_R} \quad |f| \leq \frac{B_t}{2}$$

e zero altrove.

Nella banda $[-B_t/2, B_t/2]$ lo spettro di densità di potenza del rumore ha un andamento **parabolico**, quindi le alte frequenze del segnale saranno maggiormente corrotte dal rumore rispetto alle basse. Il filtro passa basso posto dopo il detector non altera il segnale e limita il rumore consegnato in uscita nella banda $[-W, W]$, come mostrato in figura.



Dopo il filtro la potenza media del segnale demodulato è

$$S_D = f_\Delta^2 S_x$$

e la potenza media di rumore

$$N_D = 2 \int_0^W G_\xi(f) df = 2 \int_0^W \frac{\eta f^2}{2S_R} df = \frac{1}{3} \frac{\eta W^3}{S_R}$$

Pertanto

$$\left(\frac{S}{N} \right)_D = \frac{f_\Delta^2 S_x}{\frac{1}{3} \frac{\eta W^3}{S_R}} = 3 \frac{f_\Delta^2}{W^2} S_x \frac{S_R}{\eta W} = 3 D^2 S_x \gamma$$

Osservazioni:

- a parità di qualità del segnale demodulato, una FM consente di utilizzare meno potenza rispetto a una trasmissione in banda base e, quindi, anche rispetto a una DSB, SSB e AM;
- i risultati ottenuti sono validi nei limiti per cui valgono le approssimazioni compiute, cioè quando $A_c^2 \gg \sigma^2$. Come per la PM, esiste una soglia minima per $(S/N)_R$, pari a circa 10 dB;
- tale risparmio in termini di potenza di trasmissione è pagato al prezzo di una maggior larghezza di banda di trasmissione: anche qui si parla di **"scambio banda-potenza"**
- a differenza di quanto accade per la PM, D non è superiormente limitata (almeno su un piano puramente teorico), pertanto non vi sono limiti (a patto di essere sopra soglia) ai miglioramenti sulla qualità del segnale a destinazione che la FM permette di conseguire nei confronti di una trasmissione equivalente in banda base (o con modulazione SSB, DSB, AM).

NOTA: al crescere di D la banda B_t aumenta, quindi si abbassa $(S/N)_R$ (a parità di potenza in trasmissione), fino ad arrivare al livello di soglia. Questo è il limite al miglioramento ottenibile.

8.6.8 Preenfasi e deenfasi nella FM *

Le prestazioni di un sistema di trasmissione FM possono essere migliorate *enfatizzando* al trasmettitore le componenti ad alta frequenza del segnale modulante e riducendole (*de-enfatizzandole*) al ricevitore. Il miglioramento si giustifica facilmente se si tiene conto che, come visto, la densità di potenza del rumore all'uscita del rivelatore FM è parabolica: la deenfasi in ricezione ripristina il corretto valore delle componenti ad alta frequenza del segnale modulante e contemporaneamente riduce le componenti di rumore che alle stesse frequenze sono particolarmente rilevanti. Il filtro di de-enfasi viene posto nel ricevitore dopo il detector e prima del filtro passa-basso finale, mentre il filtro di pre-enfasi agisce direttamente sul segnale modulante in ingresso al trasmettitore.

I filtri di pre-enfasi, $H_{pe}(f)$, e di de-enfasi, $H_{de}(f)$, sono legati dalla relazione

$$H_{pe}(f) = \frac{1}{H_{de}(f)} \quad |f| \leq W$$

ovvero sono uno l'inverso dell'altro nella banda del segnale (al di fuori sono nulli. o comunque irrilevanti).

Il filtro di de-enfasi è solitamente un semplice filtro del primo ordine di risposta

$$H_{de}(f) = \frac{1}{1 + j(f/B_{de})}$$

La frequenza B_{de} , taglio del filtro a -3dB , è notevolmente inferiore alla banda W del segnale. Come evidente $H_{de}(f) \approx 1$ per $|f| \ll B_{de}$ e $H_{de}(f) \approx B_{de}/(jf)$ per $|f| \gg B_{de}$.

Il filtro di pre-enfasi è

$$H_{pe}(f) = 1 + j(f/B_{de})$$

che può essere approssimato, alle alte frequenze, da $H_{pe}(f) \approx (jf)/B_{de}$ per $|f| \gg B_{de}$. Quindi H_{pe} agisce, alle alte frequenze del segnale, come un filtro *differenziatore*. Ma differenziare un segnale prima di una modulazione FM equivale a modulare PM! Quindi la modulazione FM con pre-enfasi è una combuinazione di PM ed FM: essa unisce i vantaggi di entrambe le modulazioni sia per quanto riguarda le interferenze che per quanto riguarda il rumore. Tuttavia alzare le alte frequenze implica, in generale, un aumento della banda di trasmissione B_t . Tale aumento non si verifica (nel senso che è trascurabile) se lo spettro del segnale modulante ha componenti piccole alle alte frequenze, come ad esempio avviene per segnali di tipo audio.

Nel caso si usi un filtro di de-enfasi il rumore a destinazione vale

$$N_D = \int_{-W}^W |H_{de}|^2 G_\xi(f) df = \frac{\eta B_{de}^3}{S_R} \left[\left(\frac{W}{B_{de}} \right) - \arctan \left(\frac{W}{B_{de}} \right) \right]$$

Se, come usuale, $W/B_{de} \gg 1$

$$\arctan \left(\frac{W}{B_{de}} \right) \approx \frac{\pi}{2} \ll \frac{W}{B_{de}}$$

e il rumore a destinazione si puo' approssimare a

$$N_D \approx \frac{\eta B_{de}^2 W}{S_R}$$

Il rapporto segnale/rumore a destinazione diventa quindi, per una FM con pre-enfasi

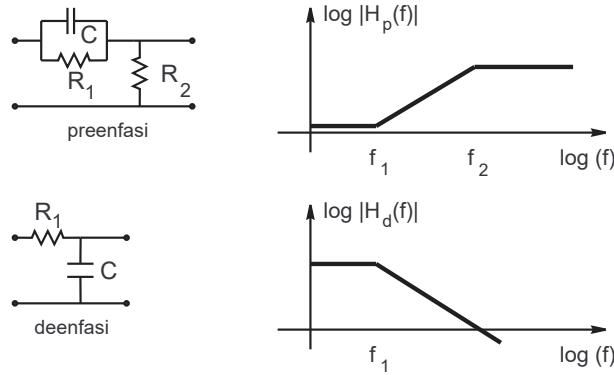
$$\left(\frac{S}{N} \right)_D \approx \left(\frac{f_\Delta}{B_{de}} \right)^2 S_x \gamma$$

che è un miglioramento di circa

$$\frac{1}{3} \left(\frac{W}{B_{de}} \right)^2$$

rispetto alla FM senza pre-enfasi.

I filtri tipici di pre-enfasi e de-enfasi sono semplici RC con costante di tempo τ_1 . In figura lo schema dei filtri ed i relativi diagrammi di Bode.



I filtri di pre-enfasi e de-enfasi hanno risposta

$$H_p(f) = \frac{1 + j(f/f_1)}{1 + j(f/f_2)} \quad H_d(f) = \frac{1}{1 + j(f/f_1)}$$

dove

$$f_1 = \frac{1}{2\pi\tau_1} = \frac{1}{2\pi R_1 C} \quad \text{e} \quad f_2 = \frac{1}{2\pi\tau_2} = \frac{R_1 + R_2}{2\pi R_1 R_2 C}$$

Come si vede l'effetto passa-alto del filtro di pre-enfasi viene troncato a f_2 : ovviamente $f_2 > W$ (per trasmissioni audio vale $f_2 = 25$ KHz).

In Italia la costante di tempo τ_1 è solitamente fissata a $50\mu s$, per cui $B_{de} = f_1 = 3.18$ KHz, mentre la banda del segnale modulante è fissata a $W = 15$ KHz.

In USA le trasmissioni FM vengono enfatizzate usando $\tau_1 = 75\mu s$ e $B_{de} = f_1 \approx 2.1$ KHz.

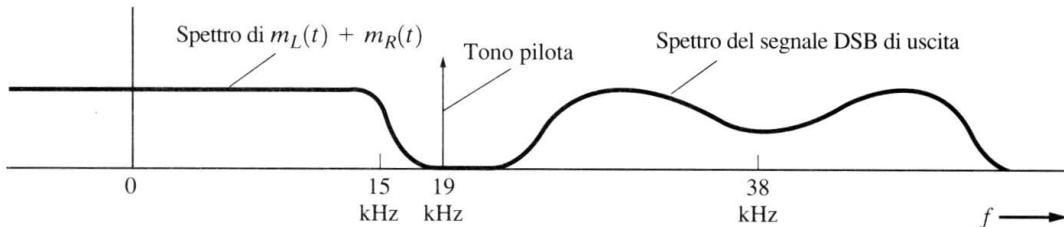
Le trasmissioni audio, come quelle in banda 88 - 108 MHz, hanno una banda di ± 75 KHz attorno alla portante.

Con tali parametri il miglioramento del rapporto segnale/rumore a destinazione è di circa 10 dB.

8.6.9 FM stereo e RDS

Per ovvi motivi la trasmissione FM stereo viene realizzata in modo che sia compatibile con il segnale mono. Per fare questo si prendono i segnali destinati ai canali destro e sinistro, entrambi di banda 50-15000 Hz, e si costruiscono due nuovi segnali mediante somma e differenza. Il segnale somma corrisponde a quello mono, il segnale differenza permette la ricostruzione dello stereo. I due segnali vengono combinati in banda base aggiungendo al segnale somma (non modificato) il segnale differenza modulato DSB con portante (soppressa) a 38 KHz. Inoltre viene aggiunto un tono pilota a 19 KHz (metà della frequenza della portante) per aiutare la demodulazione sincrona al ricevitore. Tale tono permette inoltre al ricevitore di stabilire se la trasmissione è stereo o mono.

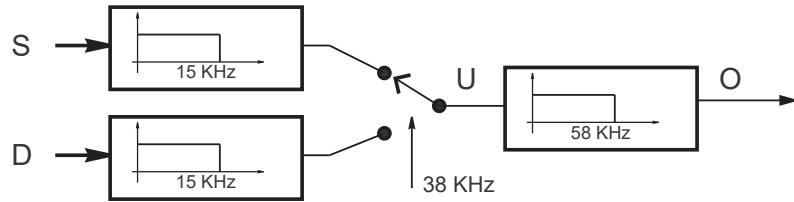
Lo spettro del segnale composito stereo è quindi



Le stazioni FM possono anche trasmettere informazioni aggiuntive (nome della stazione, titolo dei programmi ecc) attraverso il sistema RDS (Radio Data System) che prevede una ulteriore sottoporzione a 57 KHz modulata DBPSK (che vedrete forse in altri corsi, è una modulazione digitale) e permette la trasmissione alla velocità di 1200 bit/sec.

Modulatore per FM stereo*

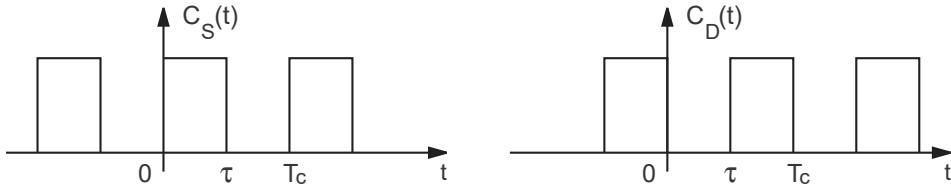
La realizzazione della FM stereo può essere effettuata mediante un semplice modulatore, il cui schema è nella figura seguente:



I due segnali sinistro (S in figura) e destro (D) vengono filtrati da un LPF con banda 15 KHz ed inviati ad un campionatore chopper a frequenza pari a 38 KHz. Ogni $T_c = 1/38K$ secondi vi sono quindi due intervalli $\tau = T_c/2$ nei quali si alternano i due canali destro e sinistro.

Il segnale a valle del chopper, $U(t)$, si può allora descrivere come la somma del canale sinistro, $x_S(t)$ moltiplicato per la forma d'onda $C_S(t)$ in figura più il canale destro, $x_D(t)$ moltiplicato per $C_D(t)$, ovvero

$$U(t) = x_S(t) \cdot C_S(t) + x_D(t) \cdot C_D(t)$$



Chiamando $x_{Sc}(t) = x_S(t) \cdot C_S(t)$ e $x_{Dc}(t) = x_D(t) \cdot C_D(t)$ ed analizzando il dominio delle frequenze si trova che

$$X_{Sc}(f) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \text{sinc}\left(\frac{k}{2}\right) e^{-j\pi k/2} X_S\left(f - \frac{k}{T_c}\right)$$

e

$$X_{Dc}(f) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \text{sinc}\left(\frac{k}{2}\right) e^{j\pi k/2} X_D\left(f - \frac{k}{T_c}\right)$$

Questo risultato è stato ricavato in sezione 2.1.6, descrivendo il campionatore chopper. Ciascun segnale chopper si compone quindi di infinite repliche pesate dello spettro del canale destro o sinistro centrate su multipli di $1/T_c = 38$ KHz. Da notare il termine di fase $e^{\pm j\pi k/2}$, dove il segno distingue i due canali: per $k = 0$ la fase è nulla, per $k = 1$ le due fasi diventano opposte, per cui in banda base si ottiene la somma mentre le prime repliche sommate rappresentano la *differenza*.

Il segnale differenza è tipicamente ad energia inferiore rispetto al segnale somma e viene inoltre scalato di un fattore $\text{sinc}(1/2) \approx 0.64$: in tale modo l'aumento di banda dovuto alla trasmissione stereo non è molto elevato.

A valle del filtro LPF finale di banda 58 KHz sopravvivono solo le repliche per $k = 0$ e $k = \pm 1$: il segnale in uscita è allora (in frequenza)

$$\begin{aligned} O(f) = & \frac{1}{2} X_S(f) + \frac{1}{2} X_D(f) - \frac{1}{2} \text{sinc}(1/2) j \left[X_S(f - \frac{1}{T_c}) - X_D(f - \frac{1}{T_c}) \right] + \\ & + \frac{1}{2} \text{sinc}(-1/2) j \left[X_S(f + \frac{1}{T_c}) - X_D(f + \frac{1}{T_c}) \right] \end{aligned}$$

che rappresenta quindi lo spettro voluto della FM stereo, a cui si aggiunge il tono pilota a 19 KHz. Da notare che la decodifica si può ovviamente realizzare usando un commutatore analogo a quello in trasmissione che separi i due canali.

FM quadrifonica*

Che io sappia non esiste uno standard ufficiale per la FM quadrifonica, tuttavia a volte l'autoradio della mia macchina riceve in quadrifonia...

Quello che sembra essere lo standard usato è il sistema *Quadraplex*, inventato da Louis Dorren nel 1969. Nel sistema *Quadraplex* ci sono due sottoportanti in più rispetto alla FM stereofonica, posizionate a 38 e 76 kHz, e lo schema di massima della trasmissione è il seguente:

- da 50 Hz a 15 kHz: canale principale, somma dei quattro segnali (LF+LB+RF+RB), per mantenere la compatibilità con i ricevitori FM mono. I quattro canali sono ovviamente Left/Front, Right/Front, Left/Back e Right/Back.
- da 23 a 53 kHz (usando una sottoportante a 19 kHz): (LF+LB) - (RF+RB). Ovvero il segnale differenza Left meno Right. In questo modo si mantiene la compatibilità con i ricevitori FM stereo.
- da 23 a 53 kHz (usando una sottoportante a 38 kHz sfasata di 90 gradi rispetto a quella a 19 kHz): (LF+RF) - (LB+RB). Ovvero il segnale differenza Front meno Back.
- da 61 a 91 kHz (usando una sottoportante a 76 kHz): (LF+RB) - (LB+RF). Ovvero il segnale differenza in diagonale.
- sottoportante a 95 kHz, con la stessa fase della sottoportante a 19 kHz, per fornire servizi aggiuntivi.

Così come il segnale stereofonico può essere considerato l’alternanza dei canali destro e sinistro alla frequenza di 38 kHz, il segnale quadrifonico alterna i quattro canali LF, RF, LB e RB con frequenza 76 kHz.

Ci sono state variazioni e differenti standard proposti, ma questo è stato scelto come lo standard per la quadrifonia negli USA.

8.6.10 Sintesi sulle modulazioni angolari

Riassumo in questa sezione le formule base che lo studente è tenuto a conoscere relativamente alle modulazioni angolari (per essere chiari: queste sono le formule da sapere se si vuole “passare” l’esame, per avere voti alti bisogna conoscere anche il resto...).

