

# Capitolo 1

## Modulazioni digitali

### 1.1 Sommario

La trasmissione di informazione tra una sorgente ed un destinatario attraverso un mezzo trasmissivo comporta sempre il ricorso alla modulazione che, a seconda della particolare applicazione, si propone di ottenere uno o più dei seguenti obiettivi:

- minimizzare la potenza irradiata.
- minimizzare la banda occupata.
- minimizzare la complessità e quindi i costi.

in questo capitolo si cercherà di evidenziare i motivi che da sempre portano a considerare la modulazione QPSK<sup>1</sup> come la più adatta a comunicazioni spaziali ad elevato data rate, in quest'ottica vengono anche introdotti concetti quali la sagomatura dell'impulso, la demodulazione con filtro adattato, i criteri di decisione e la codifica.

### 1.2 Modulazione

Molte delle moderne comunicazioni elettriche avvengono in forma digitale indipendentemente dal fatto che la sorgente sia digitale oppure analogica, il motivo di ciò è che a monte e/o a valle del canale trasmissivo il segnale subisce comunque delle elaborazioni digitali volte ad ottimizzarne le caratteristiche, pertanto la scelta di una trasmissione digitale rappresenta una soluzione di continuità.

Lo spettro di un generico segnale randomico NRZ<sup>2</sup> mostrato in Figura(1.1) evidenzia tuttavia come i segnali digitali non siano adatti alle tra-

---

<sup>1</sup>Quadrature Phase Shift Keying

<sup>2</sup>No Return to Zero

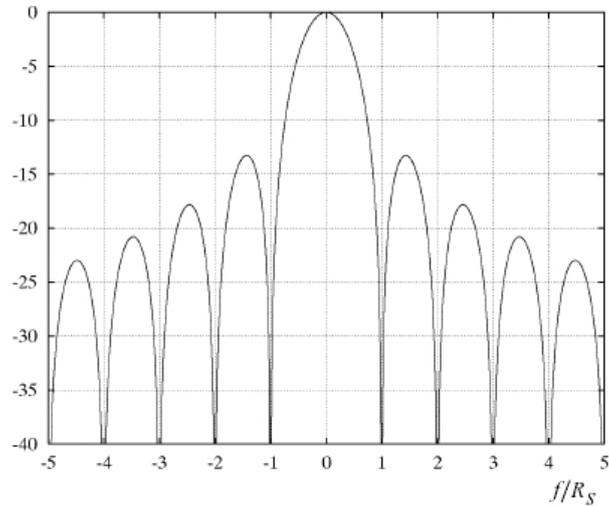


Figura 1.1: Spettro segnale NRZ

missioni in mezzi fisici reali, vi è infatti troppa potenza distribuita sui lobi laterali, essa è sia inutile ai fini della trasmissione dell'informazione che deleteria in quanto aumenta il livello del rumore in eventuali canali adiacenti. La potenza utile è invece addensata nei pressi della continua pertanto si rende necessaria una traslazione in frequenza al fine di poter inviare più sequenze digitali su di un unico mezzo trasmissivo.

### 1.2.1 Traslazione in frequenza

La traslazione in frequenza nella sua forma più semplice si ottiene moltiplicando il segnale digitale in banda base, opportunamente filtrato, per una portante analogica a frequenza  $f_c$ , ne deriva uno spettro bilatero centrato su questa frequenza come in Figura(1.2).

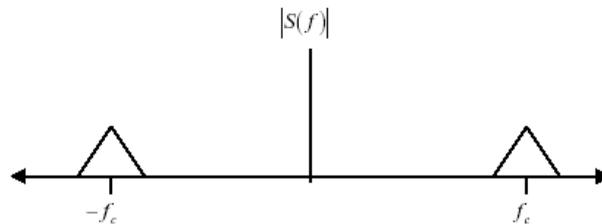


Figura 1.2: Spettro di un segnale in banda traslata

Al fine di poter confrontare sistemi operanti a frequenze diverse è importante poter rappresentare in banda base un qualsiasi segnale in banda

traslata  $s(t)$ , a tal riguardo, in riferimento a sistemi reali, è di interesse il solo asse positivo delle frequenze, descritto dal segnale analitico [2]

$$s_+(t) = s(t) + j\hat{s}(t) \quad (1.1)$$

dove

$$\hat{s}(t) = \frac{1}{\pi t} * s(t) \quad (1.2)$$

è la trasformata di Hilbert di  $s(t)$  il cui effetto nel dominio della frequenza è visualizzato in Figura(1.3), essa ruota di  $+90^\circ$  le componenti a frequenza

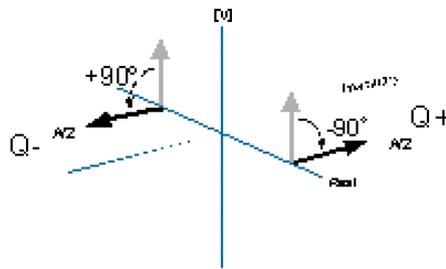


Figura 1.3: Trasformata di Hilbert nel dominio della frequenza

positiva e di  $-90^\circ$  quelle a frequenza negativa, ricordando inoltre che la traslazione in frequenza corrisponde ad una moltiplicazione per l'esponenziale complesso nel dominio del tempo si ottiene l'espressione dell'involuppo complesso  $s_l(t)$ :

$$s_l(t) = s_+(t) e^{-j2\pi f_c t} \quad (1.3)$$

Il segnale in banda traslata  $s(t)$  può pertanto essere rappresentato in tre diverse forme:

**Involuppo Complesso:**

$$s(t) = \Re [s_l(t) e^{-j2\pi f_c t}] \quad (1.4)$$

**Modulo e fase:**

$$s(t) = a(t) \cos(2\pi f_c t + \varphi(t)) \quad (1.5)$$

dove  $a(t)$  e  $\varphi(t)$  sono dei segnali reali in banda base che rappresentano rispettivamente l'ampiezza e la fase del segnale in banda traslata  $s(t)$ .

**Quadratura:**

$$s(t) = s_c(t) \cos(2\pi f_c t) - s_s(t) \sin(2\pi f_c t) \quad (1.6)$$

essendo  $s_c(t)$  e  $s_s(t)$  dei segnali in banda base denominati rispettivamente componente in fase e componente in quadratura del segnale in banda traslata  $s(t)$ .

La formulazione(1.6) è di fondamentale importanza in questa Tesi in quanto da essa si evince in maniera immediata lo schema di un possibile modulatore.

### 1.2.1.1 Spazio dei segnali

Un qualsiasi segnale reale può essere rappresentato come un vettore nello spazio dei segnali, tale spazio eredita le proprietà classiche degli spazi vettoriali definiti dall'algebra lineare, in particolare è caratterizzato da una norma:

$$\|s(t)\| = \sqrt{\langle s(t), s(t) \rangle} = \sqrt{\int_a^b |s(t)|^2 dt} \quad (1.7)$$

che consente di individuare una base ossia un gruppo di N segnali  $\psi_j$  aventi norma unitaria ed ortogonali tra loro:

$$\langle \psi_j, \psi_k \rangle = 0 \quad \forall j, k : j \neq k \quad (1.8)$$

Ogni segnale  $s_m(t)$  appartenente allo spazio dei segnali può essere rappresentato come combinazione lineare dei vettori  $\psi_j(t)$  della base

$$s_m(t) = \sum_{j=1}^N s_{mj} \psi_j(t) \quad (1.9)$$

essendo  $s_{mj}(t)$  la proiezione di  $s_m(t)$  lungo la direzione dello spazio vettoriale individuata dal  $j$ esimo vettore  $\psi_j(t)$  della base. Nel caso di uno spazio bidimensionale si ha che  $m \in (1, 2)$  pertanto tralasciando la base che è uguale per ogni segnale si ha che il generico segnale  $s_m(t)$  è individuato dal vettore  $s = [s_{m1} \ s_{m2}]$ .

