



Universidad
Carlos III de Madrid

OPENCOURSEWARE
CONSORTIUM

Teoría de la Comunicación

Grado en Ingeniería en Tecnologías de Telecomunicación

OpenCourseWare
Universidad Carlos III de Madrid
<http://ocw.uc3m.es>

MARCELINO LÁZARO

Licencia *Creative Commons*



Curso abierto de la Universidad Carlos III de Madrid para la asignatura *Teoría de la Comunicación* de los programas de los Grados en Ingeniería relacionados con las telecomunicaciones

Índice general

1. Introducción	1
1.1. Definición de un sistema de comunicaciones	1
1.2. Elementos funcionales básicos de un sistema de comunicaciones	2
1.2.1. Fuente de Información	3
1.2.2. Transmisor	3
1.2.3. Canal	4
1.2.4. Receptor	6
1.3. Sistemas de comunicaciones analógicos y digitales	6
1.4. Diseño de un sistema de comunicaciones	8
1.4.1. Calidad requerida	8
1.4.2. Recursos disponibles	9
1.4.3. Coste del sistema y tecnologías existentes	10
1.5. Objetivos y organización de la asignatura	10
1.5.1. Objetivos de la asignatura	10
1.5.2. Organización de la asignatura	11
1.5.3. Bibliografía recomendada	11
2. Ruido en los sistemas de comunicaciones	15
2.1. Probabilidad	15
2.1.1. Espacio de probabilidad	16
2.1.2. Probabilidad condicional	18
2.2. Variable aleatoria	20
2.2.1. Función de distribución	21

2.2.2.	Función de densidad de probabilidad	23
2.2.3.	Variables aleatorias de interés	25
2.2.4.	Funciones de una variable aleatoria	29
2.2.5.	Momentos estadísticos	30
2.2.6.	Variables aleatorias multidimensionales	31
2.3.	Procesos Aleatorios	37
2.3.1.	Descripción de un proceso aleatorio	39
2.3.2.	Esperanzas de los procesos (promedios estadísticos)	40
2.3.3.	Estacionariedad y cicloestacionariedad	41
2.3.4.	Ergodicidad	43
2.3.5.	Potencia y energía	45
2.3.6.	Procesos aleatorios multidimensionales (múltiples)	46
2.3.7.	Procesos aleatorios en el dominio de la frecuencia	47
2.3.8.	Procesos aleatorios estacionarios y sistemas lineales	52
2.3.9.	Suma de procesos aleatorios	56
2.4.	Modelo del ruido térmico: procesos blancos y gaussianos	57
2.4.1.	Procesos aleatorios gaussianos	58
2.4.2.	Procesos aleatorios blancos	60
2.4.3.	Modelo de ruido térmico	61
2.4.4.	Ruido filtrado y ancho de banda equivalente de ruido	62
2.4.5.	Relación señal a ruido	65
2.5.	Muestreo de procesos aleatorios limitados en banda	66
2.6.	Ejercicios	69
2.7.	Soluciones de los ejercicios	76
3.	Modulaciones Analógicas	85
3.1.	Introducción al concepto de modulación	85
3.1.1.	Notación básica y modelos de señal moduladora	87
3.2.	Modulaciones de amplitud (AM)	88
3.2.1.	AM convencional	89

3.2.2.	Doble banda lateral sin portadora (DBL)	97
3.2.3.	Modulación de banda lateral única (BLU)	103
3.2.4.	Modulación de banda lateral vestigial (BLV)	109
3.2.5.	Resumen de características y comparativa entre las distintas modulaciones de amplitud	112
3.3.	Modulaciones Angulares	114
3.3.1.	Representación de las señales FM y PM	115
3.3.2.	Índices de modulación	119
3.3.3.	Características espectrales de una modulación angular	120
3.3.4.	Modulación de las señales FM y PM	125
3.3.5.	Demodulación de las señales FM y PM	127
3.4.	Ruido en sistemas de comunicaciones analógicas	127
3.4.1.	Relación señal a ruido en una transmisión en banda base	128
3.4.2.	Efecto del ruido en modulaciones de amplitud	129
3.4.3.	Efecto del ruido en modulaciones angulares	137
3.5.	Ejercicios	139
3.6.	Soluciones de los ejercicios	145
4.	Modulación y detección en canales gaussianos	153
4.1.	Introducción	153
4.1.1.	Ventajas de los sistemas de comunicaciones digitales	153
4.1.2.	Descripción general de un sistema de comunicaciones digitales	155
4.1.3.	Diseño generico de un modulador digital y notación básica	157
4.1.4.	Transmisión a través de un canal de comunicaciones	159
4.1.5.	Diseño generico de un demodulador digital y notación básica	160
4.1.6.	Factores a considerar a la hora de elegir las M formas de onda	162
4.2.	Representación geométrica de las señales	164
4.2.1.	Espacios vectoriales	164
4.2.2.	Espacios de Hilbert para señales de energía finita	165
4.2.3.	Representación de una señal en una base del espacio vectorial	168
4.2.4.	Procedimiento de ortogonalización de Gram-Schmidt	169

4.3.	Modelo de comunicación digital	176
4.3.1.	Ejemplo para ilustrar la ventaja de la representación vectorial en el diseño de un sistema	179
4.4.	Demodulador	186
4.4.1.	Demodulador por correlación	187
4.4.2.	El filtro adaptado	188
4.4.3.	Extensión de las estructuras del demodulador para cualquier instante n . . .	193
4.4.4.	Caracterización estadística de la salida del demodulador en el caso de transmisión sobre un canal gaussiano	193
4.4.5.	Canal discreto equivalente	196
4.5.	Decisor	197
4.5.1.	Diseño del decisor - Regiones de decisión	198
4.5.2.	Obtención del decisor óptimo	198
4.5.3.	Cálculo de las probabilidades de error	205
4.5.4.	Aproximación y cotas para la probabilidad de error	222
4.5.5.	Expresiones de la probabilidad de error en función de E_s/N_0	229
4.6.	Codificador	230
4.6.1.	Diseño del codificador	230
4.6.2.	Diseño de la constelación	231
4.6.3.	Constelaciones utilizadas en sistemas de comunicaciones	237
4.6.4.	Asignación binaria - Codificación de Gray y cálculo de la BER	240
4.6.5.	Relación entre tasa de bit y tasa de símbolo	245
4.7.	Modulador	246
4.7.1.	Diseño del modulador	247
4.7.2.	Algunos ejemplos de moduladores y señales moduladas que generan	248
4.8.	Ejercicios	254
4.9.	Soluciones de los ejercicios	264
5.	Límites fundamentales en los sistemas de comunicaciones digitales	281
5.1.	Modelado de las fuentes de información	283
5.1.1.	Fuentes analógicas	283

5.1.2. Fuentes digitales	284
5.2. Modelos probabilísticos de canal	285
5.2.1. Canal gaussiano	287
5.2.2. Canal gaussiano con entrada digital	288
5.2.3. Canal digital	289
5.2.4. Canal digital binario	296
5.3. Medidas cuantitativas de información	298
5.3.1. Información y entropía	298
5.3.2. Entropía conjunta	302
5.3.3. Entropía condicional	303
5.3.4. Información mutua	305
5.3.5. Entropía diferencial e información mutua	308
5.4. Capacidad de canal	310
5.4.1. Codificación de canal para una transmisión fiable	311
5.4.2. Capacidad de canal para el canal digital	317
5.4.3. Capacidad de canal para el canal gaussiano	322
5.5. Límites en un sistema digital de comunicaciones	326
5.6. Ejercicios	330
5.7. Soluciones de los ejercicios	338
A. Tablas de interés	347

Capítulo 1

Introducción

Este capítulo presenta una introducción a la asignatura, que trata sobre los principios teóricos básicos que se aplican en el diseño y análisis de sistemas de comunicaciones. El capítulo comienza definiendo lo que es un sistema de comunicaciones, enumerando los elementos funcionales básicos que lo forman, y describiendo brevemente la principal función de cada uno de ellos. Posteriormente, se realizan algunas clasificaciones generales de los sistemas de comunicaciones en distintos tipos, y se discuten varios aspectos relacionados con el diseño y análisis de un sistema de comunicaciones, para finalizar presentando los objetivos de la asignatura y los contenidos de la misma.

1.1. Definición de un sistema de comunicaciones

Un sistema de comunicaciones tiene como fin la *transmisión de información* entre dos puntos separados por una distancia pero unidos físicamente por alguna estructura física (natural o artificial) susceptible de ser empleada para ello.

Se puede por tanto definir la transmisión como el proceso de enviar o transportar información de un punto (fuente) hasta otro punto (destino) a través de un canal o medio de transmisión, como se ilustra en la Figura 1.1.



Figura 1.1: Representación general de un sistema de comunicaciones.

El medio de transmisión puede ser cualquier elemento que permita realizar el envío de información entre una fuente y un destino: un cable de pares, una guía de onda, un cable coaxial, una fibra óptica, o la propia atmósfera (utilizando todo el espectro radioeléctrico) son algunos ejemplos habituales. Pero también se pueden considerar medios de transmisión algunos dispositivos de almacenamiento, como CD-ROMS, DVDs, etc., que permiten transportar información entre dos puntos.

En general las fuentes de información no están adecuadas para poder transmitir de forma directa la información que generan utilizando cualquier medio de transmisión; normalmente es necesario

procesar la manifestación física de la fuente para poder introducir la información que contiene de forma eficiente en el medio o canal. La mayoría de medios de transmisión están concebidos para la transmisión de señales eléctricas o electromagnéticas, mientras que la manifestación física de las fuentes no es en muchos casos de naturaleza eléctrica. Por poner ejemplo sencillo, en una fuente de audio (voz, música, etc.), la manifestación física de la información consiste en ondas de presión acústica que viajan a través del aire. Para poder transmitir este tipo de información, en primer lugar es necesario convertir la manifestación física de la información en una señal eléctrica en el lado de la fuente. Esto se hace mediante un *transductor*. El ejemplo clásico de transductor para señales de audio es un micrófono, que transforma las ondas de presión acústica en una señal eléctrica que representa la variación a lo largo del tiempo de la presión acústica. La Figura 1.2 muestra un ejemplo de señal eléctrica asociada a una señal de voz.

