

Radio Communications

Federico Casu

Novembre 04, 2023

Analog Communications

Amplitude Modulation (AM)

La modulazione d'ampiezza è la modulazione analogica più semplice che esista. Il suo funzionamento è il seguente:

$$\text{Trasmettitore: } s(t) = A_c \cdot m(t) \cdot \cos(2\pi f_c t)$$

$$\text{Ricevitore: } m(t) = A_c \cdot s(t) \cdot 2 \cos(2\pi f_c t)$$

Studiamo nel dettaglio l'aspetto del segnale trasmesso $s(t)$. In particolare, studiamo lo spettro del modulo di $s(t)$ ¹.

$$\begin{aligned} \mathcal{F}\{s(t)\} &= \mathcal{F}\{A_c \cdot m(t) \cdot \cos(2\pi f_c t)\} = \\ &= A_c \cdot \mathcal{F}\{m(t) \cdot \cos(2\pi f_c t)\} = \\ &= A_c \cdot \mathcal{F}\{m(t) \cdot \frac{e^{j2\pi f_c t} + e^{-j2\pi f_c t}}{2}\} = \\ &= A_c \cdot \mathcal{F}\{m(t) \cdot \frac{e^{j2\pi f_c t}}{2} + m(t) \cdot \frac{e^{-j2\pi f_c t}}{2}\} \\ &= \frac{A_c}{2} \cdot \mathcal{F}\{m(t) \cdot e^{j2\pi f_c t}\} + \frac{A_c}{2} \cdot \mathcal{F}\{m(t) \cdot e^{-j2\pi f_c t}\} \end{aligned}$$

Per proseguire lo studio del segnale $\mathcal{F}\{s(t)\}$ è utile ricordare il **teorema della modulazione nel tempo**:

$$\text{Per ogni } f_0 \in \mathbb{R}, \mathcal{F}\{x(t) \cdot e^{j2\pi f_0 t}\} = X(f - f_0), \text{ dove } X(f) = \mathcal{F}\{x(t)\}.$$

Possiamo sfruttare il teorema della modulazione nel tempo per calcolare $\mathcal{F}\{m(t) \cdot e^{j2\pi f_c t}\}$ e $\mathcal{F}\{m(t) \cdot e^{-j2\pi f_c t}\}$:

$$\mathcal{F}\{s(t)\} = \frac{A_c}{2} \cdot M(f - f_c) + \frac{A_c}{2} \cdot M(f + f_c)$$

Quali sono gli effetti, in frequenza, del prodotto con il coseno (Figura 1)?

¹Lo spettro del segnale $s(t) = A_c \cdot m(t) \cdot \cos(2\pi f_c t)$ è composto solo dalla componente del modulo in quanto la fase, essendo un segnale reale, è nulla in tutto il dominio della frequenza.

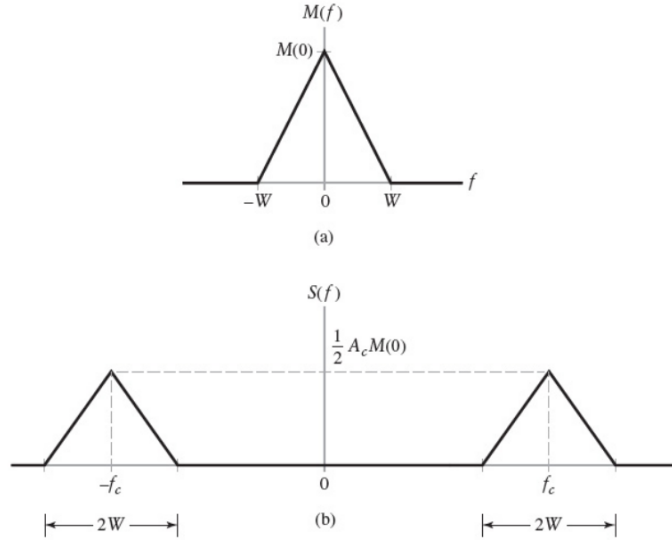


Figure 1: **(a)** Spettro del segnale $s(t)$ prima della modulazione. **(b)** Spettro del segnale $s(t)$ dopo la modulazione.

1. La **banda** del segnale modulato è *due* volte la banda del segnale in banda base. In Figura 1, la banda del segnale in banda base è W mentre la banda del segnale modulato è $2W$.
2. Grazie alla **simmetria hermitiana**², lo spettro del segnale in banda base, $M(f)$, è simmetrico rispetto all'origine. La modulazione, che non è altro che il prodotto tra il segnale $m(t)$ e $\cos(2\pi f_c t)$, produce due copie esatte del segnale in banda base di cui una centrata a frequenza $-f_c$ e l'altra centrata a frequenza f_c . Da ciò segue che anche il segnale modulato è simmetrico rispetto all'origine (la modulazione non influisce sulla simmetria hermitiana). Inoltre, il segnale modulato è simmetrico rispetto alla frequenza portante.

Traslare un segnale in banda base ad una frequenza diversa da $f = 0$ è necessario per almeno *due* motivi:

1. Se tutte le trasmissioni avvenissero in banda base allora i segnali trasmessi interferirebbero tra di loro.
2. Minore è la frequenza di trasmissione, maggiori le dimensioni dell'antenna necessari per trasmettere tali segnali.

Come si può recuperare il segnale modulante, *i.e.* il segnale che porta informazione? Sono necessarie tre operazioni:

1. Filtraggio mediante un *filtro passa banda*. Un filtro passa banda rimuove

² *Simmetria hermitiana*: se $m(t)$ è reale allora $M(-f) = M^*(f)$. $|M(-f)| = |M^*(f)|$.

tutte le componenti del segnale in input che si trovano a frequenze al di fuori dell'intervallo $[f_c - B, f_c + B]$, dove B è la banda del filtro in banda base.

2. Demodulazione. L'operazione di demodulazione “riporta” il segnale modulato nella sua versione in banda base.
3. Filtraggio mediante un *filtro passa basso*. La demodulazione produce, oltre che il segnale modulante, altre due copie del segnale: una copia centrata a $-2f_c$ e l'altra copia centrata a $2f_c$. Mediante il filtro passa basso il ricevitore è in grado di isolare la componente del segnale in banda base dalle componenti *ridondanti* in banda passante.

Demodulazione:

$$\begin{aligned} m_R(t) &= s(t) \cdot 2\cos(2\pi f_c t) = A_c \cdot m(t) \cdot \cos(2\pi f_c t) \cdot 2\cos(2\pi f_c t) = \\ &= A_c \cdot m(t) \cdot 2\left(\frac{1}{2} + \frac{\cos(2\pi 2f_c t)}{2}\right) = \\ &= A_c \cdot m(t) + A_c \cdot m(t) \cdot \cos(2\pi 2f_c t) \end{aligned}$$

$$\begin{aligned} \mathcal{F}\{m_R(t)\} &= \mathcal{F}\{A_c \cdot m(t) + A_c \cdot m(t) \cdot \cos(2\pi 2f_c t)\} = \\ &= \mathcal{F}\{A_c \cdot m(t)\} + \mathcal{F}\{A_c \cdot m(t) \cdot \cos(2\pi 2f_c t)\} = \\ &= A_c \cdot M(f) + \frac{A_c}{2}M(f - 2f_c) + \frac{A_c}{2}M(f + 2f_c) \end{aligned}$$

Analog Quadrature Amplitude Modulation (QAM)

La modulazione AM è poco efficiente. Ciò è una diretta conseguenza della *simmetria hermitiana*: il segnale modulante $m(t)$, essendo reale, quando viene modulato si sdoppia in due componenti identiche centrate a f_c e $-f_c$ per mantenere la proprietà di simmetria. La simmetria, seppur sia una proprietà matematicamente elegante, è inefficiente dal punto di vista dell'occupazione della banda: il segnale modulato occupa il doppio della banda rispetto al segnale in banda base.

I sistemi di comunicazione QAM sfruttano l'ortogonalità del seno e del coseno per raddoppiare il contenuto informativo trasmesso nell'unità di tempo.

Trasmittitore: $s(t) = s_I(t) \cdot \cos(2\pi f_c t) - s_Q(t) \cdot \sin(2\pi f_c t)$

Lato ricevitore il segnale $s(t)$ viene demodulato sia per mezzo di $2\cos(2\pi f_c t)$ sia $-2\sin(2\pi f_c t)$. Il risultato dell'operazione di demodulazione è il seguente:

$$\begin{aligned} s_I(t) &= s(t) \cdot 2\cos(2\pi f_c t) = \\ &= [s_I(t) \cdot \cos(2\pi f_c t) - s_Q(t) \cdot \sin(2\pi f_c t)] \cdot 2\cos(2\pi f_c t) = \\ &= s_I(t) \cdot \cos(2\pi f_c t) \cdot 2\cos(2\pi f_c t) - s_Q(t) \cdot \sin(2\pi f_c t) \cdot 2\cos(2\pi f_c t) \end{aligned}$$

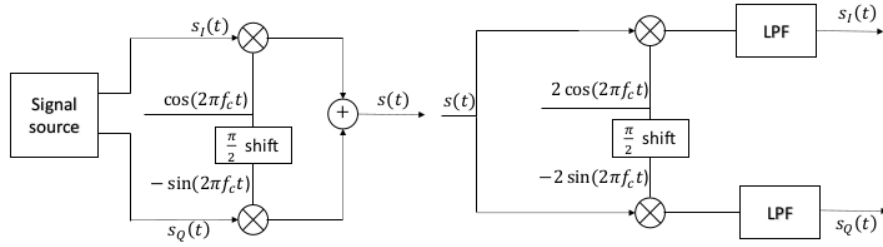


Figure 2: Schema di trasmissione/ricezione di un sistema QAM analogico.

Applicando le formule di Werner³ all'equazione di sopra otteniamo il seguente segnale:

$$\begin{aligned}
 &= s_I(t) \cdot \cos(0) + s_I(t) \cdot \cos(2\pi(2f_c)t) - s_Q(t) \cdot \sin(0) - s_Q(t) \cdot \sin(2\pi(2f_c)t) = \\
 &= s_I(t) + s_I(t) \cdot \cos(2\pi(2f_c)t) - s_Q(t) \cdot \sin(2\pi(2f_c)t)
 \end{aligned}$$

Il segnale demodulato è dunque composto da tre componenti:

- Il segnale originario $s_I(t)$, quest'ultimo in banda base.
- Una copia del segnale $s_I(t)$ traslata a $2f_c$.
- Una copia del segnale $s_Q(t)$ traslata a $2f_c$.

