

7 PRINCÍPIOS DE TRANSMISSÃO DE DADOS DIGITAIS

Durante a maior parte do século vinte, uma parcela significativa de sistemas de comunicação era de forma analógica. Contudo, ao final da década de 1990, o formato digital começou a dominar a maioria das aplicações. Não é necessário muito esforço para que percebamos a contínua migração de comunicação analógica para digital: do áudio em fita cassete para MP3 e CD, da TV analógica NTSC ou PAL-M para a digital HDTV, da telefonia tradicional para VoIP, e do vídeo em VHS para DVD. Na verdade, até o último refúgio analógico da radiodifusão enfrenta um poderoso competidor digital, na forma de rádio por satélite. Dada a importância dominadora de sistemas de comunicação digital em nosso dia a dia, nunca é cedo demais para estudar os princípios básicos e vários aspectos de comunicação de dados digitais, como faremos neste capítulo.

Este capítulo trata do problema da transmissão de dados digitais em um canal. Portanto, consideraremos que as mensagens iniciais são digitais. Iniciaremos com a consideração do caso binário, em que os dados consistem em somente dois símbolos: **1** e **0**. Aloquemos uma forma de onda (pulso) distinta a cada um desses símbolos. A resultante sequência de pulsos é transmitida através do canal. No receptor, esses pulsos são detectados e convertidos aos dados binários (**1s** e **0s**).

7.1 SISTEMAS DE COMUNICAÇÃO DIGITAL

Um sistema de comunicação digital consiste em diversos componentes, como indicado na Fig. 7.1. Nesta seção, delinearemos, de forma conceitual, a funcionalidade de cada componente em sistemas de comunicação. Os detalhes de análise e projeto serão apresentados posteriormente no capítulo, em seções dedicadas.

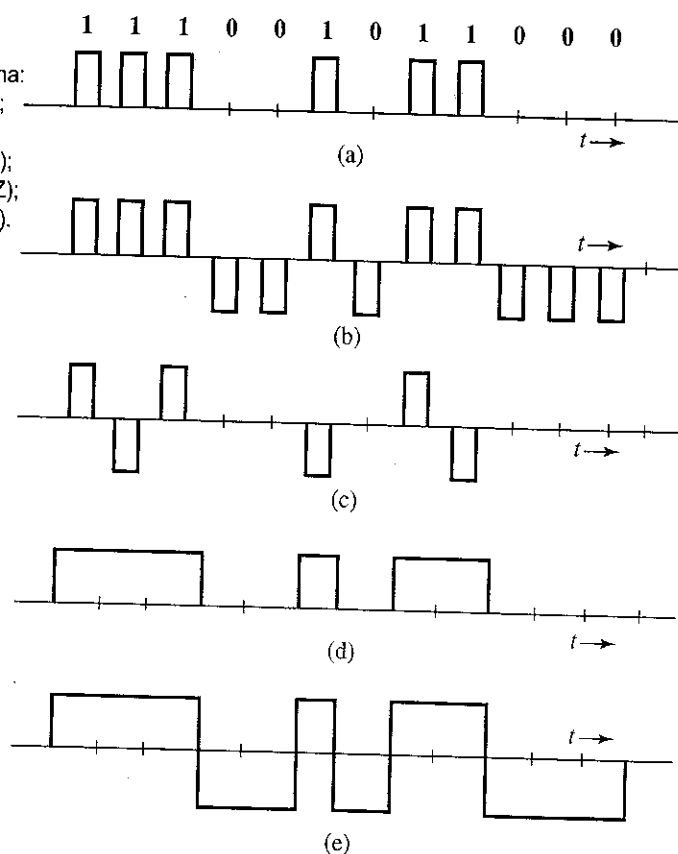
7.1.1 Fonte

A entrada de um sistema digital toma a forma de uma sequência de dígitos. A entrada pode ser a saída de um conjunto de dados, de um computador, um sinal de áudio digitalizado (PCM, DM ou LPC), fac-símile digital, HDTV, dados de telemetria, e assim por diante. Embora a maior parte da discussão neste capítulo seja restrita ao caso binário (esquemas de comunicação que usam apenas dois símbolos), o caso mais geral de comunicação M -ária, que utiliza M símbolos, também será discutido nas Seções 7.7 e 7.9.

Figura 7.1
Blocos fundamentais de um sistema de comunicação digital.



Figura 7.2
Exemplos de
códigos de linha:



7.1.2 Codificador de Linha

A saída digital de um codificador de fonte é convertida (ou codificada) em pulsos elétricos (formas de onda) para fins de transmissão no canal. Esse processo recebe a denominação **codificação de linha** ou **codificação de transmissão**. Há muitas formas possíveis de alocar formas de onda (pulsos) aos dados digitais. No caso binário (2 símbolos), por exemplo, a codificação de linha conceitualmente mais simples é **on-off** (liga-desliga), em que 1 é transmitido por um pulso $p(t)$ e 0 é transmitido por ausência de pulso (sinal nulo), como mostrado na Fig. 7.2a. Outro código de uso muito comum é o **polar**, em que 1 é transmitido por um pulso $p(t)$ e 0 é transmitido por um pulso $-p(t)$ (Fig. 7.2b). O esquema polar é o código com a maior eficiência de potência, pois, para um dado grau de imunidade ao ruído (probabilidade de erro), requer a menor potência. Outro código popular em PCM é o **bipolar**, também conhecido como **pseudoternário** ou **inversão alternada de sinal** (AMI – *alternate mark inversion*), em que 0 é codificado por ausência de pulso e 1 é codificado por um pulso $p(t)$ ou $-p(t)$, dependendo se o 1 anterior foi codificado por $-p(t)$ ou $p(t)$. Em resumo, pulsos que representam 1s sucessivos têm sinais alternados, como indicado na Fig. 7.2c. Esse código tem a vantagem de que, se ocorrer um único erro na detecção de pulsos, a sequência de pulsos recebida violará a regra bipolar e o erro poderá ser detectado imediatamente (embora não possa ser corrigido).*

Outro código de linha que, inicialmente, pareceu promissor é o duobinário (e duobinário modificado), proposto por Lender.^{1,2} Esse código é melhor que o bipolar em termos de eficiência de largura de banda. Sua versão mais proeminente, o código de linha *duobinário modificado*, tem sido aplicada em canais de leitura de hard disc drives, em transmissão óptica a 10 Gbit/s para redes metropolitanas, e na

*Isso admite não mais que um erro em sequência. Múltiplos erros em sequência podem cancelar os respectivos efeitos e, assim, não serem detectados. Contudo, a probabilidade de múltiplos erros é muito menor que a de apenas um erro. Mesmo para um erro, não podemos identificar sua localização exatamente. Portanto, esse código detecta a presença de erros isolados, mas não é capaz de corrigi-los.

primeira geração de modems para rede digital de serviços integradas (ISDN – *integrated services digital networks*). Detalhes de códigos de linha duobinários serão discutidos mais adiante neste capítulo.

Até aqui, em nossa discussão, usamos pulsos de meia largura somente para facilitar a ilustração. Podemos selecionar outras larguras. Pulsos de largura completa são usados em algumas aplicações. Sempre que pulsos de largura completa são utilizados, a amplitude do pulso é mantida em um valor constante durante toda a duração do pulso (isto é, a amplitude não retorna a zero antes do início do pulso seguinte). Por essa razão, tais esquemas são denominados **sem retorno ao zero** ou esquemas **NRZ** (*non-return-to-zero*), em contraste com esquemas com **retorno ao zero** ou **RZ** (*return-to-zero*) (Fig. 7.2a-c). A Fig. 7.2d mostra um sinal NRZ on-off e a Fig. 7.2e, um sinal NRZ polar.

7.1.3 Multiplexador

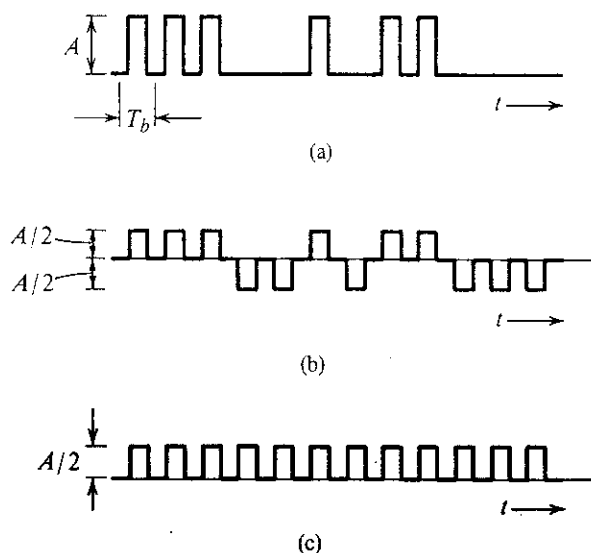
Em termos gerais, a capacidade de um canal físico (por exemplo, cabo coaxial, fibra óptica) para transmitir dados é muito maior que a taxa de dados de uma fonte. Para utilizar essa capacidade de modo eficaz, combinamos várias fontes por meio de um multiplexador digital. A multiplexação digital pode ser feita por divisão em frequência ou por divisão no tempo, como já discutimos. A divisão em código (a ser discutida no Capítulo 11) também é uma alternativa prática e eficiente. Um canal físico é, portanto, compartilhado por várias mensagens simultâneas.

7.1.4 Repetidor Regenerador

Repetidores regeneradores são usados em intervalos regularmente espaçados, ao longo de uma linha de transmissão digital, para detectar o sinal digital que chega e regenerar novos pulsos “limpos” para transmissão em mais um trecho da linha. Esse processo periódico elimina e, portanto, combate, o acúmulo de ruído e distorção de sinal na extensão da rota de transmissão. A capacidade de eliminação de ruído e efeitos de distorção de sinal desses repetidores regeneradores é uma das maiores vantagens de sistemas de comunicação digital em relação aos sistemas analógicos.

Se forem transmitidos pulsos a uma taxa de R_b pulsos por segundo, é necessário ter informação de temporização periódica – o sinal de relógio na frequência R_b Hz – para amostrar os pulsos que chegam ao receptor. Essa informação de temporização pode ser extraída do próprio sinal recebido, desde que o código de linha seja escolhido de forma adequada. A retificação do sinal polar RZ na Fig. 7.2b, por exemplo, produz um sinal de relógio periódico de frequência R_b Hz, que contém o desejado sinal periódico de temporização de frequência R_b Hz. Quando esse sinal é aplicado a um circuito ressonan-

Figura 7.3
Sinal on-off (a) é a soma de um sinal polar randômico (b) com um sinal periódico na frequência do relógio (c).



te sintonizado na frequência R_b , a saída, que é uma senoide de frequência R_b Hz, pode ser usada para a temporização. O sinal on-off pode ser expresso como a soma de um sinal periódico (na frequência do relógio) e um sinal polar ou randômico, como mostrado na Fig. 7.3. Devido à presença da componente periódica, podemos extrair a informação de temporização desse sinal com o uso de um circuito ressonante sintonizado na frequência do relógio. Um sinal bipolar, ao ser retificado, se torna um sinal on-off. Portanto, a informação de temporização pode ser extraída da mesma forma como para um sinal on-off.

O sinal de temporização (saída do circuito ressonante) é sensível ao padrão de bit de chegada. No caso on-off ou polar, um 0 é transmitido por "ausência de pulso". Assim, se houver um grande número de 0s em uma sequência (ausência de pulsos), não haverá sinal na entrada do circuito ressonante e sua saída senoidal passa a decair, originando erro na informação de temporização. Mais adiante, discutiremos formas de solucionar esse problema. Dizemos que código de linha é **transparente** quando o padrão de bits não afeta a precisão da informação de temporização. O esquema polar RZ (em que cada bit é transmitido por algum pulso) é transparente; os esquemas on-off e bipolar não são transparentes.

7.2 CODIFICAÇÃO DE LINHA

Dados digitais podem ser transmitidos por vários **códigos de linha** ou de **transmissão**. Vimos exemplos de códigos on-off, polar e bipolar. Cada código de linha tem suas vantagens e desvantagens. Entre outras propriedades desejáveis, um código de linha deve ter as seguintes:

- *Largura de banda de transmissão* tão pequena quanto possível.
- *Eficiência de potência*. Para dadas largura de banda e taxa de erros de detecção, a potência transmitida deve ser tão baixa quanto possível.
- *Capacidade de detecção e correção de erros*. É desejável que erros sejam detectados e, de preferência, corrigidos. No caso bipolar, por exemplo, um erro isolado causará violação da regra bipolar e poderá ser facilmente detectado. Códigos corretores de erros serão discutidos em detalhe no Capítulo 14.
- *Densidade espectral de potência favorável*. É desejável ter densidade espectral de potência (PSD) zero em $f = 0$ (dc), pois acoplamento e transformadores ac são usados nos repetidores.* Potência significativa nas componentes de frequências baixas também deve ser evitada, pois causa ondulação dc (*dc wander*) na sequência de pulsos quando acoplamento ac é usado.
- *Conteúdo de temporização adequado*. Deve ser possível extrair a informação de temporização ou de relógio do sinal.
- *Transparência*. A transmissão correta de um sinal digital deve ser possível, independentemente do padrão de 1s e 0s. Vimos, nos casos de códigos on-off e bipolar, que uma longa sequência de 0s pode causar problema para a extração de temporização. Um código é transparente se os dados forem codificados de modo que, para qualquer sequência possível de bits, o sinal codificado possa ser detectado com fidelidade.

7.2.1 PSD de Diversos Códigos de Linha

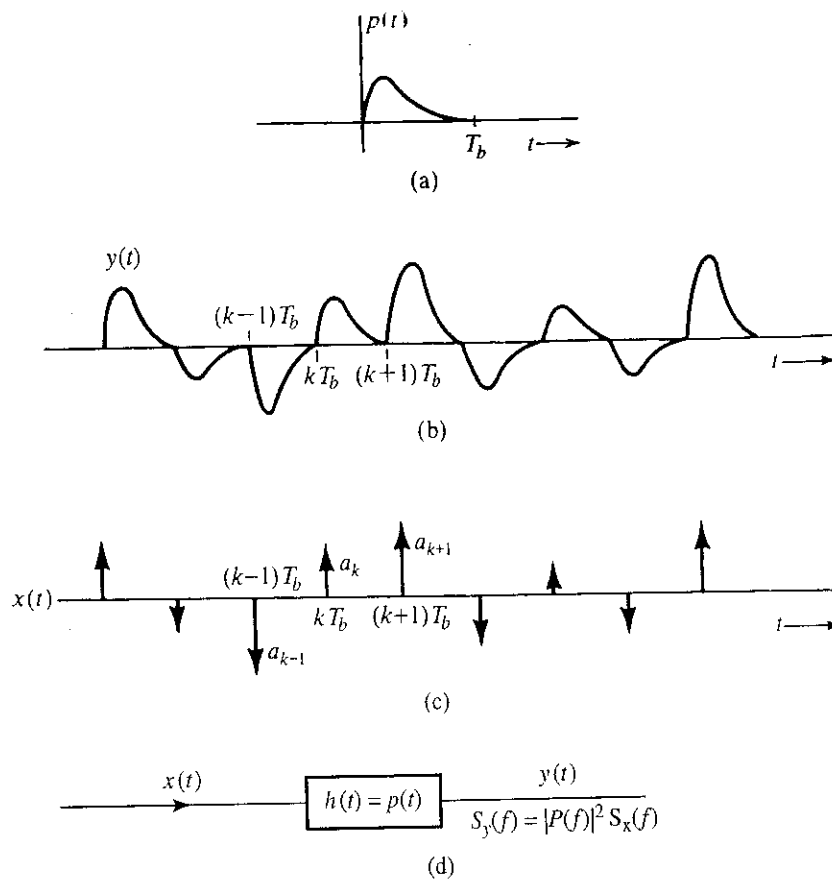
No Exemplo 3.19, discutimos o procedimento para determinar a PSD de um trem de pulsos polares. Usemos um procedimento similar para determinar uma expressão geral para PSD de sinais resultantes da modulação em banda base (codificação de linha), como ilustrado na Fig. 7.1. Em particular, aplicamos diretamente a relação entre PSD e a função de autocorrelação do sinal da modulação em banda base, dada na Seção 3.8 [Eq. (3.85)].

Na discussão a seguir, consideramos um pulso genérico $p(t)$, cuja transformada de Fourier é $P(f)$. Denotamos o símbolo do código de linha no instante de tempo k por a_k . Se a taxa de transmissão for $R_b = 1/T_b$ pulsos por segundo, o código de linha gera um trem de pulsos construído a partir do pulso

* Acoplamento ac é necessário porque as rotas dc providas por pares de cabo entre estações repetidoras são usadas para transmitir a potência necessária à operação dos repetidores.

Figura 7.4

Sinal modulado em amplitude por sequência aleatória de pulsos e sua geração por um trem de impulsos PAM.



básico $p(t)$, com amplitude a_k e iniciando no tempo $t = kT_b$; em outras palavras, o k -ésimo símbolo é transmitido como $a_k p(t - kT_b)$. A Fig. 7.4a fornece uma ilustração de um pulso especial $p(t)$, enquanto a Fig. 7.4b mostra o correspondente trem de pulsos gerado pelo codificador de linha na banda base. Como mostrado na Fig. 7.4b, contando uma sucessão de transmissões de símbolos espaçados por T_b segundos, o sinal em banda base é um trem de pulsos da forma

$$y(t) = \sum a_k p(t - kT_b) \quad (7.1)$$

Reparamos que o codificador de linha determina o símbolo $\{a_k\}$ como a amplitude do pulso $p(t - kT_b)$.

Os valores de a_k são aleatórios e dependem da entrada do codificador de linha e do próprio código de linha; $y(t)$ é um sinal modulado em amplitude de pulso (sinal PAM). Os códigos de linha on-off, polar e bipolar, são casos especiais desse trem de pulsos $y(t)$, em que a_k assume os valores 0, 1 ou -1 de modo aleatório, sujeito a algumas condições. Podemos, portanto, analisar muitos códigos de linha segundo a PSD de $y(t)$. Infelizmente, a PSD de $y(t)$ depende de a_k e de $p(t)$. Se a forma do pulso $p(t)$ for alterada, pode ser necessário refazer todo o cálculo da PSD. Essa dificuldade pode ser contornada com o simples artifício de selecionar um sinal PAM $x(t)$ que use um impulso unitário como o pulso básico $p(t)$ (Fig. 7.4c). Os impulsos ocorrem a intervalos T_b e a intensidade (área) do k -ésimo impulso é a_k . Se $x(t)$ for aplicado à entrada de um filtro cuja resposta ao impulso seja $h(t) = p(t)$ (Fig. 7.4d), a saída será o trem de pulsos $y(t)$ na Fig. 7.4b. Além disso, aplicação da Eq. (3.92) permite calcular a PSD de $y(t)$ como

$$S_y(f) = |P(f)|^2 S_x(f)$$

Essa relação nos permite determinar $S_y(f)$, a PSD de um código de linha que corresponde a uma forma de onda $p(t)$ qualquer, uma vez que conheçamos $S_x(f)$. A generalidade dessa abordagem a torna atrativa.

Agora, precisamos determinar $\mathcal{R}_x(\tau)$, a função de autocorrelação temporal do trem de impulsos $x(t)$. Isso pode ser feito com certa facilidade se considerarmos o impulso como uma forma limite de pulsos retangulares, como mostrado na Fig. 7.5a. Cada pulso tem uma largura $\epsilon \rightarrow 0$, e a altura do k -ésimo pulso é dada por

$$h_k = \frac{a_k}{\epsilon} \rightarrow \infty$$

Com isso, garantimos que a intensidade do k -ésimo impulso seja a_k , ou

$$\epsilon h_k = a_k$$

Se designarmos o correspondente trem de pulsos retangulares por $\hat{x}(t)$, então, por definição [Eq. (3.82) na Seção 3.8],

$$\mathcal{R}_{\hat{x}}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \hat{x}(t) \hat{x}(t - \tau) dt \quad (7.2)$$

Como $\mathcal{R}_{\hat{x}}(\tau)$ é uma função par de τ [Eq. (3.83)], basta que consideremos somente valores positivos de τ . Inicialmente, consideremos o caso $\tau < \epsilon$. Nesse caso, a integral na Eq. (7.2) é a área sob o sinal $\hat{x}(t)$ multiplicado por $\hat{x}(t)$ atrasado por τ ($\tau < \epsilon$). Como visto na Fig. 7.5b, a área associada ao k -ésimo pulso é $h_k^2(\epsilon - \tau)$, e

$$\begin{aligned} \mathcal{R}_{\hat{x}} &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_k h_k^2(\epsilon - \tau) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_k a_k^2 \left(\frac{\epsilon - \tau}{\epsilon^2} \right) \\ &= \frac{R_0}{\epsilon T_b} \left(1 - \frac{\tau}{\epsilon} \right) \end{aligned} \quad (7.3a)$$

em que

$$R_0 = \lim_{T \rightarrow \infty} \frac{T_b}{T} \sum_k a_k^2 \quad (7.3b)$$

Durante o intervalo de média T ($T \rightarrow \infty$), há N pulsos ($N \rightarrow \infty$), em que

$$N = \frac{T}{T_b} \quad (7.4)$$

Da Eq. (7.3b),

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k^2 \quad (7.5)$$

Observemos que a soma é feita com N pulsos. Portanto, R_0 é a média temporal do quadrado das amplitudes a_k do pulso. Usando a notação de média temporal, podemos expressar R_0 como

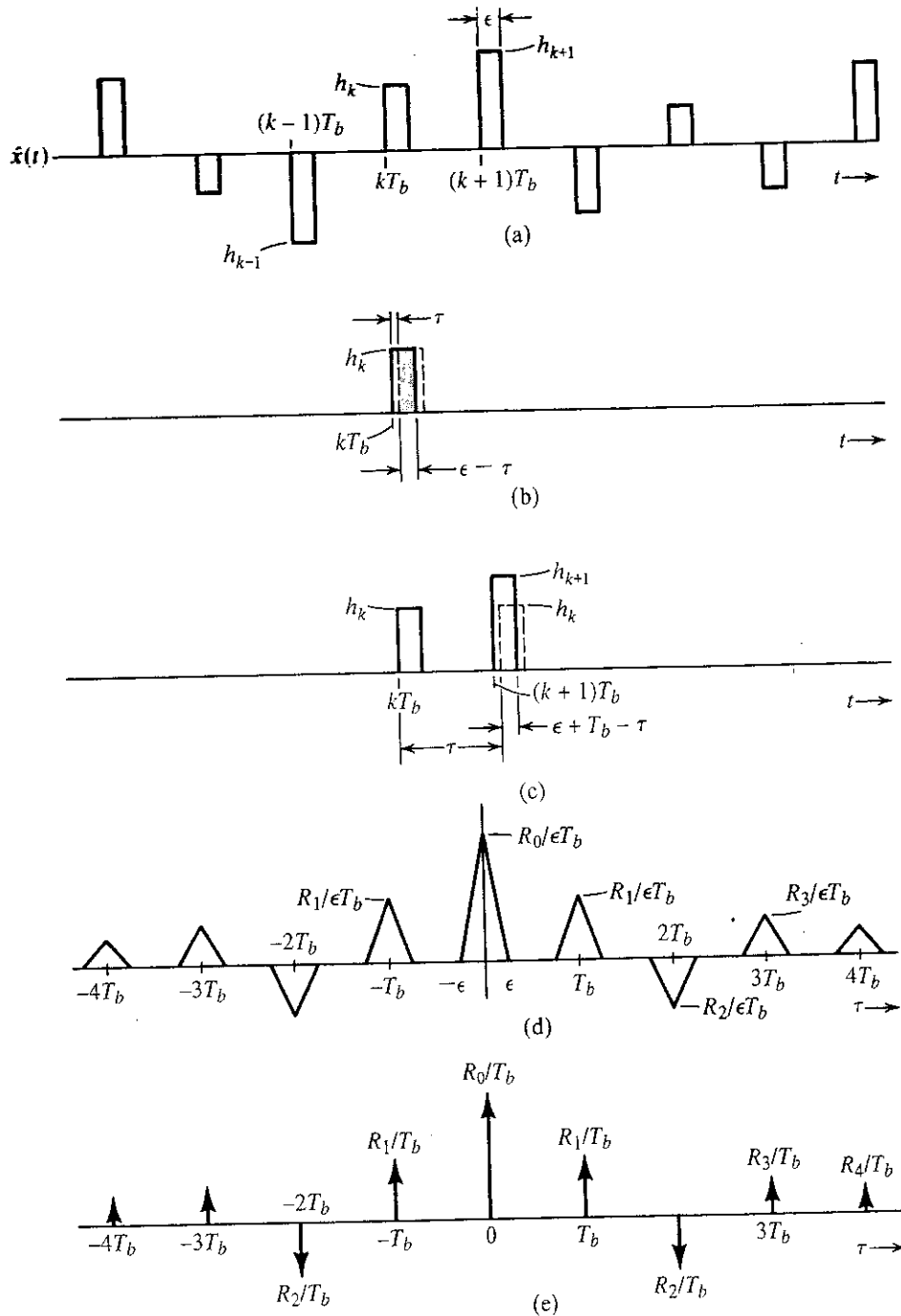
$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k^2 = \overline{a_k^2} \quad (7.6)$$

Também sabemos que $\mathcal{R}_{\hat{x}}(\tau)$ é uma função de τ [ver Eq. (3.83)]. Logo, a Eq. (7.3) pode ser expressa como

$$\mathcal{R}_{\hat{x}}(\tau) = \frac{R_0}{\epsilon T_b} \left(1 - \frac{|\tau|}{\epsilon} \right) \quad |\tau| < \epsilon \quad (7.7)$$

Figura 7.5

Dedução da PSD de um sinal PAM aleatório com largura de pulso muito estreita ϵ e altura $h_k = a_k/\epsilon$.



Este é um pulso retangular de altura $R_0/\epsilon T_b$, largura 2ϵ e centrado em $\tau = 0$ (Fig. 7.5d). Isso é esperado, pois, à medida que τ aumenta além de ϵ , não há sobreposição entre o sinal atrasado $\hat{x}(t - \tau)$ e $\hat{x}(t)$; logo, $R_{\hat{x}}(\tau) = 0$, como visto na Fig. 7.5d. Contudo, à medida que aumentamos τ ainda mais, observamos que o k -ésimo pulso de $\hat{x}(t - \tau)$ começará a se sobrepor ao $(k + 1)$ -ésimo pulso de $\hat{x}(t)$, quando τ se aproximar de T_b (Fig. 7.5c). Repetindo esse argumento, vemos que $R_{\hat{x}}(\tau)$ terá outro pulso retangular de largura 2ϵ centrado em $\tau = T_b$, com altura $R_1/\epsilon T_b$, em que

$$\begin{aligned}
 R_1 &= \lim_{T \rightarrow \infty} \frac{T_b}{T} \sum_k a_k a_{k+1} \\
 &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k a_{k+1} \\
 &= \overbrace{a_k a_{k+1}}
 \end{aligned}$$

Observemos que R_1 é obtido da multiplicação de cada intensidade de pulso (a_k) pela intensidade de seu vizinho imediato (a_{k+1}), da soma de todos esses produtos e, então, da divisão pelo número total de pulsos. Isso é claramente a média temporal do produto $a_k a_{k+1}$ e o resultado, em nossa notação, é $\overline{a_k a_{k+1}}$. Algo similar ocorre em torno de $\tau = 2T_b, 3T_b, \dots$. Assim, $\mathcal{R}_x(\tau)$ consiste em uma sequência de pulsos triangulares de largura 2ϵ , centrados em $\tau = 0, \pm T_b, \pm 2T_b, \dots$. A altura dos pulsos centrados em $\pm nT_b$ é R_n/T_b , em que

$$\begin{aligned} R_n &= \lim_{T \rightarrow \infty} \frac{T_b}{T} \sum_k a_k a_{k+n} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k a_{k+n} \\ &= \overline{a_k a_{k+n}} \end{aligned}$$

R_n é, essencialmente, a função de autocorrelação discreta dos símbolos do código de linha $\{a_k\}$.

Para calcular $\mathcal{R}_x(\tau)$, fazemos $\epsilon \rightarrow 0$ em $\mathcal{R}_x(\tau)$. À medida que $\epsilon \rightarrow 0$, a largura de cada pulso triangular $\rightarrow 0$ e a altura $\rightarrow \infty$, de modo que a área permanece finita. Assim, no limite $\epsilon \rightarrow 0$, os pulsos triangulares se tornam impulsos. Para o n -ésimo pulso centrado em nT_b , a altura é $R_n/\epsilon T_b$ e a área, R_n/T_b . Portanto, (Fig. 7.5e),

$$\mathcal{R}_x(\tau) = \frac{1}{T_b} \sum_{n=-\infty}^{\infty} R_n \delta(\tau - nT_b) \quad (7.8)$$

A PSD $S_x(f)$ é a transformada de Fourier de $\mathcal{R}_x(\tau)$. Logo,

$$S_x(f) = \frac{1}{T_b} \sum_{n=-\infty}^{\infty} R_n e^{-jn2\pi f T_b} \quad (7.9)$$

Lembrando que $R_{-n} = R_n$, [pois $\mathcal{R}(\tau)$ é uma função para de τ], temos

$$S_x(f) = \frac{1}{T_b} \left[R_0 + 2 \sum_{n=1}^{\infty} R_n \cos n2\pi f T_b \right] \quad (7.10)$$

A entrada $x(t)$ do filtro com resposta ao impulso $h(t) = p(t)$ resulta na saída $y(t)$, como mostrado na Fig. 7.4d. Se $p(t) \iff P(f)$, a função de transferência do filtro é $H(f) = P(f)$ e, segundo a Eq. (3.91),

$$S_y(f) = |P(f)|^2 S_x(f) \quad (7.11a)$$

$$= \frac{|P(f)|^2}{T_b} \left[\sum_{n=-\infty}^{\infty} R_n e^{-jn2\pi f T_b} \right] \quad (7.11b)$$

$$= \frac{|P(f)|^2}{T_b} \left[R_0 + 2 \sum_{n=1}^{\infty} R_n \cos n2\pi f T_b \right] \quad (7.11c)$$

Portanto, a PSD de um código de linha é totalmente caracterizada pelo correspondente R_n e pela seleção de forma de pulso de $P(f)$. A seguir, usaremos esse resultado geral para determinar as PSDs de vários códigos de linha específicos, a partir da determinação da autocorrelação entre símbolos.

