Appunti di Trasmissione Numerica I
Parte II - Spazio dei Segnali e Demodulazione a Minima Probabilità di Errore

Stefano Buzzi, Maurizio di Bisceglie[1]

Maggio 2004

[1]Stefano Buzzi è docente di ruolo presso l'Università degli Studi di Cassino, Dipartimento di Automazione, Elettromagnetismo, Matematica Industriale e Ingegneria dell'Informazione (DAEIMI), Via G. Di Biasio, 43, I-03043 Cassino (FR), Italia. E-mail: buzzi@unicas.it. Maurizio di Bisceglie è docente di ruolo presso l'Università degli Studi del Sannio, Facoltà di Ingegneria, Palazzo Bosco Lucarelli, C.so Garibaldi, 107, 82100 Benevento, Italia. E-mail: dibisceglie@unisannio.it

Lo Spazio dei Segnali

Nella prima parte di tali appunti ci si è soffermati sulla descrizione dello schema generale di un sistema di trasmissione numerico, e sulla descrizione dei blocchi sorgente, codificatore di sorgente e canale. In questa seconda parte analizzeremo con maggior dettaglio il blocco modulatore e demodulatore, assumendo che il canale interposto tra modulatore e demodulatore sia di tipo AWGN, ovvero si limiti ad introdurre un rumore additivo Gaussiano bianco.

Si è già detto come il modulatore numerico, nel caso che sia senza memoria[1], può essere caratterizzato da un insieme di M forme d'onda di energia, che indichiamo con l'insieme $\mathcal{S} = \{s_1(t), \ldots, s_M(t)\}$, di durata usualmente non superiore all'intervallo di segnalazione T, e da una legge biunivoca che assegna a ciascuna delle M forme d'onda una stringa di $k = \log_2 M$ simboli binari. Al fine di classificare in maniera opportuna il set $\mathcal{S}$ di forme d'onda utilizzato dal modulatore, è necessario introdurre il concetto di *spazio dei segnali* e di *dimensionalità*. All'uopo, è utile introdurre un'analogia che esiste tra lo spazio vettoriale $\mathcal{R}^N$, che è l'insieme di tutte le N-uple di numeri reali, ovvero di tutti i vettori reali N-dimensionali, e l'insieme delle funzioni al quadrato sommabile, tra i cui elementi sono comprese anche le forme d'onda dell'insieme $\mathcal{S}$.

1 Lo spazio vettoriale $\mathcal{R}^N$

Si ricordi che lo spazio vettoriale $\mathcal{R}^N$ è l'insieme delle N-uple di numeri reali. Indicheremo con lettere minuscole ed in grassetto gli elementi di tale spazio, ovvero è $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{R}^N$. Per tale spazio vettoriale valgono le seguenti proprietà.

1. L'insieme $\mathcal{R}^N$ è chiuso rispetto alle operazioni di somma e di moltiplicazione per uno scalare. In altri termini, se $\boldsymbol{x}$ e $\boldsymbol{y}$ sono elementi di $\mathcal{R}^N$, anche $\boldsymbol{x} + \boldsymbol{y}$ è un elemento di $\mathcal{R}^N$, ove $\boldsymbol{x} + \boldsymbol{y}$ è il vettore che si ottiene sommando componente per componente gli elementi di $\boldsymbol{x}$ e $\boldsymbol{y}$. Allo stesso modo, se $\alpha \in \mathcal{R}$, e $\boldsymbol{x} \in \mathcal{R}^N$, allora $\alpha\boldsymbol{x}$, che è il vettore che si ottiene da $\boldsymbol{x}$ moltiplicando ogni sua componente per lo scalare α, è un elemento di $\mathcal{R}^N$.

[1] In tali appunti saranno esaminati esclusivamente modulatori senza memoria, ad eccezione dei modulatori differenziali di cui si dirà nel seguito.

2. Sullo spazio $\mathcal{R}^N$ è definita un'operazione di prodotto scalare, che è una funzione che associa a coppie di elementi di $\mathcal{R}^N$ un numero reale:

$$< \cdot, \cdot > : \ (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{R}^N \times \mathcal{R}^N \ \longrightarrow < \boldsymbol{x}, \boldsymbol{y} > \in \mathcal{R} \ . \tag{1}$$

Indicando con x_i e y_i l'i-esima componente dei vettori $\boldsymbol{x}$ e $\boldsymbol{y}$, rispettivamente, il prodotto scalare in $\mathcal{R}^N$ è usualmente definito come

$$< \boldsymbol{x}, \boldsymbol{y} >= \sum_{i=1}^{N} x_i y_i = \boldsymbol{x}^T \boldsymbol{y} \ , \tag{2}$$

ove $(\cdot)^T$ indica trasposizione.

3. Il prodotto scalare gode delle seguenti proprietà.

   p1) $< \boldsymbol{x}, \boldsymbol{y} >=< \boldsymbol{y}, \boldsymbol{x} >$

   p2) $< \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} >=< \boldsymbol{x}, \boldsymbol{z} > + < \boldsymbol{y}, \boldsymbol{z} >$

   p3) $< \alpha \boldsymbol{x}, \boldsymbol{y} >= \alpha < \boldsymbol{x}, \boldsymbol{y} >$

   p4) $< \boldsymbol{x}, \boldsymbol{x} >\geq 0$

   p5) $< \boldsymbol{x}, \boldsymbol{x} >= 0$ se e solo se $\boldsymbol{x}$ è il vettore nullo.

4. A partire dal prodotto scalare, è possibile introdurre il concetto di ortogonalità. In particolare, diremo che $\boldsymbol{x}$ e $\boldsymbol{y}$ sono ortogonali quando il loro prodotto scalare è nullo.

5. Il prodotto scalare permette inoltre di introdurre anche una norma in $\mathcal{R}^N$; in particolare, la norma del vettore $\boldsymbol{x}$ si denota con $\|\boldsymbol{x}\|$ e si definisce come

$$\|\boldsymbol{x}\| = \sqrt{< \boldsymbol{x}, \boldsymbol{x} >} \ . \tag{3}$$

   Si noti che tale definizione coincide con l'usuale norma euclidea di $\mathcal{R}^N$.

6. E' possibile inoltre definire anche una distanza tra elementi di $\mathcal{R}^N$; in particolare, la distanza tra i vettori $\boldsymbol{x}$ e $\boldsymbol{y}$ è espressa come

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| =< \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} >^{1/2} \tag{4}$$

7. Al concetto di ortogonalità e di norma sono poi collegati il concetto di lineare indipendenza di un insieme di vettori, e la definizione del concetto di *base* e *versore*. Per brevità non ci dilunghiamo su tali aspetti e assumiamo che essi siano noti al lettore.

2 Lo spazio delle funzioni al quadrato sommabile in $[0, T]$

Consideriamo l'insieme di tutte le funzioni al quadrato sommabile in $[0, T]$. Tale insieme, che usualmente si denota col simbolo $L^2_{[0,T]}$, nella terminologia della teoria dei segnali è indicato come l'insieme dei segnali di energia a supporto in $[0, T]$. E' alquanto immediato verificare che $L^2_{[0,T]}$, analogamente allo spazio vettoriale $\mathcal{R}^N$, gode delle proprietà di chiusura rispetto alle operazioni di somma e di moltiplicazione per uno scalare. Inoltre, in tale insieme è possibile definire un prodotto scalare, ovvero un'opportuna funzione cha associa a coppie di funzioni al quadrato sommabili un numero reale. In particolare, se si considera la funzione

$$< \cdot, \cdot > : \ (a(t), b(t)) \ \in L^2_{[0,T]} \times L^2_{[0,T]} \ \longrightarrow < a(t), b(t) >= \int_0^T a(t)b(t)dt \ , \tag{5}$$

è facile rendersi conto che essa soddisfa tutte le proprietà, da p1) a p5) introdotte nel paragrafo precedente, ragion per cui la legge di corrispondenza in (5) può essere ritenuta un prodotto scalare. Dalla definizione (5) è poi immediato far discendere una definizione di norma; in particolare, detto $a(t)$ un elemento di $L^2_{[0,T]}$, la norma di $a(t)$ è data da

$$\|a(t)\| = \left[\int_0^T a^2(t)dt\right]^{1/2} \ . \tag{6}$$

Sulla base di tale definizione la norma di una funzione coincide con la radice quadrata della sua energia. Analogamente, diremo che due elementi $a(t)$ e $b(t)$ di $L^2_{[0,T]}$, sono ortogonali quando il loro prodotto scalare è nullo. Dalla definizione (5) si deduce quindi che due funzioni $a(t)$ e $b(t)$ che non siano contemporaneamente non nulle sono ortogonali. Inoltre, dalla relazione di Parseval

$$\int a(t)b(t)dt = \int A(f)B(f)df \ , \tag{7}$$

si deduce che due segnali $a(t)$ e $b(t)$ sono ortogonali se le rispettive trasformate di Fourier, $A(f)$ e $B(f)$, sono non sovrapposte in frequenza.