Grazie al filtro passa-basso siamo in grado di isolare il segnale in banda base dalle copie (quest'ultime in banda passante) ottenute in fase di demodulazione. L'analisi per il segnale $s_Q(t)$, proveniente dal ramo in *quadratura*, è del tutto analoga a quella appena compiuta per il segnale proveniente dal ramo in *fase*, *mutatis mutandis*.

Ovviamente c'è un prezzo da pagare: nei calcoli fatti sopra abbiamo assunto che i segnali sinusoidali generati dal trasmettitore sono *perfettamente sincronizzati* con i segnali generati dal ricevitore. In pratica, è quasi impossibile ottenere una sincronia di fase perfetta tra trasmettitore e ricevitore. Se è presente uno *sfasamento* allora non è più possibile ricavare una copia fedele del segnale originario a partire dal segnale trasmesso $s(t)$.

Inoltre, seppur il segnale modulato continua ad essere *reale* (quindi simmetrico rispetto l'origine), con la modulazione QAM si perde la simmetria intorno la frequenza portante. Quest'ultima è una conseguenza della modulazione mediante due segnali sinusoidali ortogonali.

³Abbiamo utilizzato $\cos(\alpha)\cos(\beta) = \frac{1}{2}\cos(\alpha + \beta) + \frac{1}{2}\cos(\alpha - \beta)$ e $\sin(\alpha)\cos(\beta) = \frac{1}{2}\sin(\alpha + \beta) + \frac{1}{2}\sin(\alpha - \beta)$.

Complex envelope

L'involuppo complesso di un segnale passa banda è una rappresentazione equivalente traslata in banda base.

Perchè studiamo l'involuppo complesso?

1. Studiare l'involuppo complesso è più semplice di studiare il segnale originale in banda passante. Questo perchè tramite l'involuppo complesso è possibile “eliminare” dal discorso la frequenza portante.
2. Simulare numericamente l'involuppo complesso di un segnale passa banda è più semplice visto che la frequenza di campionamento è molto minore rispetto alla frequenza di campionamento del segnale passa banda originale.

Ogni segnale passa banda può essere rappresentato come

$$\begin{aligned} s(t) &= \text{Re}\{\tilde{s}(t) \cdot e^{j2\pi f_c t}\} = \\ &= \text{Re}\{[s_I(t) + j \cdot s_Q(t)] \cdot [\cos(2\pi f_c t) + j \cdot \sin(2\pi f_c t)]\} = \\ &= s_I(t) \cdot \cos(2\pi f_c t) - s_Q(t) \cdot \sin(2\pi f_c t) \end{aligned}$$

dove $\tilde{s}(t) = s_I(t) + j \cdot s_Q(t)$ è l'*involuppo complesso* del segnale $s(t)$.

Frequency Modulation (FM)

Nella modulazione di frequenza il contenuto informativo è incorporato nella fase del segnale. Il segnale trasmesso è il seguente:

$$s_{\text{FM}}(t) = A_c \cdot \cos(2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau)$$

L'involuppo complesso del segnale $s_{\text{FM}}(t)$ è il seguente:

$$\tilde{s}_{\text{FM}}(t) = A_c \cdot e^{j2\pi k_f \int_{-\infty}^t m(\tau) d\tau}$$

Andiamo a studiare la *frequenza istantanea*⁴ del segnale $s_{\text{FM}}(t)$:

$$\begin{aligned} f_i(t) &= \frac{1}{2\pi} \frac{d\phi(t)}{dt} = \\ &= \frac{1}{2\pi} \frac{d}{dt} (2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau) = \\ &= f_c + k_f m(t) \end{aligned}$$

⁴La frequenza istantanea è definita come la derivata rispetto al tempo della fase del segnale, i.e. $f_i(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt}$.

A partire dalla frequenza istantanea, calcoliamo la *frequency deviation*, cioè quanto il segnale si scosta dalla frequenza portante:

$$f_d(t) = f_i(t) - f_c = k_f m(t)$$

Possiamo notare che, il contenuto informativo da trasmettere, oltre che essere incorporato nella fase del segnale, influisce sullo spettro del segnale: al variare del segnale $m(t)$, varia la frequenza del segnale portante e di conseguenza varia la sua rappresentazione nel dominio di Fourier. Le modulazioni viste in precedenza (AM e QAM) eseguivano una modulazione d'ampiezza, ovvero incorporavano l'informazione da trasmettere nell'ampiezza del segnale trasmesso. Inoltre, le modulazioni d'ampiezza producono segnali la cui occupazione di banda nel dominio della frequenza è costante nel tempo: stimare la banda di tali segnali è semplice. Con la modulazione di frequenza il gioco è ben diverso: ora la modulazione è ad *ampiezza costante*⁵ mentre ciò che varia nel tempo è la fase del segnale trasmesso. Quest'ultima conseguenza rende difficile stimare l'occupazione di banda del segnale trasmesso. Perché? Il motivo è semplice: non esiste una forma chiusa della trasformata di Fourier del segnale $s_{FM}(t)$.

Studiamo l'occupazione di banda del segnale $s_{FM}(t)$ mediante un esempio: consideriamo il seguente segnale modulante $m(t) = V_m \cos(2\pi f_m t)$.

Il segnale trasmesso dal trasmettitore FM è il seguente:

$$\begin{aligned} s_{FM}(t) &= A_c \cos(2\pi f_c t + 2\pi k_f \int_{-\infty}^t V_m \cos(2\pi f_m \tau) d\tau) = \\ &= A_c \cos(2\pi f_c t + 2\pi k_f \cdot \frac{V_m}{2\pi f_m} \sin(2\pi f_m t)) = \\ &= A_c \cos(2\pi f_c t + \frac{k_f \cdot V_m}{f_m} \sin(2\pi f_m t)) \end{aligned}$$

Definiamo due indici che ci saranno utili:

1. *Maximum frequency deviation*: $\Delta f = \max \{|f_d(t)|\} = k_f \cdot \max \{|m(t)|\}$
2. *Modulation index*: $m_f = \frac{\Delta f}{B}$, dove B è la banda del segnale modulante.

Possiamo notare che, nello sviluppo del segnale $s_{FM}(t)$, $k_f \cdot V_m$ coincide con la *maximum frequency deviation* del segnale modulante. Inoltre, ricordando che la banda in frequenza di un segnale sinusoidale è data da f_m , il segnale $s_{FM}(t)$ può essere riscritto come segue:

$$s_{FM}(t) = A_c \cos(2\pi f_c t + m_f \sin(2\pi f_m t))$$

Studiamo l'involuppo complesso del segnale $s_{FM}(t)$:

$$\tilde{s}_{FM}(t) = A_c \cdot e^{jm_f \sin(2\pi f_m t)}$$

⁵Una modulazione è detta ad *ampiezza costante* quando l'ampiezza del segnale è costante durante tutta la durata delle trasmissioni.

L'involuppo complesso $\tilde{s}_{FM}(t)$ è *periodico*. Possiamo rappresentare il segnale $\tilde{s}_{FM}(t)$ mediante somma di coefficienti di Fourier:

$$\tilde{s}_{FM}(t) = \sum_n S_n e^{j2\pi f_m n t}$$

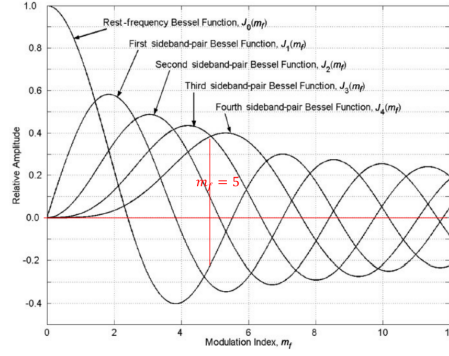


Figure 3: Possiamo notare che all'aumentare del modulo di m_f , aumenta il numero di coefficienti $J_n(m_f) \neq 0$.

Non daremo la dimostrazione ma sappiamo che i coefficienti di Fourier S_n possono essere calcolati come $J_n(m_f)$, dove $J_n(x)$ è una funzione di Bessel del primo tipo di ordine n .

$$\begin{aligned} s_{FM}(t) &= \Re\{\tilde{s}_{FM}(t)e^{j2\pi f_c t}\} = \\ &= \Re\{A_c \sum J_n(m_f) e^{j2\pi(f_c + nf_m)t}\} = \\ &= \Re\{A_c \sum J_n(m_f) [\cos(2\pi(f_c + nf_m)t) + j\sin(2\pi(f_c + nf_m)t)]\} = \\ &= A_c \sum J_n(m_f) \cos(2\pi(f_c + nf_m)t) = \\ \mathcal{F}\{s_{FM}(t)\} &= A_c \sum J_n(m_f) \cdot \left[\frac{1}{2} \delta(f - f_c - nf_m) + \frac{1}{2} \delta(f + f_c + nf_m) \right] \end{aligned}$$

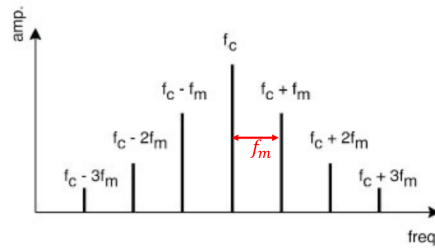


Figure 4: $\mathcal{F}\{s_{FM}(t)\}$

Lo spettro del segnale trasmesso consiste in un *treno di delta* (Figura 4): è molto difficile stimare la banda del segnale perché il contributo in frequenza del treno di delta può proseguire all'infinito! Possiamo comunque notare che:

1. Le componenti del treno di delta non nulle dipendono dai coefficienti di Fourier $J_n(m_f)$.
2. La distanza tra una componente e l'altra è data dalla frequenza del segnale modulante f_m .

Dalle considerazioni fatte sopra possiamo dedurre che la banda del segnale modulato dipende da m_f .

Carson's Rule: $B_{FM} \approx 2(m_f + 1)B = 2(\Delta f + B)$

Il segnale modulante può essere recuperato differenziando la fase del segnale modulato. Il ricevitore assume la forma presentata in Figura 5.

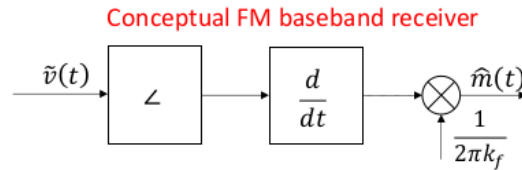


Figure 5: Ricevitore FM.

