Capítulo 6

Transmisión de datos en banda base

- Transmisión digital a través de un canal banda base limitado en frecuencia
- Señalización para canales de ancho de banda limitado
- Probabilidad de error en la detección de señales PAM
- Diseño en presencia de distorsión en el canal
- Sincronización de símbolos

En el capítulo anterior se discutió técnicas de comunicaciones digitales sobre un canal AWGN y se evaluó el desempeño del receptor óptimo en términos de probabilidad de error para varios tipos diferentes de modulación de pulsos. En este capítulo se estudiarán comunicaciones digitales sobre un canal banda base modelado con un filtro lineal como limitación de ancho de banda. Los canales con restricciones de ancho de banda más frecuentemente encontrados en la práctica son los canales telefónicos, canales de radio microondas de enlace visual y canales satelitales.

En general un canal como filtro lineal impone condiciones más restrictivas sobre el diseño de las señales de modulación. Especificamente, las señales transmitidas deben diseñarse para satisfacer la restricción de ancho de banda limitado impuesta por el canal. La restricción de ancho de banda excluye generalmente la utilización de pulsos rectangulares a la salida del modulador. En su lugar las señales transmitidas deben conformarse para restringir su ancho de banda al disponible en el canal. El diseño de señales limitadas en banda es uno de los tópicos tratados en este capítulo.

Se verá que un canal de filtro lineal distorsiona la señal transmitida. La distorsión introducida resulta en interferencia intersímbolo a la salida del modulador y conduce a un incremento en la probabilidad de error en el detector. Los dispositivos y métodos para corregir la distorsión de canal, denominados *ecualizadores*, se discutirán también.

El método básico de modulación considerado en este capítulo es PAM digital. Este tipo de modulación es particularmente adecuada para canales limitados en frecuencia debido a su alta eficiencia de ancho de banda. Otros métodos de modulación eficientes se describirán en el próximo capítulo en el contexto de metodologías de modulación pasabanda.

6.1 Transmisión digital a través de un canal banda base limitado en frecuencia

Un canal de ancho de banda limitado tal como una línea telefónica está caracterizado por un filtro lineal con respuesta impulsiva $c(t)$ y respuesta en frecuencia $C(f)$, donde

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

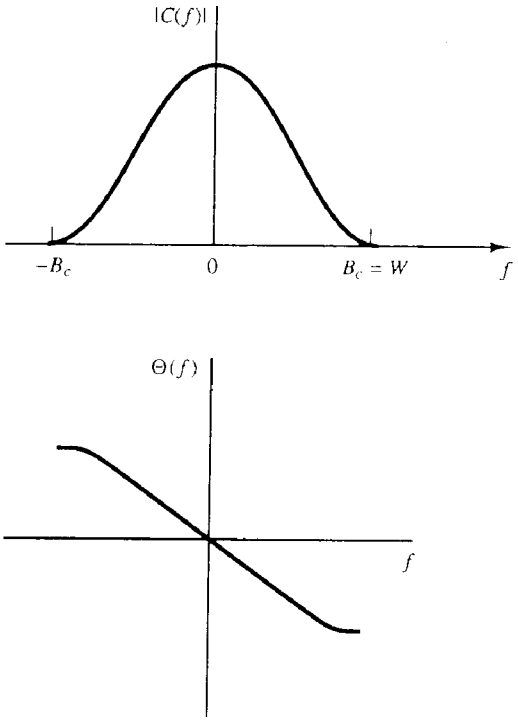

Figura 6.1: Magnitud y fase de un canal de ancho de banda limitado.

Si el canal tiene un ancho de banda limitado a B_c Hz, entonces $C(f) = 0$ para $|f| > B_c$. Cualquier componente de frecuencia a la entrada del canal que sea mayor que B_c Hz no pasará. Por esta razón es posible considerar el diseño de señales para la transmisión a través del canal que es limitado en ancho de banda a $W = B_c$ Hz como mostrado en la figura 6.1. De aqui en más W será la limitación de ancho de banda de la señal y el canal.

Es posible suponer ahora que la entrada al canal de ancho de banda limitado es una señal con forma de onda $g_T(t)$. Luego, la respuesta del canal es la convolución de $g_T(t)$ con $c(t)$, o sea

$$h(t) = \int_{-\infty}^{\infty} c(\tau)g_T(t-\tau)d\tau = c(t) * g_T(t)$$

o, expresada en el domínio frecuencia

$$H(f) = C(f)G_T(f)$$

donde $G_T(f)$ es el espectro (transformada de Fourier) de la señal $g_T(t)$ y $H(f)$ es el espectro de $h(t)$. De esta forma el canal altera o distorsiona la señal transmitida $g_T(t)$.

Es posible suponer que la señal a la salida del canal está corrompida por AWGN. Luego la señal a la entrada del demodulador es de la forma $h(t)+n(t)$, donde $n(t)$ es el AWGN. Recordando el Capítulo anterior,

un demodulador que emplea un filtro que está acoplado a la señal $h(t)$ en presencia de AWGN maximiza la relación señal a ruido (SNR) en su salida. En consecuencia, es posible pasar la señal recibida $h(t) + n(t)$ a través de un filtro con respuesta en frecuencia

$$G_R(f) = H^*(f)e^{-j2\pi t_0} \tag{6.1}$$

donde t_0 es algún retardo de tiempo nominal durante el cual se muestrea la salida del filtro.

La componente de señal a la salida del filtro acoplado en el instante de muestreo $t = t_0$ es

$$y_s(t_0) = \int_{-\infty}^{\infty} |H(f)|^2 df = \mathcal{E}_h \tag{6.2}$$

que es la energía a la salida del canal $h(t)$. La componente de ruido a la salida del filtro acoplado tiene media cero y densidad espectral de potencia

$$S_n(f) = \frac{N_0}{2}|H(f)|^2$$

En consecuencia, la potencia de ruido a la salida del filtro acoplado tiene una varianza

$$\sigma_n^2 = \int_{-\infty}^{\infty} S_n(f)df = \frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2 df = \frac{N_0\mathcal{E}_h}{2}$$

La SNR a la salida del filtro acoplado será

$$\left(\frac{S}{N}\right)_0 = \frac{\mathcal{E}_h^2}{N_0\mathcal{E}_h/2} = \frac{2\mathcal{E}_h}{N_0}$$

Este es el mismo resultado para la SNR de salida del filtro acoplado obtenido en el Capítulo anterior excepto que la energía de la señal recibida $\mathcal{E}_h$ ha reemplazado la energía de la señal transmitida $\mathcal{E}_s$. Comparado con el resultado previo la principal diferencia en este desarrollo es que la respuesta impulsiva del filtro está acoplada a la señal recibida $h(t)$ en lugar de a la señal transmitida. Notar que la implementación del filtro acoplado en el receptor requiere que $h(t)$, o equivalentemente, la respuesta impulsiva del canal $c(t)$, sea conocida en el receptor.

Ejemplo. El pulso de señal $g_T(t)$ definido por

$$g_T(t) = \frac{1}{2}\left[1 + \cos\frac{2\pi}{T}\left(t - \frac{T}{2}\right)\right], \quad 0 \le t \le T$$

se transmite a través de un canal con una respuesta en frecuencia como mostrada en la figura 6.2a. El pulso de señal se ilustra en la figura 6.2b. La salida del canal está corrompida por AWGN con densidad espectral de potencia $N_0/2$. Determinar el filtro acoplado a la señal recibida y la SNR de salida.

El problema puede resolverse más facilmente en el domínio frecuencia. Primero, el espectro del pulso de señal será

$$G_T(f) = \frac{T}{2}\frac{\sin \pi fT}{\pi fT(1 - f^2T^2)}e^{-j\pi fT} = \frac{T}{2}\frac{\mathrm{sinc}(fT)}{1 - f^2T^2}e^{-j\pi fT}$$

El espectro $|G_T(f)|^2$ se muestra en la figura 6.2c. En consecuencia,

$$H(f) = C(f)G_T(f) = \begin{cases} G_T(f), & |f| \le W \\ 0, & \text{para todo otro } f \end{cases}$$

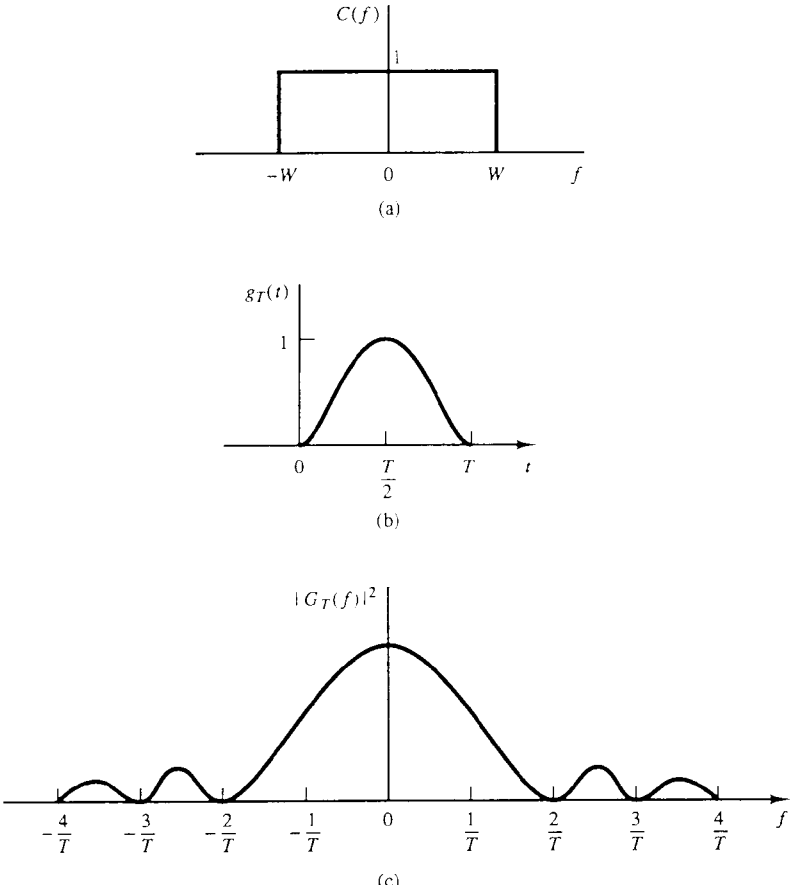

Figura 6.2: El pulso de señal mostrado en (b) transmitido por un canal ideal de ancho de banda limitado mostrado en (a). El espectro de $g_T(t)$ se muestra en (c).

Luego, la componente de energía de señal a la salida del filtro acoplado a $H(f)$ será

$$\mathcal{E}_h = \int_{-W}^{W} |G_T(f)|^2 df = \frac{1}{(2\pi)^2} \int_{-W}^{W} \frac{(\sin \pi f T)^2}{f^2(1 - f^2T^2)^2} df = \frac{T}{(2\pi)^2} \int_{-WT}^{WT} \frac{\sin^2 \pi\alpha}{\alpha^2(1-\alpha^2)^2} d\alpha$$

La varianza de la componente de ruido será

$$\sigma_n^2 = \frac{N_0}{2} \int_{-W}^{W} |G_T(f)|^2 df = \frac{N_0 \mathcal{E}_h}{2} \tag{6.3}$$

en consecuencia la SNR de salida será

$$\left(\frac{S}{N}\right)_0 = \frac{2\mathcal{E}_h}{N_0}$$

En este ejemplo se observa que la señal a la entrada del canal no está limitada en ancho de banda. En consecuencia, solo una parte de energía de la señal transmitida es recibida. La cantidad de energía de señal a la salida del filtro acoplado depende del valor del ancho de banda del canal W cuando la duración del pulso de señal es fija. El valor máximo de $\mathcal{E}_h$, obtenido cuanto $W \to \infty$ será

$$\max \mathcal{E}_h = \int_{-\infty}^{\infty} |G_T(f)|^2 df = \int_0^T g_T^2(t) dt$$

En el desarrollo anterior se consideró la transmisión y recepción de una única señal $g_T(t)$ a través de un canal de ancho de banda limitado con respuesta impulsiva $c(t)$. Se observó que el desempeño del sistema está determinado por $\mathcal{E}_h$, la energía de la señal recibida $h(t)$. Para maximizar la energía recibida es necesario asegurar que la densidad espectral de potencia de la señal transmitida se adecue a la banda de frecuencias del canal. Con este objetivo se debe estudiar la densidad espectral de potencia de la señal de entrada. Esto se discutirá en próximas secciones. El impacto de la limitación del ancho de banda del canal se refleja cuando se considera la transmisión de una secuencia de señales. Este problema se analizará en la sección siguiente.

6.1.1 Transmisión PAM en canales banda base limitados en frecuencia

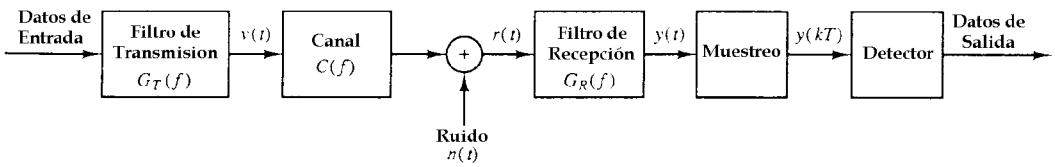

Figura 6.3: Diagrama de bloques de un sistema PAM digital.

Es posible ahora considerar el sistema de comunicaciones banda base PAM ilustrado por el diagrama funcional en bloques de la figura 6.3. El sistema consiste en un filtro de transmisión con respuesta impulsiva $g_T(t)$, el canal de filtrado lineal con AWGN, un filtro de recepción con respuesta impulsiva $g_R(t)$, un muestreador que genera muestras periodicamente a la salida del filtro receptor y un detector de símbolos. El muestreador requiere la extracción del temporizado de la señal recibida. Esta señal de temporizado sirve como un reloj que especifica los instantes de tiempo apropiados para muestrear la salida del filtro de recepción. La extracción de la señal de temporizado se considerará en la última sección de este Capítulo.

En general es posible considerar las comunicaciones digitales por medio de PAM M-ário. En consecuencia, la secuencia de datos binarios de entrada se subdivide en símbolos de k bits, y cada símbolo se mapea en un nivel de amplitud correspondiente que modula la salida del filtro de transmisión. La señal a la salida del filtro de transmisión (la entrada al canal) puede expresarse como

$$v(t) = \sum_{n=-\infty}^{\infty} a_n g_T(t - nT)$$

donde $T = k/R_b$ es el intervalo de símbolo ($1/T = R_b/k$ es la velocidad de transmisión de símbolos), R_b es la velocidad de transmisión (en bits/seg) y $\{a_n\}$ es la secuencia de niveles de amplitud correspondientes a la secuencia de bloques de k bits de información.

La salida del canal, recibida en el demodulador, puede expresarse como

$$r(t) = \sum_{n=-\infty}^{\infty} a_n h(t - nT) + n(t)$$

donde $h(t)$ es la respuesta impulsiva de la cascada del filtro transmisor y el canal, o sea $h(t) = c(t) * g_T(t)$, $c(t)$ es la respuesta impulsiva del canal y $n(t)$ representa el AWGN.

La señal recibida pasa a través de un filtro lineal de recepción con respuesta impulsiva $g_R(t)$ y respuesta en frecuencia $G_R(f)$. Si $g_R(t)$ está acoplada a $h(t)$ entonces la SNR es máxima en el instante de muestreo apropiado. La salida del filtro de recepción puede expresarse como

$$y(t) = \sum_{n=-\infty}^{\infty} a_n x(t - nT) + \nu(t)$$

donde $x(t) = h(t) * g_R(t) = g_T(t) * c(t) * g_R(t)$ y $\nu(t) = n(t) * g_R(t)$ es el ruido aditivo a la salida del filtro de recepción.

Para recuperar los símbolos de información $\{a_n\}$, la salida del filtro de recepción se muestrea periodicamente cada T segundos. De esta forma

$$y(mT) = \sum_{n=-\infty}^{\infty} a_n x(mT - nT) + \nu(mT)$$

o equivalentemente

$$y_m = \sum_{n=-\infty}^{\infty} a_n x_{m-n} + \nu_m = x_0 a_m + \sum_{n \neq m} a_n x_{m-n} + \nu_m \tag{6.4}$$

donde $x_m = x(mT)$, $\nu_m = \nu(mT)$ y $m = 0, \pm 1, \pm 2, \cdots$. Se utilizará como reloj una señal de temporizado extraída de la señal recibida, según se discutirá en la última sección de este Capítulo.

El primer término de la derecha de (6.4) es el símbolo deseado a_m, escalado por un parámetro de ganancia x_0. Cuando el filtro de recepción está acoplado a la señal recibida $h(t)$, el factor de escala será

$$x_0 = \int_{-\infty}^{\infty} h^2(t)dt = \int_{-\infty}^{\infty} |H(f)|^2 df = \int_{-W}^{W} |G_T(f)|^2 |C(f)|^2 dt = \mathcal{E}_h$$

como indicado por el desarrollo de (6.1) y (6.2). El segundo término de la derecha de (6.4) representa el efecto de los otros símbolos en el instante de muestreo $t = mT$, y se denomina *interferencia intersímbolo* (ISI). Este es un término indeseable que degrada el desempeño del sistema de comunicaciones digitales PAM. Finalmente, el tercer término, ν_m que representa el ruido aditivo es una variable aleatória Gaussiana de media cero y varianza $\sigma_\nu^2 = N_0 \mathcal{E}_h/2$, previamente obtenida en (6.3).

Mediante el diseño apropiado de los filtros de transmisión y recepción es posible satisfacer la condición $x_n = 0$ para $n \neq 0$, de forma que el término de ISI desaparezca. En este caso, el único término que puede producir errores en la secuencia digital recibida es el ruido aditivo.

En secciones próximas se considerará el diseño de los filtros de transmisión y recepción que eliminan y/o controlan la ISI. Primero, sin embargo, se determinará la densidad espectral de potencia de la señal PAM transmitida. De esta forma, se establecerá una relación entre las características espectrales de la señal PAM y los requerimientos de ancho de banda del canal para transmitir la señal.

6.1.2 Espectro de potencia de señales PAM

La señal transmitida para el caso de PAM digital tiene la siguiente forma general

$$v(t) = \sum_{n=-\infty}^{\infty} a_n g_T(t - nT)$$

donde $\{a_n\}$ es la secuencia de amplitudes correspondientes a los símbolos de información de la fuente y $g_T(t)$ es la respuesta impulsiva del filtro de transmisión. Dado que la secuencia de información $\{a_n\}$ es aleatória, $v(t)$ es una función muestra de un proceso estocástico $V(t)$. En esta sección se evaluará la densidad espectral de potencia de $V(t)$. El procedimiento a seguir será obtener primero la función autocorrelación de $V(t)$ y luego su transformada de Fourier.

Primero, el valor medio de $V(t)$ es

$$E[V(t)] = \sum_{n=-\infty}^{\infty} E[a_n] g_T(t - nT) = m_a \sum_{n=-\infty}^{\infty} g_T(t - nT)$$

donde m_a es el valor de amplitud media de la secuencia aleatória $\{a_n\}$. Notar que a pesar que m_a es una constante, el término $\sum_{n=-\infty}^{\infty} g_T(t - nT)$ es una función periódica con período T. En consecuencia, el valor medio de $V(t)$ es periódico con período T.

La función autocorrelación de $V(t)$ será

$$R_V(t+\tau,t) = E[V(t)V(t+\tau)] = \sum_{n=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} E[a_n a_l] g_T(t-nT) g_T(t+\tau-lT) \tag{6.5}$$

Se supone en general que la secuencia de información $\{a_n\}$ es estacionaria en sentido amplio (ESA) con secuencia de autocorrelación

$$R_a(n) = E[a_m a_{n+m}]$$

de forma que (6.5) puede expresarse como

$$\begin{aligned} R_V(t+\tau,t) &= \sum_{n=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} R_a(n-l) g_T(t-nT) g_T(t+\tau-lT) \\ &= \sum_{m=-\infty}^{\infty} R_a(m) \sum_{n=-\infty}^{\infty} g_T(t-nT) g_T(t+\tau-nT-mT) \end{aligned} \tag{6.6}$$

Es posible notar que la segunda sumatoria de esta ecuación es periódica con período T. En consecuencia, la función autocorrelación $R_V(t+\tau,t)$ es periódica en la variable t, o sea

$$R_V(t+T+\tau,t+T) = R_V(t+\tau,t)$$

Entonces, el proceso estocástico $V(t)$ tiene media y autocorrelación periódicas y se denomina por ello proceso *cicloestacionario.*

La densidad espectral de potencia de un proceso cicloestacionario puede determinarse promediando primero la función autocorrelación $R_V(t+\tau,t)$ sobre un período T y luego calculando la transformada de Fourier de la función autocorrelación promedio. De esta forma se tiene que

$$\begin{aligned} \tilde{R}_V(\tau) &= \frac{1}{T}\int_{-T/2}^{T/2} R_V(t+\tau,t)dt \\ &= \sum_{m=-\infty}^{\infty} R_a(m) \sum_{n=-\infty}^{\infty} \frac{1}{T}\int_{-T/2}^{T/2} g_T(t-nT) g_T(t+\tau-nT-mT)dt \\ &= \sum_{m=-\infty}^{\infty} R_a(m) \sum_{n=-\infty}^{\infty} \frac{1}{T}\int_{nT-T/2}^{nT+T/2} g_T(t) g_T(t+\tau-mT)dt \\ &= \frac{1}{T}\sum_{m=-\infty}^{\infty} R_a(m) \int_{-\infty}^{\infty} g_T(t) g_T(t+\tau-mT)dt \end{aligned} \tag{6.7}$$

Es posible interpretar la integral de (6.7) como la función autocorrelación temporal de $g_T(t)$ y definirla como

$$R_g(\tau) = \int_{-\infty}^{\infty} g_T(t) g_T(t+\tau)dt \tag{6.8}$$

Con esta definición la función autocorrelación promedio de $V(t)$ toma la forma

$$\tilde{R}_V(\tau) = \frac{1}{T}\sum_{m=-\infty}^{\infty} R_a(m) R_g(\tau-mT)$$

De esta expresión es posible observar que $\tilde{R}_V(\tau)$ tiene la forma de una sumatoria de convolución. En consecuencia la transformada de Fourier será

$$\begin{aligned} S_V(f) &= \int_{-\infty}^{\infty} \tilde{R}_V(\tau) e^{-j2\pi f\tau} d\tau \\ &= \frac{1}{T} \sum_{m=-\infty}^{\infty} R_a(m) \int_{-\infty}^{\infty} R_g(\tau - mT) e^{-j2\pi f\tau} d\tau = \frac{1}{T} S_a(f) |G_T(f)|^2 \end{aligned} \tag{6.9}$$

donde $S_a(f)$ es el espectro de potencia de la secuencia de información $\{a_n\}$, definida como

$$S_a(f) = \sum_{-\infty}^{\infty} R_a(m) e^{-j2\pi f m T} \tag{6.10}$$

y $G_T(f)$ es el espectro del filtro de transmisión. $|G_T(f)|^2$ es la transformada de Fourier de $R_g(\tau)$.

El resultado de (6.9) ilustra la dependencia en la densidad espectral de potencia $S_V(f)$ de la señal transmitida de: 1) las características espectrales del filtro de transmisión $G_T(f)$ y 2) las características espectrales de la secuencia de información $S_a(f)$. Tanto $G_T(f)$ como $S_a(f)$ puede diseñarse para controlar la forma de la densidad espectral de potencia de la señal transmitida.

Mientras la dependencia de $S_V(f)$ debido a $G_T(f)$ se entiende facilmente, el efecto de las propiedades de la autocorrelación de la secuencia de información $\{a_n\}$ es más sutil. Primero, es posible observar que para una autocorrelación arbitrária $R_a(m)$, la densidad espectral correspondiente $S_a(f)$ es periódica en frecuencia con período $1/T$. En realidad, se concluye de (6.10) que $S_a(f)$ tiene la forma de una serie de Fourier con $\{R_a(m)\}$ como los coeficientes correspondientes. En consecuencia, la secuencia de autocorrelación $\{R_a(m)\}$ es simplemente

$$R_a(m) = T \int_{-1/2T}^{1/2T} S_a(f) e^{j2\pi f m T} df$$

Segundo, es posible considerar el caso en el cual los símbolos en la secuencia $\{a_n\}$ son mutuamente no correlacionados. Entonces

$$R_a(m) = \begin{cases} \sigma_a^2 + m_a^2, & m = 0 \\ m_a^2, & m \neq 0 \end{cases}$$

donde $\sigma_a^2 = E[a_n^2] - m_a^2$ es la varianza de un símbolo de información. Sustituyendo $R_a(m)$ en (6.10) se obtiene la densidad espectral de potencia

$$S_a(f) = \sigma_a^2 + m_a^2 \sum_{m=-\infty}^{\infty} e^{-j2\pi f m T}$$

El término de la sumatoria de la derecha de esta ecuación es periódico con período $1/T$. Puede verse como la serie de Fourier de un tren periódico de impulsos donde cada impulso tiene un área $1/T$. En consecuencia, esta última ecuación puede expresarse como

$$S_a(f) = \sigma_a^2 + \frac{m_a^2}{T} \sum_{m=-\infty}^{\infty} \delta\left(f - \frac{m}{T}\right)$$

Sustituyendo ahora en la expresión de $S_V(f)$ dada por (6.9) se obtiene el resultado deseado para la densidad espectral de potencia de la señal transmitida $V(t)$ cuando la secuencia de símbolos de información no está correlacionada. O sea,

$$S_V(f) = \frac{\sigma_a^2}{T} |G_T(f)|^2 + \frac{m_a^2}{T} \sum_{m=-\infty}^{\infty} \left| G_T\left(\frac{m}{T}\right) \right|^2 \delta\left(f - \frac{m}{T}\right) \tag{6.11}$$

Esta expresión para la densidad espectral de potencia de la señal transmitida está separada adrede en dos términos para enfatizar los dos tipos diferentes de componentes espectrales. El primer término $\sigma_a^2|G_T(f)|^2/T$ es el espectro contínuo y su forma depende de $G_T(f)$. El segundo término de (6.11) consiste en componentes de frecuencia discretas espaciadas $1/T$. Cada línea espectral tiene una potencia proporcional a $|G_T(f)|^2$ evaluada en $f = m/T$. Es posible notar que las componentes de frecuencia discreta pueden eliminarse haciendo que la secuencia de símbolos de información tenga media cero. Esta condición se fuerza usualmente en los métodos de modulación digital porque las líneas espectrales discretas no son deseables. Para ser específico, la media m_a en PAM digital se hace facilmente cero seleccionando los niveles de amplitud de forma que sean igualmente probables posicionandolos simetricamente sobre el eje real. Bajo esta condición ($m_a = 0$) se tiene que

$$S_V(f) = \frac{\sigma_a^2}{T}|G_T(f)|^2 \tag{6.12}$$

De esta forma el diseñador del sistema puede controlar las características espectrales de la señal PAM digital transmitida. El ejemplo siguiente ilustra la conformación espectral resultante a partir de $g_T(t)$.