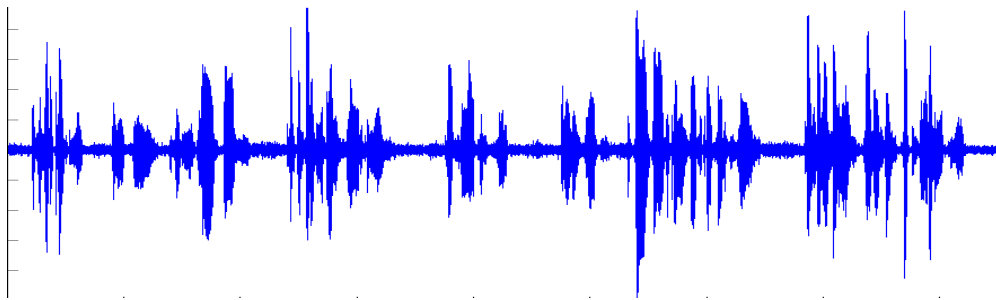


Figura 1.2: Ejemplo de señal eléctrica asociada a una señal de voz.

Una vez que la manifestación física de la información se ha convertido en una señal eléctrica adecuada para representarla, dicha señal eléctrica será procesada por el sistema de comunicaciones, y enviada a través de un cierto medio de transmisión. En el lado del receptor será de nuevo procesada para convertir la señal eléctrica recibida en la correspondiente manifestación física de la información.

1.2. Elementos funcionales básicos de un sistema de comunicaciones

En el apartado anterior se ha definido un sistema de comunicaciones y se han descrito brevemente los principales pasos a seguir para la transmisión de información entre dos puntos. En este apartado, esos pasos elementales a seguir se concretarán en la definición de los elementos funcionales básicos que van a formar parte de un sistema de comunicaciones.

Un sistema de comunicaciones está compuesto por múltiples elementos, pero desde un punto de vista funcional, los elementos básicos son los cinco que muestra la Figura 1.3

- Fuente de información
- Transmisor
- Canal
- Receptor
- Destino de la información

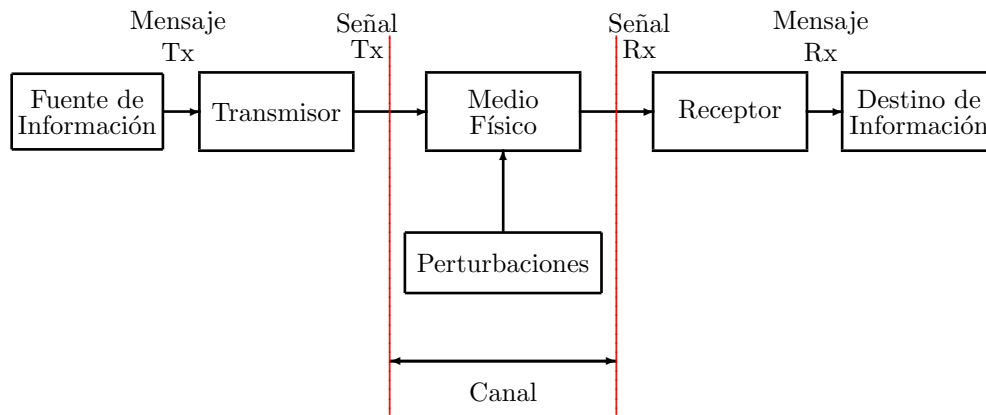


Figura 1.3: Esquema funcional de un sistema de comunicaciones.

A continuación se describe brevemente cada uno de estos elementos funcionales.

1.2.1. Fuente de Información

Como indica su nombre, la fuente de información genera la información que se pretende transmitir al otro extremo del sistema de comunicaciones. El mensaje a transmitir será la manifestación física de la información producida por la fuente. Como ya se ha comentado anteriormente, independientemente del tipo de fuente, lo habitual es que un transductor convierta la información en una señal de tipo eléctrico que la representa (por ejemplo, un micrófono convierte una señal acústica en una señal eléctrica, o una cámara de vídeo convierte una secuencia de imágenes en otra señal eléctrica).

Existen muchos tipos de fuentes de información. Si se clasifican desde el punto de vista del formato en que se representa la información, se pueden dividir en fuentes analógicas o digitales. Una fuente analógica produce una información que se manifiesta en variaciones con continuidad de algún tipo de magnitud física. Algunos ejemplos serían la presión en el aire al hablar o la variación continua de temperatura de un termómetro a lo largo del tiempo. En este tipo de fuentes el mensaje a transmitir está representado por una forma de onda continua que representa la variación de la correspondiente magnitud física, como en el ejemplo de la señal de voz de la Figura 1.2. En una fuente digital la información en cambio está representada por un conjunto de símbolos que pertenece a un alfabeto finito, que se envían secuencialmente en intervalos discretos de tiempo (se enviará un símbolo cada T segundos). Ejemplos de este tipo de fuentes serían ficheros de texto, donde el alfabeto está formado por los posibles caracteres que pueden aparecer en el texto, o los sistemas de datos binarios, donde la información se codifica como una secuencia de unos y ceros (alfabeto binario, $\{0, 1\}$ o bits).

1.2.2. Transmisor

La función del transmisor es convertir la señal de información o mensaje, independientemente de su formato, en una señal eléctrica o electromagnética adecuada para su transmisión a través del medio físico utilizado por el sistema para realizar la comunicación; habitualmente este medio se denominará genéricamente canal de comunicaciones. Para realizar de forma eficiente esta tarea se requiere un conocimiento relativo del canal: por ejemplo, es necesario saber cuanto atenúa el

medio la señal transmitida para amplificarla suficientemente o en qué rango de frecuencias permite el medio realizar la transmisión con la menor distorsión. El transmisor utilizará ese conocimiento relativo del canal para generar una señal adecuada a sus características, de forma que la señal sufra la menor distorsión posible durante su transmisión.

El proceso mediante el cual el transmisor adapta la señal a las características del canal se denomina genéricamente *modulación*. El proceso a realizar dependerá de múltiples factores, como el tipo de fuente de información, las características específicas del canal, las prestaciones requeridas, o los recursos disponibles, en especial la energía de que dispone el transmisor y el ancho de banda disponible en el canal; pero hay algunos aspectos generales que se pueden discutir en este momento. En cuanto al rango de frecuencias de la señal a transmitir, se pueden adoptar dos estrategias distintas de funcionamiento: transmisión en banda base o transmisión paso banda. En una transmisión en banda base se transmite la señal en la misma banda de frecuencias que ocupa de forma natural, que suele ser en bajas frecuencias. En una transmisión paso banda se modifica el rango de frecuencias que ocupa el espectro de la señal. La señal se procesa de tal forma que se traslada su respuesta en frecuencia a otra banda de frecuencias, centrada en torno a una cierta frecuencia central (ω_c rad/s o f_c Hz), que por distintas razones pueda ser más adecuada para la transmisión.

1.2.3. Canal

Se denomina de forma genérica como canal al medio físico que se utiliza para enviar la señal desde el transmisor al receptor. Existen distintos tipos de medios que pueden ser utilizados para transportar las señales en un sistema de comunicaciones. Los más comunes son los cables, de distintos tipos como los cables de pares utilizados en los sistemas tradicionales de telefonía, o los cables coaxiales empleados en los sistemas de televisión, las guías de onda, la fibra óptica y el espectro radioeléctrico, en el que el medio es la propia atmósfera y la energía electromagnética que transporta la información se introduce o se extrae de la misma mediante el uso de antenas. Un aspecto importante a tener en cuenta en este último medio de transmisión es que el medio es único y por tanto es compartido entre todos sus potenciales usuarios. Esto hace que en la práctica sea un recurso escaso y que el acceso al mismo esté regulado por las administraciones públicas; de otro modo las interferencias entre los distintos usuarios que trataran de hacer uso del mismo de forma no ordenada harían imposible una utilización eficiente del mismo.

Cada medio físico tiene sus propias características, pero independientemente del medio físico, un canal de comunicaciones introducirá siempre una serie de *perturbaciones* o *distorsiones* sobre las señales durante su transmisión. En el mejor de los casos, en lo que podría considerarse un escenario ideal, el canal introducirá en las señales transmitidas dos efectos: un retardo y una atenuación; ambos efectos son inherentes al propio proceso de transmisión de señales eléctricas o electromagnéticas a través de un medio:

- Las señales sufren una atenuación al propagarse por cualquier medio.
- Las señales tardan un cierto tiempo en recorrer una determinada distancia a través de cualquier medio.

Si sólo aparecen durante la transmisión estos dos efectos, la señal recibida $r(t)$ se podrá escribir en función de la señal transmitida $s(t)$ como

$$r(t) = C \cdot s(t - t_0),$$

donde la constante $C < 1$ define la atenuación y t_0 el retardo. Estos dos efectos son, como ya se ha dicho, inherentes a la transmisión de señales electromagnéticas y por tanto inevitables, pero en la mayor parte de los casos no son problemáticos. La atenuación se puede compensar con una amplificación con un factor de ganancia $G = 1/C$. En cuanto al retardo, si está dentro de unos límites razonables no suele tener implicaciones importantes, y dada la velocidad de transmisión de las señales electromagnéticas en la mayoría de las aplicaciones no supone un problema en la práctica. Por tanto, un medio que sólo incluya un retardo y una atenuación podría considerarse un sistema con una respuesta ideal.

Pero además de estos dos efectos, en la práctica aparecerán otro tipo de efectos no deseados, principalmente una distorsión lineal, una distorsión no lineal y ruido, en particular ruido térmico, que siempre estará presente en la transmisión de señales eléctricas o electromagnéticas por la inherente agitación térmica de los portadores de carga (electrones, fotones,...).