Software Defined Radio

Software Defined Radio (SDR) è una tecnologia che permette di costruire apparati radio non più strettamente realizzati come puro hardware, ovvero con circuiti e dispositivi elettronici, ma basati su un software in funzione su un PC o un sistema embedded.

In particolare vengono utilizzati ricevitori radio adatti a ricevere uno spettro di frequenze molto ampio e software appositi per gestire le funzioni di modulazione e demodulazione dei segnali. Col metodo SDR si ha un apparato riprogrammabile ogni qualvolta si vuole gestire uno standard di comunicazione diverso.

Vediamo come funziona il modulo radio (Figura 6):

1. La prima operazione compiuta dal modulo ricevitore è filtrare il segnale proveniente dall'antenna mediante un filtro passa-banda. Grazie a questa operazione, come abbiamo già studiato in precedenza, siamo in grado di isolare le frequenze del segnale centrate in f_c ed "attenuare" le componenti al di fuori dell'intervallo di frequenze $[f_c - B, f_c + B]$, dove B è la banda del segnale che siamo interessati a ricevere.
2. Il ricevitore è diverso rispetto a quello a cui siamo abituati (*homodyne*). I ricevitori di tipo *heterodyne* prendono il segnale proveniente dall'antenna

(quest'ultimo centrato alla frequenza portante f_c) e lo traslano ad una frequenza intermedia $0 < f_{if} < f_c$. Traslare il segnale ricevuto ad una frequenza inferiore alla frequenza portante è necessario per semplificare l'amplificazione del segnale. Vediamo, matematicamente, come avviene questa operazione:

$$\begin{aligned}
f_{lo} &= f_c - f_{if} \\
\bar{s}(t) &= s(t)\cos(2\pi f_{lo}t) = \\
&= [m(t)\cos(2\pi f_c t) - m(t)\sin(2\pi f_c t)]\cos(2\pi f_{lo}t) = \\
&= m(t)\cos(2\pi f_c t)\cos(2\pi f_{lo}t) - m(t)\sin(2\pi f_c t)\cos(2\pi f_{lo}t) = \\
&= m(t)\cos(2\pi[f_c - f_{lo}]t) + m(t)\cos(2\pi[f_c + f_{lo}]t) - m(t)\cos(2\pi[f_c + \frac{\pi}{2} - f_{lo}]t) - \\
&\quad - m(t)\cos(2\pi[f_c + \frac{\pi}{2} + f_{lo}]t) = \\
&= m(t)\cos(2\pi f_{if}t) + m(t)\cos(2\pi[2f_c - f_{if}]t) - m(t)\cos(2\pi[\frac{\pi}{2} + f_{if}]t) - \\
&\quad - m(t)\cos(2\pi[2f_c + \frac{\pi}{2} - f_{if}]t) = \\
&= m(t)\cos(2\pi f_{if}t) + m(t)\cos(2\pi[2f_c - f_{if}]t) - m(t)\sin(2\pi f_{if}t) - \\
&\quad - m(t)\sin(2\pi[2f_c - f_{if}]t)
\end{aligned}$$

3. Il segnale $\bar{s}(t)$ passa attraverso un filtro passa basso che rimuove le componenti del segnale a frequenza $2f_c - f_{if}$.
4. Ora il segnale può essere amplificato (*Low Noise RF Amplifier* in Figura 6).
5. Il segnale viene campionato rispettando il teorema di Shannon. In particolare, il segnale $\bar{s}(t)$ è centrato alla frequenza $f_{if} = 3.57$ MHz. La frequenza di campionamento è $f_{adc} = 28.8$ MHz. Il teorema di Shannon è ampiamente rispettato visto che $f_{adc} \gg 2(f_{if} + B)$. Il segnale campionato ha la seguente forma:

$$\bar{s}[n] = \bar{s}(nT_s) = \bar{s}(n/f_{adc}) = m(n/f_{adc})\cos(2\pi \frac{f_{if}}{f_{adc}}n) - m(n/f_{adc})\sin(2\pi \frac{f_{if}}{f_{adc}}n)$$

6. I campioni I/Q vengono prodotti in seguito alla demodulazione del segnale campionato. Le operazioni di demodulazione sono del tutto uguali a quelle viste per un sistema QAM. L'unica differenza che si può notare è nella forma dei segnali: non abbiamo segnali continui ma segnali discreti! Di conseguenza, il segnale $\bar{s}[n]$ è moltiplicato per $2\cos(2\pi \frac{f_{if}}{f_{adc}}n)$ e $-2\sin(2\pi \frac{f_{if}}{f_{adc}}n)$ per ottenere i campioni I e Q, rispettivamente.

Una volta ottenuti i campioni I e Q, da quest'ultimi andremo ad estrarre il contenuto informativo trasmesso. Come si fa? Ricordiamo che, un trasmettitore FM *imprime* il contenuto informativo nella *fase* del segnale trasmesso. In sostanza, dobbiamo studiare un modo per estrarre la fase del segnale ricevuto dai campioni I e Q.

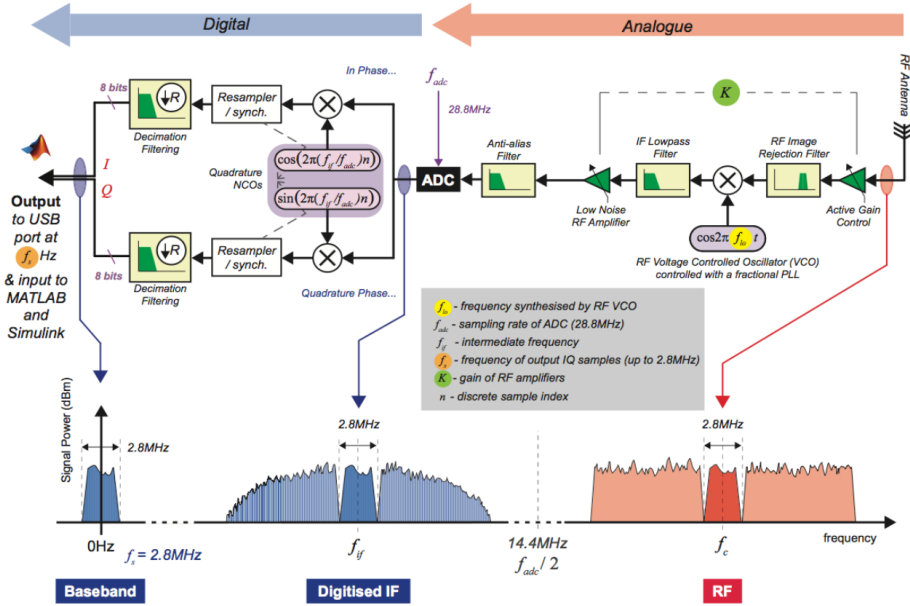


Figure 6: Schema a blocchi delle operazioni eseguite dal modulo radio in tecnologia SDR sul segnale ricevuto in banda passante.

L'involuppo complesso del segnale in entrata al modulo software è il seguente:

$$\begin{aligned}\tilde{v}[n] &= v(nT_s) = v\left(n\frac{1}{f_{adc}}\right) = \\ &= A_c e^{j2\pi k_f \int_{-\infty}^{nT_s} m(\tau) d\tau} \approx A_c e^{j2\pi k_f \sum_{k=-\infty}^n m(k)T_s}\end{aligned}$$

Visto e considerato l'approssimazione presentata sopra, notiamo che il rapporto tra un campione ed il precedente produce il seguente risultato:

$$\frac{v[n]}{v[n-1]} = A_c e^{j2\pi k_f \sum_{k=-\infty}^n m(k)T_s} \cdot A_c e^{-j2\pi k_f \sum_{k=-\infty}^{n-1} m(k)T_s} = A_c^2 e^{j2\pi k_f m(n)T_s}$$

Ecco qua: per ricavare una stima del segnale $m(t)$ è sufficiente estrarre la fase del rapporto tra un campione ed il precedente!.

$$m(nT_s) \approx \frac{1}{2\pi k_f T_s} \angle \left[\frac{v[n]}{v[n-1]} \right]$$

Digital Communications

I sistemi di comunicazione studiati fino ad ora sono *analogici*: i segnali trasmessi sono continui nel tempo. Come possiamo inviare segnali digitali? Un segnale digitale, purtroppo, è un segnale la cui banda è infinita.

Non è fattibile inviare la sequenza di bits così come si presenta (ovvero un segnale discreto). L'idea che risolve i nostri problemi è molto semplice: trasformiamo la sequenza di bits in un segnale continuo. Tale operazione è detta *interpolazione*.

Dunque, lo schema a blocchi del trasmettitore è il seguente:

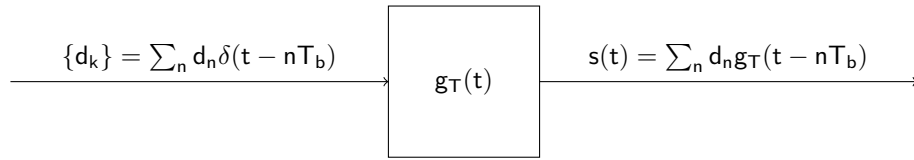


Figure 7: Interpolazione dei bit

Attenzione: in generale, per risparmiare energia è meglio trasmettere informazioni a media *nulla* ⁶. Per risolvere questo problema possiamo applicare un'ulteriore trasformazione ad ogni bit della sequenza:

$$0 \longrightarrow -1, \quad 1 \longrightarrow 1$$

Il ragionamento fatto sopra può essere esteso per un insieme di bit. Di fatto, stiamo associando ad una sotto-sequenza di bit un *simbolo*: l'unità di trasmissione, ora, non è più il bit ma è il simbolo.

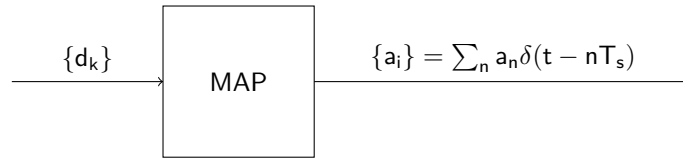


Figure 8: Mapping bits \rightarrow simbolo.

La frequenza di trasmissione dei simboli (*symbol rate*) è minore della frequenza di generazione dei bits (*bit rate*). Data la durata del bit T_b , poiché ogni simbolo mappa m bit, il tempo del simbolo T_s è:

$$T_s = m T_b \longrightarrow R_s = \frac{1}{T_s} = \frac{1}{m T_b} \leq \frac{1}{T_b} = R_b$$

⁶Si parla di segnale a media nulla perché i segnali digitali vengono modellati attraverso processi stocastici discreti. In seguito verranno studiati i processi stocastici e perché sono utili per lo studio dei sistemi di comunicazione digitali.

