

Parte II

Modulação

5 Transmissão Digital

Neste capítulo, a teoria da transmissão digital de sinais é revisada. Esta teoria fornece a base matemática para a construção de sistemas que permitam transportar, da maneira mais eficiente possível, uma informação digital através de um canal de forma de onda.

Este capítulo está organizado da seguinte forma. A Seção 5.1 desenvolve um modelo vetorial para a transmissão de sinais através de um canal limitado em banda sujeito à adição de ruído gaussiano branco. Uma extensão desse modelo para canais com outros tipos de perturbações é também discutida nesta seção. Na Seção 5.2, os parâmetros utilizados para a comparação entre diferentes sistemas de transmissão são apresentados. Alguns exemplos de sistemas que podem ser construídos a partir desta teoria são descritos na Seção 5.3.

5.1 Modelo Vetorial

Considere o modelo de um sistema de comunicações mostrado na Fig. 5.1 [39]. A cada intervalo de duração T , uma dentre M mensagens equiprováveis é escolhida para transmissão. Cada mensagem $i \in \{1, \dots, M\}$ está associada a um sinal $s_i(t)$, $0 < t \leq T$, pertencente a um conjunto $\mathcal{S} = \{s_i(t)\}_{i=1}^M$. O transmissor envia o sinal correspondente através de um canal ruidoso limitado a uma banda W , o qual transforma o sinal transmitido no sinal recebido $r(t)$. O receptor então processa o sinal recebido e decide que a mensagem \hat{i} foi transmitida.

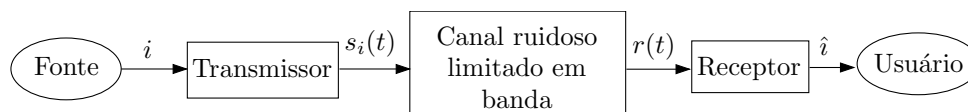


Figura 5.1: Modelo de um sistema de transmissão digital.

A abordagem tradicional para o projeto de transmissor e receptor é inicialmente substituir o canal de forma de onda por um canal vetorial (Fig. 5.2). Isso é realizado expandindo-se os sinais $s_i(t)$ em uma série de funções. Os sinais $s_i(t)$ podem então ser escritos como

$$s_i(t) = \sum_{j=1}^N x_{ij} \phi_j(t), \quad i = 1, \dots, M \quad (5-1)$$

onde os coeficientes x_{ij} são números reais, e as *funções de base* $\phi_j(t)$, $j = 1, \dots, N$, são ortonormais e limitadas a uma duração T e a uma banda W . É matematicamente possível encontrar um máximo de $N = 2TW$ funções distintas obedecendo a essas restrições.¹

Os coeficientes x_{ij} podem ser obtidos através da correlação do sinal $s_i(t)$ com as funções de base:

$$x_{ij} = \int_0^T s_i(t) \phi_j(t) dt, \quad j = 1, \dots, N. \quad (5-2)$$

Quando $s_i(t)$ é transmitido através de um canal com ruído aditivo gaussiano branco (AWGN), o sinal recebido é dado por

$$r(t) = s_i(t) + z(t) \quad (5-3)$$

onde $z(t)$ é um processo estocástico gaussiano de média nula e densidade espectral de potência $S_z(f) = N_0/2$. Mostra-se que quando o sinal recebido é decomposto nas funções de base, os coeficientes do vetor recebido são dados por

$$y_j = x_{ij} + z_j, \quad j = 1, \dots, N \quad (5-4)$$

onde z_j é uma variável aleatória gaussiana de média nula e variância $N_0/2$.