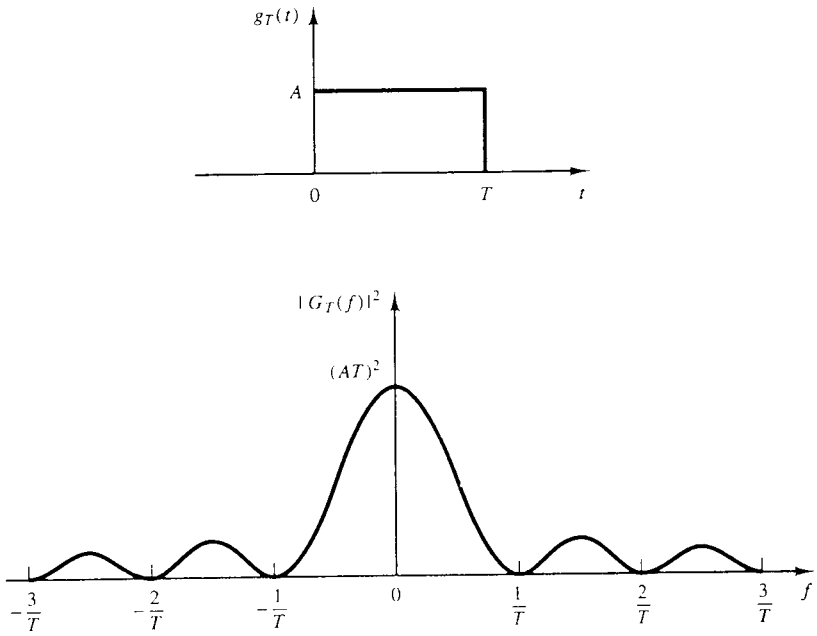

Figura 6.4: Un pulso rectangular $g_T(t)$ y su densidad espectral de energía $|G_T(f)|^2$.

Ejemplo. Determinar la densidad espectral de potencia en (6.11) cuando $g_T(t)$ es un pulso rectangular como mostrado en la figura 6.4a.
La transformada de Fourier de $g_T(t)$ es

$$G_T(f) = AT\frac{\sin \pi fT}{\pi fT}e^{-j\pi fT}$$

En consecuencia,

$$|G_T(f)|^2 = (AT)^2\left(\frac{\sin \pi fT}{\pi fT}\right)^2 = (AT)^2\text{sinc}^2(fT)$$

Este espectro se ilustra en la figura 6.4b. Es posible notar que posee ceros en frecuencia en múltiplos de $1/T$ y que decae inversamente con el cuadrado de la frecuencia. Como consecuencia de los ceros en frecuencia de

$G_T(f)$ solo una única componente de (6.11) no desaparece. De esta forma, sustituyendo $|G_T(f)|^2$ en (6.11) se obtiene

$$S_V(f) = \sigma_a^2 A^2 T \left(\frac{\sin \pi f T}{\pi f T}\right)^2 + A^2 m_a^2 \delta(f) = \sigma_a^2 A^2 T \mathrm{sinc}^2(fT) + A^2 m_a^2 \delta(f)$$

El ejemplo siguiente ilustra la conformación espectral que puede lograrse mediante operaciones realizadas sobre la secuencia de información de entrada.

Ejemplo. Considerar una secuencia binaria $\{b_n\}$ a partir de la cual se forman los símbolos $a_n = b_n + b_{n-1}$. Los $\{b_n\}$ se suponen variables aleatórias no correlacionadas con valores ± 1, cada una de media cero y varianza unitaria. Determinar la densidad espectral de potencia de la señal transmitida.
La función autocorrelación de la secuencia $\{a_n\}$ será

$$R_a(m) = E[a_n a_{n+m}] = E[(b_n + b_{n-1})(b_{n+m} + b_{n+m-1})] = \begin{cases} 2 & m = 0 \\ 1 & m = \pm 1 \\ 0 & \text{para todo otro } m \end{cases}$$

Dado que la densidad espectral de potencia de la secuencia de entrada es

$$S_a(f) = 2(1 + \cos 2\pi f T) = 4 \cos^2 \pi f T$$

y el espectro de potencia correspondiente para la señal modulada es, de (6.9),

$$S_V(f) = \frac{4}{T} |G_T(f)|^2 \cos^2 \pi f T$$

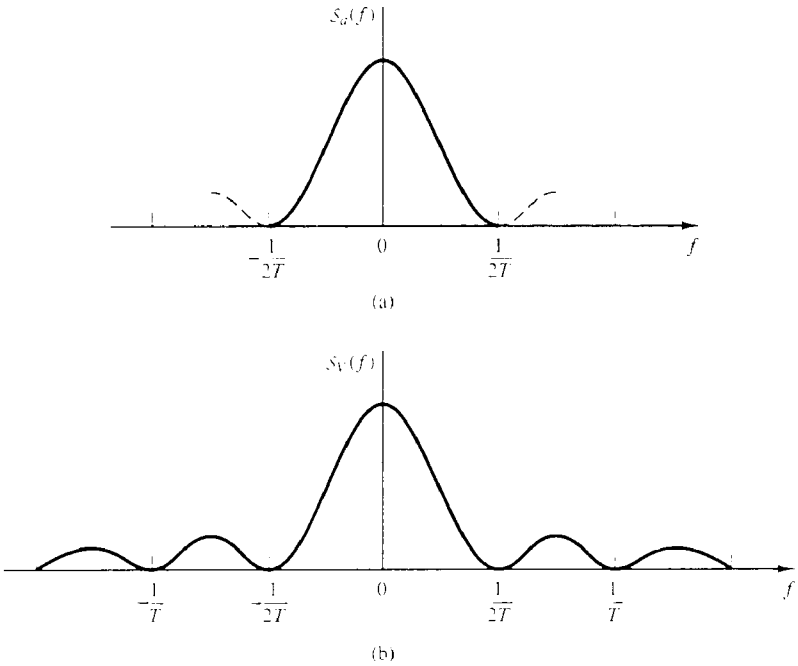

Figura 6.5: Densidad espectral de potencia para (a) la secuencia de información y (b) la señal modulada PAM.

La figura 6.5 ilustra la densidad espectral de potencia $S_a(f)$ de la secuencia de entrada y el correspondiente $S_V(f)$ cuando $G_T(f)$ es el espectro del pulso rectangular.

Como demostrado en el ejemplo, el espectro de la señal trasmitida puede conformarse por medio de una secuencia $\{a_n\}$ correlacionada a la entrada del modulador. Teniendo en cuenta del Capítulo anterior que los códigos de modulación dependientes de estados producen generalmente secuencias correlacionadas, esto provee una forma de conformación espectral. Los códigos de longitud de corrida limitada dependientes de los estados se utilizan frecuentemente con este propósito. En la sección siguiente se considerará las características espectrales de las señales moduladas producidas por esos códigos.

6.1.3 Espectro de potencia de señales con memoria

En el Capítulo anterior se demostró que los códigos de modulación dependientes de los estados resultan en señales moduladas con memoria. Tales señales fueron descriptas por cadenas de Markov, las cuales son basicamente grafos que incluyen los "estados" posibles del modulador con las probabilidades de estado correspondientes $\{p_i\}$ y transiciones de estado con las correspondientes probabilidades de transición $\{p_{ij}\}$.

La densidad espectral de potencia de señales moduladas digitalmente caracterizadas por cadenas de Markov (ver Capítulo 2) puede obtenerse siguiendo el procedimiento básico de la sección anterior. De esta forma, es posible determinar la función autocorrelación y luego evaluar la transformada de Fourier para obtener la densidad espectral de potencia. Para señales que se generan mediante una cadena de Markov con matriz de probabilidades de transición $\boldsymbol{P}$, la densidad espectral de potencia puede expresarse en la siguiente forma general

$$\begin{aligned} S(f) &= \sum_{n=-\infty}^{\infty} \left| \sum_{i=1}^{K} p_i S_i\left(\frac{n}{T}\right) \delta\left(f - \frac{n}{T}\right) \right|^2 + \frac{1}{T} \sum_{i=1}^{K} p_i |S_i'|^2 \\ &\quad + \frac{2}{T} \mathrm{Re}\left[\sum_{i=1}^{K} \sum_{j=1}^{K} p_i S_i'^*(f) S_j'(f) P_{ij}(f) \right] \end{aligned} \tag{6.13}$$

donde K es el número de estados del modulador, $S_i'(f)$ es la transformada de Fourier de la señal $s_i'(t)$, donde $s_i' = s_i(t) - \sum_{k=1}^{K} p_k s_k(t)$, y $P_{ij}(f)$ es la transformada de Fourier de la secuencia discreta en tiempo $p_{ij}(n)$, definida como

$$P_{ij}(f) = \sum_{n=1}^{\infty} p_{ij}(n) e^{-j2\pi n f T}$$

El término $p_{ij}(n)$ es la probabilidad que una señal $s_j(t)$ sea transmitida n intervalos de señalización después de la transición de $s_i(t)$. En consecuencia, $\{p_{ij}(n)\}$ son las probabilidades de transición en la matriz de transición $\boldsymbol{P}^n$. Notar que cuando $n = 1$, $\boldsymbol{P}^n = \boldsymbol{P}$, que es la matriz discutida en el Capítulo anterior.

Ejemplo. Determinar la densidad espectral de potencia de una señal NRZI.
La señal NRZI está caracterizada por la matriz de transición

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Notar que $\boldsymbol{P}^n = \boldsymbol{P}$ para todo $n > 1$. En consecuencia, con $K = 2$ estados y $g_T(t) = s_1(t) = -s_2(t)$, se obtiene

$$S(f) = \frac{(2p-1)^2}{T^2} \sum_{n=-\infty}^{\infty} \left| G_T\left(\frac{n}{T}\right) \right|^2 \delta\left(f - \frac{n}{T}\right) + \frac{4p(1-p)}{T} |G_T(f)|^2$$

donde para un pulso rectangular de amplitud A,

$$|G_T(f)|^2 = (AT)^2 \left(\frac{\sin \pi fT}{\pi fT}\right)^2 = (AT)^2 \text{sinc}^2(fT)$$

Es posible notar que cuando $p = 1/2$ (señales igualmente probables) el impulso en el espectro desaparece y $S(f)$ se reduce a

$$S(f) = \frac{1}{T}|G_T(f)|^2$$

Es posible observar también que el espectro de potencia de la señal NRZI para señales igualmente probables es idéntico a la expresión en (6.12), lo cual se aplica a una secuencia $\{a_n\}$ no correlacionada en el modulador. En consecuencia, es posible concluir que la simple operación de precodificación NRZI no resulta en una secuencia correlacionada.

Ejemplo. Determinar la densidad espectral de potencia de una señal modulada en retardo (codificación Miller).
La matriz de transición de una señal modulada en retardo es

$$\boldsymbol{P} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

Las probabilidades de estado son $p_i = 1/4$, para $i = 1, 2, 3, 4$. Las potencias de $\boldsymbol{P}$ se obtienen facilmente mediante el uso de la relación $\boldsymbol{P}^4\gamma = -\frac{1}{4}\gamma$, donde γ es la matriz correlación de la señal con elementos

$$\gamma_{ij} = \frac{1}{T}\int_0^T s_i(t)s_j(t)dt$$

y las cuatro señales $\{s_i(t), i = 1, 2, 3, 4\}$ se muestran en la figura 5.28, donde $s_3(t) = -s_2(t)$ y $s_4(t) = -s_1(t)$. Puede obtenerse facilmente que

$$\gamma = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

En consecuencia, las potencias de $\boldsymbol{P}$ pueden generarse de la relación

$$\boldsymbol{P}^{k+4}\gamma = -\frac{1}{4}\boldsymbol{P}^k\gamma, \quad k \geq 1$$

Con la ayuda de esas relaciones, la densidad espectral de potencia de una señal modulada en retardo se obtiene de (6.13), y puede expresarse en la siguiente forma

$$\begin{aligned} S(f) &= \frac{1}{2(\pi fT)^2(17 + 8\cos 8\pi fT)}[23 - 2\cos \pi fT - 22\cos 2\pi fT - 12\cos 3\pi fT + 5\cos 4\pi fT \\ &\quad +12\cos 5\pi fT + 2\cos 6\pi fT - 8\cos 7\pi fT + 2\cos 8\pi fT] \end{aligned}$$

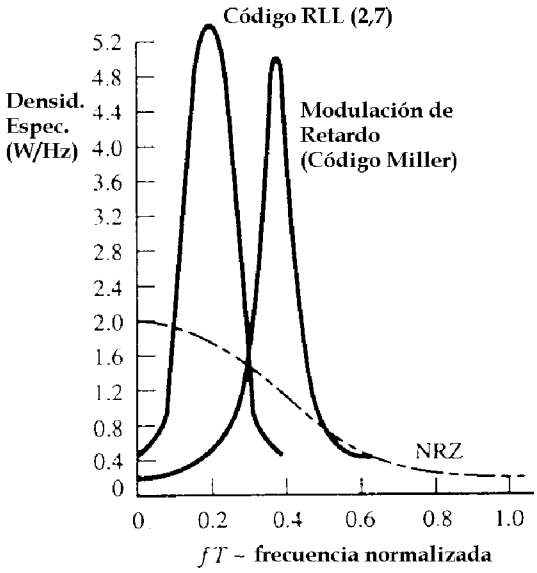

Figura 6.6: Densidad espectral de potencia de la codificación Miller (modulación de retardo), y señales banda base NRZ / NRZI.

La densidad espectral de potencia de la señal NRZI y de la señal modulada en retardo se muestran en la figura 6.6. Se observa que la señal NRZI tiene un espectro de potencia pasabajos con un pico en $f = 0$. Por otro lado, la señal modulada en retardo tiene muy poca potencia en la vecindad de $f = 0$. También tiene una densidad espectral de potencia relativamente estrecha. Esas dos características la hacen particularmente adecuada para su utilización en canales de almacenamiento de grabaciones magnéticas que utilizan cabezales sensibles al flujo para escribir o leer de un disco. También se incluye para comparación en la figura 6.6 la densidad espectral de potencia de una señal modulada generada por un código $(d, \kappa) = (2, 7)$ seguida por NRZI. El código de corrida limitada $(2, 7)$, previamente presentado en la tabla 5.3, se utiliza también en canales de grabación magnética.

Cuando no existe memoria en el método de modulación, la señal transmitida en cada intervalo de señalización es independiente de las formas de onda transmitidas en intervalos de señalización previos. La densidad espectral de potencia de la señal resultante puede aún expresarse en la forma de (6.13) si la matriz de transición de probabilidades es sustituida por

$$\boldsymbol{P} = \begin{bmatrix} p_1 & p_2 & \cdots & p_K \\ p_1 & p_2 & \cdots & p_K \\ \vdots & \vdots & & \vdots \\ p_1 & p_2 & \cdots & p_K \end{bmatrix}$$

y se impone la condición que $\boldsymbol{P}^n = \boldsymbol{P}$ para todo $n > 1$ como consecuencia de la modulación sin memoria. Bajo esas condiciones, la expresión para $S(f)$ dada por (6.13) se reduce a la forma más simple

$$\begin{aligned} S(f) &= \frac{1}{T^2} \sum_{n=-\infty}^{\infty} \left| \sum_{i=1}^{K} p_i S_i \left(\frac{n}{T}\right) \delta \left(f - \frac{n}{T}\right) \right|^2 + \frac{1}{T} \sum_{i=1}^{K} p_i (1 - p_i) |S_i'|^2 \\ &\quad + \frac{2}{T} \sum_{i=1, i<j}^{K} \sum_{j=1}^{K} p_i p_j \mathrm{Re}[S_i(f) S_j^*(f)] P_{ij}(f) \end{aligned} \tag{6.14}$$

Se observa que el resultado previo para la densidad espectral de potencia de PAM sin memoria dado por (6.11) es un caso especial de la expresión en (6.14). Especificamente, si se tienen $K = M$ señales, las

cuales son versiones escaladas en amplitud de un pulso básico $g_T(t)$, o sea $s_m(t) = a_m g_T(t)$, donde a_m es la amplitud de la señal, entonces $S_m(f) = a_m G_T(f)$ y (6.14) se convierte en

$$\begin{aligned} S(f) &= \frac{1}{T^2}\left(\sum_{i=1}^{M} a_i p_i\right)^2 \sum_{n=-\infty}^{\infty} \left|G_T\left(\frac{n}{T}\right)\right|^2 \delta\left(f-\frac{n}{T}\right) \\ &\quad + \frac{1}{T}\left[\sum_{i=1}^{M} p_i(1-p_i)a_i^2 - 2\sum_{i=1,i<j}^{M}\sum_{j=1}^{M} p_i p_j a_i a_j\right] |G_T(f)|^2 \end{aligned} \tag{6.15}$$

Si se compara (6.11) con (6.15), las dos expresiones son idénticas cuando el valor medio y la varianza de la secuencia de información toman la forma

$$\begin{aligned} m_a &= \sum_{i=1}^{M} a_i p_i \\ \sigma_a^2 &= \sum_{i=1}^{M} p_i(1-p_i)a_i^2 - 2\sum_{i=1,i<j}^{M}\sum_{j=1}^{M} p_i p_j a_i a_j \end{aligned}$$

En consecuencia (6.14) es la forma más general para la densidad espectral de potencia de señales moduladas sin memoria, dado que se aplica a señales que pueden tener diferentes formas de pulso.

6.2 Señalización para canales de ancho de banda limitado

De la sección anterior es posible recordar que la salida del filtro de transmisión para un sistema de comunicaciones PAM puede expresarse como

$$v(t) = \sum_{n=-\infty}^{\infty} a_n g_T(t-nT)$$

donde la salida del canal, o sea la señal recibida por el demodulador, puede expresarse como

$$r(t) = \sum_{n=-\infty}^{\infty} a_m h(t-nT) + n(t)$$

donde $h(t) = c(t) * g_T(t)$, $c(t)$ es la respuesta impulsiva del canal, $g_T(t)$ es la respuesta impulsiva del filtro de transmisión y $n(t)$ es una función muestra de un ruido AWGN.

En esta sección se considerará el problema de diseño de un filtro de transmisión de banda limitada. El diseño considerará en principio que no existe distorsión en el canal. Posteriormente se considerará el problema de diseño del filtro cuando el canal distorsiona la señal transmitida. Dado que $H(f) = C(f)G_T(f)$, la condición para transmisión sin distorsión es que la respuesta en frecuencia característica $C(f)$ del canal tenga magnitud constante y fase lineal sobre el ancho de banda de la señal transmitida. O sea,

$$C(f) = \begin{cases} C_0 e^{-j2\pi f t_0} & |f| \le W \\ 0, & |f| > W \end{cases}$$

donde W es el ancho de banda disponible de canal, t_0 representa un retardo arbitrário finito, definido por conveniencia igual a cero, y C_0 es un factor de ganancia constante, definido igual a uno por conveniencia. De esta forma, bajo la condición que el canal no introduzca distorsión, $H(f) = G_T(f)$ para $|f| \le W$ y cero para $|f| > W$. En consecuencia, el filtro acoplado tiene una respuesta en frecuencia $H^*(f) = G_T^*(f)$ y su salida en instantes de muestreo periódicos $t = mT$ tiene la forma

$$y(mT) = x(0)a_m + \sum_{n\neq m} a_n x(mT-nT) + \nu(mT)$$

o más simplemente

$$y_m = x_0 a_m + \sum_{n \neq m} a_n x_{m-n} + \nu_m \tag{6.16}$$

donde $x(t) = g_T(t) * g_R(t)$ y $\nu(t)$ es la respuesta de salida del filtro acoplado a un proceso de entrada AWGN $n(t)$.

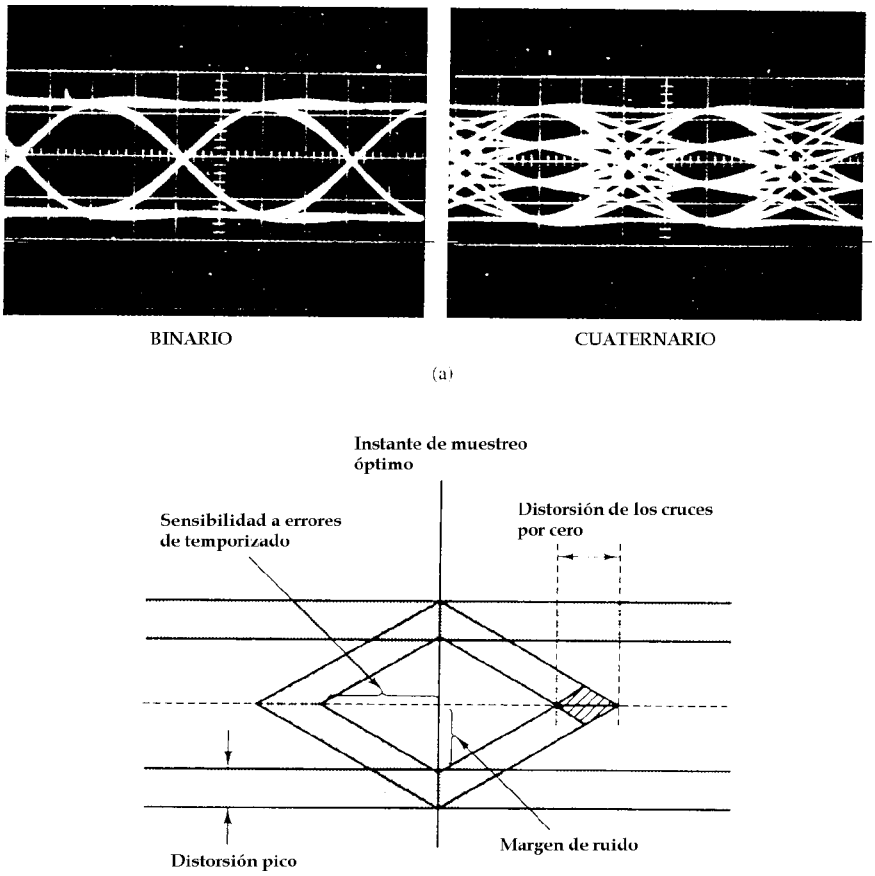

Figura 6.7: Patrón de ojo. (a) Ejemplos de patrones de ojo para ASK binaria y cuaternaria. (b) Efectos de la ISI sobre la apertura del ojo.

El término central de la derecha de (6.16) representa la ISI. La cantidad de ISI y ruido que están presentes en un sistema PAM digital pueden observarse en un osciloscopio. Especificamente, es posible mostrar la señal recibida sobre la entrada vertical con el barrido horizontal a una frecuencia $1/T$. La figura resultante en el osciloscopio se denomina *patrón de ojo* por su semejanza con el ojo humano. En la figura 6.7a se ilustran dos ejemplos de patrón de ojo para PAM binaria y PAM cuaternaria.

El efecto de la ISI es que el ojo se cierre, reduciendo de esa forma el márgen para que el ruido aditivo cause errores. La figura 6.7b ilustra el efecto de la ISI al reducir la apertura del ojo. Notar que la ISI distorsiona la posición de los cruces por cero y causa una reducción del ojo. En consecuencia, el sistema es más sensible a un error de sincronización y exibe un menor márgen contra el ruido aditivo.

A continuación se considerará el problema de diseño de señal bajo dos condiciones: 1) que no exista ISI en los instantes de muestreo y 2) que se permita una cantidad de ISI controlada.

6.2.1 Diseño de señales de ancho de banda limitado para ISI nula

Como indicado antes, la transformada de Fourier de la señal de salida del filtro receptor de un sistema general de comunicaciones digitales está dada por $X(f) = G_T(f)C(f)G_R(f)$ donde $G_T(f)$ y $G_R(f)$ son las

respuestas en frecuencia de los filtros de transmisión y recepción y $C(f)$ es la respuesta en frecuencia del canal. Se vió también que la salida del filtro receptor, muestreada en $t = mT$, está dada por

$$y_m = x(0)a_m + \sum_{n=-\infty, n\neq m}^{\infty} x(mT - nT)a_n + \nu(mT)$$

Para remover el efecto de la ISI es necesario y suficiente que $x(mT - nT) = 0$ para $n \neq m$ y $x(0) \neq 0$, donde sin pérdida de generalidad se supondrá que $x(0) = 1$. Esto significa que el sistema de comunicaciones completo se ha diseñado tal que

$$x(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

En esta sección se obtendrán condiciones necesarias y suficientes sobre $X(f)$ tal que $x(t)$ satisfaga la condición anterior. Esta condición se conoce como *Critério de forma de pulso de Nyquist* o *Condición de Nyquist para ISI cero* y se sintetiza en el siguiente teorema.

Teorema (Nyquist). La condición necesaria y suficiente para que $x(t)$ satisfaga

$$x(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \tag{6.17}$$

es que su transformada de Fourier satisfaga

$$\sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right) = T \tag{6.18}$$

Prueba: Si $x(t)$ es la transformada inversa de $X(f)$, entonces

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}\,df$$

En los instantes de muestreo $t = nT$ se tiene que

$$x(nT) = \int_{-\infty}^{\infty} X(f)e^{j2\pi fnT}\,df \tag{6.19}$$

Es posible escribir (6.19) como una suma infinita de integrales cubriendo el rango finito de $1/T$, tal que

$$\begin{aligned} x(nT) &= \sum_{m=-\infty}^{\infty} \int_{(2m-1)/2T}^{(2m+1)/2T} X(f)e^{j2\pi fnT}\,df \\ &= \sum_{m=-\infty}^{\infty} \int_{-1/2T}^{1/2T} X\left(f + \frac{m}{T}\right) e^{j2\pi fnT}\,df \\ &= \int_{-1/2T}^{1/2T} \left[\sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right)\right] e^{j2\pi fnT}\,df \\ &= \int_{-1/2T}^{1/2T} Z(f)e^{j2\pi fnT}\,df \end{aligned} \tag{6.20}$$

donde $Z(f)$ se ha definido como

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

Obviamente $Z(f)$ es una función periódica con período $1/T$, y en consecuencia puede expandirse en términos de los coeficientes $\{z_n\}$ de su serie de Fourier como

$$Z(f) = \sum_{n=-\infty}^{\infty} z_n e^{j2\pi nfT} \tag{6.21}$$

donde

$$z_n = T\int_{-1/2T}^{1/2T} Z(f)e^{-j2\pi nfT}df \tag{6.22}$$

Comparando (6.22) con (6.20) se obtiene que

$$z_n = Tx(-nT) \tag{6.23}$$

En consecuencia, la condición necesaria y suficiente para que (6.17) sea satisfecha es que

$$z_n = \begin{cases} T, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

lo cual, cuando sustituido en (6.21), produce

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

con lo que se concluye la prueba.