En esta asignatura, dado su carácter introductorio, se considerarán únicamente las distorsiones lineales y el ruido térmico, y no se tendrán en cuenta las distorsiones no lineales. En este caso, el modelo que se utilizará para la distorsión lineal es un modelo invariante en el tiempo, que vendrá caracterizado por una respuesta al impulso en el dominio temporal, $h(t)$, y la correspondiente representación de dicha respuesta en el dominio de la frecuencia, $H(j\omega)$. La relación entre la respuesta del sistema en los dominios temporal y frecuencial está dada por la transformada de Fourier

$$H(j\omega) = \mathcal{TF}\{h(t)\} = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} dt$$

$$h(t) = \mathcal{TF}^{-1}\{H(j\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) e^{+j\omega t} d\omega.$$

Considerando la distorsión lineal y el ruido, el modelo del canal de comunicaciones será el denominado modelo de canal lineal, que viene dado por la relación

$$r(t) = \int_{-\infty}^{\infty} s(\tau) \cdot h(t - \tau) d\tau + n(t) = s(t) * h(t) + n(t),$$

y que se muestra conceptualmente en el diagrama de bloques de la Figura 1.4: la salida del canal es el resultado de una distorsión lineal dada por $h(t)/H(j\omega)$ y la suma del término de ruido $n(t)$.

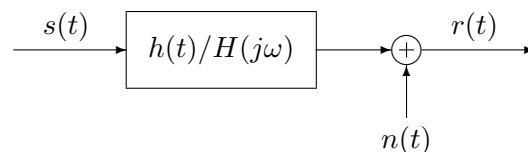


Figura 1.4: Modelo de canal lineal habitualmente utilizado para modelar el canal de comunicaciones.

Conviene recordar que el efecto del retardo y la atenuación se pueden incluir en la respuesta al impulso del sistema. En particular, un sistema ideal sin distorsión lineal de la señal y sin ruido, que sólo genera sobre la señal transmitida una atenuación dada por un factor $C < 1$ y un retardo t_0 , de forma que la señal recibida es

$$r(t) = C \cdot s(t - t_0),$$

se puede modelar a través de un sistema lineal con respuesta

$$h(t) = C \cdot \delta(t - t_0) \leftrightarrow H(j\omega) = C \cdot e^{-j\omega t_0}.$$

1.2.4. Receptor

El receptor debe recuperar la señal original de información (mensaje) a partir de la señal recibida. Como esta señal ha sufrido durante su transmisión una cierta distorsión (en el modelo de canal anterior, una distorsión lineal más el efecto del ruido), en general puede no ser posible recuperar de forma exacta la señal de información transmitida. El receptor debe diseñarse para, considerando las distorsiones que sufrirá la señal durante la transmisión, recuperar con la mayor fidelidad que sea posible la señal de información. La forma en que se mide esa fidelidad varía dependiendo del tipo de señal de información transmitida. Si la información transmitida tiene formato analógico, el objetivo será que la forma de onda de la señal recibida se parezca lo máximo posible a la de la transmitida, y la forma habitual de cuantificar este parecido es mediante la denominada relación señal a ruido (S/N, o SNR de *Signal to Noise Ratio*). Esta relación mide el cociente entre la energía o potencia de la diferencia entre señal transmitida y recibida (que se considera ruido) y la energía o potencia de la señal transmitida. Si el formato de la información transmitida es digital, el objetivo será recibir el menor número de símbolos que representan la información de forma errónea, de forma que se cuantifica la fidelidad con una probabilidad de error de símbolos (o de bits en sistemas binarios, BER de *Bit Error Rate*). En cualquier caso, el receptor en general deberá realizar las siguientes tareas:

- Demodular la señal, lo que significa deshacer todas las transformaciones que se hicieron en el transmisor para acondicionar la señal para su transmisión a través del canal de comunicaciones, como por ejemplo volver la señal a su banda de frecuencias original si se ha realizado una transmisión en paso banda.
- Minimizar el efecto del ruido sobre la señal de información.
- Compensar, si es posible, las distorsiones lineales que haya introducido el canal.

Al igual que en el transmisor, de nuevo el modo en que se realizarán las tres tareas dependerá de múltiples factores, como del formato de la información transmitida o del ancho de banda utilizado. A lo largo del curso se verá cómo se materializan estas funciones para los distintos tipos de modulaciones.

1.3. Sistemas de comunicaciones analógicos y digitales

De forma similar a cómo se clasificaron las fuentes de información en cuanto al formato de la información que generan, dando lugar a la distinción entre fuentes analógicas y fuentes digitales, los sistemas de comunicaciones se pueden también clasificar en cuanto al formato en que transportan la información en dos grandes tipos: sistemas analógicos y sistemas digitales. Un sistema de comunicaciones analógico envía la información impresa en una cierta forma de onda continua, mientras que un sistema digital enviará la información impresa en una secuencia de símbolos que se envían secuencialmente a una cierta tasa (R_s símbolos por segundo). El caso más habitual son los sistemas binarios, en los que la información está contenida en una secuencia de bits (unos y ceros) que se envían a una cierta tasa de bit (R_b bits/s).

Es importante destacar que un sistema de comunicaciones analógico/digital no está restringido a la transmisión de una fuente del mismo tipo (analógico/digital). En particular, una señal analógica puede digitalizarse, convirtiéndola en una secuencia de bits a una cierta tasa, y posteriormente volver a convertirse en analógica (procesos de conversión analógico/digital, o A/D

y digital/análogo, o D/A). La conversión analógico/digital de una señal consiste en dos pasos. Una señal analógica es continua tanto en el tiempo como en el rango de posibles amplitudes (hay continuidad en los dos ejes de la representación temporal de la señal, como se ve en la Figura 1.5). Para convertirla a formato digital la señal se discretizará tanto en tiempo como en amplitud, como se ilustra también en la figura:

- De tiempo continuo se pasa a muestras de la señal en tiempo discreto mediante un muestreo periódico

$$s[n] = s(t)|_{t=nT_m} = s(nT_m),$$

donde T_m es el intervalo de muestreo. Si la señal $s(t)$ es limitada en banda, con un ancho de banda de B Hz, el conocido teorema de Nyquist para el muestreo demuestra que no hay pérdida de información en el proceso de muestreo; esto significa que es posible recuperar la señal original a partir de sus muestras mediante una interpolación con funciones sincs (lo que es equivalente a interpolar con un filtro paso bajo del ancho de banda de la señal). Para ello, sólo es necesario garantizar que el intervalo de muestreo T_m sea lo suficientemente pequeño para garantizar que su inversa, la frecuencia de muestreo, sea al menos el doble del ancho de banda de la señal expresado en Hz

$$\frac{1}{T_m} = f_m \text{ muestras/s} \geq B \text{ Hz.}$$

- Después del muestreo, la digitalización finalizará cuantificando cada una de las muestras resultantes con n bits. En la figura se representa un ejemplo de cuantificación con 3 bits, en el que el rango dinámico de la señal se divide en 8 posibles valores (ya que 3 bits permiten representar $2^3 = 8$ valores). En este paso sí se produce una distorsión de la señal de información, ya que de los valores originales de la señal en cada muestra (círculos sin relleno en la figura) se pasa al valor cuantificado más cercano (círculos sólidos en la figura). A este efecto se le denomina ruido de cuantificación.

El proceso de conversión de una señal analógica a digital produce por tanto una secuencia de bits. La tasa de dicha secuencia vendrá dada por el producto entre la tasa de muestreo y el número de bits por muestra

$$R_b \text{ bits/s} = f_m \text{ (muestras/s)} \times n \text{ (bits/muestra).}$$

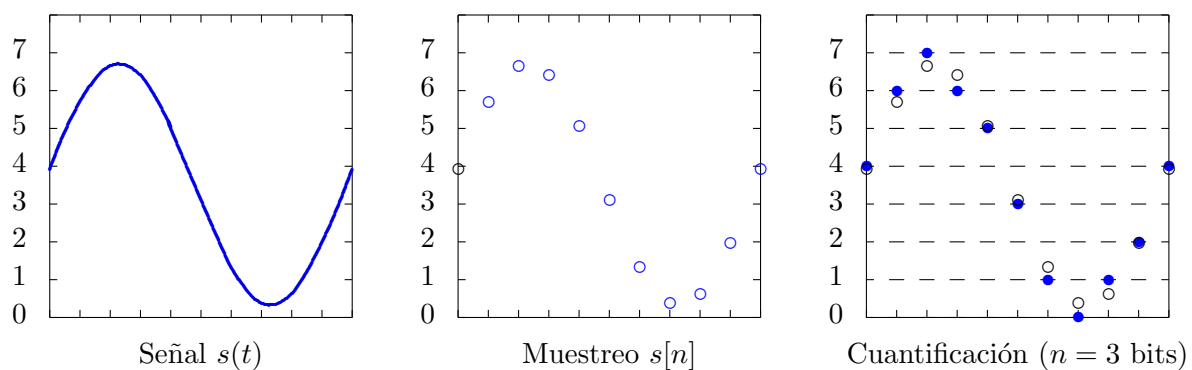


Figura 1.5: Conversión de una señal analógica a formato digital (secuencia de bits).

Aunque la conversión conlleva siempre un cierto grado de distorsión de la señal de información, debido a ese ruido de cuantificación, en muchos casos esa distorsión puede llegar a ser inapreciable.

Para ello, sólo hay que utilizar un número suficiente de bits por muestra. Si se incrementa el número de bits, el número de niveles de cuantificación será 2^n , con lo que la diferencia entre niveles dentro del rango de la señal disminuirá progresivamente a medida que se aumenta n , llegando un momento es que la distorsión debida a la cuantificación será prácticamente inapreciable al reconstruir la señal a tiempo continuo interpolando con sincs (filtrado paso bajo). El coste asociado a la disminución de la distorsión que se consigue al aumentar el número de bits por muestra es el aumento de la tasa binaria de la secuencia de bits resultante de la conversión A/D.

Actualmente la tendencia es que las fuentes de naturaleza analógica no se transmitan con un sistema de comunicaciones analógico, sino que se conviertan a formato digital, se transmitan con un sistema de comunicaciones digital y finalmente se conviertan de nuevo a formato analógico. Esto es así porque aunque cada tipo de sistema tiene sus ventajas y sus inconvenientes frente al otro tipo, en general las ventajas de los sistemas digitales tienen mayor peso que sus desventajas para la mayoría de las aplicaciones. Esto hace que en la actualidad los sistemas de comunicaciones digitales predominen claramente sobre los sistemas de comunicaciones analógicos. Cuáles son esas ventajas de los sistemas digitales sobre los sistemas analógicos que los han hecho claramente preponderantes será algo que se discutirá con detalle en el Capítulo 4, dedicado al estudio de los sistemas de comunicaciones digitales.

