

Resumindo e concluindo...

TeleTextos de bolso e de trazer por casa, suavemente, suavemente

Constelações no Espaço (em modulações digitais)

© Sílvio A. Abrantes

Departamento de Engenharia Electrotécnica e de Computadores
Faculdade de Engenharia, Universidade do Porto
Porto, Portugal
sam@fe.up.pt

Janeiro de 2009

Conteúdo

1.	Funções-base, espaços de sinal e constelações	1
1.1.	Correspondência entre energia e norma quadrática	4
1.2.	Distância euclidiana entre dois pontos	5
2.	Modulações digitais	5
2.1.	Modulação de amplitude (ASK)	7
2.2.	Modulação de fase (MPSK)	7
2.2.1.	Caso particular: BPSK	8
2.3.	Modulação simultânea de fase e amplitude (M-QAM)	9
2.4.	Modulação de frequência (MFSK)	9
2.4.1.	Caso particular: BFSK	10
2.5.	Outras constelações	10

1. Funções-base, espaços de sinal e constelações

Um *espaço ortogonal* de N dimensões é caracterizado por um conjunto de N funções linearmente independentes, as chamadas *funções-base*, $\{\psi_j(t)\}$, definidas no intervalo de tempo $0 \leq t \leq T$. As funções-base devem satisfazer a condição

$$\int_0^T \psi_j(t)\psi_k(t)dt = \begin{cases} K_j & j = k \\ 0 & j \neq k \end{cases} \quad 0 \leq t \leq T \quad j, k = 1, 2, \dots, N \quad (1)$$

onde K_j é uma constante. Quando $K_j = 1$ o espaço de sinal diz-se *ortonormado* (o.n.), senão diz-se, simplesmente, *ortogonal*. A Eq. (1) diz-nos que as funções-base são ortogonais entre si, $\int_0^T \psi_j(t)\psi_k(t)dt = 0$, para $j \neq k$, e diz-nos também que num espaço o.n. cada função-base possui uma energia unitária, $\int_0^T \psi_j^2(t)dt = 1$.

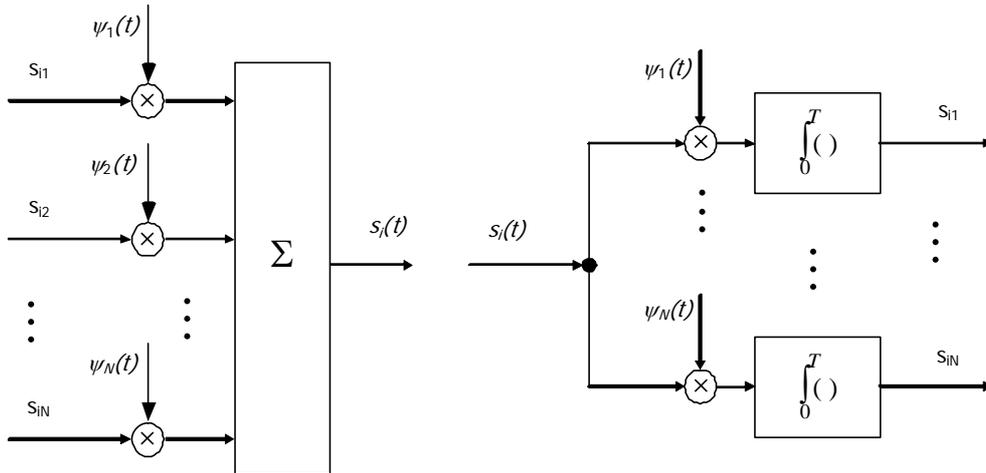
Consideremos M funções arbitrárias $s_i(t)$, com $M \geq N$. Qualquer delas pode ser obtida através de uma combinação linear das funções-base $\{\psi_j(t)\}$, desta maneira,

$$s_i(t) = \sum_{j=1}^N s_{ij} \psi_j(t) \quad i = 1, 2, \dots, M, \quad (2)$$

sendo os coeficientes s_{ij} dados por

$$s_{ij} = \frac{1}{K_j} \int_0^T s_i(t) \psi_j(t) dt \quad (3)$$

A figura seguinte ilustra as Eqs. (2) e (3) graficamente para um espaço o.n.:

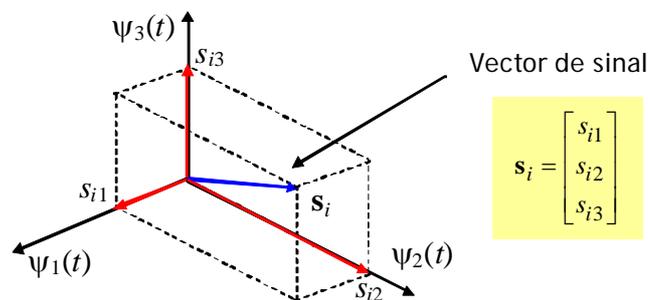


Daqui para a frente vamos considerar apenas espaços de sinal ortonormados ($K_j = 1$).