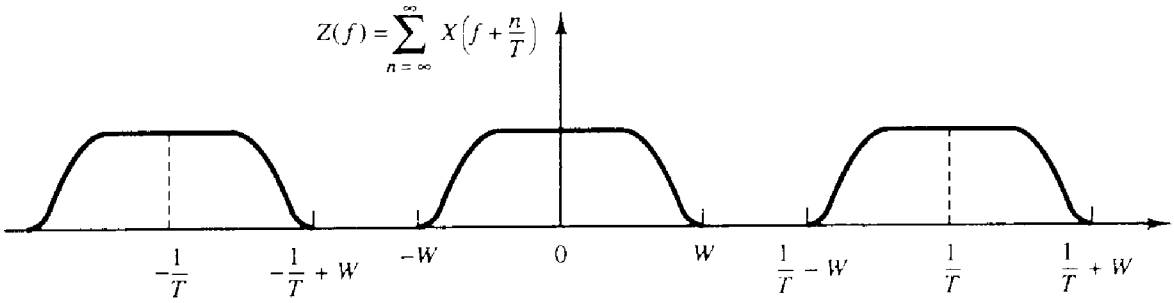

Figura 6.8: Gráfico de $Z(f)$ para el caso $T < 1/2W$.

Suponiendo ahora un canal con ancho de banda W, $C(f) = 0$ para $|f| > W$ y en consecuencia, $X(f) = 0$ para $|f| > W$, es posible distinguir los siguientes casos

1. $T < \frac{1}{2W}$, o equivalentemente $\frac{1}{T} > 2W$. Como $Z(f)$ está formada por réplicas no solapadas de $X(f)$, separadas por $1/T$ como mostrado en la figura 6.8, no existe elección que permita asegurar $Z(f) = T$ en este caso y en consecuencia no existirá forma de diseñar un sistema sin ISI.

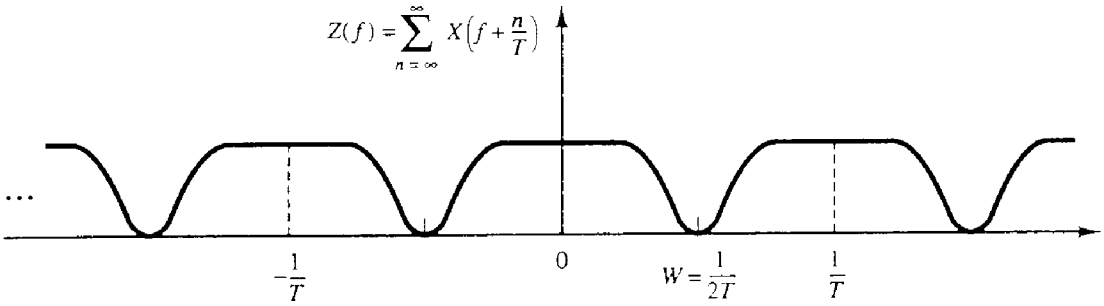

Figura 6.9: Gráfico de $Z(f)$ para el caso $T = 1/2W$.

2. $T = \frac{1}{2W}$, o equivalentemente $\frac{1}{T} = 2W$ (la frecuencia de Nyquist). En este caso las réplicas de $X(f)$, separadas $1/T$, están en el límite de solapamiento como mostrado en la figura 6.9. Es claro que en este caso existe una única $X(f)$ que resulta en $Z(f) = T$, o sea

$$X(f) = \begin{cases} T, & |f| < W \\ 0, & \forall f \end{cases}$$

de forma que

$$x(t) = \text{sinc}\left(\frac{t}{T}\right)$$

Esto significa que el menor valor de T para el cual es posible la transmisión con ISI nula es $T = \frac{1}{2W}$, y para este valor $x(t)$ debe ser la función *sinc*. La dificultad con esta elección de $x(t)$ es que se trata de una función no causal y en consecuencia no realizable. Para hacerla realizable se utiliza una versión con retardo de la misma, o sea $\text{sinc}\left(\frac{t-t_0}{T}\right)$ donde t_0 se elige tal que para $t > 0$ sea $\text{sinc}\left(\frac{t-t_0}{T}\right) \cong 0$. Obviamente que con esta elección de $x(t)$ el instante de muestreo también debe desplazarse a $mT + t_0$. Una segunda dificultad con esta forma de pulso es que la velocidad de convergencia a cero es lenta. La forma de $x(t)$ decae con una velocidad $1/t$ y en consecuencia un pequeño error de temporizado en el muestreo de la salida del filtro acoplado en el demodulador resulta en una serie infinita de componentes de ISI. Dicha serie no es absolutamente sumable y por lo tanto la sumatoria de las ISIs resultantes puede no convergir.

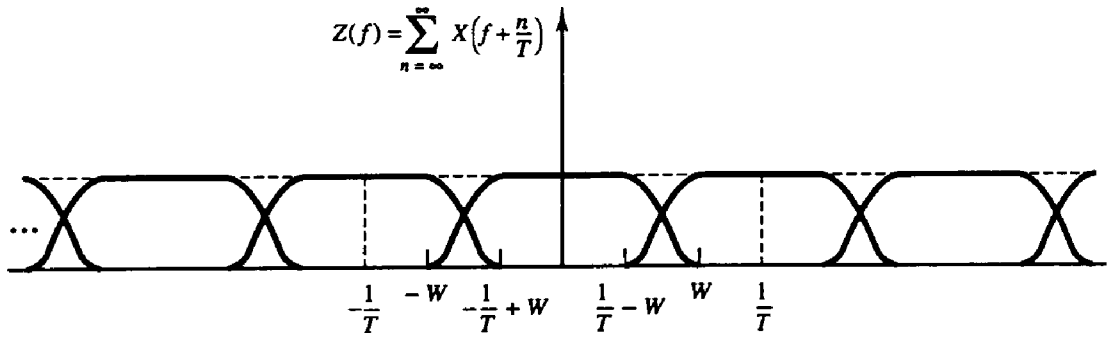

Figura 6.10: Gráfico de $Z(f)$ para el caso $T > 1/2W$.

3. $T > \frac{1}{2W}$, o equivalentemente $\frac{1}{T} < 2W$. En este caso $Z(f)$ consiste en réplicas solapadas de $X(f)$ separadas $1/T$, como mostrado en la figura 6.10. En consecuencia existen numerosas elecciones para $X(f)$ tal que $Z(f) = T$.

6.2.2 Solución práctica

Un pulso para el caso $T > \frac{1}{2T}$ que tiene propiedades espectrales deseables y que ha sido ampliamente utilizado en la práctica es el pulso *coseno elevado*. La característica en frecuencia del pulso coseno elevado está dada por

$$X_{rc}(f) = \begin{cases} T, & 0 \leq |f| \leq (1-\alpha)/2T \\ \frac{T}{2}\left[1 + \cos\frac{\pi T}{\alpha}\left(|f| - \frac{1-\alpha}{2T}\right)\right], & \frac{1-\alpha}{2T} \leq |f| \leq \frac{1+\alpha}{2T} \\ 0, & |f| > \frac{1+\alpha}{2T} \end{cases}$$

donde α se denomina *factor de caída* (rolloff) que toma valores en el rango $0 \leq \alpha \leq 1$. El ancho de banda ocupado por esta señal más allá de la frecuencia de Nyquist se denomina *ancho de banda en exceso* y se expresa usualmente como porcentaje de la frecuencia de Nyquist. Por ejemplo, cuando $\alpha = 1/2$, el ancho de banda en exceso es del 50 % y cuando $\alpha = 1$ el ancho de banda en exceso es del 100 %. El pulso con espectro coseno elevado tiene la siguiente forma

$$x(t) = \text{sinc}\left(\frac{t}{T}\right)\frac{\cos(\pi\alpha t/T)}{1 - 4\alpha^2 t^2/T^2}$$

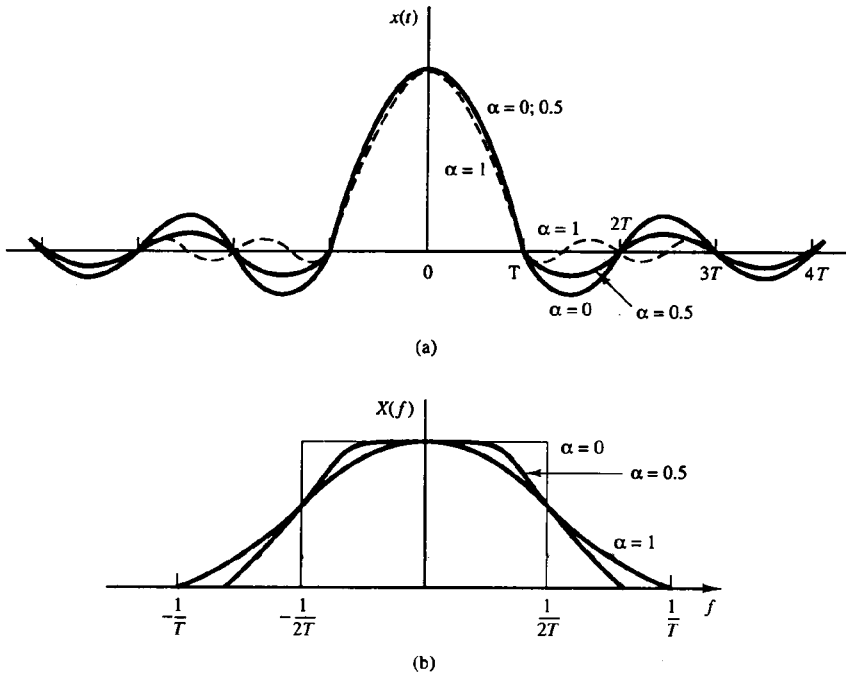

Figura 6.11: Pulsos con espectro coseno elevado.

Notar que $x(t)$ está normalizado de forma que $x(0) = 1$. La figura 6.11 ilustra el espectro del pulso coseno elevado correspondiente a $\alpha = 0, 1/2, 1$. Es posible notar que para $\alpha = 0$, el pulso se reduce a $x(t) = \text{sinc}(t/T)$ y la velocidad de transmisión será $1/T = 2W$. Cuando $\alpha = 1$ la velocidad de transmisión será $1/T = W$. En general, la cola de $x(t)$ decae con $1/t^2$ para $\alpha > 0$. En consecuencia, un error de temporizado en el muestreo conduce a una serie de componentes de ISI que convergen en el peor caso a un valor finito.

Debido a las características suaves del espectro del pulso coseno elevado es posible diseñar filtros prácticos para el transmisor y el receptor que aproximen la función transferencia total. En el caso especial en el que el canal es ideal ($C(f) = \text{rect}\left(\frac{f}{2W}\right)$), se tendrá que

$$X_{rc}(f) = G_T(f)G_R(f)$$

En ese caso, si el filtro en el receptor está acoplado al filtro de transmisión, se tendrá que $X_{rc}(f) = G_T(f)G_R(f) = |G_T(f)|^2$. Idealmente

$$G_T(f) = \sqrt{|X_{re}(f)|}e^{-j2\pi f t_0}$$

y $G_R(f) = G_T^*(f)$, donde t_0 es algún retardo nominal requerido para garantizar la realizabilidad física del filtro. De esta forma, la característica espectral del pulso coseno elevado se divide entre el filtro del transmisor y el filtro del receptor. Debe notarse que será necesario también un retardo adicional para asegurar la realizabilidad física del filtro del receptor.

6.2.3 Diseño de señales limitadas en frecuencia con ISI controlada

De la discusión del diseño de señales para ISI nula es posible notar que es necesario reducir la velocidad de transmisión de símbolos $1/T$ por debajo de la frecuencia de Nyquist de $2W$ símbolos por segundo para diseñar los filtros de transmisión y recepción. Por otro lado, si se elige relajar la condición de ISI nula, y de esta forma lograr una velocidad de transmisión de $2W$ símbolos/seg, entonces es posible permitir una cantidad de ISI controlada para lograr este objetivo.

Se vió que la condición de ISI nula es $x(nT) = 0$ para $n \neq 0$. Sin embargo, es posible suponer el diseño de la señal de banda limitada de forma de tener una cantidad controlada de ISI en algún instante. Esto significa que se permite un valor no nulo adicional en las muestras $\{x(nT)\}$. La ISI introducida es determinística o "controlada" y en consecuencia puede ser tenida en cuenta en el receptor como se discutirá en lo que sigue.

Señales de respuesta parcial

Un caso especial que conduce a filtros de transmisión y recepción fisicamente realizables (aproximadamente) está especificada por las muestras

$$x(nT) = \begin{cases} 1, & n = 0, 1 \\ 0, & \forall n \end{cases}$$

Usando ahora (6.22) se tiene que

$$z_n = \begin{cases} T, & n = 0, -1 \\ 0, & \forall n \end{cases}$$

la cual sustituida en (6.21) produce

$$Z(f) = T + Te^{-f2\pi fT}$$

Como en la sección anterior es imposible satisfacer la ecuación anterior para $T < \frac{1}{2W}$. Sin embargo, para $T = \frac{1}{2W}$ se tiene que

$$\begin{aligned} X(f) &= \begin{cases} \frac{1}{2W}\left[1 + e^{-j\pi f/W}\right], & |f| < W \\ 0, & \forall f \end{cases} \\ &= \begin{cases} \frac{1}{W}e^{-j2\pi f/W}\cos\left(\frac{\pi f}{2W}\right), & |f| < W \\ 0, & \forall f \end{cases} \end{aligned}$$

Entonces $x(t)$ estará dada por

$$x(t) = \text{sinc}(2Wt) + \text{sinc}(2Wt - 1)$$

Este pulso se denomina *señal duobinaria*, y se ilustra junto a su espectro en la figura 6.12. Es posible notar que el espectro cae a cero suavemente, lo cual significa que puede diseñarse un filtro fisicamente realizable que lo aproxime adecuadamente. De esta forma es posible alcanzar una velocidad de transmisión de $2W$ simb/seg.

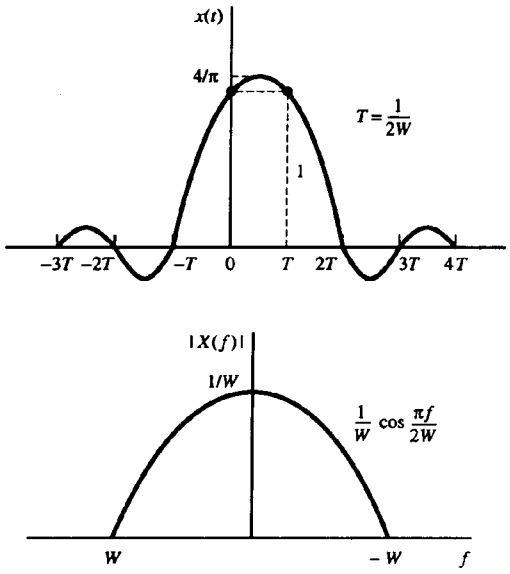

Figura 6.12: Características en el domínio tiempo y frecuencia de una señal duobinaria.

Otro caso especial que conduce a filtros de transmisión y recepción fisicamente realizables (aproximadamente) está especificada por las muestras

$$x(nT) = x(\frac{n}{2W}) = \begin{cases} 1, & n = 1 \\ -1, & n = -1 \\ 0, & \forall n \end{cases}$$

El pulso $x(t)$ correspondiente está dado por

$$x(t) = \text{sinc}(t+T)/T - \text{sinc}(t-T)/T$$

y su espectro es

$$X(f) \quad = \quad \begin{cases} \frac{1}{2W}\left(e^{j\pi f/W} - e^{-j\pi f/W}\right) = \frac{j}{W}\sin\frac{\pi f}{W}, & |f| < W \\ 0, & \forall f \end{cases}$$

Este pulso y la magnitud de su espectro se ilustran en la figura 6.13, y se denomina usualmente *señal duobinaria modificada.* Es interesante notar que el espectro de esta señal tiene un cero en el orígen, lo que la hace adecuada para la transmisión sobre un canal que no deje pasar cc.

Es posible obtener otras características de filtros fisicamente realizables seleccionando diferentes valores para las muestras $\{x(n/2W)\}$ y más de dos valores de muestras no nulas. Sin embargo, cuando se seleccionan más muestras no nulas el problema de eliminar la ISI controlada se vuelve más complicado e impráctico.

En general, la clase de señales de banda limitada que tiene la forma

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right) \frac{\sin 2\pi W(t - n/2W)}{2\pi W(t - n/2W)}$$

y su correspondiente espectro

$$X(f) \quad = \quad \begin{cases} \frac{1}{2W}\sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right) e^{-j\pi f/W}, & |f| < W \\ 0, & \forall f \end{cases}$$

se denominan *señales de respuesta parcial* ya que se introduce ISI controlada seleccionando dos o más muestras no nulas del conjunto $\{x(n/2W)\}$. Los pulsos de señal resultantes permiten la transmisión de símbolos a la velocidad de Nyquist de $2W$ simb/seg. La detección de los símbolos recibidos en presencia de ISI controlada se describirá a continuación.

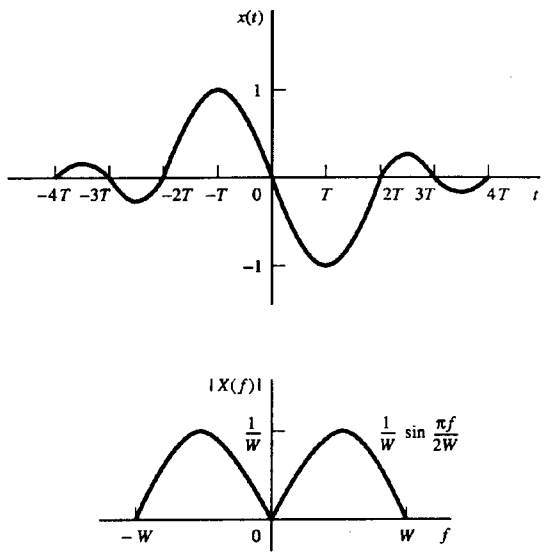

Figura 6.13: Características en el domínio tiempo y frecuencia de una señal duobinaria modificada.

6.2.4 Detección de datos para ISI controlada

En esta sección se describirán dos métodos para detectar símbolos de información en el receptor cuando la señal recibida contiene ISI controlada. Uno es un método de detección símbolo por símbolo, relativamente fácil de implementar. El segundo método se basa en el critério de Máxima Verosimilitud para detectar secuencias de símbolos. Este último método minimiza la probabilidad de error pero es un poco más complejo de implementar. En particular, es posible considerar la detección de señales de respuesta parcial duobinarias y duobinarias modificadas. En ambos casos se supondrá que la característica espectral deseada $X(f)$ para la señal de respuesta parcial se divide entre los filtros de transmisión y recepción, o sea, $|G_T(f)| = |G_R(f)| = |X(f)|^{1/2}$.

Detección símbolo por símbolo

La señal duobinaria estaba definida por $x(nT) = 1$ para $n = 0, 1$ y cero para todo otro n. En consecuencia, las muestras a la salida del filtro receptor tienen la forma

$$y_m = b_m + \nu_m = a_m + a_{m-1} + \nu_m \tag{6.24}$$

donde $\{a_m\}$ es la secuencia de amplitudes transmitidas y $\{\nu_m\}$ es una secuencia de muestras AWGN. Ignorando el ruido por el momento y considerando el caso binario donde $a_m = \pm 1$ con igual probabilidad, entonces b_m toma uno de tres posibles valores $b_m = -2, 0, 2$ con probabilidades asociadas $1/4, 1/2, 1/4$, respectivamente. Si a_{m-1} es el símbolo detectado en el intervalo $(m-1)$, su efecto sobre b_m, la señal recibida en el intervalo m puede eliminarse por substracción, permitiendo de esa forma detectar a a_m. Este proceso puede repetirse secuencialmente para cada símbolo recibido.

El principal problema con este procedimiento es que los errores que ocurren debido al ruido aditivo tienden a propagarse. Por ejemplo, si a_{m-1} está errado, su efecto sobre b_m no se elimina sino que se refuerza a través de una substracción incorrecta. En consecuencia, la detección de a_m es muy probable que esté también en error.

La propagación del error puede evitarse mediante la *precodificación* de los datos en el transmisor, en lugar de eliminar la ISI controlada mediante substracción en el receptor. La precodificación se realiza sobre la secuencia de datos binarios previamente a la modulación. A partir de la secuencia de datos $\{d_n\}$ de 0s y 1s a ser transmitida se genera una nueva secuencia $\{p_n\}$ denominada *secuencia precodificada*. Para el caso de la señal duobinaria la secuencia precodificada está definida por

$$p_m = d_m \ominus p_{m-1} \quad m = 1, 2, \cdots$$

Datos d_n		1	1	1	0	1	0	0	1	0	0	0	1	1	0	1
Precodificada p_n	0	1	0	1	1	0	0	0	1	1	1	1	0	1	1	0
Transmitida a_n	-1	1	-1	1	1	-1	-1	-1	1	1	1	1	-1	1	1	-1
Recibida b_n		0	0	0	2	0	-2	-2	0	2	2	2	0	0	2	0
Decodificada d_n		1	1	1	0	1	0	0	1	0	0	0	1	1	0	1

Tabla 6.1: Señalización binaria con pulsos duobinarios

donde el símbolo $\ominus$ es la substracción módulo 2[1]. Luego, se hace $a_m = -1$ si $p_m = 0$ y $a_m = 1$ si $p_m = 1$, o sea, $a_m = 2p_m - 1$. Esta operación de precodificación es idéntica a la descripta previamente en el contexto de la discusión de la señal NRZI.

Las muestras sin ruido a la salida del filtro receptor están dadas por

$$b_m = a_m + a_{m-1} = (2p_m - 1) + (2p_{m-1} - 1) = 2(p_m + p_{m-1} - 1)$$

En consecuencia,

$$p_m + p_{m-1} = \frac{b_m}{2} + 1$$

Dado que $d_m = p_m \oplus p_{m-1}$, se concluye que la secuencia de datos d_m se obtiene a partir de b_m como

$$d_m = \frac{b_m}{2} - 1 \pmod 2$$

En consecuencia, si $b_m = \pm 2$, $d_m = 0$ y si $b_m = 0$ entonces $d_m = 1$. Un ejemplo que ilustra las operaciones de precodificación y decodificación se muestra en la tabla 6.1. En presencia de ruido aditivo las muestras de salida del filtro receptor están dadas por (6.24). En este caso, $y_m = b_m + \nu_m$ se compara con los umbrales en $+1$ y -1. La secuencia de datos $\{d_n\}$ se obtiene de acuerdo a la siguiente regla de detección

$$d_m = \begin{cases} 1, & \text{si } -1 < y_m < 1 \\ 0, & \text{si } |y_m| \geq 1 \end{cases}$$

La extensión de PAM binaria con señalización PAM multinivel utilizando los pulsos duobinarios es directa. En este caso la secuencia $\{a_m\}$ de M amplitudes resulta en una secuencia (sin ruido) dada por

$$b_m = a_m + a_{m-1} \quad m = 1, 2, \cdots$$

la cual posee $2M - 1$ niveles posibles equiespaciados. La amplitud de los niveles está determinada por

$$a_m = 2p_m - (M - 1) \tag{6.25}$$

donde $\{p_m\}$ es la secuencia precodificada obtenida de una secuencia de datos de M niveles $\{d_m\}$ a partir de

$$p_m = d_m \ominus p_{m-1} \pmod M$$

donde los valores posibles de la secuencia $\{d_m\}$ son $0, 1, 2, \cdots, M - 1$.

En ausencia de ruido las muestras a la salida del filtro receptor pueden expresarse como

$$b_m = a_m + a_{m-1} = 2[p_m + p_{m-1} - (M - 1)]$$

[1] A pesar que es idéntica a la suma módulo 2, es conveniente ver la operación de precodificación para el caso duobinario en términos de la substracción módulo 2.

Datos d_n		0	0	1	3	1	2	0	3	3	2	0	1	0
Precodificada p_n	0	0	0	1	2	3	3	1	2	1	1	3	2	2
Transmitida a_n	-3	-3	-3	-1	1	3	3	-1	1	-1	-1	3	1	1
Recibida b_n		-6	-6	-4	0	4	6	2	0	0	-2	2	4	2
Decodificada d_n		0	0	1	3	1	2	0	3	3	2	0	1	0

Tabla 6.2: Señalización multinivel con pulsos duobinarios

En consecuencia

$$p_m + p_{m-1} = \frac{b_m}{2} + (M-1)$$

Dado que $d_m = p_m + p_{m-1} \pmod M$, se concluye que

$$d_m = \frac{b_m}{2} + (M-1) \pmod M$$

Un ejemplo ilustrando la precodificación y decodificación multinivel se muestra en la tabla 6.2.

En presencia de ruido la señal más ruido recibidos se cuantizan al nivel de señal posible más próximo y se utiliza el procedimiento anterior sobre los valores cuantizados para recuperar la secuencia de datos.