Aunque también es posible transmitir información digital utilizando sistemas de comunicaciones analógicos, esta opción es mucho menos frecuente, y en la práctica está reducida a unos pocos sistemas con necesidades muy específicas.

1.4. Diseño de un sistema de comunicaciones

Al diseñar un sistema de comunicaciones hay que considerar distintos factores. Algunos de los más importantes son los siguientes:

- Calidad o prestaciones requeridas.
- Uso de recursos.
- Coste económico.
- Tecnología disponible para la implementación.

A continuación se analizarán brevemente cada uno de estos factores.

1.4.1. Calidad requerida

Entre las especificaciones de un sistema de comunicaciones una de las más importantes es la calidad requerida para el mismo. El diseño del sistema tendrá que estar encaminado a lograr la calidad especificada con las restricciones impuestas por los recursos disponibles y con el menor coste posible, sin superar en ningún caso el coste máximo asumible para el sistema. Naturalmente, existe siempre un compromiso entre estos factores; cuanto mayores sean los recursos disponibles, mayor será la calidad que se podrá obtener.

La forma de especificar la calidad de un sistema de comunicaciones es diferente en función del tipo de sistema. Para sistemas de comunicaciones analógicos, la calidad está ligada a la *fidelidad* de

la señal recibida: la señal recibida debe ser lo más parecida que sea posible a la señal transmitida originalmente. Como se ha discutido con anterioridad, en el proceso de transmisión se producirán distorsión e interferencias, fundamentalmente ruido térmico. Esto supone que la señal recibida será distinta de la señal original. Para medir lo fiel que es una señal se suele utilizar como figura de mérito la relación señal a ruido (S/N), que nos da la relación entre la potencia de la señal y del ruido. En algunas ocasiones, al utilizar esta medida se considerará como ruido el efecto conjunto de todas las distorsiones que se producen sobre la señal transmitida, es decir, a la diferencia entre la señal transmitida y la señal recibida. En realidad se incluye el efecto del ruido y de las demás distorsiones lineales o no lineales que se producen. Por esta razón, en estas ocasiones también se le denomina relación señal a distorsión. Como es lógico, cuanto mayor sea el valor de esta relación señal a ruido (o a distorsión) mayor es la calidad del sistema.

En sistemas de comunicaciones digitales, en cambio, dado que la información está contenida en una secuencia de símbolos de un alfabeto finito (habitualmente binario, unos y ceros), la figura de mérito que se utiliza normalmente para cuantificar la calidad del sistema será la *tasa (o probabilidad) de error* de símbolo en general, o en el caso particular binario de bit (en este caso se conoce comúnmente con el acrónimo inglés BER, de *Bit Error Rate* o *Bit Error Ratio*). Naturalmente, cuanto menor sea esta tasa o probabilidad de error mayor será la calidad del sistema digital.

1.4.2. Recursos disponibles

Los recursos disponibles para su uso por el sistema de comunicaciones serán determinantes a la hora de limitar las máximas prestaciones alcanzables. Entre los recursos a considerar, los más importantes habitualmente son el ancho de banda y la energía.

El ancho de banda va a estar limitado en la práctica en la mayor parte de las aplicaciones. Bien por limitaciones físicas, dado el ancho de banda limitado que tienen algunos de los medios de transmisión habituales, como los cables, o bien por limitaciones administrativas, fundamentalmente cuando se transmite utilizando el espectro radioeléctrico, ya que el uso de este medio compartido está completamente regulado.

El ancho de banda está ligado a la calidad de la señal. En sistemas analógicos, aplicaciones con más precisión requieren mayor ancho de banda. Esto es así por dos razones principalmente. En algunos casos el ancho de banda real de las señales analógicas puede ser mayor que el ancho de banda disponible para la transmisión, por lo que hay que filtrar las señales para reducir su ancho de banda antes de la transmisión. Si se produce esta situación, cuantas más componentes frecuenciales se eliminan (a medida que disminuya el ancho de banda), mayor será la distorsión que sufre la señal de información. En otros casos, una de las formas de reducir el efecto del ruido es ensanchar el espectro de la señal de modo que haya cierta redundancia en la señal transmitida. Cuanto más se aumente el ancho de banda de la señal, mayor será la protección frente al ruido que se obtenga durante la transmisión. Esto se verá con más detalle al analizar las prestaciones de las denominadas modulaciones angulares, o modulaciones de fase y frecuencia. En sistemas digitales, el ancho de banda está relacionado con la máxima tasa de transmisión que se puede obtener con unas mínimas prestaciones, de forma que a mayor ancho de banda mayor capacidad de transmisión (transmisión a una mayor velocidad).

1.4.3. Coste del sistema y tecnologías existentes

El coste del sistema es también otro factor a considerar en el diseño del mismo, ya que habitualmente habrá un coste objetivo o un coste máximo asumible que no podrá superarse. Este coste está relacionado con las tecnologías utilizadas en la implementación de transmisores y receptores, y en la elección del medio de transmisión a utilizar. Por tanto, a la hora de decidir sobre algunas opciones de diseño el conocimiento de las posibles tecnologías a utilizar y su correspondiente coste será un aspecto importante.

Pese a esta importancia, en esta asignatura no se tratará este aspecto, que queda fuera de los objetivos de la asignatura, que como veremos en la siguiente sección tiene una orientación más teórica que práctica y se centrará en los aspectos teóricos básicos que rigen el funcionamiento de un sistema de comunicaciones.

1.5. Objetivos y organización de la asignatura

Una vez hecha una breve introducción a los sistemas de comunicaciones, en esta última sección del capítulo se van a presentar los objetivos fundamentales de la asignatura y la organización de los contenidos de la misma para lograr dichos objetivos.

1.5.1. Objetivos de la asignatura

Esta asignatura trata de establecer los principios teóricos fundamentales que se aplican en el diseño y análisis de sistemas de comunicaciones, tanto analógicos como digitales. Para ello, será fundamental la caracterización matemática de un sistema de comunicaciones y de las señales presentes en el mismo, tanto señales transmitidas como señales interferentes como el ruido. Esta caracterización será la que permita realizar el análisis de un sistema de comunicaciones, y deducir los principios teóricos que definen el diseño óptimo de cada uno de los distintos elementos funcionales que lo forman. En particular se pueden establecer los siguientes objetivos fundamentales:

- Introducir la caracterización estadística de las señales relacionadas con sistemas de comunicaciones, de las señales de información y especialmente del ruido térmico que siempre aparece en la transmisión de una señal electromagnética.
- Presentar el concepto de modulación en sistemas de comunicaciones analógicos, y estudiar los tipos de modulación más comunes: las modulaciones de amplitud y las modulaciones angulares, de fase y frecuencia.
- Formar la base del núcleo de conocimientos sobre comunicaciones digitales, presentando de forma simplificada el concepto de modulación digital, la transmisión sobre canales gaussianos, donde el principal elemento de distorsión considerado será el ruido térmico, y estudiando los principios teóricos básicos para diseñar un demodulador digital, aplicando los principios estadísticos de la teoría de la decisión y la representación vectorial de las señales. Finalmente, se utilizarán principios de la teoría de la información para obtener algunos de los límites fundamentales que se pueden alcanzar en un sistema digital de comunicaciones.

1.5.2. Organización de la asignatura

Para cumplir los objetivos presentados en el apartado anterior, la asignatura se ha organizado en los siguientes cinco capítulos:

1. Introducción

En este primer capítulo se ha hecho una breve introducción a los sistemas de comunicaciones para presentar los objetivos y la organización de la asignatura.

2. Ruido en los sistemas de comunicaciones

En este capítulo se presenta la caracterización estadística de las señales en un sistema de comunicaciones, y en particular de la señal de ruido térmico que siempre aparece presente como elemento de distorsión en la transmisión de señales eléctricas o electromagnéticas. Para realizar esta caracterización, se repasan algunos conceptos básicos de la teoría de variable aleatoria y de procesos aleatorios, para presentar el modelo estadístico comúnmente empleado para modelar el ruido térmico. Este modelo se aplicará en el cálculo de la relación señal a ruido en la transmisión de una señal de información.

3. Modulaciones analógicas

En este capítulo se presentan distintas técnicas de modulación utilizadas en sistemas analógicos de comunicaciones. En particular se presentan las variantes más conocidas de modulaciones de amplitud, y las modulaciones angulares de fase y frecuencia. Para cada tipo de modulación se presenta su descripción en el dominio del tiempo, sus características espectrales y sus requerimientos de potencia. Finalmente, se analiza el comportamiento de cada modulación frente al ruido, evaluando la relación señal a ruido obtenida con cada una de ellas.

4. Modulación y detección en canales gaussianos

Este capítulo presenta los principios básicos que rigen el diseño y análisis de un sistema digital de comunicaciones. Se introduce el concepto de modulación digital, como mecanismo para la transmisión de información digital a través de canales analógicos, y el concepto de detección como mecanismo para recuperar la información digital transmitida a partir de la señal analógica recibida a través del canal de comunicaciones. Como modelo para este canal se utilizará el modelo más sencillo: un canal aditivo gaussiano.

5. Límites fundamentales en los sistemas de comunicaciones digitales

Finalmente, el último capítulo presenta algunos de los límites fundamentales que pueden alcanzarse con un sistema de comunicaciones digitales. En particular, se estudia cómo calcular la máxima cantidad de información que se puede transmitir de forma fiable mediante un sistema digital de comunicaciones, lo que se conoce como *capacidad de canal*. La obtención de este límite se basa en la aplicación de la teoría de la información y del uso de medidas cuantitativas de información que se presentan en el capítulo como herramientas para el análisis de sistemas de comunicaciones digitales.