Definito un prodotto scalare in $L^2_{[0,T]}$, è possibile individuare una base per ogni insieme di elementi di $L^2_{[0,T]}$. In particolare, consideriamo l'insieme $\mathcal{S} = \{s_1(t), \ldots, s_M(t)\}$ delle M forme d'onda utilizzate da un modulatore. A partire da tale insieme, è possibile applicare la procedura di ortogonalizzazione di Gram-Smidth al fine di giungere ad un insieme di versori, ovvero di forme d'onda mutuamente ortogonali e a norma unitaria, che indichiamo come $\mathcal{B} = \{\psi_1(t), \ldots, \psi_L(t)\}$, tali che gli elementi di $\mathcal{S}$ possano ottenersi come combinazione lineare degli elementi della base $\mathcal{B}$. Altrimenti detto, la procedura di ortogonalizzazione conduce ad un insieme di L forme d'onda, $\psi_1(t), \psi_2(t), \ldots, \psi_L(t)$, ove $L \leq M$, e tali che

$$< \psi_i(t), \psi_j(t) >= \int_0^T \psi_i(t)\psi_j(t)dt = \delta_{i,j} = \begin{cases} 1 & \text{se} \quad i = j \ , \\ 0 & \text{se} \quad i \neq j \ . \end{cases} \tag{8}$$

Tali forme d'onda rappresentano una base per l'insieme dei segnali in $\mathcal{S}$, in quanto ciascun elemento di $\mathcal{S}$ può essere espresso tramite una combinazione lineare degli elementi di $\mathcal{B}$ tramite la relazione

$$s_i(t) = \sum_{\ell=1}^{L} < s_i(t), \psi_\ell(t) > \psi_\ell(t) \ . \tag{9}$$

Si noti che il prodotto scalare tra una generica funzione ed un versore è usualmente detto *proiezione*, ossia il prodotto scalare $< s_i(t), \psi_\ell(t) >$ rappresenta la proiezione di $s_i(t)$ lungo $\psi_\ell(t)$. La cardinalità dell'insieme $\mathcal{B}$, ovvero il parametro L, è detto *dimensionalità* dell'insieme dei segnali del modulatore, o anche *dimensionalità* della costellazione dei segnali del modulatore. L'insieme di tutte le possibili combinazioni lineari degli elementi di $\mathcal{B}$, che ovviamente coincide con l'insieme di tutte le possibili combinazioni lineari degli elementi di $\mathcal{S}$, è detto *spazio dei segnali*. E' semplice verificare che lo spazio dei segnali è un sottospazio L-dimensionale dello spazio $L^2_{[0,T]}$.

La relazione (9) dimostra che la generica forma d'onda del modulatore $s_i(t)$ può essere ottenuta a partire dalla conoscenza delle funzioni di base e delle proiezioni di $s_i(t)$ su tali funzioni di base. Alla generica forma d'onda $s_i(t)$ può quindi essere associato il vettore L-dimensionale delle sue proiezioni lungo le funzioni di base. Altrimenti detto, indichiamo con $\boldsymbol{s}_i$ il seguente vettore colonna L-dimensionale

$$\boldsymbol{s}_i = \begin{bmatrix} < s_i(t), \psi_1(t) > \\ < s_i(t), \psi_2(t) > \\ \vdots \\ < s_i(t), \psi_L(t) > \end{bmatrix} = \begin{bmatrix} s_i(1) \\ s_i(2) \\ \vdots \\ s_i(L) \end{bmatrix} \ . \tag{10}$$

Diremo che $\boldsymbol{s}_i$ è il vettore delle componenti del segnale $s_i(t)$. Ad ogni funzione appartenente allo spazio dei segnali L-dimensionale può quindi essere associato un elemento dello spazio vettoriale $\mathcal{R}^L$. Tale corrispondenza preserva il prodotto scalare. Infatti, si ha:

$$\begin{aligned} < s_i(t), s_j(t) >= \int_0^T s_i(t)s_j(t)dt = \int_0^T \sum_{m=1}^{L} s_i(m)\psi_m(t) \sum_{n=1}^{L} s_j(n)\psi_n(t)dt = \\ \sum_{m=1}^{L}\sum_{n=1}^{L} s_i(m)s_j(n) \underbrace{\int_0^T \psi_n(t)\psi_m(t)dt}_{\delta_{m,n}} = \sum_{m=1}^{L} s_i(m)s_j(m) =< \boldsymbol{s}_i, \boldsymbol{s}_j > \ . \end{aligned} \tag{11}$$

In altre parole, il prodotto scalare tra due forme d'onda appartenenti allo spazio dei segnali coincide numericamente col prodotto scalare eseguito in $\mathcal{R}^L$ tra i vettori L-dimensionali associati alle due forme d'onda. Ovviamente, la conservazione del prodotto scalare ha come diretta conseguenza la conservazione della norma e della distanza. Infatti, è immediatamente verificabile che

$$\mathcal{E}_i = \|s_i(t)\|^2 = \|\boldsymbol{s}_i\|^2 \ , \tag{12}$$

come pure che

$$d(s_i(t), s_j(t)) = \|s_i(t) - s_j(t)\| = \|\boldsymbol{s}_i - \boldsymbol{s}_j\| = d(\boldsymbol{s}_i, \boldsymbol{s}_j) \ . \tag{13}$$

Esempio

Si considerino i due segnali $s_1(t)$ ed $s_2(t)$ ad energia finita definiti in $(0, T)$ mostrati in fig. 1. Calcoliamo le proiezioni dei segnali sul riferimento individuato dai due versori $\psi_1(t)$ e $\psi_2(t)$ riportati nella fig. 2. Verifichiamo prima che $\psi_1(t)$ e $\psi_2(t)$ costituiscono una base

$$||\psi_1(t)||^2 = \int_0^T \frac{1}{T}\, dt = 1 \qquad ||\psi_2(t)||^2 = 2\int_0^{T/2} \frac{1}{T}\, dt = 1$$

$$< \psi_1(t), \psi_2(t) >= \int_0^{\frac{T}{2}} \frac{1}{T}\, dt - \int_{\frac{T}{2}}^T \frac{1}{T}\, dt = 0$$

Le componenti dei segnali $s_1(t)$ ed $s_2(t)$ corrispondono ai prodotti scalari di questi ultimi con i versori $\psi_1(t)$ e $\psi_2(t)$. Infatti

$$s_1(1) =< s_1(t), \psi_1(t) >= \int_0^{\frac{T}{2}} \frac{A}{\sqrt{T}}\, dt = \frac{A\sqrt{T}}{2}$$

$$s_1(2) =< s_1(t), \psi_2(t) >= \int_0^{\frac{T}{2}} \frac{A}{\sqrt{T}}\, dt = \frac{A\sqrt{T}}{2}$$

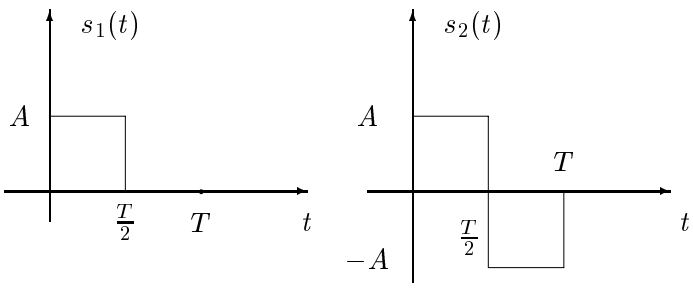

Figure 1: Grafico dei segnali $s_1(t)$ ed $s_2(t)$

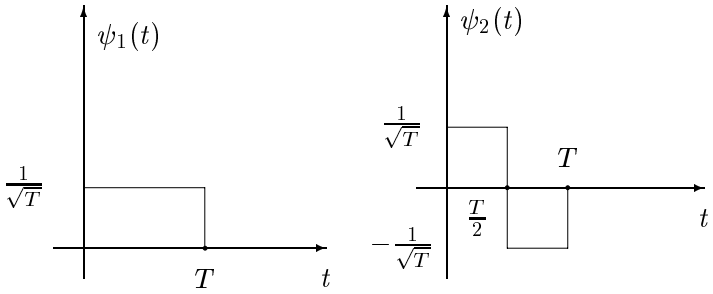

Figure 2: Grafico dei versori $\psi_1(t)$ e $\psi_2(t)$

Si noti che, data la base costituita dai segnali ortonormali $\psi_1(t)$ e $\psi_2(t)$, il segnale $s_1(t)$ è completamente rappresentato dalla coppia di proiezioni $(s_1(1), s_1(2))$ e può essere espresso tramite la somma

$$s_1(t) = s_1(1)\psi_1(t) + s_1(2)\psi_2(t) = \frac{A\sqrt{T}}{2}\psi_1(t) + \frac{A\sqrt{T}}{2}\psi_2(t)$$

La rappresentazione del segnale $s_1(t)$, tramite i versori $\psi_1(t)$ e $\psi_2(t)$ è mostrata in fig. 3. Il discorso si ripete in maniera perfettamente analoga per il segnale $s_2(t)$.