Formemos um vector-coluna de N elementos, que designaremos por \mathbf{s}_i , com os coeficientes s_{ij} :

$$\mathbf{s}_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix}$$

Num espaço de sinal de 3 dimensões obteríamos uma figura como esta:



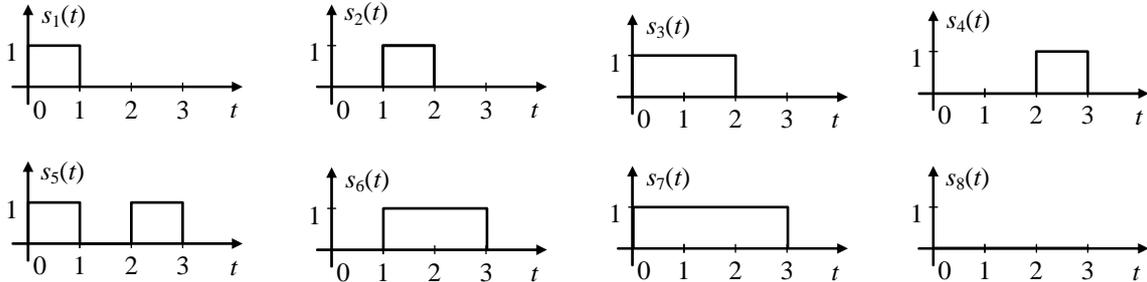
O quadrado do comprimento, ou norma, do vector é definido, naturalmente, por

$$\|\mathbf{s}_i\|^2 = \mathbf{s}_i^T \mathbf{s}_i = \sum_{j=1}^N s_{ij}^2.$$

Por outras palavras, a norma quadrática de \mathbf{s}_i é igual ao produto interno do vector consigo próprio.

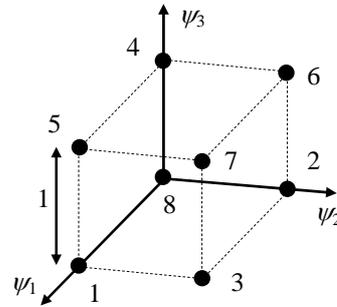
Exemplo 1: Espaço ortogonal tridimensional

Consideremos os oito sinais da figura.



Usando as equações anteriores é fácil concluir que no espaço de sinal o.n. definido pelas funções-base $\psi_1(t) = s_1(t)$, $\psi_2(t) = s_2(t)$ e $\psi_3(t) = s_4(t)$ os vectores correspondentes aos sinais têm as coordenadas dos vértices de um cubo de aresta unitária, como se mostra.

$$\begin{aligned} \mathbf{s}_1 &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \mathbf{s}_2 &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \mathbf{s}_3 &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} & \mathbf{s}_4 &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ \mathbf{s}_5 &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \mathbf{s}_6 &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} & \mathbf{s}_7 &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} & \mathbf{s}_8 &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$



Exemplo 2: Funções sinusoidais

Consideremos as seguintes quatro sinusóides de igual amplitude e frequência, fase ϕ_i diferente e duração limitada ao intervalo $0 \leq t \leq T$:

$$\begin{cases} s_i(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t - \phi_i) & 0 \leq t \leq T \\ 0 & \text{outros instantes} \end{cases} \quad \text{com} \quad \begin{cases} \phi_0 = \pi/4 \\ \phi_1 = 3\pi/4 \\ \phi_2 = -3\pi/4 \\ \phi_3 = -\pi/4 \end{cases}$$

Suponhamos que as funções-base do espaço o.n. são duas:

$$\begin{cases} \psi_1(t) = \sqrt{\frac{2}{T}} \cos 2\pi f_c t & 0 \leq t \leq T \\ 0 & \text{outros instantes} \end{cases} \quad \begin{cases} \psi_2(t) = \sqrt{\frac{2}{T}} \sin 2\pi f_c t & 0 \leq t \leq T \\ 0 & \text{outros instantes} \end{cases}$$