Il progetto di una modulazione digitale si basa quindi sulla scelta di una base ortonormale nella quale rappresentare M segnali disposti a formare una costellazione le cui caratteristiche geometriche influenzano il progetto del trasmettitore e la qualità della ricezione. La sequenza di bit da trasmettere viene suddivisa in gruppi costituiti da  $b = \log_2 M$  bits, ogni gruppo individua un simbolo tra gli M di un alfabeto cui sono associati gli M segnali della costellazione.

### 1.2.1.2 Tipologie di modulazioni satellitari

Il segnale in banda traslata espresso dall'equazione(1.5) evidenzia come l'informazione in banda base possa essere codificata sia nel modulo che nella

fase<sup>3</sup> della portante. Per applicazioni satellitari non si utilizzano modulazioni che associano informazione al modulo della portante in quanto richiedono amplificatori lineari in classe A che non sfruttano al meglio la scarsa energia disponibile, inoltre tali modulazioni sono vulnerabili alle variazioni del guadagno del canale stesso.

Nel caso di una modulazione M-PSK<sup>4</sup> utilizzando una portante a frequenza  $f_c$  e per la quale ogni simbolo da trasmettere sia caratterizzato da energia  $E_S$  si ha l'espressione del segnale modulato:

$$s_m(t) = \sqrt{\frac{2E_S}{T}} g(t) \cos\left(\frac{2\pi}{M}(m-1)\right) \cos(2\pi f_c t) - \sqrt{\frac{2E_S}{T}} g(t) \cos\left(\frac{2\pi}{M}(m-1)\right) \sin(2\pi f_c t) \quad (1.10)$$

dove  $g(t)$  è la **sagomatura d'impulso**<sup>5</sup> ed  $m = 1, 2, \dots, M$  è un intero associato al simbolo del quale è richiesta la trasmissione.

La base è costituita dalle funzioni

$$\begin{aligned} \psi_1(t) &= \sqrt{\frac{2}{T}} g(t) \cos(2\pi f_c t) \\ \psi_2(t) &= \sqrt{\frac{2}{T}} g(t) \sin(2\pi f_c t) \end{aligned} \quad (1.11)$$

pertanto ogni segnale modulato può essere rappresentato da un vettore:

$$\mathbf{s}_m = [s_{m1} \ s_{m2}] = \left[ \sqrt{E_s} \cos\left(\frac{2\pi}{M}(m-1)\right) \quad \sqrt{E_s} \sin\left(\frac{2\pi}{M}(m-1)\right) \right] \quad (1.12)$$

Le tipologie più comuni di M-PSK sono la BPSK<sup>6</sup> ( $M=2$ ) e la QPSK ( $M=4$ ), per quest'ultima l'andamento nel tempo è illustrato in Figura(1.4), dove non essendo stata applicata la sagomatura dell'impulso si ha che l'involuppo è costante, tale soluzione è di solo interesse teorico in quanto l'efficienza energetica che ne deriva non bilancia i numerosi problemi derivanti dall'insorgere dell'interferenza intersimbolica.

La costellazione QPSK è costituita da 4 punti individuati sostituendo  $m = 1, 2, 3, 4$  nella equazione(1.13),

$$s_{mj} = \int_0^T s_m(t) \psi_j(t) \quad (1.13)$$

i loro valori sono riportati in Tabella(1.1).

<sup>3</sup>e quindi nella frequenza

<sup>4</sup>Phase Shift Keying

<sup>5</sup>Sezione(1.3.3)

<sup>6</sup>Binary Phase Shift Keying

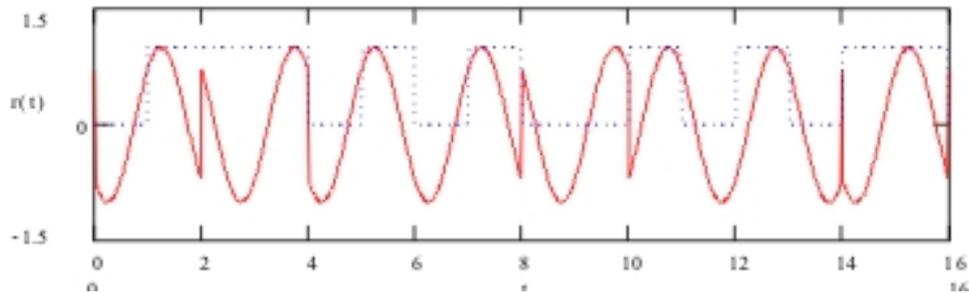


Figura 1.4: Modulazione QPSK nel tempo

Simbolo	Segnale	Coordinata I	Coordinata Q
00	$s_0$	+1	+1
01	$s_1$	-1	+1
10	$s_2$	-1	-1
11	$s_3$	+1	-1

Tabella 1.1: Punti della costellazione QPSK

Le brusche transizioni della fase in Figura(1.4) si hanno nel passaggio dal simbolo che nella costellazione di Figura(1.5a) ha coordinate  $(1,1)$  al simbolo con coordinate  $(-1,-1)$ , l'involuppo complesso passa per l'origine e compie una ampia escursione in ampiezza inadatta agli amplificatori non lineari utilizzati per ottimizzare l'impiego dell'energia nel satellite. La

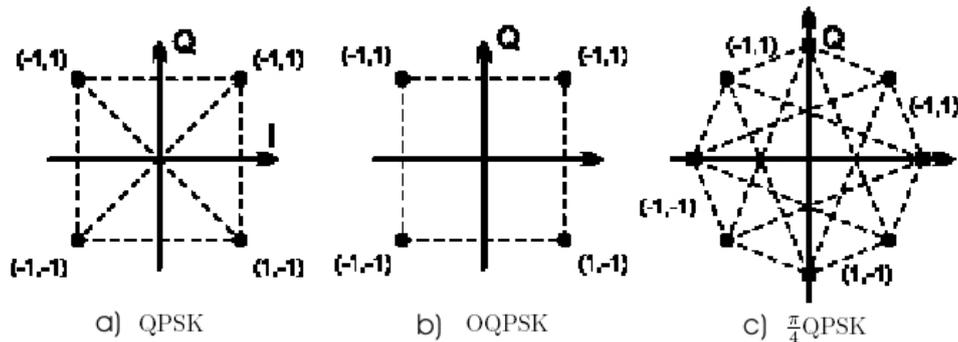


Figura 1.5: Tipologie di QPSK

modulazione OQPSK<sup>7</sup> la cui costellazione è rappresentata in Figura(1.5b) è una variante della QPSK che risolve il problema delle transizioni di fase

<sup>7</sup>Offset Quadrature Phase Shift Keying

mediante uno sfasamento temporale tra la componente in fase e quella in quadratura, esse non cambiano mai contemporaneamente pertanto la massima variazione dell'ampiezza dell'involucro complesso è di circa 3dB contro i 40dB della QPSK.