L'esempio precedente mostra come, una volta nota una base per un dato insieme di segnali, è possibile rappresentare i segnali per mezzo delle sole proiezioni sui versori della base. Ci proponiamo ora, date le M forme d'onda del modulatore $\{s_i(t)\}$, $i = 1, \ldots, M$, di determinare una base ortonormale $\{\psi_i(t)\}$, $i = 1, \ldots, L$ per la rappresentazione dei suddetti segnali. La soluzione del problema è data dal procedimento costruttivo di Gram-Schmidt. Mostreremo che $L \leq M$; se $L = M$ si dice che gli M

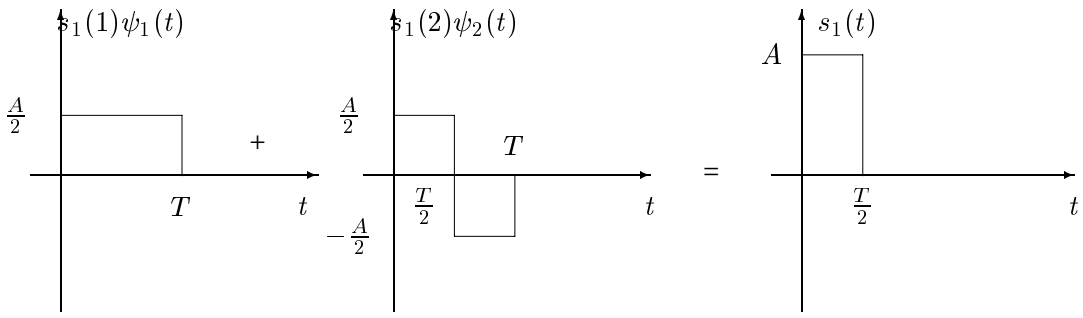

Figure 3: Rappresentazione del segnale $s_1(t)$

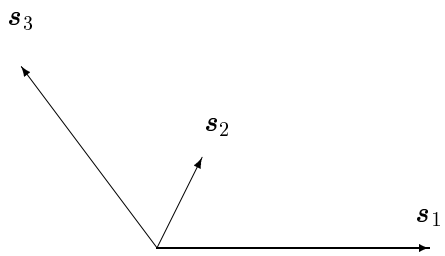

Figure 4: Vettori co-planari

segnali $\{s_i(t)\}$ sono linearmente indipendenti. Una condizione sufficiente (ma non necessaria) affinché le M forme d'onda siano linearmente indipendenti è che esse siano ortogonali.

Tuttavia, prima di illustrare la procedura di ortogonalizzazione con riferimento allo spazio $L^2_{[0,T]}$, riportiamo un esempio di applicazione con riferimento allo spazio vettoriale $\mathcal{R}^3$. Consideriamo quindi i tre vettori co-planari $\boldsymbol{s}_1$, $\boldsymbol{s}_2$ e $\boldsymbol{s}_3$ mostrati in fig. 4. L'algoritmo si articola in una serie di passi.

1.▷ Si definisce un vettore ausiliario $\boldsymbol{d}_1 = \boldsymbol{s}_1$ e si costruisce il primo vettore della base parallelo ad $\boldsymbol{s}_1$ e a norma unitaria

$$\boldsymbol{\Psi}_1 = \frac{\boldsymbol{s}_1}{\|\boldsymbol{s}_1\|}$$

2.▷ Sottraendo da $\boldsymbol{s}_2$ la sua proiezione su $\boldsymbol{\Psi}_1$ otteniamo un vettore ortogonale a $\boldsymbol{\Psi}_1$ (cfr. fig. 5).

$$\boldsymbol{d}_2 = \boldsymbol{s}_2 - s_2(1)\boldsymbol{\Psi}_1$$

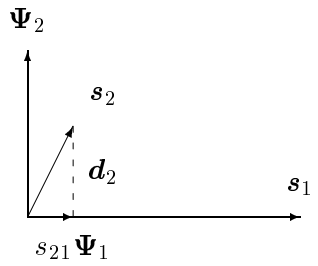

Figure 5: Generazione del vettore $\boldsymbol{d}_2$

che normalizzato fornisce il secondo versore della base

$$\boldsymbol{\Psi}_2 = \frac{\boldsymbol{d}_2}{\|\boldsymbol{d}_2\|}$$

3.▷ L'iterazione procede costruendo un terzo vettore $\boldsymbol{d}_3$ ottenuto sottraendo ad $\boldsymbol{s}_3$ le sue proiezioni lungo $\boldsymbol{\Psi}_1$ e $\boldsymbol{\Psi}_2$

$$\boldsymbol{d}_3 = \boldsymbol{s}_3 - s_3(1)\boldsymbol{\Psi}_1 - s_3(2)\boldsymbol{\Psi}_2$$

come mostrato in fig. 6. Nel caso specifico, essendo i due vettori co-planari risulta

$$\boldsymbol{s}_3 - (s_3(1)\boldsymbol{\Psi}_1 + s_3(2)\boldsymbol{\Psi}_2) = 0$$

In generale, il procedimento consiste nel calcolare il vettore ausiliario

$$\boldsymbol{d}_i = \boldsymbol{s}_i - \sum_{n=1}^{i-1} s_i(n)\boldsymbol{\Psi}_n$$

che normalizzato fornisce

$$\boldsymbol{\Psi}_i = \frac{\boldsymbol{d}_i}{\|\boldsymbol{d}_i\|}$$

In modo perfettamente analogo è possibile costruire una base per un insieme di segnali ad energia finita.

Esempio: Modulazione 4-PSK a bassa frequenza.

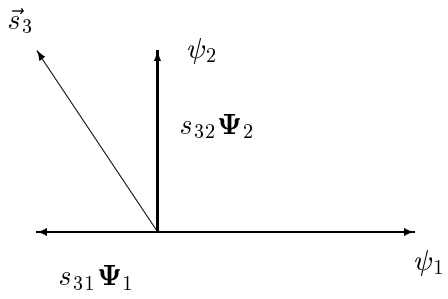

Figure 6: Generazione del vettore $\boldsymbol{d}_3$

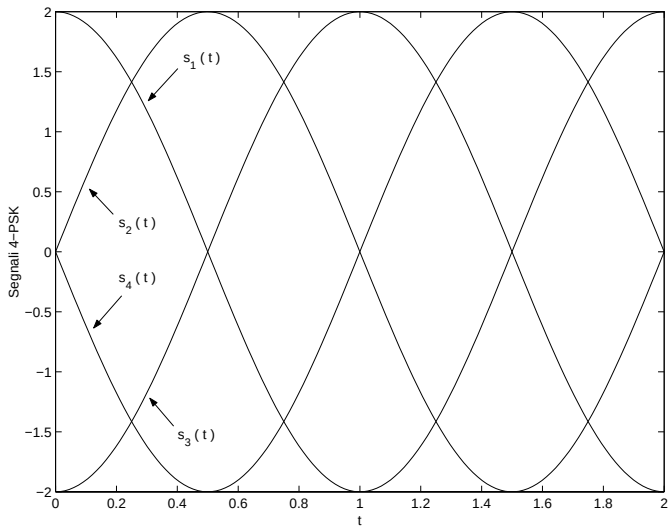

Figure 7: Forme d'onda generate per una segnalazione 4-PSK.

Consideriamo i quattro segnali

$$\begin{aligned} s_1(t) &= A\cos\left(\frac{2\pi}{T}t\right) \\ s_2(t) &= A\cos\left(\frac{2\pi}{T}t-\frac{\pi}{2}\right) = A\sin\left(\frac{2\pi}{T}t\right) \\ s_3(t) &= A\cos\left(\frac{2\pi}{T}t-\pi\right) = -A\cos\left(\frac{2\pi}{T}t\right) \\ s_4(t) &= A\cos\left(\frac{2\pi}{T}t-\frac{3\pi}{2}\right) = -A\sin\left(\frac{2\pi}{T}t\right) \end{aligned} \qquad t\in[0,T] \tag{14}$$

rappresentati in fig. 7.

Applichiamo il procedimento di Gram-Schmidt al fine di ottenere un sistema di versori ortonormali

che permetta la rappresentazione dei quattro segnali di energia $s_1(t)$, $s_2(t)$, $s_3(t)$ ed $s_4(t)$ nello spazio generato dai versori.

passo 1▹

$$\psi_1(t) = \frac{s_1(t)}{||s_1(t)||} = \frac{A\cos\left(\frac{2\pi}{T}t\right)}{\sqrt{\frac{A^2T}{2}}} = \sqrt{\frac{2}{T}}\cos\left(\frac{2\pi}{T}t\right)$$

passo 2▹

$$d_2(t) = s_2(t) - s_2(1)\psi_1(t) = A\sin\left(\frac{2\pi}{T}t\right)$$

infatti risulta

$$\begin{aligned} s_2(1) &= < s_2(t), \psi_1(t) >= \int_0^T A\sin\left(\frac{2\pi}{T}t\right)\sqrt{\frac{2}{T}}\cos\left(\frac{2\pi}{T}t\right)\, dt \\ &= A\sqrt{\frac{2}{T}}\int_0^T \sin\left(\frac{4\pi}{T}t\right)\, dt = 0 \end{aligned}$$

da cui

$$\psi_2(t) = \frac{d_2(t)}{||d_2(t)||} = \sqrt{\frac{2}{T}}\sin\left(\frac{2\pi}{T}t\right)$$

passo 3▹

$$d_3(t) = s_3(t) - s_3(1)\psi_1(t) - s_3(2)\psi_2(t) \tag{15}$$

$$\begin{aligned} s_3(1) &= < s_3(t), \psi_1(t) >= \int_0^T -A\cos\left(\frac{2\pi}{T}t\right)\frac{2}{T}\cos\left(\frac{2\pi}{T}t\right)\, dt \\ &= -\frac{A\sqrt{2}}{\sqrt{T}}\frac{T}{2} = -A\sqrt{\frac{T}{2}} \end{aligned}$$

$$s_3(2) =< s_3(t), \psi_2(t) >= \int_0^T -A\cos\left(\frac{2\pi}{T}t\right)\sqrt{\frac{2}{T}}\sin\left(\frac{2\pi}{T}t\right)\, dt = 0$$

per cui sostituendo nella (15) otteniamo

$$d_3(t) = -A\cos\left(\frac{2\pi}{T}t\right) + A\cos\left(\frac{2\pi}{T}t\right) = 0$$

$$\psi_3(t) = \frac{d_3(t)}{||d_3(t)||} = 0$$

quindi il segnale $s_3(t)$ è rappresentabile mediante i precedenti versori.