Este procedimento nos permite representar os sinais $s_i(t)$ e $r(t)$, respectivamente, como os vetores N -dimensionais $\mathbf{x}_i = (x_{i1}, \dots, x_{iN})^T$ e $\mathbf{y} = (y_1, \dots, y_N)^T$ definidos no espaço euclidiano \mathbb{R}^N . O conjunto de todos os vetores de transmissão $\mathcal{C} = \{\mathbf{x}_i\}_{i=1}^M$ é denominado *constelação* ou *código*. Cada vetor \mathbf{x}_i é também chamado de *palavra do código*. A expressão para o vetor recebido se torna portanto

$$\mathbf{y} = \mathbf{x}_i + \mathbf{z} \quad (5-5)$$

¹Note que o número de funções ortonormais de duração T que podem ser acomodadas em uma banda W depende das definições de “banda” e de “duração”, uma vez que nenhuma função é simultaneamente limitada no tempo e na frequência. A expressão $N = 2TW$ representa um limite teórico [31, 39], conveniente para fins de comparação.

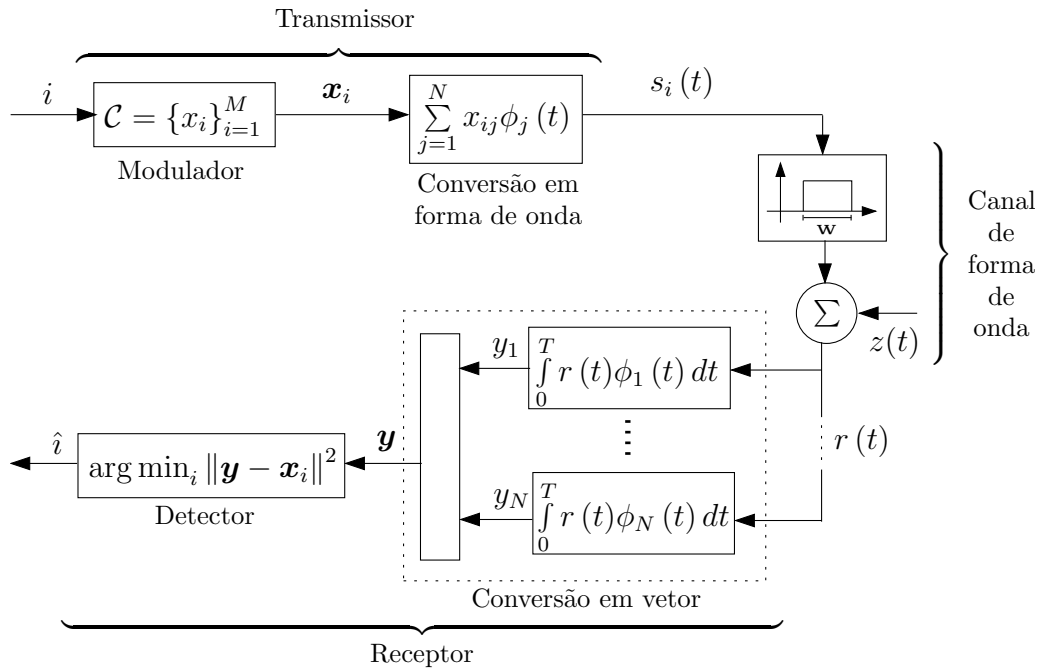


Figura 5.2: Modelo de um sistema de transmissão utilizando formas de onda associadas a vetores.

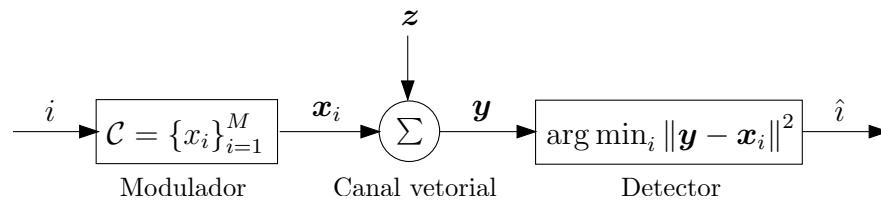


Figura 5.3: Modelo vetorial de um sistema de transmissão.

onde $\mathbf{z} = (z_1, \dots, z_N)^T$ é um vetor aleatório gaussiano com componentes independentes de média nula e variância $N_0/2$.