7.2.2 Sinalização Polar

Na sinalização polar, 1 é transmitido por um pulso $p(t)$ e 0, por um pulso $-p(t)$. Nesse caso, a_k tem igual probabilidade de ser 1 ou -1, e a_k^2 será sempre 1. Assim,

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k^2$$

Há N pulsos, e $a_k^2 = 1$ para cada um, de modo que a soma no lado direito da equação anterior é N . Logo,

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} (N) = 1 \quad (7.12a)$$

Além disso, tanto a_k como a_{k+1} são 1 ou -1. Logo, $a_k a_{k+1}$ é 1 ou -1. Como a amplitude a_k tem, em média, igual probabilidade de ser 1 e -1, considerando N termos, o produto $a_k a_{k+1}$ é igual a 1 para $N/2$ termos e igual a -1 para os restantes $N/2$ termos. Portanto,

		Possíveis valores de $a_k a_{k+1}$	
		a_k	
a_{k+1}	-1	-1	+1
	+1	+1	-1

$$R_1 = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{2} (1) + \frac{N}{2} (-1) \right] = 0 \quad (7.12b)$$

Com esse raciocínio, podemos ver que o produto $a_k a_{k+n}$ também tem igual probabilidade de ser 1 ou -1. Assim,

$$R_n = 0 \quad n \geq 1 \quad (7.12c)$$

Portanto, da Eq. (7.11c),

$$\begin{aligned} S_y(f) &= \frac{|P(f)|^2}{T_b} R_0 \\ &= \frac{|P(f)|^2}{T_b} \end{aligned} \quad (7.13)$$

Para efeitos de comparação de vários esquemas, consideremos uma forma de pulso *específica*. Seja $p(t)$ um pulso retangular de largura $T_b/2$ (pulso retangular de meia largura), ou seja,

$$p(t) = \Pi\left(\frac{t}{T_b/2}\right) = \Pi\left(\frac{2t}{T_b}\right)$$

e

$$P(f) = \frac{T_b}{2} \operatorname{sinc}\left(\frac{\pi f T_b}{2}\right) \quad (7.14)$$

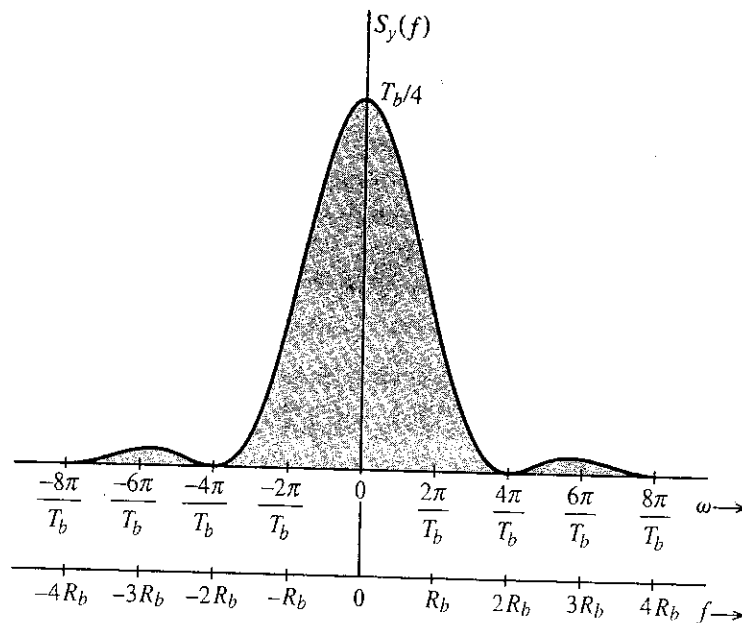
Portanto,

$$S_y(f) = \frac{T_b}{4} \operatorname{sinc}^2\left(\frac{\pi f T_b}{2}\right) \quad (7.15)$$

A Fig. 7.6 mostra o espectro $S_y(f)$. Fica claro que a maior parte da potência do sinal polar se concentra nas frequências mais baixas. Teoricamente, o espectro se torna muito estreito à medida que a frequência aumenta, mas sua largura jamais iguala zero acima de uma dada frequência. Para termos uma medida representativa de largura de banda, consideramos a *primeira frequência de nulo não dc* como a **largura de banda essencial**.*

*A rigor, a primeira frequência de nulo acima de dc nem sempre é uma boa medida da largura de banda de um sinal. A quantidade de sinal contida no lóbulo principal (primeiro) da PSD é que determina se o primeiro nulo não dc será uma medida significativa de largura de banda, como veremos mais adiante, ao compararmos a PSD de vários códigos de linha (Fig. 7.9). Na maioria dos casos práticos, essa aproximação é aceitável para os códigos de linha e formas de pulsos usados comumente.

Figura 7.6
Densidade
espectral de
potência de um
sinal polar.



Do espectro do sinal polar, a largura de banda essencial do sinal é calculada como $2R_b$ Hz (em que R_b é a frequência do relógio). Isso é 4 vezes a largura de banda teórica (largura de banda de Nyquist) exigida para transmitir R_b pulsos por segundo. O aumento da largura do pulso reduz a largura de banda (expansão no domínio do tempo resulta em compressão no domínio da frequência). Para um pulso de largura completa* (máxima largura de pulso possível), a largura de banda essencial é a metade, ou seja, R_b Hz. Esta, no entanto, ainda é o dobro da largura de banda teórica. Assim, sinalização polar não é a mais eficiente do ponto de vista de largura de banda.

Uma segunda deficiência da sinalização polar é não ter capacidade de detecção ou correção de erro. Outra desvantagem da sinalização polar é ter PSD não nula em dc ($f=0$). Isso impede o uso de acoplamento ac durante a transmissão. O modo de acoplamento ac é muito importante na prática, pois permite que transformadores e capacitores de bloqueio auxiliem no casamento de impedância e na remoção de polarização, assim como a alimentação de potência dos repetidores em linha por meio dos pares de cabo. Mais adiante, mostraremos como a PSD de um código de linha pode ser forçada a zero em dc por meio de escolha adequada da forma de $p(t)$.

Quanto aos aspectos positivos, a sinalização polar é o esquema mais eficiente do ponto de vista de consumo de potência. Para uma dada potência, podemos mostrar que a probabilidade de detecção de erro em um esquema polar é a menor dentre todas as técnicas de sinalização (Capítulo 10). A sinalização polar também é transparente, pois sempre há algum pulso (positivo ou negativo), qualquer que seja a sequência de bits. Não há uma componente discreta do relógio no espectro do sinal polar. Contudo, a retificação do sinal polar RZ produz um sinal periódico na frequência do relógio e pode ser usado prontamente para extrair a temporização.

7.2.3 Construção de um Nulo DC na PSD via Formatação de Pulso

Como $S_y(f)$, a PSD de um código de linha, contém um fator $|P(f)|^2$, podemos forçar a PSD a ter um nulo dc se selecionarmos um pulso $p(t)$ tal que $P(f)$ seja zero em dc ($f=0$). Como

$$P(f) = \int_{-\infty}^{\infty} p(t)e^{-j2\pi ft} dt$$

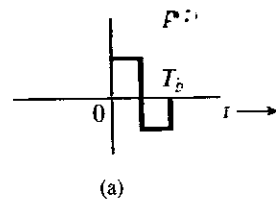
*O esquema que usa pulso de largura completa $p(t) = \Pi(t/T_b)$ é um exemplo de um esquema sem retorno ao zero (NRZ). O esquema de meia largura de pulso, por sua vez, é um exemplo de um esquema com retorno ao zero (RZ).

Figura 7.7

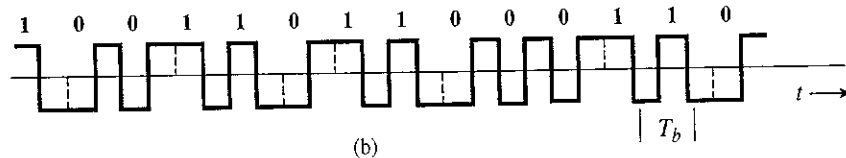
Sinal bifásico
(Manchester
ou de fase
alternada).

(a) Pulso básico
para sinalização
Manchester.

(b) Forma de
onda transmitida
para sequência
binária de
dados usando
sinalização
Manchester.



(a)



(b)

temos

$$P(0) = \int_{-\infty}^{\infty} p(t) dt$$

Portanto, se a área sob $p(t)$ for feita igual a zero, $P(0)$ será zero, e teremos um nulo dc na PSD. Para um pulso retangular, uma possível forma de $p(t)$ para obter este efeito é mostrada na Fig. 7.7a. Quando usamos esse pulso com codificação de linha polar, o sinal resultante é conhecido como sinal em **código Manchester, bifásico, ou de fase alternada**. O leitor pode usar a Eq. (7.13) para mostrar que, para esse pulso, a PSD do código de linha Manchester tem um nulo dc (Exercício 7.2-2).

7.2.4 Sinalização On-Off

Na sinalização on-off, um 1 é transmitido por um pulso $p(t)$ e um 0, pela ausência de pulso. Assim, a amplitude de pulso a_k tem igual probabilidade de ser 1 ou 0. Dentre N pulsos no intervalo de T segundos, a_k é 1 para $N/2$ pulsos e 0 para os restantes $N/2$ pulsos, em média. Portanto,

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{2}(1)^2 + \frac{N}{2}(0)^2 \right] = \frac{1}{2} \quad (7.16)$$

Para calcular o valor de R_n , precisamos considerar o produto $a_k a_{k+n}$. Como a_k e a_{k+n} têm igual probabilidade de serem 1 ou 0, o produto $a_k a_{k+n}$ tem igual probabilidade de ser 1×1 , 1×0 , 0×1 ou 0×0 , ou seja, 1, 0, 0, 0. Portanto, em média, o produto $a_k a_{k+n}$ é igual a 1 para $N/4$ termos e 0 para $3N/4$ termos, e

$$R_n = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{4}(1) + \frac{3N}{4}(0) \right] = \frac{1}{4} \quad n \geq 1 \quad (7.17)$$

Portanto [Eq. (7.9)],

$$S_x(f) = \frac{1}{2T_b} + \frac{1}{4T_b} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} e^{-jn2\pi f T_b} \quad (7.18a)$$

$$= \frac{1}{4T_b} + \frac{1}{4T_b} \sum_{n=-\infty}^{\infty} e^{-jn2\pi f T_b} \quad (7.18b)$$

A Eq. (7.18b) é obtida da Eq. (7.18a) dividindo em dois o termo $1/2T_b$ correspondente a R_0 : $1/4T_b$ fora do somatório e $1/4T_b$ dentro do somatório (correspondendo a $n = 0$). Agora, usamos a fórmula (ver a nota de rodapé para a prova*)

$$\sum_{n=-\infty}^{\infty} e^{-jn2\pi f T_b} = \frac{1}{T_b} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_b}\right)$$

A substituição deste resultado na Eq. (7.18b) leva a

$$S_x(f) = \frac{1}{4T_b} + \frac{1}{4T_b^2} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_b}\right) \quad (7.19a)$$

e à desejada PSD da forma de onda on-off $y(t)$ [da Eq. (7.11a)]

$$S_y(f) = \frac{|P(f)|^2}{4T_b} \left[1 + \frac{1}{T_b} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_b}\right) \right] \quad (7.19b)$$

Reparemos que, ao contrário do espectro contínuo da PSD da sinalização polar, a PSD do sinal on-off na Eq. (7.19b) tem uma parte discreta adicional, que pode ser anulada se a forma do pulso for escolhida de modo que

$$P\left(\frac{n}{T_b}\right) = 0 \quad n = 0, \pm 1, \dots$$

Tomando como exemplo o caso de um pulso retangular de meia largura [Eq. (7.14)],

$$S_y(f) = \frac{T_b}{16} \text{sinc}^2\left(\frac{\pi f T_b}{2}\right) \left[1 + \frac{1}{T_b} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_b}\right) \right] \quad (7.20)$$

A resultante PSD é mostrada na Fig. 7.8. A componente contínua do espectro é $(T_b/16) \text{sinc}^2(\pi f T_b/2)$. Isso é igual (exceto por um fator de escala) ao espectro do sinal polar [Eq. (7.15)]. A componente discreta é representada pelo produto de um trem de impulsos pela componente contínua $(T_b/16) \text{sinc}^2(\pi f T_b/2)$. Assim, essa componente aparece como impulsos periódicos, tendo a componente contínua como envelope. Além disso, a frequência de repetição dos impulsos é a frequência do relógio $R_b = 1/T_b$, pois sua frequência fundamental é $2\pi/T_b$ rad/s ou $1/T_b$ Hz. Esse é um resultado lógico, pois, como mostra a Fig. 7.3, um sinal on-off pode ser expresso como a soma de uma componente polar com uma componente periódica. A componente polar $y_1(t)$ é exatamente a metade do sinal polar discutido anteriormente. Logo, a PSD dessa componente é um quarto da PSD na Eq. (7.15). A componente periódica tem a frequência do relógio R_b , e consiste em componentes discretas de frequência R_b e seus harmônicos.

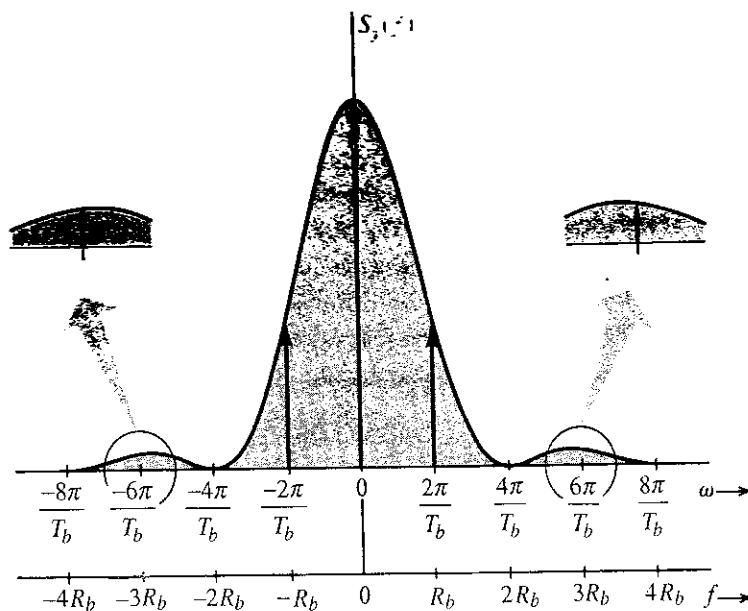
*O trem de impulsos na Fig. 3.23a do Exemplo 3.11 é $\delta_{T_b}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_b)$. Além disso, a série de Fourier para esse trem de impulsos é obtida da Eq. (2.67) como

$$\sum_{n=-\infty}^{\infty} \delta(t - nT_b) = \frac{1}{T_b} \sum_{n=-\infty}^{\infty} e^{jn2\pi R_b t} \quad R_b = \frac{1}{T_b}$$

Tomamos a transformada de Fourier dos dois lados dessa equação, e usamos o fato de que $\delta(t - nT_b) \iff e^{-jn2\pi f T_b}$ e $e^{jn2\pi R_b t} \iff \delta(f - nR_b)$. Isso leva a

$$\sum_{n=-\infty}^{\infty} e^{-jn2\pi f T_b} = \frac{1}{T_b} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_b}\right)$$

Figura 7.8
Densidade
espectral de
potência (PSD)
de um sinal
on-off.



A sinalização on-off não tem muito do que se gabar. Para uma dada potência transmitida, tem menos imunidade ao ruído do que o esquema polar, que usa um pulso positivo para 1 e um pulso negativo para 0. Isso decorre do fato de que a imunidade ao ruído depende da diferença entre as amplitudes que representam 1 e 0. Portanto, para a mesma imunidade, se a sinalização on-off usar pulsos de amplitude 2 e 0, a sinalização polar precisa apenas de pulsos com amplitudes 1 e -1. É simples mostrar que a sinalização on-off requer o dobro da potência da sinalização polar. Se um pulso de amplitude 1 ou -1 tiver energia E , o pulso de amplitude 2 tem energia $(2)^2 E = 4E$. Como $1/T_b$ dígitos são transmitidos por segundo, a potência do sinal polar é $(E)(1/T_b) = E/T_b$. Para o caso on-off, a energia de cada pulso é $4E$, embora, em média, um pulso seja transmitido na metade do tempo, enquanto nada é transmitido na outra metade. Assim, a potência média de sinal on-off é

$$\frac{1}{T_b} \left(4E \frac{1}{2} + 0 \cdot \frac{1}{2} \right) = \frac{2E}{T_b}$$

que é o dobro da potência requerida pela sinalização polar. Além disso, ao contrário do caso polar, a sinalização on-off não é transparente. Uma longa sequência de 0s (ou offs) causa ausência de sinal e pode levar a erros na extração da temporização. Adicionalmente, todas as desvantagens da sinalização polar (por exemplo, excessiva largura de banda de transmissão, espectro de potência não nulo em dc, incapacidade de detecção ou correção de erros) também estão presentes na sinalização on-off.

7.2.5 Sinalização Bipolar

O esquema de sinalização usado em PCM para redes de telefonia é denominado bipolar (pseudoternário ou inversão alternada de sinal). Um 0 é transmitido pela ausência de pulso e um 1, por um pulso $p(t)$ ou $-p(t)$, dependendo se o 1 anterior foi transmitido por $-p(t)$ ou $p(t)$. Com pulsos consecutivos alternados, podemos evitar ondulação dc e, assim, causar um nulo na PSD. A sinalização bipolar, na verdade, usa três símbolos $[p(t), 0 \text{ e } -p(t)]$ e, portanto, é uma sinalização ternária, e não binária.

Para calcular a PSD, temos

$$R_o = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k^2$$

Em média, metade dos a_k s são 0 e a outra metade é 1 ou -1, com $a_k^2 = 1$. Logo,

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{2} (\pm 1)^2 + \frac{N}{2} (0)^2 \right] = \frac{1}{2}$$

Para calcular R_1 , consideremos o produto das amplitudes de pulsos $a_k a_{k+1}$. Quatro sequências de dois bits têm igual probabilidade de ocorrência: **11**, **10**, **01**, **00**. Como o bit **0** é codificado por ausência de pulso ($a_k = 0$), o produto $a_k a_{k+1}$ é zero para as três últimas sequências. Isso significa, em média, que $3N/4$ combinações têm $a_k a_{k+1} = 0$ e apenas $N/4$ combinações têm $a_k a_{k+1}$ não nulo. Devido à regra bipolar, a sequência de bits **11** pode ser codificada somente por pulsos de polaridades opostas. Isso significa que o produto $a_k a_{k+1} = -1$ para $N/4$ combinações. Portanto,

$$R_1 = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{4} (-1) + \frac{3N}{4} (0) \right] = -\frac{1}{4}$$

Para calcular R_2 de modo similar, devemos analisar o produto $a_k a_{k+2}$. Para isso, precisamos considerar todas as possíveis combinações de três bits em sequência. Há oito combinações com igual probabilidade de ocorrência: **111**, **101**, **110**, **100**, **011**, **010**, **001**, **000**. As últimas seis combinações têm **0** como o primeiro ou o último bit; logo, $a_k a_{k+2} = 0$ para todas essas seis combinações. As duas primeiras são as únicas combinações que têm $a_k a_{k+2}$ não nulo. Dada a regra bipolar, o primeiro e o último bits na combinação **111** têm a mesma polaridade, resultando em $a_k a_{k+2} = 1$. Contudo, em **101**, o primeiro e o último bits têm polaridades opostas, resultando em $a_k a_{k+2} = -1$. Assim, em média, $a_k a_{k+2} = 1$ para $N/8$ termos, -1 para $N/8$ termos e 0 para $3N/4$ termos. Logo,

$$R_2 = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{8} (1) + \frac{N}{8} (-1) + \frac{3N}{8} (0) \right] = 0$$

Em geral,

$$R_n = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k a_{k+n}$$

Para $n > 2$, o produto $a_k a_{k+n}$ pode ser 1, -1 ou 0. Além disso, um número igual de combinações tem valores 1 e -1. Isso causa $R_n = 0$. Portanto,

$$R_n = 0 \quad n > 1$$

e [ver Eq. (7.11c)]

$$S_y(f) = \frac{|P(f)|^2}{2T_b} [1 - \cos 2\pi f T_b] \quad (7.21a)$$

$$= \frac{|P(f)|^2}{T_b} \text{sen}^2 (\pi f T_b) \quad (7.21b)$$

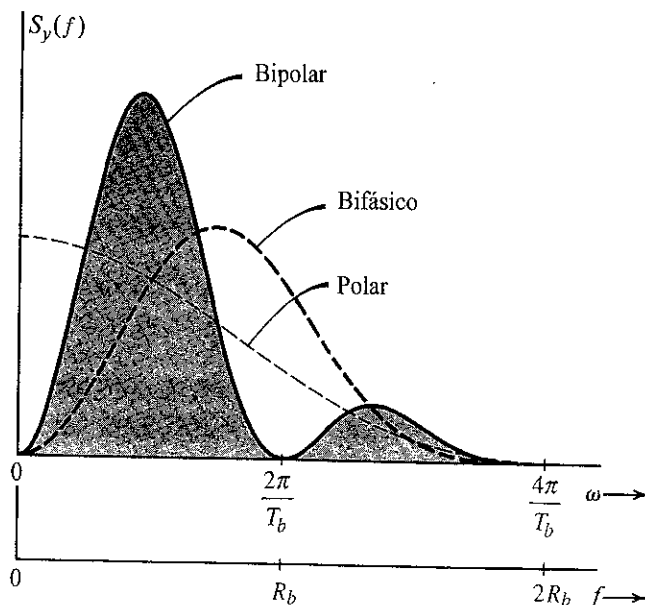
Reparemos que $S_y(f) = 0$ para $f = 0$ (dc), qualquer que seja $P(f)$. Logo, a PSD tem um nulo dc, o que é desejável para acoplamento ac. Além disso, $\text{sen}^2 (\pi f T_b) = 0$ em $f = 1/T_b$, ou seja, em $f = 1/T_b = R_b$ Hz. Portanto, independentemente de $P(f)$, fica garantida a largura de banda R_b Hz para o primeiro nulo não dc. Para o caso de pulsos de meia largura,

$$S_y(f) = \frac{T_b}{4} \text{sinc}^2 \left(\frac{\pi f T_b}{2} \right) \text{sen}^2 (\pi f T_b) \quad (7.22)$$

Isso é ilustrado na Fig. 7.9. A largura de banda essencial do sinal é R_b ($R_b = 1/T_b$), que é a metade da largura de banda do esquema polar que usa o mesmo pulso de meia largura ou da sinalização on-off e o dobro da mínima largura de banda teórica. Observamos que podemos calcular a largura de banda R_b para o caso polar (ou on-off) com pulso de largura completa. Para o caso bipolar, a largura de banda é R_b Hz, tanto para pulso de meia largura como de largura completa.

Figura 7.9

PSD de sinais bipolar, polar e bifásico normalizados para iguais potências. Pulsos retangulares de meia largura são usados.



A sinalização bipolar tem diversas vantagens: (1) seu espectro tem um nulo dc; (2) sua largura de banda não é excessiva, (3) tem capacidade de detecção de erro isolado. Isso se deve ao fato de que mesmo um erro isolado viola a regra de pulsos alternados e será detectado imediatamente. Se um sinal bipolar for retificado, obtemos um sinal on-off que tem uma componente discreta na frequência do relógio. Entre as desvantagens de um sinal bipolar está a exigência do dobro da potência (3 dB) usada por um sinal polar. Isso decorre do fato de a detecção ser essencialmente equivalente à da sinalização on-off, sob o ponto de vista de detecção: a distinção é entre $+p(t)$ ou $-p(t)$ e 0, e não entre $\pm p(t)$.

Outra desvantagem da sinalização bipolar é não ser transparente. Na prática, vários esquemas substitutos foram usados para evitar que longas de zeros lógicos permitissem que sinais de relógio se desviassem. A seguir, discutiremos dois desses esquemas.

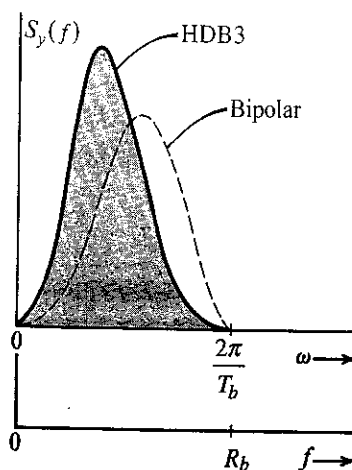
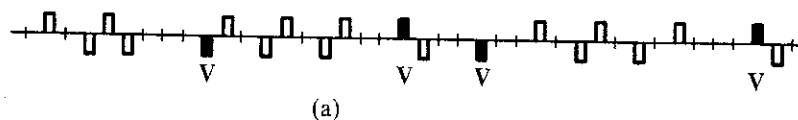
Figura 7.10

(a) Sinal HDB3
e (b) sua PSD.

Dígitos de entrada 0 1 0 1 1 1 0 0 0 0 1 0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 1 0 1 1 0 1 0 1 0 0 0 0 1

Dígitos codificados 0 1 0 1 1 1 0 0 0 V 1 0 1 1 0 1 1 0 0 V 1 0 0 V 0 0 1 0 1 1 0 1 0 1 0 0 0 V 1

Forma de onda transmitida



(b)

Sinalização Bipolar de Alta Densidade (HDB)

O esquema HDB (*high-density bipolar*) é um padrão ITU-T (antigo CCITT – Comitê Consultivo Internacional de Telefonia e Telegrafia). Nesse esquema, o problema de falta de transparência da sinalização bipolar é eliminado com a adição de pulsos quando o número de pulsos consecutivos 0s ultrapassa N . Esse código modificado é denominado **codificação bipolar de alta densidade (HDBN)**, em que N pode assumir qualquer valor 1, 2, 3, O código HDB mais importante é o formato HDB3, adotado como padrão internacional.

A ideia básica do código HDBN é que, na ocorrência de uma sequência de $N + 1$ zeros, esse grupo de zeros é substituído por uma das sequências especiais de $N + 1$ dígitos binários. Para aumentar o conteúdo de temporização do sinal, as sequências são escolhidas de modo a incluir alguns 1s binários. Os 1s incluídos deliberadamente violam a regra bipolar para facilitar a identificação da sequência subsequente. Na codificação HDB3, por exemplo, as sequências especiais usadas são **000V** e **B00V**, em que **B** = 1, que respeita a regra bipolar, e **V** = 1, que viola a regra bipolar. A escolha da sequência **000V** ou **B00V** é feita de forma que os pulsos **V** consecutivos têm sinais alternados para evitar ondulação dc e manter o nulo dc na PSD. Isso requer que a sequência **B00V** seja usada quando há um número par de 1s após a última sequência especial, e a sequência **000V** é usada quando há um número ímpar de 1s após a última sequência. A Fig. 7.10a mostra um exemplo desse código. Reparemos que, na sequência **B00V**, **B** e **V** são codificados pelo mesmo pulso. O decodificador deve verificar duas coisas: as violações bipolares e o número de 0s que antecede cada violação, para determinar se o 1 anterior também era uma substituição.

Apesar das deliberadas violações da regra bipolar, a sinalização HDB mantém a capacidade de detecção de erro. Qualquer erro isolado inserirá uma violação bipolar espúria (ou removerá uma das violações deliberadas). Isso se tornará aparente quando, na próxima violação, a alternância de violações não desaparecer. Isso também mostra que violações deliberadas podem ser detectadas, apesar de erros isolados. A Fig. 7.10b mostra a PSD de HDB3 juntamente com a de um sinal bipolar, para facilitar a comparação.³

Sinalização Binária com Substituição de N Zeros (BNZS)

Uma classe de códigos de linha similar a HDBN é a **binária com substituição de N zeros**, ou **BNZS** (*binary with N zero substitution*). Nesse código, se ocorrerem N zeros em sequência, estes são substituídos por uma das duas sequências especiais que contêm alguns 1s para aumentar o conteúdo de temporização. Há violações deliberadas da regra bipolar, assim como em HDBN. A codificação binária com substituição de oito zeros (B8ZS) é usada em sinais DS1 da hierarquia de telefonia digital, discutidos no Capítulo 6. Essa codificação substitui quaisquer sequências de oito zeros por uma sequência de uns e zeros que contêm duas violações bipolares. É improvável que uma sequência desse tipo seja afetada por erros, e qualquer uma que seja recebida é substituída por uma sequência de oito zeros antes da decodificação. A sequência empregada como substituição consiste nos padrões **000VB0VB**. Do mesmo modo, no código **B6ZS** usado em sinais DS2, uma sequência de seis zeros é substituída por **0VB0VB**; um código **B3ZS** é aplicado a sinais DS3. O código **B3ZS** é um pouco mais complexo que os outros, pois usa **B0V** ou **00V**, sendo a escolha feita de modo que o número de pulsos **B** entre pulsos **V** consecutivos seja ímpar. Esses códigos BNZS, com $N = 3, 6$ ou 8 , envolvem violações bipolares e, portanto, devem ser cuidadosamente substituídos pelas correspondentes sequências de zero no receptor.