En el caso del pulso duobinario modificado la ISI controlada está especificada por los valores $x(n/2W) = -1$, para $n = 1$ y $x(n/2W) = 1$ para $n = -1$, y cero en todo otro caso. En consecuencia, la muestra libre de ruido a la salida del filtro receptor estará dada por

$$b_m = a_m - a_{m-2}$$

donde la secuencia multinivel $\{a_n\}$ se obtiene mapeando una secuencia precodificada de acuerdo a la relación (6.25) y

$$p_m = d_m \ominus p_{m-2} \pmod M$$

A partir de esas relaciones es fácil mostrar que la regla de detección para recuperar la secuencia de datos $\{d_m\}$ a partir de $\{b_m\}$ en ausencia de ruido es

$$d_m = \frac{b_m}{2} \pmod M$$

Como se demostró antes, la precodificación de la secuencia de datos en el transmisor hace posible detectar los datos recibidos símbolo por símbolo sin necesidad de tener en cuenta los símbolos anteriores. De esta forma se evita la propagación de errores.

La regla de detección símbolo por símbolo no es un esquema de detección óptimo para señales de respuesta parcial. Independientemente de ello la detección símbolo por símbolo es relativamente simple de implementar y se utiliza en muchas aplicaciones prácticas que involucran señales duobinarias y duobinarias modificadas. Su desempeño en términos de probabilidad de error se estudiará en una sección próxima.

Detección de secuencias con Máxima Verosimilitud

De la discusión anterior es claro que las de respuesta parcial son señales con memoria. Esa memoria puede representarse convenientemente mediante una *trellis*. Por ejemplo, la trellis para una señal de respuesta parcial duobinaria para transmisión de datos binários se ilustra en la figura 6.14. Para modulación binaria esa trellis tiene dos estados, correspondientes a los dos valores posibles de salida $a_m = \pm 1$. Cada rama en la trellis se rotula mediante dos números El primer número a la izquierda es el bit de dato nuevo, o sea $a_{m+1} = \pm 1$. Este número determina la transición hacia un nuevo estado. El número a la derecha es el nivel de señal recibida.

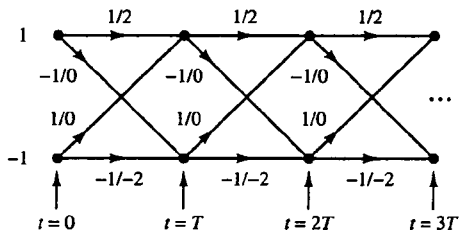

Figura 6.14: Trellis para una señal duobinaria de respuesta parcial.

La señal duobinaria tiene una memoria de longitud $L = 1$. En consecuencia, para el caso de modulación binaria la trellis tendrá $S_t = 2^L$ estados. En general, para modulación M-ária el número de estados de la trellis será M^L.

El detector de secuencias óptimo de Máxima Verosimilitud (ML) selecciona el camino más probable a través de la trellis observando la secuencia de datos recibidos $\{y_m\}$ en los instantes de muestreo $t = mT$, $m = 1, 2, \cdots$. En general cada nodo en la trellis tendrá M caminos entrantes y M métricas correspondientes. Se selecciona uno de los M caminos entrantes como el más probable, basandosé en los valores de las métricas, y los otros $M - 1$ caminos y sus métricas se descartan. El camino superviviente en cada nodo se extiende luego a M nuevos caminos, uno por cada uno de los posibles M símbolos de entrada, y se continúa el proceso de búsqueda. Este es básicamente el algoritmo de Viterbi para realizar la búsqueda sobre la trellis.

Para la clase de señales de respuesta parcial, la secuencia recibida $\{y_m,\ 1 \leq m \leq N\}$ se describe estadísticamente mediante la pdf conjunta $f(\boldsymbol{y}_N|\boldsymbol{a})$, donde $\boldsymbol{y}_N = (y_1, y_2, \cdots, y_N)^T$ y $\boldsymbol{a}_N = (a_1, a_2, \cdots, a_N)^T$ y $N > L$. Cuando el ruido es AWGN, $f(\boldsymbol{y}_N|\boldsymbol{a})$ es una pdf multivariable definida por

$$f(\boldsymbol{y}_N|\boldsymbol{a}) = \frac{1}{[2\pi \det(\boldsymbol{C})]^{1/2}} e^{\left[-\frac{1}{2}(\boldsymbol{y}_N - \boldsymbol{b}_N)^T \boldsymbol{C}^{-1} (\boldsymbol{y}_N - \boldsymbol{b}_N)\right]}$$

donde $\boldsymbol{b}_N = (b_1, b_2, \cdots, b_N)^T$ es la media del vector $\boldsymbol{y}_N$ y $\boldsymbol{C}$ es la matriz covarianza $(N \times N)$ de $\boldsymbol{y}_N$. Luego, el detector de secuencias ML selecciona la secuencia a través de la trellis que maximiza la pdf $f(\boldsymbol{y}_N|\boldsymbol{a})$.

El cómputo para hallar la secuencia más probable a través de la trellis se simplifica teniendo en cuenta el logaritmo natural de $f(\boldsymbol{y}_N|\boldsymbol{a})$. De esa forma,

$$\ln f(\boldsymbol{y}_N|\boldsymbol{a}) = -\frac{N}{2} \ln[2\pi \det(\boldsymbol{C})] - \frac{1}{2}(\boldsymbol{y}_N - \boldsymbol{b}_N)^T \boldsymbol{C}^{-1} (\boldsymbol{y}_N - \boldsymbol{b}_N)$$

Dada la secuencia $\{y_m\}$, la secuencia de datos $\{a_m\}$ que maximiza $\ln f(\boldsymbol{y}_N|\boldsymbol{a})$ es idéntica a la secuencia $\{a_m\}$ que minimiza $(\boldsymbol{y}_N - \boldsymbol{b}_N)^T \boldsymbol{C}^{-1} (\boldsymbol{y}_N - \boldsymbol{b}_N)$, o sea

$$\hat{a}_N = \arg\min_{\boldsymbol{a}_N} \left[(\boldsymbol{y}_N - \boldsymbol{b}_N)^T \boldsymbol{C}^{-1} (\boldsymbol{y}_N - \boldsymbol{b}_N)\right] \tag{6.26}$$

La búsqueda del camino de menor distancia a través de la trellis puede realizarse secuencialmente mediante el uso del algoritmo de Viterbi. Es posible considerar la señal duobinaria con modulación binaria y considerando inicialmente el estado inicial $a_0 = 1$. Luego suponiendo recibir $y_1 = a_1 + a_0 + \nu_1$ en el instante $t = T$ y $y_2 = a_2 + a_1 + \nu_2$ en el instante $t = 2T$, se tienen cuatro caminos candidatos correspondientes a $(a_1, a_2) = (1, 1)$, $(-1, 1)$, $(1, -1)$ y $(-1, -1)$. Los dos primeros caminos candidato convergen en el estado 1 en $t = 2T$. Para los dos caminos convergiendo al estado 1 se calculan las métricas $\mu_2(1, 1)$ y $\mu_2(-1, 1)$ y se selecciona el camino más probable. Un cálculo similar se realiza en el estado -1 para las secuencias $(1, -1)$ y $(-1, -1)$. De esta forma una de las dos secuencias en cada nodo se guarda y la otra se descarta. La búsqueda sobre la trellis continúa al recibir uma muestra de señal y_3 en el instante $t = 3T$, extendiendo los dos caminos supervivientes del instante $t = 2T$.

Los cálculos de las métricas se complican debido a la correlación de las muestras de ruido a la salida del filtro acoplado para el caso de una señal de respuesta parcial. Por ejemplo, en el caso de una señal duobinaria la correlación de la secuencia de ruido $\{\nu_m\}$ se verifica sobre dos muestras de señal sucesivas. En consecuencia, ν_m y ν_{m+k} están correlacionadas para $k = 1$ y no correlacionadas para $k > 1$. En general, una

señal de respuesta parcial con memoria L resultará en una secuencia de ruido correlacionada a la salida del filtro acoplado que satisface la condición $E[\nu_m \nu_{m+k}] > 0$ para $k > L$.

Si se ignora la correlación del ruido, o sea, $E[\nu_m \nu_{m+k}] = 0$ para $k > 0$, es posible introducir algunas simplificaciones en el cálculo de la métrica. Luego si por suposición $\boldsymbol{C} = \sigma_\nu^2 \boldsymbol{I}$ donde $\sigma_\nu^2 = E[\nu_m^2]$ e $\boldsymbol{I}$ es la matriz identidad de $(N \times N)$, entonces (6.26) se simplifica a

$$\begin{aligned} \hat{a}_N &= \arg\min_{\boldsymbol{a}_N} \left[(\boldsymbol{y}_N - \boldsymbol{b}_N)^T \boldsymbol{C}^{-1} (\boldsymbol{y}_N - \boldsymbol{b}_N)\right] \\ &= \arg\min_{\boldsymbol{a}_N} \left[\sum_{m=1}^{N} (y_m - b_m)^2\right] \end{aligned} \tag{6.27}$$

donde

$$b_m = \sum_{k=0}^{L} x_k a_{m-k}$$

y $x_k = x(kT)$ son los valores de las muestras de la señal de respuesta parcial. En este caso el cálculo de las métricas en cada nodo de la trellis tiene la forma

$$\mu_m(a_m) = \mu_{m-1}(a_{m-1}) + \left(y_m - \sum_{k=0}^{L} x_k a_{m-k}\right)^2 \tag{6.28}$$

donde $\mu_m(a_m)$ son las métricas en el instante $t = mT$, $\mu_{m-1}(a_{m-1})$ son las métricas en el instante $t = (m-1)T$ y el segundo término del lado derecho de (6.28) representa los nuevos incrementos sobre las métricas en base a la nueva muestra de señal recibida y_m.

Como se indicó en el Capítulo anterior, la detección de secuencias ML introduce un retardo variable al detectar cada símbolo de información recibido. En la práctica el retardo variable se evita truncando las secuencias supervivientes a los N_t símbolos más recientes, donde $N_t \gg 5L$, logrando de esta forma un retardo fijo. En el caso que M^L secuencias supervivientes difieran en el instante $t = mT$ en el símbolo a_{m-N_t}, puede elegirse el símbolo más probable de la secuencia superviviente. La pérdida de desempeño resultante de este truncamiento es despreciable si $N_t > 5L$.

Ejemplo. Expresar el cálculo de la métrica efectuado para una señal duobinaria en los instantes $t = 2T$ y $t = 3T$ en base a muestras de la señal recibida $y_m = b_m + \nu_m$ para $m = 1, 2, 3, \cdots$, ignorando la correlación del ruido.

Las métricas están dadas en general por (6.28). Habiendo recibido y_1 e y_2 y con $a_0 = 1$, las métricas de los dos caminos que convergen al estado 1 son

$$\begin{aligned} \mu_2(1,1) &= (y_1 - 2)^2 + (y_2 - 2)^2 \\ \mu_2(-1,1) &= y_1^2 + y_2^2 \end{aligned}$$

Si $\mu_2(1,1) < \mu_2(-1,1)$, se selecciona el camino $(1,1)$ como el más probable y se descarta $(-1,1)$. En caso contrário se selecciona el camino $(-1,1)$ y se descarta el camino $(1,1)$. El camino con la menor métrica se denomina superviviente. La secuencia (a_1, a_2) y la métrica correspondiente se guardan.

Un cálculo similar se realiza en el estado -1 para las secuencias $(-1,1)$ y $(-1,-1)$. Se tiene entonces

$$\begin{aligned} \mu_2(1,-1) &= (y_1 - 2)^2 + y_2^2 \\ \mu_2(-1,-1) &= y_1^2 + (y_2 + 2)^2 \end{aligned}$$

Se compara las métricas $\mu_2(1,-1)$ y $\mu_2(-1,-1)$ y se selecciona la secuencia con menor métrica.

Habiendo recibido y_3 en $t = 3T$ es posible considerar la extensión de los caminos supervivientes. Suponiendo que los dos caminos supervivientes son $(a_1, a_2) = (1, 1)$ y $(a_1, a_2) = (1, -1)$, entonces al estado 1 en $t = 3T$ convergen dos caminos $(a_1, a_2, a_3) = (1, 1, 1)$ y $(a_1, a_2, a_3) = (1, -1, 1)$. Las métricas correspondientes serán

$$\begin{aligned} \mu_3(1,1,1) &= \mu_2(1,1) + (y_3 - 2)^2 \\ \mu_3(1,-1,1) &= \mu_2(1,-1) + y_3^2 \end{aligned}$$

Comparando las métricas para esos dos caminos convergentes y seleccionando el camino con la menor métrica se obtiene el superviviente.

En forma similar, en el estado -1 en $t = 3T$ se tienen dos caminos convergentes $(1, 1, -1)$ y $(1, -1, -1)$, y sus correspondientes métricas serán

$$\begin{aligned} \mu_3(1,1,-1) &= \mu_2(1,1) + y_3^2 \\ \mu_3(1,-1,-1) &= \mu_2(1,-1) + (y_3 + 2)^2 \end{aligned}$$

Se selecciona el camino de menor métrica como el superviviente en el estado -1. Este proceso continúa hasta recibir datos adicionales en $t = kT$, $k = 4, 5, \cdots$.

6.3 Probabilidad de error para señales PAM

En esta sección se evaluará el desempeño del receptor para demodular y detectar una señal PAM M-ária en presencia de ruido AWGN en su entrada. Se considerará primero el caso en que los filtros de transmisión $G_T(f)$ y recepción $G_R(f)$ se diseñan para ISI nula. Luego se considerará el caso en el cual $G_T(f)$ y $G_R(f)$ se diseñan tal que $x(t) = g_T(t) * g_R(t)$ es o una señal duobinaria o una señal duobinaria modificada.

6.3.1 Probabilidad de error para ISI nula

En ausencia de ISI, una muestra de la señal a la salida del filtro acoplado del receptor tendrá la forma

$$y_m = x_0 a_m + \nu_m$$

donde

$$x_0 = \int_{-W}^{W} |G_T(f)|^2 df = \mathcal{E}_g$$

y ν_m es el ruido aditivo Gaussiano con media cero y varianza

$$\sigma_\nu^2 = \mathcal{E}_g N_0/2$$

En general a_m toma uno de M valores de amplitud equiespaciados con igual probabilidad. Dado un nivel particular de amplitud el problema es determinar la probabilidad de error.

El problema de evaluar la probabilidad de error para PAM digital en un canal de ancho de banda limitado AWGN en ausencia de ISI es idéntico a la evaluación de la probabilidad de error para una señal PAM M-ária como discutido en el Capítulo anterior. El resultado final obtenido de aquella derivación era

$$P_M = \frac{2(M-1)}{M} Q\left[\sqrt{\frac{2\mathcal{E}_g}{N_0}}\right]$$

Pero como $\mathcal{E}_g = 3\mathcal{E}_{av}/(M^2 - 1)$, con $\mathcal{E}_{av} = k\mathcal{E}_{bav}$ la energía promedio por símbolo y $\mathcal{E}_{bav}$ es la energía promedio por bit, entonces

$$P_M = \frac{2(M-1)}{M} Q\left[\sqrt{\frac{6(\log_2 M)\mathcal{E}_{bav}}{(M^2-1)N_0}}\right]$$

Esta es exactamente la forma de la probabilidad de error de PAM M-ária obtenida previamente en el Capítulo anterior. En el tratamiento de PAM discutido en este Capítulo se introdujo la restricción adicional de que la señal transmitida tenga un ancho de banda limitado al del canal. En consecuencia, los pulsos de señal transmitidos se diseñaron para ser de ancho de banda limitado de forma de no introducir ISI.

En contraste, no se introdujo restricciones de ancho de banda sobre las señales PAM consideradas en el Capítulo anterior. Independientemente de eso, los receptores (demoduladores y detectores) en ambos casos son óptimos (filtros acoplados) para la señal transmitida correspondiente. En consecuencia, no existirá una pérdida en el desempeño de error al considerar la restricción de ancho de banda cuando el pulso de señal se diseña para ISI nula y el canal no distorsiona la señal transmitida.

6.3.2 Probabilidad de error para señales de respuesta parcial

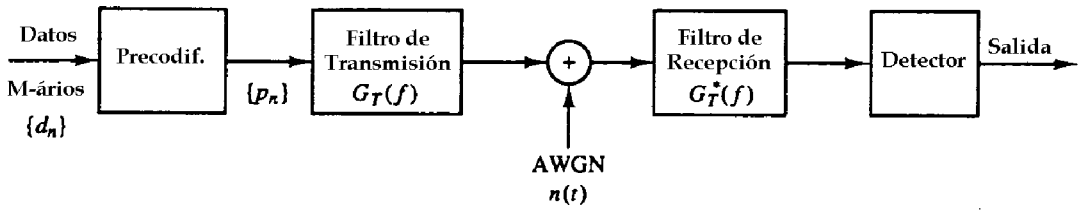

Figura 6.15: Diagrama en bloques del modulador y demodulador para señales de respuesta parcial.

En esta sección se determinará la probabilidad de error en la detección de la señalización PAM M-ária usando pulsos duobinarios y duobinarios modificados. Se supondrá que el canal es ideal de ancho de banda limitado con ruido blanco Gaussiano aditivo. El modelo para el sistema de comunicaciones se ilustra en la figura 6.15.

Se considerarán dos tipos de detectores. El primero es el detector símbolo por símbolo y el segundo es el detector de secuencias óptimo ML descripto en la sección anterior.

Detector símbolo por símbolo

Se precodifica en el transmisor la secuencia de datos M-ária como descripto anteriormente. La salida del precodificador se mapea en uno de los M posibles niveles de amplitud. La salida del filtro de transmisión con respuesta en frecuencia $G_T(f)$ será

$$v(t) = \sum_{n=-\infty}^{\infty} a_n g_T(t - nT)$$

La función de respuesta parcial $X(f)$ se divide igualmente entre los filtros de transmisión y recepción. En consecuencia, el filtro de recepción estará acoplado al pulso transmitido y la cascada de los dos filtros resulta en una característica de frecuencia dada por

$$|G_R(f)G_T(f)| = |X(f)|$$

La salida del filtro acoplado se muestrea en $t = nT = n/2W$ y las muestras se envian al decodificador. La salida del filtro acoplado en el instante de muestreo, para el caso de la señal duobinaria, puede expresarse como

$$y_m = a_m + a_{m-1} + \nu_m = b_m + \nu_m$$

donde ν_m es la componente de ruido aditivo. En forma similar, la salida del filtro acoplado para el caso de la señal duobinaria modificada será

$$y_m = a_m - a_{m-2} + \nu_m = b_m + \nu_m$$

Para transmisión binaria se supondrá $a_m = \pm d$, donde $2d$ es la distancia entre niveles de señal. Entonces, los valores correspondientes de b_m serán $(2d, 0, -2d)$. Para transmisión PAM M-ária, $a_m = \pm d, \pm 3d, \cdots, \pm(M-1)d$, y los niveles de señal recibida serán $b_m = 0, \pm 2d, \pm 4d, \cdots, \pm 2(M-1)d$. En consecuencia, el número de niveles recibidos es $2M-1$ y el factor de escala d es equivalente a $x_0 = \mathcal{E}_g$.

Los símbolos de entrada transmitidos $\{a_m\}$ se suponen equiprobables. Luego para señales duobinarias y duobinarias modificadas es posible demostrar que en ausencia de ruido los niveles de salida recibidos tienen una función densidad de probabilidad discreta triangular dada por

$$P(b = 2md) = \frac{M - |m|}{M^2}, \quad m = 0, \pm 1, \pm 2, \cdots, \pm(M-1)$$

donde b es el nivel sin ruido y $2d$ es la distancia entre niveles cualquiera adyacentes.

El canal corrompe la señal transmitida mediante el ruido AWGN de media cero y varianza $N_0/2$.

Se supone que ocurrirá un error de símbolo siempre que la magnitud del ruido aditivo exceda la distancia d. Esta suposición no considera un evento raro consistente en que una componente grande de ruido con magnitud excediendo d pueda resultar en un nivel de señal recibida que produce una decisión correcta de símbolo. La componente de ruido ν_m es Gaussiana de media cero y varianza dada por

$$\sigma_\nu^2 = \frac{N_0}{2}\int_{-W}^{W} |G_R(f)|^2 df = \frac{N_0}{2}\int_{-W}^{W} |X(f)|^2 df = 2N_0/\pi$$

para ambos tipos de señales, duobinarias y duobinarias modificadas. En consecuencia, una cota superior sobre la probabilidad de error de símbolo será

$$\begin{aligned}
P_M &< \sum_{m=-(M-2)}^{M-2} P(|y - 2md| > d \mid b = 2md)\, P(b = 2md) \\
&\quad + 2P(y + 2(M-1)d > d \mid b = -2(M-1)d)\, P(b = -2(M-1)d) \\
&= P(|y| > d \mid b = 0)\left[2\sum_{m=0}^{M-1} P(b = 2md) - P(b = 0) - P(b = -2(M-1)d)\right] \\
&= \left(1 - \frac{1}{M^2}\right) P(|y| > d \mid b = 0)
\end{aligned}$$

Pero

$$P(|y| > d \mid b = 0) = \frac{2}{\sqrt{2\pi}\sigma_\nu}\int_d^\infty e^{x^2/2\sigma_\nu} dx = 2Q\left(\sqrt{\frac{\pi d^2}{2N_0}}\right)$$

En consecuencia, la probabilidad promedio de error de símbolo tiene la siguiente cota superior

$$P_M < 2\left(1 - \frac{1}{M^2}\right) Q\left(\sqrt{\frac{\pi d^2}{2N_0}}\right) \tag{6.29}$$

El factor de escala d en esta última ecuación puede eliminarse expresando d en términos de la potencia promedio transmitida en el canal. Para una señal PAM M-ária en la cual los niveles de amplitud son equiprobables, la potencia promedio a la salida del filtro de transmisión será

$$P_{av} = \frac{E[a_m^2]}{T}\int_{-W}^{W}|G_T(f)|^2 df = \frac{E[a_m^2]}{T}\int_{-W}^{W}|X(f)|df = \frac{4}{\pi T}E[a_m^2]$$

donde $E[a_m^2]$ es el valor medio cuadrático de los M niveles de señal, o sea

$$E[a_m^2] = \frac{d^2(M^2-1)}{3}$$

En consecuencia

$$d^2 = \frac{3\pi P_{av}T}{4(M^2-1)} \tag{6.30}$$

Sustituyendo en (6.29) se obtiene la siguiente cota superior para la probabilidad de error promedio de símbolo

$$P_M < 2\left(1-\frac{1}{M^2}\right)Q\left(\sqrt{\left(\frac{\pi}{4}\right)^2\frac{6}{M^2-1}\frac{\mathcal{E}_{av}}{N_0}}\right) \tag{6.31}$$

donde $\mathcal{E}_{av}$ es la energía promedio por símbolo transmitido, la cual puede expresarse también en términos de la energía promedio por bit $\mathcal{E}_{av} = k\mathcal{E}_{bav} = (\log_2 M)\mathcal{E}_{bav}$.

La expresión en (6.31) para la probabilidad de error de PAM M-ária es válida para señales de respuesta parcial duobinarias y duobinarias modificadas. Si se compara este resultado con la probabilidad de error de PAM M-ária con ISI nula, que puede obtenerse utilizando un pulso de señal de espectro coseno elevado, es posible notar que el desempeño de la respuesta parcial duobinaria o duobinaria modificada tiene una pérdida de $(\pi/4)^2$ o 2.1 dB. Esta pérdida en la SNR se debe a que el detector para las señales de respuesta parcial realiza decisiones en base a símbolo por símbolo, ignorando de esta forma la memoria inherente contenida en la señal recibida a la entrada del receptor.

Para observar la memoria en la secuencia recibida es posible tener en cuenta la secuencia recibida, sin ruido, de la tabla 6.1 para transmisión binaria. La secuencia $\{b_m\}$ es $0, -2, 0, 2, 0, -2, 0, 2, 2, \cdots$. Notar que no es posible tener una transición de -2 a $+2$ o de $+2$ a -2 de un intervalo de símbolo para el siguiente. Por ejemplo, si el nivel de señal a la entrada del detector es -2, el próximo nivel de señal puede ser -2 o 0. En forma similar, si el nivel de señal en un instante de muestreo dado es $+2$, el nivel de señal en el próximo instante de muestreo puede ser $+2$ o 0. En otras palabras, no es posible encontrar una transición de -2 a $+2$ o visceversa entre dos muestras sucesivas recibidas del filtro acoplado. Sin embargo, un detector símbolo por símbolo no explota esta restricción, o memoria inherente, en la secuencia recibida. A continuación se discutirá el desempeño del detector de secuencias ML que explota la memoria inherente en la modulación y, en consecuencia, permite recuperar gran parte de los 2.1 dB de pérdida sufridos por el detector símbolo por símbolo.

Detector de secuencias de Máxima Verosimilitud

El detector de secuencias ML busca a través de la trellis la secuencia transmitida $\{a_m\}$ más probable, como se discutió en secciones anteriores. En cada etapa del proceso de búsqueda el detector compara las métricas de los caminos que convergen en cada uno de los nodos y selecciona el camino más probable asociado a cada nodo. El desempeño del detector puede evaluarse determinando la probabilidad de que un camino errado a través de la trellis tenga una métrica de distancia Euclidiana menor que la del camino correcto.

En general, el cálculo de la probabilidad de error exacta es extremadamente dificil. En su lugar, es posible determinar una aproximación a la probabilidad de error basada en la comparación de las métricas de dos caminos que convergen a un nodo y están separados por la menor distancia Euclidiana a los otros caminos. La aproximación se llevará a cabo para el caso de señales de respuesta parcial duobinarias.

Es posible considerar la trellis para la señal de respuesta parcial duobinaria de la figura 6.14. Se supondrá comenzar en el estado 1 en $t = 0$ y que los dos primeros símbolos transmitidos son $a_1 = 1$ y $a_2 = 1$ Luego, en $t = T$ se recibe $y_1 = 2d + \nu_1$ y en $t = 2T$ se recibe $y_2 = 2d + \nu_2$. Si el camino $(a_1, a_2) = (-1, 1)$ es más probable que $(a_1, a_2) = (1, 1)$ dados los valores recibidos y_1 e y_2, se estará introduciendo un error. Este

evento de error de camino es el evento de error de camino dominante y en consecuencia sirve como una buena aproximación a la probabilidad de error para el detector de secuencia ML.