passo 4▷

$$d_4 = s_4(t) - s_4(1)\psi_1(t) - s_4(2)\psi_2(t)$$

sviluppando i calcoli si ricava

$$s_4(1) =< s_4(t), \psi_1(t) >= 0$$

$$s_4(2) =< s_4(t), \psi_2(t) >= -A\sqrt{\frac{T}{2}}$$

$$d_4(t) = -A\sin\left(\frac{2\pi}{T}t\right) + A\sin\left(\frac{2\pi}{T}t\right) = 0$$

Dalla procedura di Gram-Schmidt deduciamo che la dimensione dello spazio dei segnali è uguale a due in quanto i segnali $s_3(t)$ ed $s_4(t)$ possono essere espressi come combinazione lineare dei versori $\psi_1(t)$ e $\psi_2(t)$. In particolare, dati i due versori $\psi_1(t) = \sqrt{\frac{2}{T}}\cos\left(\frac{2\pi}{T}t\right)$ e $\psi_2(t) = \sqrt{\frac{2}{T}}\sin\left(\frac{2\pi}{T}t\right)$, le coordinate dei quattro segnali risultano essere

$$\boldsymbol{s}_1 = \left[A\sqrt{\frac{T}{2}},\ 0\right]^T$$

$$\boldsymbol{s}_2 = \left[0,\ A\sqrt{\frac{T}{2}}\right]^T$$

$$\boldsymbol{s}_3 = \left[-A\sqrt{\frac{T}{2}},\ 0\right]^T$$

$$\boldsymbol{s}_4 = \left[0,\ -A\sqrt{\frac{T}{2}}\right]^T$$

Da ciò si deduce che i segnali sono rappresentabili in uno spazio bidimensionale e risultano situati su una circonferenza equidistanziati di un angolo pari a $\frac{\pi}{2}$ come mostrato in fig. 8.

Esempio

Dimostriamo che, posto $f_0 >> 1/T$, i due segnali

$$x(t) = \sqrt{\frac{2}{T}}\cos(2\pi f_0 t)\Pi\left(\frac{t - T/2}{T}\right)\ , \qquad y(t) = \sqrt{\frac{2}{T}}\sin(2\pi f_0 t)\Pi\left(\frac{t - T/2}{T}\right)\ . \tag{16}$$

sono due versori ortogonali. Anzitutto, cominciamo col calcolare la norma di $x(t)$ al quadrato:

$$\begin{aligned} \|x(t)\|^2 &= \frac{2}{T}\int_0^T \cos^2(2\pi f_0 t)dt = \frac{2}{T}\int_0^T \frac{1+\cos(4\pi f_0 t)}{2}dt = \\ 1 + \frac{1}{T}\int_0^T \cos(4\pi f_0 t)dt &= 1 + \frac{\sin(4\pi f_0 T)}{4\pi f_0 T} = 1 + \mathrm{sinc}(2f_0 T) \approx 1\ , \end{aligned} \tag{17}$$

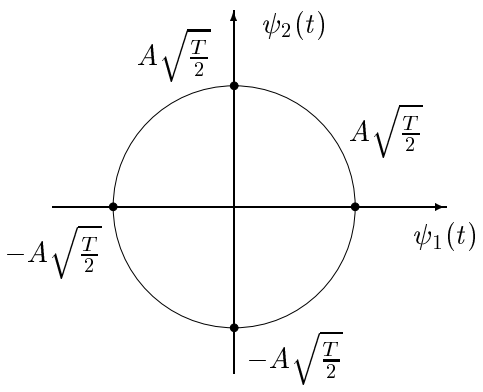

Figure 8: Rappresentazione dei vettori associati ad una costellazione di segnali di tipo 4-PSK.

ove l'ultima approssimazione è dovuta al fatto che si è assunto $f_0T >> 1$. In maniera analoga si dimostra che $\|y(t)\|^2 = 1$. Consideriamo poi il prodotto scalare

$$< x(t), y(t) >= \frac{2}{T}\int_0^T \cos(2\pi f_0 t)\sin(2\pi f_0 t) = \frac{1}{T}\int_0^T \sin(4\pi f_0 t)dt = \frac{1-\cos(4\pi f_0 T)}{4\pi f_0 T} \approx 0 \ , \tag{18}$$

ove l'ultima relazione vale in quanto, si ricorda, è $f_0T \gg 1$. I segnali in (16) che rappresentano due portanti in quadratura finestrate ad un intervallo di durata T, sono quindi due versori ortogonali.

2.1 Teorema della Dimensionalità (Landau-Pollack)

Nelle pagine precedenti si è visto come l'insieme $\mathcal{S}$ delle M forme d'onda di un modulatore definisca un sottospazio L-dimensionale dell'insieme dei segnali di energia a supporto nell'intervallo $[0, T]$. Vogliamo ora illustrare un risultato importante, noto come *Teorema di Landau e Pollack* o anche *teorema della dimensionalità*, che lega la dimensionalità di un insieme di forme d'onda alla loro occupazione in banda. Al solito, ci limiteremo a dare l'enunciato del teorema ed a giustificarlo in maniera intuitiva, rimandando a testi specialistici ulteriori approfondimenti. E' opportuno precisare che, dal momento che si stanno considerando funzioni a durata rigorosamente limitata in $[0, T]$, l'estensione in frequenza delle trasformate di Fourier di tali segnali è in linea di principio infinita. D'altra parte, nei corsi di base di Teoria dei Segnali è stato visto che, dal momento che lo spettro di un segnale di energia è infinitesimo all'infinito ed è anch'esso un segnale di energia, è ragionevole, per tali segnali, definire la banda come l'estensione dell'intervallo di frequenze all'interno del quale ricade una porzione significativa dell'energia del seg-

nale. Bene, il teorema della dimensionalità fornisce una misura approssimata del numero di forme d'onda ortogonali di durata finita pari a T e con occupazione in banda W. In particolare il teorema afferma che il numero L di forme d'onda linearmente indipendenti di durata T e banda[2] W è approssimativamente pari al prodotto WT. Altrimenti detto, si ha

$$L \approx WT \ . \tag{19}$$

Tale teorema permette quindi di determinare l'occupazione in banda dell'insieme delle forme d'onda $\mathcal{S}$ utilizzate da un modulatore numerico sulla base della durata e della dimensionalità. Inoltre, l'applicazione di tale teorema ci permette di ottenere anche una stima approssimata dell'efficienza spettrale conseguita da uno schema di modulazione. Infatti, dal momento che ad ogni forma d'onda del modulatore è associata una stringa di bit di lunghezza $\log_2 M$, ne consegue che il bit-rate R_b al quale il modulatore invia informazioni sul canale è rappresentato dal rapporto $\frac{\log_2 M}{T}$. Ma allora, sfruttando la (19) l'efficienza spettrale conseguita da un modulatore M-ario con intervallo di segnalazione T e dimensionalità della costellazione L è approssimativamente pari a

$$\frac{R_b}{W} \approx \frac{\frac{\log_2 M}{T}}{\frac{L}{T}} = \frac{\log_2 M}{L} \ . \tag{20}$$

3 Classificazione degli schemi di modulazione

In tale paragrafo sono presentati e classificati sulla base della loro dimensionalità alcuni schemi di segnalazione notevoli.

3.1 Segnalazione monodimensionale: PAM o ASK

Nella segnalazione PAM (*Pulse Amplitude Modulation*) o anche ASK (*Amplitude Shift Keying*), l'insieme di forme d'onda utilizzate dal modulatore è tale che

$$s_i(t) = A_i s(t) \ , \qquad \forall i = 1, \ldots, M \ . \tag{21}$$

In altri termini, le forme d'onda del modulatore differiscono esclusivamente per un fattore di ampiezza. Le ampiezze A_i sono quantità reali, mentre la forma d'onda $s(t)$ può essere scelta arbitrariamente. Ad

[2] L'enunciato del teorema definisce la banda come l'estensione dell'intervallo di frequenze contenente gli $11/12$ dell'energia complessiva del segnale.

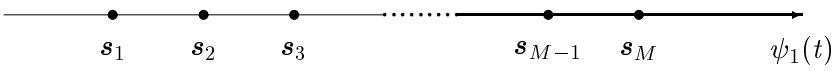

Figure 9: Rappresentazione della costellazione di segnali ASK.

esempio, per trasmissioni in banda base, $s(t)$ potrebbe essere assunto essere pari ad una finestra rettangolare di durata T e ad energia unitaria, ovvero

$$s(t) = \Pi\left(\frac{t - T/2}{T}\right) \ .$$

Nel caso invece, di maggior rilevanza pratica, di una modulazione in banda traslata, $s(t)$ può essere assunto pari ad una sinusoide finestrata, ossia

$$s(t) = \sqrt{\frac{2}{T}} \cos 2\pi f_0 t \ , \quad t \in [0, T] \ .$$

E' immediato verificare che, essendo le forme d'onda utilizzate dal modulatore tutte proporzionali ad un unico segnale $s(t)$, la dimensionalità della modulazione è in tal caso pari a 1, ovvero è $L = 1$ e $\psi_1(t) = s(t)/\|s(t)\|$. Ne consegue quindi che a ciasun segnale $s_i(t)$ può essere associato un vettore $\boldsymbol{s}_i$ che in tal caso è un semplice scalare, e la costellazione dei segnali PAM è rappresentata da un insieme di punti allineati, come illustrato in figura 9. E' da sottolineare il fatto che l'occupazione in banda del set di forme d'onda coicide in tal caso con l'occupazione in banda del segnale $s(t)$, ed è indipendente dalla cardinalità M del modulatore. Assumendo che l'occupazione in banda del segnale $s(t)$ sia approssimativamente pari all'inverso della sua durata T, l'efficienza spettrale della modulazione ASK, anche a norma della relazione (20), è espressa come

$$\frac{R_b}{W} = \log_2 M \ . \tag{22}$$

Si osservi che l'efficienza spettrale diverge all'aumentare della cardinalità M della modulazione, per cui si può concludere che la modulazione PAM è una modulazione efficiente in banda, ed è quindi indicata in quelle situazioni in cui bisogna trasmettere con elevati bit-rate in bande limitate. Vedremo nel seguito che il prezzo da pagare per tale efficienza in banda consiste nel dover accettare una inefficienza in potenza.