1.5.3. Bibliografía recomendada

Existen excelentes libros que tratan sobre los sistemas de comunicaciones. Para esta asignatura se recomiendan sólo un número reducido de ellos, que tratan de forma adecuada todos los contenidos de la asignatura, con la finalidad de evitar al estudiante una excesiva cantidad de referencias.

También con este objetivo, se ha distinguido entre lo que se ha denominado bibliografía básica, donde aparecen dos referencias que tratan todos los contenidos de la asignatura, y lo que se ha denominado bibliografía complementaria, donde se citan textos que permiten profundizar en algunos de los contenidos de la asignatura más allá de los propios objetivos de esta. A continuación se presentan estas referencias bibliográficas y se incluye un breve comentario sobre cada una de ellas.

Bibliografía básica

1. A. Artés et al. “Comunicaciones Digitales”, Pearson Educación, 2007

La primera referencia básica es este excelente libro, realizado por varios profesores universitarios, orientado hacia su uso como manual de aprendizaje, por lo que resulta muy apropiado como referencia para la asignatura. En este libro se tratan fundamentalmente los sistemas de comunicaciones digitales con un claro enfoque hacia los contenidos habitualmente tratados en las titulaciones relacionadas con la Ingeniería de Telecomunicación. El capítulo 3 cubre la mayor parte de los contenidos del capítulo 2 de la asignatura. El capítulo 4 y el capítulo 9 cubren, respectivamente, todos los contenidos de los capítulos 4 y 5 de la asignatura.

Este libro está disponible on-line a través de la página web de su primer autor, el Catedrático Antonio Artés Rodríguez, de la Universidad Carlos III de Madrid

- Disponible on-line: www.tsc.uc3m.es/~antonio/

2. J. G. Proakis, M. Salehi. “Communication Systems Engineering” (2ª Ed.), Prentice-Hall, 1994

Se trata de otro excelente texto sobre sistemas de comunicaciones, tanto digitales como analógicos. Su notación compacta y su modularidad facilitan la tarea de seguimiento de los contenidos de la asignatura pese a su diferente secuenciamiento con respecto al seguido en la asignatura. Los capítulos 4 y 5 cubren los contenidos del capítulo 2 de la asignatura; los capítulos 3 y 5 cubren los contenidos del capítulo 3 de la asignatura; y finalmente los capítulos 6 y 9 lo hacen con los del capítulo 5. Aunque el capítulo 7 cubre bastantes de los contenidos del capítulo 4 de la asignatura, en este caso la aproximación es ligeramente distinta a la seguida en la asignatura.

Bibliografía complementaria

1. A. Papoulis. “Probability, random variables, and stochastic processes”, (3ª Ed.), McGraw-Hill, 1991

Uno de los libros de referencia de las bases de la teoría de la probabilidad y los procesos estocásticos. Una excelente referencia para todos los conceptos estadísticos que se manejan en la asignatura.

2. A.B. Carlson. “Communication Systems” (2ª Ed.), McGraw-Hill, 1986

Texto clásico de introducción a la transmisión analógica y digital. Básicamente, consta de tres partes: la primera introduce los fundamentos básicos de las señales y procesos aleatorios; la segunda abarca las comunicaciones analógicas, mientras que la tercera parte está dedicada a las comunicaciones digitales. El desarrollo que realiza de las comunicaciones analógicas es sencillo e intuitivo y está ilustrado con esquemas de bloques de sistemas y circuitos eléctricos básicos.

3. S. Haykin. “An Introduction to Analog and Digital Communications”, Willey, 1989
Otro texto clásico que trata los sistemas de comunicaciones analógicos y digitales, aunque en este caso con un claro énfasis en los últimos. Es un libro interesante para el tratamiento introductorio de la teoría de la comunicación y de la naturaleza matemática de su formulación.
4. B. Sklar. “Digital communications : fundamentals and applications”, Prentice Hall, 2001
Libro avanzado sobre comunicaciones digitales, interesante para aquellos estudiantes con una formación básica en teoría de la probabilidad. Presenta una excelente introducción a los fundamentos sobre señales, espectro de las mismas, y transmisión en banda base. A partir de esta introducción, el libro presenta múltiples variantes de modulaciones y técnicas de modulación que, aunque van más allá de los objetivos de esta asignatura, pueden resultar de interés a un estudiante que la haya cursado.

Capítulo 2

Ruido en los sistemas de comunicaciones

En un sistema de comunicaciones hay distintos tipos de señales. Algunas señales son deterministas. Otras, sin embargo, tienen naturaleza aleatoria, ya que no se conoce a priori su valor concreto en un cierto instante de tiempo, sino que se conocerán únicamente sus parámetros estadísticos. A este tipo de señales pertenecen por ejemplo

- Las señales de información que se desean transmitir. Toda señal de información tiene un cierto grado de incertidumbre. Si no fuera así, no contendría realmente información (por ejemplo, si el receptor conociera de forma precisa la señal que se va a transmitir, no sería necesario transmitirla).
- El ruido térmico que siempre aparece en la transmisión de señales electromagnéticas.

Para caracterizar señales de este tipo, de las que se conocen algunas de sus características estadísticas, pero no sus valores concretos en un momento dado, se utilizan los procesos aleatorios.

Por esta razón, en este capítulo se va a utilizar la teoría de los procesos aleatorios para caracterizar señales de comunicaciones y, sobre todo, el ruido térmico, ya que este modelo será fundamental en el diseño y análisis de sistemas de comunicaciones al ser el ruido térmico una de las principales fuentes de distorsión. Se analizará también cómo afectan los sistemas lineales a los parámetros estadísticos que definen el proceso, y se verá cómo calcular la relación señal a ruido bajo diversas circunstancias en un sistema de comunicaciones.

Antes de llegar a la utilización de los procesos aleatorios como herramienta para la caracterización de señales en un sistema de comunicaciones, se realizará un breve repaso de algunos conceptos relacionados con la probabilidad, las variables aleatorias y los procesos aleatorios. Aunque estos conceptos formarían parte de la asignatura previa de *Estadística*, se incluyen aquí por completitud.

2.1. Probabilidad

En este apartado se repasarán, de forma breve, algunos de los conceptos básicos de la teoría de la probabilidad. Nos centraremos en los aspectos que son necesarios para el tratamiento de procesos aleatorios en el ámbito del modelado de señales de comunicaciones.

La teoría de la probabilidad trabaja con fenómenos que se producen de forma masiva. Hay un sin

número de ejemplos: juegos de azar, movimiento de electrones, tasas de nacimiento y muerte, etc. Y lo que la teoría de la probabilidad trata de hacer es establecer promedios para esos fenómenos. En particular, su propósito es describir y predecir estos promedios en términos de probabilidades de sucesos o eventos.

2.1.1. Espacio de probabilidad

Antes de poder definir lo que es un espacio de probabilidad, es necesario hacer varias definiciones.

Experimento aleatorio

El concepto fundamental en el que se basa la teoría de la probabilidad es el *experimento aleatorio*. Un experimento aleatorio es un experimento cuya salida no puede ser predicha con exactitud. Tirar una moneda, tirar un dado, sacar una carta de la baraja, o medir el valor de voltaje en un par de cobre, son algunos ejemplos de experimento aleatorio.

Espacio muestral (Espacio de muestras)

Todo experimento aleatorio tiene ciertos valores de salida, o posibles resultados del experimento. En el caso del lanzamiento de una moneda, que la figura que queda hacia arriba se cara o cruz, en el caso del dado, que el número de puntos de la cara que queda hacia arriba sea 1, 2, 3, 4, 5 o 6. Se define el *espacio muestral* como el conjunto de todas las posibles salidas de un experimento. Normalmente se denota con la letra griega omega Ω .

En cuanto a su naturaleza, existen dos tipos de espacios muestrales:

- Discretos, cuando el experimento tiene como posibles valores un número finito de posibles valores, o un número infinito numerable.
- No discretos (o continuos), cuando el espacio muestral corresponde a conjuntos continuos de posibles valores (o dicho de otro modo, el número de posibles valores de salida es infinito incontable).

Ejemplos de los primeros son el dado o la moneda antes mencionados. En ese caso el espacio de muestras es para la moneda cara y cruz, en el caso del dado, 1, 2, 3, 4, 5 y 6. Un ejemplo de variable aleatoria con un espacio muestral continuo es el valor del voltaje en una resistencia, que puede tomar cualquier valor dentro de un rango de valores de voltaje. En este caso el espacio de muestras es todo ese conjunto continuo de posibles valores.

Existen también espacios mixtos, con parte del espacio muestral discreto, y parte continuo, aunque no se comentarán en este capítulo.

Sucesos (Eventos)

Un suceso, o un evento, es un subconjunto del espacio de muestras sobre el que es posible definir una probabilidad. Para que esta medida de probabilidad tenga sentido, tiene que imponer una

serie de condiciones, que se verán algo más tarde. Vamos primero a ver qué es lo que se entiende por una probabilidad.

La probabilidad de un suceso E es un número, $P(E)$, no negativo, definido entre 0 y 1 ($0 \leq P(E) \leq 1$), asignado a ese evento y que describe lo probable o improbable que es dicho suceso. Este número se puede interpretar de la forma siguiente:

Si un determinado experimento se realiza un número N de veces (suponiendo que N es suficientemente grande) y el evento A ocurre N_A veces, entonces podemos decir que la probabilidad será bastante cercana a la relación N_A/N :

$$P(A) \approx \frac{N_A}{N}$$

Esta puede ser una definición intuitiva de probabilidad, es decir, que es una medida que nos indica lo frecuentemente que se produce un suceso cuando se realiza un cierto experimento.

Para el caso de espacios discretos, la idea es simple. ¿Cuál es la probabilidad de sacar un 5 al tirar un dado? Si el dado no está trucado, esta probabilidad es $1/6$. Pero para el caso de espacios continuos hay un matiz importante a tener en cuenta. Por ejemplo, ¿cuál es la probabilidad de que el voltaje en una resistencia valga 1 V? La respuesta es 0. Aunque esto puede parecer anti-intuitivo, la explicación está en que el conjunto de valores que puede tomar es infinito, así que la probabilidad de tener uno de ellos es nula. En resumen, no es posible definir una probabilidad para un valor concreto. Lo que sí es posible es definir la probabilidad de que el valor de tensión esté en un cierto intervalo, por ejemplo entre 0.99 y 1.01 voltios. Ese suceso sí tiene una probabilidad asociada.