Quando o vetor \mathbf{y} é recebido, o receptor de máxima verossimilhança decide que foi transmitida a mensagem \hat{i} associada ao vetor de \mathcal{C} que está à menor distância euclidiana de \mathbf{y} , ou seja,

$$\hat{i} = \arg \min_{i \in \{1, \dots, M\}} \|\mathbf{y} - \mathbf{x}_i\|^2. \quad (5-6)$$

Como as mensagens são equiprováveis, este receptor é também o que minimiza a probabilidade de uma decisão errada.

O modelo de transmissão que utiliza formas de onda associadas a vetores e o modelo vetorial simplificado são ilustrados nas figuras 5.2 e 5.3.

Note que o modelo apresentado nesta seção considera uma única mensagem transmitida i em um intervalo de transmissão $0 < t \leq T$, mas pode ser diretamente estendido para a transmissão de várias mensagens i_k em intervalos $(k - 1)T < t \leq kT$.

5.1.1 Canal com imperfeições

O desenvolvimento acima assume que o canal tem como única perturbação o ruído aditivo. Outros tipos de perturbações podem ser incorporados ao modelo através de uma *matriz de imperfeições* H que multiplica o vetor transmitido \mathbf{x}_i [19]. A expressão para o vetor recebido pode então ser escrita como

$$\mathbf{y} = H\mathbf{x}_i + \mathbf{z} \quad (5-7)$$

onde H é uma matriz real $n \times n$. Note que o modelo permite que a matriz H varie a cada intervalo de transmissão.

Assumindo que o estado do canal é conhecido pelo receptor (estimação de canal perfeita), a detecção de máxima verossimilhança é dada por

$$\hat{i} = \arg \min_{i \in \{1, \dots, M\}} \|\mathbf{y} - H\mathbf{x}_i\|^2. \quad (5-8)$$

Alguns tipos de imperfeições que podem ser caracterizadas por esse modelo são o *desvanecimento* e a *seletividade em frequência*. Neste trabalho, estamos apenas interessados no caso de desvanecimento sem seletividade em frequência.

Neste caso, quando as funções de base $\phi_j(t)$ são escolhidas de forma a não se sobreporem no tempo (por exemplo, o mesmo pulso definido em intervalos disjuntos), a matriz de imperfeições pode ser caracterizada como

$$H = \begin{pmatrix} h_1 & 0 & \dots & 0 \\ 0 & h_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & h_N \end{pmatrix}. \quad (5-9)$$

O canal é dito com desvanecimento *rápido* quando os coeficientes h_1, \dots, h_N podem ser diferentes, e com desvanecimento *lento* quando $H = hI$, ou seja, $h_1 = \dots = h_N = h$. Note que quando $H = I$ o modelo se reduz ao do canal AWGN.

5.2 Parâmetros de Comparação

Na escolha de um sistema de comunicações, estamos interessados no compromisso entre três grandezas: a ocupação espectral, a confiabilidade da transmissão e a razão entre a potência de sinal e o nível espectral de ruído

requerida na recepção para que se atinja tal confiabilidade. Estes parâmetros são tradicionalmente expressos através da *eficiência espectral* $\eta = R_b/W$, da probabilidade de erro de símbolo $P(e)$ e da razão E_b/N_0 , respectivamente, onde R_b é a taxa de transmissão em bits por segundo, $E_b = \mathcal{P}/R_b$ é a energia transmitida por bit e \mathcal{P} é a potência de transmissão.²

Nesta seção, desenvolvemos algumas versões desses parâmetros que são mais adequadas à comparação de códigos de diferentes dimensões N .

Taxa de informação

A *taxa de informação* de um código, em bits por uso de canal unidimensional, é definida como

$$R \triangleq \frac{1}{N} \log M \quad (\text{bits/uso de canal 1-D}). \quad (5-10)$$

A justificativa para R é baseada na eficiência espectral. Uma vez que $R_b = \frac{1}{T} \log M$ e $N = 2TW$, vemos que $\eta = R_b/W = 2R$, ou seja, a eficiência espectral é proporcional a taxa de informação R .