De las discusiones de secciones anteriores para este detector, la métrica para el camino $(a_1, a_2) = (1, 1)$ está dada por

$$\mu_2(1,1) = [y_1 - 2d \;\; y_2 - 2d]\boldsymbol{C}^{-1} \begin{bmatrix} y_1 - 2d \\ y_2 - 2d \end{bmatrix} \tag{6.32}$$

donde la matriz covarianza $\boldsymbol{C}$ está dada por

$$\boldsymbol{C} = \frac{2N_0}{\pi} \begin{bmatrix} 1 & \frac{1}{3} \\ \frac{1}{3} & 1 \end{bmatrix}$$

Para el camino $(a_1, a_2) = (-1, 1)$, la métrica correspondiente será

$$\mu_2(-1,1) = [y_1 \;\; y_2]\boldsymbol{C}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{6.33}$$

La probabilidad de un evento de error de camino es simplemente la probabilidad que la métrica $\mu_2(-1, 1)$ sea menor que la métrica $\mu_2(1, 1)$, o sea

$$P_2 = P[\mu_2(-1,1) < \mu_2(1,1)]$$

Sustituyendo $y_1 = 2d + \nu_1$ e $y_2 = 2d + \nu_2$ en (6.32) y (6.33) se obtiene que

$$P_2 = P[\nu_1 + \nu_2 < -2d]$$

Dado que ν_1 y ν_2 son variables Gaussianas (correlacionadas) de media cero, la suma también es Gaussiana de media cero. La varianza de la suma $z = \nu_1 + \nu_2$ es simplemente $\sigma_z^2 = 16N_0/2\pi$. En consecuencia

$$P_2 = P[z < -2d] = Q\left(\frac{2d}{\sigma_z}\right) = Q\left(\sqrt{\frac{4d^2}{\sigma_z^2}}\right)$$

De (6.30) se tiene que (con $M = 2$) la expresión para d^2 es

$$d^2 = \frac{\pi P_{av} T}{4} = \frac{\pi \mathcal{E}_b}{4}$$

En consecuencia, la probabilidad de un evento de error de camino será

$$P_2 = Q\left(\sqrt{\frac{1.5\pi^2}{16}\frac{2\mathcal{E}_b}{N_0}}\right)$$

Es posible notar en principio que este evento de camino de error resulta en un bit de error en la secuencia de dos bits. De esta forma, la probabilidad de error de bit es $P_2/2$. Segundo, existe una reducción en la SNR de $10\log(1.5\pi^2/16) = -0.34$ dB en relación al caso sin ISI. Esta pequeña degradación es la penalidad incurrida al cambiar por la eficiencia de ancho de banda de la señal de respuesta parcial. Finalmente, notar que el detector de secuencias ML ha recuperado 1.76 dB de los 2.11 dB de degradación inherente al detector símbolo por símbolo.

6.4 Diseño en presencia de distorsión en el canal

Anteriormente se describió un critério de diseño de señal que resulta en ISI nula a la salida del filtro receptor. Teniendo en cuenta que una señal pulso $x(t)$ satisface la condición de ISI nula en los instantes $t = nT$, $n = \pm 1, \pm 2, \cdots$, si su espectro $X(f)$ satisface la condición (6.18). De esta condición se concluye que para una transmisión libre de ruido en un canal los filtros de transmisión y recepción y la función transferencia del canal deben cumplir

$$G_T(f)C(f)G_R(f) = X_{rc}(f)$$

donde $X_{rc}(f)$ es la transformada de Fourier de un pulso coseno elevado cuyos parámetros dependen del ancho de banda del canal W y el intervalo de transmisión T. Obviamente existe un número infinito de filtros de transmisión - recepción que satisfacen la condición anterior.

En esta sección se considerará el diseño de los filtros de transmisión - recepción que maximizan la SNR a la salida del filtro de recepción. En consecuencia, en esta sección el énfasis estará orientado al diseño de un sistema de comunicaciones digitales con ISI nula y probabilidad de error mínima para un canal con distorsión. Se presentará primero una breve discusión de varios tipos de distorsión de canal y luego se obtendrán las ecuaciones de diseño para los filtros de transmisión y recepción.

Se distinguirán dos tipos de distorsión. La *distorsión de amplitud* resulta cuando la característica de amplitud $|C(f)|$ no es constante para $|f| < W$. El segundo tipo de distorsión, denominada *distorsión de fase* resulta cuando la característica de fase $\Theta_c(f)$ no es lineal en frecuencia.

Otro punto de vista de la distorsión de fase se obtiene considerando la derivada de $\Theta_c(f)$. De esta forma, se define como característica *retardo de envolvente* a

$$\tau(f) = -\frac{1}{2\pi}\frac{d\Theta_c(f)}{df}$$

Cuando $\Theta_c(f)$ es lineal con la frecuencia, el retardo de envolvente es constante para todas las frecuencias. En ese caso, todas las frecuencias en la señal transmitida pasan a través del canal con el mismo retardo fijo. En tal caso, no existe distorsión de fase. Sin embargo, cuando $\Theta_c(f)$ no es lineal, el retardo de envolvente $\tau(f)$ varía con la frecuencia y las diferentes componentes de la señal de entrada soportan diferentes retardos al pasar por el canal. En tal caso la señal transmitida ha sufrido *distorsión de retardo.*

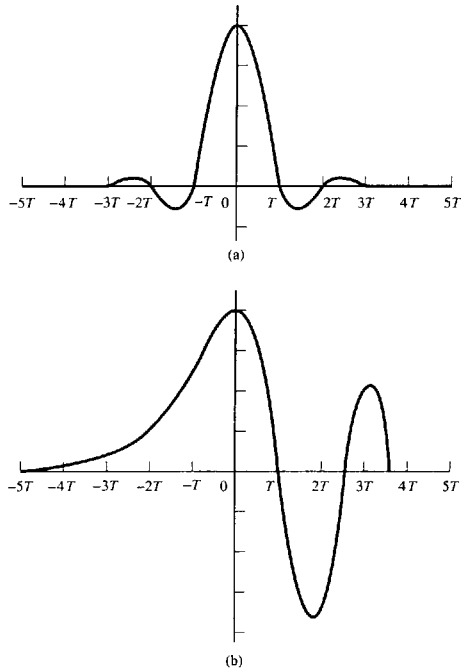

Figura 6.16: Efectos de la distorsión del canal, a la entrada (a) y la salida (b).

Tanto la distorsión de amplitud como la de retardo originan ISI en la señal recibida. Por ejemplo, es posible suponer que se ha diseñado un pulso con un espectro coseno elevado que tiene ISI nula en los instantes de muestreo, como el ilustrado en la figura 6.16a. Cuando el pulso pasa a través de un canal con amplitud constante $|C(f)| = 1$ para $|f| < W$ y una característica de fase cuadrática (retardo de envolvente lineal), el pulso recibido a la salida del canal tendrá la forma ilustrada en la figura 6.16b. Notar que los cruces periódicos por cero se han desplazado debido a la distorsión de retardo de forma que el pulso resultante resulta en ISI. En consecuencia, una secuencia de pulsos sucesivos se mezclarán unos sobre otros y los picos de los pulsos ya no serán distinguibles debido a la ISI.

Se considerarán a continuación dos problemas. Primero, se considerará el diseño de filtros óptimos de transmisión y recepción en presencia de distorsión del canal cuando las características del canal son conocidas. Segundo, se considerará el diseño de filtros especiales denominados *ecualizadores de canal*, que en forma automática y adaptiva corrigen la distorsión del canal cuando sus características son desconocidas.

6.4.1 Diseño óptimo de los filtros de transmisión y recepción

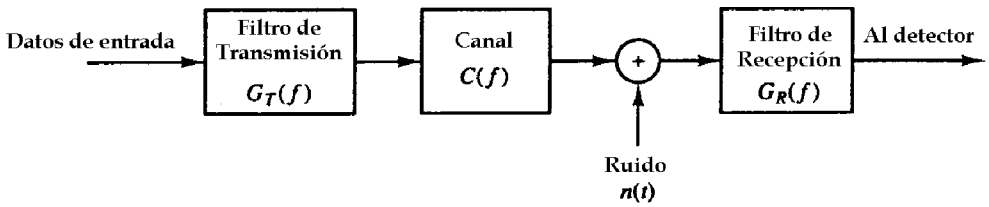

Figura 6.17: Configuración del sistema para el diseño de $G_T(f)$ y $G_R(f)$.

En esta sección se asumirá que la característica de respuesta en frecuencia del canal $C(f)$ es conocida y se considerará el problema de diseño de los filtros de transmisión y recepción que maximizan la SNR a la salida del filtro receptor y resultan en ISI nula. La figura 6.17 ilustra el sistema completo bajo estas consideraciones.

Para la componente de señal, se debe satisfacer la siguiente condición

$$G_T(f)C(f)G_R(f) = X_{rc}(f)e^{-j2\pi f t_0}, \quad |f| \leq W \tag{6.34}$$

donde $X_{rc}(f)$ es el espectro deseado coseno elevado que produce ISI nula en los instantes de muestreo, y t_0 es un retardo de tiempo necesario para asegurar la realizabilidad física de los filtros de transmisión y recepción.

El ruido a la salida del filtro de recepción puede expresarse como

$$\nu(t) = \int_{-\infty}^{\infty} n(t-\tau)g_R(\tau)d\tau$$

donde $n(t)$ es la entrada al filtro. Se supone que el ruido $n(t)$ es Gaussiano de media cero. En consecuencia, $\nu(t)$ es Gaussiano de media cero con densidad espectral de potencia

$$S_\nu(f) = S_n(f)|G_R(f)|^2$$

donde $S_n(f)$ es la densidad espectral del proceso de ruido $n(t)$.

Por simplicidad se considerará en primer lugar el caso de transmisión PAM binaria. Entonces, la salida muestreada del filtro acoplado será

$$y_m = x_0 a_m + \nu_m = a_m + \nu_m$$

donde x_0 está normalizado[2] a la unidad, $a_m = \pm d$, y ν_m representa el término de ruido, que es Gaussiano de media cero y varianza

[2]Haciendo $x_0 = 1$ y $a_m = \pm 1$, el escalamiento por x_0 se incorpora dentro del parámetro d.

$$\sigma_\nu^2 = \int_{-\infty}^{\infty} S_n(f)|G_R(f)|^2 df \tag{6.35}$$

En consecuencia, la probabilidad de error será

$$P_2 = \frac{1}{\sqrt{2\pi}} \int_{d/\sigma_\nu}^{\infty} e^{-y^2/2} dy = Q\left(\sqrt{\frac{d^2}{\sigma_\nu^2}}\right)$$

La probabilidad de error se minimiza maximizando la SNR $= d^2/\sigma_\nu^2$, ó equivalentemente, minimizando la relación de ruido a señal σ_ν^2/d^2. Pero d^2 está relacionada con la potencia de la señal transmitida por

$$P_{av} = \frac{E(a_m^2)}{T} \int_{-\infty}^{\infty} g_T^2(t)dt = \frac{d^2}{T} \int_{-\infty}^{\infty} g_T^2(t)dt$$

entonces

$$\frac{1}{d^2} = \frac{1}{P_{av}T} \int_{-\infty}^{\infty} |G_T(f)|^2 df$$

Sin embargo, $G_T(f)$ debe elegirse para satisfacer la condición de ISI nula dada por (6.34). En consecuencia

$$|G_T(f)| = \frac{|X_{rc}(f)|}{|C(f)||G_R(f)|}, \quad |f| \leq W$$

y $G_T(f) = 0$ para $|f| \leq W$, tal que

$$\frac{1}{d^2} = \frac{1}{P_{av}T} \int_{-\infty}^{\infty} \frac{|X_{rc}(f)|^2}{|C(f)|^2|G_R(f)|^2} df$$

En consecuencia, la relación ruido a señal que debe minimizarse con respecto a $|G_R(f)|$ para $|f| < W$ será

$$\frac{\sigma_\nu^2}{d^2} = \frac{1}{P_{av}T} \int_{-W}^{W} S_n(f)|G_R(f)|^2 df \int_{-W}^{W} \frac{|X_{rc}(f)|^2}{|C(f)|^2|G_R(f)|^2} df \tag{6.36}$$

La $|G_R(f)|$ óptima puede hallarse mediante el uso de la desigualdad de Schwartz, dada por

$$\int_{-\infty}^{\infty} |U_1(f)|^2 df \int_{-\infty}^{\infty} |U_2(f)|^2 df \leq \left[\int_{-\infty}^{\infty} |U_1(f)||U_2(f)| df\right]^2$$

donde $|U_1(f)|$ y $|U_2(f)|$ se definen como

$$\begin{aligned} |U_1(f)| &= |\sqrt{S_n(f)}||G_R(f)| \\ |U_2(f)| &= \frac{|X_{rc}(f)|}{|C(f)||G_R(f)|} \end{aligned}$$

El valor mínimo de (6.36) se obtiene cuando $|U_1(f)|$ es proporcional a $|U_2(f)|$, o equivalentemente cuando

$$|G_R(f)| = K \frac{|X_{rc}(f)|^{1/2}}{|C(f)|^{1/2}|S_n(f)|^{1/4}}, \quad |f| \leq W \tag{6.37}$$

donde K es una constante. El filtro de transmisión correspondiente tiene la siguiente característica de magnitud

$$|G_T(f)| = \frac{1}{K}\frac{|X_{rc}(f)|^{1/2}|S_n(f)|^{1/4}}{|C(f)|^{1/2}}, \quad |f| \leq W \tag{6.38}$$

Finalmente, con los filtros de transmisión y recepción óptimos así obtenidos, la SNR máxima será

$$\frac{d^2}{\sigma_\nu^2} = \frac{P_{av}T}{\left[\int_{-W}^{W} \frac{|X_{rc}(f)||S_n(f)|^{1/2}}{|C(f)|}df\right]^2} \tag{6.39}$$

Notar que los filtros de transmisión y recepción están especificados solo en magnitud. Las características de fase de $G_T(f)$ y $G_R(f)$ deben seleccionarse para satisfacer la condición (6.34), o sea

$$\Theta_T(f) + \Theta_c(f) + \Theta_R(f) = 2\pi f t_0 \tag{6.40}$$

donde $\Theta_T(f)$, $\Theta_c(f)$ y $\Theta_R(f)$ son las características de los filtros de transmisión, el canal y recepción, respectivamente.

En el caso especial donde el ruido aditivo a la entrada al demodulador es blanco Gaussiano con densidad espectral de potencia $N_0/2$, las características especificadas por (6.37) y (6.38) se reducen a

$$\begin{aligned} |G_R(f)| &= K_1\frac{|X_{rc}(f)|^{1/2}}{|C(f)|^{1/2}}, \quad |f| \leq W \\ |G_T(f)| &= K_2\frac{|X_{rc}(f)|^{1/2}}{|C(f)|^{1/2}}, \quad |f| \leq W \end{aligned}$$

donde K_1 y K_2 son factores de escala arbitrários. Notar que en este caso $|G_R(f)|$ es el filtro acoplado a $|G_T(f)|$. La SNR correspondiente en el detector, dada por (6.39) se reduce a

$$\frac{d^2}{\sigma_\nu^2} = \frac{2P_{av}T}{N_0}\left[\int_{-W}^{W} \frac{|X_{rc}(f)|}{|C(f)|}df\right]^{-2}$$

Ejemplo. Determinar los filtros de transmisión y recepción óptimos para un sistema de comunicaciones binário que transmite a una velocidad de 4800 bits/seg sobre un canal con respuesta en frecuencia (magnitud) dada por

$$|C(f)| = \frac{1}{\sqrt{1+\left(\frac{f}{W}\right)^2}}, \quad |f| \leq W$$

donde $W = 4800$ Hz. El ruido es AWGN con densidad espectral de potencia $N_0 = 10^{-15}$ W/Hz.
Dado que $W = 1/T = 4800$ es posible utilizar una señal pulso con espectro coseno elevado y $\alpha = 1$. De esa forma

$$X_{rc}(f) = \frac{T}{2}[1+\cos(\pi T|f|)] = T\cos^2\left(\frac{\pi|f|}{9600}\right)$$

Entonces

$$|G_T(f)| = |G_R(f)| = \left[1+\left(\frac{f}{4800}\right)^2\right]^{1/4}\cos\left(\frac{\pi|f|}{9600}\right), \quad |f| \leq 4800$$

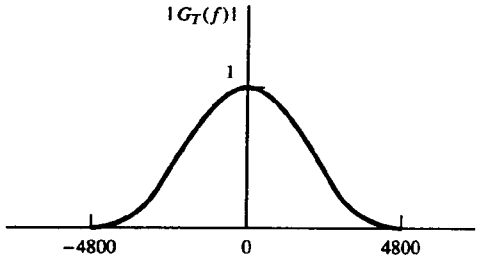

Figura 6.18: Respuesta en frecuencia del filtro de transmisión óptimo.

y $|G_T(f)| = |G_R(f)| = 0$ para toda otra frecuencia. La figura 6.18 ilustra la característica $G_T(f)$.

Es posible utilizar esos filtros óptimos para determinar la cantidad de energía transmitida $\mathcal{E}$ requerida para lograr una probabilidad de error especificada.

El procedimiento para obtener los filtros óptimos de transmisión y recepción para el caso de PAM M-ária a través de un canal no ideal $C(f)$ es idéntico al del caso de PAM binaria. Las únicas modificaciones son las expresiones para la potencia promedio transmitida, que para PAM M-ária es

$$P_{av} = \frac{E(a_m^2)}{T}\int_{-W}^{W}|G_T(f)|^2 df = \frac{(M^2-1)d^2}{3T}\int_{-W}^{W}|G_T(f)|^2 df \tag{6.41}$$

y la expresión correspondiente para la probabilidad de error, que ahora será

$$P_M = \frac{2(M-1)}{M}Q\left(\sqrt{\frac{d^2}{\sigma_\nu^2}}\right) \tag{6.42}$$

Para AWGN (6.42) puede expresarse como

$$P_M = \frac{2(M-1)}{M}Q\left(\sqrt{\frac{6\mathcal{E}_{av}}{(M^2-1)N_0}\left[\int_{-W}^{W}\frac{|X_{rc}(f)|}{|C(f)|}df\right]^{-2}}\right)$$

Finalmente, es posible observar que la pérdida debido a la distorsión del canal es

$$20\log_{10}\left[\int_{-W}^{W}\frac{|X_{rc}(f)|}{|C(f)|}df\right] \tag{6.43}$$

Notar que cuando $C(f) = 1$ para $|f| \leq W$, entonces el canal es ideal y $\int_{-W}^{W} X_{rc}(f)df = 1$, de forma que no existirán pérdidas. Por otro lado, cuando existe distorsión de amplitud, $|C(f)| < 1$ para algún rango de frecuencias en la banda $|f| \leq W$ y en consecuencia existirá una pérdida en SNR dada por (6.43). Esta pérdida es independiente de la distorsión de fase del canal, dado que la distorsión de fase ha sido compensada perfectamente siguiendo (6.40). La pérdida dada por (6.43) se debe enteramente a la distorsión de amplitud y es una medida de la amplificación del ruido resultante del filtro de recepción, el cual compensa la distorsión del canal.

6.4.2 Ecualización del canal

En la sección precedente se describió el diseño de los filtros de transmisión y recepción para transmisión PAM digital cuando las características de respuesta en frecuencia del canal son conocidas. El objetivo era

diseñar esos filtros para obtener ISI nula y SNR máxima en los instantes de muestreo. Esta metodología de diseño es apropiada cuando el canal se conoce precisamente y sus características no cambian con el tiempo.

En la práctica se encuentra frecuentemente canales cuyas características de respuesta en frecuencia son desconocidas o cambian con el tiempo. Por ejemplo, en la transmisión de datos a través de la red telefónica, el canal de comunicaciones será diferente cada vez que se digite un número porque la ruta del canal será diferente. Una vez que la conexión se realiza, sin embargo, el canal será invariante en el tiempo por un período de tiempo relativamente grande. Este es un ejemplo de un canal cuyas características no son conocidas a priori. Ejemplos de canales variantes en el tiempo son los canales de radio, tal como los de propagación ionosférica. Esos canales se describen mediante características de respuesta en frecuencia variantes en el tiempo, y sirven de ejemplo como casos donde la optimización de los filtros de transmisión y recepción no es posible.

Bajo esas circunstancias es posible diseñar el filtro transmisor para que tenga una respuesta en frecuencia de tipo raíz cuadrada coseno elevado, o sea

$$G_T(f) = \begin{cases} \sqrt{X_{rc}(f)}e^{-j2\pi f t_0}, & |f| \leq W \\ 0, & |f| \leq W \end{cases}$$

y el filtro de recepción, con respuesta en frecuencia $G_R(f)$, de forma que esté acoplada a $G_T(f)$, o sea

$$|G_T(f)||G_R(f)| = X_{rc}(f) \tag{6.44}$$

Entonces, debido a la distorsión del canal, la salida del filtro receptor será

$$y(t) = \sum_{n=-\infty}^{\infty} a_n x(t - nT) + \nu(t) \tag{6.45}$$

donde $x(t) = g_T(t) * c(t) * g_R(t)$. La salida del filtro puede muestrearse periodicamente para producir la secuencia

$$y_m = \sum_{n=-\infty}^{\infty} a_n x_{m-n} + \nu_m = x_0 a_m + \sum_{n \neq m} a_n x_{m-n} + \nu_m \tag{6.46}$$

donde $x_n = x(nT)$, $n = 0, \pm 1, \pm 2, \cdots$. El segundo término del lado derecho de (6.46) representa la ISI.

En cualquier sistema práctico es razonable suponer que la ISI afecta un número finito de símbolos. En consecuencia, es posible asumir que $x_n = 0$ para $n < -L_1$ y $n > L_2$, donde L_1 y L_2 son enteros positivos acotados. De esta forma, la ISI observada a la salida del filtro de recepción puede verse como generada haciendo pasar los datos $\{a_m\}$ a través de un filtro FIR con coeficientes $\{x_n, -L_1 \leq n \leq L_2\}$, como mostrado en la figura 6.19. Este se denomina *modelo de canal discreto equivalente*. Dado que su entrada es la secuencia de información discreta (binaria o M-ária), la salida del modelo de canal discreto puede caracterizarse como la de una máquina de estados finita corrompida por AWGN. En consecuencia, la salida sin ruido del modelo puede describirse mediante una trellis con M^L estados donde $L = L_1 + L_2$.

Detección de secuencias de Máxima Verosimilitud

El detector óptimo para la secuencia de información $\{a_m\}$, basado en la observación de la secuencia recibida $\{y_m\}$ dada por (6.46), es el detector de secuencias ML. El detector es semejante al detector de secuencias ML descripto en el contexto de la detección de señales de respuesta parcial que tienen ISI controlada. El algoritmo de Viterbi provee un método para la búsqueda del camino ML a través de la trellis. Para llevar a cabo esta búsqueda deben conocerse, o medirse mediante algún método, los coeficientes $\{x_n\}$ del modelo de canal discreto equivalente. En cada etapa de la búsqueda sobre la trellis existirán M^L secuencias supervivientes con las correspondientes M^L métricas de distancia Euclidiana asociadas a los caminos a evaluar.

Debido al incremento exponencial de la complejidad computacional del algoritmo de Viterbi con la ventana (longitud L) de la ISI, este tipo de detección es práctico solo cuando M y L son pequeños. Por ejemplo, en sistemas de telefonía celular móvil, empleados para la transmisión de señales de voz, M es

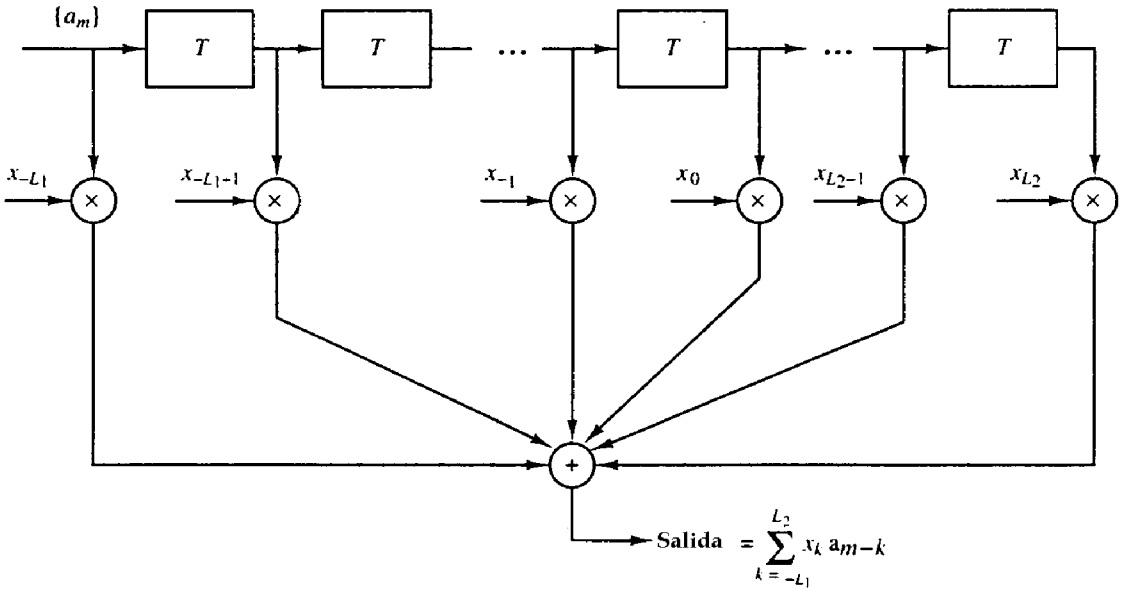

Figura 6.19: Modelo de canal discreto equivalente.

usualmente pequeño, por ejemplo: $M = 4$ y $L = 2$. En este caso, el detector de secuencias ML puede implementarse con complejidad razonable. Sin embargo, cuando M y L son grandes, el detector de secuencias ML se vuelve impráctico. En tal caso se utilizan otros métodos subóptimos, pero prácticos, para detectar la secuencia de información $\{a_m\}$ en presencia de ISI. Independientemente de ello, el desempeño del detector de secuencias ML para un canal con ISI sirve como "ideal" de comparación en relación a los métodos subóptimos. Se describirán a continuación dos métodos subóptimos.

Ecualizadores lineales

Para compensar la distorsión del canal es posible emplear un filtro lineal con parámetros ajustables. Los parámetros del filtro se ajustan en base a las mediciones de las características del canal. Estos filtros ajustables se denominan *ecualizadores de canal* o simplemente *ecualizadores*.

En canales cuyas características de respuesta en frecuencia son desconocidas, pero invariantes en el tiempo, es posible medir esas características para ajustar los parámetros del ecualizador que, uno vez ajustados, se mantendrán fijos durante la transmisión de los datos. Tales ecualizadores se denominan *preajustados*. Por otro lado, los *ecualizadores adaptivos* cambian sus parámetros en base a un ajuste periódico durante la transmisión de los datos.