Le modulazioni angolari sono la PM (modulazione di fase) e la FM (modulazione di frequenza). Il segnale modulato si esprime come:

$$y(t) = A_c \cos(\omega_c t + \theta(t))$$

dove $\theta(t)$ vale, per PM ed FM,

$$\begin{array}{lll} \text{PM} & \theta(t) = \theta_\Delta x(t) & \theta_\Delta \in [-\pi, \pi] \\ \text{FM} & \theta(t) = 2\pi f_\Delta \int_{t_0}^t x(\tau) d\tau + \theta(t_0) & t \geq t_0 \end{array}$$

e il segnale modulante $x(t)$ è normalizzato in ampiezza, $|x(t)| \leq 1$ e passabasso di banda W .

La larghezza di banda in trasmissione è, per entrambe, teoricamente illimitata. Considerando tuttavia come banda l’intervallo che contiene la quasi totalità della potenza, si arriva alle seguenti formule approssimate, valide se il segnale modulante $x(t)$ è musica o parlato (ovvero ha uno spettro concentrato sulle basse frequenze).

Per la **FM** si ha la cosiddetta regola del *Carson*:

$$\begin{aligned} B_t &\approx 2(D+2)W = 2(f_\Delta + 2W) & 2 \leq D \leq 10 \\ B_t &\approx 2(D+1)W = 2(f_\Delta + W) & D > 10 \end{aligned} \quad (8.71)$$

dove $\Delta = f_\Delta/W$.

Per la modulazione **PM** la banda può essere stimata mediante

$$B_t \approx 2(\theta_\Delta + 1)W$$

La **potenza in trasmissione** vale, per entrambe le modulazioni,

$$S_t = \frac{A_c^2}{2}$$

I rapporti segnale/rumore a destinazione per la **FM** e **PM** sono

$$\begin{aligned} \text{FM} \quad \left(\frac{S}{N} \right)_D &= 3 \frac{f_\Delta^2}{W^2} S_x \frac{S_R}{\eta W} = 3 D^2 S_x \gamma \\ \text{PM} \quad \left(\frac{S}{N} \right)_D &= \theta_\Delta^2 S_x \frac{S_R}{\eta W} = \theta_\Delta^2 S_x \gamma \end{aligned}$$

Entrambe sono valide solo se il rapporto segnale/rumore in ricezione è almeno pari a 10 dB.

Capitolo 9

Teoria della ricezione ottima

9.1 Definizione di sistema di trasmissione discreto

Un sistema di trasmissione discreto può essere rappresentato in modo abbastanza generale dal modello in figura.

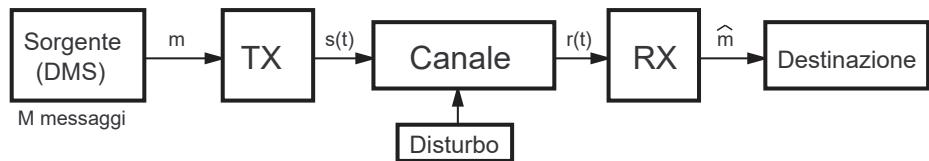


Figura 9.1: Schema di sistema di trasmissione discreto per canali a forma d'onda.

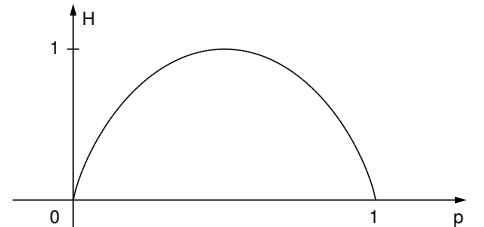
La sorgente di informazione produce un messaggio m scelto fra un insieme di messaggi $\{m_i\}$, $i = 1, \dots, M$, noto a priori. Per ogni messaggio esiste una probabilità di emissione $P(m_i)$, ovviamente $\sum_{i=1}^M P(m_i) = 1$. La sorgente è di tipo DMS, ovvero discreta e senza memoria: discreta poiché il numero M di messaggi emessi è finito e senza memoria se le probabilità di emissione ad un certo istante non dipendono dai simboli emessi in precedenza (ovvero le $P(m_i)$ sono delle costanti).

L'insieme delle probabilità $P(m_i)$ caratterizza statisticamente la sorgente. La quantità di informazione media emessa dalla sorgente si calcola mediante

$$H = - \sum_{i=1}^M P(m_i) \log_2 P(m_i)$$

H è detta entropia della sorgente. Un altro modo per definirla è l'incertezza media statistica che il messaggio è in grado di togliere nel destinatario. L'entropia si misura in numero medio di bit di informazione a simbolo emesso (bit/simbolo): il suo valore corrisponde anche al numero medio di domande a risposta binaria (si/no) necessarie per individuare ciascun simbolo emesso. L'entropia H è sempre maggiore o uguale a zero ed è limitata superiormente dal suo valore massimo $\log_2 M$, assunto quando tutti i simboli hanno la stessa probabilità di emissione (massima incertezza). Il valore 0 corrisponde ad assenza di incertezza e si verifica se un simbolo ha probabilità di emissione 1 e gli altri 0. Quindi $0 \leq H \leq \log_2 M$.

Vediamo il caso più semplice di tutti: la sorgente binaria ($M = 2$). Chiamando p la probabilità di emissione del primo simbolo (il secondo ha ovviamente probabilità $1-p$) la funzione entropia $H(p)$ assume valori fra 0 ed 1, come mostrato in figura. Da notare che il massimo è per $p = 1/2$ e vale $H(0.5) = 1$.



Trasmettitore. Il trasmettitore fa corrispondere ad ogni messaggio m_i un segnale $s_i(t)$ idoneo ad essere trasmesso sul canale.

Ricevitore. Se non ci fosse il canale il ricevitore non riceverebbe alcun segnale: in queste condizioni, senza informazioni sul messaggio trasmesso, l'unica informazione disponibile è la probabilità di emissione dei simboli $P(m_i)$, detta *probabilità a priori*.

Se, al contrario, il canale di trasmissione fosse perfetto, il ricevitore riceverebbe un segnale identico a quello trasmesso e potrebbe, mediante una operazione (detta decodifica o demodulazione) esattamente reciproca di quella effettuata da trasmettitore, stabilire senza mai sbagliare quale messaggio è stato trasmesso.

I canali di trasmissione perfetti non esistono ed il segnale ricevuto è sempre corrotto da disturbi di varia natura, come distorsioni, interferenze, rumore. Il segnale ricevuto $r(t)$ è quindi generalmente diverso dal segnale trasmesso $s(t)$, essendo a questo legato soltanto da un meccanismo di transizione probabilistica. Il ricevitore, che prima di ricevere $r(t)$ conosce solo le probabilità a priori, conserva comunque, anche dopo aver ricevuto $r(t)$, una certa quantità di incertezza. La situazione è rappresentata matematicamente dalle *probabilità a posteriori*,

$$P(m_i/r) \quad i = 1, \dots, M$$

Se il canale è così inutile che il segnale ricevuto non porta informazione su cosa è stato trasmesso vale che $P(m_i/r) = P(m_i)$, ovvero mi posso basare solo sulle probabilità a priori. Se il canale è perfetto, ovvero toglie ogni dubbio, ed indicando con m_h il segnale trasmesso, vale che $P(m_h/r) = 1$ e $P(m_i/r) = 0$ per $i \neq h$ (quindi non sbaglio mai).

Nei casi intermedi, che sono poi quelli che si incontrano in pratica, il canale porta informazione, ma non tutta quella che servirebbe per togliere ogni incertezza (pari all'entropia della sorgente, ovvero H). Resta ancora da eliminare l'incertezza a posteriori, definita da

$$H_r = - \sum_{i=1}^M P(m_i/r) \log_2 P(m_i/r)$$

che è minore di H (almeno per canali che trasmettono qualcosa). L'incertezza a posteriori varia ovviamente a seconda del particolare r ricevuto. Se si vuole una quantità media bisogna mediare statisticamente su tutti i segnali r possibili: tale media H_{rm} (che Shannon ha chiamato *equivocazione*), vale

$$H_{rm} = - \sum_{i=1}^M \sum_r P(m_i/r) \log_2 P(m_i/r) P(r) = - \sum_{i=1}^M \sum_r P(m_i, r) \log_2 P(m_i/r)$$

avendo assunto che il numero possibile di segnali ricevuti sia finito e avendo usato l'identità $P(m_i, r) = P(m_i/r)P(r)$.

La quantità di informazione media trasmessa dal canale è quindi $H_c = H - H_{rm}$.

Shannon ha mostrato che la quantità di informazione trasmissibile da un certo canale disturbato nell'unità di tempo ha un limite superiore C , che viene chiamato **capacità del canale**. Tale limite C è raggiungibile solo teoricamente: ci si può avvicinare solo mediante apparati di complessità crescente, che diventa proibitiva vicino a C . Tuttavia è fondamentale conoscere il limite teorico di ciascun canale, anche solo per capire quanto bene lo si usa.

Larghezza di banda. La larghezza di banda necessaria a trasmettere i segnali $s_i(t)$ dipende ovviamente da come questi sono fatti. Se si suppone che i segnali $s_i(t)$ siano generati a partire da N funzioni base ortonormali, come visto nel capitolo 3, è abbastanza intuitivo che più funzioni base si usano, cioè maggiore è N , più banda serve. Se T è la cadenza di emissione dei simboli dalla sorgente in secondi e B è la banda in Hz vale che

$$N \approx 2BT \quad T \approx \frac{N}{2B} \quad B \approx \frac{N}{2T}$$

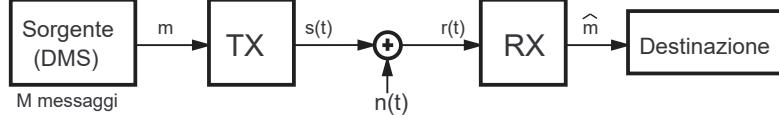
Queste sono da interpretarsi come limite: se scelgo opportunamente le N funzioni base allora posso arrivare alle relazioni sopra indicate.

In particolare, considerando $N = 1$ si ottiene che la banda richiesta vale $B \approx 1/2T$: da questo deriva che una PAM binaria può trasmettere circa 2 bit di informazione per ogni Hz di banda.

Canali ideali. Nel seguito ci occuperemo di canali ideali, ovvero canali a banda illimitata ed ampiezza unitaria, nei quali il segnale in uscita è una copia esatta di quello di ingresso (a parte l'aggiunta di rumore). Canali siffatti non esistono, sia perché ogni canale reale attenua il segnale

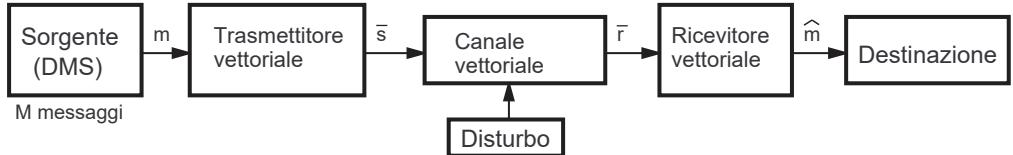
che viaggia su di esso sia perchè tutti i canali hanno banda limitata. Possiamo tuttavia accettare l'approssimazione di banda illimitata ogni volta che la banda del canale è maggiore della banda dei segnali che su di esso viaggiano. Per quanto riguarda l'attenuazione, essa agisce attenuando il segnale in transito: possiamo quindi facilmente tenerne conto inserendo un opportuno amplificatore a guadagno minore di uno (o simulando lo stesso effetto supponendo che il segnale trasmesso abbia ampiezza inferiore).

L'ultima condizione sui canali ideali è che il rumore sia additivo: lo schema che ne risulta è allora:



9.2 Canali vettoriali

Si definiscono canali vettoriali quelli usati per trasmettere messaggi mediante vettori, ossia mediante insiemi di N scalari. Visto dall'esterno, cioè prescindendo dal modo nel quale le singole componenti vengono trasmesse e recuperate, un canale vettoriale è caratterizzato dal fatto che, dato un vettore \bar{s} alla sua entrata, si ottiene in uscita un vettore \bar{r} , generalmente diverso da \bar{s} a causa di disturbi di trasmissione (rumore ed altro) di origine casuale. Un sistema di trasmissione discreto per canali vettoriali è rappresentabile come in figura.



La sorgente sceglie il messaggio m in un insieme $\{m_i\}$ di M messaggi possibili, caratterizzati ciascuno da una probabilità a priori $P(m_i)$. Il trasmettore genera, in corrispondenza del messaggio m_i , il vettore-segnale \bar{s}_i che viene inviato sul canale. Dal canale disturbato esce un vettore \bar{r} . Il ricevitore, dalla conoscenza delle probabilità a priori e del vettore ricevuto \bar{r} e basandosi sulle probabilità a posteriori $P(m_i|\bar{r})$, sceglie nell'insieme $\{m_i\}$ il messaggio \hat{m} da inviare al destinatario.

Più precisamente, una volta ricevuto un particolare vettore \bar{R} di \bar{r} , se si vuole massimizzare la probabilità $P(C)$ di prendere la decisione corretta, occorre scegliere per \hat{m} quel messaggio m_k che ha la massima probabilità a posteriori, tale cioè che

$$P(m_k|\bar{r} = \bar{R}) > P(m_i|\bar{r} = \bar{R}) \quad \text{per ogni } i \neq k \quad (9.1)$$

Infatti, per $\bar{r} = \bar{R}$ la probabilità di decisione corretta è

$$P(C|\bar{r} = \bar{R}) = P(m_k|\bar{r} = \bar{R}) \quad (9.2)$$

cosicché la probabilità di decisione corretta mediata statisticamente su tutti gli \bar{R} possibili è

$$P(C) = \int_{-\infty}^{\infty} P(C|\bar{r} = \bar{R}) f_{\bar{r}}(\bar{R}) d\bar{R} \quad (9.3)$$

essendo $f_{\bar{r}}(\bar{R})$ la densità di probabilità del vettore \bar{r} . Poichè questa è positiva, $P(C)$ è massima quando si sceglie per ogni $\bar{r} = \bar{R}$ il massimo valore di $P(C|\bar{r} = \bar{R})$. Da notare che l'integrale a secondo membro della 9.3 rappresenta in modo sintetico l'espressione

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P(C|r_1 = R_1, r_2 = R_2, \dots, r_N = R_N) f_{r_1, r_2, \dots, r_N}(R_1, R_2, \dots, R_N) dR_1 dR_2 \dots dR_N$$

Un ricevitore che opera come sopra esposto si dice a **massima probabilità a posteriori**. Poichè

$$P(m_i|\bar{r} = \bar{R}) = \frac{P(m_i) f_{\bar{r}/m_i}(\bar{R}/m_i)}{f_{\bar{r}}(\bar{R})}$$

la massimizzazione a primo membro della 9.1 equivale a massimizzare (essendo $f_{\bar{r}}(\bar{R})$ fissata)

$$P(m_i) f_{\bar{r}/m_i}(\bar{R}/m_i) = P(m_i) f_{\bar{r}/\bar{s}}(\bar{R}/\bar{s} = \bar{s}_i) \quad (9.4)$$

Spesse volte, quando le probabilità a priori non sono note (o variano...), si massimizza solo il fattore $f_{\bar{r}/m_i}(\bar{R}/m_i)$. In tal caso il ricevitore si dice **a massima verosimiglianza**: esso da la minima probabilità di errore solo quando i messaggi sono equiprobabili.

Una interpretazione geometrica di quanto precede può essere effettuata nello spazio N -dimensionale del vettore ricevuto \bar{r} . Infatti si può, per ogni valore di \bar{r} , calcolare la funzione di decisione 9.4 per tutti gli i , ed assegnare ad ciascun \bar{r} quel messaggio \hat{m}_k che rende massima la funzione di decisione. Ciò accade per tutti i valori di \bar{r} che cadono in una regione I_k dello spazio N -dimensionale considerato. Lo spazio verrà così suddiviso in M regioni disgiunte I_1, I_2, \dots, I_M (dette **regioni di decisione**) tali che, se \bar{r} cade in I_k , si sceglie $\hat{m} = m_k$. La probabilità di decisione corretta quando $m = m_i$ è data da

$$P(C/m = m_i) = P(\bar{r} \text{ in } I_k/m = m_i) \quad (9.5)$$

e la probabilità media statistica di decisione corretta su tutti gli m è

$$P(C) = \sum_{i=1}^M P(\bar{r} \text{ in } I_i/m = m_i) P(m_i) \quad (9.6)$$

9.2.1 Rumore additivo gaussiano

In questo caso il canale si limita ad aggiungere rumore gaussiano additivo alle componenti del vettore trasmesso \bar{s} , quindi

$$\bar{r} = \bar{s} + \bar{n} = (s_1 + n_1, s_2 + n_2, \dots, s_N + n_N) \quad (9.7)$$

essendo $\bar{n} = (n_1, n_2, \dots, n_N)$ il vettore rumore.