Así pues, los sucesos en experimentos con espacios muestrales discretos han de estar formados por un subconjunto del espacio muestral, incluidos sucesos de un único elemento. Y en el caso de espacios muestrales continuos, cada suceso ha de tener una probabilidad, así que hay que coger “regiones” del espacio muestral (no un único valor). Normalmente se define el campo sigma, denotado por \mathcal{B} , como la colección de los subconjuntos de Ω , es decir, como el conjunto de todos los posibles sucesos.

Algunas definiciones sobre sucesos, que pueden resultar de utilidad, son las siguientes:

- *Suceso trivial*: es el que ocurre en todo experimento, es decir, que su probabilidad es 1. Ejemplo, Ω .
- *Conjunto nulo* (\emptyset): El que no tiene ningún elemento.
- *Unión de sucesos* ($E_1 \cup E_2$): es el suceso que ocurre cuando sucede E_1 , E_2 o ambos.
- *Intersección de sucesos* ($E_1 \cap E_2$): el evento que ocurre cuando los eventos E_1 y E_2 se producen al mismo tiempo.
- *Complemento de un suceso* (E^c): es el espacio muestral menos el propio suceso.
- *Eventos exclusivos o disjuntos*: aquellos para los que $E_1 \cap E_2 = \emptyset$. Para ellos se cumple que $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

Espacio de probabilidad

El espacio de probabilidad se define como el triplete (Ω, \mathcal{B}, P) ; es decir, el espacio muestral, el espacio con los distintos sucesos y la medida de probabilidad que nos indica la probabilidad de cada suceso. Algunas de las propiedades que tiene que cumplir la medida de probabilidad sobre sucesos son las siguientes:

1. $P(E^c) = 1 - P(E)$.
2. $P(\emptyset) = 0$.
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.
4. Si $E_1 \subset E_2$ entonces $P(E_1) \leq P(E_2)$.

2.1.2. Probabilidad condicional

Supongamos que existen dos sucesos, E_1 y E_2 , definidos sobre el mismo espacio de probabilidad con sus correspondientes probabilidades $P(E_1)$ y $P(E_2)$. A estas probabilidades a veces se las conoce como probabilidades *a priori* de cada suceso. Si se sabe que uno de los eventos se ha producido, por ejemplo E_2 , esto nos puede proporcionar cierta información sobre el otro suceso, que cambia su probabilidad a priori (sin conocer que se haya producido ningún evento). A esta nueva probabilidad se le denomina probabilidad condicional, o condicionada. La probabilidad condicional del suceso E_1 dado el suceso E_2 , denotada como $P(E_1|E_2)$ se define como:

$$P(E_1|E_2) = \begin{cases} \frac{P(E_1 \cap E_2)}{P(E_2)}, & P(E_2) \neq 0 \\ 0, & P(E_2) = 0 \end{cases}.$$

Ejemplo

Se lanza un dado no cargado, y se definen los siguientes sucesos

- E_1 : resultado mayor que 3
- E_2 : resultado par

Las probabilidades de cada uno de estos sucesos se calcula de forma muy sencilla sumando las probabilidades de cada uno de los posibles valores de salida del espacio muestral que forman parte del suceso

$$P(E_1) = P(4) + P(5) + P(6) = \frac{1}{2}$$

$$P(E_2) = P(2) + P(4) + P(6) = \frac{1}{2}$$

$$P(E_1 \cap E_2) = P(4) + P(6) = \frac{1}{3}$$

La probabilidad de $E_1|E_2$ es

$$P(E_1|E_2) = \frac{1/3}{1/2} = \frac{2}{3}$$

Se comprueba que el resultado obtenido coincide con la probabilidad de tener un 4 o un 6 cuando el espacio muestral es el suceso E_2 . El hecho de conocer que el resultado es par, altera las probabilidades de tener un resultado mayor que 3 cuando no hay ninguna información previa.

Sucesos estadísticamente independientes

De la probabilidad condicional se deriva una importante definición estadística. Si ocurre que $P(E_1|E_2) = P(E_1)$ esto significa que el conocimiento de E_2 no aporta información sobre E_1 y por tanto no cambia su probabilidad con respecto a la probabilidad a priori (sin el conocimiento de que se ha producido E_2). En este caso, se dice que los dos sucesos son *estadísticamente independientes*.

Formalmente, se dice que dos sucesos son *estadísticamente independientes* cuando las probabilidades condicionales coinciden con las probabilidades a priori

$$P(E_1|E_2) = P(E_1), \quad P(E_2|E_1) = P(E_2).$$

Dada la relación entre probabilidades a priori, y probabilidades condicionales a través de la probabilidad de la intersección, la probabilidad de la intersección de dos sucesos estadísticamente independientes es igual al producto de las probabilidades de cada suceso

$$P(E_1 \cap E_2) = P(E_1) \times P(E_2).$$

Teorema de la probabilidad total

Si los sucesos E_i , con $i = 1, \dots, N$ forman una partición del espacio muestral Ω , lo que quiere decir que se cumplen las siguientes condiciones

- $\cup_{i=1}^N E_i = \Omega$
- $E_i \cap E_j = \emptyset$ para todo $i \neq j$

es decir, que la unión de los sucesos forma todo el espacio muestral, siendo los sucesos disjuntos entre sí, entonces, si para un suceso A se dispone de las probabilidades condicionales $P(A|E_i)$ para todos los sucesos que forman la partición, $i = 1, \dots, N$, la probabilidad $P(A)$ se obtiene mediante el *teorema de la probabilidad total*

$$P(A) = \sum_{i=1}^N P(A|E_i)P(E_i).$$

Regla de Bayes

Por otro lado, la *Regla de Bayes* (aunque su idea se debe a Bayes, finalmente la formuló Laplace) nos dice que las probabilidades condicionales de los sucesos de la partición dado A , $P(E_i|A)$, se obtienen mediante la siguiente expresión

$$P(E_i|A) = \frac{P(A|E_i)P(E_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{\sum_{j=1}^N P(A|E_j)P(E_j)}.$$

2.2. Variable aleatoria

Una *variable aleatoria* (*v.a.*) no es más que una función que asigna un número a cada una de las posibles salidas de un experimento aleatorio, es decir, a cada uno de los elementos del espacio muestral. En esta sección el estudio se centrará en variables aleatorias reales, para las que el número asignado a cada posible salida del experimento aleatorio pertenece al conjunto de números reales.

$$\begin{aligned} \Omega &\rightarrow \mathbb{R} \\ \lambda \in \Omega &\rightarrow X(\lambda) \in \mathbb{R} \end{aligned}$$

Por tanto, una v.a. (real) mapea los resultados de un experimento aleatorio en la recta real.

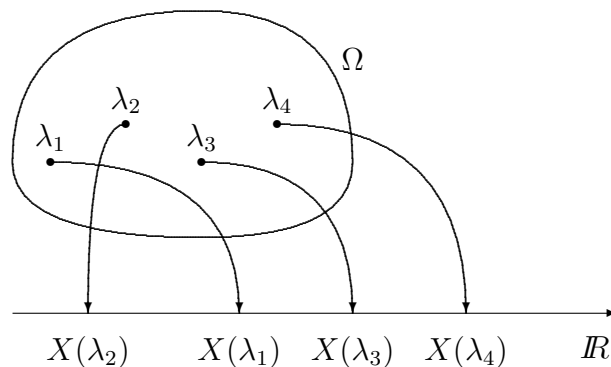


Figura 2.1: Variable aleatoria como un mapeo de Ω a \mathbb{R} .

Por ejemplo, en el experimento lanzar un dado aparece ya una asignación de forma implícita (el número de puntos que hay en la cara que queda hacia arriba). En otros casos, como en el caso de lanzar una moneda, es posible asignar un número a la cara y uno a la cruz (por ejemplo cara $\equiv 0$, cruz $\equiv 1$). Las variables aleatorias normalmente se denotan con mayúscula X , Y , y no se suele expresar la dependencia implícita con los elementos del espacio muestral del experimento aleatorio, λ_i . De nuevo, al clasificar en cuanto al tipo de valores que puede tomar, vamos a tener principalmente dos categorías de variable aleatoria:

- Discreta: número finito de valores.
- Continua: rango continuo de valores (uno o varios intervalos).

En cuanto a los valores en los que se traduce la salida del experimento aleatorio, se denomina *rango* (o *recorrido*) de una v.a. al conjunto de números reales que tienen asociado un resultado del espacio muestral, es decir:

$$\mathcal{A}_X = \{x \in \mathbb{R} : \exists \lambda \in \Omega \text{ tal que } X(\lambda) = x\}.$$

En el caso de variables aleatorias discretas, también se le denomina en ocasiones *alfabeto* de la variable aleatoria.

Probabilísticamente, una variable aleatoria se caracteriza habitualmente mediante dos funciones (que están ligadas entre sí):

- Función de distribución, $F_X(x)$.
- Función de densidad de probabilidad, $f_X(x)$.

A continuación se describe cada una de estas funciones.

2.2.1. Función de distribución

La *función de distribución (FD)* de una variable aleatoria se define como

$$F_X(x) = P(X \leq x),$$

es decir, como la probabilidad de que la variable aleatoria X tome un valor menor o igual que el argumento x . Las principales propiedades de la función de distribución son las siguientes:

1. $0 \leq F_X(x) \leq 1$.
2. $x_1 < x_2 \rightarrow F_X(x_1) \leq F_X(x_2)$ ($F_X(x)$ es no decreciente).
3. $F_X(-\infty) = 0$ y $F_X(\infty) = 1$ ($\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow \infty} F_X(x) = 1$).
4. $F_X(x^+) = F_X(x)$ ($F_X(x)$ es continua por la derecha).
5. $F_X(b) - F_X(a) = P(a < X \leq b)$.