3.2 Segnalazioni bidimensionali

In una segnalazione bidimensionale il set di forme d'onda utilizzato dal modulatore può essere espresso come combinazione lineare di due versori ortonormali $\psi_1(t)$ e $\psi_2(t)$. In particolare, si ha

$$s_i(t) = A_i(1)\psi_1(t) + A_i(2)\psi_2(t) \ , \quad i = 1, \ldots, M \ . \tag{23}$$

Un esempio di segnalazione bidimensionale è dato dalla segnalazione PSK (*Phase Shift Keying*) in cui gli elementi di $\mathcal{S}$ sono delle sinusoidi finestrate che differiscono per il valore della fase iniziale, ovvero è

$$s_i(t) = \sqrt{\frac{2\mathcal{E}}{T}} \cos\left(2\pi f_0 t + \phi_i\right) \Pi\left(\frac{t - T/2}{T}\right) \ . \tag{24}$$

Per dimostrare che il set di segnali definito dalla (24) è bidimensionale è sufficiente sviluppare il coseno con la formula di addizione. Si ha, infatti:

$$s_i(t) = \sqrt{\mathcal{E}}\cos(\phi_i)\left[\sqrt{\frac{2}{T}}\cos(2\pi f_0 t)\right] + \sqrt{\mathcal{E}}\sin(\phi_i)\left[-\sqrt{\frac{2}{T}}\sin(2\pi f_0 t)\right] \ , \quad t \in [0, T] \ . \tag{25}$$

Da tale relazione si riconosce che il generico segnale $s_i(t)$ può essere espresso come combinazione lineare dei due versori

$$\begin{cases} \psi_1(t) = \sqrt{\frac{2}{T}}\cos(2\pi f_0 t)\Pi\left(\frac{t - T/2}{T}\right) \ , \\ \psi_2(t) = -\sqrt{\frac{2}{T}}\sin(2\pi f_0 t)\Pi\left(\frac{t - T/2}{T}\right) \ . \end{cases} \tag{26}$$

Alla forma d'onda $s_i(t)$ può quindi essere associato il vettore bidimensionale $\boldsymbol{s}_i = [\sqrt{\mathcal{E}}\cos(\phi_i),\ \sqrt{\mathcal{E}}\sin(\phi_i)]^T$. La costellazione dei segnali PSK è quindi rappresentata da un insieme di punti giacenti su una circonferenza di raggio $\sqrt{\mathcal{E}}$ in un piano bidimensionale. Assumendo che l'occupazione in banda dei versori $\psi_1(t)$ e $\psi_2(t)$ possa essere assunta pari all'inverso della loro durata T, è immediato verificare che l'efficienza spettrale per la segnalazione PSK è pari a $\log_2 M$. Anche la segnalazione PSK, quindi, è una segnalazione efficiente in banda e non in potenza.

3.3 Segnalazione ortogonale

Nella segnalazione ortogonale le forme d'onda utilizzate dal modulatore sono ortogonali tra loro, ovvero vale la relazione

$$< s_i(t), s_j(t) >= \int s_i(t)s_j(t)dt = 0 \quad \text{per } i \neq j \ . \tag{27}$$

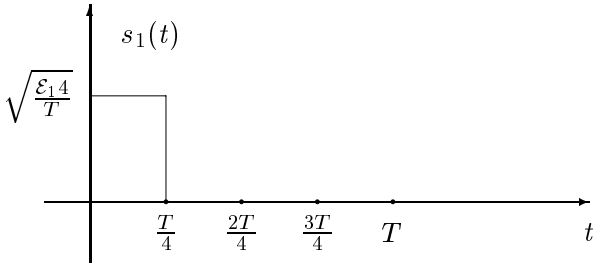

Figure 10: Forma d'onda $s_1(t)$ usata nella segnalazione 4-PPM.

Ne consegue quindi che per tale segnalazione è $L = M$, ovvero la dimensionalità della modulazione coicide con la cardinalità del modulatore. In particolare, gli elementi della base $\mathcal{B}$ sono espressi come

$$\psi_\ell(t) = \frac{s_\ell(t)}{\|s_\ell(t)\|} \ , \qquad \ell = 1, \ldots, L \ , \tag{28}$$

mentre il vettore L-dimensionale associato al generico segnale $s_i(t)$ è dato da

$$\boldsymbol{s}_i = \left[\underbrace{0\,,\ldots,0}_{i-1}, \|s_i(t)\|\,, \underbrace{0\,,\ldots,0}_{M-i} \right]^T \ . \tag{29}$$

Un esempio di segnalazione ortogonale è dato dalla segnalazione PPM (*Pulse Position Modulation*), in cui l'informazione è affidata alla posizione che l'impulso trasmesso occupa all'interno dell'intervallo $[0, T]$. Ad esempio, un modulatore PPM può utilizzare le forme d'onda

$$s_i(t) = \sqrt{\frac{\mathcal{E}_i M}{T}}\, \Pi\left(\frac{t - \frac{T}{2M} - (i-1)T/M}{T/M} \right) \ , \qquad i = 1, \ldots, M \ . \tag{30}$$

Altrimenti detto, l'intervallo di segnalazione è suddiviso in M sottointervalli di durata T/M, e ciascuna delle forme d'onda del modulatore è non nulla solo all'interno di tale sottointervallo. Nella figura 10 è riportata la forma d'onda $s_1(t)$ per il caso $M = 4$.

Altro esempio di segnalazione ortogonale è la modulazione FSK (*Frequency Shift Keying*), in cui si sceglie

$$s_i(t) = \sqrt{\frac{2\mathcal{E}}{T}} \cos\left(2\pi f_0 t + 2\pi i \triangle_f t\right) \qquad 0 \le t \le T \qquad i = 1\,, \ldots\,, M \ . \tag{31}$$

In altri termini, le forme d'onda del modulatore sono delle sinusoidi, finestrate all'intervallo $[0, T]$, a frequenze diverse tra loro. L'informazione quindi, è associata alla frequenza della sinusoide trasmessa.

In verità, la segnalazione FSK in (31) non può essere sempre considerata una segnalazione ortogonale, in quanto, come illustrato nel seguito, i segnali FSK sono ortogonali soltanto per opportuni valori dell'offset frequenziale $\triangle_f$. Infatti, si consideri il prodotto scalare

$$\begin{aligned}
< s_i(t), s_j(t) > &= \int_0^T \frac{2\mathcal{E}}{T} \cos\left(2\pi f_0 t + 2\pi i \triangle_f t\right) \cos\left(2\pi f_0 t + 2\pi j \triangle_f t\right)\, dt \\
&= \frac{\mathcal{E}}{T} \int_0^T \cos\left[4\pi f_0 t + 2\pi \triangle_f t(i+j)\right]\, dt + \\
&+ \frac{\mathcal{E}}{T} \int_0^T \cos\left[2\pi (i-j) \triangle_f t\right]\, dt \\
&= \frac{\mathcal{E}}{T} \frac{\alpha}{4\pi f_0 + 2\pi \triangle_f (i+j)} + \frac{\mathcal{E}}{T} \frac{\sin\left(2\pi (i-j) \triangle_f T\right)}{2\pi (i-j) \triangle_f} \qquad (32) \\
&= \mathcal{E} \mathrm{sinc}\left(2(i-j) \triangle_f T\right) \qquad (33)
\end{aligned}$$

dove nella (33) si é trascurato il primo termine della (32) in quanto $\ll 1$ per $f_0 \gg \frac{1}{T}$. Affinché sia soddisfatta la condizione di ortogonalità è necessario quindi imporre che sia

$$2\triangle_f (i-j) T = m$$

con m pari ad un qualsiasi intero non nullo. Ne consegue quindi che i segnali FSK sono ortogonali solo se la spaziatura tra le frequenze portanti delle forme d'onda utilizzate sia un multiplo intero di $1/(2T)$. In particolare, la minima spaziatura necessaria a garantire l'ortogonalità delle forme d'onda è pari a $1/(2T)$. Vedremo nel seguito che, se si assume che le forme d'onda utilizzate dal modulatore non sono tutte in fase, ovvero si assume che la generica $s_i(t)$ abbia una propria fase iniziale θ_i, allora la condizione di ortogonalità richiede una spaziatura tra le frequenze che sia un multiplo di $1/T$.