In questo caso l'evento ($\bar{r} = \bar{R}/\bar{s} = \bar{s}_1$) nell'eq. 9.4 corrisponde all'evento ($\bar{n} = \bar{R} - \bar{s}_1/\bar{s} = \bar{s}_1$) e quindi

$$f_{\bar{r}/\bar{s}}(\bar{R}/\bar{s} = \bar{s}_1) = f_{\bar{n}/\bar{s}}(\bar{R} - \bar{s}_1/\bar{s} = \bar{s}_1) \quad (9.8)$$

Se poi \bar{n} e \bar{s} sono statisticamente indipendenti, come spesso accade, vale che $f_{\bar{n}/\bar{s}} = f_{\bar{n}}$ e la funzione di decisione 9.4 diventa

$$P(m_i) f_{\bar{n}}(\bar{R} - \bar{s}_i) \quad (9.9)$$

Questa formula diventa particolarmente semplice da calcolare quando le N componenti del vettore rumore \bar{n} sono fra loro statisticamente indipendenti, a valor medio nullo e tutte con pari varianza σ_n^2 : in tal caso

$$f_{\bar{n}}(\bar{\alpha}) = \frac{1}{(2\pi\sigma_n^2)^{N/2}} e^{-\frac{\sum_{i=1}^N \alpha_i^2}{2\sigma_n^2}} = \frac{1}{(2\pi\sigma_n^2)^{N/2}} e^{-\frac{|\bar{\alpha}|^2}{2\sigma_n^2}} \quad (9.10)$$

in questo caso la funzione di decisione si scrive

$$P(m_i) e^{-\frac{|\bar{R} - \bar{s}_i|^2}{2\sigma_n^2}} \quad (9.11)$$

che viene massimizzata rendendo minima l'espressione

$$|\bar{R} - \bar{s}_i|^2 - 2\sigma_n^2 \ln P(m_i) \quad (9.12)$$

Le regioni di decisione sono facilmente individuate ricordando che $|\bar{R} - \bar{s}_i|^2$ è il quadrato della distanza fra gli estremi di \bar{R} e di \bar{s}_i . Se i segnali sono equiprobabili occorre scegliere, per ogni \bar{R} ricevuto, il segnale \bar{s}_i più vicino a \bar{R} (a distanza vettoriale minima) (vedi figura 9.2).

9.3 Trasformazione in canale vettoriale di un canale per forme d'onda

Un sistema di trasmissione per messaggi discreti, utilizzando un normale canale a forme d'onda, è illustrato in figura 9.1. Da capitolo 3 sappiamo che l'insieme degli M segnali $s_i(t)$, $i = 1, \dots, M$, a energia finita e limitati ad un intervallo temporale T , può essere rappresentato mediante la combinazione lineare di N funzioni ortonormali ϕ_k , essendo $N \leq M$ la dimensione dello spazio generato dai segnali $s_i(t)$. Quindi ad ogni segnale $s_i(t)$ è associato un vettore di coefficienti in uno spazio N -dimensionale (la sua immagine). Quindi i segnali $s_i(t)$ possono essere generati come visto in sezione 3.3.1: in tal caso il trasmettitore fa corrispondere in primo luogo al messaggio m_i il vettore $\bar{s}_i = (s_{i1}, s_{i2}, \dots, s_{iN})$ a partire dal quale viene poi sintetizzato il segnale $s_i(t)$.

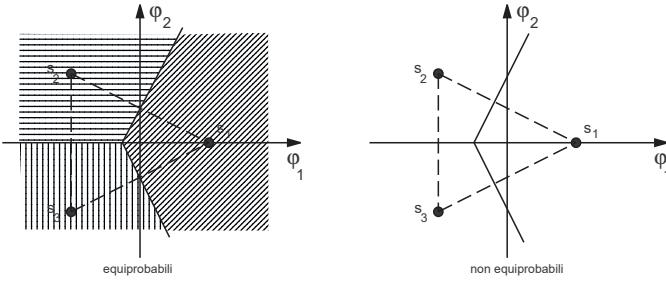


Figura 9.2: Regioni di decisione per tre segnali: a sinistra sono equiprobabili, a destra no, e vale che $P(s_1) > P(s_2) = P(s_3)$.

Analogamente il segnale $r(t)$ può essere rappresentato mediante combinazione lineare di N_r funzioni base ortonormali (essendo N_r la dimensione dello spazio dei possibili segnali ricevuti) e cioè nuovamente mediante un vettore \bar{r} le cui componenti possono essere ricavate come visto in sezione 3.3.2. Il sistema di in figura 9.1 può quindi essere attuato come mostrato in in figura 9.3. Si noti che le funzioni base necessarie per rappresentare $r(t)$ possono essere diverse, anche in numero ($N_r \neq N$) rispetto a quelle usate per rappresentare i segnali $s_i(t)$.

Se, in figura 9.3, il blocco di sintesi e di analisi vengono conglobati nel canale di trasmissione, si ottiene un canale vettoriale completamente caratterizzato dal meccanismo di transizione probabilistico da \bar{s} a \bar{r} .

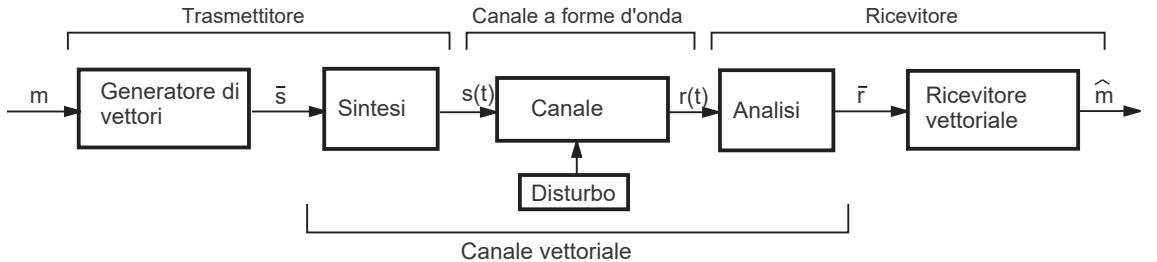


Figura 9.3: Canale vettoriale e a forme d'onda.

9.3.1 Rumore additivo gaussiano

In generale, lo studio di un canale vettoriale generico ha notevoli difficoltà. Se però il canale per forme d'onda è quello nella figura a lato, ovvero un canale ideale che aggiunge un rumore gaussiano bianco a media nulla e indipendente dai segnali trasmessi, sono possibili interessanti semplificazioni.

Il rumore $n_w(t)$ ha densità spettrale bilatera $\eta/2$. In questo caso si può dimostrare che il vettore \bar{r} che contiene tutta l'informazione che $r(t)$ può dare per la scelta di \hat{m} ha la stessa dimensione di \bar{s} ed è ottenibile da $r(t)$ usando le stesse funzioni ortonormali utilizzate per rappresentare i segnali $s_i(t)$. Attenzione alla frase precedente: se uso per l'analisi di $r(t)$ le stesse funzioni base usate per la sintesi di $s_i(t)$ NON rappresento in modo completo $r(t)$ (il che richiederebbe un numero di funzioni base infinito). Ma la parte che non rappresento è inutile per la scelta che il ricevitore deve compiere (ovvero non conoscere le componenti mancanti non cambia la probabilità di decisione corretta).

Per dimostrare quanto affermato, consideriamo l'uscita del blocco di analisi di figura 9.3 se esso usa le stesse funzioni base ϕ_k , $k = 1, \dots, N$, usate per la sintesi dei segnali e se l'ingresso è $r(t) = s(t) + n_w(t)$. Il vettore di uscita ha componenti $\bar{r} = (r_1, r_2, \dots, r_N)$, ciascuna delle quali si ottiene da

$$r_k = \int_{-\infty}^{\infty} r(t) \phi_k(t) dt = \int_{-\infty}^{\infty} s(t) \phi_k(t) dt + \int_{-\infty}^{\infty} n_w(t) \phi_k(t) dt = s_k + n_k$$

$$r_k = \int_{-\infty}^{\infty} r(t) \phi_k(t) dt = \int_{-\infty}^{\infty} s(t) \phi_k(t) dt + \int_{-\infty}^{\infty} n_w(t) \phi_k(t) dt = s_k + n_k \quad (9.13)$$

Conoscere \bar{r} equivale alla conoscenza di un segnale ricevuto $r_1(t)$ dato da

$$r_1(t) = \sum_{k=1}^N s_k \phi_k(t) + \sum_{k=1}^N n_k \phi_k(t) = s(t) + n(t) \quad (9.14)$$

avendo chiamato $n(t) = \sum_{k=1}^N n_k \phi_k(t)$. Dovrebbe essere chiaro che $n(t) \neq n_w(t)$, quindi

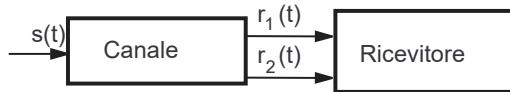
$$r(t) = s(t) + n_w(t) = s(t) + n(t) + [n_w(t) - n(t)] = r_1(t) + r_2(t) \quad (9.15)$$

essendo ovviamente $r_2(t) = n_w(t) - n(t)$.

Quindi utilizzare il blocco di analisi così definito equivale a considerare del segnale ricevuto la sola parte $r_1(t)$, trascurando $r_2(t)$. Occorre quindi esaminare se la conoscenza di $r_2(t)$ sia importante per la determinazione di \hat{m} . A tale scopo dimostriamo il teorema riportato nella sezione che segue.

9.3.2 Teorema dell'irrilevanza

Sia dato un sistema di trasmissione come quello in figura, nel quale il canale mette a disposizione del ricevitore due segnali ricevuti, $r_1(t)$ e $r_2(t)$.



Evidentemente, visto che ora il ricevitore ha a disposizione due segnali, dovrà massimizzare una probabilità a posteriori che dipende da entrambi, ovvero $P(s/r_1, r_2)$ (vedi eq. 9.1).

Il segnale $r_2(t)$ è ovviamente irrilevante se

$$P(s/r_1, r_2) = P(s/r_1) \quad (9.16)$$

Poichè, applicando ripetutamente $P(A/B) = P(AB)/P(B)$,

$$P(s/r_1, r_2) P(r_2/r_1) = \frac{P(sr_1r_2)}{P(r_1r_2)} \frac{P(r_2r_1)}{P(r_1)} = \frac{P(sr_1r_2)}{P(r_1)} = P(s, r_2/r_1)$$

e

$$P(r_2/r_1, s) P(s/r_1) = \frac{P(sr_1r_2)}{P(r_1s)} \frac{P(sr_1)}{P(r_1)} = \frac{P(sr_1r_2)}{P(r_1)} = P(s, r_2/r_1)$$

si ha

$$\frac{P(s/r_1, r_2)}{P(s/r_1)} = \frac{P(r_2/r_1, s)}{P(r_2/r_1)} \quad (9.17)$$

l'eq. 9.16 equivale a

$$P(r_2/r_1, s) = P(r_2/r_1) \quad (9.18)$$

Le eq. 9.16 e 9.18 sono condizioni necessarie e sufficienti per l'irrilevanza di r_2 . Quindi è sufficiente che r_2 sia indipendente sia da r_1 che da s .

Dimostrazione dell'irrilevanza per rumore AWGN*

Dato il canale visto in sezione 9.3.1, sappiamo che $r_2(t)$ è indipendente dal segnale $s(t)$, infatti $r_2(t) = n_w(t) - n(t)$.

Se riusciremo a dimostrare che $r_2(t)$ è indipendente anche da $n(t)$ avremo dimostrato quanto vogliamo, cioè che r_2 è indipendente sia da s che da $r_1(t) = s(t) + n(t)$.

Poichè $n_w(t)$ è un processo aleatorio gaussiano, sono anche gaussiani i p.a. $n(t)$ e $r_2(t)$: questi ultimi sono quindi statisticamente indipendenti se e solo se, per qualsiasi τ e t , la covarianza delle v.a. $n(\tau)$ e $r_2(t)$ vale zero (v.a. gaussiane scorrelate sono indipendenti, vedi sez 6.10.3), ovvero se

$$E\{n(\tau)r_2(t)\} - E\{n(\tau)\} E\{r_2(t)\} = E\{n(\tau)r_2(t)\} = 0 \quad (9.19)$$

che può essere scritta, ricordando che $n(t) = \sum_{k=1}^N n_k \phi_k(t)$ come

$$E\left\{r_2(t) \sum_{k=1}^N n_k \phi_k(\tau)\right\} = \sum_{k=1}^N \phi_k(\tau) E\{n_k r_2(t)\} = 0 \quad (9.20)$$

dove le funzioni ϕ_k sono state portate fuori dall'operatore di media non avendo componenti aleatorie. Per dimostrare l'eq. 9.19 basta quindi dimostrare che

$$E\{n_k r_2(t)\} = 0 \quad \text{per ogni } k \text{ e } t \quad (9.21)$$

Siccome $r_2(t) = n_w(t) - n(t)$, $n(t) = \sum_{k=1}^N n_k \phi_k(t)$ e $n_k = \int_{-\infty}^{\infty} n_w(t) \phi_k(t) dt$ si ha che

$$\begin{aligned} E\{n_k r_2(t)\} &= E\{n_k n_w(t)\} - E\{n_k n(t)\} = \\ &= \int_{-\infty}^{\infty} E\{n_w(t) n_w(\tau)\} \phi_k(\tau) d\tau - \sum_{i=1}^N E\{n_k n_i\} \phi_i(t) \end{aligned} \quad (9.22)$$

$E\{n_w(t) n_w(\tau)\}$ è l'autocorrelazione di $n_w(t)$ calcolata in $t - \tau$, ovvero $R_{n_w}(t - \tau)$, e poichè $n_w(t)$ è un rumore gaussiano bianco (con autocorrelazione impulsiva, vedi eq. 7.69), l'integrale nella 9.22 vale

$$\int_{-\infty}^{\infty} R_{n_w}(t - \tau) \phi_k(\tau) d\tau = \frac{\eta}{2} \int_{-\infty}^{\infty} \delta(t - \tau) \phi_k(\tau) d\tau = \frac{\eta}{2} \phi_k(t) \quad (9.23)$$

Per quanto riguarda la somatoria nella 9.22 si può calcolare osservando che

$$\begin{aligned} E\{n_k n_i\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E\{n_w(\tau) n_w(\xi)\} \phi_k(\tau) \phi_i(\xi) d\tau d\xi = \\ &= \frac{\eta}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\tau - \xi) \phi_k(\tau) \phi_i(\xi) d\tau d\xi = \\ &= \frac{\eta}{2} \int_{-\infty}^{\infty} \phi_k(\tau) \phi_i(\tau) d\tau \end{aligned}$$

e quindi, ricordando l'ortonormalità delle $\phi_h(t)$,

$$\sum_{i=1}^N E\{n_k n_i\} \phi_i(t) = \frac{\eta}{2} \phi_k(t) \quad (9.24)$$

I risultati ottenuti in 9.23 e 9.24 comportano l'annullarsi della 9.22 e perciò l'irrilevanza di $r_2(t)$.
C.V.D.

9.4 Ricevitore ottimo per canali a forme d'onda

L'irrilevanza di $r_2(t)$ appena dimostrata assicura la legittimità dell'adozione del canale equivalente vettoriale mostrato in figura 9.4 (B) al posto del canale a forme d'onda visto in sezione 9.3.1. Nella stessa figura è mostrata (in A) una possibile realizzazione del canale vettoriale mediante forme d'onda. Le funzioni $\phi_k(t)$ si suppongono definite fra zero e T (e nulle altrove). Una diversa realizzazione è possibile usando, al posto dei filtri adattati, un banco di correlatori (vedi sezioni 3.3.1 e 3.3.2).

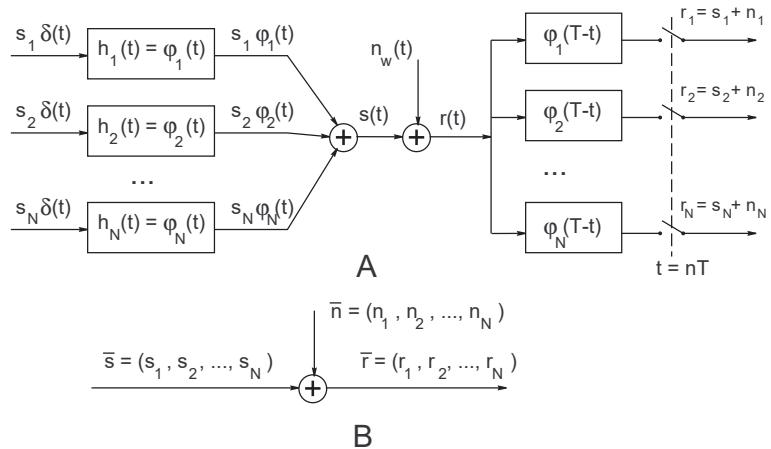


Figura 9.4: Canale vettoriale (A) e sua schematizzazione (B).