La modulazione  $\frac{\pi}{4}$ QPSK [1] risolve il medesimo problema ruotando di  $\frac{\pi}{4}$  la costellazione ad ogni simbolo come in Figura(1.5c), in tal modo la escursione di fase massima è di  $135^\circ$ , intermedia tra i  $180^\circ$  della QPSK convenzionale ed i  $90^\circ$  della OQPSK, rispetto a quest'ultima però si ha il vantaggio di poter utilizzare la demodulazione incoerente, l'informazione è infatti associata all'ampiezza della variazione di fase ed alla direzione in cui avviene la rotazione pertanto non è necessario che il ricevitore conosca la fase in anticipo visto che il riferimento per la demodulazione del simbolo attuale è costituito dal simbolo precedente. Il ricevitore  $\frac{\pi}{4}$ QPSK ha un'architettura più semplice ma la probabilità d'errore è maggiore di circa 3dB rispetto all'equivalente ricevitore coerente in quanto sono possibili due sorgenti d'errore, un simbolo corrotto oppure un riferimento errato.

Una diretta estensione della modulazione M-PSK è la M-QAM<sup>8</sup>, le loro costellazioni coincidono nel caso di  $M = 4$  mentre per valori diversi differiscono profondamente come evidenziato in Figura(1.6) nel caso di  $M = 16$ .

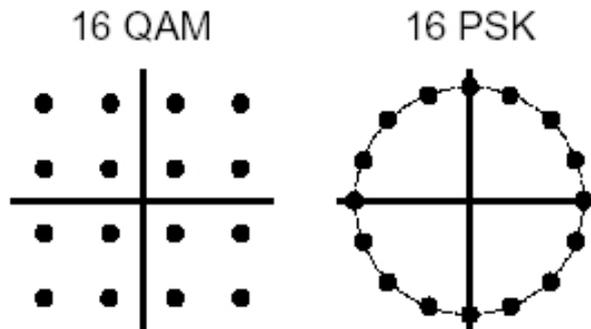


Figura 1.6: Confronto costellazioni M-QAM e M-PSK

Per entrambe l'efficienza spettrale è  $\eta_s = 4$  in quanto ad ogni simbolo sono associati 4 bit mentre nella QPSK soltanto due, i segnali della costellazione QAM tuttavia non hanno tutti la stessa energia pertanto, per i motivi precedentemente esposti, non sono adatti ad applicazioni spaziali.

Dalla costellazione della 16-PSK si può osservare come i segnali siano ravvicinati tra loro molto più che nella QPSK dove se ne ha uno in ogni quadrante, ciò determina un aumento della probabilità di scambiare un

<sup>8</sup>Quadrature Amplitude Modulation

segnale per uno dei segnali adiacenti pertanto occorre aumentare la potenza al fine di ottenere la stessa probabilità d'errore della QPSK.

L'altra grande famiglia di modulazioni digitali utilizzata in ambito spaziale è quella che interessa le variazioni della frequenza della portante e che va sotto il nome di M-FSK<sup>9</sup> [7], continua o discontinua a seconda che la variazione tra le M frequenze associate ai simboli da trasmettere avvenga con continuità di fase o meno. In Figura(1.7) è riportato l'andamento tem-

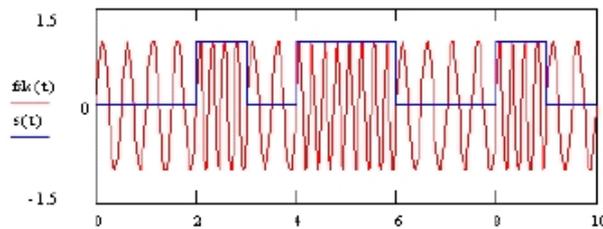


Figura 1.7: Andamento temporale BFSK

porale di una M-FSK incoerente per il caso di  $M = 2$ , essa si realizza con due oscillatori, a seconda del simbolo da trasmettere il modulatore emette una portante oppure l'altra.

Le discontinuità nella fase che si hanno nella M-FSK incoerente determinano una banda molto ampia e pertanto una scarsa efficienza spettrale, per questo motivo si sono affermate le tecniche a fase continua CPFSK<sup>10</sup> il cui generico segnale modulato è descritto dalla:

$$s_m(t) = \sqrt{\frac{2E_b}{T_b}} \cos\left(2\pi f_c t + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau\right) \quad (1.14)$$

dove la fase del segnale modulato è continua in quanto non direttamente proporzionale al segnale modulante digitale  $m(t)$  ma al suo integrale. L'indice di modulazione  $k_f = \frac{2\Delta f}{R_b}$  determina la separazione tra i toni, essa è minima nel caso della MSK<sup>11</sup> per la quale si ha  $k_f = 0,5$ . L'andamento della fase nel tempo per la MSK è rappresentato in Figura(1.8), ad ogni bit 1 da trasmettere viene associato un aumento della fase pari a  $+90^\circ$  mentre ad ogni bit 0 si associa una diminuzione di  $-90^\circ$  della stessa.

Una derivazione della MSK è la GMSK<sup>12</sup> che riduce ulteriormente la banda filtrando la modulante digitale con un filtro gaussiano avente risposta

<sup>9</sup>Frequency Shift Keying

<sup>10</sup>Continuous Phase Frequency Shift Keying

<sup>11</sup>Minimum Shift Keying

<sup>12</sup>Gaussian Minimum Shift Keying

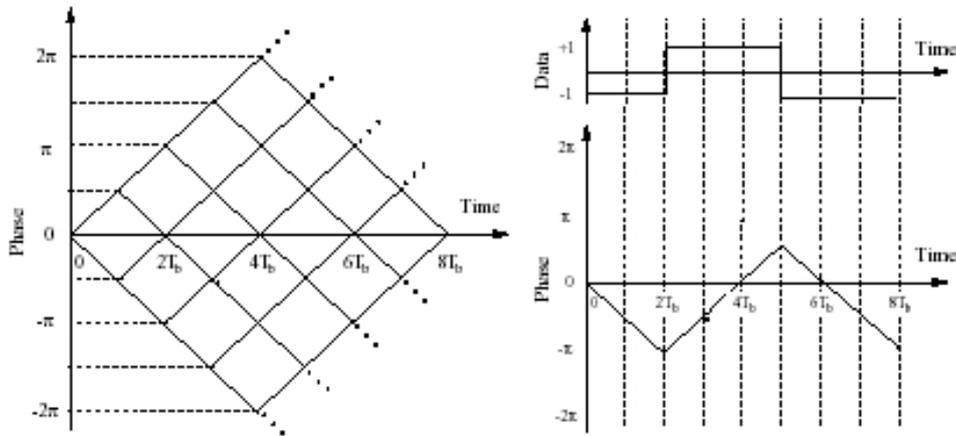


Figura 1.8: Andamento temporale fase del segnale MSK

all'impulso

$$h_G(t) = \frac{\sqrt{\pi}}{\alpha} e^{-\frac{\pi^2}{\alpha^2} t^2} \tag{1.15}$$

rappresentata in Figura(1.9),

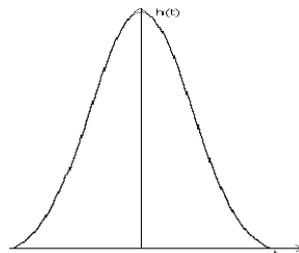


Figura 1.9: Risposta all'impulso del filtro gaussiano

esso è completamente descritto dal prodotto BT dove T è l'intervallo di simbolo e B la banda a 3dB determinata da  $\alpha = \frac{1,8}{B}$ , quanto più BT diminuisce tanto più lo spettro diviene compatto comportando però un aumento dell'interferenza intersimbolica<sup>13</sup>.