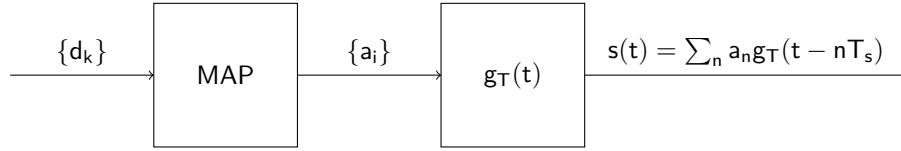


Figure 9: Interpolazione dei simboli

Pulse Amplitude Modulation (PAM)

La **Pulse Amplitude Modulation (PAM)** è un tipo di modulazione in cui l'informazione (espressa in bits) è codificata nell'ampiezza di una serie di segnali. La modulazione PAM è ottenuta attraverso due operazioni:

1. Il mapping dei bits in simboli, *i.e.* ad ogni simbolo trasmesso è associata biunivocamente una sottosequenza di bits estratta dalla sequenza di bits da trasmettere. Il mapping dei bits è svolto da un modulo, detto *mapper*, che prende in input i bits $\{d_k\}$ prodotti dalla sorgente e restituisce una sequenza di simboli $\{a_i\}$.
2. La conversione del segnale discreto composto dalla sequenza dei simboli in segnale analogico mediante l'utilizzo di un filtro passa-basso. Tale filtro è detto *impulso sagomatore*. Il segnale analogico ottenuto dall'interpolazione dei simboli assume la seguente forma: $\tilde{s}_{PAM}(t) = \sum_n a_n g_T(t - nT_s)$.
3. La modulazione del segnale, *i.e.* la conversione del segnale $\tilde{s}_{PAM}(t)$ in banda passante. Dunque, il segnale in uscita da un trasmettitore PAM è $s_{PAM} = \sum_n a_n g_T(t - nT_s) \cos(2\pi f_c t)$.

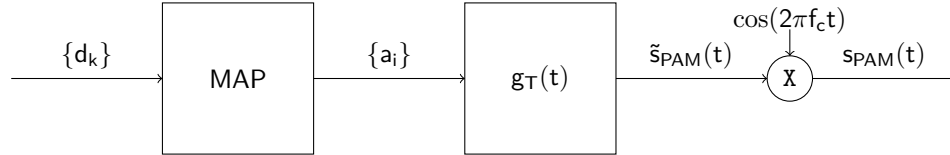


Figure 10: PAM - Trasmettitore

Il segnale PAM è equivalente al segnale AM-DSB, in cui il segnale modulante (e l'involuppo complesso) è

$$m(t) = \sum_n a_n g_T(t - nT_s) = \tilde{s}_{PAM}(t)$$

D'ora in avanti ci occuperemo di studiare il sistema PAM prendendo in considerazione l'involuppo complesso del segnale $s_{PAM}(t)$. Per semplicità, ometteremo il simbolo \sim : $s_{PAM}(t) = \sum_n a_n g_T(t - nT_s)$.

Dal punto di vista del ricevitore, il segnale $s_{PAM}(t)$ è un *processo stocastico*. Per capire l'occupazione di banda del segnale $s_{PAM}(t)$ dovremo studiarne il suo spettro

in frequenza. Invece, quando si ha a che fare con processi aleatori, lo studio della banda viene compiuto analizzando lo spettro della densità di potenza.

Densità spettrale di potenza segnale PAM

$$\text{DSP}_{\text{PAM}}(f) = \frac{S_a(f)}{T_s} |G_T(f)|^2$$

dove T_s è il periodo di emissione di un simbolo, $S_a(f)$ è la densità spettrale di potenza della sequenza dei simboli $\{a_i\}$ e $G_T(f) = \mathcal{F}\{g_T(t)\}$.

La densità spettrale di potenza della sequenza di simboli $\{a_n\}$ può essere calcolata sfruttando il teorema di *Wiener-Kintchine*: la sequenza $\{a_n\}$ è un processo stocastico discreto, stazionario e indipendente per cui le ipotesi del teorema sono rispettate.

$$S_a(f) = \mathcal{F}\{R_{aa}(m)\}$$

$$(1) \quad R_{aa}(m) = E\{a_i a_{i+m}\} = \text{indipendenza} = \begin{cases} E\{a_i^2\} & \text{se } m = 0 \\ E\{a_i\} E\{a_{i+m}\} & \text{altrimenti} \end{cases}$$

$$= \text{stazionarietà (in senso lato)} = \begin{cases} E\{a_i^2\} & \text{se } m = 0 \\ E\{a_i\} E\{a_i\} & \text{altrimenti} \end{cases}$$

$$= E\{a_i^2\} \delta[m] + \sum_{k \neq 0} E\{a_i\}^2 \delta[m+k]$$

$$(2) \quad E\{a_i\} = \sum_{n=1}^M a_n \frac{1}{M} = \frac{1}{M} \sum_{n=1}^M (2n-1-M) =$$

$$= \frac{2}{M} \frac{M(M+1)}{2} - \sum_{n=1}^M \frac{M+1}{M} = (M+1) - (M+1) = 0 \rightarrow E\{a_i\}^2 = 0$$

$$(3) \quad R_{aa}(m) = E\{a_i^2\} \delta[m]$$

$$(4) \quad S_a(f) = \mathcal{F}\{R_{aa}(m)\} = E\{a_i^2\} = A$$

Densità spettrale di potenza segnale PAM standard

$$\text{DSP}_{\text{PAM}}(f) = \frac{E\{a_i^2\}}{T_s} |G_T(f)|^2$$

Dall'ultima espressione ottenuta per $\text{DSP}_{\text{PAM}}(f)$, è facile notare che la banda del segnale dipende dal filtro passa-basso usato per interpolare la sequenza di simboli. Come ben sappiamo, se un segnale ha una durata finita del tempo allora occupa una banda infinita in frequenza (e viceversa). Come scegliamo l'impulso sagomatore?

Come al solito, ci troviamo di fronte ad un compromesso:

- spettro compatto \rightarrow grande quantità di interferenza nel dominio temporale (Scelta estrema: **rect** nel dominio delle frequenze e **sinc** nel tempo).
- spettro ampio \rightarrow la maggior parte dell'energia del simbolo è contenuta all'interno di un intervallo di simbolo (Scelta estrema: **rect** nel dominio temporale e **sinc** nelle frequenze).

PAM - Ricevitore. Prima di studiare come è strutturato il modulo ricevitore in una PAM, studiamo il sistema nella sua interezza introducendo le caratteristiche del *canale*.

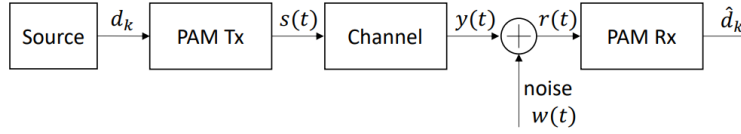


Figure 11: Schema a blocchi rappresentante un sistema di comunicazione PAM.

- Il canale di propagazione è generalmente modellato come un filtro lineare tempo-invariante (LTI) con risposta impulsiva $h(t)$. Quando il canale è ideale, allora $h(t) = \delta(t)$.
- La componente di rumore $w(t)$ è un processo stazionario gaussiano, a media nulla e bianco, con densità spettrale di potenza $S_N(f) = \frac{N_0}{2}$.

Attenzione: lo schema a blocchi rappresentato in Figura 11 è un modello equivalente in banda base (il segnale $s(t)$ è, in realtà, l'involuppo complesso del segnale trasmesso da T_x). Il segnale in entrata al ricevitore R_x è il seguente:

$$\begin{aligned}
 r(t) &= y(t) + w_{BP}(t) = \\
 &= \tilde{s}(t) \otimes h(t) + n(t) \\
 &= m(t)\cos(2\pi f_c t) \otimes h(t) + w_{BP}(t) \\
 &= \sum_n a_n g_T(t - nT_s) \cos(2\pi f_c t) \otimes h(t) + w_{BP}(t)
 \end{aligned}$$

dove $w_{BP}(t)$ è la versione traslata in banda passante del rumore $w(t)$, *i.e.*, $w_{BP}(t) = w(t)\cos(2\pi f_c t)$.

L'equivalente in banda base del segnale $r(t)$ è il seguente:

$$\begin{aligned}
 r(t) &= y(t) + w(t) = \\
 &= s(t) \otimes h(t) + w(t) \\
 &= m(t) \otimes h(t) + w(t) \\
 &= \sum_n a_n g_T(t - nT_s) \otimes h(t) + w(t)
 \end{aligned}$$

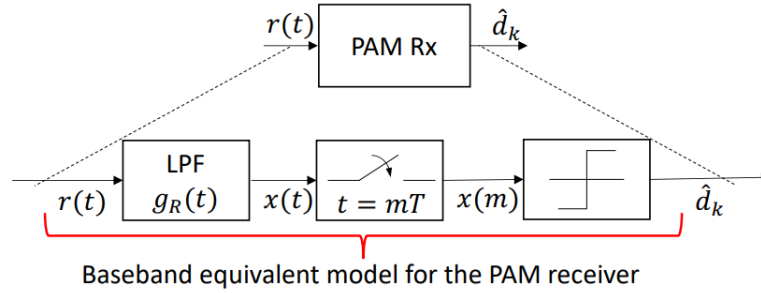


Figure 12: PAM - Ricevitore equivalente in banda base.

Il ricevitore **Rx** PAM, per recuperare i bits trasmessi con $s(t)$, esegue le seguenti operazioni:

1. Filtra il rumore e le interferenze dal segnale ricevuto $r(t)$. Il filtraggio viene eseguito da un filtro passa-basso lineare e tempo-invariante (LTI) $g_R(t)$.
2. Campiona il segnale filtrato $x(t)$. Il periodo di campionamento coincide con il periodo del simbolo, *i.e.*, T_s .
3. Recupera i bit trasmessi dai campioni del segnale $x[m]$. Tale operazione viene eseguita da un modulo chiamato *decisore*.