Poichè il rumore $n_w(t)$ è gaussiano a media nulla e con densità spettrale bilatera $G_{n_w}(f) = \eta/2$ e i filtri in figura 9.4 (A) sono ad energia unitaria, le N componenti del vettore \bar{n} sono v.a. gaussiane indipendenti (fra di loro e dai segnali) a media nulla e con varianza pari a $\sigma_n^2 = \eta/2$. La funzione di decisione di eq. 9.12 diventa quindi

$$|\bar{r} - \bar{s}_i|^2 = \eta \ln P(m_i) \quad (9.25)$$

Il ricevitore vettoriale deve calcolare la 9.25 per $i = 1, 2, \dots, M$ e scegliere come segnale trasmesso quello che rende minimo il valore della 9.25 stessa. Per eseguire il calcolo si può notare che

$$|\bar{r} - \bar{s}_i|^2 = \sum_{k=1}^N (r_k - s_{ik})^2 = \sum_{k=1}^N (r_k^2 - 2r_k s_{ik} + s_{ik}^2) = |\bar{r}|^2 - 2\bar{r} \cdot \bar{s}_i + |\bar{s}_i|^2$$

quindi è sufficiente calcolare l'espressione

$$\bar{r} \cdot \bar{s}_i + c_i \quad \text{dove} \quad c_i = \frac{1}{2} (\eta \ln P(m_i) - |\bar{s}_i|^2) \quad (9.26)$$

e scegliere in segnale s_i che la rende **massima**.

Il ricevitore vettoriale ottimo è allora quello di figura 9.5, dove il blocco denominato matrice di pesatura ha M uscite, ciascuna delle quali vale $\bar{r} \cdot \bar{s}_i = \sum_{k=1}^N r_k s_{ik}$, $i = 1, 2, \dots, M$.

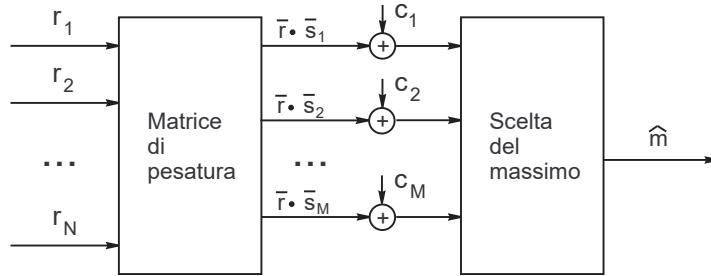


Figura 9.5: Ricevitore vettoriale ottimo.

Il ricevitore ottimo per forme d'onda è mostrato in figura 9.6. Tale ricevitore può essere sostituito da quello in figura 9.7, nel quale i filtri adattati alle funzioni base sono sostituiti da filtri direttamente adattati ai segnali. In questo modo non esiste più la matrice di pesatura, il che semplifica il ricevitore, ma i filtri sono ora M anziché N e quindi il suo impiego è conveniente solo se il numero di segnali M non è molto maggiore del numero di funzioni base N .

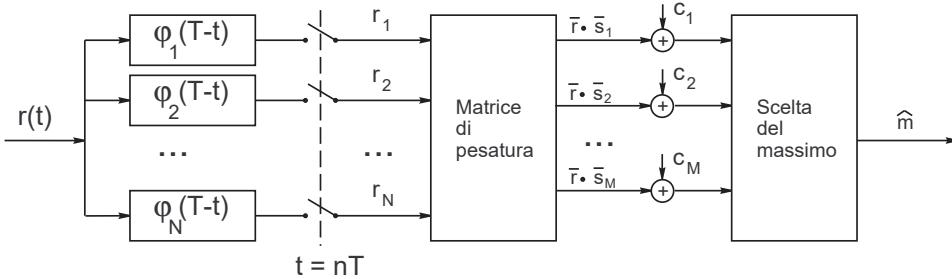


Figura 9.6: Ricevitore vettoriale ottimo per forme d'onda con filtri adattati alle funzioni base.

Una giustificazione più fisica della struttura dei ricevitori ottimi sopra illustrati si può avere considerando la seguente proprietà del filtro adattato, implicita nella trattazione precedentemente svolta: se viene ricevuto un segnale $s_1 \phi_1(t)$ (in cui $\phi_1(t)$ è noto e definito fra 0 e T e s_1 è uno scalare da determinare) immerso in un rumore bianco, il filtro adattato a $\phi_1(t)$ seguito da un campionatore all'istante T è il miglior circuito per estrarre il parametro s_1 , vale a dire ottenere s_1 con il minimo disturbo (medio) sovrapposto.

9.5 Calcolo delle prestazioni di particolari insiemi di segnali

Abbiamo visto come un canale caratterizzato semplicemente da rumore additivo gaussiano bianco sia trasformabile in un equivalente canale vettoriale qualora i segnali trasmessi e ricevuti vengano

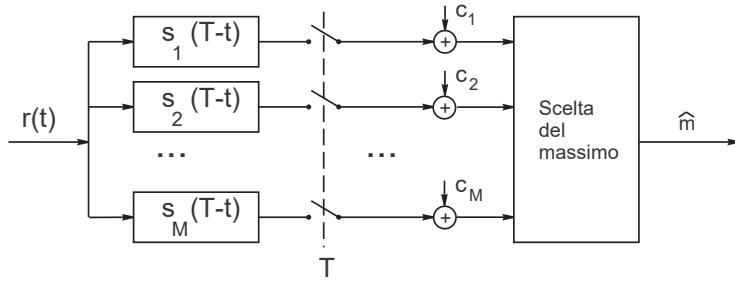


Figura 9.7: Ricevitore vettoriale ottimo per forme d'onda con filtri adattati ai segnali.

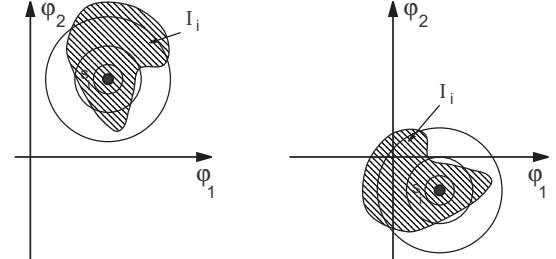
rappresentati come combinazione lineare di appropriate funzioni base ortonormali ed il ricevitore utilizzi un banco di filtri adattati alle stesse funzioni base (o, in modo equivalente, un insieme di correlatori). Ricordando quanto visto, si può notare una fondamentale proprietà: poiché sul canale vettoriale equivalente il vettore rumore \bar{n} ha componenti che sono sempre le stesse, qualunque sia l'insieme di funzioni base ortonormali adottate, la qualità della trasmissione di un vettore N -dimensionale è *indipendente dal particolare insieme di N funzioni base ortonormali adottate*, ma dipende solo dalle componenti (s_1, s_2, \dots, s_N) che caratterizzano il vettore \bar{s} .

In questa sezione si valuterà la probabilità di errore per insiemi interessanti di segnali vettoriali. A tale scopo si stabiliranno innanzitutto le regioni di decisione I_i discusse in sezione 9.2 e si calcolerà la probabilità media di errore (o di decisione corretta) mediante l'equazione 9.6.

Prima di procedere è opportuno premettere qualche osservazione sulle caratteristiche dei segnali-vettori in rapporto alle regioni di decisione.

Consideriamo per semplicità uno spazio bidimensionale, il punto rappresentativo del segnale \bar{s}_i (vedi figura) e supponiamo che la corrispondente regione di decisione sia I_i .

Il rumore che il canale aggiunge a \bar{s}_i ha distribuzione di probabilità con simmetria sferica: i cerchi concentrici a \bar{s}_i sono perciò linee di equiprobabilità del rumore aggiunto a \bar{s}_i .



Si nota immediatamente che se \bar{s}_i viene traslato e la regione di decisione ruotata attorno ad esso, la probabilità condizionata di decisione corretta $P(C/m = m_i)$ non cambia. Quindi, dato un insieme di segnali e fissate le relative regioni di decisione, la probabilità di errore resta la stessa comunque venga traslato e ruotato rispetto agli assi cartesiani l'insieme dei punti rappresentativi dei segnali (e naturalmente le relative regioni di decisione), mantenendone rigida la posizione relativa. Con questa procedura si possono trovare infiniti insiemi di segnali, equivalenti nei riguardi della probabilità di errore, anche se eventualmente caratterizzati da diversa energia (da ricordare che la distanza dall'origine del punto che rappresenta un segnale è la radice della sua energia).

L'energia media di un insieme di M segnali è

$$E_m = \sum_{i=1}^M E_i P(m_i) = \sum_{i=1}^M |\bar{s}_i|^2 P(m_i) \quad (9.27)$$

Questa energia equivale al momento di inerzia rispetto all'origine di masse di valore $P(m_i)$ posti nei punti \bar{s}_i . D'altra parte il momento di inerzia è minimo se preso rispetto al centro di massa, e ciò indica come si può, a parità di probabilità di errore, minimizzare l'energia media statistica. Infatti il centro di massa di un insieme di masse $P(m_i)$ poste agli estremi dei vettori \bar{s}_i si trova all'estremo del vettore \bar{a}

$$\bar{a} = \sum_{i=1}^M P(m_i) \bar{s}_i = E[\bar{s}] \quad (9.28)$$

Allora l'insieme di segnali $\{\bar{s}_i'\}$ a minima energia media equivalente all'insieme $\{\bar{s}_i\}$ si ottiene ponendo

$$\{\bar{s}_i'\} = \{\bar{s}_i\} - \bar{a} \quad (9.29)$$

Passiamo ora a calcolare la probabilità di errore di particolari insiemi di segnali, privilegiando insiemi *rettangolari*, ovvero tali che i punti rappresentativi dei segnali siano disposti sui vertici o lati di un rettangolo N -dimensionale.

9.5.1 Segnali binari

Il più generale insieme di segnali binari è riportato in figura 9.8 a). Un insieme equivalente come probabilità di errore è in figura 9.8 b): quest'ultimo è anche ottimo sotto l'aspetto dell'energia media richiesta nel caso in cui $P(m_1) = P(m_2) = 1/2$. Per segnali **equiprobabili** è evidente che le regioni di decisione corrispondono ai punti a minima distanza dai segnali e dunque ai semipiani positivo e negativo in figura b). Da notare che l'insieme in b) richiede una sola funzione base, anziché due come l'insieme a).

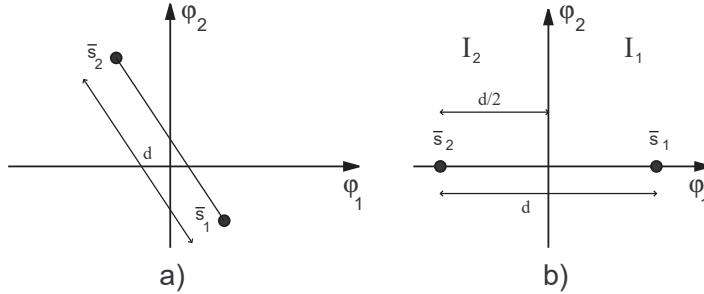


Figura 9.8: Segnali binari.

La probabilità di errore se viene trasmesso il segnale \bar{s}_2 vale, per i segnali in b),

$$P(E/m_2) = P(\bar{R} \text{ in } I_1/m_2) = P(n_1 > \frac{d}{2}) \quad (9.30)$$

essendo n_1 l'unica componente del vettore \bar{n} . Quindi

$$P(E/m_2) = \int_{d/2}^{\infty} \frac{1}{\sqrt{\pi\eta}} e^{-\frac{\alpha^2}{\eta}} d\alpha \quad (9.31)$$

ponendo $\gamma = \alpha/\sqrt{(\eta/2)}$

$$P(E/m_2) = \int_{d/\sqrt{2\eta}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\gamma^2}{2}} d\gamma = \mathbf{Q}\left(\frac{d}{\sqrt{2\eta}}\right) \quad (9.32)$$

dove $\mathbf{Q}()$ è la funzione vista in sezione 5.14. Poichè la configurazione è simmetrica ne risulta che la probabilità di errore se viene trasmesso il segnale \bar{s}_1 , cioè $P(E/m_1)$, è la stessa e quindi la probabilità di errore media vale

$$P(E) = \mathbf{Q}\left(\frac{d}{\sqrt{2\eta}}\right) \quad (9.33)$$

La formula 9.33 vale per qualunque insieme di due segnali separati da una distanza d , come ad esempio quelli di figura 9.8 a). Nel caso di segnali a minima energia media vale che $d/2 = \sqrt{E_s}$ e quindi

$$P(E) = \mathbf{Q}\left(\sqrt{\frac{E_s}{\eta/2}}\right) \quad (9.34)$$

Le configurazioni come in figura 9.8 b), composte da due segnali con baricentro nell'origine e posti a uguale distanza ma direzione opposta rispetto all'origine, si indicano come **segnali antipodali** e, come ricordato, sono ottimi nello sfruttare l'energia disponibile.

Se invece di due segnali antipodali si usano due **segnali ortogonali** ed equienergetici, come in figura 9.9 a), vale che $d = \sqrt{2 E_s}$ e quindi la probabilità di errore è

$$P(E) = \mathbf{Q}\left(\sqrt{\frac{E_s}{\eta}}\right) \quad (9.35)$$

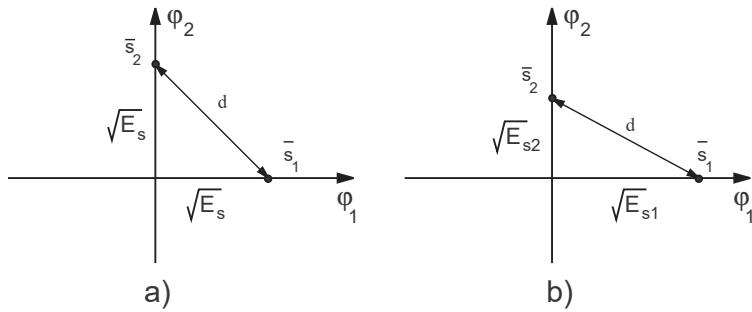


Figura 9.9: Segnali binari ortogonali.

A parità di tasso di errore i segnali ortogonali richiedono quindi energia doppia (3 dB in più) rispetto ai segnali antipodali.

Lo stesso risultato si ottiene se i due segnali ortogonali non hanno la stessa energia, come in figura 9.9 b). L'energia media vale infatti

$$E_s = \frac{1}{2} E_{s1} + \frac{1}{2} E_{s2} = \frac{1}{2} (E_{s1} + E_{s2})$$

e la distanza è $d = \sqrt{E_{s1} + E_{s2}} = \sqrt{2E_s}$. Come caso particolare di due segnali ortogonali si può considerare quello in cui uno dei due segnali ha energia nulla (è uguale a zero).

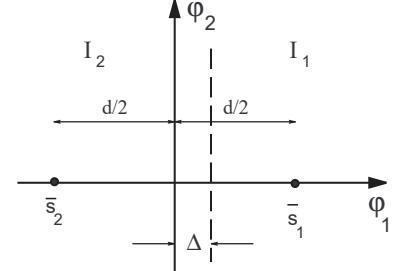
Segnali binari non equiprobabili

Analizziamo ora il caso di segnali binari a diversa probabilità.

Visto che il solo parametro che conta è la reciproca distanza, consideriamo una configurazione comoda da trattare, ovvero i segnali antipodali di figura. Le regioni di decisione vengono ovviamente modificate, essendo la linea verticale a distanza Δ (positiva o negativa) dall'origine il nuovo confine fra I_1 e I_2 . Vediamo di calcolare il valore di Δ . Le regioni di decisione sono definite come visto in eq. 9.25, ovvero da

$$\min_i [|\bar{r} - \bar{s}_i|^2 - \eta \ln P(m_i)]$$

Bisogna trovare il valore di $\bar{r} = (\Delta)$ tale che



e quindi

$$\Delta = \frac{\eta}{2d} \ln \frac{P(m_2)}{P(m_1)} \quad (9.36)$$

la probabilità di errore vale

$$P(E) = P(m_1) \mathbf{Q} \left(\frac{d - 2\Delta}{\sqrt{2\eta}} \right) + P(m_2) \mathbf{Q} \left(\frac{d + 2\Delta}{\sqrt{2\eta}} \right) \quad (9.37)$$

Se i segnali sono equiprobabili $\Delta = 0$ e la 9.37 fornisce lo stesso valore della 9.33. La 9.37 resta valida per qualsiasi coppia di segnali in posizione arbitraria, purché a distanza d . Anche la 9.36 resta valida, a patto di interpretare Δ come lo spostamento dal punto intermedio fra \bar{s}_1 e \bar{s}_2 attraverso cui passa il confine fra le regioni di decisione.

Da notare infine come la configurazione in figura non sia quella ottima dal punto di vista della minimizzazione dell'energia. L'ottimo si ottiene spostando le immagini dei due segnali della quantità $\bar{a} = (d/2)P(m_1) - (d/2)P(m_2)$, in modo che il baricentro cada nell'origine.

9.5.2 Segnali posti ai vertici di un ipercubo

Consideriamo i 2^N segnali, che supporremo equiprobabili, posti ai vertici di un ipercubo, con centro nell'origine, in uno spazio N -dimensionale. In figura 9.10 sono mostrate due configurazioni: a sinistra quattro segnali ai vertici di un quadrato in due dimensioni e a destra otto segnali ai vertici di un cubo tridimensionale.