La Figura(1.10) illustra un confronto spettrale tra le modulazioni MSK, GMSK e QPSK, lo spettro di quest'ultima è il più stretto per quel che riguarda la banda utile ma presenta molta energia distribuita inutilmente sui lobi secondari, la modulazione più efficiente sotto questo punto di vista è la GMSK che, per tal motivo, è utilizzata nella telefonia cellulare GSM

<sup>13</sup>Sezione(1.3.3)

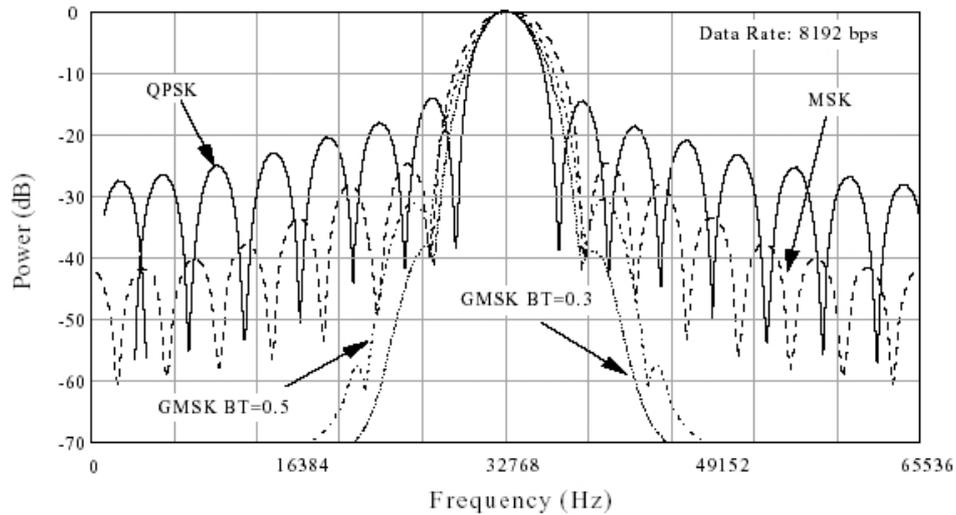


Figura 1.10: Confronto spettri QPSK, MSK e GMSK

dove i canali sono molto vicini tra loro e si richiede elevata autonomia delle batterie.

### 1.3 Principi di demodulazione

Questa Tesi è volta alla realizzazione di un modulatore QPSK, ciò implica tuttavia la conoscenza almeno dei principi della demodulazione che vengono qui brevemente esposti tralasciando il caso in cui il segnale in ingresso al ricevitore presenti una deriva di fase, frequenza o temporizzazione.

Un generico ricevitore digitale è costituito dalla cascata di due circuiti, un demodulatore ed un decisore, il primo si occupa di estrarre dal segnale modulato affetto da rumore le componenti nelle direzioni dei vettori  $\psi_i$  della base dello spazio dei segnali utilizzato, compito poi del decisore è di scegliere quale tra i segnali appartenenti alla costellazione trasmessa sia da associare al segnale ricevuto.

#### 1.3.1 Demodulatore

Il segnale  $r(t)$  in ingresso al ricevitore è la somma del segnale modulato  $s_m(t)$  e del rumore  $n(t)$  introdotto dal canale, quest'ultimo può essere espresso come combinazione lineare dei vettori  $\psi_i$  della base<sup>14</sup>, si ha

<sup>14</sup>la parte di  $n(t)$  che non è rappresentabile come combinazione lineare degli  $\psi_i$  è irrilevante nel processo di decisione

pertanto:

$$\begin{aligned}
 r_j(t) &= \int_0^T r(t) \psi_j(t) dt \\
 &= \int_0^T s_m(t) \psi_j(t) dt + \int_0^T n(t) \psi_j(t) dt \\
 &= s_{mj}(t) + n_j(t)
 \end{aligned} \tag{1.16}$$

questa equazione descrive in maniera immediata il demodulatore che può essere realizzato con un banco di correlatori ciascuno dei quali correla il segnale ricevuto  $r(t)$  con uno dei vettori della base restituendone la proiezione in quella direzione, tale demodulatore è descritto in Figura(1.11).

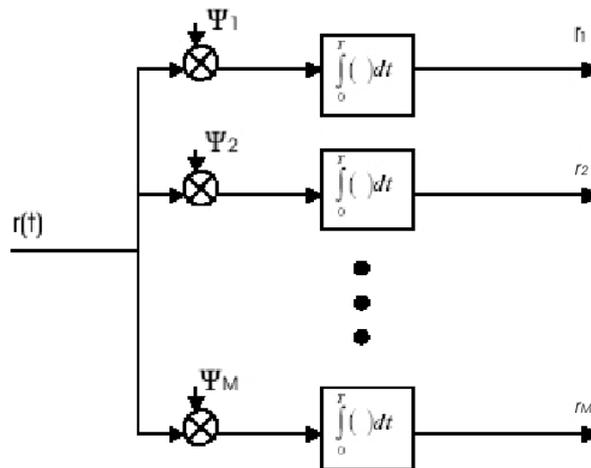


Figura 1.11: Demodulatore con banco di correlatori

Un demodulatore ben più semplice può essere dedotto dalla medesima equazione(1.16) scrivendo la correlazione come una convoluzione:

$$\begin{aligned}
 r_j &= \int_0^T r(\tau) \psi_j(\tau) d\tau \\
 &= \left[ \int_0^T r(\tau) \psi_j(T-t+\tau) d\tau \right] \Big|_{t=T} \\
 &= [r(t) * \psi_j(T-t)] \Big|_{t=T}
 \end{aligned} \tag{1.17}$$

ne consegue che la componente  $r_j(t)$  del segnale ricevuto può essere calcolata campionando al tempo  $t = T$  l'uscita di un filtro avente risposta all'impulso  $h(t) = \psi_j(T-t)$ . Il demodulatore basato su questo banco di filtri che si dicono adattati ai segnali della base è illustrato in Figura(1.12), per esso una considerazione importante è che l'uscita dai filtri adattati è

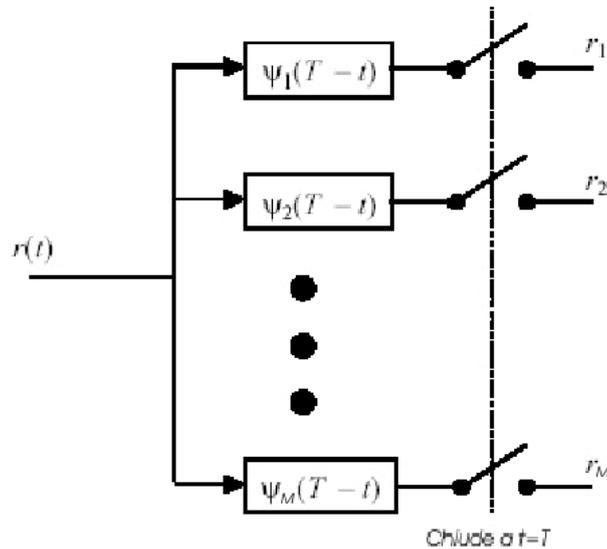


Figura 1.12: Demodulatore con filtro adattato

campionata soltanto a  $t = T$ , essendo  $T$  il tempo di simbolo, e quindi non è richiesto che essa sia uguale all'impulso trasmesso in ogni istante, è sufficiente che lo sia nel solo istante di campionamento.