Para calcular otras probabilidades incluyendo o no los límites del intervalo

$$\begin{aligned} P(a \leq X \leq b) &= F_X(b) - F_X(a^-). \\ P(a < X < b) &= F_X(b^-) - F_X(a). \\ P(a \leq X < b) &= F_X(b^-) - F_X(a^-). \end{aligned}$$

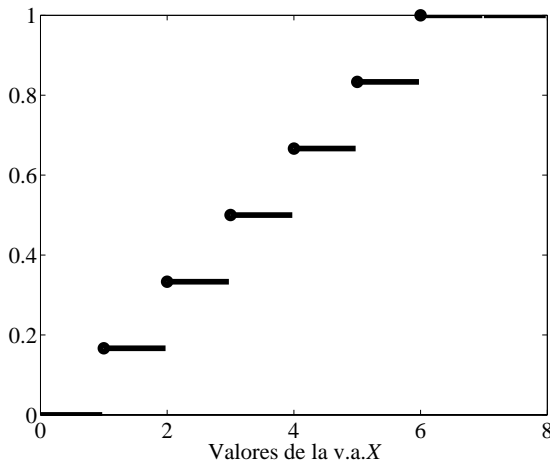
6. $P(X = a) = F_X(a) - F_X(a^-)$.
7. $P(X > x) = 1 - F_X(x)$.

En las expresiones anteriores, se ha utilizado como notación

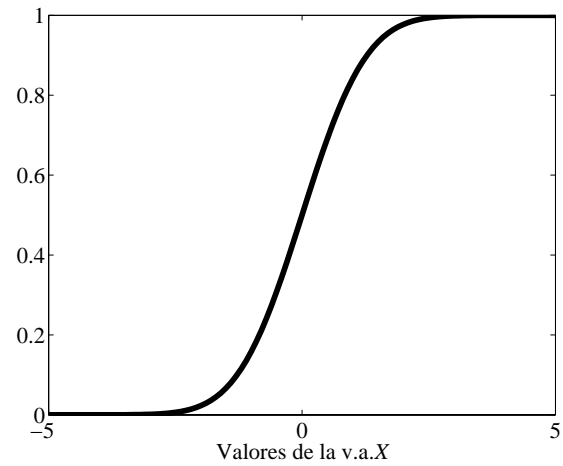
$$F_X(x^\pm) = \lim_{\varepsilon \rightarrow 0} F_X(x \pm \varepsilon).$$

Esta distinción $F_X(x^\pm)$ se realiza para tener en cuenta el caso particular de funciones de distribución para v.a. discretas, para las que $F_X(x_i^-) \neq F_X(x_i)$, siendo $\{x_i\}_{i=1}^N$ el conjunto discreto de valores que forman el rango de X . En general, para variables aleatorias continuas $F_X(x) = F_X(x^-)$, lo que implica que la probabilidad de tomar un valor concreto, $P(X = a) = 0$. En cualquier caso, tanto para variables discretas como continuas, se cumple $F_X(x) = F_X(x^+)$ (ver propiedad 4).

Para variables aleatorias discretas $F_X(x)$ es una función escalonada, con discontinuidades en los valores discretos que forman el rango de la variable aleatoria. Para una variable continua tiene una variación continua. La Figura 2.2 muestra ejemplos de función de distribución discreta, en este caso el experimento lanzar un dado, y continua.



(a) Discreta



(b) Continua

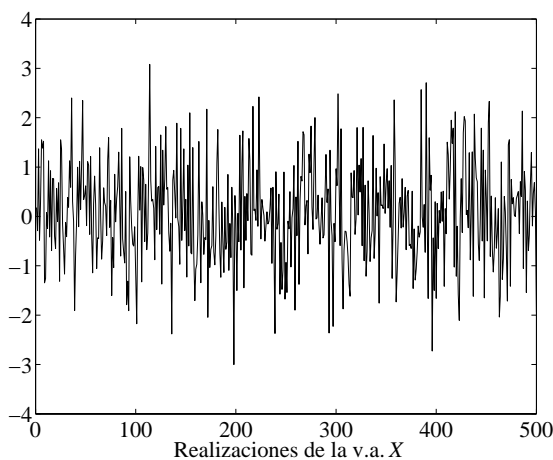
Figura 2.2: Ejemplos de función de distribución de v.a. discreta y v.a. continua.

Interpretación frecuencial (probabilística)

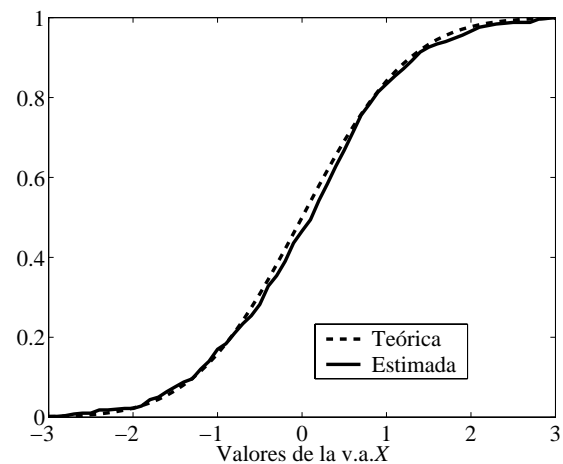
Para presentar una interpretación empírica, constructiva, de la función de distribución, podemos escribir:

$$F_X(x) = P(X \leq x) = \lim_{n \rightarrow \infty} \frac{n_x}{n},$$

donde n es el número de realizaciones del experimento aleatorio, y n_x es el número de resultados para los cuales $X \leq x$. Obviamente no podremos nunca realizar un número infinito de experimentos, pero podemos realizar una estima a partir de un número limitado de los mismos. La Figura 2.3 muestra 500 realizaciones de un experimento y la estima realizada de este modo comparada con la función de distribución teórica.



(a) Realizaciones



(b) Estima de la función de distribución

Figura 2.3: Estima de la función de distribución mediante su interpretación frecuencial.

2.2.2. Función de densidad de probabilidad

La otra función empleada para caracterizar una variable aleatoria es la *función de densidad de probabilidad* (f.d.p.), que se denota como $f_X(x)$. La función de densidad de probabilidad se define como la derivada de la función de distribución

$$f_X(x) = \frac{d}{dx}F_X(x).$$

Esta función indica como se distribuye la probabilidad de la variable aleatoria. Sus principales propiedades son las siguientes:

1. $f_X(x) \geq 0$.
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$.
3. $\int_{a^+}^{b^+} f_X(x)dx = P(a < X \leq b)$.
4. En general, $P(X \in A) = \int_A f_X(x)dx$.
5. $F_X(x) = \int_{-\infty}^{x^+} f_X(u)du$.

En el caso de variables continuas tiene una variación continua, y en el caso de variables discretas, la f.d.p. incluye impulsos situados en los valores discretos que puede tomar la variable (la derivada de una función con escalones). El valor en cada uno de esos valores discretos corresponde a la probabilidad de que la variable aleatoria tome dicho valor.

El matiz a^+ sirve para tratar las señales discretas. En este caso, el impulso está situado en a , e integrar desde a^+ no lo incluye. Para variables continuas podemos utilizar directamente a .

En el caso de variables discretas, en ocasiones en lugar de trabajar con la f.d.p., se trabaja con la *función masa de probabilidad*, o a veces los llamados *puntos de masa*. En el caso de una variable discreta, sólo unos valores concretos $\{x_i\}_{i=1}^N$ son posibles. En ese caso se define la función masa de probabilidad o puntos de masa como

$$p_i = P(X = x_i).$$

En este caso se cumple que

1. $p_i \geq 0$.
2. $\sum_{i=1}^N p_i = 1$.

La diferencia con la f.d.p. es que se suele representar en función de i en lugar de con respecto a x_i , pero conceptualmente es lo mismo.

En otras ocasiones, para variables aleatorias discretas, una vez conocido el espacio muestral $\{x_i\}_{i=1}^N$, las probabilidades de cada uno de los valores de dicho espacio se denotan como $p_X(x_i)$.

En este curso en general se trabajará con la f.d.p., pero cuando se trabaje con variables aleatorias discretas, en ocasiones en lugar de utilizar la notación $f_X(x)$ se utilizará la notación $p_X(x_i)$.

Interpretación frecuencial

Para dar una interpretación empírica de la f.d.p., podemos definir la función densidad de probabilidad como

$$f_X(x) = \lim_{\Delta_x \rightarrow 0} \frac{P(x \leq X \leq x + \Delta_x)}{\Delta_x},$$

es decir

$$f_X(x) = \frac{\text{Probabilidad de un intervalo}}{\text{Longitud del intervalo}} = \text{Densidad de Probabilidad},$$

cuando la longitud del intervalo se lleva al límite infinitesimal. Utilizando la definición frecuencial de la probabilidad,

$$f_X(x) = \lim_{\Delta_x \rightarrow 0} \left\{ \frac{1}{\Delta_x} \lim_{n \rightarrow \infty} \frac{n_x}{n} \right\},$$

donde n es el número de realizaciones del experimento aleatorio, y n_x es el número de resultados para los cuales $x \leq X < x + \Delta_x$.

Esto es equivalente a hacer un histograma, que consiste en dividir la recta real en intervalos de anchura Δ_x y levantar una barra vertical con la frecuencia relativa de cada intervalo. En este caso, se puede comprobar que un histograma tiende a la función densidad de probabilidad cuando el número de realizaciones crece y la longitud del intervalo disminuye. La Figura 2.4 muestra un histograma con un valor $\Delta_x = 0,2$ realizado a partir de 1000 realizaciones y lo compara con la función densidad de probabilidad teórica para una distribución gaussiana de media nula.