Studiamo il segnale in uscita dal filtro passa-basso $g_R(t)$ (Figura 13):

$$\begin{aligned}
 x(t) &= r(t) \otimes g_R(t) + w(t) \otimes g_R(t) = \\
 &= s(t) \otimes h(t) \otimes g_R(t) + w(t) \otimes g_R(t) = \\
 &= \sum_{n=0} a_n g_T(t - nT_s) \otimes h(t) \otimes g_R(t) + w(t) \otimes g_R(t)
 \end{aligned}$$

Definiamo $n(t) = w(t) \otimes g_R(t)$, ovvero il rumore gaussiano (ora non più bianco) presente nel segnale in uscita dal filtro di ricezione. Inoltre, se definiamo $g(t) = g_T(t) \otimes h(t) \otimes g_R(t)$, il segnale $x(t)$ può essere riscritto come:

$$x(t) = \sum_{n=0} a_n g(t - nT_s) + n(t)$$

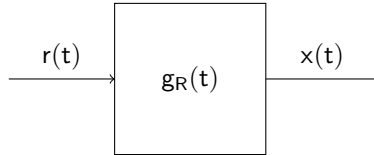


Figure 13: PAM: filtro passa-basso.

Ora vogliamo studiare l'aspetto del segnale in uscita dal campionatore, *i.e.*

$$x[m] = x(t)|_{t=mT_s}:$$

$$x[m] = \sum_n a_n g(mT_s - nT_s) + n(mT_s)$$

Per semplicità, per ora non consideriamo il rumore presente nel segnale $x[m]$. In particolare prendiamo in considerazione il seguente segnale:

$$x[m] = \sum_{n=-\infty}^{+\infty} a_n g(mT_s - nT_s) = \sum_{n=-\infty}^{+\infty} a_n g((m-n)T_s)$$

Definiamo l'indice $k = m - n$:

$$x[m] = \sum_{k=-\infty}^{+\infty} a_{m-k} g(kT_s) = a_m g(0) + \sum_{k, k \neq 0} a_{m-k} g(kT_s)$$

Attenzione: il segnale $x[m]$ è composto da due componenti ma soltanto una di queste può essere utilizzata per ricostruire il simbolo trasmesso all'istante mT_s :

1. $a_m g(0)$ componente utile.
2. $\sum_{k, k \neq 0} a_{m-k} g(kT_s)$ componente non utile: *interferenza intersimbolica* (ISI).

Da queste considerazioni nasce il *criterio di Nyquist*: vorremo che la risposta nel tempo del filtro $g(t)$ sia:

$$g(kT_s) = \begin{cases} 1 & \text{se } k = 0 \\ 0 & \text{altrimenti} \end{cases}$$

Da cui segue $g(t) = \delta(t)$. Se il criterio di Nyquist è rispettato allora, all'entrata del decisore, avremo la seguente sequenza:

$$x[m] = a_m + n(mT_s)$$

Esiste una condizione che traduce il criterio di Nyquist in frequenza? Sì, eccola qui:

1. Consideriamo il segnale $G(f) = \mathcal{F}\{g(t)\}$.
2. Visto che campionare nel tempo causa periodicità nel dominio della frequenza, la trasformata di Fourier del segnale $g(m) = g(t)|_{t=mT_s}$ è periodica.

$$\mathcal{F}\{g(m)\} = \frac{1}{T_s} \sum_k G(f - \frac{k}{T_s})$$

Ora non ci rimane che sfruttare ciò che ci dice il criterio di Nyquist: $g(t) = \delta(t) \rightarrow \mathcal{F}\{g(t)\} = \mathcal{F}\{\delta(t)\} = 1 \forall f$.

$$\mathcal{F}\{g(t)\} = \frac{1}{T_s} \sum_k G(f - \frac{k}{T_s}) = 1$$

oppure, equivalentemente:

$$\mathcal{F}\{g(t)\} = \sum_k G(f - \frac{k}{T_s}) = T_s$$

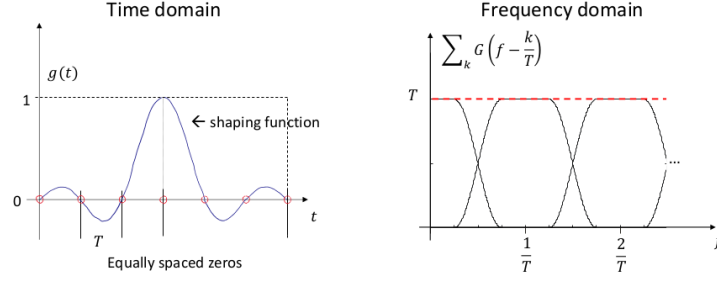


Figure 14: Criterio di Nyquist (nel tempo ed in frequenza).

Prima di poter studiare le caratteristiche in termini di banda, potenza ed energia per simbolo trasmesso del segnale PAM dobbiamo studiare come implementare un filtro che permette di 1) avere zero-ISI e 2) massimizzare il *signal-to-noise* ratio.

Raised Cosine Filters - La funzione coseno rialzato è spesso utilizzata per l'implementazione di filtri passa-basso per la modulazione digitale grazie alla sua capacità di annullare l'ISI. Il nome deriva dal fatto che la porzione non nulla dello spettro in frequenza della sua forma più semplice ($\beta = 1$) è una funzione coseno, 'rialzata' di 1.

La banda del filtro è data dalla seguente espressione:

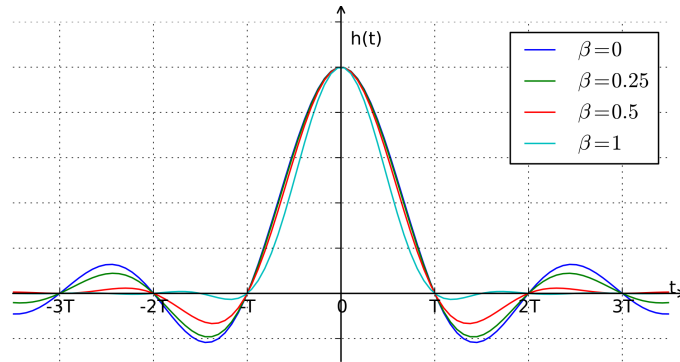
$$B_{RC}(\beta) = \frac{\beta + 1}{2T_s}$$

dove T_s è il periodo di campionamento e β è detto fattore *roll-off*. Il fattore roll-off è un parametro di design e ci dice quanto rapidamente la risposta del filtro diventa nulla.

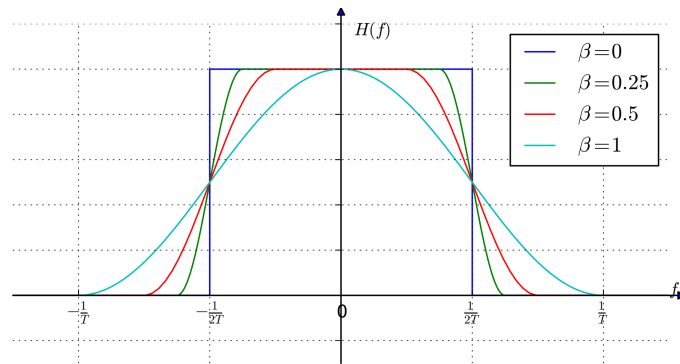
Additive white Gaussian noise - Trascurando, per il momento, l'effetto del canale, l'altro principale problema nel ricevitore è la presenza di rumore gaussiano bianco. La densità spettrale di potenza del rumore $w(t)$ è costante in tutto il dominio della frequenza: $S_w(f) = \frac{N_0}{2}$. La densità spettrale dell'involuppo complesso di $w(t)$, ovvero $\tilde{w}(t) = w^I(t) + jw^Q(t)$, è la seguente:

$$S_{\tilde{w}}(f) = 2N_0 \rightarrow S_{\tilde{w}}^{(I)}(f) = S_{\tilde{w}}^{(Q)}(f) = N_0$$

L'azione del filtro $g_R(t)$ cambia l'aspetto del rumore $w(t)$: il risultato del filtraggio è sempre un processo aleatorio stazionario e gaussiano ma la densità spettrale di



(a) Dominio del tempo.



(b) Dominio della frequenza

Figure 15: Raised Cosine Filter

potenza non è più costante! La componente di rumore in uscita dal filtro $g_R(t)$ non è più bianca:

$$S_n(f) = S_w(f) |G_R(f)|^2 = \frac{N_0}{2} |G_R(f)|^2$$

$$S_{\tilde{n}}(f) = S_{\tilde{w}}(f) |G_R(f)|^2 = 2N_0 |G_R(f)|^2 \quad (\text{inviluppo complesso})$$

Matched filters

Il filtro $g_R(t)$ che **massimizza** il rapporto signal-to-noise (quindi minimizza l'effetto del rumore) è dato da:

$$g_R(t) = g_T(-t) \longrightarrow G_R(f) = G_T^*(f)$$

Se $g_T(t)$ è *pari*, allora $g_R(t) = g_T(t)$.

Vediamo l'idea sul quale si basa il concetto di matched filters:

1. Prendiamo in esame la Figura 16. Supponiamo di avere il filtro $G_R(f) = \text{rect}(f)$. Come è facile notare, il filtro riesce a tagliare tutte le componenti di rumore che si trovano al di fuori della banda del segnale $S(f)$.
2. Tale approccio non è ottimale: quando il segnale $S(f)$ ha ampiezza minore del rumore (esempio: nell'intervallo $[0.25, 0.50]$), il compito del decisore diventa molto arduo perchè la componente di rumore è molto maggiore della componente informativa che si vorrebbe estrarre.
3. Analizziamo la Figura 17. Il filtro è una esatta copia del segnale ricevuto (quindi è un'esatta copia dell'impulso sagomatore). Cosa possiamo notare? Nelle zone in cui il segnale è "meno forte" del rumore, il filtro riesce ad attenuare l'effetto del rumore, di fatto massimizzando il rapporto **signal-to-noise**.