Il filtro adattato è il demodulatore che consente di ottenere il più alto SNR<sup>15</sup>, la sua risposta in frequenza è infatti:

$$H_{RX}(f) = H_{TX}^*(f) e^{-j2\pi ft} \quad (1.18)$$

che corrisponde a ruotare tutte le componenti del segnale d'ingresso in modo che abbiano la stessa fase come illustrato graficamente in Figura(1.13), si ottiene così che i moduli delle singole componenti del segnale si sommano in maniera costruttiva massimizzando l'energia del segnale mentre quella del rumore rimane invariata.

### 1.3.2 Decisore

L'uscita del demodulatore lungo il  $j$ esimo vettore della base dello spazio dei segnali è somma di una componente deterministica e di una componente aleatoria gaussiana generata dall'elaborazione di un processo di rumore gaussiano, la componente deterministica individua con precisione nello spazio dei segnali il segnale trasmesso mentre la componente gaussiana provoca una deviazione rispetto ad esso la cui ampiezza è individuata dal valore della densità spettrale di rumore  $\frac{N_0}{2}$ .

<sup>15</sup>Signal to Noise Ratio

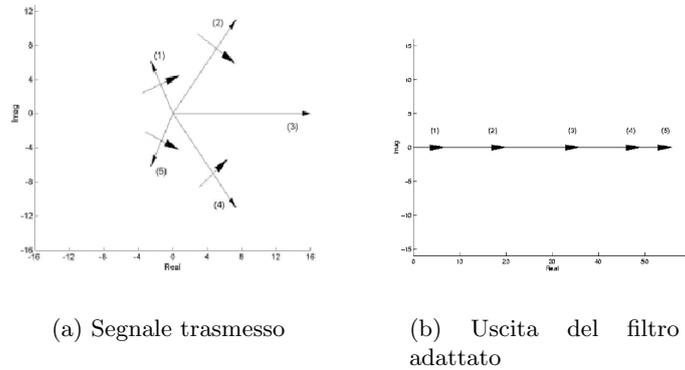


Figura 1.13: Effetto filtro adattato

Il processo di decisione [13] che associa il segnale demodulato  $r_0$  ad uno dei vettori della costellazione può essere implementato secondo il criterio ML<sup>16</sup> oppure il MAP<sup>17</sup>:

**ML:** consiste nello scegliere il vettore  $s_j$  della costellazione che ha la maggior probabilità di coincidere col segnale  $s_m$  emesso.

**MAP:** è uguale al criterio ML ma tiene anche conto del fatto che non necessariamente tutti i segnali appartenenti alla costellazione hanno la stessa probabilità di essere emessi.

Se i segnali della costellazione sono equiprobabili allora i due criteri coincidono e il decisore non fa altro che calcolare le distanze tra il segnale demodulato  $r_0$  e tutti i segnali  $s_j$  della costellazione optando poi a favore del più vicino, in sostanza quindi il decisore suddivide lo spazio dei segnali in  $M$  regioni dette di decisione, ciascuna contiene un punto  $s_j$  della costellazione e tutti i punti che sono più vicini ad esso che non agli altri punti della costellazione. Nel caso si utilizzi il criterio MAP le regioni si ampliano per quei punti della costellazione che hanno maggiore probabilità di essere emessi a discapito delle regioni che contengono i punti con minore probabilità.

Le probabilità d'errore [8] di BPSK e QPSK coincidono in quanto la QPSK è data dall'unione di due costellazioni BPSK ortogonali tra loro, I due segnali  $s_1 = \sqrt{E_b}$  e  $s_2 = -\sqrt{E_b}$  della costellazione BPSK sono rappresentati in Figura(1.14) insieme alla densità di probabilità condizionate:

$$p(r|s_1) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r - \sqrt{E_b})^2}{N_0}} \quad (1.19)$$

<sup>16</sup>Maximum Likelihood

<sup>17</sup>Maximum A Posteriori probability

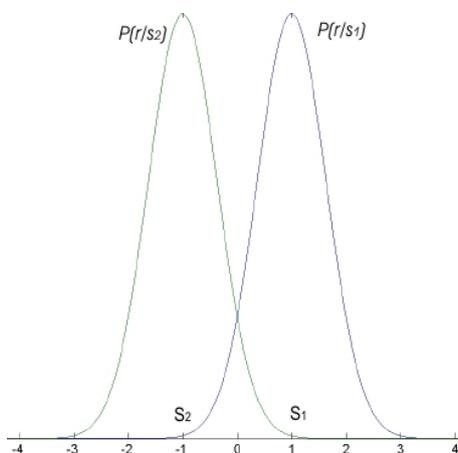


Figura 1.14: Costellazione BPSK e densità di probabilità condizionate

$$p(r|s_2) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r+\sqrt{E_b})^2}{N_0}} \quad (1.20)$$

Nell'ipotesi che sia stato emesso  $s_1(t)$  la probabilità d'errore si ottiene integrando la gaussiana da  $-\infty$  a 0 infatti quella è la regione di decisione associata al simbolo  $s_2$ , si ha:

$$\begin{aligned} P(e|s_1) &= \int_{-\infty}^0 P(r|s_1) dr \\ &= \frac{1}{\sqrt{\pi N_0}} \int_{-\infty}^0 e^{-\frac{(r-\sqrt{E_b})^2}{N_0}} dr \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{\frac{2E_b}{N_0}}} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\sqrt{\frac{2E_b}{N_0}}}^{+\infty} e^{-\frac{x^2}{2}} dx \\ &= Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q\left(\sqrt{\frac{d_{12}^2}{2N_0}}\right) = \frac{1}{2} \operatorname{erfc}\left(\frac{E_b}{N_0}\right) \end{aligned} \quad (1.21)$$

essendo  $\operatorname{erfc}(x)$  la funzione complementare di errore

$$\operatorname{erfc}(x) = \frac{2}{\pi} \int_x^{+\infty} e^{-y^2} dy \quad (1.22)$$

Il medesimo risultato si ha per  $P(e|s_2)$  e, per via della simmetria della costellazione, entrambe coincidono con la probabilità d'errore della mo-

dulazione BPSK e quindi anche della QPSK che tuttavia ha un'efficienza spettrale doppia in quanto ad ogni simbolo associa due bit e non uno.

L'ultima formulazione della (1.21) consente due osservazioni importanti:

1. la probabilità d'errore dipende unicamente dal rapporto  $\frac{E_b}{N_0}$  denominato rapporto segnale/rumore per bit e non da altre caratteristiche del segnale o del rumore.
2. quanto più sono ravvicinati i simboli della costellazione tanto più peggiora la probabilità d'errore, questo è il motivo per il quale alla 16-PSK si preferisce la QPSK anche se l'efficienza spettrale è inferiore.

La BER<sup>18</sup> esprime lo stesso concetto della probabilità d'errore ma in maniera più immediata, essa è infatti il rapporto tra il numero dei bit ricevuti in maniera errata ed il numero totale dei bit trasmessi, una analoga misura è la SER<sup>19</sup> che esprime invece il rapporto tra il numero dei simboli ricevuti in maniera errata ed il numero totale dei simboli trasmessi, esse coincidono per la QPSK a patto di utilizzare la codifica di Gray la quale associa a segnali adiacenti della costellazione gruppi di bit che differiscono tra loro soltanto per un bit.

### 1.3.3 Interferenza intersimbolica

L'equazione delle onde per un mezzo debolmente disomogeneo e non dissipativo

$$\nabla^2 \mathbf{E} + \kappa_0^2 n^2(\mathbf{r}) \mathbf{E} = 0 \quad (1.23)$$

è funzione dell'indice di rifrazione  $n(r) = \sqrt{\epsilon'(\mathbf{r})}$  essendo  $\epsilon'$  la parte reale della costante dielettrica del mezzo.