Há muitos outros códigos (de linha) de transmissão, em número demasiadamente grande para serem listados aqui. Uma lista de códigos e referências apropriadas pode ser encontrada em Bylanski e Ingram.³

7.3 FORMATAÇÃO DE PULSO

A PSD $S_y(f)$ de um sinal digital $y(t)$ pode ser controlada por uma escolha do código de linha ou por $P(f)$, a forma do pulso. Na última seção, discutimos como a PSD é controlada por um código de linha.

Na presente seção, examinaremos como $S_y(f)$ é influenciada pela forma do pulso $p(t)$, e aprenderemos como formatar um pulso $p(t)$ para obter uma $S_y(f)$ desejada. A PSD $S_y(f)$ é forte e diretamente influenciada pela forma do pulso $p(t)$, pois $S_y(f)$ contém o termo $|P(f)|^2$. Assim, em comparação com a natureza do código de linha, a forma do pulso é um fator mais direto e poderoso para formatar a PSD $S_y(f)$.

7.3.1 Interferências Intersimbólicas (ISI) e Efeitos

Na última seção, por conveniência, usamos um simples pulso retangular de meia largura $p(t)$. A rigor, nesse caso, a largura de banda $S_y(f)$ é infinita, pois $P(f)$ tem largura de banda infinita. Contudo, vimos que a largura de banda essencial de $S_y(f)$ era finita. Por exemplo, a maior parte da potência de um sinal bipolar está contida na largura de banda essencial entre 0 e R_b Hz. Notemos, no entanto, que a PSD é pequena, mas não nula, no intervalo $f > R_b$ Hz. Portanto, quando um sinal desse tipo é transmitido em um canal de largura de banda R_b Hz, uma porção significativa de seu espectro é transmitida, enquanto uma parcela pequena é suprimida. Nas Seções 3.5 e 3.6, vimos que uma distorção espectral como essa tendia a espalhar o pulso (dispersão). Espalhamento de um pulso além da janela de tempo T_b a ele alocada causará interferência com pulsos vizinhos. Isso é conhecido como **interferência intersimbólica** ou **ISI** (*intersymbol interference*).

ISI não é ruído. ISI é causada por canais não ideais que não são livres de distorção em toda a largura de banda do sinal. No caso de um pulso retangular de meia largura, a largura de banda do sinal é, estritamente falando, infinita. A ISI, como uma manifestação de distorção do canal, pode causar erros na detecção do pulso, caso seja suficientemente grande.

Para resolver a dificuldade associada à ISI, redefinamos nosso problema. Precisamos transmitir um pulso a cada T_b segundos, sendo $a_k p(t - kT_b)$ o k -ésimo pulso. O canal tem largura de banda finita, e devemos detectar a amplitude de pulso a_k corretamente (ou seja, sem ISI). Em nossa discussão até aqui, consideramos pulsos limitados no tempo. Como tais pulsos não podem ser limitados em frequência, parte de seus espectros é suprimida por um canal de largura de banda finita. Isso causa distorção dos pulsos (que se espalham) e, em consequência, a ISI. Podemos, inicialmente, tentar resolver essa dificuldade usando pulsos que tenham largura de banda finita, de modo que possam ser transmitidos intactos pelo canal de banda finita. No entanto, pulsos limitados em frequência não podem ser limitados no tempo. Obviamente, vários pulsos se sobreporão e causarão ISI. Portanto, quer usemos pulsos limitados no tempo ou pulsos limitados em frequência, parece que a ISI não pode ser evitada: é inerente à transmissão em largura de banda finita. Por sorte, existe uma escapatória desse beco sem saída. Amplitudes de pulso podem ser detectadas corretamente apesar de espalhamento (ou sobreposição) dos pulsos, desde que não haja ISI nos instantes de tomada de decisão. Isso pode ser alcançado por meio de pulsos limitados em frequência com forma apropriada. Para eliminar ISI, Nyquist propôs três diferentes critérios para formatação de pulsos,⁴ sendo permitida sobreposição de pulsos. Todavia, os pulsos são formatados de modo a causar interferência zero (ou controlada) em todos os outros pulsos nos instantes de tomada de decisão. Assim, ao limitar a exigência de não interferência somente nos instantes de tomada de decisão, eliminamos a necessidade de não sobreposição total dos pulsos. Consideraremos apenas os dois primeiros critérios. O terceiro é menos útil,⁵ e, por conseguinte, não será analisado aqui.

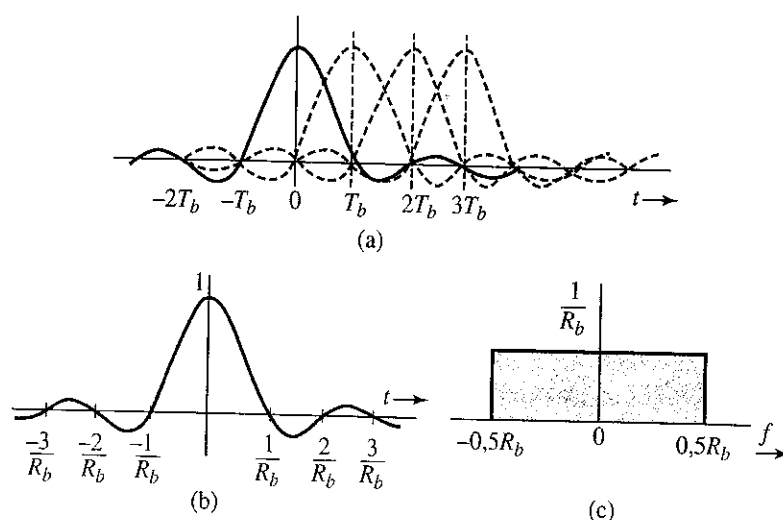
7.3.2 Primeiro Critério de Nyquist para ISI Nula

No primeiro método, Nyquist obtém ISI zero ao escolher uma forma de pulso que tenha amplitude não nula no centro (digamos, em $t = 0$) e amplitudes nulas em $t = \pm nT_b$ ($n = 1, 2, 3, \dots$), em que T_b é a separação entre pulsos transmitidos sucessivos (Fig. 7.11a). Assim,

$$p(t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm nT_b \end{cases} \quad \left(T_b = \frac{1}{R_b} \right) \quad (7.23)$$

Um pulso que satisfaça esse critério causa ISI zero nos centros de todos os outros pulsos, ou instantes de sinalização, como mostrado na Fig. 7.11a, na qual ilustramos vários pulsos sucessivos (linha tra-

Figura 7.11
Pulso com
mínima largura
de banda que
satisfaz o
primeiro critério
de Nyquist e
seu espectro.



cejada), centrados em $t = 0, T_b, 2T_b, 3T_b, \dots$ ($T_b = 1/R_b$). Por conveniência, mostramos todos os pulsos como positivos.* Fica claro que, desta figura, as amostras em $t = 0, T_b, 2T_b, 3T_b, \dots$ consistem na amplitude de apenas um pulso (centrado no instante de amostragem), sem interferência dos outros pulsos.

A transmissão de R_b bit/s requer uma mínima largura de banda teórica de $R_b/2$ Hz. Seria muito bom se a mínima largura de banda de um pulso que satisfizesse o critério de Nyquist fosse $R_b/2$ Hz. Será possível determinar um pulso $p(t)$ desse tipo? Já resolvemos esse problema (Exemplo 6.1 com $B = R_b/2$), e mostramos que existe um (e somente um) pulso que atende o critério de Nyquist (7.23) e tem largura de banda de $R_b/2$ Hz. Esse pulso $p(t) = \text{sinc}(\pi R_b t)$ (Fig. 7.11b) tem a propriedade

$$\text{sinc}(\pi R_b t) = \begin{cases} 1 & t = 0 \\ 0 & t = \pm nT_b \end{cases} \quad \left(T_b = \frac{1}{R_b}\right) \quad (7.24a)$$

Além disso, a transformada de Fourier desse pulso é

$$P(f) = \frac{1}{R_b} \Pi\left(\frac{f}{R_b}\right) \quad (7.24b)$$

que tem largura de banda $R_b/2$ Hz, como visto na Fig. 7.11c. Podemos usar esse pulso para transmitir uma taxa de R_b pulsos por segundo sem ISI, em uma largura de banda de apenas $R_b/2$.

Esse esquema mostra que podemos alcançar o limite teórico de desempenho se usarmos um pulso sinc. Lamentavelmente, esse pulso é impraticável, pois tem início em $-\infty$. Teremos de esperar um tempo infinito para gerá-lo. Qualquer tentativa de truncá-lo aumentaria sua largura de banda além dos $R_b/2$ Hz. Contudo, mesmo se esse pulso fosse realizável, teria uma característica indesejável: decairia a uma taxa demasiadamente lenta de $1/t$. Isto causaria alguns sérios problemas práticos. Por exemplo, se a taxa de dados nominal R_b bit/s requerida para esse esquema variasse um pouco, a amplitude de um pulso não se anularia nos centros dos outros pulsos. Como os pulsos decaem com $1/t$ apenas, a interferência cumulativa no centro de um dado pulso devido a todos os outros seria da forma $\sum(1/n)$. É um fato bem conhecido que essa série infinita não converge e pode produzir um valor muito alto. Um resultado similar ocorreria se tudo fosse perfeito no transmissor, mas a taxa de amostragem no receptor se desvia da taxa de R_b Hz. A mesma coisa aconteceria se os instantes de amostragem se desviassem um pouco devido à incerteza temporal nos pulsos, o que é inevitável mesmo nos sistemas mais sofisticados. Portanto, esse esquema falhará, a menos que tudo seja perfeito, o que é uma impossibi-

*Na verdade, um pulso correspondendo a 0 seria negativo. Contudo, considerar todos os pulsos como positivos não afeta o raciocínio. Mostrar pulsos negativos tornaria a figura desnecessariamente confusa.

lidade prática. E tudo isso decorre do fato de $\text{sinc}(\pi R_b t)$ decair de modo demasiadamente lento (com $1/t$). A solução consiste em encontrar um pulso $p(t)$ que satisfaça a Eq. (7.23), mas decaia mais rapidamente que $1/t$. Nyquist mostrou que um pulso como esse requer uma largura de banda $kR_b/2$, com $1 \leq k \leq 2$.

Isso pode ser provado da seguinte forma: consideremos $p(t) \iff P(f)$, em que a largura de banda de $P(f)$ está no intervalo $(R_b/2, R_b)$ (Fig. 7.12a). O desejado pulso $p(t)$ satisfaz a Eq. (7.23). Se amostrarmos $p(t)$ a cada T_b segundos, multiplicando-o por $\delta_{T_b}(t)$ (um trem de impulsos), devido à propriedade (7.23), todas as amostras, exceto a feita na origem, serão iguais a zero. Assim, o sinal amostrado $\bar{p}(t)$ será

$$\bar{p}(t) = p(t)\delta_{T_b}(t) = \delta(t) \quad (7.25)$$

Pela análise da Eq. (6.4) do Capítulo 6, sabemos que o espectro de um sinal amostrado $\bar{p}(t)$ é $(1/T_b)$ vezes o espectro de $p(t)$ repetido periodicamente à frequência de amostragem R_b . Portanto, a transformada de Fourier dos dois lados da Eq. (7.25) produz

$$\frac{1}{T_b} \sum_{n=-\infty}^{\infty} P(f - nR_b) = 1 \quad \text{em que} \quad R_b = \frac{1}{T_b} \quad (7.26)$$

ou

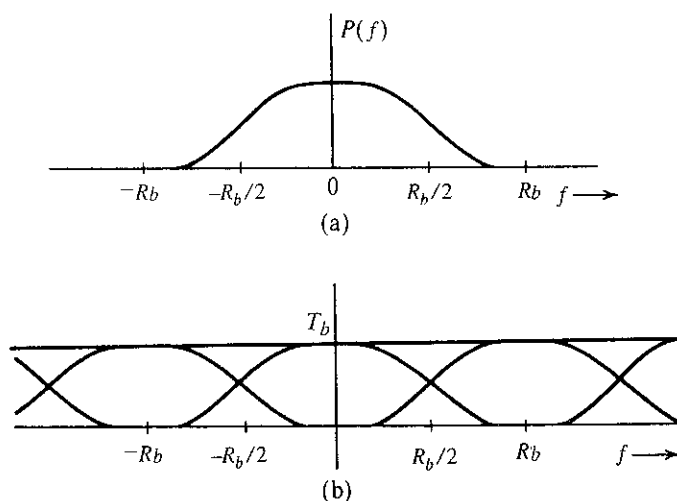
$$\sum_{n=-\infty}^{\infty} P(f - nR_b) = T_b \quad (7.27)$$

Assim, a soma dos espectros formados com a repetição de $P(f)$ à frequência R_b é uma constante T_b , como mostrado na Fig. 7.12b.*

Consideremos o espectro na Fig. 7.12b, em um intervalo $0 < f < R_b$. Esse intervalo envolve apenas dois termos, $P(f)$ e $P(f - R_b)$, do somatório na Eq. (7.27). Logo,

$$P(f) + P(f - R_b) = T_b \quad 0 < f < R_b$$

Figura 7.12
Dedução do
critério Nyquist
para ISI zero.



*Observemos que, se $R_b > 2B$, em que B é a largura de banda (em hertz) de $P(f)$, as repetições de $P(f)$ não se sobrepõem, e a condição (7.27) não pode ser satisfeita. Para $R_b = 2B$, a condição é satisfeita somente para o caso ideal passabaixas $P(f)$ [$p(t) = \text{sinc}(\pi R_b t)$], que não é realizável. Logo, devemos ter $B > R_b/2$.

Seja $x = f - R_b/2$; portanto,

$$P(x + 0,5R_b) + P(x - 0,5R_b) = T_b \quad |x| < 0,5R_b \quad (7.28a)$$

ou

$$P\left(x + \frac{R_b}{2}\right) + P\left(x - \frac{R_b}{2}\right) = T_b \quad |x| < 0,5R_b \quad (7.28b)$$

Usando a propriedade da simetria conjugada [Eq. (3.11)] na Eq. (7.28), temos

$$P\left(\frac{R_b}{2} + x\right) + P^*\left(\frac{R_b}{2} - x\right) = T_b \quad |x| < 0,5R_b \quad (7.29)$$

Se escolhermos $P(f)$ com valores reais e positivos, basta que $|P(f)|$ satisfaça a Eq. (7.29). Como $|P(f)|$ é real, a Eq. (7.29) implica em

$$\left|P\left(\frac{R_b}{2} + x\right)\right| + \left|P\left(\frac{R_b}{2} - x\right)\right| = T_b \quad |x| < 0,5R_b \quad (7.30)$$

Portanto, $|P(f)|$ deve ter a forma mostrada na Fig. 7.13. Essa curva tem simetria ímpar em relação ao conjunto de eixos que se cruzam no ponto α [ponto na curva de $|P(f)|$ em $f = R_b/2$]. Isso requer

$$|P(0,5R_b)| = 0,5|P(0)|$$

A largura de banda de $P(f)$ em hertz é $0,5R_b + f_x$, em que f_x é a largura de banda além da mínima largura de banda $R_b/2$. Seja r a razão entre o excesso de largura de banda f_x e a mínima largura de banda teórica $R_b/2$:

$$\begin{aligned} r &= \frac{\text{excesso de largura de banda}}{\text{mínima largura de banda teórica}} \\ &= \frac{f_x}{0,5R_b} \\ &= 2f_x/T_b \end{aligned} \quad (7.31)$$

Observemos que f_x não pode ser maior que $R_b/2$,

$$0 \leq r \leq 1 \quad (7.32)$$

Figura 7.13
Espectro
vestigial
(cosseno
levantado).

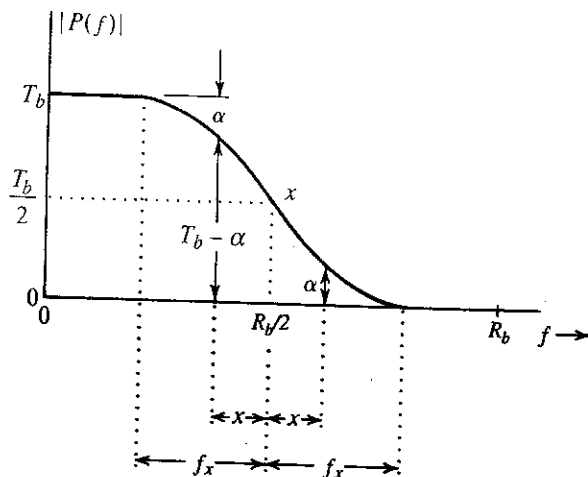
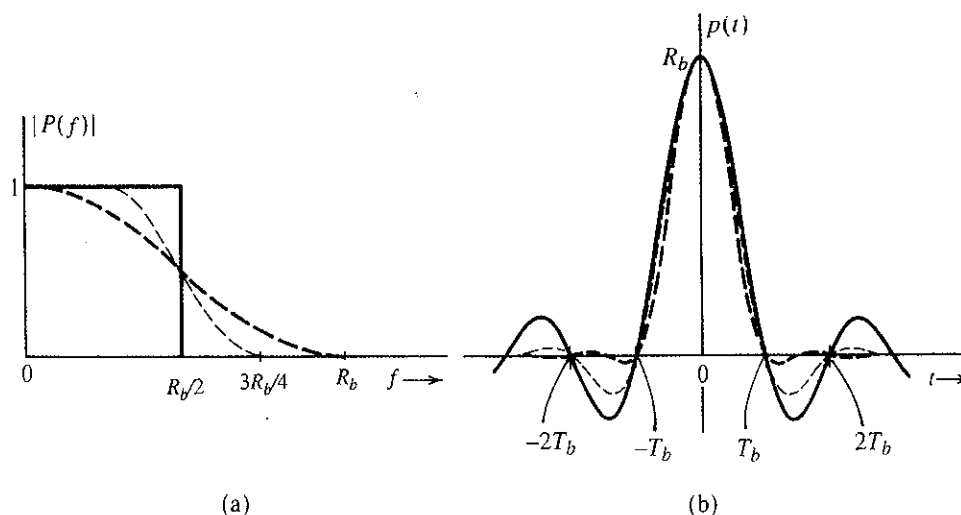


Figura 7.14

Pulsos que satisfazem o primeiro critério de Nyquist: linha cheia, $f_x = 0$ ($r = 0$) ideal; linha tracejada fina, $f_x = R_b/4$; linha tracejada grossa, $f_x = R_b/2$ ($r = 1$).



Em termos da frequência f , mínima largura de banda teórica é $R_b/2$ Hz, e o excesso de largura de banda é $f_x = rR_b/2$. Portanto, a largura de banda de $P(f)$ é

$$B_T = \frac{R_b}{2} + \frac{rR_b}{2} = \frac{(1+r)R_b}{2} \quad (7.33)$$

A constante r é denominada **fator de decaimento** (*roll-off*), e é também expressa em porcentagem. Por exemplo, se $P(f)$ for um espectro que segue o primeiro critério de Nyquist e cuja largura de banda seja 50% maior que o mínimo teórico, seu fator de decaimento será $r = 0,5$ ou 50%.

A modulação em banda lateral vestigial discutida na Seção 4.5 [Eq. (4.26)] requer um filtro cuja resposta de amplitude tenha essas mesmas características. Por essa razão, nos referimos ao espectro de $P(f)$ nas Eqs. (7.29) e (7.30) como **espectro vestigial**. O pulso $p(t)$ na Eq. (7.23) tem ISI zero nos centros de todos os outros pulsos transmitidos à taxa de R_b pulsos por segundo. Um pulso $p(t)$ que cause ISI zero nos centros de todos os outros pulsos (ou instantes de sinalização) é o pulso do primeiro critério de Nyquist. Mostramos que um pulso com espectro vestigial [Eq. (7.29) ou Eq. (7.30)] satisfaz o primeiro critério de Nyquist para ISI zero.

Como $0 \leq r < 1$, a largura de banda de $P(f)$ fica restrita ao intervalo entre $R_b/2$ e R_b Hz. O pulso $p(t)$ pode ser gerado como a resposta ao impulso unitário de um filtro com função de transferência $P(f)$. Contudo, como $P(f) = 0$ em uma banda de frequências, o critério de Paley-Wiener é violado e, portanto, o filtro é irrealizável. Todavia, a característica de decaimento vestigial é gradual e pode ser bem aproximada por um filtro prático. Uma família de espectros que satisfazem o primeiro critério de Nyquist é

$$P(f) = \begin{cases} 1, & |f| < \frac{R_b}{2} - f_x \\ \frac{1}{2} \left[1 - \sin \pi \left(\frac{f - R_b/2}{2f_x} \right) \right], & \left| f - \frac{R_b}{2} \right| < f_x \\ 0, & |f| > \frac{R_b}{2} + f_x \end{cases} \quad (7.34)$$

A Fig. 7.14a mostra três curvas dessa família, correspondendo a $f_x = 0$ ($r = 0$), $f_x = R_b/4$ ($r = 0,5$) e $f_x = R_b/2$ ($r = 1$). As respectivas respostas ao impulso são mostradas na Fig. 7.14b. Pode ser visto que o aumento de f_x (ou de r) melhora $p(t)$; ou seja, um decaimento mais gradual reduz a natureza oscilatória de $p(t)$ e faz com que $p(t)$ decaia mais rapidamente no domínio do tempo. Para o caso do valor máximo $f_x = R_b/2$ ($r = 1$), a Eq. (7.34) se reduz a

$$P(f) = \frac{1}{2} (1 + \cos \pi f T_b) \Pi \left(\frac{f}{2R_b} \right) \quad (7.35a)$$

$$= \cos^2 \left(\frac{\pi f T_b}{2} \right) \Pi \left(\frac{f T_b}{2} \right) \quad (7.35b)$$

A característica da Eq. (7.34) é conhecida na literatura como **cosseno levantado**, pois representa um cosseno levantado por sua amplitude de pico. A Eq. (7.35) também é conhecida como característica de **decaimento de cosseno amplo** (*full-cosine roll-off*). A transformada de Fourier inversa desse espectro é prontamente calculada como (Exercício 7.3-8)

$$p(t) = R_b \frac{\cos \pi R_b t}{1 - 4R_b^2 t^2} \text{sinc}(\pi R_b t) \quad (7.36)$$

Esse pulso é mostrado na Fig. 7.14b ($r = 1$). Podemos fazer várias observações importantes sobre o pulso cosseno levantado. Primeira, esse pulso, cuja largura de banda é R_b Hz, tem valor R_b em $t = 0$ e é zero não apenas em todos os outros instantes de amostragens, mas também nos pontos médios entre todos os instantes de sinalização. Segunda, o pulso decai rapidamente, com $1/t^3$. Como resultado, o pulso cosseno levantado é relativamente insensível a desvios de R_b , da taxa de amostragem, incerteza temporal e assim por diante. Além disso, o filtro de geração de pulso com função de transferência $P(f)$ [Eq. (7.35b)] é realizável em boa aproximação. A característica de fase desse filtro é praticamente linear, de modo que nenhuma equalização de fase adicional se faz necessária.

Vale lembrar que os pulsos recebidos na entrada do detector é que devem ter a forma para ISI zero. Na prática, como o canal não é ideal (sem distorção), os pulsos transmitidos devem ser formatados para que, após passarem pelo canal com função de transferência $H_c(f)$, sejam recebidos com a forma adequada (como pulsos cosseno levantado). Portanto, o pulso transmitido $p_i(t)$ deve satisfazer

$$P_i(f)H_c(f) = P(f)$$

em que $P(f)$ tem o espectro vestigial da Eq. (7.30). Por conveniência, a função de transferência $H_c(f)$ como um canal também pode incluir um filtro receptor projetado para rejeitar interferências e outros ruídos fora da banda.

Exemplo 7.1 Determinemos a taxa de transmissão de pulsos em termos da largura de banda de transmissão B_T e do fator de decaimento r . Consideremos um esquema que use o primeiro critério de Nyquist.

Da Eq. (7.33)

$$R_b = \frac{2}{1+r} B_T$$

Como $0 \leq r \leq 1$, a taxa de transmissão de pulsos varia de $2B_T$ a B_T , dependendo da escolha de r . Um menor valor de r produz uma taxa de sinalização mais alta. Contudo, o pulso $p(t)$ decai lentamente, criando o mesmo problema discutido em relação ao pulso sinc. Para o pulso cosseno levantado $r = 1$ e $R_b = B_T$, alcançamos a metade da máxima taxa teórica. Mas, o pulso decai mais rapidamente, com $1/t^3$ e é menos vulnerável à ISI.

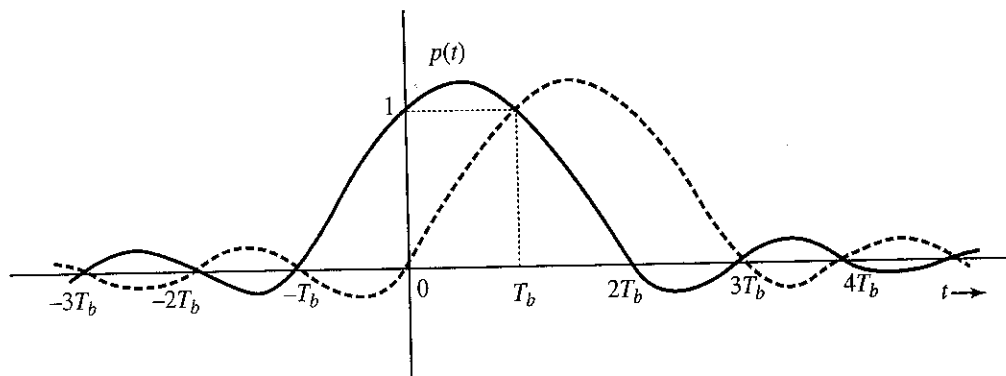
7.3.3 ISI Controlada ou Sinalização de Resposta Parcial

Pulsos que seguem o critério de Nyquist requerem uma largura de banda maior que o mínimo teórico. Se desejarmos reduzir a largura de banda do pulso, devemos encontrar uma forma de alargar o pulso $p(t)$ (mais largo o pulso, menor sua largura de banda). O aumento da largura do pulso pode resultar em interferência (ISI) em pulsos vizinhos. No entanto, na transmissão binária com apenas dois símbolos, pode ser possível remover ou cancelar um grau conhecido e controlado de ISI, pois somente uns poucos padrões de interferência são possíveis.

Consideremos um pulso especificado por (Fig. 7.15):

$$p(nT_b) = \begin{cases} 1 & n = 0, 1 \\ 0 & \text{para todos outros valores de } n \end{cases} \quad (7.37)$$

Figura 7.15
Comunicação
usando ISI
controlada ou o
segundo critério
de Nyquist.



Isso leva a um grau conhecido e controlado de ISI causada pelo k -ésimo pulso no próximo pulso transmitido. Usemos sinalização polar com esse pulso. Assim, 1 é transmitido por $p(t)$ e 0, pelo pulso $-p(t)$. O sinal recebido é amostrado em $t = nT_b$, e o pulso $p(t)$ tem valor zero em todos os n , exceto em $n = 0$ e 1, quando vale 1 (Fig. 7.15). Fica claro que esse pulso causa ISI zero em todos os outros pulsos, exceto o que o sucede. Consideremos dois pulsos sucessivos localizados em 0 e T_b , respectivamente. Se os dois pulsos forem positivos, o valor amostrado do sinal resultante em $t = T_b$ será 2. Se os dois pulsos forem negativos, o valor amostrado será -2. Contudo, se os dois pulsos tiverem polaridades opostas, o valor amostrado será zero. Como somente esses três valores são possíveis, o valor da amostra do sinal permite que tomemos a decisão correta nos instantes de amostragem. A regra de decisão é a seguinte: se o valor da amostra for positivo, o bit corrente é 1 e o bit anterior também é 1. Se o valor da amostra for negativo, o bit atual é 0 e o bit anterior também é 0. Se o valor da amostra for zero, o bit atual é o oposto do anterior. O conhecimento do bit anterior permite a determinação do bit corrente.

A Tabela 7.1 mostra uma sequência de bits transmitidos, os valores das amostras do sinal recebido $x(t)$ (admitindo a ausência de erros devido a ruído do canal) e a decisão do detector. Esse exemplo também indica a propriedade de detecção de erro desse esquema. Um exame das amostras da forma de onda $y(t)$ na Tabela 7.1 revela que sempre há um número par de amostras de valor zero entre duas amostras de valores completos de mesma polaridade, e um número ímpar de amostras de valor zero entre duas amostras de valores completos de polaridades opostas. Assim, o primeiro valor amostrado de $x(t)$ é 2, e o próximo valor completo de amostra (da quarta amostra) é 2. Entre essas duas amostras de valores completos de mesma polaridade, há um número par (ou seja, 2) de amostras de valor zero. Se um dos valores de amostra for detectado erroneamente, essa regra será violada e o erro, detectado.