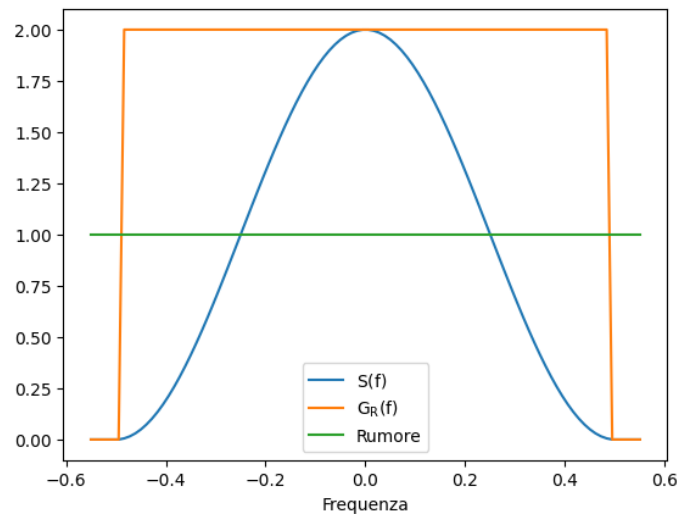


Figure 16

Root Raised Cosine Filter - I filtri RRC sono filtri la cui risposta in frequenza è la radice quadrata di un coseno rialzato, *i.e.*, $H_{RRC}(\beta, f) = \sqrt{H_{RC}(\beta, f)}$. Perché gli studiamo? Perché i filtri RRC sono delle implementazioni utili a

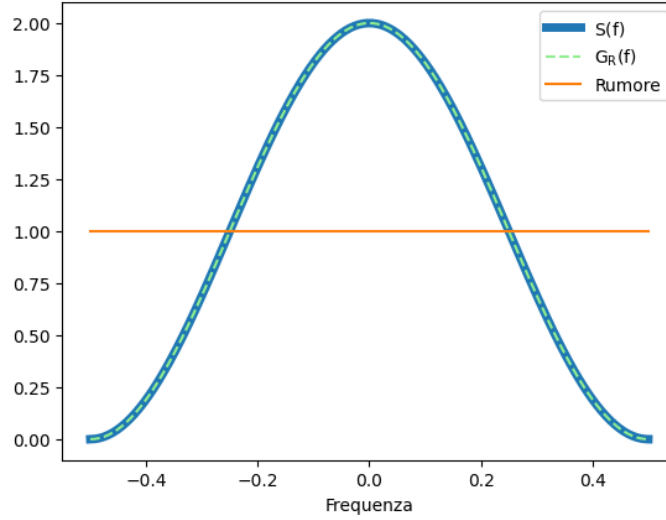


Figure 17

rispettare il criterio dei *matched filters*! Vediamo perché:

$$G_T(f) = G_R(f) = H_{RRC}(\beta, f)$$

$$\begin{aligned} (1) \text{ Nyquist Criterion: } g(t) &= g_T(t) * h(t) * g_R(t) = \\ &= h_{RRC}(\beta, t) * \delta(t) * h_{RRC}(\beta, t) = \\ &= h_{RRC}(\beta, t) * h_{RRC}(\beta, t) \end{aligned}$$

$$\begin{aligned} \mathcal{F}\{h_{RRC}(\beta, t) * h_{RRC}(\beta, t)\} &= H_{RRC}(\beta, f) \cdot H_{RRC}(\beta, f) = H_{RRC}(\beta, f)^2 \\ &= (\sqrt{H_{RC}(\beta, f)})^2 = H_{RC}(\beta, f) \rightarrow \text{Nyquist criterion OK} \end{aligned}$$

$$(2) \text{ Matched filters: } H_{RRC}(\beta, f) \text{ is even} \rightarrow G_R(f) = G_T^*(f) = G_T(f) \text{ OK!}$$

Banda di un segnale PAM con filtraggio RRC - La densità spettrale di potenza del segnale PAM, come visto in precedenza, è data dalla seguente espressione:

$$S_s(f) = \frac{1}{T_s} S_a(f) |G_T(f)|^2$$

Ora, viste e considerate le proprietà della sequenza di simboli $\{a_n\}$, possiamo dire che la banda di un segnale PAM dipende solamente dalla banda dell'impulso

sagomatore. Vediamo perchè:

$$\begin{aligned}
R_{aa}(m) &= E\{a_i a_{i+m}\} \rightarrow \text{indipendenza} \rightarrow E\{a_i\} E\{a_{i+m}\} = \\
&= \begin{cases} E\{a_i^2\} & m = 0 \\ E\{a_i\}^2 & m \neq 0 \rightarrow \text{stazionarietà} \end{cases} \\
&= E\{a_i^2\} \delta[m] + \sum_{k \neq 0} E\{a_i\}^2 \delta[m+k]
\end{aligned}$$

$$E\{a_i\} = 0 \rightarrow E\{a_i\}^2 = 0$$

$$E\{a_i^2\} = \frac{M^2 - 1}{3}$$

$$DSP_{PAM}(f) = \frac{E\{a_i^2\}}{T_s} |G_T(f)|^2$$

Supponiamo di implementare l'impulso sagomatore mediante l'utilizzo di $h_{RRC}(\beta, t)$. Da ciò segue

$$G_T(f) = H_{RCC}(\beta, f) \rightarrow DSP_{PAM}(f) = \frac{E\{a_i^2\}}{T_s} |H_{RRC}(f)|^2 = \frac{E\{a_i^2\}}{T_s} H_{RC}(f)$$

La banda del segnale $H_{RC}(\beta, f)$ è data dalla seguente espressione:

$$\text{Banda del segnale } H_{RC}(\beta, f) = \frac{\beta + 1}{2T_s}$$

Visto che abbiamo usato il segnale $H_{RC}(\beta, f)$, quest'ultimo in banda base, per ottenere l'occupazione del segnale PAM in banda passante è sufficiente raddoppiare la banda del segnale in banda base, *cioè*,

$$\text{Banda del segnale } H_{RC}^{BP}(\beta, f) = \frac{\beta + 1}{T_s}$$

Attenzione: la banda del segnale PAM dipende dal *bit-rate*. Vediamo come:

$$f_s = \frac{1}{T_s} \text{ (symbol rate)}$$

$$T_s = \log_2 M \cdot T_b$$

$$f_s = \frac{1}{\log_2 M \cdot T_b} = \frac{R_b}{\log_2 M}$$

$$\text{Banda del segnale } H_{RC}(\beta, f) = (\beta + 1) \frac{R_b}{\log_2 M}$$

Potenza di un segnale PAM con filtraggio RRC - Esistono due modi per calcolare la potenza:

1. $\int_{-\infty}^{+\infty} \text{DSP}_{\text{PAM}}(f) df$
2. $R_{ss}(0)$

Ovviamente la via più facile da percorrere è la seconda. In particolare, grazie al teorema di Wiener-Kintchine e alla biunivocità della trasformata continua di Fourier, sappiamo che l'antitrasformata della densità spettrale di potenza del segnale PAM è proprio l'autocorrelazione del segnale.

$$\mathcal{F}^{-1}\{\text{DSP}_{\text{PAM}}(f)\} = \mathcal{F}^{-1}\left\{\frac{E\{a_i^2\}}{T_s} |G_T(f)|^2\right\} \quad (1)$$

Supponiamo di applicare il principio di *matched filters* e di utilizzare $H_{\text{RRC}}(\beta, f)$ come impulso sagomatore e filtro di ricezione. La densità spettrale di potenza del segnale PAM è dunque la seguente:

$$\text{DSP}_{\text{PAM}}(f) = \frac{E\{a_i^2\}}{T_s} H_{\text{RC}}(\beta, f) \quad (2)$$

Sostituiamo l'espressione ottenuta nell'Equazione 2 dentro l'Equazione 1. Il risultato è il seguente:

$$\mathcal{F}^{-1}\{\text{DSP}_{\text{PAM}}(f)\} = \mathcal{F}^{-1}\left\{\frac{E\{a_i^2\}}{T_s} H_{\text{RC}}(\beta, f)\right\} = \frac{E\{a_i^2\}}{T_s} h_{\text{RC}}(\beta, t)$$

La potenza del segnale PAM può essere calcolata come

$$\frac{E\{a_i^2\}}{T_s} h_{\text{RC}}(\beta, 0) = \frac{E\{a_i^2\}}{T_s} = \frac{M^2 - 1}{3T_s}$$

Energia per simbolo trasmesso - L'energia per simbolo è calcolata come la potenza del segnale a cui si moltiplica la durata del simbolo, cioè

$$E_s = P_s \cdot T_s = \frac{E\{a_i^2\}}{2T_s} \cdot T_s = E\{a_i^2\} = \frac{M^2 - 1}{6}$$

Strategia di decisione - La nostra variabile decisionale è $x[m] = a_m + n[m]$, dove

1. a_m è il simbolo trasmesso all'istante mT_s in assenza di ISI. A tale risultato ci siamo arrivati grazie a $g_T(t) = g_R(t) = h_{\text{RRC}}(\beta, t)$ e all'ipotesi di canale ideale, $h(t) = \delta(t)$.
2. $n[m]$ è una variabile aleatoria ottenuta in seguito al campionamento del rumore filtrato. In altre parole, $n[m] = n(t)|_{t=mT_s} = [w(t) * g_R(t)]|_{t=mT_s}$.

Visto che $n(t)$ è un processo aleatorio gaussiano a media nulla, allora $n[m]$ è una variabile aleatoria gaussiana a media nulla. La varianza σ_n^2 della variabile aleatoria $n[m]$ può essere calcolata come $\sigma_n^2 = E\{n[m]^2\} - E\{n[m]\}^2$. Visto che $n[m]$ è una variabile aleatoria a media nulla, allora la varianza σ_n^2 coincide con il valore quadratico medio $E\{n[m]^2\}$.

$$E\{n[m]^2\} = E\{n[m]n^*[m]\} = E\{n[m]n^*[m+k]\}_{k=0} = R_{nn}(k)|_{k=0}$$

$$R_{nn}(k)|_{k=0} = \int_{-\infty}^{+\infty} S_n(f) df = 2N_0 \int_{-\infty}^{+\infty} |G_R(f)|^2 df$$

L'ultima osservazione da fare è la seguente, visto che il filtro di ricezione è $h_{RRC}(\beta, t)$, allora sappiamo calcolare $\int_{-\infty}^{+\infty} |G_R(f)|^2 df$:

$$\int_{-\infty}^{+\infty} |G_R(f)|^2 df = \int_{-\infty}^{+\infty} H_{RC}(f) df = \int_{-\infty}^{+\infty} H_{RC}(f) e^{j2\pi f t} |_{t=0} = h_{RC}(\beta, 0) = 1$$

Ora che abbiamo caratterizzato, *statisticamente*, la variabile decisionale $x[m]$, possiamo definire la strategia decisionale che implementa il decisore.

Optimal decision strategy

La strategia ottima è quella che sceglie il simbolo della costellazione $a^{(i)}$ che massimizza la probabilità $P(a^{(i)} | x[m])$.

$$\hat{a}_m = \arg \max \{P(a^{(i)} | x[m])\}$$

Calcolare la probabilità $P(a^{(i)} | x[m])$ è molto difficile. Possiamo sfruttare il teorema di Bayes:

$$P(a^{(i)} | x[m]) = \frac{P(x[m] | a^{(i)})P(a^{(i)})}{P(x[m])}$$

Se i simboli $a^{(i)}$ sono equiprobabili, allora possiamo notare che la probabilità condizionata $P(a^{(i)} | x[m])$ è proporzionale alla probabilità condizionata $P(x[m] | a^{(i)})$: massimizzare $P(x[m] | a^{(i)})$ equivale a massimizzare $P(a^{(i)} | x[m])$!