Ipotizzando per la (1.23) una soluzione espressa nella forma della espansione asintotica di Luneburg-Kline

$$\mathbf{E}(\mathbf{r}) = e^{-j\kappa_0 \Phi(\mathbf{r})} \sum_{m=0}^{\infty} \frac{\mathbf{E}_m(\mathbf{r})}{(j\kappa_0)^m} \quad (1.24)$$

si perviene all'importante risultato secondo il quale nel caso in cui l'indice di rifrazione del mezzo trasmissivo dipende dalla frequenza si ha che un impulso trasmesso subisce un allargamento temporalmente e per il principio della conservazione dell'energia diminuisce in ampiezza [11]. Il fenomeno viene denominato ISI<sup>20</sup> ed è particolarmente deleterio nelle trasmissioni di impulsi digitali reali come quelli tratteggiati in Figura(1.15a), essi vengono deformati come in Figura(1.15b) pertanto negli istanti di campionamento

<sup>18</sup>Bit Error Rate

<sup>19</sup>Symbol Error Rate

<sup>20</sup>Inter Symbol Interference

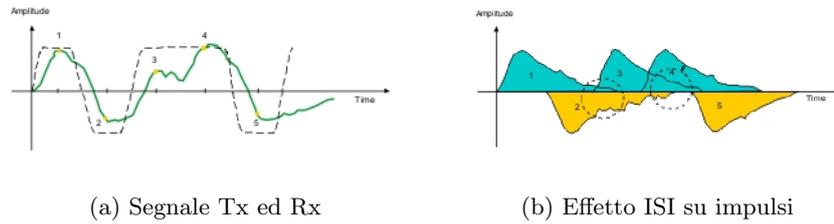


Figura 1.15: Effetto interferenza intersimbolica

del ricevitore si hanno contributi derivanti da più impulsi il che può generare errori come ad esempio nell'istante (3) dove il segnale ricevuto ha ampiezza dimezzata rispetto al simbolo trasmesso.

Il problema dell'interferenza intersimbolica è stato individuato sin dalle prime trasmissioni transoceaniche di Morse, una soluzione intuitiva fu quella di operare al massimo data rate in grado di consentire una ricezione di buona qualità<sup>21</sup>, tuttavia l'evolversi delle tecnologia e la sempre più stringente necessità di trasmettere in tempi brevi grandi moli di informazione ha portato ad una soluzione più ingegnosa.

Ricordando che il ricevitore basato sul filtro adattato richiede l'uguaglianza tra l'impulso trasmesso e quello ricevuto soltanto nell'istante di campionamento ne consegue che si può scegliere una forma dell'impulso qualsiasi a patto che si annulli in tutti i multipli dell'istante di campionamento tranne uno, il suo andamento temporale è pertanto:

$$x(nT) = \begin{cases} 1 & \text{se } n = 0 \\ 0 & \text{se } n \neq 0 \end{cases} \quad (1.25)$$

in virtù della condizione di Nyquist [8] per l'annullamento dell'ISI si ha che gli impulsi  $x(t)$  che rispettano la (1.25) sono tutti quelli la cui trasformata di Fourier risponde al vincolo

$$\sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right) = T \quad (1.26)$$

dove  $W$  è la banda del canale.

I casi possibili sono tre:

$T < \frac{1}{2W}$ : le repliche di  $X(f)$  non si sovrappongono pertanto non c'è modo di rispettare il criterio di Nyquist e quindi l'utilizzo di queste tipologie di impulsi porta a trasmissioni affette da ISI.

$T = \frac{1}{2W}$ : questa condizione implica che le repliche spettrali si tocchino solo in un punto come in Figura(1.17), l'unico spettro che soddisfa questa

<sup>21</sup>circa 200Baud per il telegrafo

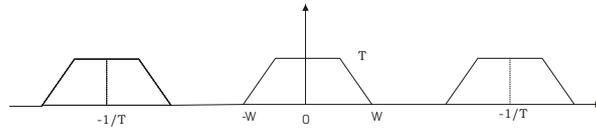


Figura 1.16:  $\sum_{m=-\infty}^{\infty} X(f + \frac{m}{T})$  per  $T < \frac{1}{2W}$

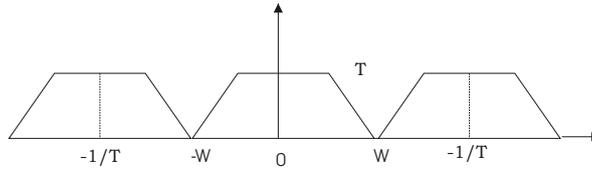


Figura 1.17:  $\sum_{m=-\infty}^{\infty} X(f + \frac{m}{T})$  per  $T = \frac{1}{2W}$

condizione è descritto dalla

$$X(f) = \begin{cases} T & \text{se } |f| < W \\ 0 & \text{altrimenti} \end{cases} \quad (1.27)$$

si tratta di un rettangolo in frequenza cui corrisponde la

$$x(t) = \frac{\sin(\pi \frac{t}{T})}{\pi \frac{t}{T}} \equiv \text{sinc}\left(\pi \frac{t}{T}\right) \quad (1.28)$$

rappresentata in Figura(1.18), essa presenta la banda minima ma è praticamente irrealizzabile per via della anticausalità e delle lunghe code che si estendono illimitatamente e decadono come  $\frac{1}{x}$ .

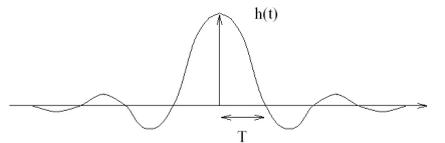


Figura 1.18:  $\text{sinc}\left(\pi \frac{t}{T}\right)$

$T > \frac{1}{2W}$ : le repliche spettrali in questo caso si sovrappongono, vi sono quindi molti impulsi che rispettano il criterio di Nyquist, la famiglia più

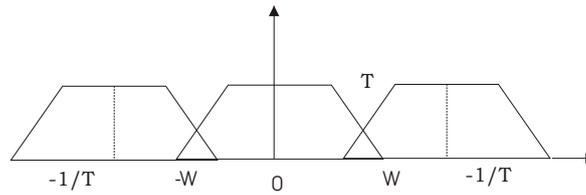


Figura 1.19:  $\sum_{m=-\infty}^{\infty} X(f + \frac{m}{T})$  per  $T > \frac{1}{2W}$

utilizzata è quella degli spettri a coseno rialzato descritti dalla

$$H_{rc}(f) = \begin{cases} T & \text{se } 0 \leq |f| \leq \frac{(1+\alpha)}{2T} \\ \frac{T}{2} \left[ 1 + \cos \frac{\pi T}{\alpha} \left( |f| - \frac{(1-\alpha)}{2T} \right) \right] & \text{se } \frac{(1+\alpha)}{2T} \leq |f| \leq \frac{(1-\alpha)}{2T} \\ 0 & \text{se } 0 \leq |f| \geq \frac{(1-\alpha)}{2T} \end{cases} \quad (1.29)$$

dove  $\alpha$  è il fattore di roll-off ed assume valori compresi tra 0 ed 1, in particolare per  $\alpha = 0$  la banda occupata dall'impulso è proprio la minima di Nyquist<sup>22</sup>, al crescere di  $\alpha$  essa aumenta sino a raddoppiare per  $\alpha = 1$ . L'individuazione degli spettri a coseno rialzato parte dalla antitrasformata di Fourier:

$$h(t) = \frac{\sin\left(\frac{2\pi t}{T_s}\right) \cos(2\pi \alpha t)}{2\pi T \left[ 1 - \left(\frac{2\alpha t}{\pi}\right)^2 \right]} \quad (1.30)$$

che evidenzia come la *sinc* viene corretta con il coseno al fine di migliorarne le caratteristiche, rilevante l'osservazione che le code dell'impulso sagomato pur continuando ad estendersi illimitatamente decadono come  $\frac{1}{x^3}$  pertanto molto più rapidamente che non nel caso in cui l'impulso sagomato sia una *sinc*.