Energia média

A *energia média por palavra por dimensão* de um código é definida como

$$E \triangleq \frac{1}{NM} \sum_{i=1}^M \|\mathbf{x}_i\|^2 \quad (5-11)$$

Como as palavras \mathbf{x}_i são usadas com igual probabilidade, a energia média por palavra pode ser escrita como $E_s = \frac{1}{M} \sum_{i=1}^M \|\mathbf{x}_i\|^2$. Uma vez que $E_b = E_s/\log M$, encontramos que $E_b = NE/\log M = \frac{1}{R}E$, isto é, entre códigos com a mesma taxa R , aqueles que apresentam menor E contém menor energia por bit.

²Note que não utilizamos o parâmetro mais comum para medida de confiabilidade, a probabilidade de erro de bit $P_b(e)$, porque seu valor depende fortemente do mapeamento utilizado entre os bits de informação e os vetores \mathbf{x}_i . Enquanto, em algumas modulações, um mapeamento Gray pode ser utilizado, ou ainda, uma expressão simples para $P_b(e)$ existe, o mesmo não necessariamente acontece em modulações genéricas. Em todo caso, as probabilidades de erro de símbolo e de bit estão, obviamente, relacionadas: $P(e)/\log M \leq P_b(e) \leq P(e)$.

Distância mínima

A *distância euclidiana mínima* entre as palavras de um código é definida como

$$d_{\min} \triangleq \min_{i,j, i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\| \quad (5-12)$$

Para valores de razão sinal-ruído moderados a altos, uma aproximação satisfatória da probabilidade de erro de um sistema pode ser obtida a partir da distância mínima entre as palavras [24]. Assim, consideramos a distância mínima como nossa medida de confiabilidade neste trabalho.³

Para ilustrar a utilização desses parâmetros, suponha, por exemplo, que a confiabilidade de um sistema é dada por uma função de $d_{\min}^2/2N_0$, isto é

$$P(e) = f\left(\frac{d_{\min}^2}{2N_0}\right). \quad (5-13)$$

Aplicando as definições acima, obtemos

$$P(e) = f\left(\frac{d_{\min}^2}{2E} \frac{E}{N_0}\right) = f\left(\frac{d_{\min}^2}{2E} \frac{E_b}{N_0} R\right) = f\left(\gamma \frac{E_b}{N_0}\right) \quad (5-14)$$

onde

$$\gamma = \frac{d_{\min}^2}{2E} R. \quad (5-15)$$

Conclui-se que, para uma mesma taxa R (e, conseqüentemente, uma mesma eficiência espectral η), códigos com maior razão d_{\min}^2/E apresentam melhor desempenho.

Assim, com relação ao projeto de sistemas, estamos interessados em encontrar códigos que, para uma dada taxa R e uma distância mínima d_{\min} , possuem o menor valor de E possível; isto é, códigos que requerem o menor valor da razão E_b/N_0 tal que sejam obtidas a confiabilidade especificada por $d_{\min}^2/2N_0$ e a eficiência espectral especificada por R .

³Note que, em alguns casos, d_{\min} é um parâmetro ainda mais significativo que $P(e)$. Considere um sistema com um código corretor de erros projetado com o objetivo de se beneficiar da modulação utilizada (um código de Ungerboeck [36], por exemplo). Nesse caso, é fácil ver que o desempenho do sistema é muito mais influenciado pela distância entre os símbolos da modulação codificada do que pela probabilidade de erro de símbolo (ou de bits) da modulação sem codificação.