Tabela 7.1

Bits transmitidos e amostras recebidas na sinalização com ISI controlada

Sequência de informação	1	1	0	1	1	0	0	0	1	0	1	1	1
Amostra $y(kT_b)$	1	2	0	0	2	0	-2	-2	0	0	0	2	2
Sequência detectada	1	1	0	1	1	0	0	0	1	0	1	1	1

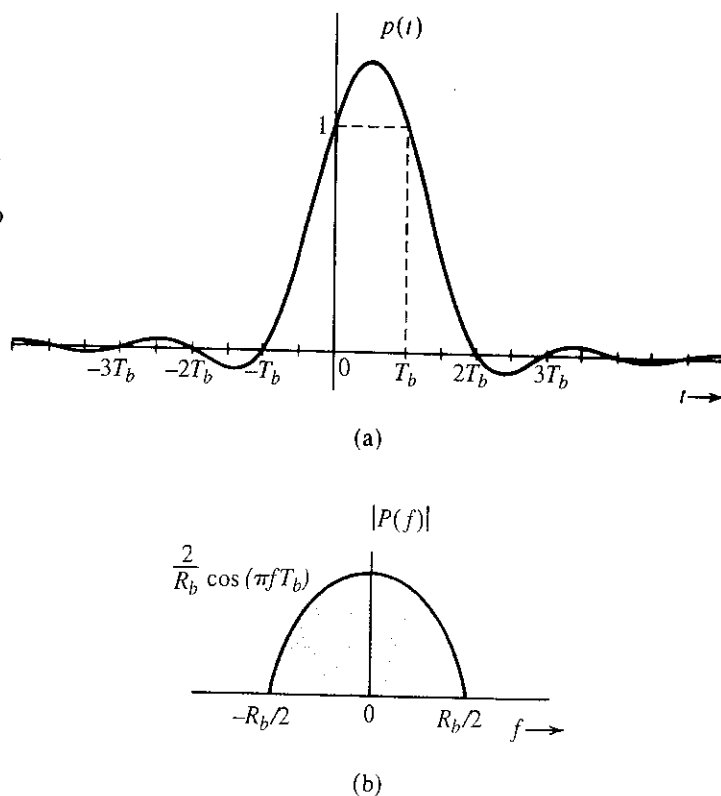
O pulso $p(t)$ vai a zero em $t = -T_b$ e $2T_b$, resultando em uma largura de pulso (do primeiro lóbulo) 50% maior que a do pulso do primeiro critério. Tal alargamento do pulso no domínio do tempo reduz a largura de banda do mesmo. Esse é o segundo critério proposto por Nyquist. Esse esquema de ISI controlada também é conhecido como esquema de **resposta parcial** ou **correlativa**. Um pulso que satisfaça o segundo critério na Eq. (7.37) também é conhecido como **pulso duobinário**.

7.3.4 Exemplo de Pulso Duobinário

Se restringirmos a largura de pulso a $R_b/2$, seguindo o procedimento do Exemplo 7.1, podemos mostrar que (Exercício 7.3-9) somente o pulso $p(t)$ dado a seguir satisfaz a exigência na Eq. (7.37) para o pulso duobinário:

$$p(t) = \frac{\text{sen}(\pi R_b t)}{\pi R_b t(1 - R_b t)} \quad (7.38)$$

Figura 7.16
(a) Mínima largura de banda de pulso que satisfaz o critério de pulso duobinário e (b) seu espectro de amplitude.



A transformada de Fourier $P(f)$ do pulso $p(t)$ é dada por (Exercício 7.3-9):

$$P(f) = \frac{2}{R_b} \cos\left(\frac{\pi f}{R_b}\right) \Pi\left(\frac{f}{R_b}\right) e^{-j\pi f/R_b} \quad (7.39)$$

O pulso $p(t)$ e seu espectro de amplitude $|P(f)|$ são mostrados na Fig. 7.16.* Esse pulso transmite dados binários à taxa de R_b bit/s e tem a mínima largura de banda teórica de $R_b/2$ Hz. A Eq. (7.38) mostra que esse pulso decai rapidamente com o tempo, com $1/t^2$. Esse pulso não é idealmente realizável, pois $p(t)$ é não causal e tem duração infinita [pois $P(f)$ é limitado em frequência]. Contudo, decai rapidamente (com $1/t^2$) e, portanto, pode ser bem aproximado.

Pode parecer surpreendente que possamos alcançar a taxa teórica com o pulso duobinário. Na verdade, é uma ilusão. A taxa teórica de transmissão é 2 porções independentes de informação por segundo por hertz de largura de banda. Alcançamos essa taxa para informação binária. Aqui está o problema! Uma porção de informação binária não se qualifica como uma porção independente de informação, pois não pode assumir um valor arbitrário: deve ser selecionada de um conjunto finito. O pulso duobinário falharia se os pulsos fossem verdadeiramente porções independentes de informação, ou seja, se os pulsos pudessem assumir amplitudes arbitrárias. O esquema funciona somente porque os pulsos binários assumem valores finitos conhecidos e, portanto, existe apenas um número finito (conhecido) de padrões de interferência entre pulsos, o que permite a determinação correta das amplitudes dos pulsos, apesar da ocorrência de interferência.

7.3.5 Relações de Pulsos entre Sinalizações com ISI Zero, Duobinária e Duobinária Modificada

Agora, podemos estabelecer a relação simples entre um pulso $p_a(t)$ que satisfaz o primeiro critério de Nyquist (ISI zero) e um pulso duobinário $p_b(t)$ (com ISI controlada). Das Eqs. (7.23) e (7.37), fica

*O espectro de fase é linear: $\theta_p(f) = -\pi f T_b$.

claro que $p_a(kT_b)$ e $p_b(kT_b)$ diferem somente para $k = 1$. Para todos os outros valores do inteiro k , esses pulsos têm amostras com valores iguais. Portanto, podemos construir um pulso $p_b(t)$ de $p_a(t)$, da seguinte forma

$$p_b(t) = p_a(t) + p_a(t - T_b)$$

Essa adição é a sinalização com ISI “controlada” ou resposta parcial, deliberadamente introduzida para reduzir a exigência de largura de banda. Para determinar o efeito da sinalização “duobinária” na largura de banda espectral, consideremos a relação entre os dois pulsos no domínio da frequência:

$$P_b(f) = P_a(f)[1 + e^{-j2\pi f T_b}] \quad (7.40a)$$

$$|P_b(f)| = |P_a(f)|\sqrt{2(1 + \cos(2\pi f T_b))}2|\cos(\pi f T_b)| \quad (7.40b)$$

Podemos ver que a sinalização de resposta parcial, na verdade, força um nulo de $P(f)$ em $2\pi f T_b = \pi$ ou $f = 0,5/T_b$. Portanto, conceitualmente, podemos ver como a sinalização de resposta parcial oferece uma oportunidade adicional para reformatar a PSD ou largura de banda de transmissão. De fato, a sinalização duobinária, ao forçar um nulo na frequência $0,5/T_b$, faz com que sua largura de banda essencial seja a mínima largura de banda de transmissão necessária para uma taxa de dados de $1/T_b$ (como discutido na Seção 6.1.3).

Na verdade, muitos canais físicos, como uma gravação magnética, têm ganho zero em dc. Por tanto, não faz sentido que o sinal em banda base tenha qualquer componente dc em sua PSD. Uma modificação da sinalização de resposta parcial é comumente adotada para forçar um nulo em dc. Um exemplo notável é a chamada sinalização **duobinária modificada**, que requer

$$p_c(nT_b) = \begin{cases} 1 & n = -1 \\ -1 & n = 1 \\ 0 & \text{para todos outros inteiros } n \end{cases} \quad (7.41)$$

Um raciocínio similar indica que $p_c(t)$ pode ser gerado de qualquer pulso $p_a(t)$ que satisfaça o primeiro critério de Nyquist, via

$$p_c(t) = p_a(t + T_b) - p_a(t - T_b)$$

De modo equivalente, no domínio da frequência, o pulso duobinário é

$$P_c(f) = 2jP_a(f)\sin(2\pi f T_b)$$

que usa $\sin(2\pi f T_b)$ para forçar um nulo em dc e se conformar à limitação física do canal.

7.3.6 Detecção de Sinalização Duobinária e Codificação Diferencial

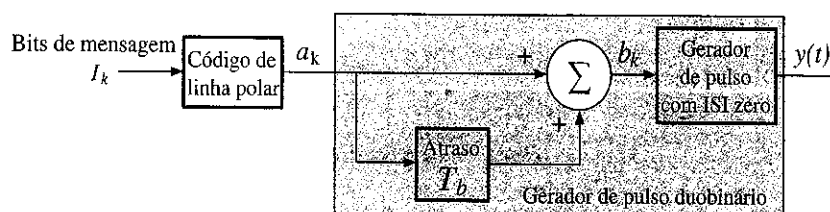
O diagrama básico do transmissor para o método de ISI controlada da sinalização duobinária é mostrado na Fig. 7.17. Agora, examinemos em mais detalhe a relação entre todos os símbolos de dados na banda base e o procedimento de detecção. Para o bit de mensagem binária $I_k = 0$ ou 1, símbolos polares são simplesmente

$$a_k = 2I_k - 1$$

Com ISI controlada, as amostras do sinal de transmissão $y(t)$ são

$$y(kT_b) = b_k = a_k + a_{k-1} \quad (7.42)$$

Figura 7.17
Sinalização
duobinária
equivalente.



Para o receptor, a questão é como detectar I_k de $y(kT_b)$ ou b_k . Essa questão pode ser respondida considerando, primeiro, todos os valores possíveis de b_k ou $y(kT_b)$. Como $a_k = \pm 1$, $b_k = 0, \pm 2$. Da Eq. (7.42), fica evidente que

$$\begin{aligned} b_k = 2 &\Rightarrow a_k = 1 && \text{ou } I_k = 1 \\ b_k = -2 &\Rightarrow a_k = -1 && \text{ou } I_k = 0 \\ b_k = 0 &\Rightarrow a_k = -a_{k-1} && \text{ou } I_k = 1 - I_{k-1} \end{aligned} \quad (7.43)$$

Portanto, um simples detector de sinalização duobinária deve, primeiro, detectar todos os bits I_k correspondentes a $b_k = \pm 2$. Os remanescentes $\{b_k\}$ são amostras de valor zero que implicam transição; ou seja, o dígito corrente é 1, e o anterior, 0 ou vice-versa. Isso significa que a detecção de dígitos deve ser baseada no dígito anterior. Um exemplo dessa detecção dígito a dígito foi mostrado na Tabela 7.1. A desvantagem do método de detecção na Eq. (7.43) é que, quando $y(kT_b) = 0$, a decisão para o bit corrente depende da decisão para o bit anterior. Se o dígito anterior tiver sido detectado de modo incorreto, o erro tende a se propagar, até que surja um valor de amostra ± 2 . Para mitigar esse problema de propagação de erro, aplicamos um mecanismo eficiente, conhecido como **codificação diferencial**.

A Fig. 7.18 ilustra um gerador de sinal duobinário, no qual há um codificador diferencial adicional antes da geração do pulso de resposta parcial. Como mostrado na Fig. 7.18, a codificação diferencial é uma etapa muito simples que modifica a relação entre o código de linha e os bits de mensagem. A codificação diferencial gera uma nova sequência binária

$$p_k = I_k \oplus p_{k-1} \quad \text{módulo 2}$$

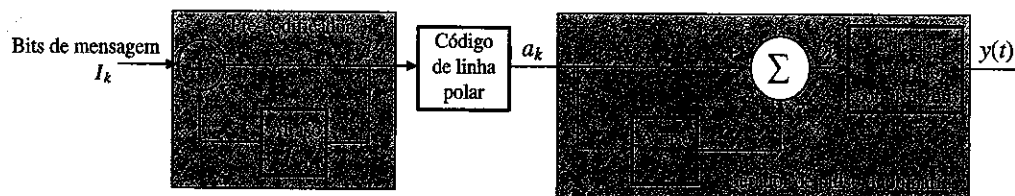
supondo que o estado inicial do pré-codificador seja $p_0 = 0$ ou $p_0 = 1$. A saída do pré-codificador é aplicada à entrada de um codificador de linha polar e gera

$$a_k = 2p_k - 1$$

Devido à sinalização duobinária $b_k = a_k + a_{k-1}$ e ao gerador de pulso com ISI zero, as amostras do sinal recebido $y(t)$ sem ruído se tornam

$$\begin{aligned} y(kT_b) &= b_k = a_k + a_{k-1} \\ &= 2(p_k + p_{k-1}) - 2 \\ &= 2(p_{k-1} \oplus I_k + p_{k-1} - 1) \\ &= \begin{cases} 2(1 - I_k) & p_{k-1} = 1 \\ 2(I_k - 1) & p_{k-1} = 0 \end{cases} \end{aligned} \quad (7.44)$$

Figura 7.18
Sinalização
duobinária com
codificação
diferencial.



A partir da Eq. (7.44), podemos resumir a relação direta entre os bits de mensagem e os valores de amostras como

$$y(kT_b) = \begin{cases} 0 & I_k = 1 \\ \pm 2 & I_k = 0 \end{cases} \quad (7.45)$$

Essa relação funciona como uma base para um algoritmo de detecção símbolo a símbolo. Em poucas palavras, o algoritmo de decisão se baseia na amostra corrente $y(kT_b)$; quando não há ruído, $y(kT_b) = b_k$ e a decisão do receptor é

$$I_k = \frac{2 - |y(kT_b)|}{2} \quad (7.46)$$

Portanto, a incorporação da codificação diferencial à sinalização duobinária não apenas simplifica a regra de decisão, mas também torna a decisão independente do dígito anterior e elimina a propagação de erro. Na Tabela 7.2, o exemplo da Tabela 7.1 é recalculado com a codificação diferencial. A relação de decodificação da Eq. (7.45) fica clara nesse exemplo.

Tabela 7.2

Sinalização duobinária com codificação diferencial

Tempo k	0	1	2	3	4	5	6	7	8	9	10	11	12	13
I_k		1	1	0	1	1	0	0	0	1	0	1	1	1
p_k	0	1	0	0	1	0	0	0	0	1	1	0	1	0
a_k	-1	1	-1	-1	1	-1	-1	-1	-1	1	1	-1	1	-1
b_k		0	0	-2	0	0	-2	-2	-2	0	2	0	0	0
Bits detectados		1	1	0	1	1	0	0	0	1	0	1	1	1

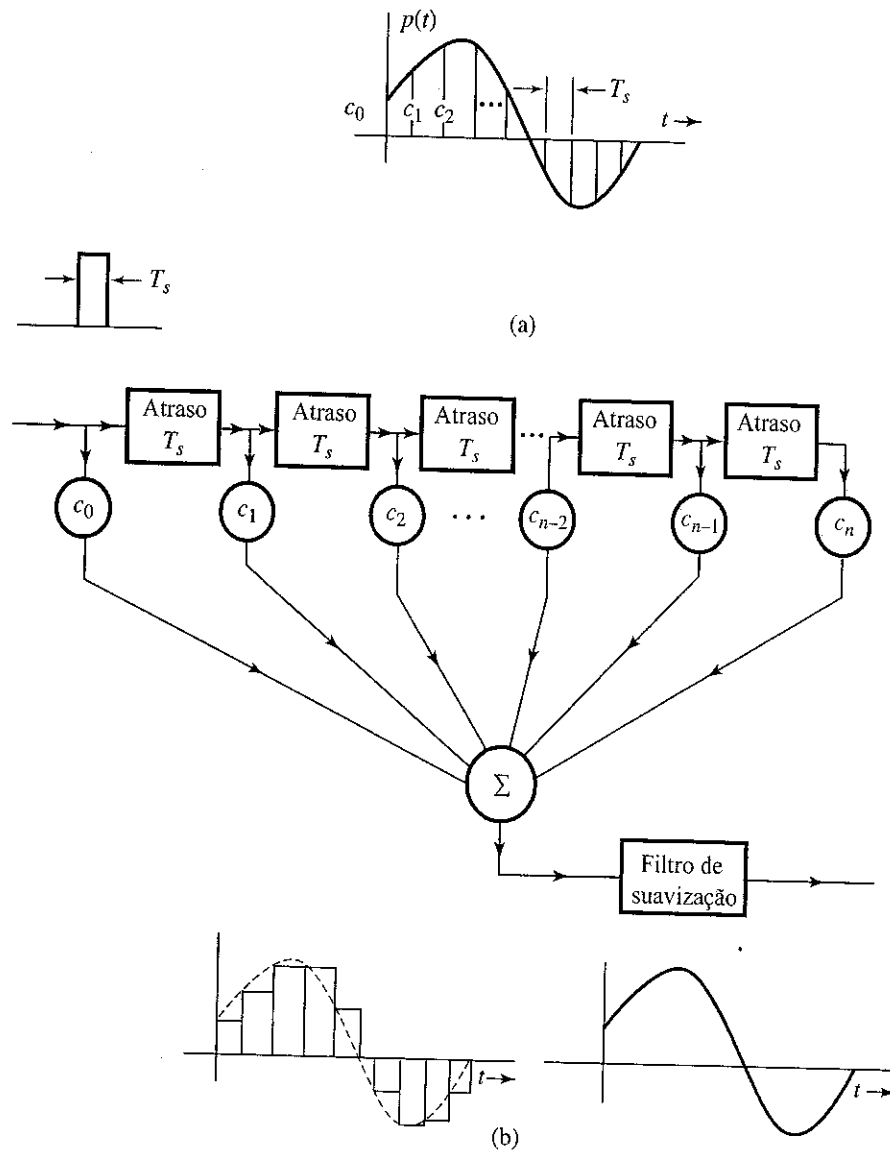
A codificação diferencial definida para símbolos de informação binária pode ser convenientemente generalizada para símbolos não binários. Quando os símbolos de informação I_k são M -ários, a única modificação a ser feita no bloco de codificação diferencial é a substituição de “módulo 2” por “módulo M ”. De modo similar, outra sinalização de resposta parcial generalizada, como a duobinária modificada, também está sujeita ao problema de propagação de erro na detecção. Um tipo adequado de codificação diferencial pode ser adotado para evitar a propagação de erro.

7.3.7 Geração de Pulsos

Um pulso $p(t)$ que satisfaça um critério de Nyquist pode ser gerado como a resposta ao impulso unitário de um filtro com função de transferência $P(f)$. Isso nem sempre é fácil. Um método melhor consiste em gerar a forma de onda diretamente, com o uso de um filtro transversal (linha de retardo com derivação) discutido a seguir. O pulso $p(t)$ a ser gerado é amostrado com um intervalo de amostragem T_s suficientemente pequeno (Fig. 7.19a), e os ganhos da derivação do filtro são ajustados proporcionalmente aos valores amostrados em sequência, como indicado na Fig. 7.19b. Quando um pulso retangular estreito, de largura T_s (intervalo de amostragem), é aplicado à entrada do filtro transversal, a saída será uma aproximação em degrau de $p(t)$. Essa saída, ao ser aplicada a um filtro passa-baixas, é suavizada. A aproximação pode ser melhorada com a redução do intervalo de amostragem T_s .

Vale ressaltar, mais uma vez, que os pulsos que chegam à entrada do detector no receptor devem atender o desejado critério de Nyquist. Portanto, os pulsos transmitidos devem ser formatados de modo que, após passagem pelo canal, sejam recebidos no formato (de Nyquist) desejado. Na prática, no entanto, não é necessário que pulsos sejam rigidamente formatados no transmissor. A formatação final pode ser efetuada por um equalizador no receptor, como discutido mais adiante (Seção 7.5).

Figura 7.19
Geração de pulso por filtro transversal.



7.4 EMBARALHAMENTO (SCRAMBLING) DE DADOS

Em geral, um embaralhador tende a aumentar o caráter aleatório de dados através da remoção de seqüências de 1s ou 0s. O embaralhamento pode ser útil na extração da temporização, por remover longas seqüências de 0s de dados binários. Contudo, embaralhadores são usados principalmente para evitar acesso não autorizado aos dados, sendo otimizados para esse fim. Essa otimização pode resultar na geração de uma longa seqüência de zeros nos dados. A rede digital deve ser capaz de tratar de tais longas seqüências de zeros com o emprego de técnicas de substituição discutidas na Seção 7.2.

A Fig. 7.20 mostra típicos embaralhador e desembaralhador. O embaralhador consiste em um registrador de deslocamento com realimentação (*feedback shift register*), e o correspondente desembaralhador tem um registrador de deslocamento com antecipação (*feedforward shift register*), como indicado na Fig. 7.20. Cada estágio do registrador de deslocamento atrasa um bit em uma unidade. Para analisar o embaralhador e o correspondente desembaralhador, consideremos uma seqüência de saída T do embaralhador (Fig. 7.20a). Se S for a seqüência de entrada do embaralhador,

$$S \oplus D^3T \oplus D^5T = T \quad (7.47)$$

Como a soma em módulo 2 de qualquer sequência com ela própria é zero, $D^8 \oplus D^8 = 0$ e

$$F^2 = D^6 \oplus D^{10}$$

Do mesmo modo,

$$F^3 = (D^6 \oplus D^{10})(D^3 \oplus D^5) = D^9 \oplus D^{11} \oplus D^{13} \oplus D^{15}$$

e assim por diante. Logo [Eq. (7.49)],

$$T = (1 \oplus D^3 \oplus D^5 \oplus D^6 \oplus D^9 \oplus D^{10} \oplus D^{11} \oplus D^{12} \oplus D^{13} \oplus D^{15} \dots)S$$

Como $D^n S$ é simplesmente a sequência S atrasada em n bits, vários termos na equação anterior correspondem às seguintes sequências:

$$S = 101010100000111$$

$$D^3 S = 000101010100000111$$

$$D^5 S = 00000101010100000111$$

$$D^6 S = 000000101010100000111$$

$$D^9 S = 000000000101010100000111$$

$$D^{10} S = 0000000000101010100000111$$

$$D^{11} S = 00000000000101010100000111$$

$$D^{12} S = 000000000000101010100000111$$

$$D^{13} S = 0000000000000101010100000111$$

$$D^{15} S = 000000000000000101010100000111$$

$$T = 101110001101001$$

Notemos que a sequência de entrada contém a sequência periódica **10101010...**, assim como uma longa sequência de **0s**. A saída do embaralhador remove a componente periódica, assim como a longa sequência de **0s**. A sequência de entrada tem 15 dígitos. A saída do embaralhador é mostrada somente até o 15º dígito, pois todos os outros dígitos além do 15º dependem dos dígitos de entrada além do 15º, que não foram dados.

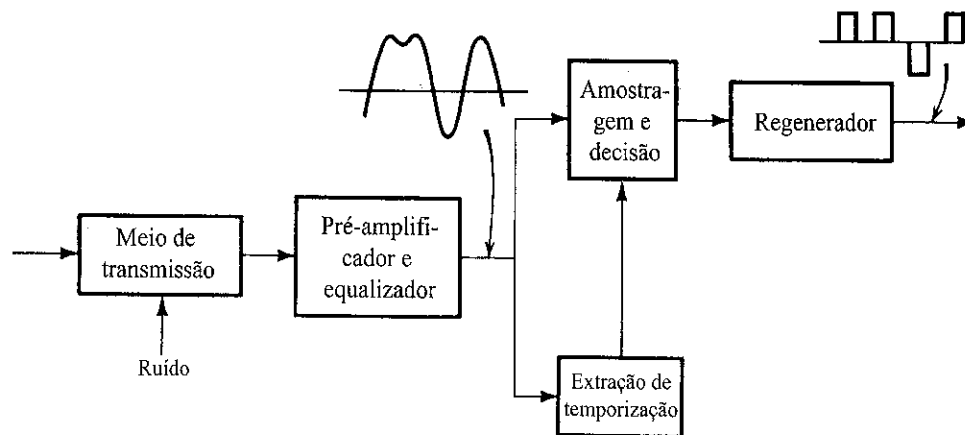
O leitor pode comprovar que a saída do desembaralhador é, de fato, S , se a sequência T anterior for aplicada à sua entrada.

7.5 RECEPTORES DIGITAIS E REPETIDORES REGENERATIVOS

Basicamente, um receptor ou repetidor regenerativo executa três funções: (1) reformata pulsos recebidos por meio de um equalizador, (2) extrai informação de temporização necessária para amostrar os pulsos recebidos nos instantes ótimos, e (3) decide que símbolo foi detectado, com base em amostras de pulsos. O repetidor mostrado na Fig. 7.21 consiste em um receptor mais um "regenerador". Um receptor completo pode incluir ainda a provisão para separar potência dc e sinais ac. Isso é normalmente feito por acoplamento ac dos sinais por transformador, e desviando dc do transformador para o circuito de alimentação de potência.*

*O repetidor, em geral, inclui circuito para proteger a eletrônica do regenerador de transientes de alta tensão induzidos por surtos de potência ou descargas elétricas. Enrolamentos especiais de transformador podem ser fornecidos para acoplar sinais espúrios a cabos dedicados.

Figura 7.21
Repetidor
regenerador.



7.5.1 Equalizadores

Um trem de pulsos é atenuado e distorcido pelo meio de transmissão. A atenuação pode ser compensada pelo pré-amplificador, enquanto a distorção é compensada pelo equalizador. A distorção de canal tem a forma de dispersão, causada pela atenuação de certas *componentes em frequências críticas* do trem de pulsos de dados. Teoricamente, um equalizador deve ter uma característica de frequência igual ao inverso da característica do meio de transmissão. Isso restaurará as componentes em frequências críticas e eliminará a dispersão dos pulsos. Lamentavelmente, isso também realça o ruído de canal recebido, pois amplifica as componentes desse nas frequências críticas. Esse fenômeno indesejável é conhecido como *amplificação de ruído*.

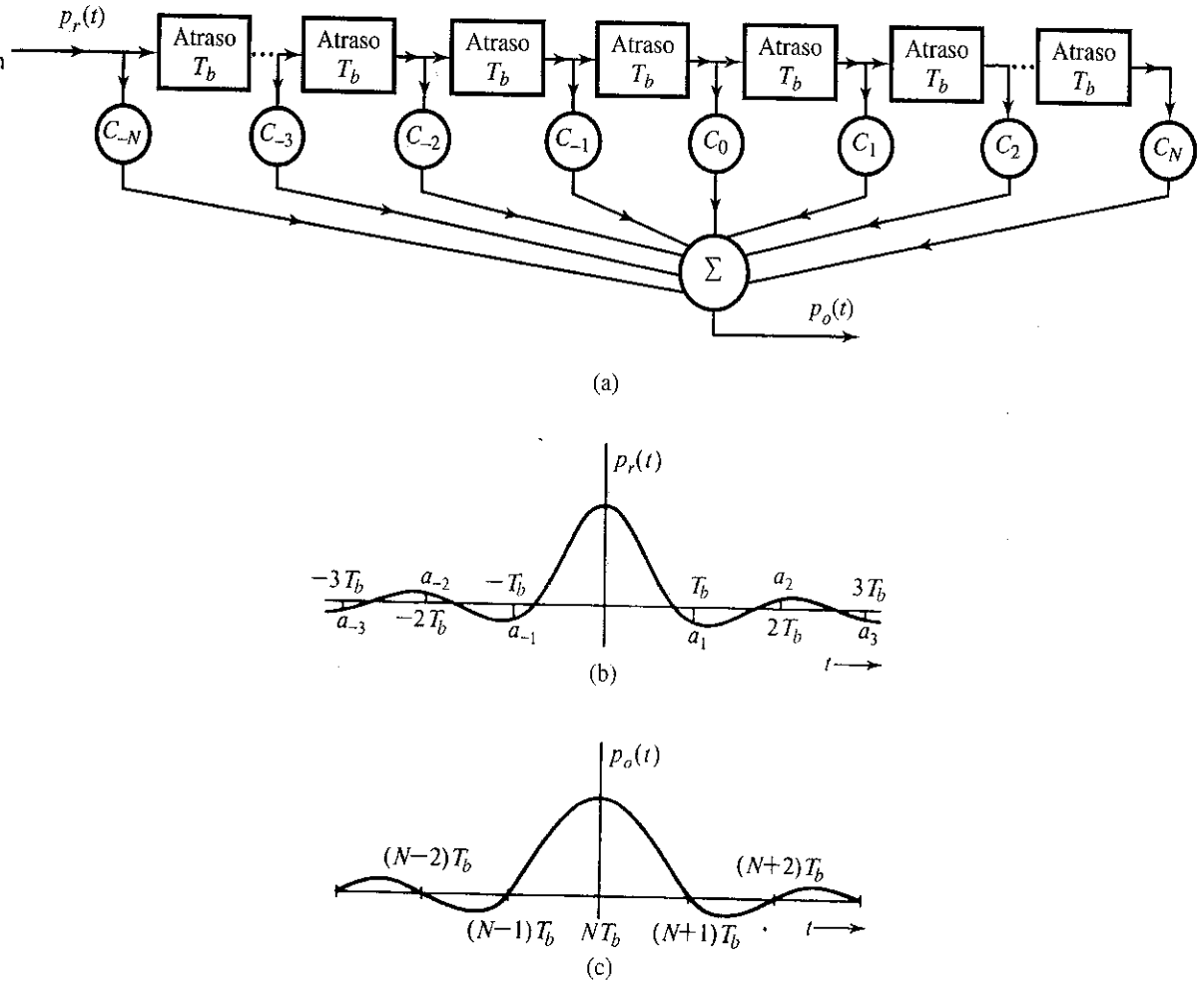
Para sinais digitais, no entanto, a equalização completa não é necessária, pois o detector precisa tomar decisões relativamente simples – como decidir se o pulso é positivo ou negativo (ou se o pulso está presente ou ausente). Portanto, considerável dispersão de pulsos pode ser tolerada. Dispersão de pulsos resulta em ISI e, por conseguinte, em aumento no erro de detecção. O aumento de ruído resultante da equalização (que amplifica as frequências altas) também aumenta a probabilidade de erro de detecção. Por essa razão, o projeto de um equalizador ótimo envolve um inevitável equilíbrio entre redução de ISI e redução de ruído de canal. Uma escolha acertada das características de equalização é uma questão central em todos os sistemas de comunicação digital bem projetados.⁶

Equalizador com Forçamento a Zero

A eliminação ou minimização da ISI (interferência) entre pulsos vizinhos para todo t não é, de fato, necessária. Basta eliminar ou minimizar a interferência entre pulsos vizinhos em seus respectivos *instantes de amostragem*, pois a decisão do receptor é baseada somente em valores de amostras. Esse tipo de equalização (abrandada) pode ser feito com equalizadores que usem a estrutura de filtro transversal discutida anteriormente. Ao contrário de filtros tradicionais, equalizadores a filtros transversais são facilmente ajustáveis para compensar efeitos de canais diferentes ou que variem lentamente no tempo. O objetivo de projeto é forçar que o pulso de saída do equalizador tenha ISI zero nos instantes de amostragem (tomada de decisão). Em outras palavras, os pulsos de saída do equalizador satisfazem o critério de Nyquist ou o de ISI controlada. O atraso temporal T entre derivações sucessivas é escolhido como T_b , o intervalo entre pulsos.