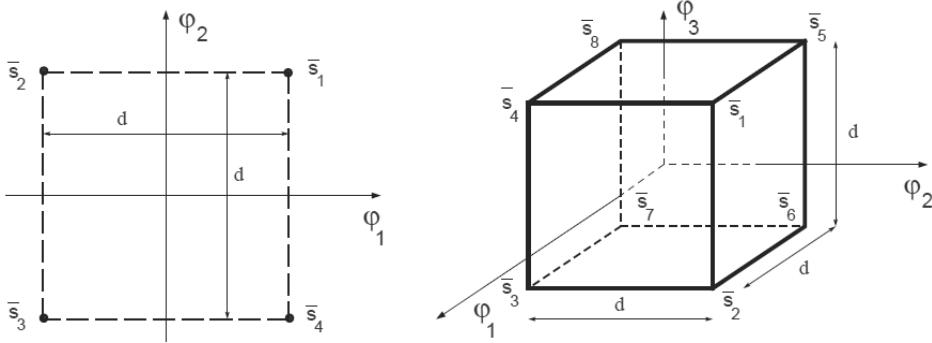


Figura 9.10: Segnali binari ortogonali.

Ciascuna delle N componenti del generico vettore \bar{s}_i , $i = 1, 2, \dots, 2^N$, può valere solo $+d/2$ oppure $-d/2$. Supponiamo che sia stato trasmesso il segnale \bar{s}_1 e che esso abbia tutte le componenti pari a $+d/2$ (come nell'esempio in figura). Chiaramente, essendo tutti i segnali equiprobabili e a pari distanza dall'origine, la regione di decisione I_1 associata al vettore \bar{s}_1 consiste nella regione dello spazio N -dimensionale che ha tutte le coordinate positive (ad esempio il quadrante in alto a destra per lo spazio bidimensionale). La probabilità di decisione corretta se viene trasmesso \bar{s}_1 è quindi pari alla probabilità che il vettore ricevuto \bar{r} sia in I_1 , il che accade se tutte le N componenti del vettore rumore \bar{n} sono maggiori di $-d/2$:

$$P(C/m_1) = P(n_k > -\frac{d}{2}, \forall k \in [1, N]) = \prod_{k=1}^N P\left(n_k > -\frac{d}{2}\right) \quad (9.38)$$

Ma

$$P\left(n_k > -\frac{d}{2}\right) = 1 - P\left(n_k < -\frac{d}{2}\right) = 1 - P\left(n_k > \frac{d}{2}\right) = 1 - \mathbf{Q}\left(\frac{d}{\sqrt{2\eta}}\right) \quad (9.39)$$

in accordo con quanto visto per il caso binario. Quindi

$$P(C/m_1) = \left[1 - \mathbf{Q}\left(\frac{d}{\sqrt{2\eta}}\right)\right]^N \quad (9.40)$$

La stessa probabilità di decisione corretta condizionata vale anche per tutti gli altri segnali, quindi la probabilità di decisione corretta non condizionata è ancora

$$P(C) = \left[1 - \mathbf{Q}\left(\frac{d}{\sqrt{2\eta}}\right)\right]^N \quad (9.41)$$

Questa formula corrisponde al fatto che per ogni segnale occorre prendere N decisioni binarie indipendenti, una per coordinata, e che tutte queste decisioni devono essere corrette perché la decisione sia corretta nel suo insieme.

Considerando ora l'energia dei segnali si nota subito come i segnali (equiprobabili) posti ai vertici di un ipercubo siano in configurazione ottima (il baricentro coincide con l'origine) e che l'energia di ciascuno vale

$$|\bar{s}_i|^2 = E_s = N \frac{d^2}{4} \quad (9.42)$$

Quindi la 9.41 può anche essere scritta come

$$P(C) = \left[1 - \mathbf{Q}\left(\sqrt{\frac{2E_s}{N\eta}}\right)\right]^N \quad (9.43)$$

9.5.3 Segnali ortogonali

I segnali in questa configurazione sono tanti quante le funzioni base utilizzate ($N = M$) e ciascuno dei vettori che li rappresentano ha una sola componente diversa da zero. Inoltre sono tutti equienergetici (oltre che equiprobabili), ovvero possono essere scritti nella forma $s_i(t) = E_s \varphi_i(t)$, $i = 1, \dots, N$. In figura 9.11 un esempio di segnali ortogonali in due e tre dimensioni.

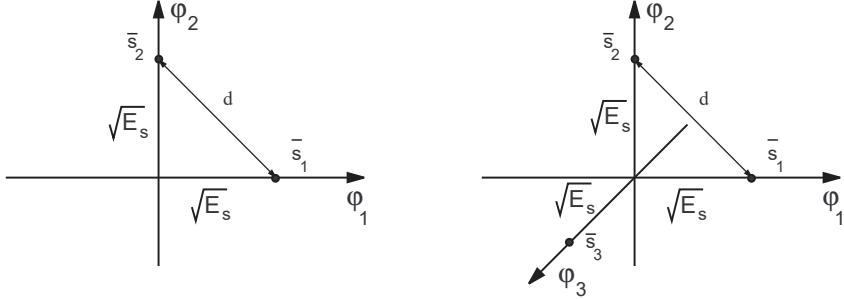


Figura 9.11: Segnali ortogonali in due e tre dimensioni.

Se è stato trasmesso \bar{s}_k esso viene correttamente rilevato se la componente r_k del vettore ricevuto \bar{r} è maggiore di tutte le altre componenti, ovvero

$$P(C/m_k) = P(r_j < r_k \quad \forall j \neq k) \quad (9.44)$$

dove

$$r_k = \sqrt{E_s} + n_k \quad \text{e} \quad r_j = n_j \quad \text{per} \quad j \neq k \quad (9.45)$$

Supponiamo che sia stato ricevuto un particolare vettore $\bar{r} = \bar{R}$, in tal caso

$$P(C/m_k, r_k = R_k) = P(n_j < R_k \quad \forall j \neq k) = [P(n < R_k)]^{N-1} \quad (9.46)$$

avendo indicato con n la generica componente del vettore \bar{n} , visto che sono statisticamente indistinguibili e tutte con densità di probabilità $f_n(n) = 1/\sqrt{\pi\eta} e^{-n^2/\eta}$.

Poiché $f_{r_k}(R_k) = f_n(R_k - \sqrt{E_s})$ e dato che la probabilità di decisione corretta è la stessa per tutti i segnali

$$P(C) = \int_{-\infty}^{\infty} f_n(R_k - \sqrt{E_s}) \left[\int_{-\infty}^{R_k} f_n(\alpha) d\alpha \right]^{N-1} dR_k \quad (9.47)$$

La corrispondente probabilità di errore vale ovviamente $P(E) = 1 - P(C)$, ma entrambe le probabilità non possono in generale essere scritte mediante una formula semplice (tranne che nel caso binario, come visto).

È possibile valutare numericamente la formula 9.47: il risultato in termini di $P(E)$ per diversi valori di M è in figura 9.12. La $P(E)$ è valutata in funzione di $S/(\eta r_b)$, essendo $S = E_s/T$ la potenza media dei segnali e $r_b = \log_2 M/T$ il numero di bit al secondo trasmessi (bit-rate). Il valore di $S/(\eta r_b)$, a livello di rumore fissato, ci dice quanta potenza stiamo usando per trasmettere un bit al secondo, ed è uno dei parametri fondamentali nella progettazione di sistemi di trasmissione discreti. Si può notare che, fissando un valore per $S/(\eta r_b)$, la probabilità di errore può essere resa piccola a piacere aumentando M , purchè $S/(\eta r_b) > \ln 2$. Tale limite è rappresentato dalla linea tratteggiata in figura 9.12. Quindi, se $M \rightarrow \infty$, la trasmissione con segnali ortognali e ricezione ottima tende alla trasmissione senza errori per ogni bit-rate

$$r_b < \frac{S}{\eta \ln 2} = C_\infty \quad (9.48)$$

dove C_∞ è la massima capacità di un canale AWGN (come vedrete in corsi seguenti).

Da notare tuttavia che al crescere di M crescono sia la banda necessaria alla trasmissione che la complessità del ricevitore (che è proporzionale a M). Inoltre la figura 9.12 rivela anche che se fissiamo $M \gg 1$ esiste un effetto soglia, nel senso che la probabilità di errore cresce molto rapidamente se $S/(\eta r_b)$ scende oltre una certo limite.

Come visto, la probabilità di errore non è calcolabile in modo semplice: se ne può tuttavia calcolare facilmente un limite superiore. Supponiamo che sia stato trasmesso il segnale \bar{s}_k : la probabilità di

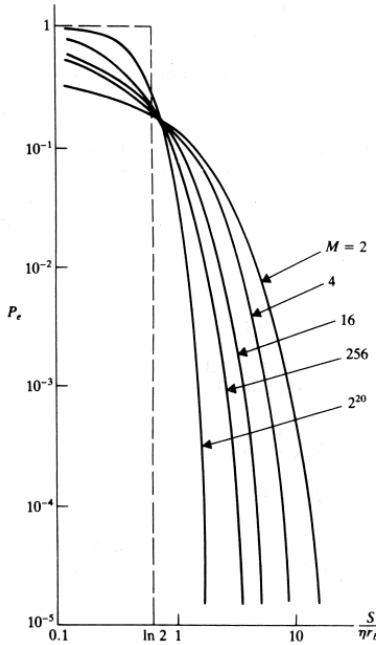


Figura 9.12: Probabilità di errore per M segnali ortogonali (da [Car86] p. 652). In ascissa il rapporto segnale/rumore per bit $S/(\eta r_b)$, con $S = E_s/T$.

errore è la probabilità che il segnale ricevuto \bar{r} sia più vicino ad almeno un altro segnale che a \bar{s}_k . D'altra parte la probabilità che \bar{r} sia più vicino al segnale \bar{s}_j che a \bar{s}_k per un qualsiasi $j \neq k$, ovvero che vi sia un errore che porta da \bar{s}_k a \bar{s}_j , è solita probabilità di errore del caso binario, che dipende solo dalla distanza fra i segnali

$$P_{Err}(\bar{s}_k \rightarrow \bar{s}_j) = Q\left(\frac{|\bar{s}_k - \bar{s}_j|}{\sqrt{2}\eta}\right) \quad (9.49)$$

Se gli eventi $\bar{s}_k \rightarrow \bar{s}_j$ per tutti $j \neq k$ fossero indipendenti la probabilità di errore dato \bar{s}_k sarebbe

$$\sum_{j=1, j \neq k}^M Q\left(\frac{|\bar{s}_k - \bar{s}_j|}{\sqrt{2}\eta}\right)$$

poichè non sono indipendenti vale che

$$P(E/m_k) \leq \sum_{j=1, j \neq k}^M Q\left(\frac{|\bar{s}_k - \bar{s}_j|}{\sqrt{2}\eta}\right) \quad (9.50)$$

Il risultato in 9.50 vale per qualsiasi insieme di segnali, comunque disposti. Nel caso di segnali ortogonali si ha che la distanza fra qualsiasi coppia di segnali vale $d = \sqrt{2E_s}$, qualsiasi sia la dimensione $N \geq 2$ e quindi

$$P(E) \leq (N-1)Q\left(\sqrt{\frac{E_s}{\eta}}\right) \quad (9.51)$$

9.5.4 Segnali “simplex”

Un insieme di segnali ortogonali non è evidentemente a energia minima (come già visto nel caso binario). Per ottenere un insieme a energia minima è necessario sottrarre ad ogni vettore-segnale un vettore

$$\bar{a} = E\{\bar{s}_i\} = \frac{1}{N} \sum_{i=1}^M \bar{s}_i = \frac{\sqrt{E_s}}{N} \sum_{i=1}^N \bar{\phi}_i \quad \text{dove } N = M$$

la cui intensità (modulo) è $\sqrt{E_s/N}$ e la cui direzione è la bisettrice dell'iperangolo formato dai vari assi coordinati. Si può anche rilevare che \bar{a} è ortogonale all'iperpiano passante per gli estremi degli \bar{s}_i e rappresenta la distanza dall'origine di tale iperpiano. La figura 9.13 illustra la situazione nel

caso $N = 3$. I vettori \bar{s}'_i giacciono pertanto tutti nello stesso iperpiano a $N - 1$ dimensioni e perciò possono essere rappresentati con una base $N - 1$ dimensionale. Da notare che la distanza reciproca fra i vari segnali non cambia (e quindi non cambia la probabilità di errore), mentre cambia la loro distanza dall'origine (cioè la loro energia), pur rimanendo equienergetici (in figura 9.13 a destra i segnali sono sui vertici di un triangolo equilatero e il loro baricentro è nell'origine).

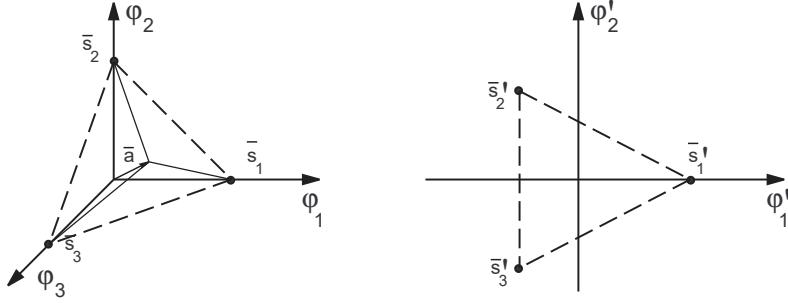


Figura 9.13: Da segnali ortogonali in tre dimensioni (funzioni base) a simplex in due dimensioni.

L'energia del generico segnale simplex \bar{s}'_i vale

$$|\bar{s}'_i|^2 = \bar{s}'_i \cdot \bar{s}'_i = (\bar{s}_i - \bar{a}) \cdot (\bar{s}_i - \bar{a}) = |\bar{s}_i|^2 - 2\bar{a} \cdot \bar{s}_i + |\bar{a}|^2 = E_s - 2\frac{E_s}{N} + \frac{E_s}{N}$$

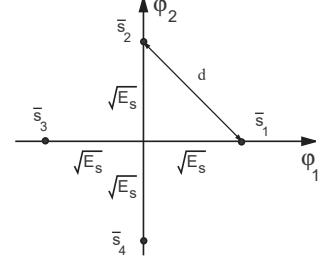
essendo E_s l'energia dei segnali ortogonali e $N = M$ il loro numero. L'energia di ciascun segnale simplex in funzione dell'energia di quelli ortogonali vale quindi

$$E_{s'} = E_s (1 - \frac{1}{M}) \quad (9.52)$$

A parità di probabilità di errore l'energia necessaria è quindi $(1 - 1/M)$ volte inferiore, e cioè la metà nel caso di $M = 2$: per M grande il risparmio di energia tende a diventare trascurabile, il che significa che i segnali ortogonali tendono a diventare ottimi.

9.5.5 Segnali biortogonali

Un insieme di segnali biortogonali si ottiene da un insieme di segnali ortogonali aggiungendo a questi ultimi gli stessi cambiati di segno. Si ha dunque $M = 2N$, come illustrato dall'esempio in figura per $n = 2$. La probabilità di decisione corretta si trova seguendo lo stesso ragionamento che ha portato all'eq. 9.47 per i segnali ortogonali, ovvero



$$\begin{aligned} P(C) &= \int_0^\infty f_n(R_k - \sqrt{E_s}) \left[\int_{-R_k}^{R_k} f_n(\alpha) d\alpha \right]^{N-1} dR_k = \\ &= \int_0^\infty f_n(R_k - \sqrt{E_s}) \left[1 - 2 \int_{R_k}^\infty f_n(\alpha) d\alpha \right]^{\frac{M}{2}-1} dR_k \end{aligned} \quad (9.53)$$

Si può dimostrare che quando M è grande e la probabilità di errore è piccola la prestazione dei segnali biortogonali tende ad essere la stessa di quelli ortogonali, nonostante il dimezzamento del numero di dimensioni richieste.

Anche per i segnali biortogonali si può trovare agevolmente l'espressione di un limite superiore per la probabilità di errore, seguendo la stessa procedura adottata per i segnali ortogonali. Infatti, a partire dalla 9.50, si ricava che

$$P(E) \leq (M - 2) Q \left(\sqrt{\frac{E_s}{\eta}} \right) + Q \left(\sqrt{\frac{2E_s}{\eta}} \right) \quad (9.54)$$

Capitolo 10

PAM con rumore

La PAM è stata introdotta nel capitolo 2: questo capitolo, che assume la conoscenza di quanto già visto, aggiunge il calcolo delle prestazioni in presenza di rumore.