5.3

Exemplos

Diversos tipos de sistemas podem ser obtidos de acordo com a escolha das funções de base e da correspondente constelação. Por exemplo, considere $N = 2$ e as funções de base

$$\phi_1(t) = \phi^{(I)}(t) = \sqrt{2} \bar{g}(t) \cos(2\pi f_c t + \theta), \quad 0 < t \leq T \quad (5-16)$$

$$\phi_2(t) = \phi^{(Q)}(t) = -\sqrt{2} \bar{g}(t) \sin(2\pi f_c t + \theta), \quad 0 < t \leq T. \quad (5-17)$$

onde $\bar{g}(t)$ é o pulso formatador normalizado para energia unitária $\int_{-\infty}^{+\infty} |\bar{g}(t)|^2 dt = 1$. Esse é um sistema do tipo AM-PM. Uma modulação específica pode ser obtida combinando-se um sistema “base” com uma constelação ou código \mathcal{C} . Por exemplo, se

$$\mathcal{C} = \{(a, a)^T, (a, -a)^T, (-a, a)^T, (-a, -a)^T\} \quad (5-18)$$

então a modulação resultante é um QPSK.

Múltiplos intervalos

Um outro sistema de interesse (que pode ser entendido como um extensão do sistema descrito no exemplo anterior) pode ser obtido agrupando-se N_T intervalos de transmissão (consecutivos) de duração τ em um novo intervalo de transmissão de duração T . Isto é, consideramos $N = 2N_T$ e

$$\begin{aligned} \phi_{2j-1}(t) &= \phi_j^{(I)}(t) \\ &= \sqrt{2} \bar{g}(t) \cos(2\pi f_c(t - (j-1)\tau) + \theta), \quad (j-1)\tau < t \leq j\tau \end{aligned} \quad (5-19)$$

$$\begin{aligned} \phi_{2j}(t) &= \phi_j^{(Q)}(t) \\ &= -\sqrt{2} \bar{g}(t) \sin(2\pi f_c(t - (j-1)\tau) + \theta), \quad (j-1)\tau < t \leq j\tau \end{aligned} \quad (5-20)$$

onde $\tau = T/N_T$ e $j = 1, \dots, N_T$. Uma palavra \mathbf{x}_i desse sistema tem os seguintes componentes:⁴

$$\begin{matrix} & 1 & 2 & 3 & \dots & N_T \\ \begin{matrix} I \\ Q \end{matrix} & \begin{pmatrix} x_1 & x_3 & x_5 & \dots & x_{N-1} \\ x_2 & x_4 & x_6 & \dots & x_N \end{pmatrix} \end{matrix}. \quad (5-21)$$

⁴Note que \mathbf{x}_i é um vetor $N \times 1$. Os coeficientes x_1, \dots, x_N são apresentados em forma de matriz apenas para destacar sua formação.

Um outro sistema de interesse pode ser obtido a partir do sistema acima simplesmente descartando-se⁵ as funções em quadratura $\phi_j^{(Q)}(t)$. Nesse caso, temos que $N = N_T$ e as palavras do código tem componentes que correspondem a

$$I \begin{pmatrix} 1 & 2 & 3 & \dots & N_T \\ x_1 & x_2 & x_3 & \dots & x_N \end{pmatrix}. \quad (5-22)$$

Se, por exemplo, $N = 4$ e

$$\mathcal{C} = \{(a, 0, 0, 0)^T, (0, a, 0, 0)^T, (0, 0, a, 0)^T, (0, 0, 0, a)^T\} \quad (5-23)$$

então obtemos uma modulação PPM (*pulse position modulation*).

Múltiplas portadoras

Outra possibilidade de extensão é a extensão freqüencial: consideramos $N = N_W$ funções de base em N_W freqüências adjacentes com espaçamento W/N_W , ou seja,

$$\phi_j(t) = \phi_j^{(I)}(t) = \sqrt{2} \bar{g}(t) \cos(2\pi f_j t + \theta), \quad 0 < t \leq T \quad (5-24)$$

onde $f_j - f_{j-1} = W/N_W = 1/2T$ e $j = 1, \dots, N_W$. Nesse caso, as palavras \mathbf{x}_i tem os seguintes componentes

$$I \begin{pmatrix} 1 & 2 & 3 & \dots & N_W \\ x_1 & x_2 & x_3 & \dots & x_N \end{pmatrix}. \quad (5-25)$$