Inicialmente, tomemos os ganhos das derivações $c_0 = 1$ e $c_k = 0$, para todos os valores de k , no filtro transversal na Fig. 7.22a. Com isso, a saída do filtro será igual à entrada atrasada por NT_b . Para um pulso $p_r(t)$ (Fig. 7.22b) na entrada do filtro transversal com essa especificação de derivação, a saída do filtro $p_o(t)$ será exatamente $p_r(t - NT_b)$, ou seja, $p_r(t)$ atrasado por NT_b . Esse atraso não tem efeito prático em nosso sistema de comunicação e não é relevante para nossa discussão. Assim, por conveniência, ignoraremos esse atraso. Isso significa que $p_r(t)$ na Fig. 7.22b também representa a saída do filtro

Figura 7.22
Análise de
equalizador com
forçamento a
zero.



$p_o(t)$ para essa configuração de derivações ($c_0 = 1$ e $c_k = 0, k \neq 0$). Exigimos que esse pulso satisfaça o critério de Nyquist ou o critério de ISI controlada, como for o caso. Da Fig. 7.22b, vemos que as amplitudes de pulso a_1, a_{-1} , e a_2 em $T_b, -T_b$ e $2T_b$, respectivamente, não são desprezíveis. Ajustando os ganhos das derivações (c_k), geramos pulsos deslocados adicionais de amplitudes apropriadas que forçarão o resultante pulso de saída a ter os valores desejados em $t = 0, \pm T_b, \pm 2T_b, \dots$

A saída $p_o(t)$ (Fig. 7.22c) é a soma de pulsos da forma $c_k p_r(t - kT_b)$ (ignorando o atraso NT_b). Assim,

$$p_o(t) = \sum_{n=-N}^N c_n p_r(t - nT_b) \quad (7.50)$$

As amostras de $p_o(t)$ em $t = kT_b$ são

$$p_o(kT_b) = \sum_{n=-N}^N c_n p_r(kT_b - nT_b) \quad k = 0, \pm 1, \pm 2, \pm 3, \dots \quad (7.51a)$$

Usando uma notação mais conveniente $p_r[k]$ para denotar $p_r(kT_b)$, e $p_o[k]$ para denotar $p_o(kT_b)$, a Eq. (7.51a) pode ser expressa como

$$p_o[k] = \sum_{n=-N}^N c_n p_r[k - n] \quad k = 0, \pm 1, \pm 2, \pm 3, \dots \quad (7.51b)$$

O primeiro critério de Nyquist requer amostras $p_o[k] = 0$ para $k \neq 0$ e $p_o[k] = 1$ para $k = 0$. Substituindo esses valores na Eq. (7.51b), obtemos um conjunto de infinitas equações simultâneas em termos de $2N + 1$ variáveis. Obviamente, a solução dessas equações não é possível. Contudo, se especificarmos os valores de $p_o[k]$ somente em $2N + 1$ pontos como

$$p_o[k] = \begin{cases} 1 & k = 0 \\ 0 & k = \pm 1, \pm 2, \dots, \pm N \end{cases} \quad (7.52)$$

existirá uma única solução. Isso assegura que um pulso terá interferência zero nos instantes de amostragem de N pulsos anteriores e N pulsos posteriores. Como as amplitudes dos pulsos decaem rapidamente, em geral, interferência além do N -ésimo pulso não é significativa para $N > 2$. A substituição da condição (7.52) na Eq. (7.51b) produz um conjunto de $2N + 1$ equações simultâneas para $2N + 1$ incógnitas. Essas $2N + 1$ equações podem ser reescritas em forma matricial como

$$\underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{p}_o} = \underbrace{\begin{bmatrix} p_r[0] & p_r[-1] & \cdots & p_r[-2N+1] & p_r[-2N] \\ p_r[1] & p_r[0] & \cdots & p_r[-2N+2] & p_r[-2N+1] \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ p_r[2N-1] & p_r[2N-2] & \cdots & p_r[0] & p_r[-1] \\ p_r[2N] & p_r[2N-1] & \cdots & p_r[1] & p_r[0] \end{bmatrix}}_{\mathbf{P}_r} \underbrace{\begin{bmatrix} c_{-N} \\ c_{-N+1} \\ \vdots \\ c_{-1} \\ c_0 \\ c_1 \\ \vdots \\ c_{N-1} \\ c_N \end{bmatrix}}_{\mathbf{c}} \quad (7.53)$$

Nessa expressão compacta, a matriz \mathbf{P}_r de ordem $(2N + 1) \times (2N + 1)$ tem entradas idênticas ao longo de todas as diagonais. Esse tipo de matriz é conhecida como matriz de Toeplitz e é comumente encontrado na descrição de relações convolutivas. Uma matriz de Toeplitz é totalmente determinada pelas primeiras linha e coluna, tem algumas propriedades interessantes e admite algoritmos mais simples para o cálculo de sua inversa (por exemplo, os métodos de Trench⁷). Os ganhos c_k das derivações podem ser obtidos da solução dessa equação, com o cálculo da inversa da matriz \mathbf{P}_r ,

$$\mathbf{c} = \mathbf{P}_r^{-1} \mathbf{p}_o$$

Exemplo 7.3 Para o pulso recebido $p_r(t)$ na Fig. 7.22b, sejam

$$\begin{aligned} p_r[0] &= 1 \\ p_r[1] &= -0,3 & p_r[2] &= 0,1 \\ p_r[-1] &= -0,2 & p_r[-2] &= 0,05 \end{aligned}$$

Projetemos um equalizador de três derivações ($N = 1$).

Substituindo os valores anteriores na Eq. (7.53), obtemos

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -0,2 & 0,05 \\ -0,3 & 1 & -0,2 \\ 0,1 & -0,3 & 1 \end{bmatrix} \begin{bmatrix} c_{-1} \\ c_0 \\ c_1 \end{bmatrix} \quad (7.54)$$

A solução desse conjunto de equações produz $c_{-1} = 0,210$, $c_0 = 1,13$ e $c_1 = 0,318$. Essa configuração de derivações assegura $p_o[0] = 1$ e $p_o[-1] = p_o[1] = 0$. O gráfico da saída ideal $p_o(t)$ é mostrado na Fig. 7.22c.

Vale notar que o equalizador determinado da Eq. (7.53) pode garantir somente a condição de ISI zero da Eq. (7.52). Em outras palavras, a ISI é zero somente para $k = 0, \pm 1, \dots, \pm N$. Na verdade, para outros valores de k , é muito comum que as amostras $p_o(kT_b) \neq 0$, indicando ISI residual. Por exemplo, consideremos o problema de equalizador do Exemplo 7.3. As amostras do pulso equalizado têm ISI zero para $k = -1, 0, 1$. Contudo, de

$$p_o[k] = \sum_{n=-N}^N c_n p_r[k-n]$$

podemos ver que os parâmetros do equalizado de três derivações com forçamento a zero resultarão em

$$\begin{aligned} p_o[-3] &= 0,010 & p_o[-2] &= 0,0145 & p_o[2] &= 0,0176 \\ p_o[3] &= 0,0318 & p_o[k] &= 0 & k &= 0, \pm 1, \pm 4, \dots \end{aligned}$$

Fica claro, portanto, que nem toda ISI foi removida, devido a essas quatro amostras não nulas do pulso de saída do equalizador. Na verdade, como temos apenas $2N + 1$ ($N = 1$ no Exemplo 7.3) parâmetros no equalizador, é impossível forçar $p_o[k] = 0$ para todo k , a menos que $N = \infty$. Isso significa que não poderemos projetar um equalizador prático com um número finito de derivações que alcance ISI zero ideal. Todavia, quando N é suficientemente grande, os valores de amostras não nulas residuais serão pequenos, indicando que a maior parte da ISI foi suprimida.

Método do Mínimo Erro Quadrático Médio

Na prática, uma abordagem alternativa consiste em minimizar a diferença quadrática média entre a resposta de saída do equalizador $p_o[k]$ e a resposta desejada com ISI zero. Isso é conhecido como método do mínimo erro quadrático médio (MMSE – *minimum mean square error*) para o projeto de equalizadores com filtros transversais. O método MMSE não tenta forçar que as amostras de pulsos sejam zero em $2N$ pontos, mas busca minimizar o erro quadrático médio considerando um conjunto de amostras de saída. O método envolve mais equações simultâneas. Assim, devemos determinar os valores de derivações do equalizador para minimizar o erro quadrático (médio) em uma janela maior $[-K, K]$:

$$\text{MSE} \triangleq \frac{1}{2K+1} \sum_{k=-K}^K (p_o[k] - \delta[k])^2$$

em que usamos uma função conhecida como delta de Kronecker

$$\delta[k] = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}$$

A solução desse problema de minimização é convenientemente representada em forma matricial como

$$\mathbf{c} = \mathbf{P}_r^\dagger \mathbf{p}_o$$

em que \mathbf{P}_r^\dagger representa a pseudoinversa de Moore-Penrose da matriz não quadrada \mathbf{P}_r de ordem $(2K+1) \times (2N+1)$. O projeto MMSE leva a um equalizador mais robusto para a redução de ISI.

Equalização Adaptativa e Outros Equalizadores Mais Genéricos

A estrutura do filtro equalizador descrita aqui tem a forma mais simples. Sistemas de comunicação digital práticos, muitas vezes, aplicam estruturas equalizadoras mais sofisticadas e algoritmos de equalização mais avançados.⁶ Devido às necessárias ferramentas probabilísticas, adiaremos a discussão desse tema especializado de equalização para o Capítulo 12.

7.5.2 Extração de Temporização

O sinal digital recebido deve ser amostrado em instantes precisos. Isso requer um sinal de relógio no receptor em sincronismo com o sinal do relógio no transmissor (**sincronização de símbolo ou de bit**), atrasado pela resposta do canal. Há três métodos genéricos de sincronização:

1. Derivação de um padrão primário ou secundário (por exemplo, transmissor e receptor escravizados à fonte mestra de temporização).
2. Transmissão de um sinal de sincronização separado (relógio piloto).
3. Autossincronização, em que a informação de temporização é extraída do próprio sinal recebido.

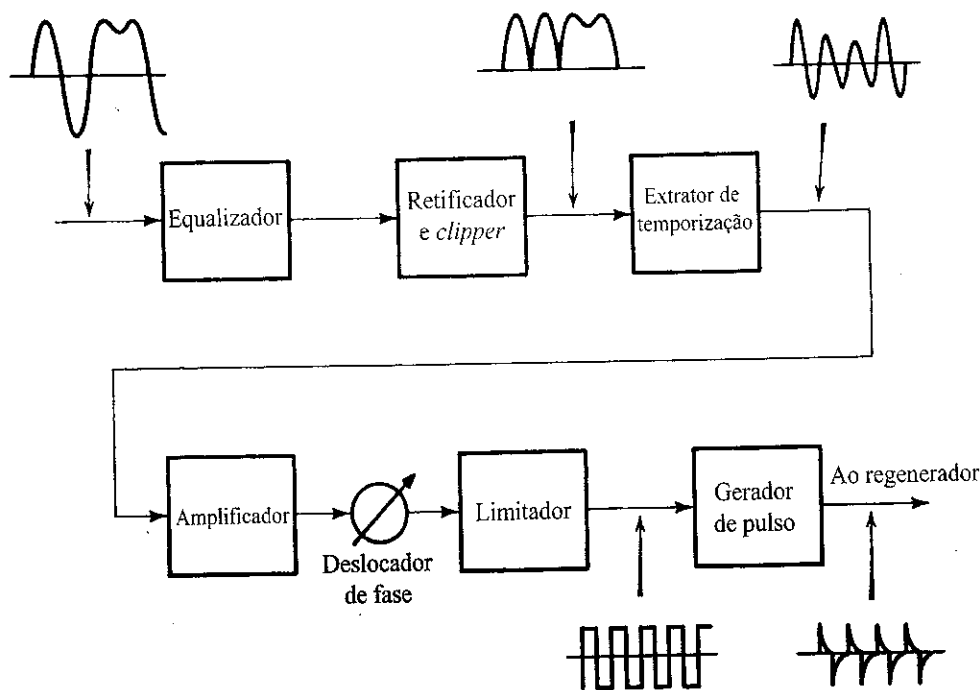
Devido ao alto custo, o primeiro método é adequado a grandes volumes de dados e sistemas de comunicação de alta velocidade. O segundo método, em que parte da capacidade do canal é usada para transmitir a informação de temporização, é adequado a situações em que a capacidade disponível é grande em comparação com a taxa de dados, e quando há disponibilidade de potência de transmissão. O terceiro é um método muito eficiente de extração de temporização ou recuperação de relógio, pois a temporização é extraída do próprio sinal de mensagem recebido. Um exemplo do método de autossincronização será discutido a seguir.

Já mostramos que um sinal digital, como o sinal on-off (Fig. 7.3a), contém uma componente discreta da frequência do relógio (Fig. 7.3c). Portanto, quando o sinal binário on-off é aplicado a um circuito ressonante sintonizado na frequência do relógio, o sinal de saída é o desejado sinal do relógio.

Nem todos os sinais binários contêm uma componente discreta da frequência [Eq. (7.21) ou Fig. 7.9]. Nesses casos, a extração da temporização pode ser possível com o uso de um *dispositivo não linear* para gerar um tom de frequência relacionado ao relógio de temporização. No caso bipolar, por exemplo, uma simples retificação converte um sinal bipolar em um sinal on-off, que pode ser prontamente utilizado para extrair a temporização.

Pequenos desvios aleatórios dos pulsos recebidos em relação à localização ideal (conhecidos como **incerteza temporal**) sempre estão presentes, mesmo nos mais sofisticados sistemas. Embora a fonte emita pulsos nos instantes corretos, operações subsequentes durante a transmissão (como deslocamento Doppler) tendem a forçar os pulsos a se desviarem de suas posições originais. O Q do circui-

Figura 7.23
Extração de
temporização.



to sintonizado usado para extração da temporização deve ser grande o bastante para prover adequada supressão da incerteza temporal, tornando-a suficientemente pequena para atender os requisitos de estabilidade. Durante os intervalos em que não há pulso na entrada, a oscilação continua, devido ao efeito de volante (*flywheel effect*) do circuito de alto Q . Não obstante, a saída do oscilador é sensível ao padrão de entrada; por exemplo, durante uma longa sequência de 1s, a amplitude de saída aumentará, ao passo que durante uma longa sequência de 0s, diminuirá. Isso introduz incerteza adicional ao sinal de temporização extraído.

O completo extrator de temporização e gerador de pulso de temporização para o caso bipolar são ilustrados na Fig. 7.23. A saída senoidal do oscilador (extração de temporização) é aplicada ao deslocador de fase, que ajusta a fase do sinal de temporização de modo que os pulsos de temporização ocorram nos pontos máximos. Esse método é usado para recuperar o relógio em cada um dos regeneradores em um sistema PCM. As incertezas introduzidas por sucessivos regeneradores se somam e, após um certo número de regeneradores, se torna necessário o uso de um regenerador com um sistema mais sofisticado de recuperação do relógio, como um sistema de *malha de captura de fase* (PLL - *phase locked loop*).

Incerteza temporal

Variações nas posições dos pulsos ou instantes de amostragem causam incerteza temporal (*timing jitter*). Isso advém de várias causas, algumas das quais dependem do padrão de pulsos sendo transmitido, enquanto outras não. As primeiras são cumulativas ao longo da cadeia de repetidores regenerativos, pois todos os repetidores são afetados da mesma forma. As outras formas de incerteza são aleatórias de regenerador para regenerador e, portanto, tende a haver um cancelamento parcial de efeitos mútuos ao longo de um enlace de grande distância. Formas aleatórias de incerteza temporal são causadas por ruído, interferências e desvio de sintonia de circuitos de relógio. Incerteza dependente de padrão resulta de desvios de relógio, conversão amplitude-fase em circuitos de relógio e ISI, que altera a posição do pico do sinal de entrada segundo o padrão. É possível mostrar que o valor rms da incerteza ao longo de uma cadeia de N repetidores aumenta com \sqrt{N} .

O acúmulo de incerteza em um enlace digital pode ser reduzido com o emprego de armazenagem elástica (*buffer*) e temporização da sequência de dígitos sob o controle de uma malha de captura de fase altamente estável. Em um longo enlace digital, se faz necessária a redução de incertezas a cada 300 km, para manter a incerteza máxima dentro de limites razoáveis.

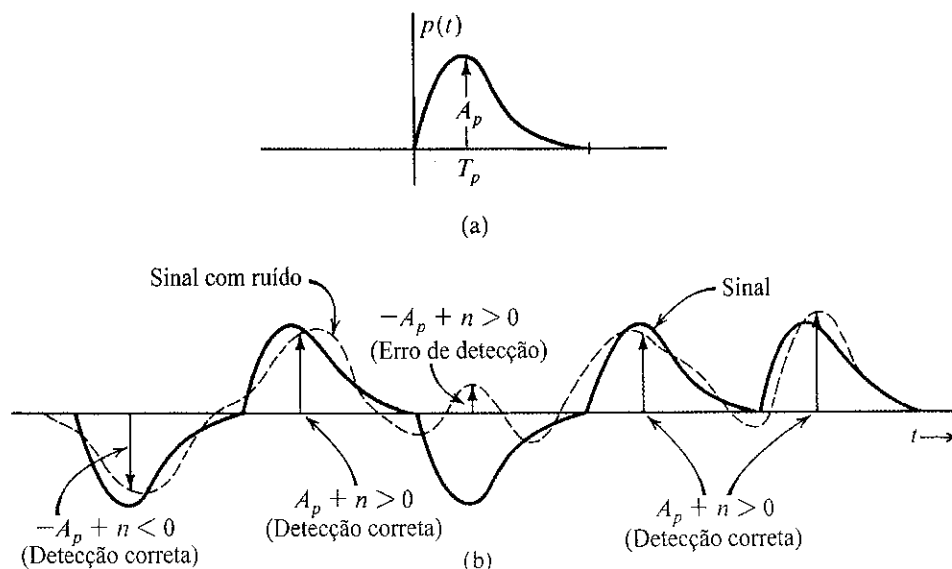
7.5.3 Detecção de Erro

Uma vez que a transmissão tenha passado pelo equalizador, a detecção pode ser feita no detector que amostra o sinal recebido com base no relógio fornecido pelo extrator de temporização. O sinal recebido no detector consiste no trem de pulsos equalizados mais um ruído de canal aleatório. O ruído pode causar erro na detecção de pulso. Consideremos, por exemplo, o caso da transmissão polar que usa um pulso básico $p(t)$ (Fig. 7.24a). Esse pulso tem uma amplitude de pico A_p . Um típico trem de pulsos recebidos é mostrado na Fig. 7.24b. Os pulsos são amostrados em seus valores de pico. Se o ruído estivesse ausente, a amostra do pulso positivo (correspondendo a 1) seria A_p e a do pulso negativo (correspondendo a 0), $-A_p$.* Devido ao ruído, os valores dessas amostras passariam a $\pm A_p + n$, em que n é a amplitude do ruído aleatório (Fig. 7.24b). Dada a simetria da situação, o limiar de detecção é zero; ou seja, se o valor da amostra de pulso for positivo, o dígito é detectado como 1; se o valor da amostra for negativo, o dígito é detectado como 0.

A decisão do detector em declarar 1 ou 0 poderia ser tomada prontamente a partir da amostra de pulso, não fosse o valor do ruído n aleatório, o que significa que seu valor exato é imprevisível. O valor do ruído pode ser grande ou pequeno, assim como pode ser negativo ou positivo. É possível que 1 seja transmitido e que, no instante de amostragem, n tenha um valor negativo grande. Isso tornaria o valor da amostra $A_p + n$ pequeno ou, até mesmo, negativo. Se, por outro lado, um 0 for transmitido e

*Assumindo ISI zero.

Figura 7.24
Probabilidade
de erro na
detecção por
limiar.



n tiver um valor positivo grande no instante de amostragem, o valor da amostra $-A_p + n$ pode ser positivo e o dígito, detectado como 1. Isso fica claro na Fig. 7.24b.

O desempenho de um sistema de comunicação digital é, em geral, especificado pelo número médio de erros de detecção. Por exemplo, se dois telefones celulares (receptores) no mesmo local tentam detectar a mesma transmissão da estação radiobase, o telefone celular com menor erros de detecção será o melhor receptor. É provável que esse telefone tenha menor perda de chamadas e menos dificuldade para receber uma fala clara. Contudo, como o ruído é aleatório, algumas vezes um telefone celular pode ser melhor; outras, o segundo telefone celular pode ter menos erros. A real medida do desempenho do receptor é a razão média entre o número de erros e o número total de dados transmitidos. Assim, a comparação razoável de desempenho é a probabilidade de ocorrência de erro de detecção, ou **probabilidade de erro de detecção**.

Como a análise precisa e o cálculo dessa probabilidade de erro requer conhecimento e ferramentas da teoria da probabilidade, adiaremos a análise de erros até introdução de probabilidade no Capítulo 8. Mais adiante, no Capítulo 10, discutiremos em detalhe a análise de probabilidade de erro de diferentes sistemas de comunicação digital, considerando diversos modelos de ruído, assim como projetos de sistemas para combater ruídos variados. Por exemplo, o ruído gaussiano pode caracterizar o ruído aleatório de canal proveniente de efeitos térmicos e interferência (*cross talk*) intrassistema. Detectores ótimos podem ser projetados para minimizar a probabilidade de erro devido a ruído gaussiano. Contudo, transientes de comutação, descargas elétricas, chaveamento de carga em linhas de alta tensão, e outros eventos singulares causam pulsos de ruído de alta intensidade e curta duração, que contaminam os pares de cabo que transportam sinais digitais. Esses pulsos, denominados coletivamente **ruído impulsional**, são de difícil combate e constituem a prevalecente fonte de erros em ambientes externos aos sistemas digitais. Erros, portanto, praticamente jamais ocorrem de modo isolado, mas em rajadas de até centenas de cada vez. Para corrigir uma rajada de erros, lançamos mão de **códigos especiais para corrigir erros em rajadas**, descritos no Capítulo 14.

7.6 DIAGRAMAS DE OLHO: UMA FERRAMENTA ÚTIL

Na seção anterior, estudamos o efeito de ruído e ISI de canal na detecção de transmissões digitais. Descrevemos, também, o projeto de equalizadores para compensar distorção de canal e explicamos o processo de extração de temporização. A seguir, apresentamos uma prática ferramenta de engenharia

conhecida como **diagrama de olho**. O diagrama de olho é de fácil geração, e aplicado por engenheiros a sinais recebidos, pois facilita o exame visual da gravidade da ISI, da precisão da extração de temporização, da imunidade ao ruído e de outros fatores importantes.

Para gerar um diagrama de olho, precisamos apenas de um osciloscópio básico. Dado um sinal em banda base na saída do canal

$$y(t) = \sum a_k p(t - kT_b)$$

o mesmo pode ser aplicado à entrada vertical do osciloscópio. A base de tempo do osciloscópio é disparada à mesma taxa $1/T_b$ de chegada de pulsos, produzindo uma varredura que dura exatamente T_b segundos, a duração de um símbolo de dado transmitido a_k . O osciloscópio mostra a superposição de diversos traços de comprimento T_b da saída do canal $y(t)$. O que aparece no osciloscópio são simplesmente cortes do sinal de entrada (entrada vertical), feitos a cada T_b e superpostos uns aos outros. O padrão resultante no osciloscópio tem a aparência de um olho humano, daí a denominação diagrama de olho. De modo mais genérico, também podemos aplicar uma varredura temporal com duração de m intervalos de símbolos ou mT_b . O padrão no osciloscópio passa, então, a ser composto por cortes do sinal de entrada (entrada vertical), feitos a cada mT_b e superpostos uns aos outros. O osciloscópio exibirá um diagrama de olho com largura mT_b e tem o formato de m olhos em uma linha horizontal.

A seguir, apresentamos um exemplo. Consideremos a transmissão de um sinal binário por pulsos polares NRZ (Fig. 7.25a). Os correspondentes diagramas de olho são mostrados na Fig. 7.25b para bases de dados T_b e $2T_b$, respectivamente. Nesse exemplo, o canal tem largura de banda infinita para passar o pulso NRZ e não há distorção de canal. Assim, obtemos diagramas de olho com olho(s) totalmente aberto(s). Podemos, ainda, considerar uma saída de canal com o mesmo código de linha polar, mas um formato diferente de pulsos (RZ), como mostrado na Fig. 7.25c. Os resultantes diagramas de olho são mostrados na Fig. 7.25d. Nesse caso, o olho está bem aberto apenas no ponto médio do intervalo de duração de cada pulso. Com extração de temporização adequada, para melhor imunidade ao ruído no ponto de decisão (Seção 7.5.3), o receptor deve amostrar o sinal recebido exatamente no ponto médio, em que o olho está totalmente aberto. O ponto médio do olho representa o melhor instante de amostragem para cada pulso, pois a amplitude é máxima, sem interferência de pulsos vizinhos (ISI zero).

Figura 7.25
Diagrama de olho.

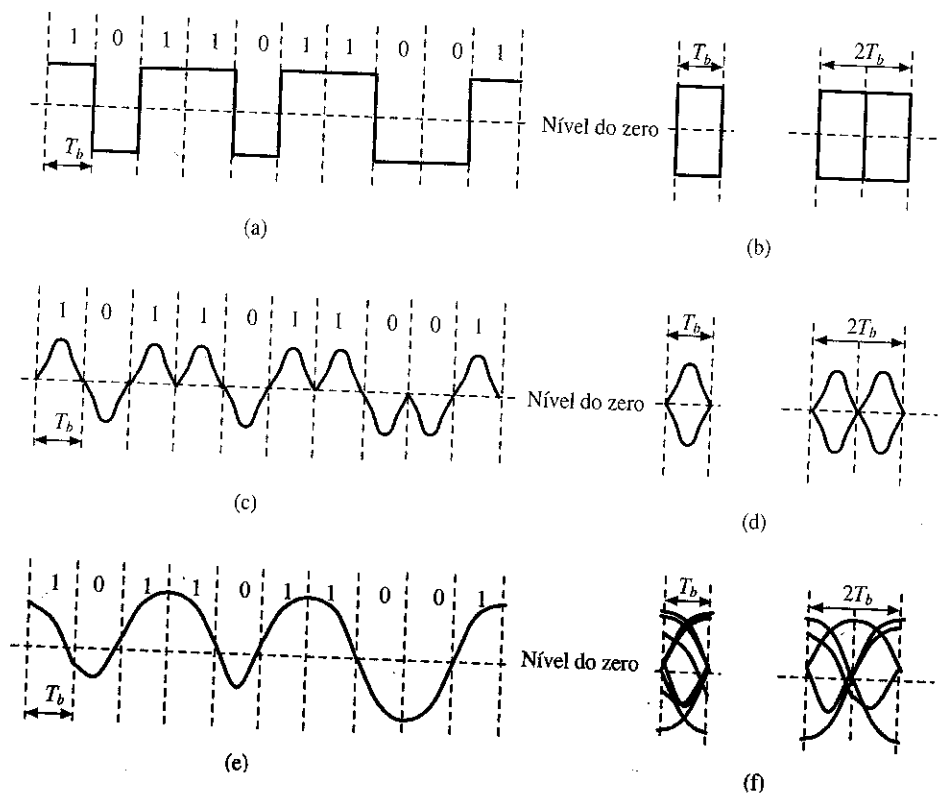
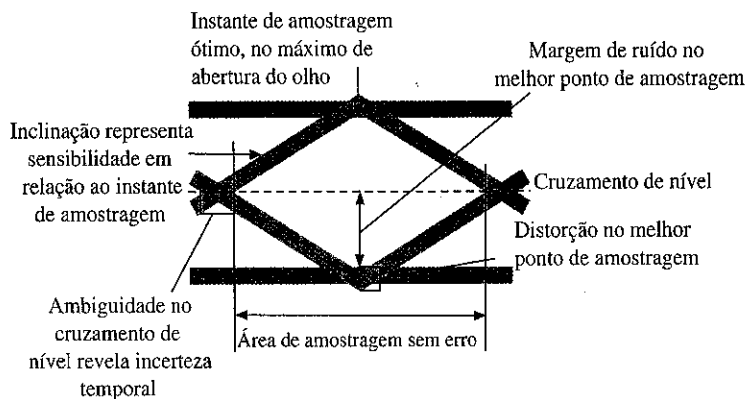


Figura 7.26
Leitura de um
diagrama de
olho.