Riassunto delle puntate precedenti

La PAM (*pulse amplitude modulation*) è un sistema di modulazione che utilizza sempre la stessa forma d'onda, ripetuta ogni D secondi, per trasmettere informazione. Tale informazione può essere sia continua che di tipo discreto: la versione più diffusa è la PAM discreta, la sola che verrà trattata in questo capitolo. L'informazione viaggia sull'ampiezza delle forme d'onda. Ciascuna forma d'onda porta una quantità fissa di informazione: se i livelli possibili sono solo due (PAM binaria, la più semplice possibile) viene trasportato un bit di informazione, se sono di più (PAM a M livelli) l'informazione che viene trasportata da ciascuna forma d'onda è pari a $\log_2 M$ bit (per ovvi motivi M è scelto quasi sempre fra le potenze di due).

L'insieme delle funzioni forma il segnale trasmesso, che viene normalmente rappresentato da

$$x(t) = \sum_k a_k p(t - kD) \quad (10.1)$$

dove l'ampiezza modulante a_k rappresenta il k -esimo simbolo nella sequenza di messaggi (a_k può assumere uno fra M valori possibili prefissati). L'indice k varia di solito da $-\infty$ a ∞ , a rappresentare una trasmissione di durata infinita.

La funzione $p(t)$ è l'impulso non modulato, che può assumere diverse differenti forme, ma deve rispettare la condizione

$$p(t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm D, \pm 2D, \dots, \pm kD \end{cases} \quad \forall k > 0 \quad (10.2)$$

Tale condizione ci assicura che possiamo ricostruire il messaggio campionando $x(t)$ periodicamente per $t = KD$ (K intero qualsiasi), visto che, a K fissato,

$$x(KD) = \sum_k a_k p(KD - kD) = a_K \quad (10.3)$$

Da notare che D rappresenta il tempo di ripetizione dell'impulso $p(t)$, non la sua durata, che varia, a seconda dei sistemi, da una frazione di D all'infinito.

Il valore $1/D$ è il numero di impulsi (simboli) al secondo, ovvero la *symbol rate* o *signaling rate*, che si misura in *baud* (cioè simboli al secondo). Nel caso di PAM binaria ogni simbolo porta una quantità di informazione pari a un bit, nel caso di PAM a M livelli i bit a simbolo sono $\log_2 M$: la misura della quantità di informazione trasportata è

$$r = \frac{\log_2 M}{D} \quad (10.4)$$

dove r , detta *bit rate*, si misura in bit al secondo (bit/sec) o *bps*. Nel caso di PAM binaria si usa spesso indicare con $T_b = D$ il tempo di ripetizione degli impulsi e col simbolo $r_b = 1/T_b$ la bit rate. Da ricordare i formati PAM binari e l'analisi compiuta per la trasmissione su canali a banda stretta.

10.1 PAM con rumore

Inizieremo l'analisi dell'effetto del rumore in termini di probabilità di errore dal caso binario e generalizzeremo poi il risultato a M livelli. Per questa trattazione assumeremo un canale privo di distorsioni, in modo che non vi sia ISI. Assumeremo inoltre che il rumore sia additivo, gaussiano, bianco, a media nulla e indipendente dal segnale.

10.1.1 Probabilità di errore per il caso binario

La figura 10.1 mostra il ricevitore PAM binario in banda base. Il segnale $x(t)$ viene trasmesso dal canale senza distorsioni ma con l'aggiunta di un rumore AWGN con densità spettrale bilatera $G_n(f) = \eta/2$. Il filtro passa-basso di ricezione riduce il rumore fuori dalla banda del segnale senza introdurre ISI. Il campionatore campiona ad ogni istante t_k un valore

$$y(t_k) = a_k + N(t_k) \quad (10.5)$$

Gli istanti t_k sono a distanza multipla di D . Il valore campionato viene confrontato con una soglia fissa V : se è maggiore si sceglie 1, sennò 0. Il valore ricevuto \hat{x}_k corrisponde quindi al simbolo trasmesso ad eccezione di occasionali errori.

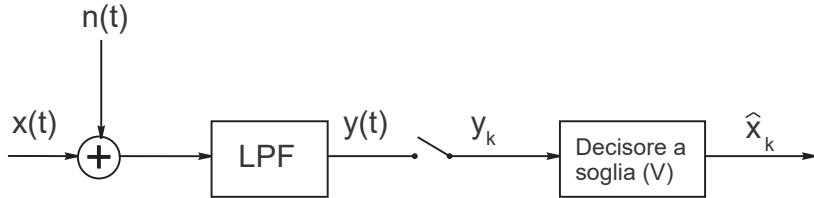


Figura 10.1: Ricevitore PAM binario in banda base.

Per iniziare la nostra analisi supponiamo che $x(t)$ sia un segnale unipolare nel quale $a_k = A$ rappresenta il messaggio (bit) 1 e $a_k = 0$ rappresenta lo zero. Intuitivamente, la soglia dovrà essere scelta fra i due valori trasmessi, $0 < V < A$. Sia y la v.a. che si ottiene campionando $y(t)$: la sua densità di probabilità $f_y(y)$ dipenderà sia da quale simbolo è stato trasmesso (ovvero dalla presenza o assenza dell'impulso PAM) che dalla densità $f_N(N)$ del rumore $N(t)$, essendo $N(t)$ il rumore $n(t)$ filtrato dal filtro di ricezione LPF. Il campione k è quindi

$$y_k = a_k + N_k$$

essendo a_k una v.a. che vale 0 oppure A con probabilità note e N_k una v.a. gaussiana indipendente da a_k , ovvero con statistiche che non dipendono da cosa è stato trasmesso: $f_{N_k}(N_k) = f_N(N)$.

Se è stato trasmesso zero ($a_k = 0$) il campionatore campiona solo il rumore: indicando con H_0 tale ipotesi si ha che la densità condizionata $f_y(y/H_0)$ vale

$$f_y(y/H_0) = f_N(y) \quad (10.6)$$

Se invece è stato trasmesso uno ($a_k = A$), ed indichiamo con H_1 tale evento

$$f_y(y/H_1) = f_N(y - A) \quad (10.7)$$

Il decisore a soglia agisce secondo la seguente strategia:

decide per H_0 ($a_k = 0$) se $y < V$

decide per H_1 ($a_k = A$) se $y > V$

Il caso $y = V$ viene ignorato essendo a probabilità nulla. Le probabilità di errore nei due casi sono allora

$$\begin{aligned} P_{e0} &\equiv P(y > V/H_0) = \int_V^\infty f_y(y/H_0) dy \\ P_{e1} &\equiv P(y < V/H_1) = \int_{-\infty}^V f_y(y/H_1) dy \end{aligned} \quad (10.8)$$

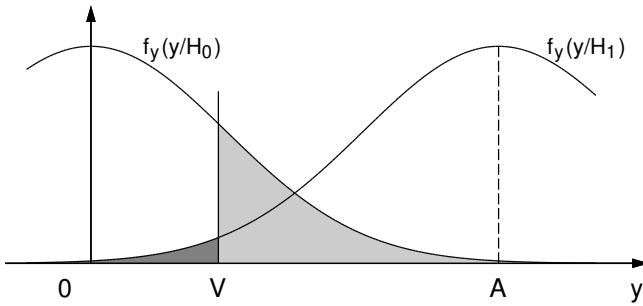


Figura 10.2: Densità condizionate e probabilità di errore a soglia (V) fissata.

Le due probabilità di errore sono indicate in figura 10.2 dalle aree delle code delle rispettive densità condizionate per una generica soglia V : l'area più scura è P_{e1} , quella più chiara P_{e0}

L'interpretazione in termini di aree chiarisce il significato del valore di soglia V a parità degli altri fattori: ridurre V significa ridurre P_{e1} ed aumentare P_{e0} , aumentare V agisce in senso opposto.

Ovviamente la soglia deve essere fissata in modo da rendere minima la probabilità di errore media, cioè

$$P_e = P(H_0) P_{e0} + P(H_1) P_{e1} \quad (10.9)$$

La soglia ottima V_{ott} è quella che minimizza la 10.9 e si può trovare imponendo $dP_e/dV = 0$ ed utilizzando le eq. 10.8:

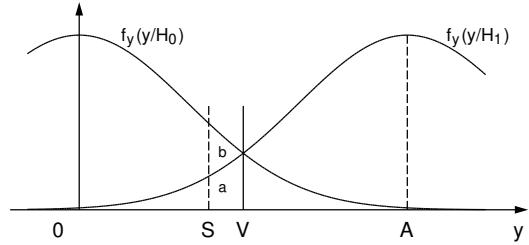
$$P(H_0) f_y(V_{ott}/H_0) = P(H_1) f_y(V_{ott}/H_1) \quad (10.10)$$

Nel caso di simboli equiprobabili, $P(H_0) = P(H_1)$, che è il caso di gran lunga più comune, la 10.10 si riduce a

$$f_y(V_{ott}/H_0) = f_y(V_{ott}/H_1) \quad (10.11)$$

La soglia ottima corrisponde quindi al punto nel quale le due densità condizionate si intersecano, come anche si può intuire osservando la fig. 10.2.

Per chiarire meglio il concetto si osservi la figura a lato: V è la soglia ottima, che corrisponde al punto ove le densità si intersecano. Spostando V in diversa posizione (ad esempio in S) una delle due aree che rappresentano la probabilità di errore diminuisce della parte a mentre l'altra aumenta di $a + b$. Quindi la probabilità di errore media aumenta.



Se il rumore filtrato $N(t)$ è di tipo gaussiano a media nulla e varianza σ_n^2 , ovvero ha densità

$$f_N(n) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{n^2}{2\sigma_n^2}} \quad (10.12)$$

dalle eq.10.8 si ottiene

$$\begin{aligned} P_{e0} &= \int_V^\infty f_N(y) dy = Q\left(\frac{V}{\sigma_n}\right) \\ P_{e1} &= \int_{-\infty}^{-V} f_N(y - A) dy = Q\left(\frac{A - V}{\sigma_n}\right) \end{aligned} \quad (10.13)$$

Poiché $f_N(n)$ ha simmetria pari, le due densità condizionate si intersecano a metà strada, ovvero $V_{ott} = A/2$ (ovviamente sempre nel caso di segnali equiprobabili). Quindi la probabilità di errore diventa

$$P_e = Q\left(\frac{A}{2\sigma_n}\right) \quad (10.14)$$

Il risultato appena raggiunto riguarda segnali unipolari, ma resta valido anche nel caso polare se i livelli della PAM sono $a_k = \pm A/2$ (in realtà rimane valido per qualsiasi configurazione nella quale la differenza delle ampiezze per zero e per uno vale A). Naturalmente la soglia ottima V_{ott} deve essere fissata a metà strada fra i livelli (se i segnali sono equiprobabili), ovvero $V_{ott} = 0$ per segnali polari. Il vantaggio dei segnali polari riguarda la potenza media in trasmissione: per evidenziare questo fatto possiamo esprimere la probabilità di errore in funzione della potenza media in ricezione dei segnali,

ossia S_R . Se assumiamo segnali equiprobabili e impulsi rettangolari di durata D , allora si ha che $S_R = A^2/2$ per segnali unipolari e $S_R = A^2/4$ per segnali polari. Quindi, a parità di S_R ,

$$A = \begin{cases} \sqrt{2S_R} & \text{per segnali unipolari} \\ \sqrt{4S_R} & \text{per segnali polari} \end{cases} \quad (10.15)$$

dove il fattore $\sqrt{2}$ può abbassare notevolmente la probabilità di errore.

Poichè il rumore ha media nulla, la sua varianza σ_n^2 è la potenza media del rumore N all'uscita del filtro di ricezione. Quindi possiamo esprimere il fattore $\frac{A}{2\sigma_n}$ anche in termini di rapporto segnale-rumore in ricezione

$$\left(\frac{A}{2\sigma_n}\right)^2 = \frac{A^2}{4\sigma_n^2} = \begin{cases} \frac{1}{2}(S/N)_R & \text{per segnali unipolari} \\ (S/N)_R & \text{per segnali polari} \end{cases} \quad (10.16)$$

La potenza N del rumore dipende dalla banda passante del filtro di ricezione, che a sua volta è direttamente proporzionale alla frequenza con la quale i segnali sono inviati sul canale. Chiamando $r_b = 1/D$ il numero di bit al secondo inviati mediante PAM binaria (bit-rate) si ha che, per lasciar passare impulsi rettangolari di durata D , la banda richiesta deve essere $B_N \geq r_b/2$, cosicchè la potenza di rumore vale

$$N = \eta B_N \geq \frac{\eta r_b}{2} \quad (10.17)$$

Questo dimostra che per mantenere la stessa probabilità di errore aumentando r_b è necessario aumentare la potenza del segnale.

10.1.2 Probabilità di errore per il caso di PAM a M livelli

La PAM binaria offre la massima resistenza al rumore per ogni $(S/N)_R$, visto che ha solo due livelli (e non si può trasmettere informazione con meno di due livelli!). La PAM a più livelli richiede, a parità di probabilità di errore, più potenza in trasmissione ma necessita di meno banda sul canale, dato che il numero di impulsi al secondo (baud-rate) è inferiore. Di conseguenza la PAM a più livelli è più adatta a trasmettere quando la banda è limitata ma il rapporto segnale/rumore è relativamente grande, come ad esempio sui canali di tipo vocale (telefonici analogici).

Calcoliamo ora la probabilità di errore per una PAM a M livelli con rumore AWGN, supponendo di usare una segnalazione polare con numero di livelli pari ed equispaziati (che è il caso più diffuso). I livelli utilizzati sono

$$a_k = \pm \frac{A}{2}, \pm \frac{3A}{2}, \dots, \pm \frac{(M-1)A}{2} \quad (10.18)$$

Assumiamo inoltre che i segnali siano equiprobabili, ciascuno con probabilità $1/M$: la probabilità di errore vale quindi

$$P_e = \frac{1}{M} (P_{e0} + P_{e1} + \dots + P_{eM}) \quad (10.19)$$

La figura 10.3 illustra il caso di PAM polare quaternaria ($M = 4$) con rumore gaussiano. La regola di decisione richiede tre livelli di soglia, indicati in figura a $y = -A$, $y = 0$ e $y = A$. Queste sono le soglie che minimizzano la probabilità di errore P_e , ma la probabilità di errore non è la stessa per tutti i simboli. Per i due livelli estremi, pari a $\pm 3A/2$ otteniamo lo stesso valore già calcolato nel caso binario, ovvero

$$P_{e0} = P_{e3} = Q\left(\frac{A}{2\sigma_n}\right) \quad (10.20)$$

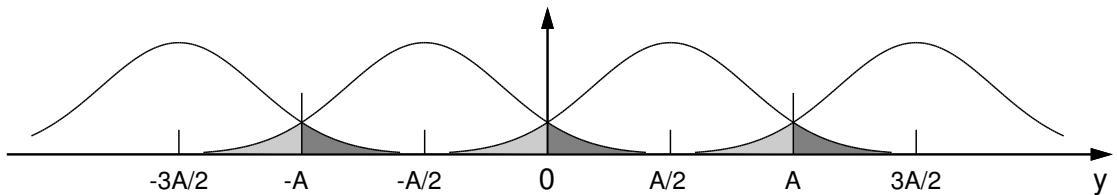


Figura 10.3: Densità condizionate e probabilità di errore con rumore gaussiano per PAM quaternaria.

Per i livelli interni, in generale in numero pari a $M - 2$, la probabilità di errore è doppia, visto che si può sbagliare in due direzioni, anzichè in una sola (ciascuna probabilità di errore si ottiene sommando due code di gaussiana, come mostrato in figura). Quindi

$$P_{e1} = P_{e2} = 2 \mathbf{Q} \left(\frac{A}{2\sigma_n} \right) \quad (10.21)$$

La probabilità di errore media per PAM quaternaria è allora

$$P_e = \frac{1}{4} \left[\mathbf{Q} \left(\frac{A}{2\sigma_n} \right) + \mathbf{Q} \left(\frac{A}{2\sigma_n} \right) + 2 \mathbf{Q} \left(\frac{A}{2\sigma_n} \right) + 2 \mathbf{Q} \left(\frac{A}{2\sigma_n} \right) \right] = \frac{3}{2} \mathbf{Q} \left(\frac{A}{2\sigma_n} \right) \quad (10.22)$$

ovvero maggiore del 50% rispetto al caso binario con la stessa spaziatura fra i livelli.

Per una PAM a M livelli polari spaziati come in 10.18, analoghi ragionamenti portano a

$$P_e = 2 \left(1 - \frac{1}{M} \right) \mathbf{Q} \left(\frac{A}{2\sigma_n} \right) \quad (10.23)$$

che si riduce ovviamente alle formula già viste per $M = 2$ e $M = 4$.

Per $M \gg 2$ la probabilità di errore media tende quindi al doppio di quella ottenibile nel caso binario con la stessa spaziatura fra i livelli.