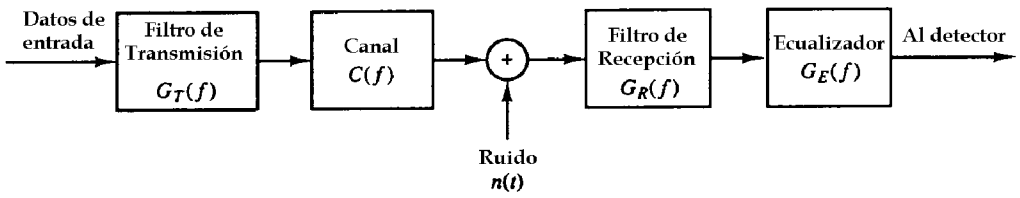

Figura 6.20: Diagrama en bloques de un sistema con ecualizador.

Se considerará las características de diseño para un ecualizador lineal desde el punto de vista del domínio frecuencia. La figura 6.20 muestra un diagrama de bloques de un sistema que utiliza un filtro lineal como ecualizador de canal.

El demodular consiste en un filtro de recepción con respuesta en frecuencia $G_R(f)$ en cascada con un filtro de ecualización de canal con respuesta en frecuencia $G_E(f)$. Dado que $G_R(f)$ está acoplada a $G_T(f)$ y estos filtros están diseñados para satisfacer (6.44), $|G_E(f)|$ debe compensar la distorsión del canal. En consecuencia, la respuesta en frecuencia del ecualizador debe ser igual a la inversa de la respuesta del canal, o sea

$$G_E(f) = \frac{1}{C(f)} = \frac{1}{|C(f)|} e^{-j\Theta_c(f)} \quad |f| \leq W$$

donde $|G_E(f)| = 1/|C(f)|$ y la característica de fase del ecualizador será $\Theta_E(f) = -\Theta_c(f)$.

Es posible notar que este filtro elimina completamente la ISI causada por el canal. Dado que anula la ISI en los instantes de muestreo $t = nT$, el ecualizador se denomina *ecualizador de ISI nula.* De esta forma, la entrada al detector tiene la forma

$$y_m = a_m + \nu_m$$

donde ν_m es la componente de ruido AWGN con varianza dada por

$$\sigma_\nu^2 = \int_{-\infty}^{\infty} S_n(f)|G_R(f)|^2|G_E(f)|^2 df = \int_{-W}^{W} \frac{S_n(f)|X_{rc}(f)|}{|C(f)|^2} df$$

donde $S_n(f)$ es la densidad espectral de potencia del ruido. Cuando el ruido es blanco $S_n(f) = N_0/2$ y la varianza se vuelve

$$\sigma_\nu^2 = \frac{N_0}{2} \int_{-W}^{W} \frac{|X_{rc}(f)|}{|C(f)|^2} df \tag{6.47}$$

Notar que la varianza del ruido a la salida del ecualizador de ISI nula es en general mayor que la varianza del ruido a la salida del filtro de recepción óptimo $|G_R(f)|$, dado por (6.35), para el caso que el canal era conocido.

Ejemplo. El canal dado en el ejemplo anterior se procesa con un ecualizador de ISI nula. Suponiendo que los filtros de transmisión y recepción satisfacen (6.44), determinar el valor de la varianza del ruido en los instantes de muestreo y la probabilidad de error.

Cuando el ruido es blanco, la varianza del ruido a la salida del ecualizador (entrada al detector) estará dada por (6.47). En consecuencia,

$$\begin{aligned} \sigma_\nu^2 &= \frac{N_0}{2} \int_{-W}^{W} \frac{|X_{rc}(f)|}{|C(f)|^2} df = \frac{TN_0}{2} \int_{-W}^{W} \left[1 + \left(\frac{f}{W}\right)\right] \cos^2 \frac{\pi|f|}{2W} df \\ &= N_0 \int_0^1 (1 + x^2) \cos^2 \frac{\pi x}{2} dx = \left(\frac{2}{3} - \frac{1}{\pi^2}\right) N_0 \end{aligned}$$

Por otro lado, de (6.41) se obtiene la potencia promedio transmitida dada por

$$P_{av} = \frac{(M^2-1)d^2}{3T} \int_{-W}^{W} |G_T(f)|^2 df = \frac{(M^2-1)d^2}{3T} \int_{-W}^{W} |X_{rc}(f)| df = \frac{(M^2-1)d^2}{3T}$$

La expresión general para la probabilidad de error está dada por (6.42). En consecuencia se tiene que

$$P_M = \frac{2(M-1)}{M} Q\left(\sqrt{\frac{3P_{av}T}{(M^2-1)(2/3 - 1/\pi^2)N_0}}\right)$$

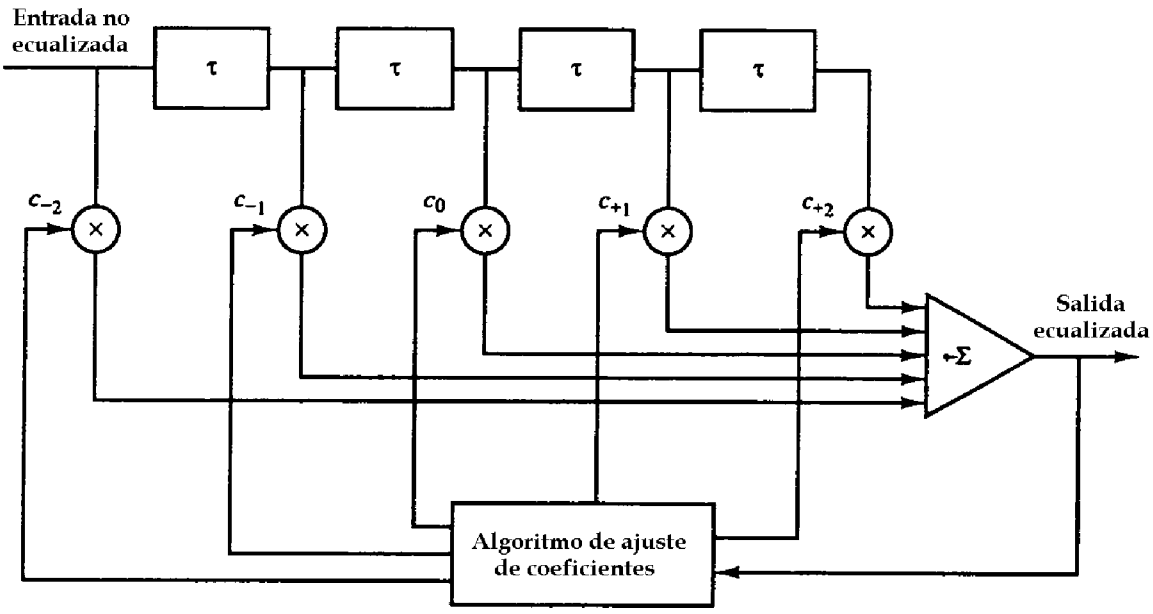

Figura 6.21: Filtro transversal lineal.

Si el canal fuera ideal, el argumento de la función Q sería $6P_{av}T/(M^2-1)N_0$. En consecuencia, la pérdida de desempeño debido al canal no ideal ecualizado está dada por el factor $2(\frac{2}{3}-\frac{1}{\pi^2}) = 1.133$ ó 0.54 dB.

Es posible también considerar el diseño de un ecualizador lineal desde el punto de vista del domínio tiempo. Se observó anteriormente que en canales reales la ISI está limitada un número finito L de muestras. En consecuencia, el ecualizador de canal práctico se aproxima mediante un filtro de respuesta finita al impulso con coeficientes ajustables $\{c_n\}$ como ilustrado en la figura 6.21. El retardo τ entre coeficientes adyacentes puede seleccionarse tan grande como T, el intervalo de símbolo, en cuyo caso el ecualizador FIR se denomina *ecualizador de espaciamiento de símbolo.* En este caso la entrada al ecualizador es la secuencia muestreada dada por (6.46). Sin embargo, es posible notar que cuando $1/T < 2W$, las frecuencias por encima de $1/T$ en la señal recibida se solaparán con las frecuencias por debajo de $1/T$. En este caso, El ecualizador compensa una señal distorsionada por el canal con aliasing.

Por otro lado, cuando el retardo de tiempo τ entre coeficientes adyacentes se elige tal que $1/\tau \geq 2W > 1/T$ no ocurrirá aliasing y, en consecuencia, el ecualizador de canal compensará la verdadera distorsión del canal. Dado que $\tau < T$, el ecualizador se denomina *de espaciamiento fraccionario.* En la práctica τ se elige usualmente $\tau = T/2$. Notar que en este caso la frecuencia de muestreo a la salida del filtro $G_R(f)$ es $\frac{2}{T}$.

La respuesta impulsiva del ecualizador FIR será

$$g_E(t) = \sum_{n=-N}^{N} c_n \delta(t - n\tau)$$

y la correspondiente respuesta en frecuencia

$$G_E(f) = \sum_{n=-N}^{N} c_n e^{-j2\pi f n\tau}$$

donde $\{c_n\}$ son los $(2N+1)$ coeficientes del ecualizador y N se elige suficientemente grande como para que

la ventana de ecualización corresponda a la de ISI, o sea: $2N+1 \geq L$. Dado que $X(f) = G_T(f)C(f)G_R(f)$ y $x(t)$ es la señal pulso correspondiente a $X(f)$, entonces el pulso de salida ecualizado será

$$q(t) = \sum_{n=-N}^{N} c_n x(t - n\tau)$$

Ahora es posible aplicar la condición de ISI nula a las muestras de $q(t)$ en los instantes $t = mT$. Esas muestras serán

$$q(mT) = \sum_{n=-N}^{N} c_n x(mT - n\tau), \quad m = 0, \pm 1, \cdots, \pm N$$

Dado que existen $2N+1$ coeficientes del ecualizador es posible controlar solo $2N+1$ valores muestreados de $q(t)$. Especificamente, es posible forzar las siguientes condiciones

$$q(mT) = \sum_{n=-N}^{N} c_n x(mT - n\tau) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2, \cdots, \pm N \end{cases} \tag{6.48}$$

lo cual puede expresarse en forma matricial como $\boldsymbol{X}\boldsymbol{c} = \boldsymbol{q}$, donde $\boldsymbol{X}$ es una matriz de $(2N+1) \times (2N+1)$ con elementos $x(mT - n\tau)$, $\boldsymbol{c}$ es el vector de $(2N+1)$ coeficientes y $\boldsymbol{q}$ es un vector columna de dimensión $(2N+1)$ con solo un elemento no nulo. De esta forma, se obtiene un conjunto de $2N+1$ ecuaciones lineales para los coeficientes del ecualizador de ISI nula.

Debe enfatizarse que el ecualizador de ISI nula FIR no elimina completamente la ISI debido a su longitud finita. Sin embargo, cuando N aumenta la ISI residual puede reducirse y en el límite para $N \to \infty$, la ISI se elimina completamente.

Ejemplo. Considerar un pulso de canal distorsionado $x(t)$ a la entrada del ecualizador dado por

$$x(t) = \frac{1}{1 + \left(\frac{2t}{T}\right)^2}$$

donde $1/T$ es la velocidad de transmisión de símbolos. El pulso se muestrea a una velocidad $2/T$ y se procesa con un ecualizador de ISI nula. Determinar los coeficientes de un ecualizador de ISI nula de 5 coeficientes.

De acuerdo a (6.48) el ecualizador de ISI nula debe satisfacer las siguientes ecuaciones

$$q(mT) = \sum_{n=-2}^{2} c_n x(mT - nT/2) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2 \end{cases}$$

La matriz $\boldsymbol{X}$ con elementos $x(mT - nT/2)$ está dada por

$$\boldsymbol{X} = \begin{bmatrix} \frac{1}{5} & \frac{1}{10} & \frac{1}{17} & \frac{1}{26} & \frac{1}{37} \\ 1 & \frac{1}{2} & \frac{1}{5} & \frac{1}{10} & \frac{1}{17} \\ \frac{1}{5} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{17} & \frac{1}{10} & \frac{1}{5} & \frac{1}{2} & 1 \\ \frac{1}{37} & \frac{1}{26} & \frac{1}{17} & \frac{1}{10} & \frac{1}{5} \end{bmatrix}$$

El vector de coeficientes $\boldsymbol{c}$ y el vector $\boldsymbol{q}$ están dados por

$$\boldsymbol{c} = \begin{bmatrix} c_{-2} \\ c_{-1} \\ c_0 \\ c_1 \\ c_2 \end{bmatrix} \qquad \boldsymbol{q} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Entonces, la solución del sistema $\boldsymbol{X}\boldsymbol{c} = \boldsymbol{q}$ determina la solución

$$\boldsymbol{c}_{opt} = \boldsymbol{X}^{-1}\boldsymbol{q} = \begin{bmatrix} -2.2 \\ 4.9 \\ -3 \\ 4.9 \\ -2.2 \end{bmatrix}$$

Un problema del ecualizador de ISI nula es que ignora la presencia de ruido aditivo. En consecuencia, su utilización puede resultar en una amplificación de ruido significativa. Esto puede verse facilmente notando que en un rango de frecuencias donde $C(f)$ es pequeña el ecualizador $G_E(f) = 1/C(f)$ compensa colocando una gran ganancia. De esa forma el ruido en ese rango de frecuencias se amplifica. Una alternativa es relajar la condición de ISI nula y seleccionar la característica del ecualizador de canal para minimizar la potencia combinada en la ISI residual y el ruido aditivo a la salida del ecualizador. Un ecualizador de canal optimizado con ese objetivo se basa en el critério de error medio cuadrático mínimo (MMSE).

Para discutir conceptos al respecto, es posible considerar la salida del ecualizador FIR corrompida por ruido,

$$z(t) = \sum_{n=-N}^{N} c_n y(t - n\tau)$$

donde $y(t)$ es la entrada al ecualizador, dada por (6.45). Para la salida muestreada en $t = mT$ se tiene que

$$z(mT) = \sum_{n=-N}^{N} c_n y(mT - n\tau)$$

La respuesta deseada a la salida del ecualizador en $t = mT$ es el símbolo transmitido a_m. El error se define como la diferencia entre a_m y $z(mT)$. Luego, el error medio cuadrático (MSE) entre la muestra de salida $z(mT)$ y el valor deseado a_m será

$$\begin{aligned} MSE &= E[z(mT) - a_m]^2 = E\left[\sum_{n=-N}^{N} c_n y(t - n\tau) - a_m\right]^2 \\ &= \sum_{n=-N}^{N}\sum_{k=-N}^{N} c_n c_k R_Y(n-k) - 2\sum_{n=-N}^{N} c_k R_{AY}(k) + E[a_m]^2 \end{aligned} \tag{6.49}$$

donde las correlaciones se definen como sigue

$$\begin{aligned} R_Y(n-k) &= E[y(mT - n\tau)y(mT - n\tau)] \\ R_{AY}(n-k) &= E[y(mT - n\tau)a_m] \end{aligned}$$

y la esperanza opera sobre la secuencia de información $\{a_m\}$ y el ruido aditivo.

La solución MMSE se obtiene derivando (6.49) con respecto a los coeficientes del ecualizador $\{c_n\}$. De esta forma se obtienen las condiciones necesarias para el MMSE dadas por

$$\sum_{n=-N}^{N} c_n R_y(n-k) = R_{AY}(k), \qquad k = 0, \pm 1, \pm 2, \cdots, \pm N. \tag{6.50}$$

Esto determina $(2N+1)$ ecuaciones lineales en los coeficientes del ecualizador. En contraste a la solución de ISI nula descripta anteriormente, esas ecuaciones dependen de las propiedades estadísticas del ruido y de la ISI a através de la autocorrelación $R_Y(n)$.

En la práctica la autocorrelación $R_Y(n)$ y la correlación cruzada $R_{AY}(n)$ no son conocidas. Sin embargo, pueden estimarse transmitiendo una señal de prueba por el canal y utilizando los promedios temporales

$$\begin{aligned}\hat{R}_Y(n) &= \frac{1}{K}\sum_{k=1}^{K} y(kT - n\tau)y(kT) \\ \hat{R}_{AY}(n-k) &= \frac{1}{K}\sum_{k=1}^{K} y(kT - n\tau)a_k\end{aligned}$$

en lugar de los promedios estadísticos para obtener los coeficientes dados por (6.50).

Ecualizadores adaptivos

Se mostró que los coeficientes de un ecualizador lineal pueden determinarse resolviendo un sistema de ecuaciones lineales. Con el critério de ISI nula, el sistema de ecuaciones estaba dado por (6.48). Por otro lado, si el critério de optimización estaba basado en la minimización del MSE, los coeficientes del ecualizador óptimo se obtienen resolviendo el conjunto de ecuaciones de (6.50).

En ambos casos, es posible expresar el conjunto de ecuaciones lineales en la forma matricial general siguiente

$$\boldsymbol{B}\boldsymbol{c} = \boldsymbol{d}$$

donde $\boldsymbol{B}$ es una matriz de $(2N+1)\times(2N+1)$, $\boldsymbol{c}$ es un vector columna asociado a los coeficientes del ecualizador y $\boldsymbol{d}$ es un vector columna de dimensión $(2N+1)$. Obviamente la solución de esta ecuación es

$$\boldsymbol{c}_{opt} = \boldsymbol{B}^{-1}\boldsymbol{d} \tag{6.51}$$

En implementaciones prácticas de ecualizadores, la solución de (6.51) se obtiene usualmente mediante un método iterativo para evitar el cálculo explícito de la inversa de la matriz $\boldsymbol{B}$. El procedimiento iterativo más simple es el método *steepest descent*, en el cual se comienza eligiendo arbitrariamente el vector de coeficientes $\boldsymbol{c}_0$. Esta elección inicial de coeficientes corresponde a un punto sobre la función critério que está siendo optimizada. Por ejemplo, en el caso del critério MSE, $\boldsymbol{c}_0$ corresponde a un punto sobre la superfície cuadrática en el espacio $(2N+1)$-dimensional de coeficientes. Se calcula el vector gradiente, definido como $\boldsymbol{g}_0$, o sea la derivada del MSE con respecto a los $2N+1$ coeficientes del ecualizador, y cada coeficiente se ajusta en la dirección opuesta a su dirección de gradiente correspondiente. El cambio en el coeficiente j-ésimo es proporcional al tamaño de la componente j del gradiente.

Por ejemplo, el vector gradiente $\boldsymbol{g}_k$ para el critério MSE hallado en base a las derivadas del MSE con respecto a cada uno de los $2N+1$ coeficientes será

$$\boldsymbol{g}_k = \boldsymbol{B}\boldsymbol{c}_k - \boldsymbol{d}, \quad k = 0, 1, 2, \cdots \tag{6.52}$$

Entonces el vector de coeficientes $\boldsymbol{c}_k$ se ajusta de acuerdo con la siguiente relación

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k - \Delta\boldsymbol{g}_k$$

donde Δ es el *parámetro de tamaño del paso* asociado al procedimiento iterativo. Para asegurar la convergencia del proceso iterativo Δ debe elegirse pequeño y positivo. En tal caso, el vector gradiente $\boldsymbol{g}_k$ converge a cero, o sea $\boldsymbol{g}_k \to \boldsymbol{0}$ para $k \to \infty$, y el vector de coeficientes $\boldsymbol{c}_k \to \boldsymbol{c}_{opt}$ como ilustrado en la figura 6.22 en base a un ejemplo de dos coeficientes. En general, la convergencia de los coeficientes a la solución óptima $\boldsymbol{c}_{opt}$ no puede lograrse con un número finito de iteraciones con el método de steepest descent. Sin embargo, dicha solución óptima puede aproximarse tanto cuanto se desee en unos cientos de iteraciones. En sistemas de comunicaciones digitales que emplean ecualizadores de canal, cada iteración corresponde a un intervalo de tiempo para enviar un símbolo y, en consecuencia, unos cientos de iteraciones para lograr la convergencia a un valor próximo del óptimo corresponden a una fracción de segundo.

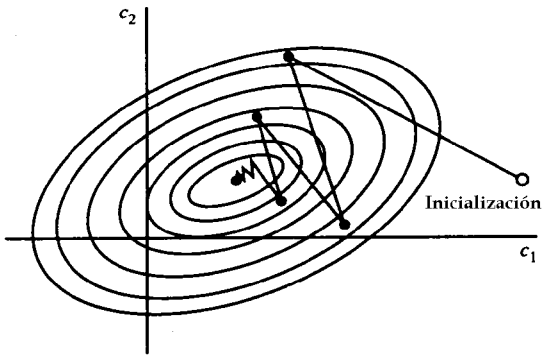

Figura 6.22: Ejemplo de convergencia de un algoritmo de gradiente.

La ecualización de canal adaptiva se requiere en canales cuyas características cambian en el tiempo. En tal caso, la ISI varía con el tiempo. El ecualizador de canal debe seguir tales variaciones en la respuesta del canal y ajustar sus coeficientes para reducir la ISI. En el contexto de la discusión anterior, los coeficientes óptimos $\boldsymbol{c}_{opt}$ varían en el tiempo debido a variaciones en la matriz $\boldsymbol{B}$ y, para el caso del critério MSE, variaciones en el tiempo en el vector $\boldsymbol{d}$. Bajo esas condiciones, el método iterativo descripto puede modificarse para utilizar estimaciones de las componentes del gradiente. De esa forma, el algoritmo para ajustar los coeficientes del ecualizador puede expresarse como

$$\hat{\boldsymbol{c}}_{k+1} = \hat{\boldsymbol{c}}_k - \Delta \hat{\boldsymbol{g}}_k \tag{6.53}$$

donde $\hat{\boldsymbol{g}}_k$ es una estimación del vector gradiente $\boldsymbol{g}_k$ y $\hat{\boldsymbol{c}}_k$ es la estimación del vector de coeficientes.

En el caso del critério MSE, el vector gradiente $\boldsymbol{g}_k$ dado por (6.52) puede expresarse también como sigue

$$\boldsymbol{g}_k = -E[e_k \boldsymbol{y}_k] \tag{6.54}$$

donde e_k es la *señal de error* $e_k = a_k - z_k$ definida como la diferencia entre la salida deseada del ecualizor a_k en el instante k y la salida real z_k, e $\boldsymbol{y}_k$ es el vector columna de $2N+1$ valores de señal recibidos contenidos en el ecualizador en el instante k. En base a (6.54), una estimación del gradiente $\hat{\boldsymbol{g}}_k$ se calcula como sigue

$$\hat{\boldsymbol{g}}_k = -e_k \boldsymbol{y}_k \tag{6.55}$$

de forma que sustituyendo en (6.53) se obtiene el algoritmo adaptivo para optimizar los coeficientes en base al critério MSE como sigue

$$\hat{\boldsymbol{c}}_{k+1} = \hat{\boldsymbol{c}}_k + \Delta e_k \boldsymbol{y}_k \tag{6.56}$$

Dado que se utiliza una estimación del vector gradiente en (6.56), el procedimiento se denomina *algoritmo de gradiente estocástico*. Se conoce también como *algoritmo Least Mean Square* (LMS).

En la figura 6.23 se ilustra un diagrama de bloques de un ecualizador adaptivo que ajusta los coeficientes de acuerdo con (6.56). Notar que para formar la señal error e_k se utiliza la diferencia entre la respuesta deseada del ecualizador a_k y la salida real z_k. Esta señal error está escalada por el parámetro de tamaño de paso Δ y esto a su vez multiplicado por los valores de la señal recibida $\{y(kT - n\tau)\}$ en los $2N+1$ coeficientes. Los productos $\Delta e_k y(kT - n\tau)$ en los $(2N+1)$ coeficientes se obtienen para ajustar los coeficientes de acuerdo con (6.56). Este cálculo se repite por cada nueva muestra recibida. De esta forma los coeficientes del ecualizador se ajustan a la velocidad de símbolos.

Inicialmente, el ecualizador adaptivo se entrena mediante la transmisión de una secuencia seudoaleatória conocida $\{a_m\}$ enviada a través del canal. En el demodulador, el ecualizador emplea la secuencia conocida para ajustar sus coeficientes. Después del ajuste inicial, el ecualizador adaptivo conmuta de un *modo de entrenamiento* a un *modo dirigido por la decisión*, en el cual las decisiones a la salida del detector son

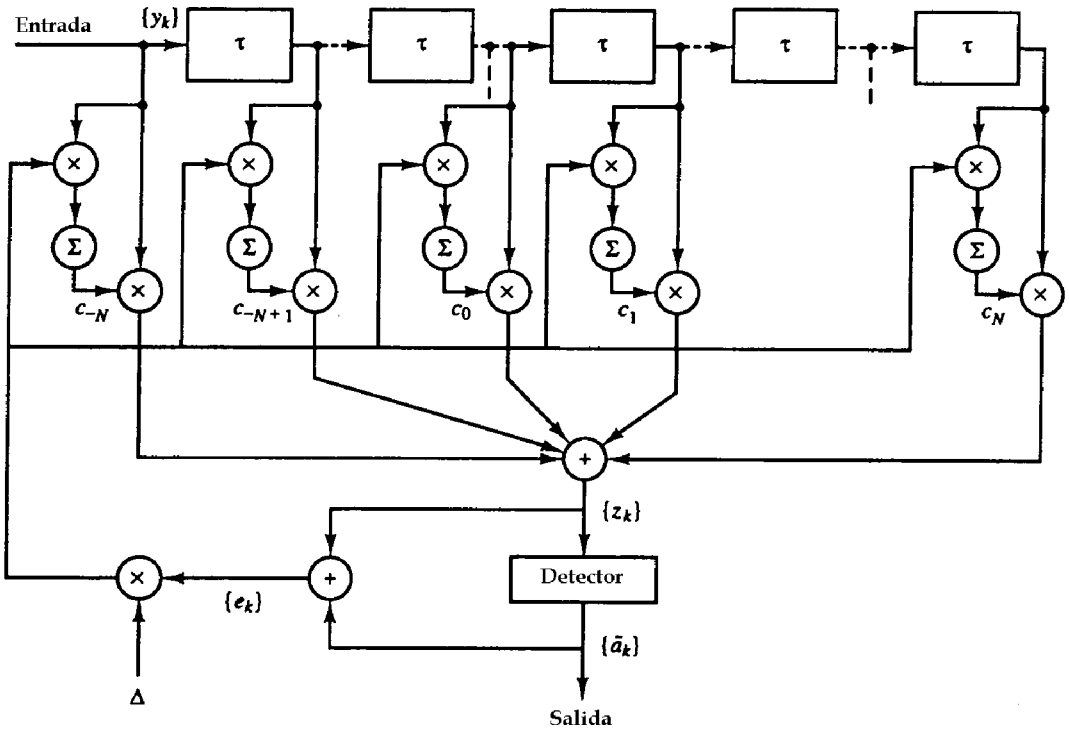

Figura 6.23: Ecualizador lineal adaptivo basado en el critério MSE.

suficientemente confiables de forma que la señal de error se forma ahora calculando la diferencia entre la salida del detector y la salida ecualizada, o sea

$$e_k = \tilde{a}_k - z_k$$

donde $\tilde{a}_k$ es la salida del detector. En general, los errores de decisión no ocurren frecuentemente y, en consecuencia, deberían tener efectos pequeños sobre el desempeño del algoritmo de seguimiento dado por (6.56).