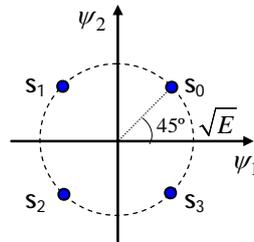
Como é a constelação daquelas quatro sinusóides? Ora desenvolvamos a expressão

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t - \phi_i):$$

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t - \phi_i) = \sqrt{\frac{2E}{T}} \cos 2\pi f_c t \cos \phi_i + \sqrt{\frac{2E}{T}} \sin 2\pi f_c t \sin \phi_i =$$

$$= (\sqrt{E} \cos \phi_i) \psi_1(t) + (\sqrt{E} \sin \phi_i) \psi_2(t)$$

Já estamos a ver que $s_{i1} = \sqrt{E} \cos \phi_i$ e $s_{i2} = \sqrt{E} \sin \phi_i$, ou $\mathbf{s}_i = \begin{bmatrix} \sqrt{E} \cos \phi_i \\ \sqrt{E} \sin \phi_i \end{bmatrix}$, a que corresponde a constelação de quatro pontos seguinte:



Qual é a energia de qualquer dos sinais sinusoidais no intervalo $0 \leq t \leq T$? É igual à sua potência vezes o intervalo de tempo, claro:

$$E_i = \int_0^T s_i^2(t) dt = \frac{(\sqrt{2E/T})^2}{2} T = E.$$

E a que distância estão os pontos da constelação da origem dos eixos? Estão todos a \sqrt{E} , como vemos. Parece, portanto, que a energia de um sinal é igual ao quadrado da distância do seu ponto à origem... De facto é sempre assim, como se mostra a seguir.

1.1. Correspondência entre energia e norma quadrática

Sabemos que, por definição, a energia do sinal $s_i(t)$ no intervalo $0 \leq t \leq T$ é dada por $E_i = \int_0^T s_i^2(t) dt$. Ora, tendo em conta a Eq. (2) obtemos sucessivamente

$$E_i = \int_0^T s_i^2(t) dt = \int_0^T \left[\sum_{j=1}^N s_{ij} \psi_j(t) \right] \left[\sum_{k=1}^N s_{ik} \psi_k(t) \right] dt =$$

$$= \sum_{j=1}^N \sum_{k=1}^N s_{ij} s_{ik} \int_0^T \psi_j(t) \psi_k(t) dt$$

Segundo a Eq. (1) o integral é nulo se $j \neq k$ e igual a 1 num espaço o.n. se $j = k$. Então concluímos que

$$E_i = \sum_{j=1}^N s_{ij}^2 = \|\mathbf{s}_i\|^2.$$

Esta é uma conclusão muito importante:

A energia de um sinal é igual à norma quadrática do vector que o representa.

Idêntico procedimento nos conduziria às seguintes expressões:

(correlação entre sinais) $\int_0^T s_i(t)s_k(t)dt = \mathbf{s}_i^T \mathbf{s}_k$

(coeficiente de correlação) $\rho = \frac{1}{\sqrt{E_i E_k}} \int_0^T s_i(t)s_k(t)dt = \frac{\mathbf{s}_i^T \mathbf{s}_k}{\|\mathbf{s}_i\| \|\mathbf{s}_k\|} = \cos \theta_{ik}$

Na expressão precedente θ_{ik} é o ângulo entre vectores. Note-se que normalmente é mais fácil usar os vectores do que os sinais para calcular o coeficiente de correlação ρ .

1.2. Distância euclidiana entre dois pontos

Imaginemos dois pontos representados pelos vectores s_i e s_k . A distância quadrática entre eles - também chamada *distância euclidiana quadrática* - é igual a

$$d_{ik}^2 = \|\mathbf{s}_i - \mathbf{s}_k\|^2 = \sum_{j=1}^N (s_{ij} - s_{kj})^2 = \int_0^T [s_i(t) - s_k(t)]^2 dt .$$

O que representa o integral da expressão? Representa a energia, E_d , do sinal-diferença $s_i(t) - s_k(t)$. Então $d_{ik}^2 = E_d$. Obtivemos nova conclusão importante:

A energia da diferença de dois sinais é igual à distância quadrática entre os correspondentes pontos do espaço ortonormado.

Com esta conclusão vai-se tornando aparente que é mais simples trabalhar com pontos e vectores do que com as formas de onda originais.