La struttura del Demodulatore

Dopo aver illustrato la modalità di funzionamento del modulatore, occupiamoci ora del blocco di demodulazione. Il demodulatore deve essenzialmente assolvere il compito di *rivelare* la forma d'onda trasmessa, ovvero, sulla base dell'osservazione del segnale ricevuto, deve determinare quale tra le M possibili forme d'onda è stata trasmessa dal modulatore. Cominciamo col supporre che il modulatore si limiti a trasmettere un'unica forma d'onda nell'intervallo $[0, T]$, ovvero consideriamo il caso di trasmissione isolata di un singolo simbolo M-ario di informazione. Ci focalizzeremo inoltre sul modello di

canale AWGN non distorcente, ovvero assumeremo che il canale si limiti ad introdurre sul segnale trasmesso un disturbo di tipo additivo. A valle della trasmissione di una possibile forma d'onda del modulatore, ad esempio $s_i(t)$, il segnale ricevuto $r(t)$ è quindi esprimibile come:

$$r(t) = s_i(t) + n(t) \ , \tag{34}$$

ove $n(t)$ è assunto essere una funzione membro di un processo aleatorio Gaussiano bianco a densità spettrale di potenza (PSD) pari a $\mathcal{N}_0/2$. Naturalmente, a causa del rumore introdotto dal canale, non sarà in generale possibile riuscire a rivelare correttamente qual'è stata la forma d'onda inviata sul canale, e, quindi, potranno verificarsi degli *errori di rivelazione*. Al fine di minimizzare l'occorrenza di tali errori, è di interesse ricercare la strategia di demodulazione che renda minima la probabilità di errore. Altrimenti detto, siamo interessati a determinare quali sono le elaborazioni da effettuare sul segnale ricevuto $r(t)$ in modo tale che la forma d'onda trasmessa possa essere individuata con una probabilità di errore minima.

Il primo passo verso l'individuazione della regola di decisione a minima probabilità di errore (o, equivalentemente, a massima probabilità di corretta decisione), consiste nella discretizzazione del segnale ricevuto $r(t)$. Al riguardo, dal momento che le forme d'onda trasmesse sul canale sono perfettamente rappresentate dal vettore delle loro proiezioni sugli elementi della base $\mathcal{B}$ dello spazio dei segnali, si considera il vettore delle proiezioni di $r(t)$ lungo tale base. Si ottiene quindi il seguente vettore L-dimensionale:

$$\boldsymbol{r} = \begin{bmatrix} < r(t), \psi_1(t) > \\ < r(t), \psi_2(t) > \\ \vdots \\ < r(t), \psi_L(t) > \end{bmatrix} . \tag{35}$$

Posto

$$\boldsymbol{n} = \begin{bmatrix} < n(t), \psi_1(t) > \\ < n(t), \psi_2(t) > \\ \vdots \\ < n(t), \psi_L(t) > \end{bmatrix} , \tag{36}$$

è immediato giungere alla relazione

$$\boldsymbol{r} = \boldsymbol{s}_i + \boldsymbol{n} \ . \tag{37}$$

Si noti che l'operazione di discretizzazione introduce un errore di rappresentazione del rumore $n(t)$. Infatti, mentre gli elementi di $\mathcal{B}$ costituiscono una base per le forme d'onda del modulatore, e, quindi, la discretizzazione di $s_i(t)$ non introduce alcuna perdita di informazione in quanto trattasi di trasformazione reversibile, lo stesso non può dirsi con riferimento al rumore $n(t)$. In altre parole, il processo di discretizzazione che fa passare da $n(t)$ a $\boldsymbol{n}$ non è reversibile, in quanto è

$$n(t) \neq \sum_{\ell=1}^{L} < n(t), \psi_\ell(t) > \psi_\ell(t) \ . \tag{38}$$

Tuttavia, dal momento che il rumore è gaussiano e bianco, si può dimostrare che la porzione di rumore persa nel processo di discretizzazione è *irrilevante* ai fini del processo di rivelazione. Ovvero, dal momento che è possibile dimostrare che la parte di rumore persa nel processo di discretizzazione, ossia la quantità $n(t) - \sum_{\ell=1}^{L} < n(t), \psi_\ell(t) > \psi_\ell(t)$, è statisticamente indipendente dal vettore $\boldsymbol{n}$, la discretizzazione del rumore comporta la perdita di informazioni non utili ai fini della rivelazione. Altrimenti detto, è possibile dimostrare che elaborare il vettore $\boldsymbol{r}$ non comporta alcuna perdita di prestazioni rispetto all'elaborazione del segnale tempo-continuo $r(t)$.

Valutiamo ora le statistiche del vettore $\boldsymbol{n}$. Anzitutto, le componenti di $\boldsymbol{n}$, essendo il risultato di elaborazioni lineari sul processo aleatorio $n(t)$, sono variabili aleatorie gaussiane. La media di tali variabili è nulla. Infatti, indicando con $n(k)$ la k-esima componente del vettore $\boldsymbol{n}$, si ha che

$$E[n(k)] = E\left[\int_0^T n(t)\psi_k(t)\,dt\right] = \int_0^T E[n(t)]\psi_k(t)\,dt = 0 \ .$$

Essendo le medie statistiche tutte nulle, la covarianza tra la k-esima e la h-esima componente di $\boldsymbol{n}$ si può calcolare come

$$\begin{aligned}
E[n(k)n(h)] &= E\left\{\int_0^T n(t)\psi_k(t)\,dt \int_0^T n(\tau)\psi_h(\tau)\,d\tau\right\} \\
&= \int_0^T \int_0^T E[n(t)n(\tau)]\psi_k(t)\psi_h(\tau)\,dt\ d\tau \\
&= \int_0^T \int_0^T r_n(t-\tau)\psi_k(t)\psi_h(\tau)\,dt\ d\tau \\
&= \int_0^T \int_0^T \frac{N_0}{2}\delta(t-\tau)\psi_k(t)\psi_h(\tau)\,d\tau\ dt \\
&= \frac{\mathcal{N}_0}{2}\int_0^T \psi_k(\tau)\psi_h(\tau)\,d\tau \\
&= \frac{\mathcal{N}_0}{2}\delta_{k,h} \ ,
\end{aligned} \tag{39}$$

ove si è indicato con $r_n(t-\tau)$ la funzione di autocorrelazione di $n(t)$. In definitiva, quindi, le variabili aleatorie $n(k), k = 1, \ldots, L$, risultano essere mutuamente incorrelate, gaussiane, a media nulla e varianza $\mathcal{N}_0/2$. La matrice di covarianza del vettore $\boldsymbol{n}$ è quindi espressa come

$$E\left[\boldsymbol{n}\boldsymbol{n}^T\right] = \frac{\mathcal{N}_0}{2}\boldsymbol{I}_L \ , \tag{40}$$

ove $\boldsymbol{I}_L$ denota la matrice identica di ordine L.

Assumendo che le forme d'onda del modulatore abbiano una distribuzione di probabilità uniforme, ovvero che il modulatore possa inviare con eguale probabilità (pari a $1/M$) ciascuna delle M forme d'onda a sua disposizione, la strategia di rivelazione ottima, ovvero la regola di decisione che permette di conseguire la minima probabilità di errore, si dimostra essere[3] quella della *massima verosimiglianza*. Tale regola consiste nello scegliere, quale segnale trasmesso, quello che rende massima la densità di probabilità del vettore osservato $\boldsymbol{r}$ condizionatamente al simbolo trasmesso. Altrimenti detto, la regola di decisione ottima è

$$\text{decidi per } s_i(t) \text{ se } \quad f_{\boldsymbol{r}|\boldsymbol{s}_i}(\boldsymbol{r}|\boldsymbol{s}_i) > \max_{k\neq i} f_{\boldsymbol{r}|\boldsymbol{s}_k}(\boldsymbol{r}|\boldsymbol{s}_k) \ . \tag{41}$$

Condizionatamente al vettore $\boldsymbol{s}_i$, ovvero nell'ipotesi che sia stato trasmesso $s_i(t)$, il vettore ricevuto $\boldsymbol{r}$ è un vettore aleatorio Gaussiano a media $\boldsymbol{s}_i$ e con matrice di covarianza $\frac{\mathcal{N}_0}{2}\boldsymbol{I}_L$, quindi la pdf condizionale $f_{\boldsymbol{r}|\boldsymbol{s}_i}(\boldsymbol{r}|\boldsymbol{s}_i)$ è data da

$$f_{\boldsymbol{r}|\boldsymbol{s}_i}(\boldsymbol{r}|\boldsymbol{s}_i) = \frac{1}{(\pi\mathcal{N}_0)^{L/2}} \exp\left\{-\frac{1}{\mathcal{N}_0}\sum_{k=1}^{L}(\boldsymbol{r}(k)-\boldsymbol{s}_i(k))^2\right\} = \frac{1}{(\pi\mathcal{N}_0)^{L/2}} \exp\left\{-\frac{\|\boldsymbol{r}-\boldsymbol{s}_i\|^2}{\mathcal{N}_0}\right\} \ . \tag{42}$$

Sulla base di tale relazione, è facile rendersi conto che massimizzare la verosimiglianza equivale a minimizzare la distanza euclidea tra il vettore osservato $\boldsymbol{r}$ ed i vettori associati alle possibili forme d'onda trasmesse sul canale. La regola a massima verosimiglianza (41) è quindi equivalente alla seguente regola di decisione a minima distanza

$$\text{decidi per } s_i(t) \text{ se } \quad \|\boldsymbol{r}-\boldsymbol{s}_i\|^2 < \min_{k\neq i} \|\boldsymbol{r}-\boldsymbol{s}_k\|^2 \ . \tag{43}$$

In altre parole, il ricevitore ottimo a minima probabilità di errore decide per quel simbolo il cui vettore rappresentativo si trova più vicino al vettore osservato $\boldsymbol{r}$ di quanto lo siano gli altri segnali. Ad esempio,

[3] La dimostrazione di tale risultato non è riportata per motivi di semplicità e brevità.