Maximum likelihood decision strategy

La strategia ottima è quella che sceglie il simbolo della costellazione $a^{(i)}$ che massimizza la probabilità $P(x[m] | a^{(i)})$.

$$\hat{a}_m = \arg \max \{P(x[m] | a^{(i)})\}$$

Come si applica la strategia di decisione *maximum likelihood* al campione $x[m]$? Per prima cosa, studiamo la densità di probabilità della variabile $x[m]$.

$$x[m] = x_I[m] + jx_Q[m] = a_m + n_I[m] + jn_Q[m]$$

Visto che il segnale PAM è *reale* (il contenuto informativo è trasmesso mediante il ramo in fase del sistema), possiamo considerare solamente la parte reale della variabile decisionale $x[m]$:

$$x_I[m] = \Re\{x[m]\} = \Re\{a_m + n_I[m] + jn_Q[m]\} = a_m + n_I[m]$$

Ecco fatto! Diamo uno sguardo a $x_I[m]$:

1. a_m è una costante (non è una variabile aleatoria).
2. $n_I[m]$ è una variabile aleatoria *gaussiana*, a media nulla e con varianza $\sigma_I^2 = N_0$.

Ecco che abbiamo trovato la densità di probabilità della variabile $x_I[m]$:

$$f(x_I[m]) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_I[m] - a_m)^2}{2\sigma^2}\right)$$

Massimizzare $P(x_I[m] | a^{(i)})$ equivale a minimizzare la distanza $|x_I[m] - a^{(i)}|$ visto che la variabile decisionale è $\mathcal{N}(a^{(i)}, \sigma^2 = N_0)$.

Adottando il criterio della massima verosimiglianza, possiamo suddividere lo spazio del segnale in regioni di decisione, dove la regione $Z^{(i)}$ è l'insieme di punti che sono più vicini al simbolo $a^{(i)}$ che a qualsiasi altro simbolo.

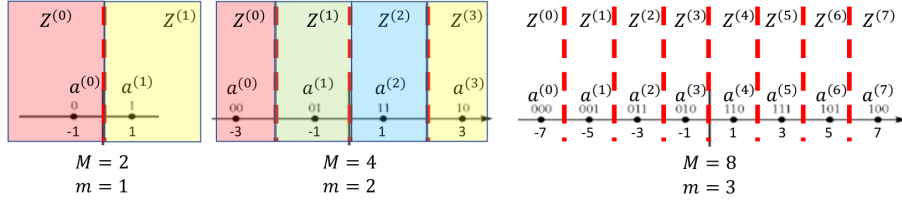


Figure 18: Strategia di decisione PAM (simboli equiprobabili). Regioni di decisione.

PAM Error Probability

La strategia di decisione presentata in precedenza, seppur *ottima*, comunque può produrre degli errori a causa del rumore. Vediamo un esempio:

- Supponiamo di avere una 2-PAM la cui costellazione è composta dai simboli 1 e -1 , equiprobabili, associati a 0 e 1, rispettivamente.

- Supponiamo che il trasmettitore abbia inviato il simbolo $\mathbf{a}_m = -1$ (quindi il bit inviato risulta essere 0). Al ricevitore, a causa del “forte” rumore, la variabile decisionale $x[m]$ risulta essere 0.75.
- Il decisore implementa la strategia di decisione a *minima distanza*:
 - $|x[m] - \mathbf{a}^{(0)}| = |0.75 + 1| = 1.75$
 - $|x[m] - \mathbf{a}^{(1)}| = |0.75 - 1| = 0.25$
- Il decisore sceglie il simbolo $\hat{\mathbf{a}}_m = 1$: errore!

Definiamo la probabilità d’errore del simbolo $\mathbf{a}^{(i)}$, $P(e | \mathbf{a}^{(i)})$, come la probabilità che la variabile decisionale $x[m]$ non cada nella regione $Z^{(i)}$, sapendo di aver trasmesso $\mathbf{a}^{(i)}$, cioè:

$$\begin{aligned} P(e | \mathbf{a}^{(i)}) &= P(x[m] \notin Z^{(i)} | \mathbf{a}_m = \mathbf{a}^{(i)}) = \\ &= 1 - P(x[m] \in Z^{(i)} | \mathbf{a}_m = \mathbf{a}^{(i)}) \end{aligned}$$

Se il decisore implementa una strategia *maximum likelihood*, lo scopo del decisore è massimizzare la probabilità $P(x[m] \in Z^{(i)} | \mathbf{a}_m = \mathbf{a}^{(i)})$: quest’ultima rappresenta la probabilità di successo del decisore.

La probabilità d’errore totale può essere calcolata come la media di tutte le probabilità d’errore dei simboli $\mathbf{a}^{(i)}$, cioè:

$$P_e = \frac{1}{M} \sum_{i=0}^{M-1} P(e | \mathbf{a}^{(i)}) \approx \lim_{N^{(s)} \rightarrow \infty} \frac{N_e^{(s)}}{N^{(s)}}$$

In precedenza abbiamo studiato che la variabile decisionale $x[m]$ è una variabile gaussiana $\mathcal{N}(\mathbf{a}_m, \sigma^2 = N_0)$. Calcolare la probabilità d’errore $P(e | \mathbf{a}^{(i)})$ equivale a calcolare l’area sottesa dalla funzione $f(x[m] | \mathbf{a}^{(i)})$ che non appartiene alla regione $Z^{(i)}$.

Supponiamo di voler calcolare la probabilità d’errore associata al simbolo $\mathbf{a}^{(0)} = -1$ in una 2-PAM. Il decisore a minima distanza di una 2-PAM utilizza una sola soglia di decisione, $t = 0$. Dunque, calcolare la probabilità d’errore associata al simbolo $\mathbf{a}^{(0)} = -1$ equivale a calcolare l’area sottesa dalla funzione $f(x[m] | \mathbf{a}^{(0)})$ nell’intervallo $[0, \infty]$ (la soglia di decisione è $t = 0$).

$$\begin{aligned} P(e | \mathbf{a}^{(0)}) &= \int_t^\infty f(x | \mathbf{a}^{(0)}) dx = \\ &= \int_t^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mathbf{a}^{(0)})^2}{2\sigma^2}} dx = \end{aligned}$$

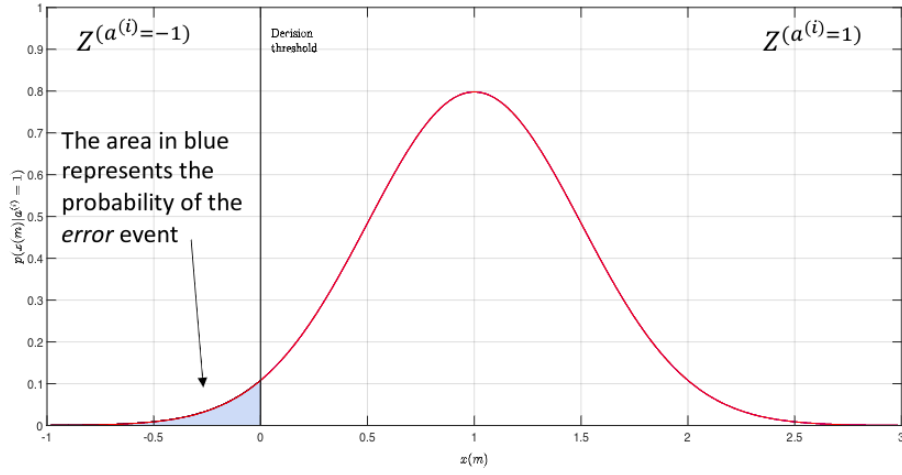


Figure 19: Probabilità di errore associata al simbolo $a^{(1)} = 1$

Applichiamo la sostituzione $y = \frac{x-a^{(0)}}{\sigma}$. Segue:

$$\begin{aligned} dy &= \frac{dx}{\sigma} \rightarrow dx = \sigma dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\frac{t-a^{(i)}}{\sigma}}^{\infty} e^{-\frac{y^2}{2}} dx \sigma = \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{t-a^{(i)}}{\sigma}}^{\infty} e^{-\frac{y^2}{2}} dy \end{aligned}$$

Notiamo che $f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ non è altro che la distribuzione di probabilità di una variabile aleatoria Normale, $\mathcal{N}(0, 1)$.

Definiamo la funzione coda $Q(x)$ come:

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Alcune proprietà utili della funzione coda:

1. $Q(x)|_{x \rightarrow -\infty} = 1$
2. $Q(x)|_{x=0} = \frac{1}{2}$ (è simmetrica rispetto l'origine).
3. $Q(-x) = 1 - Q(x)$

Dunque, la probabilità d'errore $P(e | a^{(0)})$ può essere calcolata come:

$$P(e | a^{(0)}) = Q\left(\frac{t - a^{(0)}}{\sigma}\right) = Q\left(\frac{0 - (-1)}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right)$$

Come si calcola la probabilità d'errore $P(e | a^{(1)})$? Semplice: il decisore compie un errore quando il campione ricevuto è minore di $t = 0$ se il simbolo trasmesso in origine era $a_m = a^{(1)} = 1$: da ciò segue che la probabilità d'errore $P(e | a^{(1)})$ si può calcolare come l'area della funzione $f(x|m | a^{(1)})$ nell'intervallo $[-\infty, 0]$.

$$\begin{aligned} P(e | a^{(1)}) &= \int_{-\infty}^0 f(x | a^{(1)}) dx = \\ &= \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-1)^2}{2\sigma^2}} dx = \end{aligned}$$

Applichiamo la sostituzione $y = \frac{x-1}{\sigma}$. Segue:

$$\begin{aligned} dy &= \frac{dx}{\sigma} \rightarrow dx = \sigma dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\frac{1}{\sigma}}^{\infty} e^{-\frac{y^2}{2}} \sigma dy = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\frac{1}{\sigma}}^{\infty} e^{-\frac{y^2}{2}} dy = Q\left(\frac{-1}{\sigma}\right) \end{aligned}$$

Ricordando che il decisore è simmetrico, allora $Q(\frac{1}{\sigma}) = Q(\frac{-1}{\sigma})$:

$$P(e | a^{(1)}) = Q\left(\frac{-1}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right)$$

La probabilità d'errore totale, dunque, può essere calcolata mediante la funzione $Q(x)$:

$$\begin{aligned} P_e^{2-PAM} &= \frac{1}{2}P(e | a^{(0)}) + \frac{1}{2}P(e | a^{(1)}) = \\ &= \frac{1}{2}Q\left(\frac{1}{\sigma}\right) + \frac{1}{2}Q\left(\frac{1}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right) \end{aligned}$$