2. Modulações digitais

As modulações digitais são designadas por siglas em inglês:

- Amplitude: ASK ("Amplitude Shift-Keying")
- Fase: PSK ("Phase Shift-Keying")
- Frequência: FSK ("Frequency Shift-Keying")
- Fase/amplitude: QAM ("Quadrature Amplitude Modulation")

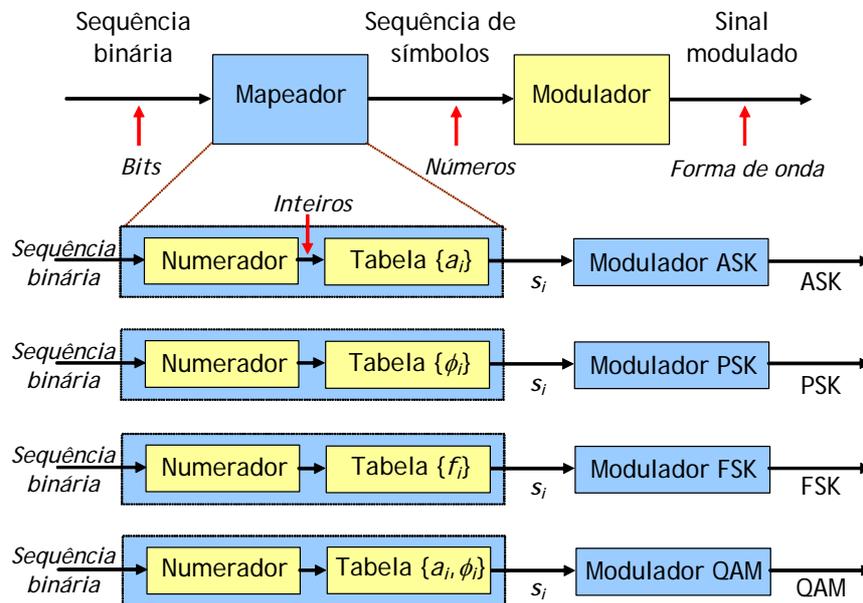
Ao contrário do que sucede nas modulações analógicas, o sinal modulador numa modulação digital é um sinal discreto, ou seja, é um sinal que, em cada intervalo de tempo (de símbolo), toma apenas um de M valores possíveis. O sinal modulador é, assim, uma sequência de valores - sim, números! - extraídos de um *alfabeto* de tamanho M . Esta sequência de valores, ou símbolos, é obtida a partir de uma sequência original de bits: se agruparmos os bits dois a dois obtemos símbolos com quatro valores possíveis que, por exemplo, podem representar frequências, como aqui:

Binário	01	11	10	11	00	01	01	00	11	10	10	00
Decimal	1	3	2	3	0	1	1	0	3	2	2	0
Frequências	800	1600	1200	1600	400	800	800	400	1600	1200	1200	400

se os agruparmos três a três obtemos símbolos com oito valores possíveis que, por exemplo, podem ser ângulos, como aqui:

Binário	011	110	110	001	010	011	101	000
Decimal	3	6	6	1	2	3	5	0
Fases	90°	180°	180°	45°	135°	90°	270°	0°

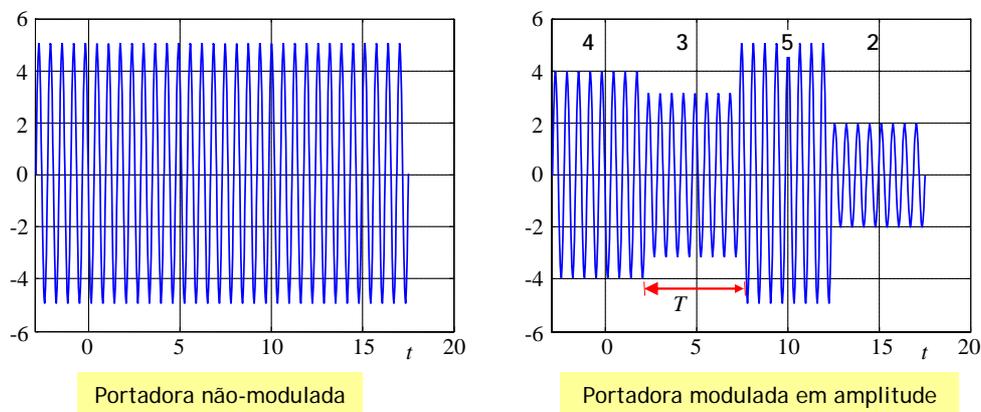
e assim por diante (é claro que se não agruparmos os bits temos dois valores apenas), ou seja, se agruparmos os bits k a k obtemos símbolos de um alfabeto de $M = 2^k$ valores. Portanto, consoante se trate de modulação de amplitude, fase ou frequência assim a sequência binária é mapeada numa sequência numerada de M amplitudes $\{a_i\}$, fases $\{\phi_i\}$ ou frequências $\{f_i\}$, respectivamente, como se esquematiza na figura seguinte:



Vale a pena realçar esta ideia:

É a sequência de símbolos que vai modular a onda portadora, não a sequência de bits original.

Vamos considerar que cada símbolo dura T segundos, intervalo de tempo que está relacionado com o intervalo de bit T_b através de $T = kT_b$, naturalmente. A sequência de símbolos vai fazer variar uma ou mais características da onda portadora, que é sempre uma onda sinusoidal de alta frequência (à frequência chamemos f_c). Qualquer que seja a modulação usada, sabemos que em cada intervalo de T segundos a característica da portadora que transporta a informação corresponde a um certo símbolo. Dito de outro modo, em cada intervalo de T segundos a amplitude, a fase ou a frequência toma um dos M valores possíveis, como na figura seguinte para 8ASK. Por exemplo, com $M = 8$ e FSK (ou seja, 8FSK) a portadora pode assumir uma de 8 frequências possíveis em cada intervalo de símbolo T .



Como temos M formas de onda diferentes em cada T segundos podemos exprimi-las à custa de $N \leq M$ funções-base, que suporemos sempre sinusoidais, de duração limitada aos mesmos T segundos e de energia unitária - o que quer dizer que o espaço de sinal é ortonormado. As correspondentes constelações terão sempre, é claro, M pontos. Abreviando:

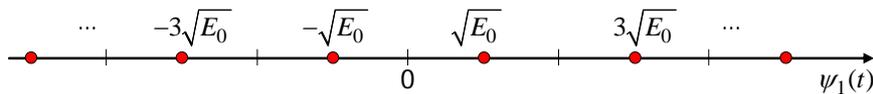
- Em ASK o espaço de sinal é unidimensional (só precisamos de uma função-base).
- Em MPSK e em QAM o espaço de sinal é bidimensional, excepto se se tratar de PSK binária, onde o espaço é unidimensional.
- Em MFSK o espaço de sinal tem M dimensões (precisamos de M funções-base).

$$N = \begin{cases} 1 & \text{ASK, BPSK} \\ 2 & \text{MPSK, QAM, BFSK} \\ M & \text{MFSK} \end{cases}$$

Em seguida vamos passar em revista cada modulação no que diz respeito às funções-base, formas de onda e constelações. Resumindo e concluindo...

2.1. Modulação de amplitude (ASK)

- Funções-base (uma): $\psi_1(t) = \sqrt{\frac{2}{T}} \cos 2\pi f_c t$, $0 \leq t \leq T$.
- Forma de onda: $s_i(t) = a_i \sqrt{E_0} \psi_1(t)$, $0 \leq t \leq T$, com $a_i = \pm 1, \pm 3, \dots$
- Coordenadas dos pontos da constelação: $\mathbf{s}_i = [a_i \sqrt{E_0}]$.
- Constelação unidimensional:



2.2. Modulação de fase (MPSK)

- Funções-base (duas):

$$\begin{aligned} \psi_1(t) &= \sqrt{\frac{2}{T}} \cos 2\pi f_c t \\ \psi_2(t) &= \sqrt{\frac{2}{T}} \sin 2\pi f_c t \end{aligned} \quad 0 \leq t \leq T$$

- Forma de onda:

$$\begin{aligned} s_i(t) &= \sqrt{\frac{2E}{T}} \cos(2\pi f_c t - \phi_i) = \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t - \frac{2\pi i}{M}\right) = \\ &= \underbrace{(\sqrt{E} \cos \phi_i)}_{s_{i1}} \psi_1(t) + \underbrace{(\sqrt{E} \sin \phi_i)}_{s_{i2}} \psi_2(t) \end{aligned} \quad \begin{aligned} &0 \leq t \leq T \\ &i = 1, 2, \dots, M \end{aligned}$$

- Coordenadas dos pontos da constelação:

$$\mathbf{s}_i = \begin{bmatrix} \sqrt{E} \cos \phi_i \\ \sqrt{E} \sin \phi_i \end{bmatrix}.$$

São pontos numa circunferência de raio \sqrt{E} centrada na origem.