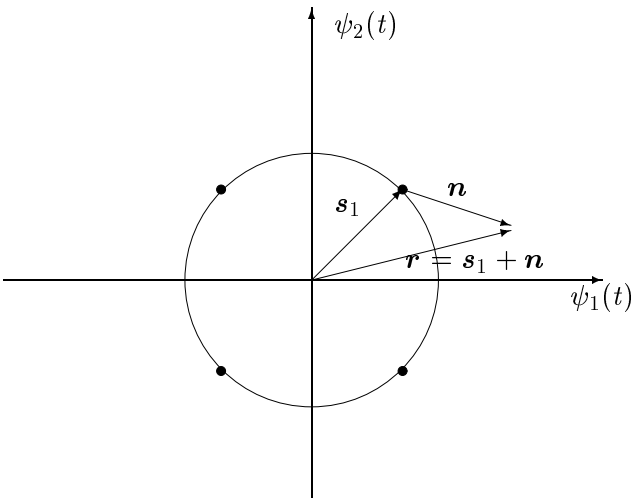

Figure 11: Rappresentazione del vettore $\boldsymbol{r}$ nel caso che sia stato trasmesso $s_1(t)$ ed il rumore è tale da non causare una rivelazione errata.

nella figura 11 è rappresentata la situazione in cui la modulazione utilizzata è una 4-PSK ed il vettore $\boldsymbol{r}$ ha entrambe le coordinate positive e si trova quindi nel primo quadrante. In tale situazione, dal momento che il punto $\boldsymbol{s}_1$ si trova alla distanza minore da $\boldsymbol{r}$, il ricevitore ottimo deciderà per il simbolo $s_1(t)$. Si tenga presente che la regola a minima distanza è la strategia di rivelazione che minimizza la probabilità di errore, ma che in ogni caso non la rende nulla. Ad esempio, potrebbe verificarsi la situazione rappresentata in Fig. 12, in cui è stata trasmessa la forma d'onda $s_4(t)$, ma si è avuta una realizzazione di rumore $\boldsymbol{n}$ tale che la somma $\boldsymbol{s}_4 + \boldsymbol{n}$ da luogo ad un vettore osservato che si trova nel primo quadrante, per cui il ricevitore a minima distanza deciderà erroneamente per $\boldsymbol{s}_1$. La regola di decisione a minima distanza fa si che l'insieme $\mathcal{R}^L$ delle L-uple di numeri reali possa essere suddiviso in M *regioni di decisione*. Altrimenti detto, è possibile associare a ciascun vettore $\boldsymbol{s}_i$ l'insieme di tutti i punti che sono più vicini a $\boldsymbol{s}_i$ che non agli altri vettori $\boldsymbol{s}_j$, con $j = 1, \ldots, M$ e $j \neq i$. Tale insieme è detto regione di decisione associata al segnale $\boldsymbol{s}_i(t)$. Facendo riferimento alla segnalazione 4-PSK rappresentata in Fig. 11, è banale verificare che le regioni di decisione sono costituite dai quattro quadranti. Il demodulatore, quindi, non deve far altro che stabilire in quale regione di decisione cade il vettore $\boldsymbol{r}$ rappresentativo del segnale ricevuto $r(t)$. Un possibile schema a blocchi implementativo della regola di decisione a minima distanza è riportato in Fig. 13. Considerato che vale la relazione

$$\|\boldsymbol{r} - \boldsymbol{s}_i\|^2 = \|\boldsymbol{r}\|^2 + \|\boldsymbol{s}_i\|^2 - 2\boldsymbol{r}^T\boldsymbol{s}_i \ , \tag{44}$$

è immediato dimostrare che la regola di decisione (43) può anche essere espressa in maniera equivalente

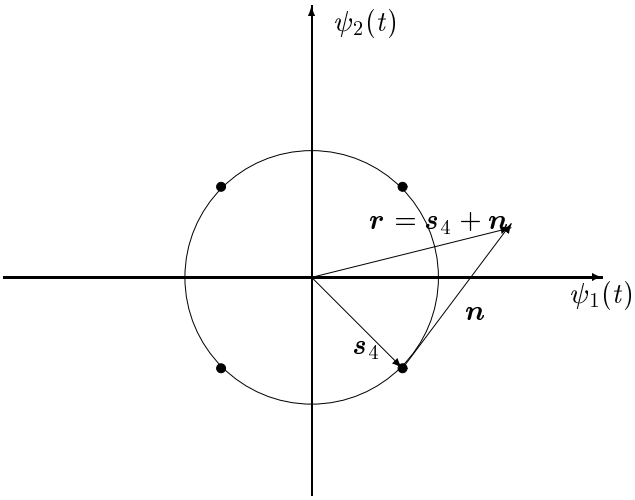

Figure 12: Rappresentazione del vettore $\boldsymbol{r}$ nel caso che sia stato trasmesso $s_4(t)$ ed il rumore è tale da causare una rivelazione errata.

come

$$\text{decidi per } s_i(t) \text{ se } \quad \boldsymbol{r}^T \boldsymbol{s}_i - \frac{\mathcal{E}_i}{2} > \max_{k \neq i} \boldsymbol{r}^T \boldsymbol{s}_k - \frac{\mathcal{E}_k}{2} \ . \tag{45}$$

Ricordando poi che il prodotto scalare tra i vettori L-dimensionali coincide numericamente col prodotto scalare tra le corrispondenti forme d'onda, la regola di decisione a minima probabilità di errore può essere espressa anche come

$$\text{decidi per } s_i(t) \text{ se } \quad \int_0^T r(t) s_i(t) dt - \frac{\mathcal{E}_i}{2} > \max_{k \neq i} \int_0^T r(t) s_k(t) dt - \frac{\mathcal{E}_k}{2} \ . \tag{46}$$

Uno schema a blocchi alternativo del demodulatore, in accordo alla regola di decisione (46), è rappresentato in Fig. 14. Si noti che lo schema di Fig. 13 è costituito da L rami e da uno stadio combinatorio, mentre quello di Fig. 14 è costituito da M rami e da un semplice selettore del massimo. Per modulazioni monodimensionali e bidimensionali, quindi, è conveniente utilizzare lo schema di figura 13, mentre, nel caso di modulazioni ortogonali, ove $L = M$, entrambi gli schemi hanno la stessa complessità. Dal punto di vista realizzativo, inoltre, è importante evidenziare che gli schemi di figura 13 e 14 possono essere implementati in modo più efficiente mediante un filtro lineare tempo invariante (LTI) di opportuna risposta impulsiva. In particolare, come rappresentato in fig. 15, la moltiplicazione per la forma d'onda $\psi_\ell(t)$ e la successiva integrazione sull'intervallo $[0, T]$, può essere sostituita da un filtro LTI di risposta impulsiva $h(t) = \psi_\ell(t - T)$ seguito da un campionatore all'istante $t = T$. Infatti, detta $y(t)$ l'uscita del filtro, si ha

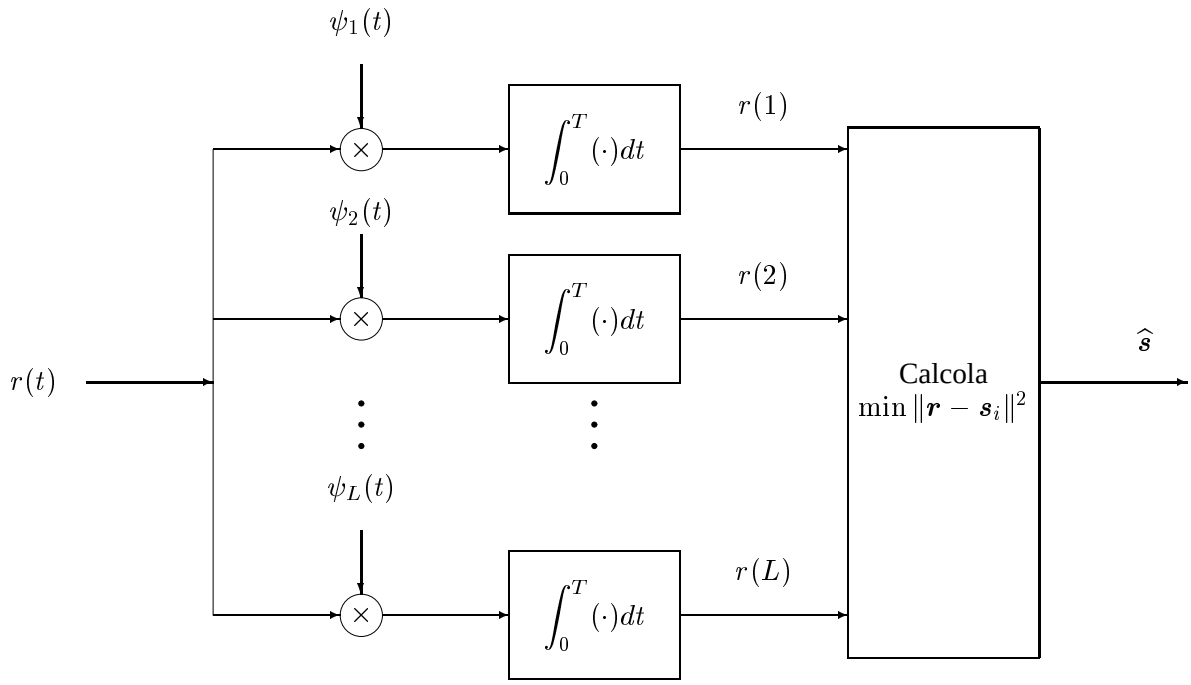

Figure 13: Schema a blocchi del demodulatore ottimo.

che

$$y(t) = \psi_\ell(T-t) * r(t) = \int \psi_\ell(T-\tau) r(t-\tau) d\tau \ , \tag{47}$$

e, campionando il segnale $y(t)$ all'istante $t = T$, si ha

$$y(T) = \int \psi_\ell(T-\tau) r(T-\tau) d\tau = \int_0^T \psi_\ell(t) r(t) dt \ , \tag{48}$$

ove l'integrale è esteso all'intervallo $[0, T]$ e non a tutto l'asse reale perché la funzione $\psi_\ell(t)$ è non nulla solo per $t \in [0, T]$. Il filtro di risposta impulsiva $\psi_\ell(T-t)$ è usualmente detto *filtro adattato* (in inglese *matched filter*) alla forma d'onda $\psi_\ell(t)$. In maniera del tutto analoga è immediato verificare che il blocco di moltiplicazione per la generica forma d'onda $s_i(t)$ e la successiva integrazione sull'intervallo $[0, T]$ può essere sostituita da un filtraggio adattato ad $s_i(t)$ (ovvero da un filtro LTI di risposta impulsiva $s_i(T-t)$) seguito da un campionatore all'istante $t = T$. Gli schemi realizzativi delle figure 13 e 14 possono essere rispettivamente sostituiti dagli schemi 16 e 17.