Como regla práctica para seleccionar el parámetro de tamaño de paso que asegure convergencia y adecuada capacidad de seguimiento en canales lentamente variantes en el tiempo es posible adoptar el siguiente

$$\Delta = \frac{1}{5(2N+1)P_R}$$

donde P_R es la potencia de la señal recibida (incluído el ruido) que puede estimarse también iterativamente.

Las características de convergencia del algoritmo de gradiente estocástico se ilustran en la figura 6.24. Esos gráficos fueron obtenidos mediante simulación de un ecualizador adaptivo de 11 coeficientes operando un canal con una cantidad de ISI relativamente modesta. La potencia de señal recibida P_R fué normalizada a la unidad. Con la regla anterior para la determinación del parámetro de tamaño del paso se obtuvo $\Delta = 0.018$. El efecto de hacer muy grande Δ se ilustra con saltos grandes obtenidos en el MSE cuando $\Delta = 0.115$. Cuando Δ se disminuye la convergencia se hace algo más lenta, pero se logra un MSE menor, indicando que los coeficientes estimados están próximos al $\boldsymbol{c}_{opt}$.

Aunque se ha descripto con algún detalle la operación de un ecualizador adaptivo optimizado en base al critério MSE, la operación de un ecualizador en base al critério de ISI nula es muy similar. La principal diferencia reside en el método de estimar el vector gradiente $\boldsymbol{g}_k$ en cada iteración. Un diagrama de bloques de un ecualizar adaptivo de ISI nula se ilustra en la figura 6.25.

Ecualizador con realimentación de decisión

Los ecualizadores lineales descriptos previamente son muy efectivos en canales tal como líneas telefónicas donde la ISI no es muy severa. La severidad de la ISI está directamente relacionada con sus características

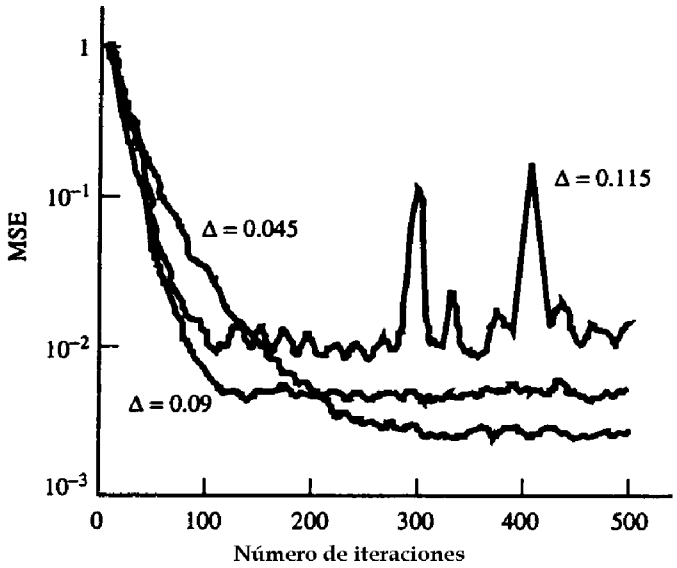

Figura 6.24: Convergencia inicial del algoritmo LMS con diferentes tamaño de paso.

espectrales y no necesariamente con la ventana temporal de la ISI. Por ejemplo, es posible considerar la ISI resultante de dos canales, que se ilustran en la figura 6.26. La ventana temporal de la ISI para el canal A es de 5 intervalos de símbolo sobre cada lado de la componente de señal deseada, la cual tiene un valor 0.72. Por otro lado, la ventana temporal de ISI en el canal B es de un intervalo de símbolo a cada lado de la componente de señal deseada, la cual tiene un valor 0.815. La energía de la respuesta total se normalizó a la unidad en ambos canales.

A pesar de la ventana temporal de ISI más corta, el canal B resulta en ISI más severa. Esto se evidencia en las características de respuesta en frecuencia de esos canales, las cuales se ilustran en la figura 6.27. Es posible observar que el canal B tiene una franja de frecuencias nulas en $f = 1/2T$, mientras que esto no ocurre en el caso del canal A. En consecuencia, un ecualizador lineal introducirá una gran ganancia en ese rango de frecuencias. De esta forma, el ruido en el canal B será mucho más amplificado que en el canal A. Esto implica que el desempeño del ecualizador lineal para el canal B será más pobre que para el canal A. En consecuencia, la limitación básica de un ecualizador lineal es que funciona inadecuadamente para canales con bandas de muy baja ganancia. Tales canales se encuentran frecuentemente en radio comunicaciones, tal como la transmisión ionosférica a frecuencias por debajo de 30 MHz y canales de radio móviles, como aquellos utilizados en las comunicaciones de radio celulares.

Un *ecualizador con realimentación de decisión* (DFE) es un ecualizador no lineal que utiliza decisiones previas para eliminar la ISI causada por símbolos previamente detectados sobre el símbolo que ha ser detectado. Un diagrama en bloques simplificado de un DFE se muestra en la figura 6.28. El DFE está compuesto por dos filtros. El primer filtro se denomina *filtro hacia adelante* (feedforward) y es generalmente un filtro FIR de espaciamiento fraccionario con coeficientes ajustables. Este filtro es idéntico en forma al ecualizador linear descripto anteriormente. Su entrada es la señal filtrada recibida $y(t)$. El segundo filtro es un *filtro de realimentación* (feedback). Está implementado como un filtro FIR con coeficientes ajustables de espaciamiento de símbolo. Su entrada es el conjunto de símbolos previamente detectados. La salida del filtro de realimentación se sustrae de la salida del filtro feedforward para formar la entrada al detector. De esta forma se tiene que

$$z_m = \sum_{n=1}^{N_1} c_n y(mT - n\tau) - \sum_{n=1}^{N_2} b_n \tilde{a}_{m-n}$$

donde $\{c_n\}$ y $\{b_n\}$ son los coeficientes ajustables de los filtros feedforward y feedback, respectivamente, $\tilde{a}_{m-n}$, $n = 1, 2, \cdots, N_2$ son los símbolos previamente detectados, N_1 es la longitud del filtro feedforward y N_2 es la longitud del filtro feedback. En base a la entrada z_m, el detector determina cual de los posibles símbolos transmitidos está más próximo en distancia a la señal de entrada z_m. De esta forma, el detector

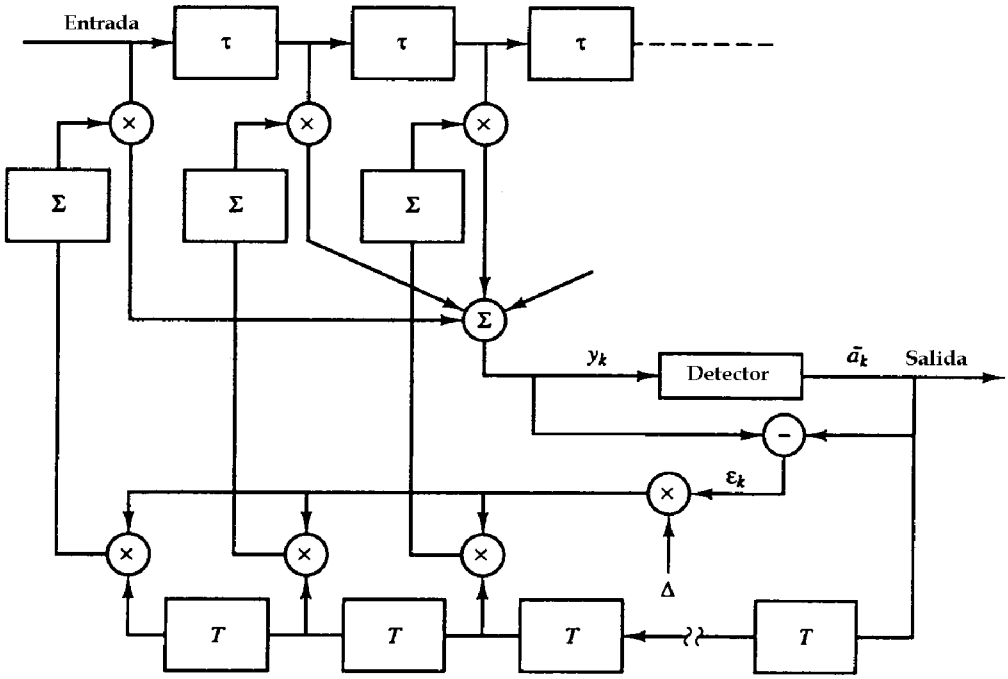

Figura 6.25: Ecualizador adaptivo de ISI nula.

realiza su decisión y obtiene como salida $\tilde{a}_m$. Lo que hace al DFE no lineal es la característica no lineal del detector, el cual provee la entrada al filtro feedback.

Los coeficientes de los filtros feedback y forward se seleccionan para optimizar alguna medida de desempeño deseada. Para implementar un DFE adaptivo, por simplicidad matemática, se aplica usualmente el critério MSE con un algoritmo de gradiente estocástico. La figura 6.29 ilustra el diagrama en bloques de un DFE adaptivo cuyos coeficientes se ajustan por medio de un algoritmo LMS. La figura 6.30 ilustra el desempeño en probabilidad de error del DFE, obtenida para transmisión PAM binaria con el canal B. La ganancia en desempeño comparada con la obtenida con un ecualizador lineal es claramente evidente.

Se debe mencionar que los errores de decisión del detector introducidos al filtro feedback tienen un efecto pequeño en el desempeño del DFE. En general debido a ellos se introduce una pérdida de desempeño de 1-2 dB para tazas de error debajo de 10^{-2}. Por eso los errores de decisión en el filtro feedback no tienen consecuencias catastróficas.

A pesar que el DFE es mejor que un ecualizador lineal, no es un ecualizador óptimo desde el punto de vista de la minimización de la probabilidad de error. Como indicado anteriormente, el detector óptimo en un sistema de comunicaciones digitales en presencia de ISI es un detector de secuencias ML. Este detector es particularmente adecuado para canales con ISI severa, cuando la ventana temporal de ISI es de unos pocos símbolos. Por ejemplo, la figura 6.31 ilustra el desempeño de probabilidad de error del algoritmo de Viterbi para una señal PAM binaria transmitida a través del canal B (ver figura 6.26). Con el propósito de comparación se ilustra también la probabilidad de error para un DFE. Ambos resultados se obtuvieron mediante simulaciones computacionales. Se observa que el desempeño del detector de secuencias de ML es alrededor de 4.5 dB mejor que el desempeño del DFE para una probabilidad de error de 10^{-4}. En consecuencia, este es un ejemplo donde el detector de secuencias ML provee una ganancia en desempeño significativa en un canal con una ventana de ISI relativamente pequeña.

En conclusión, se discutió aspectos relacionados con ecualizadores adaptivos, utilizados ampliamente en sistemas de comunicaciones digitales de alta velocidad para canales telefónicos. Los modems para líneas telefónicas de alta velocidad (para velocidades por encima de 2400 bits/seg) generalmente incluyen un ecualizador adaptivo implementado con un filtro FIR con coeficientes que se ajustan en base a un critério de MMSE. Dependiendo de la velocidad de los datos, el ecualizador contempla generalmente una ventana de entre 20 y 70 símbolos. Para el ajuste de los coeficientes del ecualizador adaptivo se utiliza un algoritmo LMS dado por (6.56).

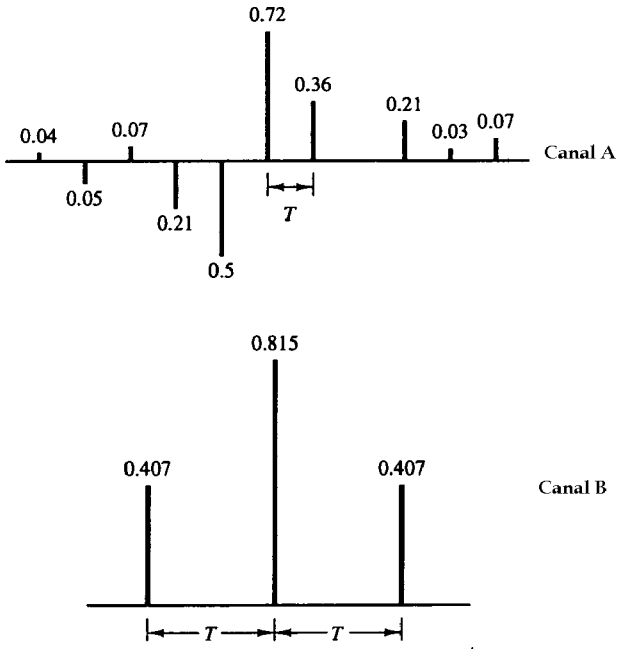

Figura 6.26: Dos canales con ISI.

6.5 Sincronización de símbolos

En un sistema de comunicaciones digital, la salida del filtro de recepción $y(t)$ debe muestrearse periodicamente a la velocidad de símbolos en los instantes precisos de muestreo $t_m = mT + \tau_0$, donde T es el intervalo de símbolo y τ_0 es un retardo de tiempo nominal que tiene en cuenta el tiempo de propagación de la señal del transmisor al receptor. Para realizar este muestreo periódico se requiere una señal de reloj en el receptor. El proceso de extraer tal señal de reloj en el receptor se denomina *sincronización de símbolos.*

La sincronización de símbolos es una de las funciones más críticas que se realizan en el receptor de un sistema de comunicaciones digital sincrónico. Debe notarse que el receptor debe conocer no solo la frecuencia $(1/T)$ a la cual se muestrean las salidas de los filtros acoplados sino que también debe conocer cuando tomar muestras dentro de cada intervalo de símbolo. La elección del instante de muestreo dentro del intervalo de símbolo de duración T se denomina *fase del temporizado.*

Visto en relación con los patrones de ojo mostrados en la figura 6.7, la mejor fase de temporizado corresponde al instante de tiempo dentro del intervalo de símbolo donde el patrón de ojo tiene un máximo. En un sistema de comunicaciones práctico el reloj receptor debe ajustarse continuamente en frecuencia $(1/T)$ y en fase de temporizado τ_0 para compensar los desplazamientos de frecuencia entre los osciladores utilizados en los relojes de transmisión y recepción y, de esta forma, optimizar los instantes de tiempo de muestreo a la salida del filtro acoplado o correlador.

La sincronización de símbolos puede lograrse de varias formas. En algunos sistemas de comunicaciones los relojes de transmisión y recepción están sincronizados con un reloj maestro, que provee una señal de temporizado muy precisa. En este caso, el receptor debe estimar y compensar el retardo de tiempo relativo entre las señales transmitidas y recibidas. Ese puede ser el caso en sistemas de comunicaciones de radio que operan en la banda VLF (debajo de 30 kHz), donde las señales de reloj precisas son transmitidas a partir de una estación de radio maestra.

Otro método para lograr la sincronización de símbolos es que el transmisor emita simultaneamente la frecuencia de reloj $1/T$ junto con la señal de información. El receptor debe simplemente utilizar un filtro de banda estrecha sintonizado a la frecuencia del reloj de transmisión y, de esa forma, extraer la señal de reloj para el muestreo. Esta aproximación tiene la ventaja de ser simple de implementar. Existen, sin embargo, varias desventajas. Una es que el transmisor debe asignar cierta porción de su potencia disponible para la transmisión de la señal de reloj. Otra es que una pequeña fracción del ancho de banda del canal debe asignarse a la transmisión de la señal de reloj. A pesar de esas desventajas, este método se utiliza frecuentemente en sistemas de transmisión telefónica que emplean grandes anchos de banda para transmitir

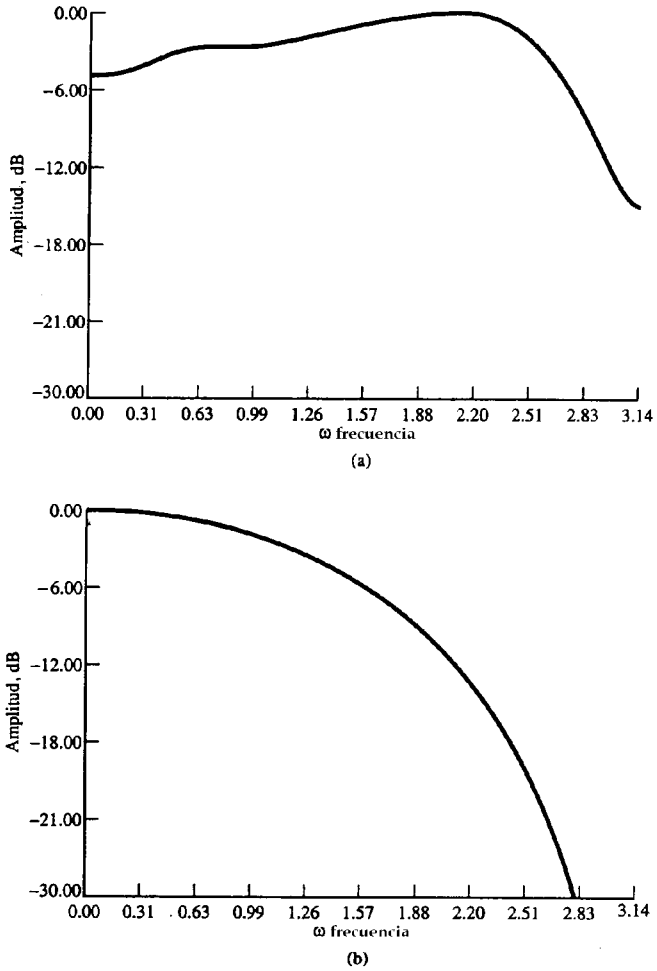

Figura 6.27: Espectro de amplitud para (a) el canal A (figura 6.26a) y (b) el canal B (figura 6.26b).

las señales asociadas a varios usuarios. En tal caso la transmisión de la señal de reloj se comparte en la demodulación de las señales entre los varios usuarios. A través de este uso compartido de la señal de reloj la penalidad en potencia transmitida y ancho de banda asignado se reduce proporcionalmente al número de usuarios.

Una señal de reloj puede extraerse también de la señal de datos recibida. Existen cierto número de métodos diferentes que puede utilizarse en el receptor para lograr la autosincronización. A continuación se discutirán cuatro aproximaciones al problema de lograr sincronización de símbolos a partir de la señal recibida. Se asumirá que la señal recibida a la salida del filtro receptor $g_R(t)$ tiene la siguiente forma general

$$y(t) = \sum_{n=-\infty}^{\infty} a_n x(t - nT - \tau_0) + \nu(t)$$

donde $x(t) = g_T(t) * c(t) * g_R(t)$, $\{a_n\}$ es la secuencia de símbolos de información (binaria o M-ária), $\nu(t)$ representa el ruido a la salida del filtro de recepción, y τ_0 representa la fase de temporizado. Se desea obtener una señal de reloj con una frecuencia igual a la velocidad de transmisión $1/T$ y una fase de temporizado τ_0. La secuencia de datos $\{a_n\}$ se supone una secuencia estacionaria de media cero con elementos estadisticamente i.i.d. En consecuencia, la señal PAM resultante

$$v(t) = \sum_{n=-\infty}^{\infty} a_n x(t - nT - \tau_0)$$

es un proceso estocástico cicloestacionario de media cero.

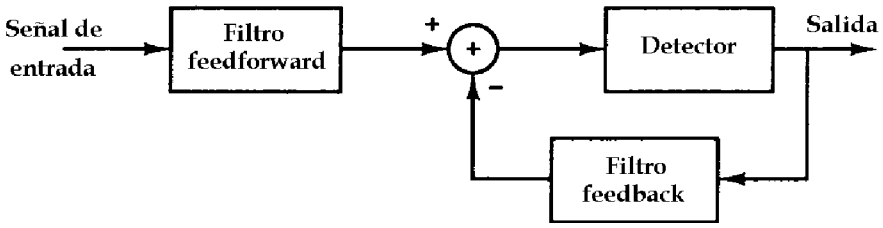

Figura 6.28: Diagrama en bloques del DFE.

6.5.1 Métodos de línea espectral

Dado que la componente de señal $v(t)$ en la señal recibida $y(t)$ es periódica con período T, es posible recuperar una señal de reloj con frecuencia $1/T$ filtrando una componente de señal en $f = 1/T$. Es posible observar, sin embargo, que $E[v(t)] = 0$, dado que $E(a_n) = 0$. En consecuencia, $v(t)$ no puede utilizarse directamente para generar una frecuencia en $f = 1/T$. Por otro lado, es posible realizar una operación no lineal sobre $v(t)$ para generar potencia en $f = 1/T$ y sus armónicas.

Se considerará por ejemplo una no linealidad de ley cuadrática. En base al cuadrado de la señal $v(t)$ y su valor esperado con respecto a la secuencia de datos $\{a_n\}$ se obtiene

$$E[v^2(t)] = E\left[\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} a_n a_m x(t-mT-\tau_0)x(t-nT-\tau_0)\right] \tag{6.57}$$

$$= \sigma_a^2 \sum_{n=-\infty}^{\infty} x^2(t-nT-\tau_0) \tag{6.58}$$

donde $\sigma_a^2 = E[a_n^2]$. Dado que $E[v^2(t)] > 0$, es posible usar $v^2(t)$ para generar la componente de frecuencia deseada.

Se utilizará la Fórmula de Suma de Poisson[3] para expresar (6.58) en la forma de una serie de Fourier, entonces

$$E[v^2(t)] = \sigma_a^2 \sum_{n=-\infty}^{\infty} x^2(t-nT-\tau_0) = \frac{\sigma_a^2}{T}\sum_m c_m e^{j2\pi m(t-\tau_0)/T} \tag{6.59}$$

donde

$$c_m = \int_{-\infty}^{\infty} X(f)X(m/T-f)df \tag{6.60}$$

El espectro de la señal transmitida está confinado, por diseño, a las frecuencias por debajo de $1/T$. En consecuencia, $X(f) = 0$ para $|f| > 1/T$ y entonces, existirán solamente tres términos no nulos en (6.59) para $m = 0, \pm 1$. De esa forma, la señal $v^2(t)$ contiene una componente de cc y una componente en la frecuencia $1/T$.

El desarrollo anterior sugiere elevar al cuadrado la señal $y(t)$ a la salida del filtro de recepción y filtrar $y^2(t)$ con un filtro de banda estrecha $B(f)$ sintonizado a la velocidad de símbolos $1/T$. Haciendo la respuesta del filtro $B(1/T) = 1$, se tiene que

$$E[v^2(t)] = \frac{\sigma_a^2}{T}\mathrm{Re}\left[c_1 e^{j2\pi(t-\tau_0)/T}\right] = \frac{\sigma_a^2}{T}c_1 \cos\frac{2\pi}{T}(t-\tau_0) \tag{6.61}$$

[3] A partir de $\sum_{-\infty}^{\infty}\delta(t-nT) = \frac{1}{T}\sum_{-\infty}^{\infty} e^{jn(2\pi t)/T}$ es posible mostrar que para cualquier $x(t)$ y T, $\sum_{-\infty}^{\infty} x(t-nT) = \frac{1}{T}\sum_{-\infty}^{\infty} X\left(\frac{n}{T}\right)e^{jn(2\pi t)/T}$, de donde para $t = 0$ se obtiene la Fórmula de Suma de Poisson deseada.

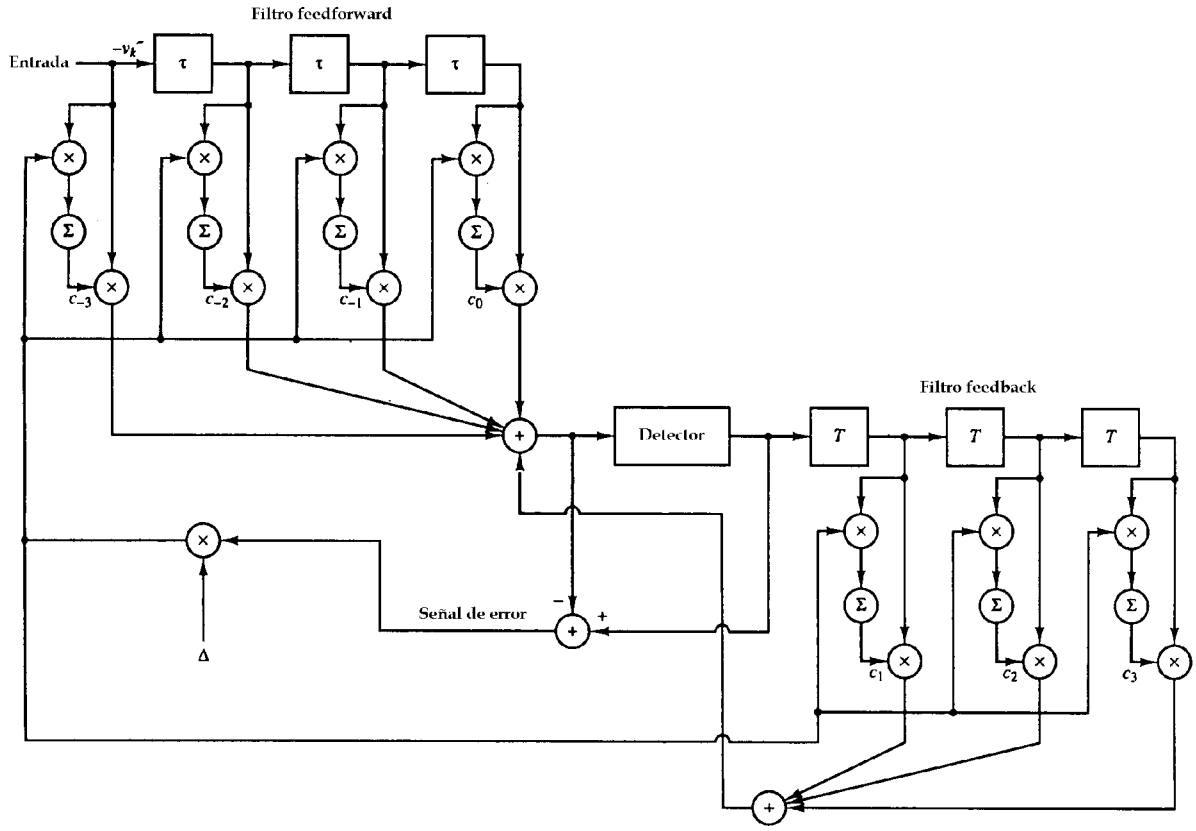

Figura 6.29: DFE adaptivo.

tal que la señal de temporizado es una sinusoide con una fase de $-2\pi\tau_0/T$, asumiendo que $X(f)$ sea real. Es posible usar los cruces por cero de la señal de temporizado como una indicación de los instantes de muestreo correctos. Sin embargo, los cruces por cero de la señal dada por (6.61) ocurren en

$$\frac{2\pi}{T}(t-\tau_0) = (4k+1)\frac{\pi}{2}$$

o equivalentemente

$$t = kT + \tau_0 + \frac{T}{4}$$

los cuales están desplazados en tiempo en $T/4$ en relación a los cruces por cero deseados. En un sistema práctico el desplazamiento de temporizado puede compensarse facilmente mediante un circuito de reloj relativamente simple o mediante el diseño del filtro pasabanda $B(f)$, de forma que tenga un desplazamiento de fase de $\pi/2$ en $f = 1/T$. La figura 6.32 ilustra este método para la generación de la señal de temporizado en el receptor.