Nelle applicazioni pratiche si utilizza la sagomatura d'impulso con spettro a coseno rialzato, la risposta in frequenza e nel tempo è illustrata al variare di  $\alpha$  in Figura(1.20) si può osservare come quanto più  $\alpha$  tende a zero, tanto più hanno ampiezza elevata i lobi laterali, ciò determina forti escursioni dell'involuppo complesso come mostrato in Figura(1.21), pertanto per consentire un utilizzo ottimale degli amplificatori, nelle applicazioni spaziali si utilizza il valore  $\alpha = 0,35$ .

Alla luce dei concetti esposti per il filtro adattato, si comprende come sia sensato suddividere la sagomatura a coseno rialzato tra il trasmettitore ed il ricevitore secondo la

$$H_{rc}(f) = H_{Tx}(f) H_{Rx}(f) \quad (1.31)$$

<sup>22</sup> ossia  $\frac{R_s}{2}$

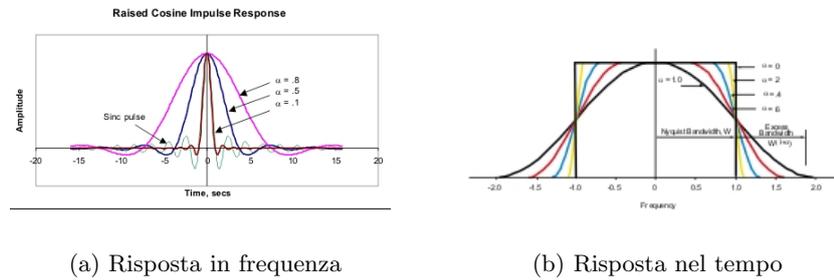


Figura 1.20: Sagomatura d'impulso a coseno rialzato

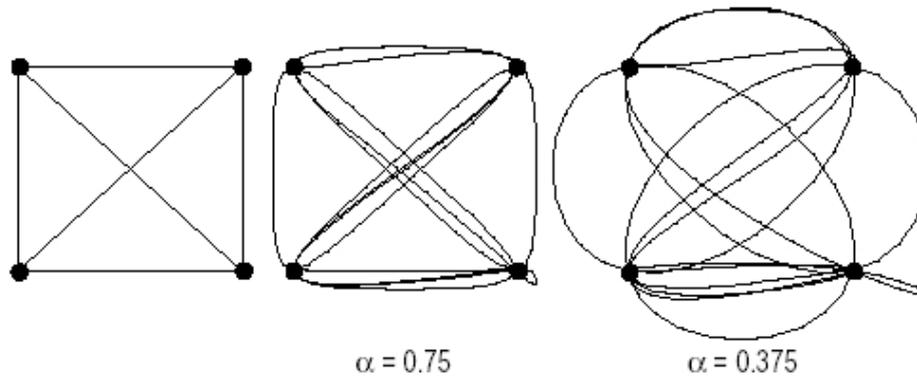


Figura 1.21: Diagrammi vettoriali QPSK con sagomatura d'impulso

in entrambe sono infatti necessari dei filtri<sup>23</sup> che possono essere progettati in modo da eseguire anche la sagomatura dell'impulso, il modulo della risposta in frequenza per entrambe gli SRRC<sup>24</sup> sarà pertanto:

$$|H_{Tx}(f)| = |H_{Rx}(f)| = \sqrt{H_{rc}(f)} \quad (1.32)$$

### 1.3.4 Principi di codifica

Tralasciando l'eventuale conversione in digitale di una sorgente analogica, si ha che una generica trasmissione numerica può essere realizzata come in figura(1.22) dove la codifica di sorgente si propone di rimuovere le dipendenze tra i simboli da trasmettere al fine di ridurre il numero come nel caso degli algoritmi che nei computer vengono utilizzati per la compressione dei file.

<sup>23</sup>il filtro in trasmissione riduce le emissioni nelle bande adiacenti mentre quello in ricezione effettua la demodulazione ed elimina il rumore esterno alla banda utile.

<sup>24</sup>Square Root Raised Cosine

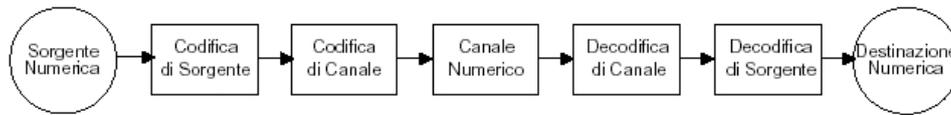


Figura 1.22: Trasmissione numerica

La codifica di canale si basa sull'introduzione di simboli addizionali nella sequenza trasmessa, essi sono scelti in modo da presentare una qualche correlazione tra loro consentendo così la rivelazione degli errori e conseguentemente la riduzione della probabilità di errore intrinseca del canale. Le principali tecniche di correzione degli errori sono due, la ARQ<sup>25</sup> e la FEC<sup>26</sup>:

**ARQ:** suddivide in blocchi l'informazione da trasmettere, se in ricezione si individua la presenza di errori in un blocco ne viene richiesta la ritrasmissione, necessita di un canale duplex ed ha lo svantaggio che anche un singolo errore comporta la ritrasmissione dell'intero blocco in cui è contenuto.

**FEC:** non richiede un canale duplex ma il numero di simboli ridondanti da aggiungere è molto maggiore rispetto all'ARQ, ne consegue che il data rate effettivo si riduce ma, per il teorema di Shannon sulla [Capacità trasmissiva](#)<sup>27</sup> si riduce anche la potenza necessaria per ottenere una trasmissione numerica praticamente esente da errori.

In ambito satellitare si utilizza la codifica di canale di tipo FEC, in una o più delle seguenti forme:

**Codifica a blocchi:** la più importante è la Reed-Salomon (204,188), essa aggiunge 16 bytes di ridondanza per ogni 188 bytes da trasmettere, in tal modo si possono correggere sino a 8 errori e la BER richiesta passa da  $10^{-4}$  a più di  $10^{-10}$ . Il principio è che con l'aggiunta di bit si ottengono delle combinazioni le quali possono essere ricevute soltanto in caso di errore, una volta che esso è stato rilevato si sceglie nell'alfabeto dei possibili simboli trasmessi quello che presenta la distanza di Hamming<sup>28</sup> minore rispetto al simbolo ricevuto.

**Codifica convoluzionale:** viene specificata con i tre parametri (n,k,m) dove k è il numero di bit in ingresso ed n il numero di bit in uscita dal codificatore, entrambe assumono valori compresi tra 1 ed 8, m è invece il numero di registri utilizzati ed è compreso tra 2 e 10.