Consideremos, agora, um canal com distorção ou com largura de banda finita, ou ambos. Após passagem pelo canal não ideal, o sinal polar NRZ da Fig. 7.25a se torna a forma da onda na Fig. 7.25e. Os pulsos de sinal recebidos não são mais retangulares, mas arredondados, distorcidos e alargados. Os diagramas de olho não são mais totalmente abertos, como mostrado na Fig. 7.25f. Nesse caso, a ISI não é zero. Portanto, em cada traço, os valores do pulso nos respectivos instantes de amostragem se desviarão dos valores ideais por um grau variável, causando um borrão e resultando em um padrão de olho parcialmente fechado.

Na presença de ruído de canal, o olho tenderá a se fechar em todos os casos. Ruído fraco causa um fechamento proporcionalmente pequeno do olho. O limiar de decisão, em relação ao símbolo (1 ou 0) transmitido, é o ponto médio do olho.* Observemos que, para ISI zero, o sistema pode tolerar ruído de até metade da abertura vertical do olho. Qualquer valor de ruído maior que esse limite pode causar erro de decisão caso seu sinal seja oposto ao do símbolo de dado. Como a ISI reduz a abertura do olho, também reduz a tolerância ao ruído. O diagrama de olho é usado, ainda, para determinar especificação ótima de derivações do equalizador. Derivações são ajustadas para obter o máximo de aberturas vertical e horizontal de olho.

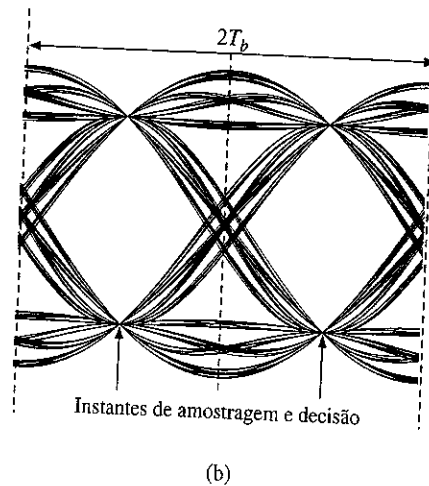
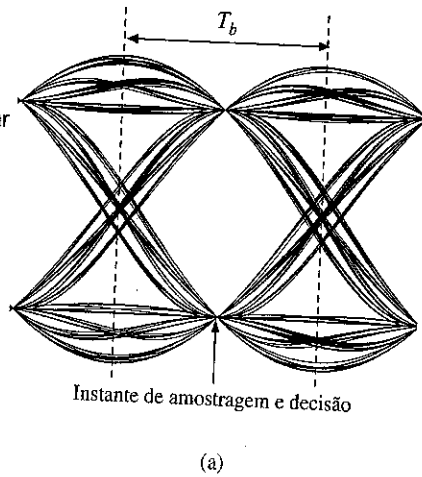
O diagrama de olho é uma ferramenta muito eficaz para análise de sinais em experimentos em tempo real. Não apenas é de operação simples, mas fornece informação rica e importante sobre a qualidade e suscetibilidade do sinal digital recebido. Do típico diagrama de olho ilustrado na Fig. 7.26, podemos extrair diversas medidas importantes relativas à qualidade do sinal:

- *Ponto de máxima abertura.* O grau de abertura no instante de amostragem e decisão indica a quantidade de ruído que o detector pode tolerar sem cometer um erro. Essa quantidade é conhecida como *margem de ruído*. O instante de máxima abertura do olho indica o instante ótimo de amostragem e decisão.
- *Sensibilidade à incerteza temporal.* A largura do olho indica o intervalo de tempo em que uma decisão correta pode ser feita, sendo desejável ter um olho com máxima abertura horizontal. Se o instante de tomada de decisão se desviar do instante em que o olho tem máxima abertura vertical, a margem de tolerância ao ruído fica reduzida. Isso causa maior probabilidade de erro na detecção de pulsos. A inclinação do olho mostra quão rápido a tolerância ao ruído é reduzida e, em consequência, a sensibilidade da tolerância da decisão ao ruído em relação a variações no instante de amostragem. A inclinação demonstra efeitos da incerteza temporal.
- *Incerteza (temporal) no cruzamento de nível.* Tipicamente, receptores práticos extraem informação de temporização sobre a taxa de pulso e sobre o relógio de amostragem a partir do cruzamento de nível (zero) da forma de onda do sinal recebido. A variação do cruzamento de nível pode ser vista da largura das esquinas do olho. Essa medida fornece informação sobre a incerteza temporal a que o receptor estará sujeito.

*Isso se aplica ao caso de decisão em dois níveis [por exemplo, quando $p(t)$ e $-p(t)$ são usados para 1 e 0, respectivamente]. No caso de decisão em três níveis (por exemplo, sinalização bipolar), haverá dois limiares.

Figura 7.27

Diagramas de olho de um sistema de sinalização polar usando um pulso cosseno levantado, com fator de decaimento $r = 0,5$; (a) largura de $2T_b$, 2 períodos de símbolo, e deslocamento temporal $T_b/2$; (b) sem deslocamento temporal.



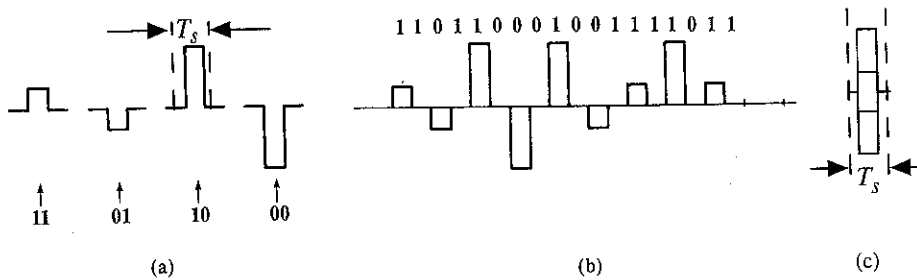
Por fim, vejamos um exemplo prático de diagrama de olho para uma forma de onda de sinalização polar. Para esse caso, selecionemos um pulso com decaimento cosseno que satisfaça o primeiro critério de Nyquist para ISI zero. Escolhamos o fator de decaimento como $r = 0,5$. O diagrama de olho é mostrado na Fig. 7.27, com uma base de tempo $2T_b$. Na verdade, para um mesmo sinal, o diagrama de olho pode adquirir aparências distintas, dependendo dos valores dos deslocamentos temporais aplicados (pontos iniciais). A Fig. 7.27a ilustra o diagrama de olho dessa forma de onda de sinalização polar para um deslocamento temporal $T_b/2$, enquanto a Fig. 7.27b mostra o diagrama de olho sem a aplicação de deslocamento temporal. Uma comparação dos dois diagramas deixa claro que os mesmos guardam uma simples relação de deslocamento circular horizontal. Observando a máxima abertura de olho, vemos que esse sinal em banda base tem ISI zero, confirmando a principal característica do pulso cosseno levantado. Como o primeiro critério de Nyquist não impõe nenhuma condição ao cruzamento do zero do pulso, o diagrama de olho indica a probabilidade de ocorrência de incerteza temporal.

7.7 PAM: SINALIZAÇÃO M-ÁRIA EM BANDA BASE PARA TAXA DE DADOS MAIS ELEVADA

Independentemente do código de linha usado, modulações binárias em banda base têm algo em comum: todas transmitem um bit de informação no intervalo de T_b segundos, ou seja, a uma taxa de $1/T_b$ bits por segundo. Caso o transmissor deseje enviar seus bits a uma taxa muito mais elevada, T_b deve ser reduzido. Por exemplo, para aumentar a taxa de bits por um fator M , T_b deve ser reduzido pelo mesmo fator; contudo, um alto preço será pago em termos de largura de banda. Como demonstramos na Fig. 7.9, a largura de banda de uma modulação em banda base é proporcional à taxa de pulsos $1/T_b$. A redução de T_b por um fator M certamente aumentará a necessária largura de banda de canal pelo mesmo fator M . Afortunadamente, a redução de T_b não é a única forma de aumentar a taxa de dados. Uma solução muito prática consiste em permitir que cada pulso transporte múltiplos bits. A seguir, explicaremos esse conceito.

Para que cada símbolo de transmissão transporte mais bits no intervalo de tempo T_b , devem existir mais de dois símbolos. Aumentando o número de símbolos para M , asseguramos que a informação transmitida por símbolo também aumenta com M . Por exemplo, quando $M = 4$ (sinalização 4-ária ou quaternária), temos quatro símbolos básicos, ou pulsos, disponíveis para comunicação (Fig. 7.28a). Uma sequência de dois dígitos binários pode ser transmitida por apenas um símbolo 4-ário,

Figura 7.28
Sinalização PAM 4-ária: (a) quatro símbolos RZ; (b) transmissão em banda base; (c) diagrama de olho RZ 4-ário.



pois uma sequência de dois bits pode formar apenas quatro possíveis sequências (11, 10, 01 e 00). Como dispomos de quatro símbolos distintos, podemos alocar um deles a cada uma dessas combinações (Fig. 7.28a). Cada símbolo ocupa um intervalo de tempo T_s . Um exemplo de sinalização para uma sequência curta é dada na Fig. 7.28b e o diagrama de olho 4-ário é mostrado na Fig. 7.28c.

Essa sinalização nos permite transmitir cada par de bits por um pulso 4-ário (Fig. 7.28b). Logo, para transmitir n bits, precisamos somente de $(n/2)$ pulsos 4-ários. Isso significa que um símbolo 4-ário pode transmitir a informação de dois dígitos binários. Além disso, como três bits formam $2 \times 2 \times 2 = 8$ combinações, um grupo de três bits pode ser transmitido por um símbolo 8-ário. De modo similar, um grupo de quatro bits podem ser transmitidos por um símbolo 16-ário. Em geral, a informação I_M transmitida por um símbolo M -ário é

$$I_M = \log_2 M \text{ bits} \quad (7.55)$$

Isso significa que podemos aumentar a taxa de transmissão de informação aumentando o valor de M .

Essa sinalização M -ária especial é conhecida como **modulação por amplitude de pulso (PAM – pulse amplitude modulation)**, pois a informação de dados é transportada pela variação da amplitude do pulso. Vale observar que a modulação por amplitude de pulso é apenas uma das várias possibilidades de sinalização M -ária. Há um número infinito de possibilidades. Todavia, apenas algumas poucas são, de fato, eficazes no combate ao ruído e na redução de largura de banda e consumo de potência. Uma discussão mais detalhada de outros esquemas de sinalização M -ária será apresentada um pouco mais adiante, na Seção 7.9.

Na maioria dos projetos de sistemas, sempre há um preço a ser pago por ganho possível. O preço pago por PAM para aumentar a taxa de dados é o consumo de potência. À medida que M aumenta, a potência transmitida também aumenta com M . Isso se deve ao fato de que, para termos imunidade ao ruído, a mínima separação entre amplitudes de pulsos deve ser comparável à do pulso binário. Portanto, amplitudes de pulso aumentam com M (Fig. 7.28). Podemos mostrar que a potência transmitida aumenta com M^2 (Exercício 7.7-5). Assim, para aumentar a taxa de comunicação por um fator $\log_2 M$, a potência necessária aumenta com M^2 . Como a largura de banda de transmissão depende somente da taxa de pulsos, e não das amplitudes dos pulsos, a largura de banda independe de M . Para ilustrar esse ponto, usemos o exemplo de análise de PSD a seguir.

Exemplo 7.4 Determinemos a PSD da sinalização quaternária (4-ária) em banda base dada na Fig. 7.28, considerando que os bits de mensagem 1 e 0 têm igual probabilidade de ocorrência.

O código de linha 4-ário tem quatro símbolos distintos, correspondentes às quatro combinações de dois bits de mensagem. Um possível mapeamento é

$$a_k = \begin{cases} -3 & \text{bits de mensagem 00} \\ -1 & \text{bits de mensagem 01} \\ +1 & \text{bits de mensagem 10} \\ +3 & \text{bits de mensagem 11} \end{cases} \quad (7.56)$$

Portanto, todos os quatro valores de a_k têm igual probabilidade de ocorrência, cada um com uma chance de 1 em 4. Recordemos que

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k^2$$

No somatório, 1/4 de a_k será ± 1 e ± 3 . Logo,

$$R_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{4}(-3)^2 + \frac{N}{4}(-1)^2 + \frac{N}{4}(1)^2 + \frac{N}{4}(3)^2 \right] = 5$$

Para $n > 0$, devemos determinar

$$R_n = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_k a_k a_{k+n}$$

Para calcular esse valor médio, construamos uma tabela com todos os possíveis valores do produto $a_k a_{k+n}$:

Possíveis valores de $a_k a_{k+n}$

$a_k \backslash a_{k+n}$	-3	1	+1	+3
-3	9	3	-3	-9
-1	3	1	-1	-3
+1	-3	-1	1	3
+3	-9	-3	3	9

Da listagem anterior de todos os possíveis produtos $a_k a_{k+n}$, vemos que cada produto no somatório $a_k a_{k+n}$ pode assumir qualquer um dos seguintes seis valores: $\pm 1, \pm 3, \pm 9$. Observamos que $(\pm 1, \pm 9)$ têm igual probabilidade de ocorrência (1 em 8). Por sua vez, ± 3 também têm igual probabilidade de ocorrência (1 em 4). Assim, obtemos

$$R_n = \lim_{N \rightarrow \infty} \frac{1}{N} \left[\frac{N}{8}(-9) + \frac{N}{8}(+9) + \frac{N}{8}(-1) + \frac{N}{8}(+1) + \frac{N}{4}(-3) + \frac{N}{4}(+3) \right] = 0$$

Logo,

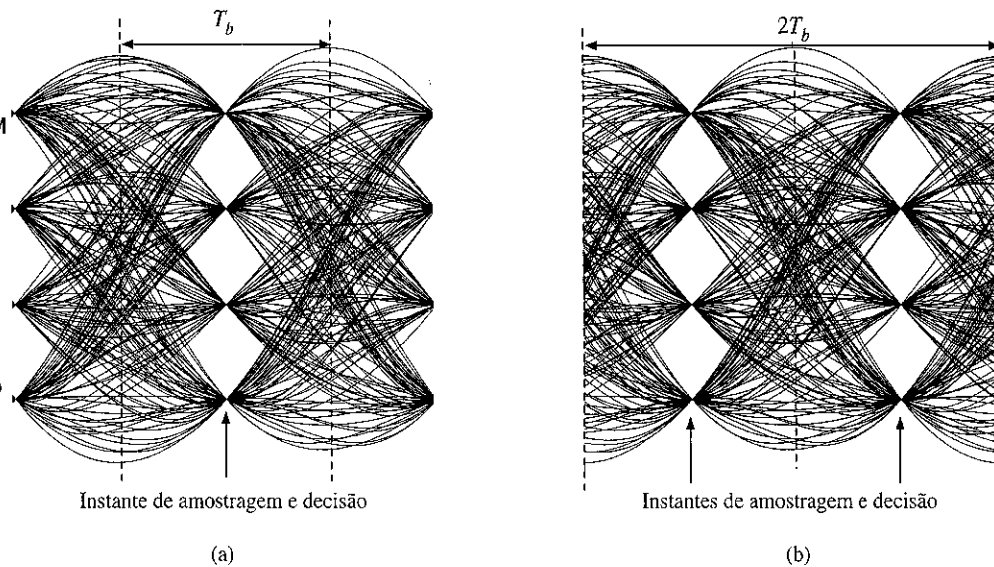
$$S_x(f) = \frac{5}{T_s} \implies S_y(f) = \frac{5}{T_s} |P(f)|^2$$

Portanto, o código de linha *M*-ário gera a mesma forma de PSD que a sinalização polar binária. A única diferença é que utiliza 5 vezes a potência de sinal original.

Embora a maioria da rede terrestre de telefonia digital empregue codificação binária, a porção da malha de assinante da rede digital de serviços integrados (ISDN) usa o código quaternário, 2B1Q, semelhante à Fig. 7.28a. São usados pulsos NRZ para transmitir 160 kbit/s de dados a uma taxa de **baud** (taxa de pulsos ou de símbolos) de 80 kbit/s. Dos vários códigos de linha examinados pelo comitê de padronização de ANSI, 2B1Q é o que provê maior redução da taxa de baud na planta de cabos locais, ambiente ruidoso e sujeito a interferências.

Formatação de Pulso e Diagramas de Olho em PAM: Neste caso, podemos usar pulsos que satisfaçam o critério de Nyquist, pois tais pulsos têm zero ISI nos pontos de amostragem e, portanto, suas amplitudes podem ser detectadas corretamente se forem amostradas nos centros dos pulsos. Também podemos usar ISI controlada (sinalização de resposta parcial) para sinalização *M*-ária.⁸

Figura 7.29
Diagramas de olho de um sistema de sinalização PAM 4-ária usando pulso cosseno levantado com fator de decaimento $r = 0,5$: (a) com dois períodos de símbolo $2T_b$ e deslocamento temporal de $T_b/2$; (b) sem deslocamento temporal.



Diagramas de olho também podem ser gerados para PAM M -ária usando o mesmo método empregado em modulações binárias. Devido à sinalização em múltiplos níveis, o diagrama de olho deve ter M níveis nos instantes de amostragem ótima, mesmo quando a ISI for zero. Aqui, geramos o exemplo prático de diagrama de olho para um sinal PAM de quatro níveis que usa o mesmo pulso cosseno levantado com fator de decaimento $r = 0,5$, como no diagrama de olho na Fig. 7.27. Os correspondentes diagramas de olho com deslocamentos temporais de $T_b/2$ e 0 são dados nas Fig. 7.29 a e b, respectivamente. Mais uma vez, nenhuma ISI é observada nos instantes de amostragem. Os diagramas de olho mostram claramente quatro valores de sinais igualmente separados, sem ISI nos pontos de amostragem ótima.

7.8 SISTEMAS DIGITAIS COM PORTADORA

Até aqui, discutimos sistemas digitais em banda base, em que sinais são transmitidos diretamente, sem nenhum deslocamento em frequência. Como sinais em banda base têm muita potência nas frequências baixas, são adequados à transmissão por um par de fios ou por um cabo coaxial. Uma boa parte da comunicação moderna é conduzida dessa forma. Contudo, sinais em banda base não podem ser transmitidos em enlaces de rádio ou de satélites, pois isso exigiria antenas demasiadamente grandes, e impraticáveis, para irradiar eficientemente o espectro de baixa frequência desses sinais. Portanto, para essas aplicações, o espectro do sinal deve ser deslocado a uma faixa de frequências altas. Um deslocamento do espectro para frequências mais altas também se faz necessário para transmitir várias mensagens simultaneamente, com compartilhamento da grande largura de banda de um meio de transmissão. Como vimos no Capítulo 4, o espectro de um sinal pode ser deslocado para uma frequência mais alta usando o sinal em banda base para modular uma senoide (portadora) de alta frequência.

Na transmissão e recepção de sinais digitais com portadora, precisamos de um modulador e de um demodulador para transmitir e receber dados. Para comunicação em duas direções (duplex), esses dois dispositivos, **modulador** e **demodulador**, são comumente empacotados em uma unidade que recebe o nome de **modem**.

7.8.1 Modulações Binárias Básicas com Portadora

Há duas formas básicas de modulação de portadora: modulação em amplitude e modulação em ângulo. Na modulação em amplitude, a amplitude da portadora é variada em proporção ao sinal modulante (ou seja, o sinal em banda base). Isso é ilustrado na Fig. 7.30. Uma portadora não modulada $\cos \omega_c t$ é

Figura 7.30

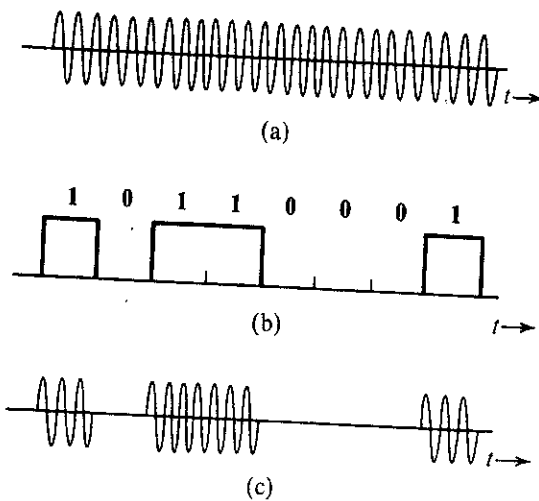
(a) Portadora

 $\cos \omega_c t$.

(b) Sinal

modulante $m(t)$.

(c) ASK: sinal

modulado $m(t)$ $\cos \omega_c t$.

mostrada na Fig. 7.30a. O sinal on-off em banda base $m(t)$ (sinal modulante) é mostrado na Fig. 7.30b. Segundo a Eq. (7.1), o sinal modulante pode ser escrito como

$$m(t) = \sum a_k p(t - kT_b), \quad \text{em que} \quad p(t) = \Pi\left(\frac{t - T_b/2}{T_b}\right)$$

O código de linha $a_k = 0, 1$ é on-off. Quando a amplitude da portadora é variada em proporção a $m(t)$, podemos escrever o sinal modulado com portadora como

$$\varphi_{\text{ASK}}(t) = m(t) \cos \omega_c t \quad (7.57)$$

que é mostrado na Fig. 7.30c. Reparemos que o sinal modulado ainda é um sinal on-off. Esse esquema de modulação para transmissão de dados binários é conhecido como **chaveamento on-off (OOK)** – *on-off keying* ou **modulação por chaveamento de amplitude (ASK)** – *amplitude shift keying*.

O sinal em banda base $m(t)$ pode, obviamente, usar um pulso $p(t)$ diferente do retangular mostrado no exemplo da Fig. 7.30. Isso gerará um sinal ASK que não tem amplitude constante durante a transmissão de 1 ($a_k = 1$).

Se o sinal em banda base $m(t)$ fosse polar (Fig. 7.31a), o correspondente sinal modulado $m(t) \cos \omega_c t$ teria a aparência ilustrada na Fig. 7.31b. Nesse caso, se $p(t)$ for o pulso básico, transmitimos 1 por um pulso $p(t) \cos \omega_c t$ e 0, por $-p(t) \cos \omega_c t = p(t) \cos(\omega_c t + \pi)$. Os dois pulsos, portanto, estão defasados de π radianos. A informação reside na fase ou no sinal do pulso. Por essa razão, esse esquema é conhecido como **chaveamento por deslocamento de fase (PSK)** – *phase shift keying*. Reparemos que a transmissão ainda é polar. Na verdade, assim como ASK, o sinal PSK modulado com portadora tem a mesma forma

$$\varphi_{\text{PSK}}(t) = m(t) \cos \omega_c t \quad m(t) = \sum a_k p(t - kT_b) \quad (7.58)$$

com a diferença que o código de linha é polar, $a_k = \pm 1$.

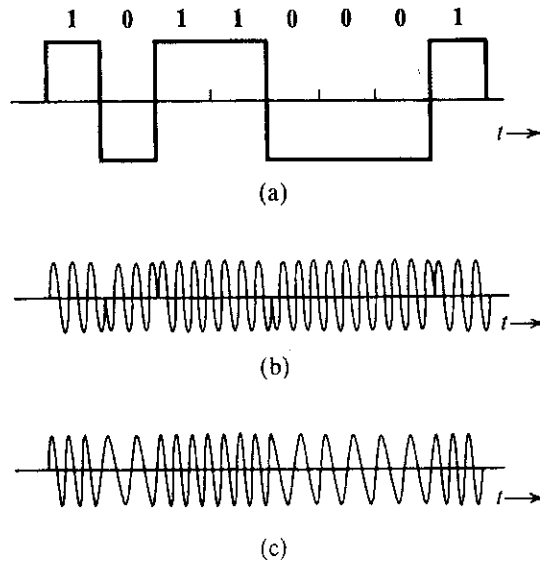
Quando dados são transmitidos por meio da variação da frequência, temos o caso de **chaveamento por deslocamento de frequência (FSK)** – *frequency shift keying*, como ilustrado na Fig. 7.31c. Um 0 é transmitido por um pulso de frequência ω_{c0} e um 1, por um pulso de frequência ω_{c1} . A informação sobre o dado transmitido reside na frequência portadora. O sinal FSK pode ser visto como a soma de dois sinais ASK entrelaçados, um com frequência portadora ω_{c0} e outro com frequência portadora ω_{c1} . Podemos usar a expressão de ASK binário da Eq. (7.57) e escrever o sinal FSK como

$$\varphi_{\text{FSK}}(t) = \sum a_k p(t - kT_b) \cos \omega_{c1} t + \sum (1 - a_k) p(t - kT_b) \cos \omega_{c0} t$$

em que $a_k = 0, 1$ é on-off. Assim, o sinal FSK é uma superposição de dois sinais AM com diferentes frequências portadoras e amplitudes diferentes, mas complementares.

Figura 7.31

(a) Sinal modulante $m(t)$.
 (b) PSK: sinal modulado $m(t) \cos \omega_c t$.
 (c) FSK: sinal modulado.



Na prática, a ASK como um esquema on-off é comumente usado em sistemas de comunicação por fibra óptica, na forma de modulação da intensidade do laser. A PSK é comumente utilizado em sistemas de comunicação por satélite e também foi usado nos primeiros modems telefônicos (2400 e 4800 bit/s). Quanto à FSK, em 1962, a AT&T desenvolveu um dos primeiros modems para linha telefônica, chamado 103A, e usava a FSK para transmitir 300 bit/s em duas frequências, 1070 e 1270 Hz; a recepção era FSK em 2025 e 2225 Hz.

7.8.2 PSD de Modulação Digital com Portadora

Acabamos de mostrar que os esquemas de modulação binária com portadora ASK, PSK e FSK podem ser escritos na forma $m(t) \cos \omega_c t$. Para determinar a PSD dos sinais ASK, PSK e FSK, é conveniente que, primeiro, determinemos a relação entre a PSD de $m(t)$ e a PSD do sinal modulado

$$\varphi(t) = m(t) \cos \omega_c t$$

Recordemos, da Eq. (3.80), que a PSD de $\varphi(t)$ é

$$S_\varphi(f) = \lim_{T \rightarrow \infty} \frac{|\Psi_T(f)|^2}{T}$$

em que Ψ_T é a transformada de Fourier do sinal truncado

$$\begin{aligned} \varphi_T(t) &= \varphi(t)[u(t + T/2) - u(t - T/2)] \\ &= m(t)[u(t + T/2) - u(t - T/2)] \cos \omega_c t \\ &= m_T(t) \cos \omega_c t \end{aligned} \quad (7.59)$$

Aqui, $m_T(t)$ é o sinal em banda base truncado, com transformada de Fourier $M_T(f)$. Aplicando a propriedade de translação em frequência [Eq. (3.36)], temos

$$\Psi_T(f) = \frac{1}{2} [M_T(f - f_c) + M_T(f + f_c)]$$

Por conseguinte, a PSD do sinal modulado com portadora $\varphi(t)$ é

$$S_\varphi(f) = \lim_{T \rightarrow \infty} \frac{1}{4} \frac{|M_T(f + f_c) + M_T(f - f_c)|^2}{T}$$

Como $M(f)$ é um sinal em banda base, $M_T(f + f_c)$ e $M_T(f - f_c)$ têm sobreposição zero à medida que $T \rightarrow \infty$, desde que f_c seja maior que a largura de banda de $M(f)$. Portanto, concluímos que

$$\begin{aligned} S_\phi(f) &= \lim_{T \rightarrow \infty} \frac{1}{4} \left[\frac{|M_T(f + f_c)|^2}{T} + \frac{|M_T(f - f_c)|^2}{T} \right] \\ &= \frac{1}{4} S_M(f + f_c) + \frac{1}{4} S_M(f - f_c) \end{aligned} \quad (7.60)$$

Em outras palavras, para uma frequência portadora escolhida de modo apropriado, a modulação causa um deslocamento na PSD do sinal em banda base.