El ruido aditivo que corrompe la señal causará generalmente fluctuaciones en los cruces por cero de la señal deseada. El efecto de las fluctuaciones dependerá de la amplitud c_1 de la señal de temporizado sinusoidal media dada por (6.61). Notar que la amplitud de señal c_1 es proporcional a la pendiente de la señal de temporizado en la vecindad del cruce por cero, como mostrado en la figura 6.33. En consecuencia, cuanto mayor sea la amplitud c_1, mayor será la pendiente y de esa forma los errores de temporizado debidos al ruido serán menores. De (6.60) se observa que c_1 depende de la cantidad de solapamiento espectral de

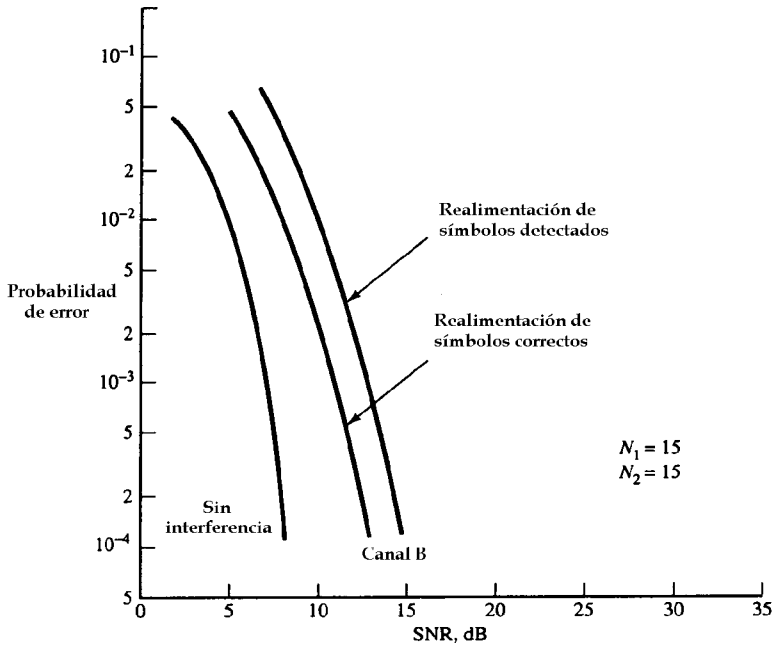

Figura 6.30: Desempeño del DFE con y sin error de propagación.

$X(f)$ y $X(1/T - f)$. De esta forma, c_1 depende de la cantidad en que el ancho de banda de $X(f)$ excede el ancho de banda de Nyquist $1/2T$, o sea, c_1 depende del ancho de banda en exceso. Si el exceso de ancho de banda es cero, o sea $X(f) = 0$ para $|f| > 1/2T$, entonces $c_1 = 0$, y este método no permite obtener una señal de temporizado. Si el exceso de ancho de banda es grande, por ejemplo de 50 % a 100 %, la amplitud de la señal de temporizado será suficientemente grande para producir estimaciones del temporizado de símbolos suficientemente precisas.

6.5.2 Sincronizador por simetría (temprano-tarde)

Existe otro método para generar una señal de temporizado en el receptor que explota las propiedades de simetría de la señal a la salida del filtro acoplado o correlador. Para describir este método es posible considerar el pulso rectangular $s(t)$, $0 \leq t \leq T$, mostrado en la figura 6.34a. La salida del filtro acoplado a $s(t)$ logra su máximo valor en el instante $t = T$, como mostrado en la figura 6.34b. De esta forma, la salida del filtro acoplado es la función autocorrelación temporal del pulso $s(t)$. Por supuesto, esto es cierto para una forma de pulso arbitrária. Claramente, el instante adecuado para muestrear la salida del filtro acoplado para obtener una salida máxima es en $t = T$, o sea, en el pico de la función correlación.

En presencia de ruido, la identificación del valor pico de la señal es generalmente dificil. En lugar de muestrear la señal en el pico, es posible suponer el muestreo antes, en $t = T - \delta T$ y después en $t = T + \delta T$. Los valores absolutos de las muestras tempranas $|y[m(t - \delta T)]|$ y las muestras tardías $|y[m(t + \delta T)]|$ serán menores (en el promedio en presencia de ruido) que las muestras en el valor pico $|y(mT)|$. Dado que la función autocorrelación es par con respecto al valor de muestra óptimo en $t = T$, los valores absolutos de la función correlación en $t = T - \delta T$ y $t = T + \delta T$ son iguales. Bajo esta condición, el instante de muestreo apropiado es el punto medio entre $t = T - \delta T$ y $t = T + \delta T$. Esta condición es la base para el *sincronizador de símbolos temprano - tarde.*

La figura 6.35 ilustra el diagrama en bloques de un sincronizador temprano - tarde. En esta figura se utilizan correladores en lugar de filtros acoplados equivalentes. Los dos correladores integran sobre un intervalo de símbolo T, pero un correlador comienza a integrar δT más temprano en relación al instante de muestreo óptimo estimado y el otro integrador comienza a integrar δT tarde en relación al instante de muestreo óptimo estimado. Se forma una señal de error en base a la diferencia entre los valores absolutos de las dos salidas de los integradores. Para suavizar el ruido que perturba a las muestras de señal, la señal de error se pasa a través de un filtro pasabajos. Si el temporizado está lejos del instante de muestreo óptimo, la señal de error promedio a la salida del filtro pasabajos no es cero, y la señal de reloj se retarda o avanza,

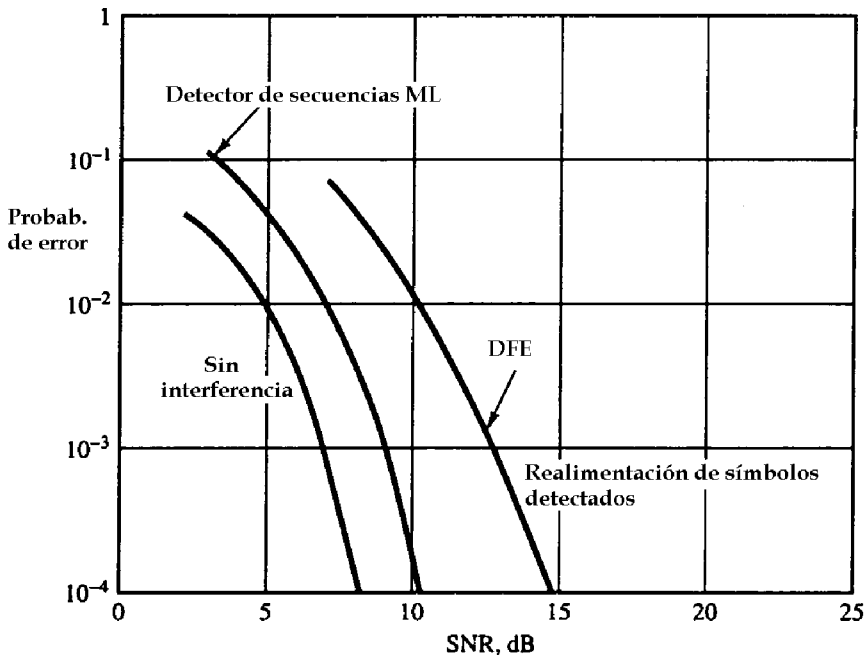

Figura 6.31: Desempeño del detector de Viterbi y el DFE para el canal B.

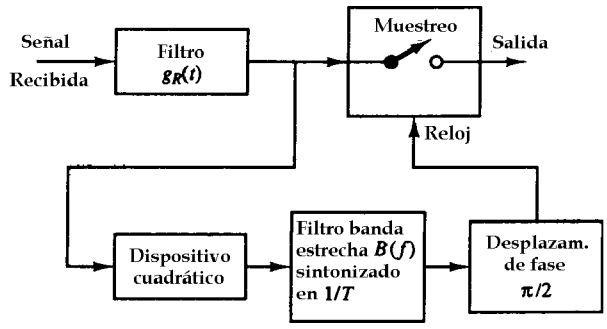

Figura 6.32: Temporizado de símbolos en base a un método de línea espectral.

dependiendo del signo del error. De esta forma, la señal de error suavizada se utiliza para alimentar un oscilador controlado por tensión (VCO), cuya salida es la señal de reloj deseada para ser utilizada en el muestreo. La salida del VCO se utiliza también como señal de reloj para un generador de forma de onda de símbolos que permite obtener la misma forma básica del pulso a la salida del filtro de transmisión. La forma de onda de pulso se avanza o retarda y luego se alimenta a los dos correladores, como mostrado en la figura 6.35. Notar que si los pulsos de señal son rectangulares, no existe necesidad de un generador de pulsos de señal dentro del lazo.

Se observa que el sincronizador temprano - tarde es basicamente un sistema de control de lazo cerrado cuyo ancho de banda es relativamente estrecho comparado con la velocidad de símbolos $1/T$. El ancho de banda del lazo determina la calidad de la estimación del temporizado. Un lazo de banda estrecha provee mejor promediación sobre el ruido aditivo y, de esa forma, mejora la calidad de los instantes de muestreo estimados, previendo que el retardo de propagación del canal y el oscilador del reloj en el transmisor no se desplazen con el tiempo. Por otro lado, si el retardo de propagación del canal está cambiando con el tiempo y/o el reloj del transmisor está desplazandose en el tiempo, entonces el ancho de banda del lazo debe incrementarse para proveer un seguimiento más rápido de las variaciones en el tiempo del temporizado de símbolo.

En el modo de seguimiento, los dos correladores están afectados por símbolos adyacentes. Sin embargo, si la secuencia de símbolos de información tiene media cero, como es el caso para PAM y otros formatos de

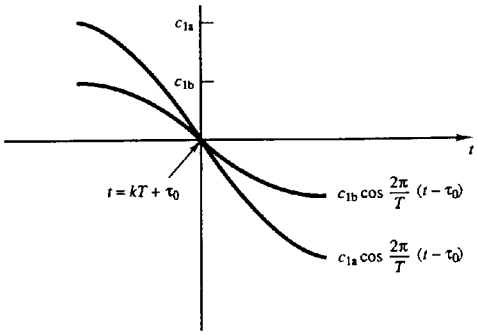

Figura 6.33: Ilustración de la pendiente de una sinusoide en el cruce por cero en función de la amplitud.

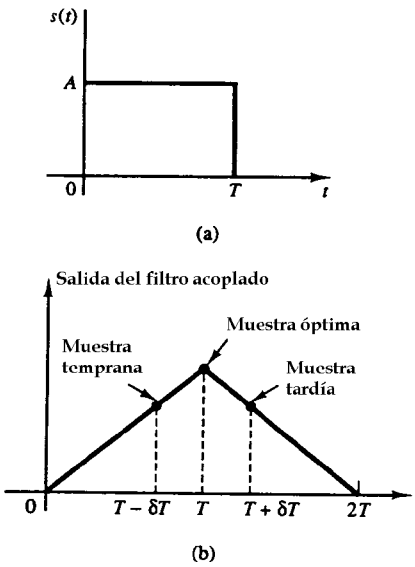

Figura 6.34: (a) Señal pulso rectangular, (b) la salida de su filtro acoplado.

señalización, la contribución a la salida de los correladores a partir de símbolos adyacentes tiene promedio cero en el filtro pasabajos.

Una realización equivalente del sincronizador temprano - tarde algo más simple de implementar se muestra en la figura 6.36. En este caso el reloj del VCO se avanza y se retarda en δT, y esas señales de reloj se utilizan para muestrear las salidas de los dos correladores.

6.5.3 Método del error medio cuadrático mínimo (*)

Otra aproximación al problema de sincronización a partir de la señal recibida está basada en la minimización del MSE entre la salida del filtro de recepción y los símbolos deseados. En la práctica el MSE, dado por

$$MSE = E\{[y_m(\tau_0) - a_m]^2\}$$

se aproxima por el promedio temporal de errores al cuadrado dado por

$$< MSE >= \sum_m [y_m(\tau_0) - \hat{a}_m]^2 \tag{6.62}$$

donde

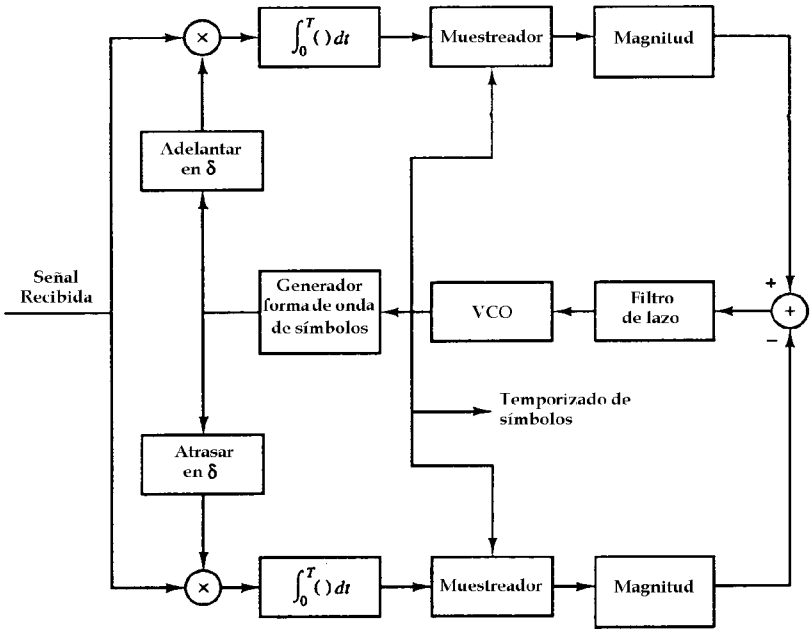

Figura 6.35: Diagrama en bloques de un sincronizador temprano - tarde.

$$y_m(\tau_0) = \sum_n a_n x(mT - nT - \tau_0) + \nu(mT) \tag{6.63}$$

y $\hat{a}_m$ es el símbolo a la salida del detector.

El MSE mínimo con respecto a la fase de temporizado τ_0 se obtiene derivando (6.62) con respecto a τ_0. De esta forma, se obtiene la condición necesaria

$$\sum_m [y_m(\tau_0) - \hat{a}_m] \frac{dy_m(\tau_0)}{d\tau_0} = 0$$

Una interpretación de esta condición necesaria es que el instante de muestreo óptimo corresponde a la condición en que la señal de error $[y_m(\tau_0) - \hat{a}_m]$ no esté correlacionada con la derivada $dy_m(\tau_0)/d\tau_0$. Dado que la salida del detector se utiliza en la formación de la señal de error, este método de estimación de la fase de temporizado se dice que está *dirigido por la decisión.*

La figura 6.37 ilustra la implementación de un sistema que está basado en la condición anterior. Notar que la operación de sumatoria se implementa como un filtor pasabajos, el cual promedia cierto número de símbolos. El tiempo promedio es rigurosamente igual al recíproco del ancho de banda del filtro. La salida del filtro se conduce a un oscilador controlado por tensión (VCO), lo cual provee la mejor estimación MSE de la fase de temporizado τ_0.

6.5.4 Método de Máxima Verosimilitud (*)

En el critério ML el temporizado óptimo de símbolo se obtiene maximizando la *función verosimilitud* dada por

$$\Lambda(\tau_0) = \sum_m a_m y_m(\tau_0)$$

donde $y_m(\tau_0)$ es la salida muestreada del filtro de recepción dada por (6.63). Desde un punto de vista matemático, la función verosimilitud es proporcional a la probabilidad de la señal recibida condicionada por

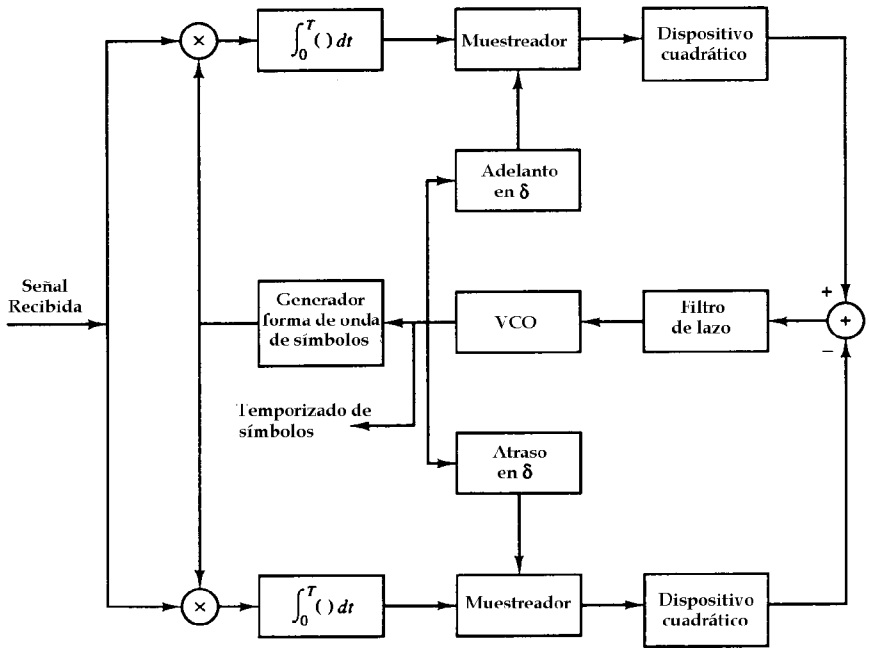

Figura 6.36: Diagrama en bloques de forma alternativa de un sincronizador temprano - tarde.

la señal transmitida conocida. Físicamente, $\Lambda(\tau_0)$ es simplemente la salida del filtro acoplado o correlador en el receptor promediada sobre cierto número de símbolos.

Una condición necesaria para que τ_0 sea una estimación ML es que

$$\frac{d\Lambda(\tau_0)}{d\tau_0} = \sum_m a_m \frac{dy_m(\tau_0)}{d\tau_0} = 0$$

Este resultado sugiere la implementación de un lazo de seguimiento como mostrado en la figura 6.38. Se observa que el producto de la salida del detector $\hat{a}_m$ con la entrada $dy_m(\tau_0)/d\tau_0$ se promedia mediante un filtro pasabajos cuya salida se conduce al VCO. Dado que la salida del detector se utiliza en el método de estimación, la estimación $\hat{\tau}$ está dirigida por las decisiones.

Como alternativa a la utilización de los símbolos de salida del detector es posible usar un *método de decisión no dirigida* que no requiera el conocimiento de los símbolos de información. Este método se basa en la promediación sobre la estadística de los símbolos. Por ejemplo, es posible elevar al cuadrado la salida del filtro de recepción y maximizar la función

$$\Lambda_2(\tau_0) = \sum_m y_m^2(\tau_0)$$

con respecto a τ_0. De esta forma se obtiene

$$\frac{d\Lambda_2(\tau_0)}{d\tau_0} = 2\sum_m y_m(\tau_0)\frac{dy_m(\tau_0)}{d\tau_0} = 0$$

La condición para el τ_0 óptimo de esta ecuación puede satisfacerse mediante la implementación mostrada en la figura 6.39. En este caso, no existe necesidad de conocer la secuencia de datos $\{a_m\}$. En consecuencia, el método no está dirigido por las decisiones.

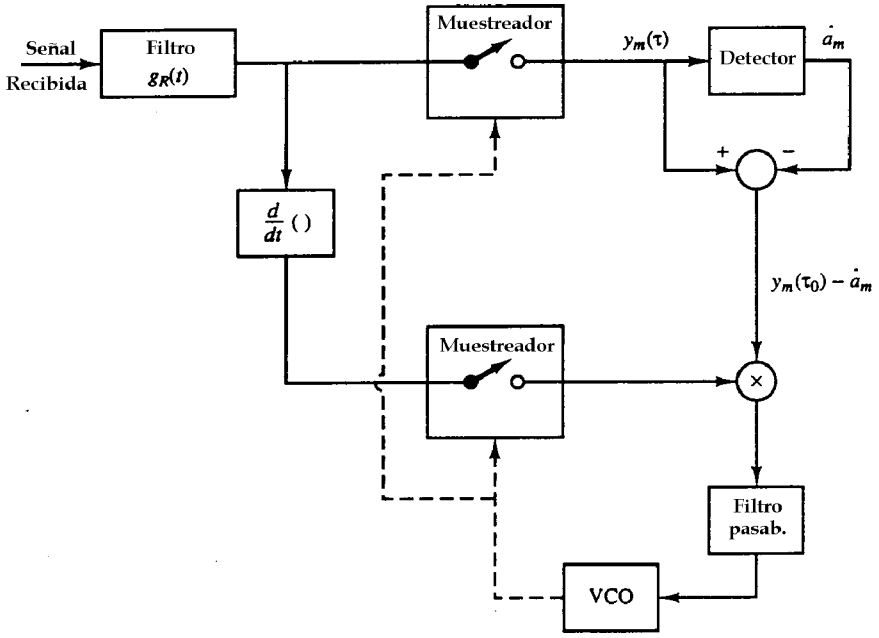

Figura 6.37: Recuperación del temporizado en base a la minimización del MSE.

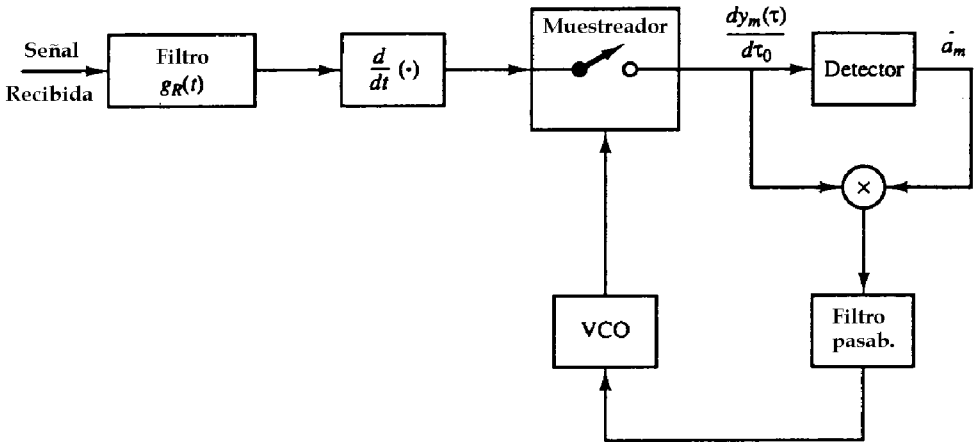

Figura 6.38: Recuperación del temporizado ML dirigida por las decisiones para señales PAM banda base.

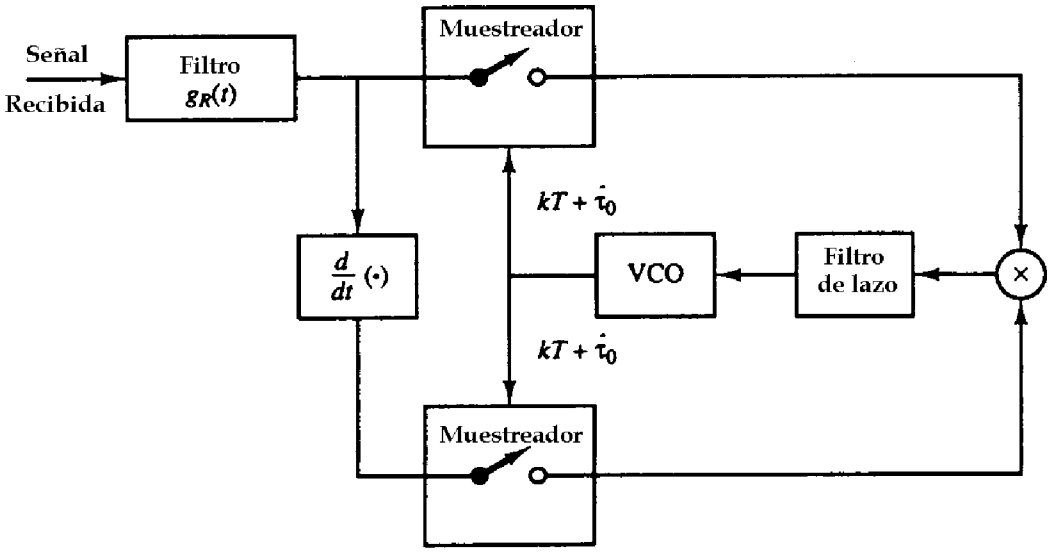

Figura 6.39: Estimación del temporizado para señales PAM banda base no dirigido por las decisiones.