Novamente, se $N = 4$ e

$$\mathcal{C} = \{(a, 0, 0, 0)^T, (0, a, 0, 0)^T, (0, 0, a, 0)^T, (0, 0, 0, a)^T\} \quad (5-26)$$

então obtemos uma modulação 4-FSK. Note que os códigos em (5-23) e (5-26) são exatamente os mesmos; devem, portanto, possuir o mesmo desempenho em canal AWGN (*cf.* seção anterior). Este exemplo ilustra uma propriedade interessante dessa formalização: na avaliação do desempenho de uma modulação, não é necessário considerar a base de funções utilizada, basta considerar o *código*.

Se incluirmos no sistema acima as funções em quadratura $\phi_j^{(Q)}(t)$, obteremos um novo sistema com $N = 2N_W$ funções de base ocupando um

⁵Observa-se que, para um mesmo N , este sistema ocupa o dobro da banda do sistema anterior, pois metade das dimensões são “desperdiçadas.”

total de N_W frequências igualmente espaçadas de W/N_W , isto é,

$$\phi_{2j-1}(t) = \phi_j^{(I)}(t) = \sqrt{2} \bar{g}(t) \cos(2\pi f_j t + \theta), \quad 0 < t \leq T \quad (5-27)$$

$$\phi_{2j}(t) = \phi_j^{(Q)}(t) = -\sqrt{2} \bar{g}(t) \sin(2\pi f_j t + \theta), \quad 0 < t \leq T. \quad (5-28)$$

onde $f_j - f_{j-1} = W/N_W = 1/T$ e $j = 1, \dots, N_W$. Note que a faixa de separação frequencial aumenta em relação ao sistema anterior de forma a permitir que as funções em fase e quadratura sejam ortogonais. As palavras desse sistema tem componentes que podem ser representados de forma análoga a (5-21), como

$$\begin{matrix} & 1 & 2 & 3 & \dots & N_W \\ I & \left(\begin{matrix} x_1 & x_3 & x_5 & \dots & x_{N-1} \end{matrix} \right) \\ Q & \left(\begin{matrix} x_2 & x_4 & x_6 & \dots & x_N \end{matrix} \right) \end{matrix}. \quad (5-29)$$

Este tipo de sistema corresponde ao conhecido sistema OFDM. Como exemplo, suponha uma constelação para o QPSK dada por

$$\mathcal{A} = \{(a, a)^T, (a, -a)^T, (-a, a)^T, (-a, -a)^T\}. \quad (5-30)$$

O código para uma modulação OFDM/QPSK (isto é, um sistema OFDM com modulação de portadora QPSK) pode ser obtido como

$$\mathcal{C} = \overbrace{\mathcal{A} \times \mathcal{A} \times \mathcal{A} \times \dots \times \mathcal{A}}^{N_W}. \quad (5-31)$$

Uma vez desenvolvidas as extensões no tempo, na frequência e em quadratura, é fácil ver que outras duas combinações ainda podem ser obtidas: extensão no tempo e na frequência de uma portadora em fase, e extensão no tempo e na frequência de um par de portadoras em fase e quadratura. No primeiro caso, temos $N = 2TW = N_T N_W$ funções de base, e no segundo caso, $N = 2TW = 2N_T N_W$, sendo que em ambos os casos as funções têm duração $\tau = T/N_T$ e espaçamento de frequências $f_j - f_{j-1} = W/N_W$.

A discussão apresentada nesta seção visa ilustrar como as funções de base podem ser escolhidas de acordo com o sistema que se pretende obter. A partir de um mesmo código \mathcal{C} , diferentes sistemas podem ser obtidos de acordo com a base de funções utilizada, a qual pode ser escolhida tendo em vista a aplicação pretendida. No próximo capítulo, a não ser quando mencionado o contrário, desconsideramos a escolha da base de funções e nos restringimos apenas ao código utilizado.