Appendice A

Formule utili

Numeri complessi e fasori

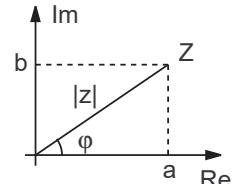
Un numero complesso z può essere rappresentato in due modi:

$$\begin{aligned} z &= a + jb && a \text{ parte reale, } b \text{ parte immaginaria} \\ z &= |z|e^{j\varphi} && |z| \text{ modulo, } \varphi \text{ fase.} \end{aligned}$$

Le relazioni che legano le due rappresentazioni sono:

$$\begin{aligned} |z| &= \sqrt{a^2 + b^2} & \varphi &= \arctan\left(\frac{b}{a}\right) \\ a &= |z| \cos \varphi & b &= |z| \sin \varphi \end{aligned}$$

Si definisce z complesso coniugato il numero $z^* = a - jb = |z|e^{-j\varphi}$. Ovviamente $z \cdot z^* = |z|^2$.



Formule di Eulero

$$e^{\pm j\varphi} = \cos \varphi \pm j \sin \varphi \quad \cos \varphi = \frac{e^{j\varphi} + e^{-j\varphi}}{2} \quad \sin \varphi = \frac{e^{j\varphi} - e^{-j\varphi}}{2j} \quad (\text{A.1})$$

Ovvietà

$$\log_a b = c \iff a^c = b$$

Formule trigonometriche

$$\begin{aligned} \sin^2 \alpha + \cos^2 \alpha &= 1 \\ \sin(\alpha \pm \beta) &= \sin(\alpha) \cos(\beta) \pm \cos(\alpha) \sin(\beta) \\ \cos(\alpha \pm \beta) &= \cos(\alpha) \cos(\beta) \mp \sin(\alpha) \sin(\beta) \\ \tan(\alpha \pm \beta) &= \frac{\tan(\alpha) \pm \tan(\beta)}{1 \mp \tan(\alpha) \tan(\beta)} \\ \sin(2\alpha) &= 2 \sin(\alpha) \cos(\alpha) \\ \cos(2\alpha) &= \cos^2 \alpha - \sin^2 \alpha = 1 - 2 \sin^2 \alpha = 2 \cos^2 \alpha - 1 \\ \tan(2\alpha) &= \frac{2 \tan(\alpha)}{1 - \tan^2(\alpha)} \\ \sin(\alpha) + \sin(\beta) &= 2 \sin\left(\frac{\alpha + \beta}{2}\right) \cos\left(\frac{\alpha - \beta}{2}\right) \\ \sin(\alpha) - \sin(\beta) &= 2 \cos\left(\frac{\alpha + \beta}{2}\right) \sin\left(\frac{\alpha - \beta}{2}\right) \\ \cos(\alpha) + \cos(\beta) &= 2 \cos\left(\frac{\alpha + \beta}{2}\right) \cos\left(\frac{\alpha - \beta}{2}\right) \\ \cos(\alpha) - \cos(\beta) &= -2 \sin\left(\frac{\alpha + \beta}{2}\right) \sin\left(\frac{\alpha - \beta}{2}\right) \\ \sin(\alpha) \sin(\beta) &= \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)] \end{aligned}$$

$$\begin{aligned}
 \cos(\alpha) \cos(\beta) &= \frac{1}{2} [\cos(\alpha + \beta) + \cos(\alpha - \beta)] \\
 \sin(\alpha) \cos(\beta) &= \frac{1}{2} [\sin(\alpha + \beta) + \sin(\alpha - \beta)] \\
 \cos^2(\alpha) &= \frac{1}{2} [1 + \cos(2\alpha)] \\
 \sin^2(\alpha) &= \frac{1}{2} [1 - \cos(2\alpha)] \\
 \sin(\alpha) \cos(\alpha) &= \frac{1}{2} \sin(2\alpha)
 \end{aligned}$$

Best fit lineare

Dati m punti $x_i, y_i, i = 1, \dots, m$ trovare la miglior retta

$$F(x) = a + bx$$

nel senso dell'errore quadratico medio. Si deve minimizzare

$$\rho(a, b) = \sum_{i=1}^m r_i^2 = \sum_{i=1}^m [y_i - (bx_i + a)]^2 = \sum_{i=1}^m y_i^2 + (bx_i + a)^2 - 2(bx_i + a)y_i$$

L'errore ρ è minimo per i valori a e b tali che

$$\begin{aligned}
 \frac{\delta \rho}{\delta b} &= 0 & \frac{\delta \rho}{\delta a} &= 0 \\
 \frac{\delta \rho}{\delta b} &= \sum_{i=1}^m [2bx_i^2 + 2ax_i - 2x_i y_i] = 0 \\
 \frac{\delta \rho}{\delta a} &= \sum_{i=1}^m [2a + 2x_i b - 2y_i] = 0
 \end{aligned}$$

ponendo

$$S_{xx} = \sum_{i=1}^m x_i^2 \quad S_x = \sum_{i=1}^m x_i \quad S_{xy} = \sum_{i=1}^m x_i y_i \quad S_y = \sum_{i=1}^m y_i$$

si ha

$$S_{xx}b + S_xa = S_{xy} \quad S_xb + ma = S_y$$

e quindi

$$b = \frac{S_x S_y - m S_{xy}}{S_x^2 - m S_{xx}} \quad a = \frac{S_x S_{xy} - S_{xx} S_y}{S_x^2 - m S_{xx}}$$

Simboli in Teoria degli Insiemi

\emptyset	Insieme vuoto
N	Insieme dei numeri naturali: $N = \{0, 1, 2, 3, \dots\}$
Z	Insieme dei numeri interi: $Z = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$
Q	Insieme dei numeri razionali (rapporto fra due numeri interi)
R	Insieme dei numeri reali
C	Insieme dei numeri complessi
\cup	Unione (anche indicata con $+$, $A \cup B = A + B$)
\cap	Intersezione (anche indicata con il prodotto, $A \cap B = AB$)
\in	Appartiene (specifica che un elemento appartiene ad un insieme)
\notin	Non appartiene (specifica che un elemento non appartiene ad un insieme)
\subset	Sottoinsieme proprio
\subseteq	Sottoinsieme (eventualmente i due insiemi possono coincidere)
\overline{A}	\overline{A} è l'insieme complementare dell'insieme A

Appendice B

Nomi illustri

Jean Baptiste Joseph Fourier: (Auxerre, 21 marzo 1768 - Parigi, 16 maggio 1830) matematico e fisico francese.

La sua istruzione si compì dapprima dai Benedettini, poi in una scuola militare. Partecipò alla Rivoluzione francese, rischiando di essere ghigliottinato durante il Terrore, salvato dalla caduta di Robespierre. Entrò quindi nella École Normale Supérieure, dove ebbe come professori, tra gli altri, Joseph-Louis Lagrange e Pierre-Simon Laplace. Succedette anche a quest'ultimo nel ruolo di professore alla École Polytechnique nel 1797.

Fourier partecipò alla campagna d'Egitto di Napoleone nel 1798, e ricoprì un importante ruolo di diplomatico in quel paese. Al suo ritorno in Francia nel 1801, fu nominato da Napoleone prefetto dell'Isère. Fu quindi a Grenoble che condusse i suoi esperimenti sulla propagazione del calore che gli consentirono di modellizzare l'evoluzione della temperatura per mezzo di serie trigonometriche. Questi lavori furono pubblicati nel 1822 in Teoria analitica del calore, ma furono molto contestati, specialmente da Pierre-Simon Laplace e Joseph-Louis Lagrange. Entrò all'Accademia delle Scienze nel 1817.

Tra i suoi maggiori contributi figurano la teorizzazione della serie di Fourier e la conseguente trasformata di Fourier in matematica e la formulazione dell'equazione generale della conduzione termica, denominata legge di Fourier, in termodinamica. Il cratere Fourier sulla Luna è intitolato a lui.

[da Wikipedia]



Heinrich Rudolf Hertz: (Amburgo, 22 febbraio 1857 - Bonn, 1 gennaio 1894) fisico tedesco.

Per primo dimostrò l'esistenza delle onde elettromagnetiche con un apparato di sua costruzione, il dipolo hertziano, in grado di emettere onde radio. In suo onore, nel sistema internazionale la frequenza è misurata in hertz.

Nel 1885 ricevette la cattedra all'università di Karlsruhe e nello stesso periodo scoprì le onde elettromagnetiche.

In seguito ad un primo esperimento eseguito da Michelson nel 1881 (anticipatore del più celebre esperimento di Michelson-Morley del 1887) che escludeva l'esistenza dell'etere, egli riformulò le equazioni di Maxwell per tenere conto della novità.

Con un esperimento egli dimostrò che dei segnali elettrici potevano essere inviati attraverso l'aria, come già predetto da James Clerk Maxwell e Michael Faraday e pose le basi per l'invenzione della radio. Egli scoprì inoltre l'effetto fotoelettrico (la cui spiegazione teorica fu successivamente elaborata da Albert Einstein) osservando che oggetti elettricamente carichi perdevano la carica se esposti alla luce ultravioletta. Morì per avvelenamento del sangue all'età di 37 anni a Bonn, in Germania.

Suo nipote Gustav Ludwig Hertz vinse il premio Nobel per la fisica nel 1925 e suo figlio, Carl Hellmuth Hertz inventò l'ecografia medica.

[da Wikipedia]



Harry Nyquist: (Nilsby, 7 febbraio 1889 - Harlingen, 4 aprile 1976) fisico statunitense.

Ricercatore presso i Bell Telephone Laboratories, diede un grande contributo alla tecnologia delle telecomunicazioni e dei controlli automatici.

Suoi sono il teorema del campionamento, che definisce la frequenza di campionamento necessaria per riprodurre un segnale analogico campionato, ed il Criterio di Nyquist, che permette di determinare la stabilità di un sistema retroazionato dall'analisi in frequenza dello stesso sistema a ciclo aperto o attraverso il diagramma di Nyquist.

[da Wikipedia]



Claude Elwood Shannon: (Petoskey, 30 aprile 1916 - Medford, 24 febbraio 2001) ingegnere e matematico USA, definito "il padre della teoria dell'informazione". Con la tesi che presentò al master del 1938, "Un'analisi simbolica dei relè e dei circuiti", Shannon dimostrò, utilizzando un circuito elettrico dotato di un interruttore, che il fluire di un segnale elettrico attraverso una rete di interruttori - che possono essere "accesi/spentti" - segue esattamente le regole dell'algebra di Boole. Shannon pose così la base teorica dei sistemi di codificazione e trasmissione digitale dell'informazione.

Dopo aver lavorato sulla genetica nel laboratorio di Cold Spring Harbor, Shannon si dedicò al suo dottorato di ricerca (PhD) al MIT, dove collaborò alla costruzione dell'analizzatore differenziale di Vannevar Bush, un calcolatore analogico. La sua tesi di dottorato, risalente al 1940, è intitolata "Un'algebra per la genetica teorica". L'anno dopo cominciò a lavorare ai laboratori Bell, alternandosi con l'insegnamento al MIT a partire dagli anni Cinquanta.



Nei laboratori della Bell, dove rimase fino al 1972, Claude Shannon si occupò di problemi d'ingegneria a livello matematico. Partì dal collegamento tra entropia e disordine e dalla considerazione che la specificazione degli aspetti d'ordine di un sistema richiede informazione. Le sue ricerche lo portarono alla conclusione che l'informazione poteva essere definita nei termini del logaritmo della probabilità, cambiato di segno.

Durante la seconda guerra mondiale il Pentagono chiese a Shannon di realizzare delle ricerche sulla possibilità di "guidare" i missili. Nel 1948 Shannon pubblicò in due parti il saggio Una teoria matematica della comunicazione. In questo lavoro si concentrava sul problema di ricostruire con un certo grado di certezza le informazioni trasmesse da un mittente. Fu in questa ricerca che Shannon coniò la parola **bit**, per designare l'unità elementare di informazione. La sua teoria dell'informazione pose le basi per progettare sistemi informatici partendo dal presupposto che l'importante era cercare di memorizzare delle informazioni in modo da poterle anche trasferire e collegare tra loro.

Nel 1949 Shannon pubblicò un altro notevole articolo, La teoria della comunicazione nei sistemi crittografici, con cui praticamente fondò la teoria matematica della crittografia. È inoltre riconosciuto come il "padre" del *teorema del campionamento*.

Nel 1956 fu eletto membro dell'Accademia Nazionale delle Scienze. Dal 1958 al 1978 fu professore al MIT.

[da Wikipedia]

Andrej Nikolaevi Kolmogorov: (Tambov, 25 aprile 1903 Mosca, 20 ottobre 1987) matematico sovietico.

Tra i più importanti e influenti matematici del XX secolo, compì importanti progressi in diversi campi accademici, tra cui la teoria delle probabilità, la topologia, la logica intuizionista, la turbolenza, la meccanica classica e la complessità computazionale. A dispetto della considerevole importanza della sua Scuola matematica per lo sforzo bellico durante la seconda guerra mondiale, fu uno dei matematici sovietici esclusi dalla ricerca scientifica in ambito militare, a causa della sua convivenza, a partire dal 1929, col compagno, matematico anch'esso, Pavel Aleksandrov. Si devono a lui l'introduzione della definizione di insieme limitato e gli assiomi del calcolo probabilistico.

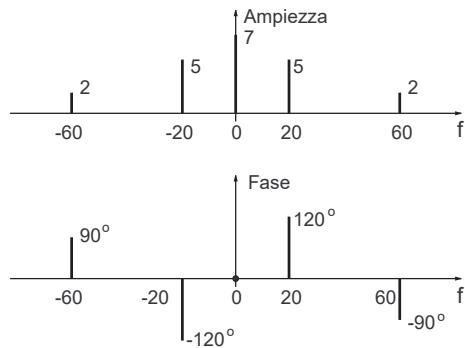


[da Wikipedia]

Appendice C

Soluzione agli esercizi proposti

C.1 Soluzione dell'esercizio 1.1



C.2 Soluzione dell'esercizio 1.5

Le trasformate di Fourier dei due filtri sono:

$$\prod \left(\frac{f}{f_0} \right) \longleftrightarrow f_0 \operatorname{sinc}(f_0 t)$$

$$\prod \left(\frac{f}{f_0/2} \right) \longleftrightarrow \frac{f_0}{2} \operatorname{sinc} \left(\frac{f_0}{2} t \right)$$

L'area di entrambe, dalla 1.71, è pari a uno. L'area deve essere unitaria per non amplificare o attenuare il segnale: supponendo il segnale costante solo se l'area è unitaria l'uscita della convoluzione è ancora la stessa costante.

C.3 Soluzione dell'esercizio 7.7

Vale che

$$E\{a_k\} = 0 \quad E\{a_k^2\} = V^2 \quad E\{a_j a_k\} = 0 \quad j \neq k$$

Quindi

$$R_x(\tau) = E\{a_j a_k\} P(A) + E\{a_k^2\}[1 - P(A)] = V^2 \left[1 - \frac{|\tau|}{T}\right]$$

per $|\tau| \leq T$ e zero altrove.

C.4 Soluzione dell'esercizio 7.8

Vale che

$$\begin{aligned} E\{a_k\} &= \frac{1}{2}0 + \frac{1}{2}A = \frac{A}{2} & E\{a_k^2\} &= \frac{1}{2}0 + \frac{1}{2}A^2 = \frac{A^2}{2} \\ E\{a_j a_k\} &= E\{a_j\} E\{a_k\} = \frac{A^2}{4} & j \neq k \end{aligned}$$

Quindi

$$R_x(\tau) = E\{a_j a_k\} P(A) + E\{a_k^2\}[1 - P(A)] = \frac{A^2}{4}P(A) + \frac{A^2}{2}[1 - P(A)]$$

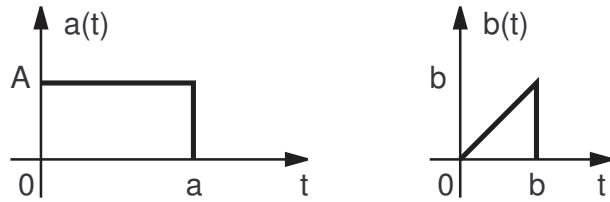
ovvero

$$R_x(\tau) = \frac{A^2}{4} \frac{|\tau|}{T} + \frac{A^2}{2} \left[1 - \frac{|\tau|}{T}\right]$$

per $|\tau| \leq T$ e $A^2/4$ altrove.