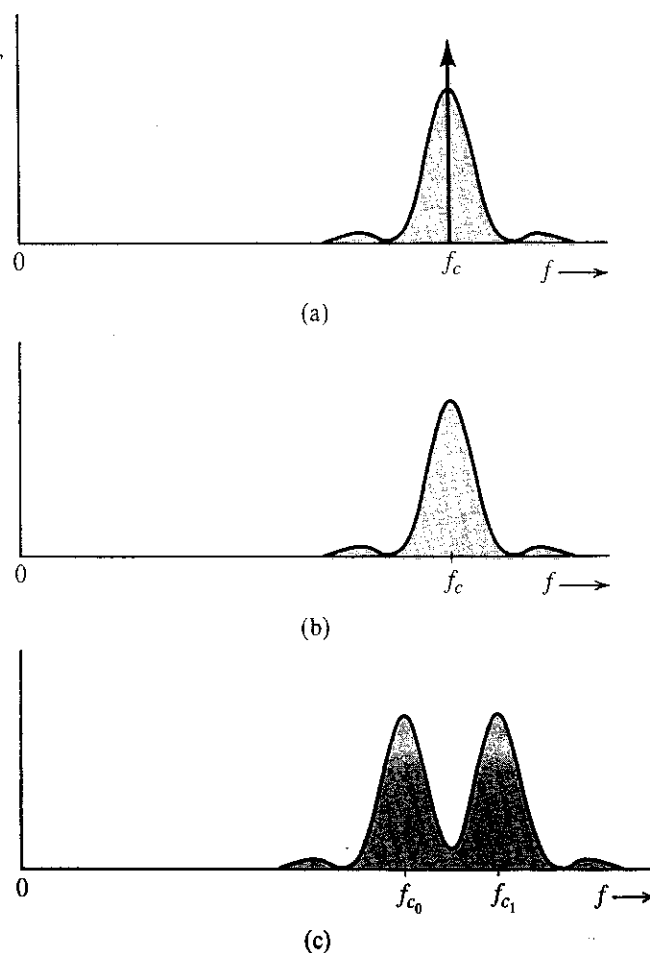
O sinal ASK na Fig. 7.30c se encaixa nesse modelo, com $m(t)$ sendo um sinal on-off (usando um pulso NRZ ou de largura completa). Portanto, a PSD do sinal ASK é igual à de um sinal on-off (Fig. 7.4b) deslocada para $\pm f_c$, como mostrado na Fig. 7.32a. Recordemos que, usando um pulso retangular de largura completa $p(t)$,

$$P\left(\frac{n}{T_b}\right) = 0 \quad n = \pm 1, \pm 2, \dots$$

Nesse caso, a PSD do sinal on-off na Fig. 7.30b não tem componentes discretas, exceto em dc. Em consequência, o espectro ASK tem componentes discretas apenas em ω_c .

O sinal PSK também se encaixa nessa descrição de modulação, em que $m(t)$ é um sinal polar que usa um pulso NRZ de largura completa. Portanto, a PSD de um sinal PSK é igual à do sinal polar em banda base deslocada para $\pm \omega_c$, como mostrado na Fig. 7.32b. Reparemos que essa PSD tem a mesma forma (com diferente fator de escala) que a PSD do sinal ASK, sem suas componentes discretas.

Figura 7.32
PSD de (a) ASK,
(b) PSK e (c)
FSK.



Por fim, mostramos que o sinal FSK pode ser visto como a soma de dois sinais ASK entrelaçados usando o pulso de largura completa. Portanto, o espectro FSK é a soma de dois espectros ASK nas frequências ω_{c_0} e ω_{c_1} , como mostrado na Fig. 7.32c. Podemos mostrar que, com escolha adequada de ω_{c_0} e ω_{c_1} , e com manutenção da continuidade de fase durante a translação em frequência, as componentes discretas podem ser eliminadas em ω_{c_0} e ω_{c_1} . Com isso, nenhuma componente discreta aparece nesse espectro. É importante ressaltar que a largura de banda de FSK é maior que as de ASK ou PSK.

Como observado anteriormente, a sinalização polar é o esquema mais eficiente do ponto de vista de potência. PSK, por ser polar, requer 3 dB de potência menor que ASK (ou FSK), para uma mesma imunidade ao ruído, ou seja, para uma mesma probabilidade de erro na detecção de pulsos.

É claro que podemos também modular o esquema bipolar ou qualquer outro discutido anteriormente. Reparemos que, na Fig. 7.30 e na Fig. 7.31, usamos pulso retangular NRZ apenas para facilitar a ilustração. Na prática, pulsos em banda base podem ser formatados especificamente para eliminar a ISI.

7.8.3 Relações entre Modulações Analógica e Digital com Portadora

Existe uma relação natural e clara entre ASK e AM, pois a mensagem de informação é diretamente refletida na variação da amplitude do sinal modulado. Devido à sua amplitude não negativa, a ASK é, em essência, um sinal AM com índice de modulação $\mu = 1$. Existe uma conexão similar entre FSK e FM. FSK é simplesmente um sinal FM com um número limitado de frequências instantâneas.

A relação entre PSK e modulação em ângulo é um pouco mais sutil. Para PSK, o sinal modulado pode ser escrito como

$$\varphi_{\text{PSK}}(t) = A \cos(\omega_c t + \theta_k) \quad kT_b \leq t < kT_b + T_b$$

Esse sinal pode, portanto, ser relacionado a PM. Contudo, um exame mais detalhado do sinal PSK revela que, devido à fase constante θ_k , sua frequência instantânea não varia. Na verdade, fazendo $a_k = A \cos \theta_k$ e $b_k = -A \sin \theta_k$, podemos reescrever o sinal PSK como

$$\begin{aligned} \varphi_{\text{PSK}}(t) &= A \cos \theta_k \cos \omega_c t - A \sin \theta_k \sin \omega_c t \\ &= a_k \cos \omega_c t + b_k \sin \omega_c t \quad kT_b \leq t < kT_b + T_b \end{aligned} \quad (7.61)$$

Da Eq. (7.61), reconhecemos uma forte semelhança com a representação do sinal QAM dada na Seção 4.4. Portanto, uma modulação digital PSK guarda uma relação próxima com o sinal analógico QAM. Em particular, para PSK binário, $\theta = 0, \pi$. Assim, PSK binário pode ser escrito como

$$\pm A \cos \omega_c t$$

Isso é, efetivamente, uma manifestação digital da modulação em amplitude DSB-SC. Na verdade, como veremos mais adiante, se permitirmos que a_k assumam valores em múltiplos níveis e tomarmos $b_k = 0$, podemos gerar outra modulação digital com portadora conhecida como modulação por amplitude de pulso (ou PAM), que pode transportar múltiplos bits durante cada intervalo de tempo T_b de modulação.

Como estudado no Capítulo 4, a modulação em amplitude DSB-SC tem melhor eficiência de potência que AM. Portanto, a PSK binária tem melhor eficiência de potência que a ASK. Em termos de uso de largura de banda, podemos ver, de suas relações com modulações analógicas, que a ASK e a PSK têm idênticas larguras de banda, enquanto a FSK requer maior largura de banda. Essas observações corroboram, de modo intuitivo, os resultados de PSD na Fig. 7.32.

7.8.4 Demodulação

A demodulação de sinais modulados digitais é similar à de sinais modulados analógicos. Devido às relações entre ASK e AM, entre FSK e FM e entre PSK e QAM (ou DSB-SC AM), as diferentes técnicas empregadas para modulações analógicas podem ser diretamente aplicadas aos equivalentes digitais.

Detecção ASK

Assim como a AM, a ASK (Fig. 7.30c) pode ser demodulada de forma coerente (para detecção síncrona) ou incoerente (para detecção de envelope). O detector coerente requer equipamento mais elaborado e tem desempenho superior, especialmente quando a potência de sinal (e, portanto, a SNR) é baixa. Para SNR mais alta, o detector de envelope tem desempenho quase tão bom como o detector coerente. Por conseguinte, a detecção coerente não é muito utilizada para ASK, pois violaria o propósito do esquema (simplicidade de detecção). Se pudermos tirar proveito de um detector síncrono, é preferível o uso de PSK, que tem melhor eficiência de potência que ASK.

Detecção FSK

Novamente, o sinal FSK binário pode ser visto como dois sinais ASK entrelaçados, com frequências portadoras ω_{c0} e ω_{c1} (Fig. 7.32c). Portanto, a FSK pode ser detectada de forma coerente ou incoerente. Na detecção incoerente, o sinal que chega é aplicado a um par de filtros sintonizados em ω_{c0} e ω_{c1} . Cada filtro é seguido por um detector de envelope (Fig. 7.33a). As saídas dos dois detectores de envelope são amostradas e comparadas. Se um 0 for transmitido por um pulso de frequência ω_{c0} , esse pulso aparecerá na saída do filtro sintonizado em ω_{c0} . Praticamente nenhum sinal aparecerá na saída do filtro sintonizado em ω_{c1} . Assim, a amostra da saída do detector de envelope que segue o filtro ω_{c0} será maior que a amostra da saída do detector de envelope que segue o filtro ω_{c1} , e o receptor decide que um 0 foi transmitido. No caso de um 1, ocorre o oposto.

É óbvio que a FSK também pode ser detectada coerentemente com a geração de duas frequências de referência ω_{c0} e ω_{c1} , para os dois demoduladores, para a demodulação do sinal recebido e a

Figura 7.33
(a) Detecção incoerente de FSK. (b) Detecção coerente de FSK.

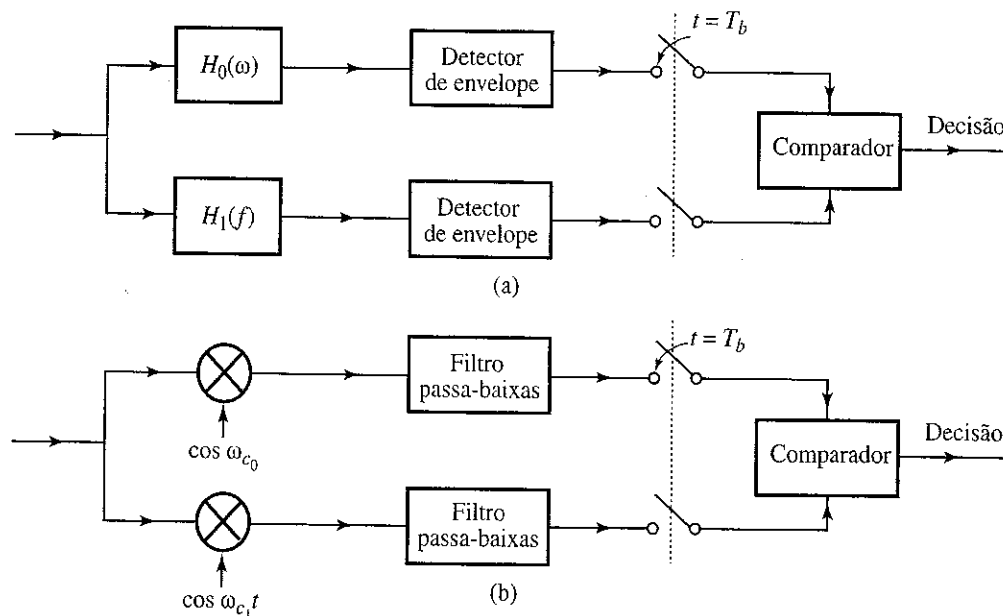
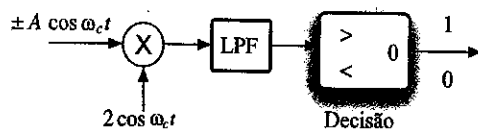


Figura 7.34
Detector PSK binário coerente (similar a um demodulador DSB-SC).



subsequente comparação das saídas dos dois demoduladores, como ilustrado na Fig. 7.33b. Assim, o detector FSK coerente deve gerar duas portadoras em sincronia com as portadoras de modulação. Novamente, esse demodulador complexo viola o propósito da FSK, projetado principalmente para detecção mais simples e não coerente. Na prática, a detecção FSK coerente não é utilizada.

Deteção PSK

Em PSK binário, um 1 é transmitido por um pulso $A \cos \omega_c t$ e um 0, por um pulso $-A \cos \omega_c t$ (Fig. 7.31b). A informação em sinais PSK reside na fase da portadora. Assim como em DSB-SC, esses sinais não podem ser demodulados por detecção de envelope, pois o envelope é o mesmo para 1 e 0 (Fig. 7.31b). O detector coerente da modulação PSK binária é mostrado na Fig. 7.34. A detecção coerente é similar à usada para sinais analógicos. Métodos de aquisição de portadora foram discutidos na Seção 4.8.

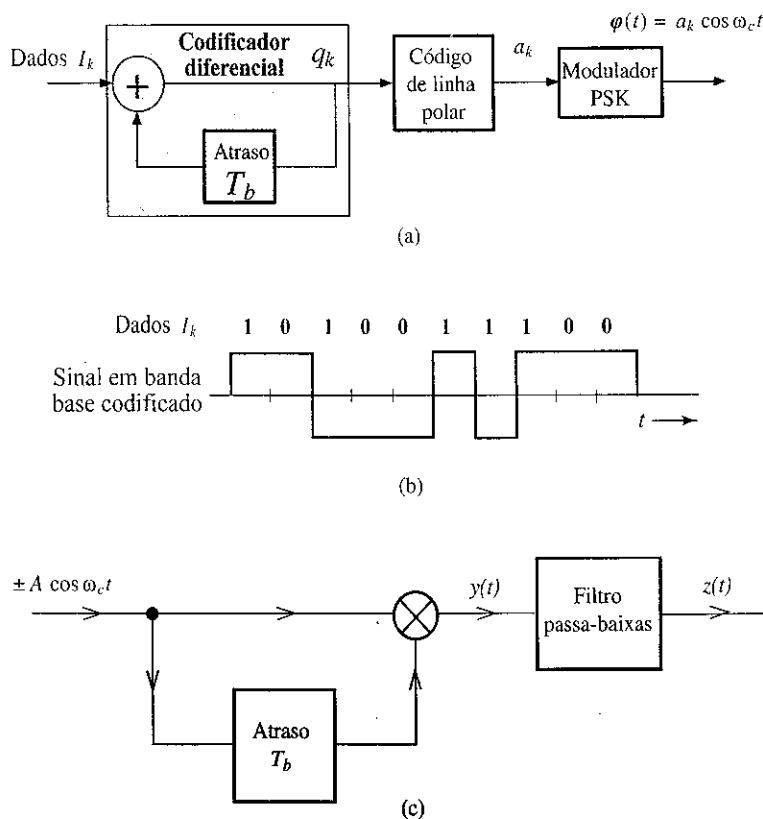
PSK Diferencial

Embora a detecção de envelope não possa ser usada para PSK, ainda é possível explorar o número finito de valores de fase de modulação para detecção incoerente. De fato, sinais PSK podem ser demodulados de forma incoerente por meio de um engenhoso método conhecido como **PSK diferencial** ou DPSK. O princípio da detecção diferencial é que o receptor detecte a mudança de fase relativa entre sucessivas fases moduladas θ_k e θ_{k-1} . Como os valores de fase em PSK são finitos (iguais a 0 e π em PSK binário), o transmissor pode codificar o dado de informação na diferença de fase $\theta_k - \theta_{k-1}$. Por exemplo, uma diferença de fase zero representa 0, enquanto uma diferença de fase π significa 1.

Essa técnica é conhecida como **codificação diferencial** (antes da modulação). Em um código diferencial, um 0 é codificado pelo mesmo pulso usado para codificar o bit de dado anterior (sem transição), e 1 é codificado pelo negativo do pulso usado para codificar o bit de dado anterior (transição). A codificação diferencial é de simples implementação, como mostrado na Fig. 7.35b. Reparemos que a adição é em módulo 2. O sinal codificado é mostrado na Fig. 7.35b. Desse modo, uma transição na sequência de pulsos do código de linha indica 1 e nenhuma transição, 0. O sinal modulado consiste em pulsos

$$A \cos(\omega_c t + \theta_k) = \pm A \cos \omega_c t$$

Figura 7.35
(a) Codificação diferencial;
(b) sinal codificado;
(c) receptor PSK diferencial.



Se o bit de dado for 0, o pulso corrente e o anterior têm a mesma polaridade ou fase; os dois pulsos são iguais a $A \cos \omega_c t$ ou iguais a $-A \cos \omega_c t$. Se o bit de dado for 1, o pulso corrente e o anterior têm polaridades ou fases opostas; se o pulso corrente for $A \cos \omega_c t$, o anterior é $-A \cos \omega_c t$, ou vice-versa.

Na demodulação da DPSK (Fig. 7.35c), evitamos a geração de uma portadora local observando que o próprio sinal modulado recebido é uma portadora ($\pm A \cos \omega_c t$), com uma possível ambiguidade de sinal. Para demodulação, no lugar da portadora, podemos usar o sinal recebido atrasado por T_b (um intervalo de bit). Se o pulso recebido for igual ao anterior, o produto dos dois será $y(t) = A^2 \cos^2 \omega_c t = (A^2/2)(1 + \cos 2\omega_c t)$, e a saída do filtro passa-baixas, $z(t) = A^2/2$. Imediatamente, detectamos o pulso corrente como um 0. Se o pulso recebido e o anterior tiverem polaridades opostas, $y(t) = -A^2 \cos^2 \omega_c t$ e $z(t) = -A^2/2$, e o bit corrente é detectado como 1. A Tabela 7.3 ilustra um exemplo específico de codificação e decodificação.

Tabela 7.3

Codificação diferencial e detecção de DPSK binário

Tempo k	0	1	2	3	4	5	6	7	8	9	10
I_k		1	0	1	0	0	1	1	1	0	0
Q_k	0	1	1	0	0	0	1	0	1	1	1
Código de linha a_k	-1	1	1	-1	-1	-1	1	-1	1	1	1
θ_k	π	0	0	π	π	π	0	π	0	0	0
$\theta_k - \theta_{k-1}$		π		π			π	π	π	0	0
Bits detectados		1	0	1	0	0	1	1	1	0	0

Em termos de complexidade de demodulação, ASK, FSK e DPSK podem ser todos detectados incoerentemente, sem uma portadora síncrona no receptor. A PSK, por sua vez, deve ser detectada coerentemente. Detecção incoerente, no entanto, tem um preço em termos de imunidade ao ruído. Do ponto de vista de imunidade ao ruído, a PSK coerente é superior a todos os outros esquemas. A PSK também requer menor largura de banda do que a FSK (Fig. 7.32). Uma discussão quantitativa sobre esse tema será feita no Capítulo 10.

7.9 MODULAÇÃO DIGITAL M-ÁRIA COM PORTADORA

As modulações digitais binárias ASK, FSK e PSK com portadora transmitem um bit de informação por intervalo de T_b segundos, o que corresponde a uma taxa de bits de $1/T_b$ bit/s. Assim como no caso da transmissão de sinais digitais em banda base, maior taxa de transmissão de bits pode ser alcançada com a redução de T_b ou com aplicação de sinalização M -ária; a primeira opção requer mais largura de banda e a segunda, mais potência. Na maioria dos sistemas de comunicação, largura de banda é estritamente limitada. Assim, para conservar largura de banda, uma forma eficaz de aumentar a taxa de transmissão de dados consiste em generalizar a modulação binária com o emprego de sinalização M -ária. Especificamente, podemos aplicar modulações ASK de M níveis, FSK de M frequências e PSK de M fases.

ASK M -ária e Detecção Incoerente

A ASK M -ária é uma generalização muito simples da modulação ASK binária. Em vez de enviar apenas

$$\varphi(t) = 0 \text{ para } 0 \quad \text{e} \quad \varphi(t) = A \cos \omega_c t \text{ para } 1$$

a modulação ASK M -ária pode enviar $\log_2 M$ bits por vez ao transmitir, por exemplo,

$$\varphi(t) = 0, A \cos \omega_c t, 2A \cos \omega_c t, \dots, (M-1)A \cos \omega_c t$$

Esse ainda é um sinal AM que usa M amplitudes distintas e um índice de modulação $\mu = 1$. A largura de banda é igual à do sinal ASK binário, enquanto a potência cresce proporcionalmente com M^2 . A demodulação pode ser feita por detecção de envelope ou por detecção coerente.

FSK M -ária e Sinalização Ortogonal

A FSK M -ária é gerada com a seleção de uma senoide do conjunto $\{A \cos 2\pi f_i t, i = 1, \dots, M\}$ para transmitir um padrão particular de $\log_2 M$ bits. Em geral, com FSK, podemos projetar um incremento de frequência δf e definir

$$f_m = f_1 + (m - 1)\delta f \quad m = 1, 2, \dots, M$$

Para essa modulação FSK com igual separação entre frequências, o desvio de frequência (na análise do sinal FM) é

$$\Delta f = \frac{f_M - f_1}{2} = \frac{1}{2}(M - 1)\delta f$$

Fica claro que a seleção do conjunto de frequências $\{f_i\}$ determina o desempenho e a largura de banda da modulação FSK. Se δf for escolhido demasiadamente grande, o sinal FSK M -ário usará excessiva largura de banda. Se, por outro lado, δf for escolhido demasiadamente pequeno, no intervalo de tempo de T_b segundos praticamente não haverá distinção entre diferentes símbolos FSK e o receptor será incapaz de distinguir os símbolos de modo confiável. Assim, grande valor de δf leva a desperdício de largura de banda, enquanto pequeno δf favorece a suscetibilidade a erros de detecção devido ao ruído de transmissão e à interferência.

O projeto da modulação FSK M -ária requer a determinação de um valor suficientemente pequeno de δf , de modo que cada símbolo FSK $A \cos \omega_i t$ seja distinguível de todos os outros símbolos FSK. Uma solução para esse problema de projeto de sinais FSK pode ser encontrada na discussão do espaço ortogonal de sinais na Seção 2.6.2. Se pudermos projetar símbolos FSK que sejam ortogonais em T_b com a seleção de um pequeno valor para δf (ou Δf), os sinais FSK serão verdadeiramente distintos no intervalo T_b e o consumo de largura de banda será pequeno.

Para determinar o mínimo valor de δf que leva a um conjunto ortogonal de sinais FSK, a condição de ortogonalidade, segundo a Seção 2.6.2, requer que

$$\int_0^{T_b} A \cos(2\pi f_m t) A \cos(2\pi f_n t) dt = 0 \quad m \neq n \quad (7.62)$$

Podemos usar essa condição para determinar o mínimo δf . Primeiro, calculemos

$$\begin{aligned} \int_0^{T_b} A \cos(2\pi f_m t) A \cos(2\pi f_n t) dt &= \frac{A^2}{2} \int_0^{T_b} [\cos 2\pi(f_m + f_n)t + \cos 2\pi(f_m - f_n)t] dt \\ &= \frac{A^2}{2} T_b \frac{\sin 2\pi(f_m + f_n)T_b}{2\pi(f_m + f_n)T_b} + \frac{A^2}{2} T_b \frac{\sin 2\pi(f_m - f_n)T_b}{2\pi(f_m - f_n)T_b} \end{aligned} \quad (7.63)$$

Como, em modulações práticas, $(f_m + f_n)T_b$ é muito grande (em geral, não menor que 10^3), o primeiro termo na Eq. (7.63) é efetivamente zero e desprezível. Assim, a condição de ortogonalidade se reduz à condição que, para quaisquer inteiros $m \neq n$,

$$\frac{A^2}{2} \frac{\sin 2\pi(f_m - f_n)T_b}{2\pi(f_m - f_n)} = 0$$

Como $f_m = f_1 + (m - 1)\delta f$, para mútua ortogonalidade, temos

$$\sin[2\pi(m - n)\delta f T_b] = 0 \quad m \neq n$$

Dessa condição, fica claro que o menor δf que satisfaz a condição de mútua ortogonalidade é

$$\delta f = \frac{1}{2T_b} \text{ Hz}$$

Essa escolha de mínima separação de frequência é conhecida como FSK *de mínimo deslocamento*. Por formar um conjunto ortogonal de símbolos, essa escolha também é conhecida como sinalização ortogonal.

Podemos descrever a modulação FSK *de mínimo deslocamento* geometricamente, aplicando o conceito de funções de base ortogonais da Seção 2.6. Sejam

$$\psi_i(t) = \sqrt{\frac{2}{T_b}} \cos 2\pi \left(f_1 + \frac{i-1}{2T_b} \right) t \quad i = 1, 2, \dots, M$$

Podemos comprovar com facilidade que

$$\int_0^{T_b} \psi_m(t) \psi_n(t) dt = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Assim, cada símbolo FSK pode ser escrito como

$$A \cos 2\pi f_m t = A \sqrt{\frac{T_b}{2}} \psi_m(t) \quad m = 1, 2, \dots, M$$

A relação geométrica entre os dois símbolos FSK para $M = 2$ é capturada na Fig. 7.36.

A demodulação de sinais FSK M -ários segue os mesmos princípios que a demodulação de sinais FSK binários. Com a generalização dos demoduladores FSK binários na Fig. 7.33, podemos aplicar um banco de M detectores coerentes ou incoerentes para sinais FSK M -ários antes de tomar uma decisão com base no ramo detector mais intenso.

Anteriormente, na análise da PSD de modulações em banda base, mostramos que, no intervalo de símbolo T_b , a largura de banda de um sinal digital em banda base podia ser aproximada por $1/T_b$. Portanto, para FSK de deslocamento mínimo, $\Delta f = (M-1)/(4T_b)$ e, segundo a regra de Carson, a largura de banda é, aproximadamente,

$$2(\Delta f + B) = \frac{M-3}{2T_b}$$

Na verdade, podemos mostrar que, em geral, a largura de banda de um esquema M -ário ortogonal é M vezes a do esquema binário [Seção 10.7, Eq. (10.123)]. Portanto, em um esquema M -ário ortogonal, a taxa de comunicação aumenta por um fator $\log_2 M$, ao custo de um aumento na largura de banda de transmissão por um fator M . No esquema ortogonal, para uma comparável imunidade ao ruído, a potência de transmissão praticamente independe de M . Por conseguinte, em contraste com a modulação ASK M -ária, a FSK M -ária não requer maior potência de transmissão. Contudo, seu requisito de largura de banda aumenta quase linearmente com M (em comparação com FSK binária ou ASK M -ária).

Figura 7.36
Símbolos FSK
binários no
espaço
ortogonal
bidimensional
de sinais.

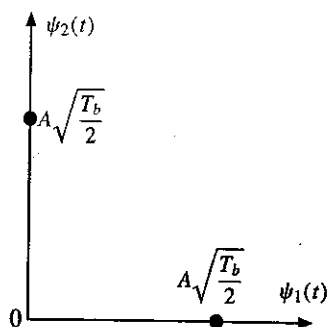
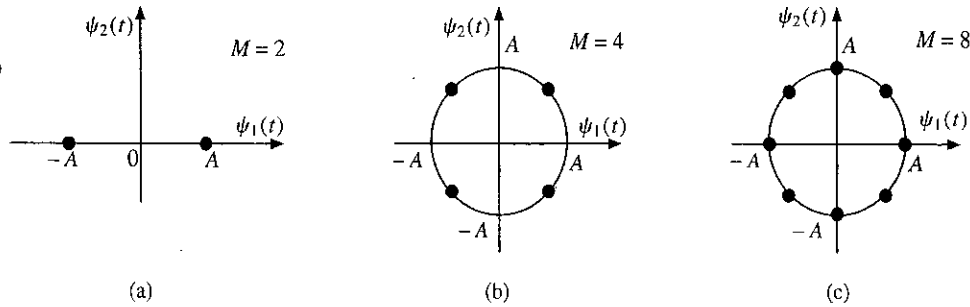


Figura 7.37
Símbolos PSK
 M -ário no espaço
ortogonal de
sinais: (a) $M = 2$;
(b) $M = 4$;
(c) $M = 8$.



PSK, PAM e QAM M -árias

Com uma pequena modificação na Eq. (7.61), sinais PSK genéricos podem ser escritos no formato

$$\varphi_{\text{PSK}}(t) = a_m \sqrt{\frac{2}{T_b}} \cos \omega_c t + b_m \sqrt{\frac{2}{T_b}} \sin \omega_c t \quad 0 \leq t < T_b \quad (7.64a)$$

em que $a_m = A \cos \theta_m$ e $b_m = -A \sin \theta_m$. Na verdade, com base na análise na Seção 2.6, $\sqrt{2/T_b} \cos \omega_c t$ e $\sqrt{2/T_b} \sin \omega_c t$ são mutuamente ortogonais. Além disso, são normalizados no intervalo $[0, T_b]$. Como resultado, podemos representar todos os símbolos PSK em um sinal bidimensional de sinais com funções de base

$$\psi_1(t) = \sqrt{\frac{2}{T_b}} \cos \omega_c t \quad \psi_2(t) = \sqrt{\frac{2}{T_b}} \sin \omega_c t$$

tais que

$$\varphi_{\text{PSK}}(t) = a_m \psi_1(t) + b_m \psi_2(t) \quad (7.64b)$$

Podemos ilustrar geometricamente a relação entre símbolos PSK no espaço de sinais (Fig. 7.37). A Eq. (7.64) significa que modulações PSK podem ser representadas como uma modulação QAM. Na verdade, como o sinal é PSK, os pontos de sinal devem satisfazer a condição especial

$$\begin{aligned} a_m^2 + b_m^2 &= A^2 \cos^2 \theta_m + (-A)^2 \sin^2 \theta_m \\ &= A^2 = \text{constante} \end{aligned} \quad (7.64c)$$

Em outras palavras, todos os pontos de sinal devem estar em uma circunferência de raio A . Na prática, todos os pontos de sinal são escolhidos para serem igualmente espaçados, com o propósito de obter melhor imunidade ao ruído. Portanto, para a sinalização PSK M -ária, os ângulos são escolhidos como uniformemente espaçados:

$$\theta_m = \theta_0 + \frac{2\pi}{M}(m-1) \quad m = 1, 2, \dots, M$$

A sinalização PSK especial com $M=4$ é um formato de modulação digital extremamente popular e poderoso.* Esse sinal é a soma de dois sinais PSK binários de mesma frequência, um com portadora $\cos \omega_c t$ (em fase) e outro com portadora $\sin \omega_c t$ (em quadratura). Por essa razão, este sinal também é conhecido como **PSK em quadratura (QPSK – quadrature PSK)**. Podemos transmitir e receber os dois sinais no mesmo canal, dobrando a taxa de transmissão.

*QPSK tem diversas variações práticas, incluindo QPSK offset.