Calcoliamo la probabilità d'errore totale per una 4 - PAM:

- I simboli trasmessi sono $a^{(i)} = 2i - 3$, $i = 0, \dots, 3$.
- Se i simboli sono equiprobabili allora possiamo implementare un decisore a *minima distanza*. In particolare, le soglie $t^{(i)}$ del decisore coincidono con il punto intermedio di ciascuno segmento $a^{(i)}a^{(i+1)}$, $0 < i < 3$. In altre parole, le soglie di decisione saranno $t^{(0)} = -2$, $t^{(1)} = 0$ e $t^{(2)} = 2$.
- $P(e | a^{(0)}) = Q\left(\frac{|t^{(0)} - a^{(0)}|}{\sigma}\right) = Q\left(\frac{|-2 - (-3)|}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right)$
- $P(e | a^{(3)}) = Q\left(\frac{|t^{(2)} - a^{(3)}|}{\sigma}\right) = Q\left(\frac{|2 - 3|}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right)$
- $P(e | a^{(1)}) = Q\left(\frac{|t^{(0)} - a^{(1)}|}{\sigma}\right) + Q\left(\frac{|t^{(1)} - a^{(1)}|}{\sigma}\right) = 2Q\left(\frac{1}{\sigma}\right)$

- $P(e | a^{(2)}) = Q\left(\frac{|t^{(1)} - a^{(2)}|}{\sigma}\right) + Q\left(\frac{|t^{(2)} - a^{(2)}|}{\sigma}\right) = 2Q\left(\frac{1}{\sigma}\right)$

La probabilità totale d'errore per una 4-PAM è la seguente:

$$P_e^{4-PAM} = \frac{1}{4} \sum_{i=0}^3 P(e | a^{(i)}) = \frac{3}{2} Q\left(\frac{1}{\sigma}\right)$$

PAM error probability as a function of SNR

2-PAM: $E_s = \frac{2^2-1}{6} = \frac{1}{2} \Rightarrow 2E_s = 1$ and $\sigma^2 = N_0$, and $\frac{1}{\sigma} = \sqrt{\frac{1}{\sigma^2}} = \sqrt{\frac{2E_s}{N_0}}$.

$$P_e^{(2-PAM)} = Q\left(\sqrt{\frac{2E_s}{N_0}}\right)$$

4-PAM: $E_s = \frac{4^2-1}{6} = \frac{5}{2} \Rightarrow \frac{2}{5}E_s = 1$, and $\frac{1}{\sigma} = \sqrt{\frac{1}{\sigma^2}} = \sqrt{\frac{2E_s}{5N_0}}$.

$$P_e^{(4-PAM)} = \frac{3}{2} Q\left(\sqrt{\frac{2E_s}{5N_0}}\right)$$

PAM bit error probability

Abbiamo definito la probabilità d'errore totale come il limite del rapporto tra il numero di simboli errati ricevuti e il numero di simboli inviati. Se il nostro sistema è ben progettato (Gray mapping) possiamo assumere che (in media) un errore sul simbolo si traduce in un solo bit errato.

$$P_e^{(b)} = \lim_{N^{(b)} \rightarrow \infty} \frac{N_e^{(b)}}{N^{(b)}} \approx \lim_{N^{(s)} \rightarrow \infty} \frac{N_e^{(s)}}{\log_2(M) N^{(s)}} = \frac{1}{\log_2(M)} \lim_{N^{(s)} \rightarrow \infty} \frac{N_e^{(s)}}{N^{(s)}} = \frac{1}{\log_2(M)} P_e$$

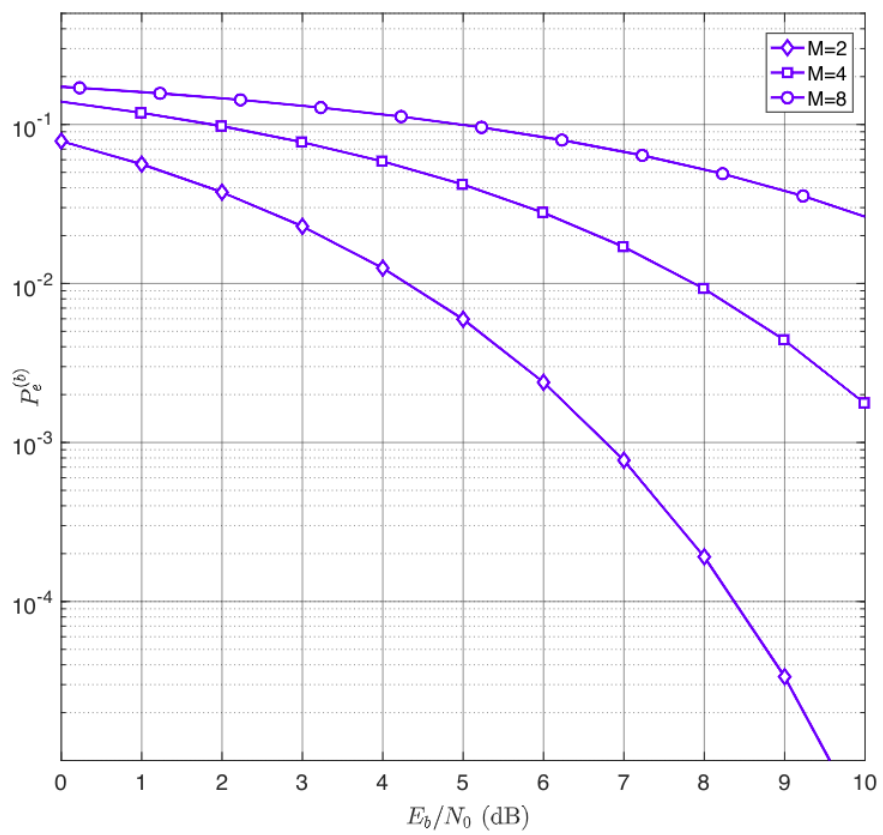


Figure 20: Andamento della probabilità d'errore sul bit in funzione del rapporto segnale/rumore.

(Appendice A) Processi Stocastici

Si definisce *processo stocastico* (o *processo aleatorio*) una famiglia di variabili aleatorie $\{X(t), t \in T \subseteq \mathbb{R}_+\}$ dipendenti da una variabile, che nel nostro caso è il *tempo*, definite su uno spazio campione Ω . Fissando un istante di tempo t_0 , è possibile individuare valori generalmente differenti, ognuno relativo a una determinata realizzazione e quindi ad un elemento dello spazio campione: $X(t_0)$ è allora una variabile aleatoria e rappresenta la “fotografia” del processo stocastico in un determinato istante. Quindi, rispetto a una semplice variabile aleatoria, $X(t_0)$ fornisce anche un’informazione relativa all’evoluzione temporale.

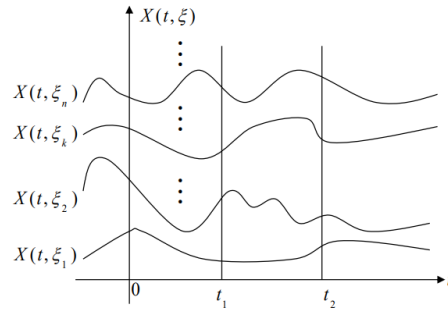


Figure 21: Realizzazioni di un processo aleatorio.

Ogni realizzazione $X(t_i)$ di un processo aleatorio può essere descritta attraverso le stesse funzioni che descrivono una variabile aleatoria: *distribuzione di probabilità* e *densità di probabilità*.

Distribuzione di probabilità:

$$F_X(x, t) = P\{X(t) \leq x\}$$

Densità di probabilità:

$$f_X(x, t) = \frac{d}{dx} F_X(x, t)$$

Valor medio:

$$\mu_X(t_0) = E\{X(t_0)\} = \int_{-\infty}^{+\infty} x f_X(x, t_0) dx$$

Autocorrelazione:

$$R_{XX}(t_1, t_2) = E\{X(t_1)X^*(t_2)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2^* f_{X(t_1), X(t_2)}(x_1, x_2) dx_1 dx_2$$

Stazionarietà - Un *processo stazionario* è un processo stocastico la cui distribuzione di probabilità congiunta non cambia se viene traslata nel tempo:

$$F_{X(t_1), \dots, X(t_n)}(x_1, \dots, x_n) = F_{X(t_1+\tau), \dots, X(t_n+\tau)}(x_1, \dots, x_n) \quad \forall \tau \in \mathbb{R}$$

Esistono diversi *ordini* di stazionarietà. Esempio: **Stazionarietà del 1° ordine** ($n = 1$). Un processo $X(t)$ è detto stazionario del primo ordine se le proprietà statistiche del processo sono uguali alle proprietà statistiche del processo $X(t + c)$, per ogni c . Da tale proprietà possiamo dire che un processo stazionario del 1° ordine ha valore medio costante e indipendente da t :

$$f_X(x, t) = f_X(x, t + c) \quad \forall c \in \mathbb{R} \longrightarrow E\{X(t)\} = E\{X(t + c)\} = \mu_X$$

Possiamo utilizzare una definizione più ampia di stazionarietà. Un processo $X(t)$ è detto **stazionario in senso lato** (wide sense stationarity, **WSS**) se vengono soddisfatte solo queste due condizioni:

- $E\{X(t)\} = \mu \quad \forall t \in \mathbb{R}$
- $R_{XX}(t_1, t_2) = R_{XX}(t_2 - t_1)$

Per un processo WSS, la media è costante e la funzione di autocorrelazione dipende solo dalla differenza tra gli indici di tempo.

Densità spettrale di potenza - La densità spettrale di potenza $S_X(f)$ di un processo stocastico WSS $X(t)$ descrive la distribuzione di potenza nelle componenti di frequenza che compongono il segnale. $S_X(f)$ è misurata in watt per hertz (W/Hz). Possiamo calcolare $S_X(f)$ sfruttando il teorema di *Wiener-Kintchine*.

Teorema di Wiener-Kintchine

Enunciato

$$S_X(f) = \mathcal{F}\{R_{XX}(\tau)\} = \int_{-\infty}^{+\infty} R_{XX}(\tau) e^{-j2\pi f\tau} d\tau$$

La potenza P_X del segnale $X(t)$ può essere calcolata integrando la densità spettrale di potenza $S_X(f)$:

$$P_X = \int_{-\infty}^{+\infty} S_X(f) df$$