C.5 Soluzione dell'esercizio su convoluzione

Siano $a(t)$ e $b(t)$ due segnali a supporto limitato come in figura ($a \geq b$).



$$a(t) = A[U(t) - U(t-a)]$$

$$b(t) = t[U(t) - U(t-b)]$$

La loro convoluzione, calcolata graficamente, vale:

$$a(t) * b(t) = \begin{cases} 0 & t < 0 \\ \frac{At^2}{2} & 0 \leq t \leq b \\ \frac{Ab^2}{2} - \frac{A(t-a)^2}{2} & b < t \leq a \\ 0 & t > a+b \end{cases} \quad (\text{C.1})$$

Il calcolo mediante integrali è:

$$\begin{aligned} a(t) * b(t) &= \int_{-\infty}^{\infty} a(\tau)b(t-\tau) d\tau = \int_{-\infty}^{\infty} a(t-\tau)b(\tau) d\tau = \\ &= \int_{-\infty}^{\infty} A[U(t-\tau) - U(t-\tau-a)] \tau [U(\tau) - U(\tau-b)] d\tau = \\ &= \int_{-\infty}^{\infty} A \tau [U(t-\tau)U(\tau) - U(t-\tau)U(\tau-b) - U(t-\tau-a)U(\tau) + U(t-\tau-a)U(\tau-b)] d\tau = \end{aligned}$$

spezzando gli integrali e risolvendoli separatamente si ottiene

$$\begin{aligned} \int_{-\infty}^{\infty} A \tau U(t-\tau)U(\tau) d\tau &= A \int_0^t \tau d\tau = \frac{At^2}{2} \quad t \geq 0 \\ \int_{-\infty}^{\infty} A \tau U(t-\tau)U(\tau-b) d\tau &= A \int_b^t \tau d\tau = \frac{At^2}{2} - \frac{Ab^2}{2} \quad t \geq b \\ \int_{-\infty}^{\infty} A \tau U(t-\tau-a)U(\tau) d\tau &= A \int_0^{t-a} \tau d\tau = \frac{A(t-a)^2}{2} \quad t \geq a \\ \int_{-\infty}^{\infty} A \tau U(t-\tau-a)U(\tau-b) d\tau &= A \int_b^{t-a} \tau d\tau = \frac{A(t-a)^2}{2} - \frac{Ab^2}{2} \quad t \geq a+b \end{aligned}$$

Sommandoli con i giusti segni nei giusti intervalli si ottiene il risultato già visto.

C.6 Ancora Lotto

Si supponga di scegliere N numeri fra 1 e 90, tutti diversi fra loro. Indichiamo tali numeri come n_i , $i = 1, \dots, N$. Sia $\{E_i\}$ per $i = 1, \dots, N$ l'evento "viene estratto il numero i-esimo" (cioè n_i).

La probabilità che un certo numero venga estratto è, come visto, $P(E_i) = 1/18 \forall i$.

La probabilità che vengano estratti (entrambi) i due numeri n_i e n_k (definiti a priori) vale $P(E_i E_k)$, $i \neq k$. Gli eventi E_i ed E_k non sono indipendenti, infatti $P(E_i/E_k) = 4/89 \neq P(E_i)$. Siccome

$$P(E_i/E_k) = 4/89 = \frac{P(E_i E_k)}{P(E_k)} = \frac{P(E_i E_k)}{1/18}$$

si ricava che

$$P(E_i E_k) = (4/89) \cdot (1/18) = 1/400.5$$

che è la probabilità di indovinare un ambo, già calcolata con diverso metodo.

La probabilità che venga estratto almeno uno fra i due numeri n_i e n_k , $i \neq k$, si ottiene da

$$P(E_i + E_k) = P(E_i) + P(E_k) - P(E_i E_k) = \frac{1}{18} + \frac{1}{18} - \frac{1}{400.5} \approx 0.1086$$

Consideriamo ora tre numeri, n_i , n_k e n_j con $i \neq k \neq j$. I tre eventi E_i , E_k ed E_j non sono indipendenti, infatti, usando il risultato precedente,

$$P(E_i/E_k E_j) = 3/88 = \frac{P(E_i E_k E_j)}{P(E_k E_j)} = \frac{P(E_i E_k E_j)}{1/400.5}$$

quindi

$$P(E_i E_k E_j) = (3/88) \cdot (1/400.5) = 1/11748$$

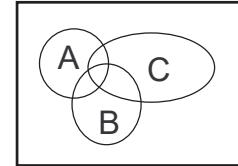
che è la probabilità di indovinare un terno, già calcolata con diverso metodo.

La probabilità che venga estratto almeno uno fra i tre numeri n_i , n_k e n_j , $i \neq k \neq j$, si ottiene da

$$P(E_i + E_k + E_j) = P(E_i) + P(E_k) + P(E_j) - P(E_i E_k) - P(E_i E_j) - P(E_k E_j) + P(E_i E_k E_j)$$

Infatti, con riferimento alla figura a fianco,

$$\begin{aligned} P(A + B + C) &= P(A) + P(B) + P(C) + \\ &\quad - P(AB) - P(AC) - P(CB) + \\ &\quad + P(ABC) \end{aligned}$$



Quindi

$$P(E_i + E_k + E_j) = \frac{3}{18} - \frac{3}{400.5} + \frac{1}{11748} \approx 0.1593$$

Considerando quattro numeri, in notazione semplificata, si ottiene

$$P(E_4/E_1 E_2 E_3) = 2/87 = \frac{P(E_1 E_2 E_3 E_4)}{P(E_1 E_2 E_3)} = \frac{P(E_1 E_2 E_3 E_4)}{1/11748}$$

e quindi

$$P(E_1 E_2 E_3 E_4) = \frac{1}{86} \cdot \frac{1}{11748} = \frac{1}{511038}$$

che è la probabilità di indovinare una quaterna, già calcolata con diverso metodo.

Con cinque numeri

$$P(E_5/E_1 E_2 E_3 E_4) = 1/86 = \frac{P(E_1 E_2 E_3 E_4 E_5)}{P(E_1 E_2 E_3 E_4)} = \frac{P(E_1 E_2 E_3 E_4 E_5)}{1/11748}$$

e quindi

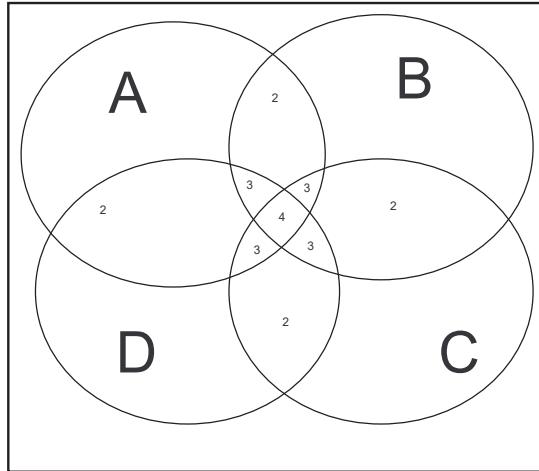
$$P(E_1 E_2 E_3 E_4 E_5) = \frac{1}{86} \cdot \frac{1}{511038} = \frac{1}{43949268}$$

che è la probabilità di indovinare una quintina, già calcolata con diverso metodo.

Ovviamente la probabilità che vengano estratti sei o più numeri insieme è pari a zero, visto che se ne estraggono cinque per volta. La probabilità che almeno uno di N numeri venga estratto diventa

sempre più complicata, ma la strada da seguire è quella indicata. Per quattro eventi vale infatti (se non mi sono sbagliato)

$$\begin{aligned} P(A + B + C + D) &= P(A) + P(B) + P(C) + P(D) + \\ &\quad - P(AB) - P(AC) - P(AD) - P(BC) - P(BD) - P(CD) + \\ &\quad + P(ABC) + P(ABD) + P(ACD) + P(BCD) + \\ &\quad - P(ABCD) \end{aligned}$$



In figura i numeri rappresentano quante volte sono contate le aree da $P(A) + P(B) + P(C) + P(D)$. Sostituendo i valori già calcolati

$$P(E_1 + E_2 + E_3 + E_4) = \frac{4}{18} - \frac{6}{400.5} + \frac{4}{11748} - \frac{1}{511038} \approx 0.20758$$

che è la probabilità che almeno uno fra quattro numeri venga estratto.

A lezione è stato posto il problema di trovare la probabilità che almeno uno fra dieci numeri venga estratto. Benché sia possibile seguire la strada indicata, le relazioni da scrivere diventano più complesse della pazienza di cui dispongo. Tuttavia è possibile calcolare la probabilità di almeno un estratto su 10 numeri in modo diverso, ovvero programmando un elaboratore elettronico. Il numero di possibili uscite dell'esperimento "estrazione del lotto" è infatti finito, e quindi è possibile enumerarle e contare i risultati favorevoli. Il numero di risultati favorevoli diviso quelli possibili è la probabilità cercata (si usa qui il fatto che gli eventi elementari sono equiprobabili).

Valutiamo quanti sono i possibili risultati: il primo estratto può essere uno fra 90 numeri, per ciascun primo estratto il secondo estratto può essere uno fra 89 numeri e così via. Il numero N_T di possibili uscite (la dimensione dello spazio) è allora

$$N_T = 90 \cdot 89 \cdot 88 \cdot 87 \cdot 86 = 5\,273\,912\,160$$

Tanto per verificare che il risultato sia giusto, vediamo quante diverse quinte possono uscire. Cinque numeri possono essere posti in cinque spazi in $5! = 120$ modi diversi: il numero di quinte possibili è allora $5273912160/120 = 43949268$, in accordo con quanto già calcolato.

Tutte le possibili uscite si possono ottenere mediante programma innestando cinque cicli e controllando, in quello più interno, se si verifica la condizione cercata. Da notare che il numero di uscite è *maggiore* del massimo numero rappresentabile da un intero a 32 bit, così come il numero di casi favorevoli (che però si può rappresentare con un intero a 32 bit senza segno). Nel programma ho utilizzato una semplice classe per gestire numeri interi maggiori di 32 bit, ma ho poi verificato che i risultati si possono anche accumulare in numeri a virgola mobile in doppia precisione (64 bit) senza errori e senza che si verifichi perdita di precisione (cosa tutt'altro che ovvia...). La funzione che esegue il calcolo con accumulatori *double* in linguaggio **C** (o **C++**) è riportata nel seguito.

Il risultato è: numero di casi favorevoli 2 389 110 240, per una probabilità

$$P\{\text{Almeno un estratto su dieci numeri}\} = \frac{2\,389\,110\,240}{5\,273\,912\,160} \approx 0.4530$$

Su di un vecchio Athlon 1900 ed usando l'accumulo su classe (più lento) il tempo di calcolo è di circa 150 secondi.

```

void Lotto_A()
{
    int e1, e2, e3, e4, e5, i;
    int u[5];
    double dT = 0.0, dH = 0.0;

    for(e1=1; e1 <= 90; e1++) {
        u[0] = e1;
        for(e2=1; e2 <= 90; e2++) {
            u[1] = e2; if( e2 == e1 ) continue;
            for(e3=1; e3 <= 90; e3++) {
                u[2] = e3; if( e3 == e1 || e3 == e2 ) continue;
                for(e4=1; e4 <= 90; e4++) {
                    u[3] = e4; if( e4 == e1 || e4 == e2 || e4 == e3 ) continue;
                    for(e5=1; e5 <= 90; e5++) {
                        u[4] = e5;
                        if( e5 == e1 || e5 == e2 || e5 == e3 || e5 == e4 ) continue;
                        dT += 1.0;
                        bool hit = false;
                        for(i=0; i < 5; i++) {
                            if( u[i] >= 1 && u[i] <= 10 ) hit = true;           // 10 numeri fra 1 e 10
                        }
                        if( hit ) dH += 1.0;
                    }
                }
            }
        }
    }
    printf("Prob. di indovinare almeno 1 di 10 numeri al lotto %f",dH/dT);
}

```

A lezione, quando ho posto il problema, avevo stimato una probabilità di circa il 45%: come ho fatto? La probabilità che uno dei dieci numeri NON venga estratto è evidentemente $P(\bar{E}_1) = 17/18$. Se gli eventi fossero indipendenti, la probabilità che nessuno venga estratto sarebbe $(17/18)^{10}$ e la probabilità di almeno un estratto sarebbe

$$P(\cdot) = 1 - \left(\frac{17}{18}\right)^{10} \approx 0.43537$$

Questo risultato è sbagliato, visto che gli eventi non sono indipendenti. Sono però "poco" dipendenti: per una verifica empirica proviamo a calcolare la probabilità che almeno uno fra 90 numeri venga estratto. Essa deve ovviamente essere uno, ma seguendo il metodo errato viene

$$1 - \left(\frac{17}{18}\right)^{90} \approx 0.9942$$

che è sbagliato, ma non molto...

Approfondiamo: la probabilità che un numero non venga estratto è $85/90 = 17/18$.

$$P(\bar{E}_2/\bar{E}_1) = \frac{84}{89} = \frac{P(\bar{E}_1\bar{E}_2)}{P(\bar{E}_1)}$$

$$P(\bar{E}_1\bar{E}_2) = \frac{84}{89} \cdot \frac{17}{18} \approx 0.89139 \neq P(\bar{E}_1)P(\bar{E}_2) = \frac{17}{18} \cdot \frac{17}{18} \approx 0.89198$$

se fossero uguali gli eventi sarebbero indipendenti, il fatto che siano simili ci indica che considerarli indipendenti è sbagliato, ma l'errore che si commette non è molto alto...

Una soluzione in forma chiusa esiste (come trovato dallo studente Mattia Desana): il numero di uscite possibili è $\binom{90}{5}$, il numero di uscite sfavorevoli è $\binom{80}{5}$, quindi

$$P\{\text{Almeno un estratto su 10 n.}\} = 1 - \frac{\binom{80}{5}}{\binom{90}{5}} = 1 - \frac{80!}{5!75!} \frac{5!85!}{90!} = \frac{2389110240}{5273912160} \approx 0.4530$$

Una diversa soluzione (da Gabriele Zereik, ma...). Chiamiamo E_1 la probabilità che il primo estratto sia uno dei 10 numeri scelti e, ovviamente, \bar{E}_1 la probabilità che non lo sia (ed analogamente per le posizioni fino a 5).

la probabilità che il primo estratto sia uno dei 10 numeri è $P(E_1) = 10/90$. La probabilità che il secondo estratto sia uno dei 10 numeri vale

$$P(E_2) = \frac{10}{89} \cdot \frac{80}{90} + \frac{9}{89} \cdot \frac{10}{90}$$

che si ottiene dal teorema della probabilità totale

$$P(E_2) = P(E_2/\bar{E}_1) P(\bar{E}_1) + P(E_2/E_1) P(E_1)$$

La probabilità dell'evento unione (ovvero che *almeno* uno dei dieci numeri compaia nelle prime due estrazioni vale

$$P(E_1 + E_2) = P(E_1) + P(E_2) - P(E_1 E_2)$$

dove

$$P(E_1 E_2) = P(E_2 E_1) = P(E_2/E_1) P(E_1) = \frac{9}{89} \cdot \frac{10}{90}$$

Quindi

$$P(E_1 + E_2) = \frac{10}{90} + \frac{10}{89} \cdot \frac{80}{90} + \frac{9}{89} \cdot \frac{10}{90} - \frac{9}{89} \cdot \frac{10}{90} = \frac{10}{90} + \frac{10}{89} \cdot \frac{80}{90}$$

Iterando il procedimento si arriva a

$$P(E_1 + E_2 + E_3 + E_4 + E_5) = \frac{10}{90} + \frac{10}{89} \cdot \frac{80}{90} + \frac{10}{88} \cdot \frac{80}{89} \cdot \frac{79}{90} + \frac{10}{87} \cdot \frac{80}{89} \cdot \frac{79}{88} \cdot \frac{78}{87} + \frac{10}{86} \cdot \frac{80}{89} \cdot \frac{79}{88} \cdot \frac{78}{87} \cdot \frac{77}{86}$$

e si ottiene il solito risultato.

Interessante il fatto che ciascuno dei 5 termini è la probabilità che venga estratto uno dei dieci numeri nella posizione corrispondente se non ne sono stati estratti in precedenza...

Bibliografia

- [Car86] A. Bruce Carlson. *Communication systems*. McGraw–Hill, New York, 1986.
- [Cou02] Leon W. Couch. *Fondamenti di telecomunicazioni*. Apogeo, 2002.
- [Dev08] Keith Devlin. *La lettera di Pascal*. Rizzoli, 2008.
- [EB77a] S. Benedetto E. Biglieri. *Teoria dei segnali determinati*. Boringhieri, 1977.
- [EB77b] S. Benedetto E. Biglieri. *Teoria delle probabilità e variabili casuali*. Quaderni di Elettronica, Bollati Boringhieri, 1977.
- [EC02] C. Galdi E. Conte. *Appunti di Teoria dei Fenomeni Aleatori*. Università di Napoli Federico II, 2002.
- [GG06] G. Pavan G. Galati. *Teoria dei Fenomeni Aleatori*. TexMat Tor Vergata, 2006.
- [Men79] Umberto Mengali. *Teoria dei sistemi di comunicazione*. TS Università, 1979.
- [Pap62] A. Papoulis. *The Fourier Integral and its Applications*. McGraw–Hill, 1962.
- [Pap77] A. Papoulis. *Signal Analysis*. McGraw–Hill, 1977.
- [Pap91] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw–Hill, 1991.
- [Sak68] D. J. Sakrison. *Communication Theory: Transmission of Waveforms and Digital Information*. John Wiley & Sons, 1968.