Para generalizar PSK e alcançar taxa de dados ainda mais alta, podemos ver que a representação PSK da Eq. (7.64) é um caso especial da modulação em amplitude em quadratura (QAM) discutida no Capítulo 4 (Fig. 4.19). A única diferença reside na exigência de PSK de que o sinal modulado tenha uma magnitude (módulo) constante A . Na verdade, o formato de modulação QAM, muito mais flexível e geral, também pode ser convenientemente usado para modulação digital. O sinal transmitido por um sistema QAM M -ário pode ser escrito como

$$\begin{aligned} p_i(t) &= a_i p(t) \cos \omega_c t + b_i p(t) \sin \omega_c t \\ &= r_i p(t) \cos (\omega_c t - \theta_i) \quad i = 1, 2, \dots, M \end{aligned}$$

em que

$$r_i = \sqrt{a_i^2 + b_i^2} \quad \text{e} \quad \theta_i = \tan^{-1} \frac{b_i}{a_i} \quad (7.65)$$

e $p(t)$ é um pulso em banda base adequadamente formatado. A escolha mais simples para $p(t)$ seria um pulso retangular

$$p(t) = \sqrt{\frac{2}{T_b}} [u(t) - u(t - T_b)]$$

Pulsos melhores podem, certamente, ser aplicados para conservar largura de banda.

A Fig. 7.38 mostra o modulador e o demodulador QAM. Cada um dos dois sinais $m_1(t)$ e $m_2(t)$ é uma sequência de pulsos \sqrt{M} -ários em banda base. Os dois sinais são modulados por duas portadoras de mesma frequência, mas em quadratura de fase. O sinal QAM digital $p_i(t)$ pode ser gerado por meio de QAM com $m_1(t) = a_i p(t)$ e $m_2(t) = b_i p(t)$. Os sinais $m_1(t)$ e $m_2(t)$ são sinais PAM em banda base. O diagrama de olho do sinal QAM consiste na componente em fase $m_1(t)$ e na componente em quadratura $m_2(t)$. Ambas exibem o diagrama de olho do sinal PAM M -ário em banda base, como discutido anteriormente na Seção 7.6.

A representação geométrica de sinais QAM M -ários pode ser obtida como extensão do espaço de sinais PSK com a simples remoção da condição de módulo constante na Eq. (7.64c). Uma escolha prática e muito popular de r_i e θ_i para $M = 16$ é ilustrada graficamente na Fig. 7.38b. O pulso transmitido $p_i(t)$ pode assumir 16 formas distintas e, portanto, é um pulso 16-ário. Como $M = 16$, cada pulso pode transmitir a informação de $\log_2 16 = 4$ dígitos binários. Isso pode ser feito da seguinte forma: há 16 possíveis sequências de quatro dígitos binários e há 16 combinações (a_i, b_i) na Fig. 7.38b. Assim, cada possível sequência de quatro bits é transmitida por um particular (a_i, b_i) ou (r_i, θ_i) . Portanto, um pulso de sinal $r_i p(t) \cos (\omega_c t - \theta_i)$ transmite quatro bits. Em comparação com PSK binário (ou BPSK), a taxa de bits do sinal QAM 16-ário é quatro vezes maior, sem aumentar a largura de banda. A taxa de transmissão pode ser aumentada ainda mais com o aumento do valor de M .

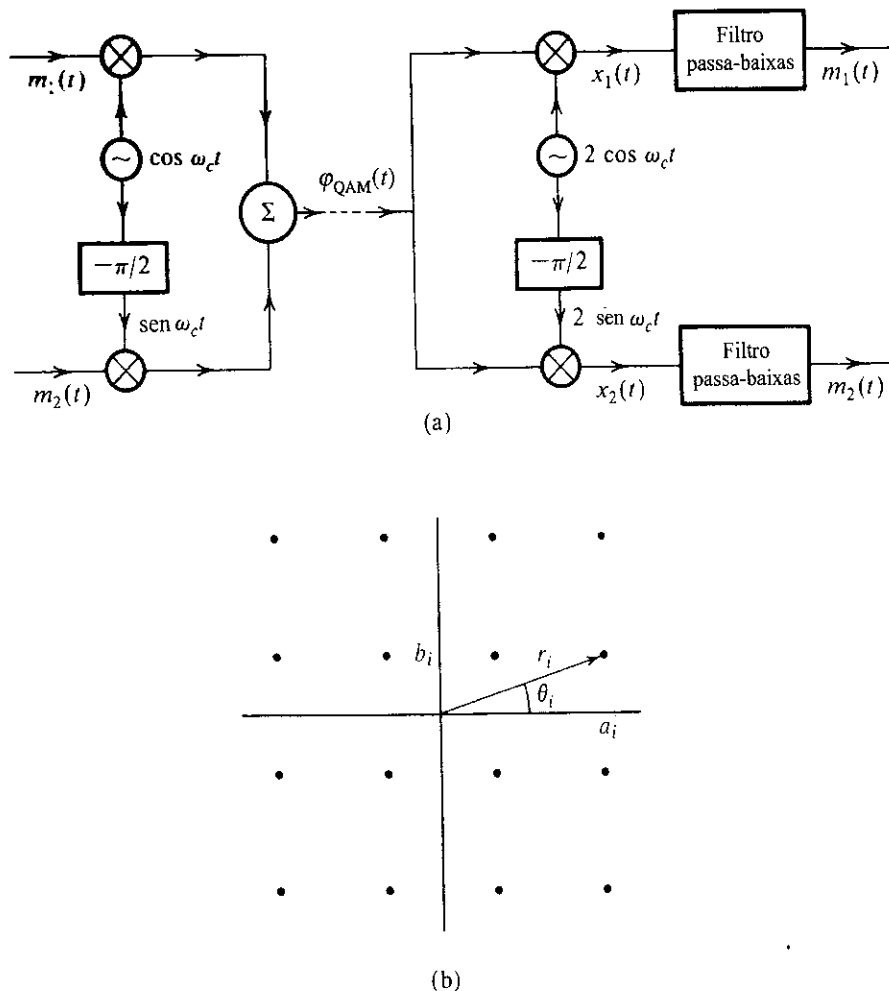
Modulação e demodulação podem ser efetuadas com o sistema ilustrado na Fig. 7.38a. As entradas são $m_1(t) = a_i p(t)$ e $m_2(t) = b_i p(t)$. As duas saídas do demodulador são $a_i p(t)$ e $b_i p(t)$. O conhecimento de (a_i, b_i) permite a determinação de quatro bits transmitidos. Uma análise mais detalhada de sinais QAM 16-ários em canal ruidoso é feita na Seção 10.6 [Eq. (10.104)]. O valor prático da sinalização QAM-16 fica evidente quando consideramos a grande variedade de suas aplicações. Na verdade, a QAM-16 é usada em modems de dados telefônicos/fax V.32 (9600 bit/s), em modems de cabo de alta velocidade e na moderna difusão da televisão digital via satélite.

Reparemos que, se desabilitarmos a sequência de dados que modula $\sin \omega_c t$ em QAM, todos os pontos de sinalização poderão ser reduzidos a uma única dimensão. Fazendo $m_2(t) = 0$, a modulação QAM passa a

$$p_i(t) = a_i p(t) \cos \omega_c t, \quad t \in [0, T_b]$$

Figura 7.38

(a) QAM ou multiplexação em quadratura e (b) QAM de 16 pontos ($M = 16$).



Que se degenera em modulação por amplitude de pulso ou PAM. Uma comparação da expressão do sinal $p_i(t)$ com a do sinal DSB-SC analógico deixa claro que PAM é a versão digital de DSB-SC. Assim como um sinal QAM analógico é formado pela superposição de dois sinais com modulação em amplitude DSB-SC em quadratura de fase, o sinal QAM digital consiste em dois sinais PAM, cada um com \sqrt{M} níveis de sinalização. Como na relação entre as modulações DSB-SC analógica e QAM, a PAM requer a mesma largura de banda que QAM. No entanto, a PAM é muito menos eficiente, pois necessita de M níveis de sinalização de modulação em uma dimensão, enquanto QAM requer apenas com \sqrt{M} níveis de sinalização em cada uma das duas dimensões QAM ortogonais.

Negociação entre Potência e Largura de Banda

No Capítulo 10, discutiremos outros tipos de sinalização M -ária. A natureza da negociação entre largura de banda de transmissão e potência transmitida (ou SNR) depende da escolha do esquema M -ário. Por exemplo, na sinalização ortogonal, a potência transmitida praticamente independe de M , mas a largura de banda de transmissão aumenta com M . No caso PAM, em contraste, a potência transmitida aumenta aproximadamente com M^2 , enquanto a largura de banda permanece constante. Assim, a sinalização M -ária nos permite grande flexibilidade na negociação entre potência de sinal (ou SNR) e largura de banda de transmissão. A escolha do sistema apropriado depende de circunstâncias particulares. Por exemplo, o uso de sinalização QAM é apropriado se a largura de banda for limitada (como em linha telefônicas), e o uso de sinalização ortogonal é adequado quando há restrição de potência (como em comunicações espaciais).

7.10 EXERCÍCIOS COM O MATLAB

Nesta seção, apresentamos programas MATLAB para gerar diagramas de olho. O primeiro passo consiste na especificação das formas básicas de pulso em PAM. Os quatro programas curtos dados a seguir são usados para gerar pulsos NRZ, RZ, meia senoide e cosseno levantado.

```
% (pnrz.m)
% gera um pulso retangular de largura T
% Uso function pout=pnrz(T);
function pout=prect(T);
pout=ones(1,T);
end
```

```
% (prz.m)
% gera um pulso retangular de largura T/2
% Uso function pout=prz(T);
function pout=prz(T);
pout=[zeros(1,T/4) ones(1,T/2) zeros(1,T/4)];
end
```

```
% (psine.m)
% gera pulso senoidal de largura T
%
function pout=psine(T);
pout=sin(pi*[0:T-1]/T);
end
```

```
% (prcos.m)
% Uso y=prcos(rollfac,length, T)
function y=prcos(rollfac,length, T)
% rollfac = 0 a 1 é o fator de decaimento
% length é o comprimento unilateral do pulso em termos de T
% length = 2T+1;
% T é a taxa de superamostragem
y=rcosfir(rollfac, length, T,1, 'normal');
end
```

O primeiro programa (binary_eye.m) usa quatro diferentes formas de pulso para gerar diagramas de olho de sinalização polar binária.

```
% (binary_eye.m)
% gera e traça diagrama de olho
%
clear;clf;
```



```

data = sign(randn(1,400)); % Gera 400 bits aleatórios
Tau=64; % Define o período de símbolo
dataup=upsample(data, Tau); % Gera trem de impulsos
yrz=conv(dataup,prz(Tau)); % Sinal polar com retorno ao zero
yrz=yrz(1:end-Tau+1);
ynrz=conv(dataup,prnz(Tau)); % Sinal polar sem retorno ao zero
ynrz=ynrz(1:end-Tau+1);
ysine=conv(dataup,psine(Tau)); % Sinal polar de meia senoide
ysine=ysine(1:end-Tau+1);
Td=4; % trunca cosseno levantada em 4 períodos
yrcos=conv(dataup,prcos(0.5,Td,Tau)); % fator de decaimento = 0,5
yrcos=yrcos(2*Td*Tau:end-2*Td*Tau+1); % gera trem de pulsos RC
eye1=eyediagram(yrz,2*Tau,Tau,Tau/2);title('Diagrama de olho RZ');
eye2=eyediagram(ynrz,2*Tau,Tau,Tau/2);title('Diagrama de olho NRZ');
eye3=eyediagram(ysine,2*Tau,Tau,Tau/2);title('Diagrama de olho de
meia senoide');
eye4=eyediagram(yrcos,2*Tau,Tau); title('Diagrama de olho de cosseno
levantado');

```

O segundo programa (Mary-eye.m) usa quatro diferentes formas de pulso para gerar diagramas de sinalização PAM de quatro níveis.

```

% (Mary_eye.m)
% gera e traça diagrama de olho
%
%
clear;clf;
data = sign(randn(1,400))+2* sign(randn(1,400)); % Gera 400 símbo-
los PAM
Tau=64; % Define o período de símbolo
dataup=upsample(data, Tau); % Gera trem de impulsos
yrz=conv(dataup,prz(Tau)); % Sinal polar com retorno ao zero
yrz=yrz(1:end-Tau+1);
ynrz=conv(dataup,prnz(Tau)); % Sinal polar sem retorno ao zero
ynrz=ynrz(1:end-Tau+1);
ysine=conv(dataup,psine(Tau)); % Sinal polar de meia senoide
ysine=ysine(1:end-Tau+1);
Td=4; % trunca cosseno levantada em 4 períodos
yrcos=conv(dataup,prcos(0.5,Td,Tau)); % fator de decaimento = 0,5
yrcos=yrcos(2*Td*Tau:end-2*Td*Tau+1); % gera trem de pulsos RC
eye1=eyediagram(yrz,2*Tau,Tau,Tau/2);title('Diagrama de olho RZ');
eye2=eyediagram(ynrz,2*Tau,Tau,Tau/2);title('Diagrama de olho NRZ');
eye3=eyediagram(ysine,2*Tau,Tau,Tau/2);title('Diagrama de olho de meia
senoide');
eye4=eyediagram(yrcos,2*Tau,Tau); title('Diagrama de olho de cosseno
levantado');

```

REFERÊNCIAS

1. A. Lender, "Duobinary Technique for High Speed Data Transmission," *IEEE Trans. Commun. Electron.*, vol. CE-82, pp. 214-218, May 1963.

2. A. Lender, "Correlative Level Coding for Binary-Data Transmission," *IEEE Spectrum*, vol. 3, no. 2, pp. 104-115, Feb. 1966.
3. P. Bylanski and D. G. W. Ingram, *Digital Transmission Systems*, Peter Peregrinus Ltd., Hertshire, England, 1976.
4. H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *AIEE Trans.*, vol. 47, p. 817, April 1928.
5. E. D. Sunde, *Communication Systems Engineering Technology*, Wiley, New York, 1969.
6. R. W. Lucky and H. R. Rudin, "Generalized Automatic Equalization for Communication Channels," *IEEE Int. Commun. Conf.*, vol. 22, 1966.
7. W. F. Trench, "An Algorithm for the Inversion of Finite Toeplitz Matrices," *J. SIAM*, vol. 12, pp. 515-522, Sept. 1964.
8. A. Lender, Chapter 7, in *Digital Communications: Microwave Applications*, K. Feher, Ed., Prentice-Hall, Englewood Cliffs, NJ, 1981.

EXERCÍCIOS

7.2-1 Considere um pulso retangular de largura completa

$$p(t) = \Pi(t/T_b)$$

- (a) Determine a PSD para as sinalizações polar, on-off e bipolar.
- (b) Esboce o gráfico de cada PSD e determine sua largura de banda. Para cada sinalização, compare a largura de banda com a do caso em que $p(t)$ é um pulso retangular de meia largura.

- 7.2-2 (a) Uma sequência aleatória de dados binários **110100101...** é transmitida com o uso do código de linha Manchester (bifásico) com o pulso $p(t)$ mostrado na Fig. 7.7a. Esboce o gráfico da forma de onda $y(t)$.
- (b) Calcule $S_y(f)$, a PSD de um sinal Manchester (bifásico) na parte (a), supondo que 1 e 0 têm igual probabilidade de ocorrência. Esboce o gráfico desta PSD e determine sua largura de banda.

7.2-3 Para a forma de pulso

$$p(t) = \Pi\left(\frac{t}{0,5T_b}\right)$$

use o código diferencial (Fig. 7.18) para determinar a PSD para um sinal binário. Determine a PSD $S_y(f)$.

- 7.2-4 O código de linha **duobinário** proposto por Lender também é terciário, como o bipolar, mas requer somente metade da largura de banda de bipolar. Na prática, o código duobinário é realizado indiretamente com o uso de uma forma especial de pulso, como discutido na Seção 7.3 (Fig. 7.18). Nesse código, um 0 é transmitido por ausência de pulso e um 1, por um pulso $p(t)$ ou $-p(t)$ segundo a regra: um 1 é codificado pelo mesmo pulso usado para o 1 anterior, se houver um número par de 0s entre eles; um 1 é codificado por um pulso com polaridade oposta à do pulso usado para o 1 anterior, se houver um número ímpar de 0s entre eles. Um número 0 é considerado um número par. Como o bipolar, esse código também tem uma capacidade de detecção de erro isolado, pois a recepção correta implica que, entre sucessivos pulsos de mesma polaridade, deve ocorrer um número par de 0s, e entre sucessivos pulsos de polaridades opostas, um número ímpar de 0s.

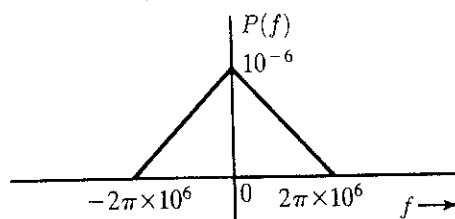
- (a) Admitindo um pulso retangular de meia largura, esboce o gráfico do sinal duobinário $y(t)$ para a sequência binária aleatória

1110001101001010...

- (b) Determine R_0 , R_1 e R_2 para este código. Admita (se preferir, prove) que $R_n = 0$ para todo $n > 2$. Determine a PSD para esse código de linha (considerando um pulso de meia largura) e esboce seu gráfico. Mostre que sua largura de banda é $R_b/2$, metade da largura de banda do código bipolar.

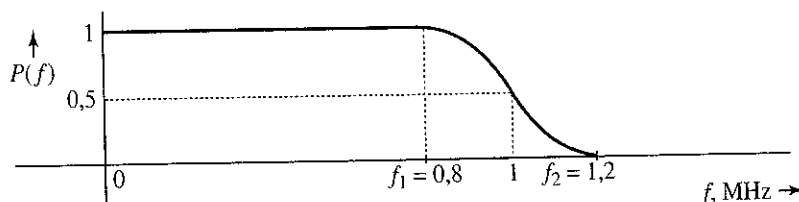
- 7.3-1 Dados, a uma taxa de 6 kbit/s, devem ser transmitidos em uma linha privada com largura de banda de 4 kHz, usando pulsos que satisfaçam o critério de Nyquist. Determine o máximo valor do fator de decaimento r que pode ser usado.
- 7.3-2 Em um certo sistema de telemetria, há oito medidas analógicas, cada uma com largura de banda de 2 kHz. Amostras desses sinais são multiplexadas por divisão no tempo, quantizadas e codificadas em binário. O erro nas amplitudes das amostras não pode ultrapassar 1% da amplitude de pico.
- (a) Determine L , o número de níveis de quantização.
- (b) Determine a largura de banda de transmissão B_T se forem usados pulsos que satisfaçam o critério de Nyquist, com fator de decaimento $r = 0,2$. A taxa de amostragem deve ser pelo menos 25% acima da taxa de Nyquist.
- 7.3-3 Uma linha privada de telefonia, com largura de banda de 3 kHz, é usada para transmitir dados binários. Calcule a taxa de dados (em bits por segundo) que pode ser transmitida se for usado:
- (a) Sinal polar com pulsos retangulares de meia largura.
- (b) Sinal polar com pulsos retangulares de largura completa.
- (c) Sinal polar com pulsos que satisfaçam o critério de Nyquist, com fator de decaimento $r = 0,25$.
- (d) Sinal bipolar com pulsos retangulares de meia largura.
- (e) Sinal bipolar com pulsos retangulares de largura completa.
- 7.3-4 A transformada de Fourier $P(f)$ do pulso básico $p(t)$ usado em certos sistemas de comunicação binária é mostrado na Fig. E.7.3-4.
- (a) A partir da forma de $P(f)$, determine a que taxa de pulsos esse pulso satisfaria o critério de Nyquist.
- (b) Determine $p(t)$ e comprove que esse pulso satisfaz (ou não) o critério de Nyquist.
- (c) Se o pulso não satisfizer o critério de Nyquist, qual é a taxa de transmissão (em bits por segundo) e qual é o valor do fator de decaimento?

Figura E.7.3-4



- 7.3-5 Um pulso $p(t)$ cujo espectro $P(f)$ é mostrado na Fig. E.7.3-5 satisfaz o critério de Nyquist. Se $f_1 = 0,8$ MHz e $f_2 = 1,2$ MHz, determine a máxima taxa em que dados binários podem ser transmitidos por esse pulso usando o critério de Nyquist. Qual é o valor do fator de decaimento?

Figura E.7.3-5



- 7.3-6 Dados binários, à taxa de 1 Mbit/s, devem ser transmitidos com uso de pulsos que satisfaçam o critério de Nyquist, sendo $P(f)$ mostrada na Fig. E.7.3-5. As frequências f_1 e f_2 do espectro são ajustáveis. O canal disponível para transmissão desses dados tem largura de banda de 700 kHz. Determine f_1 e f_2 e o fator de decaimento.
- 7.3-7 Mostre que a transformada de Fourier inversa de $P(f)$ na Eq. (7.39) é, de fato, o pulso $p(t)$ dado na Eq. (7.38), que satisfaz o segundo critério de Nyquist.

Sugestão: Use a Eq. (3.32) para determinar a transformada inversa de Fourier de $P(f)$ na Eq. (7.39) e expresse $\text{sinc}(x)$ na forma $\text{sen } x/x$.

- 7.3-8 Mostre que a transformada de Fourier inversa de $P(f)$ (o espectro do pulso cosseno levantado) na Eq. (7.35) é o pulso $p(t)$ dado na Eq. (7.36).

Sugestão: Use a Eq. (3.32) para determinar a transformada inversa de Fourier de $P(f)$ na Eq. (7.35) e expresse $\text{sinc}(x)$ na forma $\text{sen } x/x$.

- 7.3-9 Mostre que existe um (e somente um) pulso $p(t)$ de largura de banda $R_b/2$ Hz que satisfaz o segundo critério de Nyquist [Eq. (7.37)]. Mostre que esse pulso é dado por

$$p(t) = \{\text{sinc}(\pi R_b t) + \text{sinc}[\pi R_b(t - T_b)]\} = \frac{\text{sen}(\pi R_b t)}{\pi R_b t(1 - R_b t)}$$

e que sua transformada de Fourier é $P(f)$ dada na Eq. (7.39).

Sugestão: Para um pulso de largura de banda $R_b/2$ Hz, o intervalo de Nyquist é $1/R_b = T_b$, e a condição (7.37) dá os valores das amostras de Nyquist em $t = \pm nT_b$. Para construir $p(t)$, use a fórmula de interpolação [Eq. (6.10)] com $B = R_b/2$, $T_s = T_b$. Na determinação de $P(f)$, use $(1 + e^{-j2\pi f T_b}) = e^{-j\pi f T_b} (e^{j\pi f T_b} + e^{-j\pi f T_b})$.

- 7.3-10 Em uma transmissão de dados binários usando pulsos duobinários, valores de amostras foram lidos como:

120 -2 -200 -202002000 -2

(a) Determine se há qualquer erro na detecção.

(b) Se não houver erro de detecção, determine a sequência de bits recebida.

- 7.3-11 Em uma transmissão de dados binários usando pulsos duobinários, valores de amostras foram lidos como:

12000 -200 -20200 -20220 -2

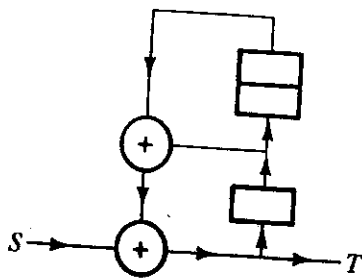
(a) Determine se há qualquer erro na detecção.

(b) Você é capaz de deduzir a correta sequência de dígitos transmitida? Há mais de uma sequência correta possível. Forneça tantas sequências corretas quanto possível, admitindo que a ocorrência de mais de um erro de detecção é extremamente provável.

- 7.4-1 No Exemplo 7.2, quando a sequência $S = 101010100000111$ foi aplicada à entrada do desembaralhador na Fig. 7.20a, a saída T foi determinada como 101110001101001 . Comprove que, quando essa sequência T é aplicada à entrada do desembaralhador na Fig. 7.20b, a saída é a sequência original de entrada, $S = 101010100000111$.

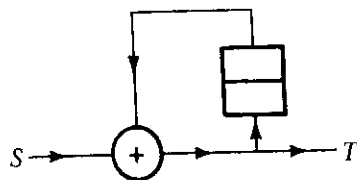
- 7.4-2 Projete um desembaralhador para o embaralhador da Fig. E.7.4-2. Se uma sequência $S = 101010100000111$ for aplicada à entrada do embaralhador, determine a sequência de saída T . Comprove que, se essa sequência T for aplicada à entrada do desembaralhador, a saída será a sequência S .

Figura E.7.4-2



- 7.4-3 Repita o Exercício 7.4-2 para o caso em que o desembaralhador mostrado na Fig. E.7.4-3 é concatenado com o embaralhador na Fig. E.7.4-2 para formar um embaralhador composto.

Figura E.7.4-3



7.5-1 Em certo sistema de comunicação binária que usa pulsos que satisfazem o critério de Nyquist, um pulso recebido $p_r(t)$ (Fig. 7.22a) tem os seguintes valores de amostras não nulas

$$\begin{aligned} p_r(0) &= 1 \\ p_r(T_b) &= 0,1 & p_r(-T_b) &= 0,3 \\ p_r(2T_b) &= -0,02 & p_r(-2T_b) &= -0,07 \end{aligned}$$

- Determine a configuração das derivações de um equalizador de três derivações e forçamento a zero.
- Usando o equalizador na parte (a), determine a ISI residual não nula.

7.7-1 Em um esquema PAM com $M = 16$,

- Determine a mínima largura de banda necessária para transmitir dados a uma taxa de 12.000 bits/s com ISI zero.
- Determine a largura de banda de transmissão se pulsos que satisfazem o critério de Nyquist com fator de decaimento $r = 0,2$ forem usados para transmitir os dados.

7.7-2 Um sinal de áudio de largura de banda de 4 kHz é amostrado a uma taxa 25% acima da taxa de Nyquist, e quantizado. O erro de quantização não deve exceder 0,1% da amplitude de pico do sinal. As resultantes amostras quantizadas são, então, codificadas e transmitidas por pulsos 4-ários.

- Determine o número mínimo de pulsos 4-ários para codificar cada amostra.
- Determine a largura mínima de banda necessária à transmissão destes dados com ISI zero.
- Se pulsos 4-ários que satisfaçam o critério de Nyquist com 25% de decaimento forem usados para transmitir estes dados, determine a largura de banda de transmissão.

7.7-3 Dados binários são transmitidos em um certo canal a uma taxa de R_b bit/s. Para reduzir a largura da banda, foi decidido o uso de sinalização PAM 16-ária para a transmissão desses dados.

- Qual é o fator de redução de largura de banda?
- Qual é o fator de aumento da potência de transmissão, admitindo que a mínima separação entre amplitudes de pulso seja a mesma nos dois casos.

Sugestão: Tome amplitudes de pulso como $\pm A/2, \pm 3A/2, \pm 5A/2, \pm 7A/2, \dots, \pm 15A/2$, de modo que a mínima separação entre os vários níveis de amplitude seja A (como no caso binário, com pulsos $\pm A/2$). Suponha que todos os 16 níveis tenham igual probabilidade de ocorrência. Recorde, ainda, que a multiplicação de um pulso por uma constante k aumenta a energia do mesmo k^2 vezes.

7.7-4 Um sinal de áudio de largura de banda de 10 kHz é amostrado à taxa de 24 kHz, quantizado em 256 níveis e codificado por pulsos PAM M -ários que satisfazem o critério de Nyquist com fator de decaimento $r = 0,2$. Uma largura de banda de 30 kHz é disponível para a transmissão dos dados. Determine o melhor valor de M .

7.7-5 Considere o caso de transmissão binária via sinalização polar que usa pulsos retangulares de meia largura, de amplitudes $A/2$ e $-A/2$. A taxa de dados é R_b bit/s.

- Qual é a mínima largura de banda e a potência transmitida.
- Esses dados devem ser transmitidos por pulsos retangulares M -ários de meia largura e amplitudes

$$\pm A/2, \pm 3A/2, \pm 5A/2, \dots, \pm [(M-1)/2]A$$

Repare que, para manter a mesma imunidade ao ruído, a mínima separação entre amplitudes de pulso é A . Se todos os pulsos M -ários tiverem igual probabilidade de ocorrência, mostre que a potência transmitida será

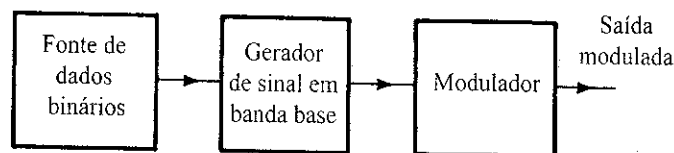
$$P = \frac{(M^2 - 1)A^2}{24 \log_2 M}$$

Determine, ainda, a largura de banda.

7.8-1 A Fig. E7.8-1 mostra um esquema de transmissão de dados binários. O gerador de sinal em banda base usa pulsos de largura completa e sinalização polar. A taxa de dados é 1 Mbit/s.

- (a) Se o modulador gerar um sinal PSK, qual é a largura de banda da saída modulada?
- (b) Se o modulador gerar FSK com diferença $f_{c1} - f_{c0} = 100$ kHz (Fig. 7.32c), determine a largura de banda do sinal modulado.

Figura E.7.8-1.



- 7.8-2** Refaça o Exercício 7.8-1 com pulsos de largura completa que satisfazem o critério de Nyquist com $r = 0,2$.
- 7.8-3** Refaça o Exercício 7.8-1 para um esquema multiamplitude com $M = 4$ (sinalização PAM com pulsos de largura completa). Em FSK [Exercício 7.8-1, parte (b)], admita que níveis sucessivos de amplitude sejam transmitidos por frequências separadas por 100 kHz.