Fino a questo momento, è stato considerato unicamente il caso che il modulatore si limitasse a

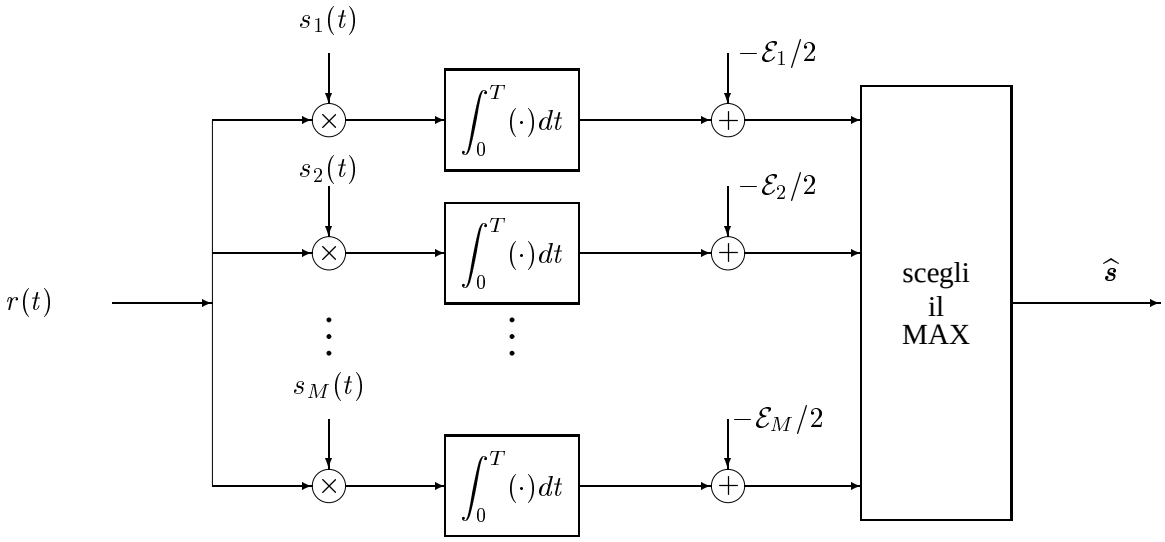

Figure 14: Schema alternativo del demodulatore a massima verosimiglianza.

trasmettere una tra le M forme d'onda a sua disposizione nell'intervallo $[0, T]$. E' chiaro tuttavia che, nella realtà, il modulatore si troverà a lavorare per molteplici intervalli di segnalazione consecutivi, e che quindi invierà una tra le forme d'onda a sua disposizione anche negli intervalli di segnalazione successivi all'intervallo $[0, T]$. Ci chiediamo quindi in che modo il demodulatore deve operare al fine di essere in grado di rivelare un simbolo trasmesso ogni T secondi. Cominciamo quindi a chiederci quale sia la regola di decisione da implementare nell'm-esimo intervallo di segnalazione $[mT, (m+1)T]$. In tale intervallo, il modulatore può inviare sul canale le M forme d'onda

$$s_1(t - mT), s_2(t - mT), \ldots, s_M(t - mT) \ ,$$

per cui è facile convincersi che il demodulatore dovrà implementare la seguente regola di decisione

$$\text{decidi per } s_i(t - mT) \text{ se } \int_{mT}^{(m+1)T} r(t)s_i(t-mT)dt - \frac{\mathcal{E}_i}{2} > \max_{k \neq i} \int_{mT}^{(m+1)T} r(t)s_k(t-mT)dt - \frac{\mathcal{E}_k}{2} \ . \tag{49}$$

Il termine

$$\int_{mT}^{(m+1)T} r(t)s_i(t - mT)$$

può essere ottenuto considerando l'uscita del filtro adattato a $s_i(t)$ campionato all'istante $t = (m+1)T$. Infatti, inviando il segnale ricevuto $r(t)$ in ingresso ad un filtro LTI di risposta $s_i(T - t)$, l'uscita, che indichiamo con $y(t)$, è espressa come

$$y(t) = \int s_i(T - \tau)r(t - \tau)d\tau \ . \tag{50}$$

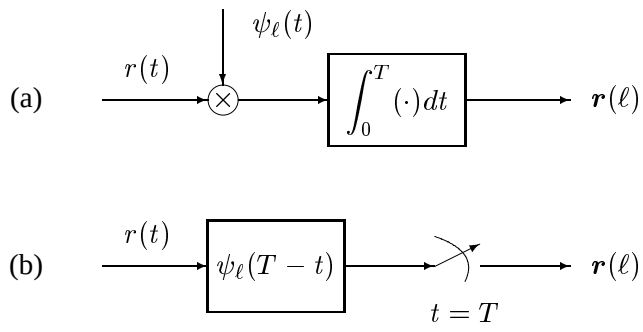

Figure 15: La struttura rappresentata in (a) è equivalente a quella in (b).

Considerando il campione di $y(t)$ all'istante $t = (m+1)T$ si ha

$$y((m+1)T) = \int s_i(T-\tau)r((m+1)T-\tau)d\tau = \int r(\lambda)s_i(\lambda-mT)d\lambda = \int_{mT}^{(m+1)T} r(t)s_i(t-mT)dt \ . \tag{51}$$

Ne consegue, quindi, che gli schemi di figura 16 e 17 possono essere utilizzati per rivelare non un'unica forma d'onda, ma bensì successioni di forme d'onda inviate sul canale dal modulatore in intervalli di segnalazione consecutivi, a patto di campionare le uscite dei filtri adattati non solo all'istante $t = T$ ma ogni T secondi, ovvero agli istanti multipli dell'intervallo di segnalazione base T.

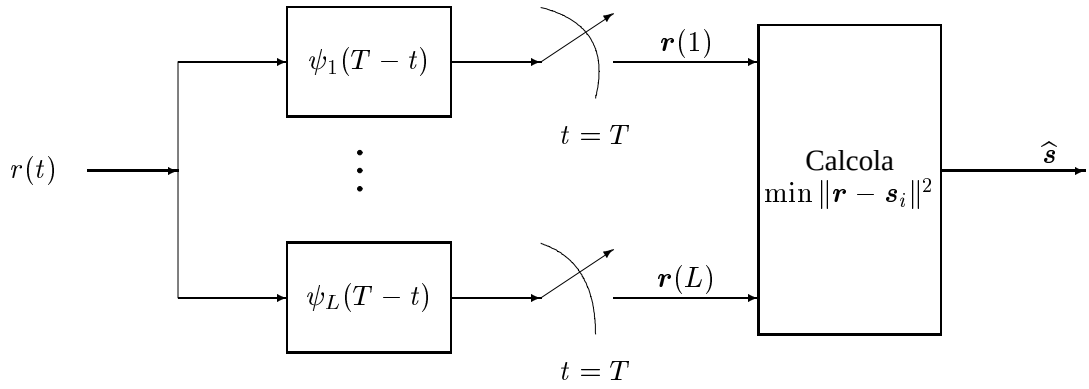

Figure 16: Schema alternativo del demodulatore a massima verosimiglianza.

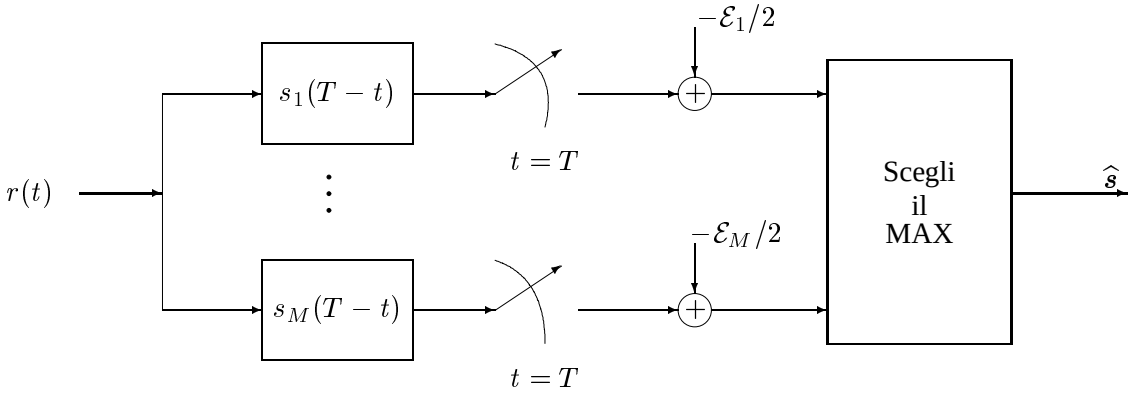

Figure 17: Schema alternativo del demodulatore a massima verosimiglianza.