- Constelação bidimensional (dois exemplos):

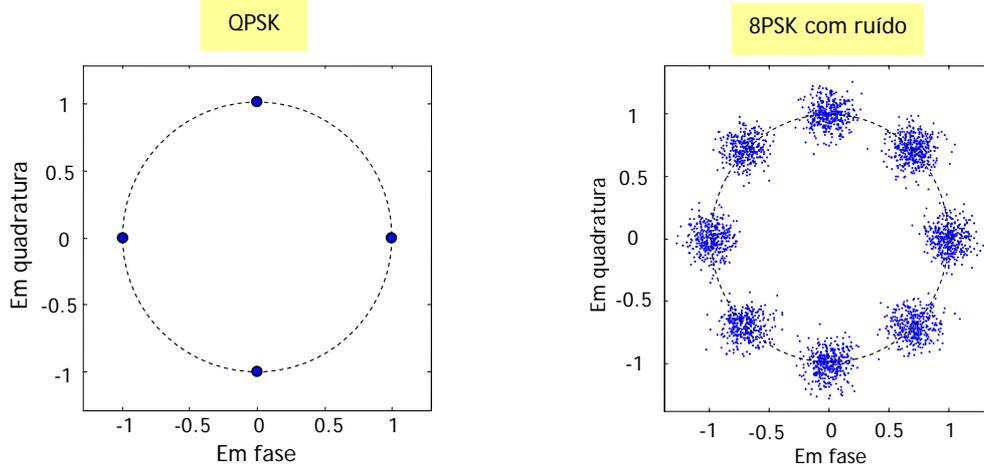
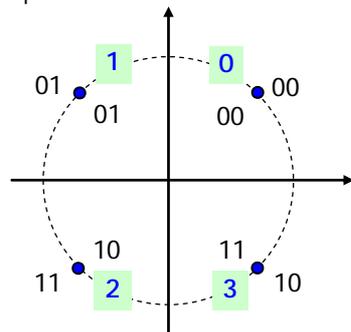


Fig. 1 Constelações de QPSK e 8PSK.

**Exemplo 3:
Mapeamentos em QPSK**

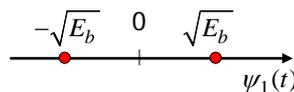
Na figura seguinte é apresentada a constelação de QPSK com dois mapeamentos bits-símbolos. O mapeamento exterior – mapeamento de Gray – é preferível ao mapeamento natural pois minimiza a probabilidade de bit errado¹.



Mapeamento interior: natural
Mapeamento exterior: Gray

2.2.1. Caso particular: BPSK

- Funções-base (uma): $\psi_1(t) = \sqrt{\frac{2}{T_b}} \cos 2\pi f_c t$, $0 \leq t \leq T_b$ (não esquecer que agora $T = T_b$)
- Forma de onda: $s_i(t) = \pm \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_c t = \pm \sqrt{E_b} \psi_1(t)$, $0 \leq t \leq T_b$.
- Coordenadas dos dois pontos da constelação: $\mathbf{s}_i = [\pm \sqrt{E_b}]$.
- Constelação unidimensional:



¹ Este é um tópico que poderá ser abordado noutra TeleTexto.

2.3. Modulação simultânea de fase e amplitude (M-QAM)

- Funções-base (duas):

$$\begin{aligned}\psi_1(t) &= \sqrt{\frac{2}{T}} \cos 2\pi f_c t \\ \psi_2(t) &= \sqrt{\frac{2}{T}} \sin 2\pi f_c t\end{aligned} \quad 0 \leq t \leq T$$

São as mesmas funções-base de MPSK.

- Forma de onda:

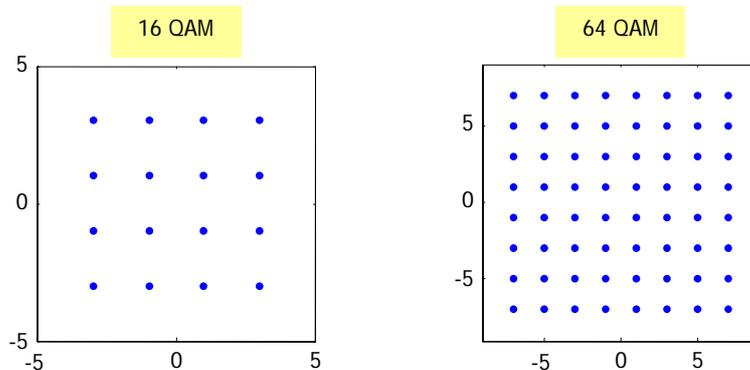
$$\begin{aligned}s_i(t) &= \sqrt{\frac{2E_0}{T}} a_i \cos 2\pi f_c t + \sqrt{\frac{2E_0}{T}} b_i \sin 2\pi f_c t = \\ &= \underbrace{a_i \sqrt{E_0}}_{s_{i1}} \psi_1(t) + \underbrace{b_i \sqrt{E_0}}_{s_{i2}} \psi_2(t)\end{aligned} \quad 0 \leq t \leq T$$

a_i e b_i têm valores independentes $\pm 1, \pm 3, \dots, \pm(L-1)$, com $L = \sqrt{M}$, e E_0 é metade da energia do ponto mais próximo da origem dos eixos.

- Coordenadas dos pontos da constelação:

$$\mathbf{s}_i = \begin{bmatrix} a_i \sqrt{E_0} \\ b_i \sqrt{E_0} \end{bmatrix} \quad \begin{array}{l} a_i = \pm 1, \pm 3, \dots \\ b_i = \pm 1, \pm 3, \dots \end{array} \quad (\text{pontos numa grelha quadrada})$$

- Constelação bidimensional (dois exemplos):



2.4. Modulação de frequência (MFSK)

- Funções-base (M): $\psi_i(t) = \sqrt{\frac{2}{T}} \cos 2\pi f_i t$, $0 \leq t \leq T$, $i = 1, 2, \dots, M$ (com $f_{i+1} - f_i = 1/T$)
- Forma de onda: $s_i(t) = \sqrt{\frac{2E}{T}} \cos 2\pi f_i t = \sqrt{E} \psi_i(t)$, $0 \leq t \leq T$, $i = 1, 2, \dots, M$

- Coordenadas dos pontos da constelação:

$$\mathbf{s}_i = \begin{bmatrix} 0 \\ \vdots \\ \sqrt{E} \\ \vdots \\ 0 \end{bmatrix} \quad (\text{elemento não-nulo na posição } i)$$

- Constelação: M -dimensional. Cada ponto está situado num dos M eixos à distância \sqrt{E} da origem.

2.4.1. Caso particular: BFSK

- Funções-base (duas):

$$\psi_1(t) = \sqrt{\frac{2}{T_b}} \cos 2\pi f_1 t \quad 0 \leq t \leq T_b \quad (\text{não esquecer que agora } T = T_b)$$

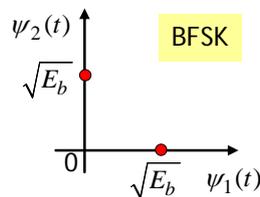
$$\psi_2(t) = \sqrt{\frac{2}{T_b}} \cos 2\pi f_2 t$$

- Forma de onda: $s_i(t) = \sqrt{\frac{2E_b}{T_b}} \cos 2\pi f_i t = \sqrt{E_b} \psi_i(t) \quad 0 \leq t \leq T_b \quad (\text{com } f_2 - f_1 = \Delta f = 1/T_b)$
 $i = 1, 2$

- Coordenadas dos pontos da constelação:

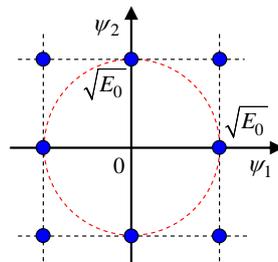
$$\mathbf{s}_1 = \begin{bmatrix} \sqrt{E_b} \\ 0 \end{bmatrix} \quad \mathbf{s}_2 = \begin{bmatrix} 0 \\ \sqrt{E_b} \end{bmatrix}$$

- Constelação bidimensional:



2.5. Outras constelações

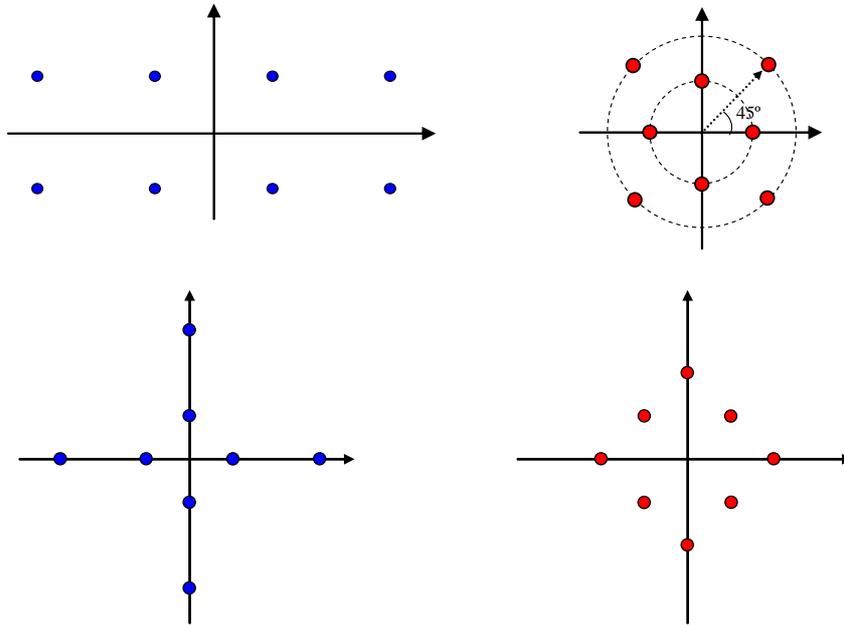
As constelações que vimos até agora não são as únicas que podemos usar. É certo que apresentam uma geometria regular mas em casos particulares pode ser interessante ou até conveniente utilizar outras. Um exemplo é apresentado na figura seguinte.



Relativamente à constelação 8PSK que se obteria se todos os pontos estivessem na circunferência tracejada podemos apontar uma vantagem e um inconveniente. A vantagem é que, estando os quatro pontos dos cantos mais afastados dos outros, a probabilidade do decisor se enganar nos símbolos é maior. A desvantagem é que, pelo mesmo motivo apontado aos pontos dos

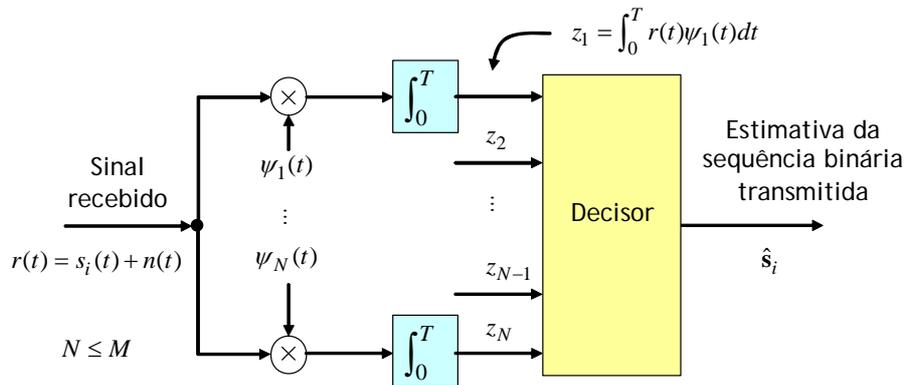
cantos, estes possuem uma maior energia e com isso elevam a energia média do sinal, o que não é, normalmente, desejável.

Eis outras constelações, para terminar:



3. Epílogo

O sinal modulado atravessa o canal e chega ao receptor, que tenta estimar a sequência de símbolos enviada. Como fazê-lo? Por exemplo, usando detecção coerente, ou seja, aquela que obriga o receptor a conhecer a frequência e a fase do sinal recebido. O receptor coerente dispõe de um ou mais correlacionadores como os da figura seguinte:



Sem ruído as saídas dos integradores nos sucessivos instantes de amostragem apresentam-nos as coordenadas dos pontos da constelação:

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix}$$

Com ruído já não obtemos esses pontos mas sim nuvens de pontos à sua volta, como na Fig. 1 (porquê à sua volta? Porque como o valor médio do ruído é nulo, os valores médios das saídas acabam por ser os pontos da constelação).

A pergunta que fica é: como estimar o símbolo s_i a partir do sinal recebido $r(t)$? Ou doutra maneira: como estimar o símbolo s_i a partir do vector recebido \mathbf{z} ? Observemos a figura seguinte. Se recebermos o vector \mathbf{z} assinalado que ponto da constelação devemos escolher? O ponto 0? Por estar mais perto de \mathbf{z} ? E se os símbolos não forem equiprováveis, continuaremos a escolher o ponto 0? As respostas ficam para outro texto...