<sup>25</sup> Automatic ReQuest for repeat

<sup>26</sup> Forward Error Correction

<sup>27</sup> Sezione(1.3.5)

<sup>28</sup> è il numero di bit diversi tra due simboli

Un esempio di codificatore è riportato in Figura(1.23), esso produce 3

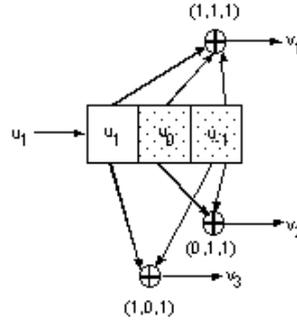


Figura 1.23: Codificatore (3, 1, 3)

bit di uscita per ogni bit di ingresso quindi il ritmo di codifica è  $1/3$ , ogni bit d'uscita è generato quale somma di alcuni dei bit presenti negli  $m = 3$  registri di memoria, il criterio di selezione di questi polinomi generatori determina la qualità della codifica. Laddove si desideri un ritmo di codifica variabile si ricorre alla versione *Punctured*, in particolare se si richiede la massima qualità vengono utilizzati tutte le  $n$  uscite del codificatore convoluzionale, altrimenti soltanto una parte di esse.

**Interleaving:** le codifiche precedenti hanno prestazioni eccellenti ma in presenza di un gruppo di errori ravvicinati non sono in grado di ricostruire la sequenza trasmessa, il problema si risolve mescolando i bit della sequenza da trasmettere ad esempio caricando una ROM lungo le righe e leggendone il contenuto lungo le colonne.

### 1.3.5 Capacità trasmissiva

Per confrontare differenti tipi di modulazione e comprendere come l'effetto dei codici possa migliorare la qualità di una trasmissione numerica è molto utile il teorema [3] seguente:

**Teorema 1 (di Shannon)** *Un qualsiasi canale caratterizzato da una velocità di segnalazione  $R_s = \frac{1}{T_s}$ , una varianza<sup>29</sup> di rumore al ricevitore  $\sigma_S^2$  ed una varianza del segnale ricevuto  $\sigma_N^2$  pone un limite, denominato capacità di canale  $C$ , al massimo flusso informativo che transita in esso:*

$$C = \frac{1}{2T_s} \log_2 \left( 1 + \frac{\sigma_S^2}{\sigma_N^2} \right) \quad (\text{bit/s}) \quad (1.33)$$

<sup>29</sup>potenza

nel caso di rumore gaussiano additivo bianco con densità spettrale di potenza  $S_W(f) = \frac{N_0}{2}$  si ha che la potenza di rumore è  $P_W = \sigma_W^2 = N_0B$  inoltre se la sorgente è gaussiana e limitata nella banda B può essere rappresentata con  $2B$  campioni<sup>30</sup> ciascuno con la medesima potenza  $P_S = \sigma_S^2$ , ne deriva che l'equazione(1.33) può essere riscritta nella forma:

$$C = B \log_2 \left( 1 + \frac{P_S}{N_0B} \right) \quad (\text{bit/s}) \quad (1.34)$$

dalla quale si deduce che a parità di rumore se si vuole aumentare la capacità trasmissiva di un canale conviene aumentare la banda B, un aumento della potenza ha infatti minore effetto essendo mitigato dal logaritmo.

E' interessante una rappresentazione grafica dell'equazione(1.34) nel cosiddetto *piano di Shannon*, si giunge ad essa partendo dall'ipotesi che il ritmo binario in trasmissione sia uguale alla capacità del canale, si ha cioè  $C = R_b = \frac{1}{T_b}$  pertanto

$$\frac{C}{B} = \log_2 \left( 1 + \frac{R_b P_S T_b}{N_0 B} \right) = \log_2 \left( 1 + \frac{C P_S T_b}{B N_0} \right) = \log_2 \left( 1 + \frac{C E_b}{B N_0} \right) \quad (1.35)$$

da cui si ottiene

$$\frac{E_b}{N_0} = \frac{2^{\frac{C}{B}} - 1}{\frac{C}{B}} \quad (1.36)$$

che rappresenta la curva limite nel piano di Shannon illustrato in Figura(1.24), la regione sottostante è quella permessa ed in essa giacciono i punti corrispondenti alle varie modulazioni, quanto più essi si avvicinano alla curva limite riuscendo a mantenere basso il rapporto  $\frac{E_b}{N_0}$  tanto più la modulazione è efficiente.

I punti corrispondenti alle modulazioni si ottengono imponendo per tutte una determinata probabilità d'errore,  $10^{-5}$  per il piano in Figura(1.24), ad essa nel caso della QPSK corrisponde  $\frac{E_b}{N_0} = 9,6\text{dB}$  mentre l'efficienza spettrale è  $\eta_S = \frac{R_b}{B} = 2$  in quanto ad ogni coppia di bit si associa un simbolo, le coordinate del punto sono pertanto (9,6 , 2). La modulazione BPSK presenta lo stesso  $\frac{E_b}{N_0}$  e quindi la stessa efficienza energetica ma l'efficienza spettrale vale 1 pertanto sotto questo punto di vista è più lontana dalla curva limite e quindi le sue prestazioni sono inferiori. Nel caso di una QPSK con codifica convoluzionale, caratterizzata da un ritmo di codifica 1/2 e  $k = 7$  bit, si ha un guadagno di codifica di 5,2dB rispetto alla QPSK non codificata mentre l'efficienza spettrale vale 1 pertanto le coordinate del punto sono (4,4 , 1), la forte vicinanza alla curva limite giustifica il crescente interesse per i *turbocodici*. E' interessante osservare come le modulazioni di

<sup>30</sup>per il teorema del campionamento

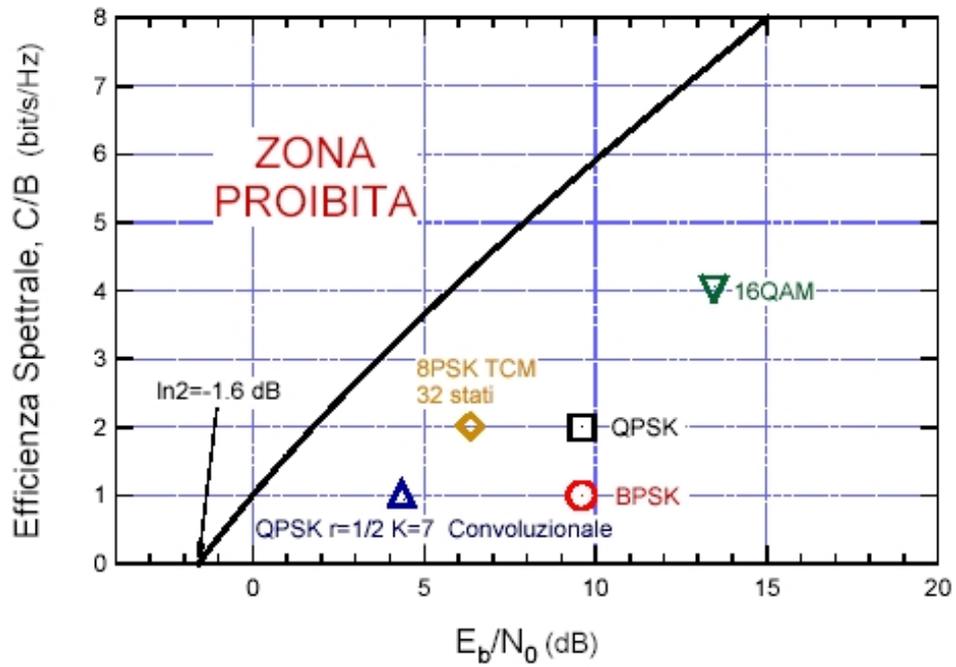


Figura 1.24: Piano di Shannon

tipo M-QAM con costellazione quadrata incrementano l'efficienza spettrale della trasmissione, ma penalizzano l'efficienza energetica di circa 6 dB per ogni quadruplicazione dei punti della costellazione.