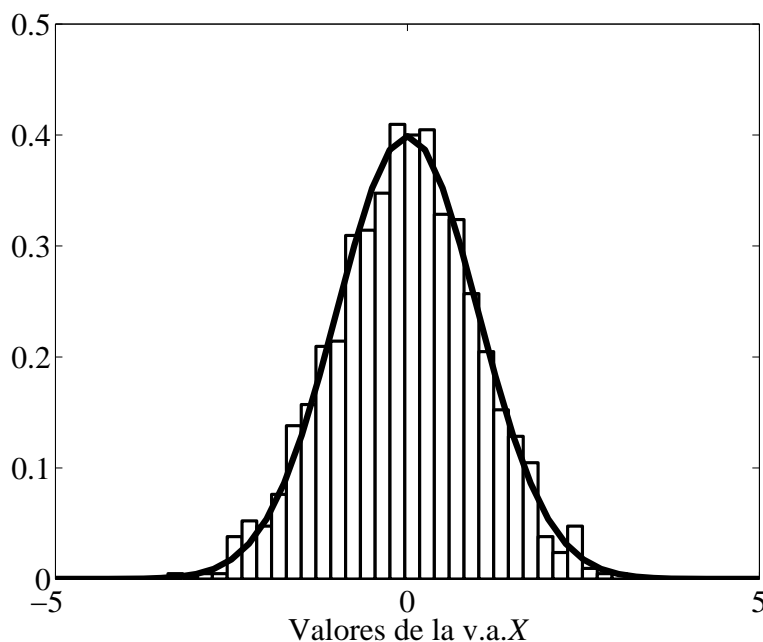


Figura 2.4: Aproximación de la f.d.p. mediante un histograma.

2.2.3. Variables aleatorias de interés

A continuación vamos a ver las variables aleatorias más frecuentemente utilizadas en comunicaciones.

Variable aleatoria de Bernoulli

Esta es una variable aleatoria discreta que toma dos valores, 1 y 0, con probabilidades

- $P(1) = p,$
- $P(0) = 1 - p,$

respectivamente.

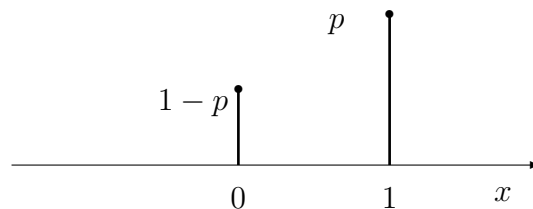


Figura 2.5: $f_X(x)$ de una v.a. de Bernoulli.

Se trata de una distribución con un parámetro, en este caso p . Su función densidad de probabilidad es, obviamente:

$$f_X(x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \\ 0, & \text{en otro caso} \end{cases} .$$

Una variable aleatoria de Bernoulli es un buen modelo para

- *Generador de datos binario.* En este caso, lo normal es que el parámetro p valga $1/2$, es decir, que los 1's y los 0's sean equiprobables.
- *Modelo de errores.* Por otro lado, en cualquier transmisión sobre un canal de comunicaciones se van a producir errores. Un error se puede modelar como la suma módulo-2 (XOR) del bit de entrada con un 1. Por tanto, este tipo de variables también se pueden emplear para modelar errores. En este caso, el parámetro p es precisamente la tasa de errores.

Variable aleatoria binomial

Es también una variable aleatoria discreta. Esta variable modela el número de 1's en una secuencia de n experimentos de Bernoulli independientes, con lo que tiene dos parámetros, n y p . Su función densidad de probabilidad es la siguiente:

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1 - p)^{n-x}, & 0 \leq x \leq n \text{ y } x \in \mathbb{Z} \\ 0, & \text{en otro caso} \end{cases} .$$

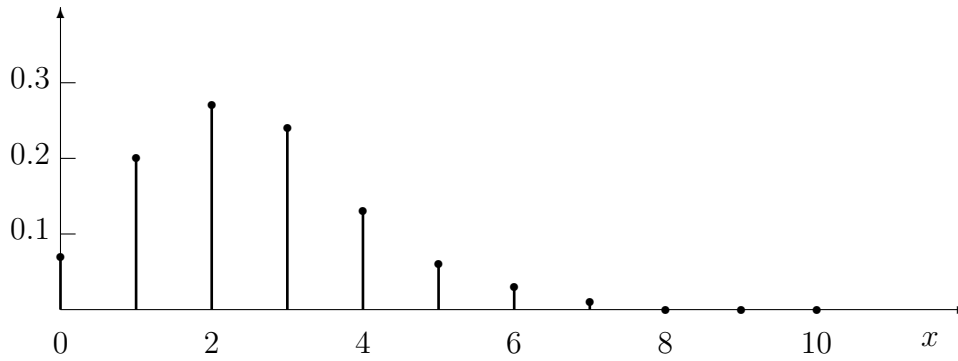


Figura 2.6: $f_X(x)$ de una v.a. binomial.

Esta variable se puede utilizar, por ejemplo, para modelar el *número total de bits recibidos con error* cuando una secuencia de n bits es transmitida a través de un canal con probabilidad de error de bit p .

Variable aleatoria uniforme

Esta es una variable aleatoria continua de dos parámetros, a y b , que toma valores en el intervalo (a,b) con la misma probabilidad para intervalos de igual longitud. Su función de densidad de probabilidad es

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{en otro caso} \end{cases} .$$

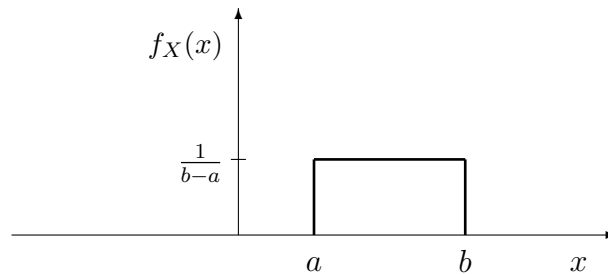


Figura 2.7: $f_X(x)$ de una v.a. uniforme.

En ocasiones se utiliza la notación $\mathcal{U}(a,b)$ para denotar una distribución uniforme entre a y b . Este modelo se utiliza para variables continuas con rango conocido para las cuales nada más se conoce. Por ejemplo, para modelar una *fase aleatoria en una senoide*, se suele emplear una v.a. uniforme entre 0 y 2π .

Variable aleatoria gaussiana o normal

Se trata de una variable aleatoria continua con dos parámetros, μ y σ . Su función densidad de probabilidad es una gaussiana de media μ y varianza σ^2 (o lo que es lo mismo, desviación típica σ),

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} .$$

En ocasiones se denota como $\mathcal{N}(\mu, \sigma^2)$. La gaussiana es la v.a. más importante y la más utilizada

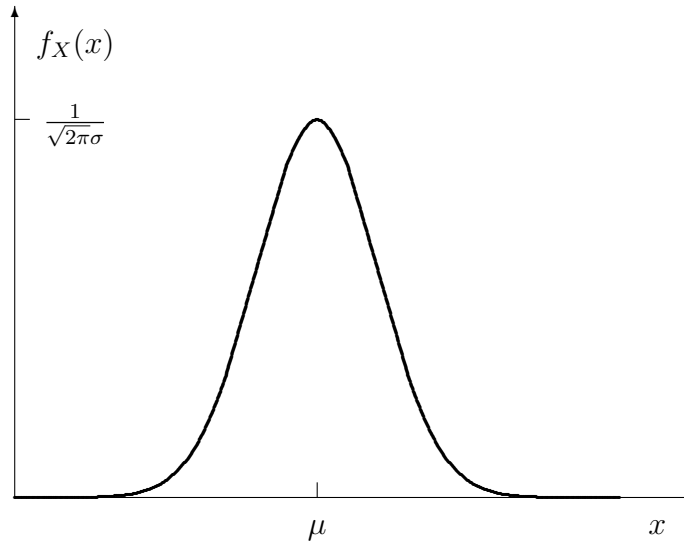


Figura 2.8: Función densidad de probabilidad para una v.a. gaussiana

sin duda en comunicaciones. La principal razón es que el ruido térmico, que es la mayor fuente de ruido en los sistemas de comunicaciones, tiene una distribución gaussiana.

La función de distribución, $F_X(x)$, para una v.a. gaussiana de media nula y varianza unidad se denota comúnmente como $\Phi(x)$

$$\Phi(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

Una función relacionada con esta función de distribución, que se utiliza con mucha frecuencia es la función $Q(x)$, que se define a partir de la función de distribución como

$$Q(x) = 1 - \Phi(x) = P(X > x)$$

lo que proporciona la probabilidad de que la variable aleatoria gaussiana de media nula y varianza unidad tome valores mayores que el argumento de la función. Esta función, definida formalmente como

$$Q(x) = \int_x^{+\infty} f_X(z) dz = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

no tiene solución analítica ($\Phi(x)$ tampoco la tiene), pero es posible calcularla de forma numérica y normalmente se presenta tabulada para sus valores positivos, tal y como se muestra en la Tabla A.3, del Apéndice A. En la Figura 2.9 se muestra la interpretación gráfica del valor de esta función como el área debajo de la curva de la distribución gaussiana $\mathcal{N}(0, 1)$ para valores positivos y negativos del argumento x .

A partir de esta figura es fácil extraer algunas de las propiedades propiedades de esta función

1. $Q(0) = \frac{1}{2}$.
2. $Q(+\infty) = 0$.
3. $Q(-x) = 1 - Q(x)$.

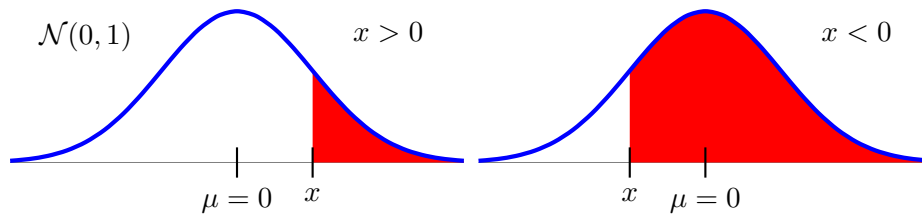


Figura 2.9: Interpretación gráfica de $Q(x)$ para valores positivos y negativos de x .

Debido a la simetría $Q(-x) = 1 - Q(x)$, las tablas de esta función se suelen presentar únicamente para valores positivos del argumento de la función, dado que para valores negativos se puede obtener a partir de esa relación.

Para una distribución con media μ y varianza σ^2 , i.e. $\mathcal{N}(\mu, \sigma^2)$, un simple cambio de variable sirve para estimar $P(X > x)$ a través de la función $Q(x)$ como

$$P(X > x) = Q\left(\frac{x - \mu}{\sigma}\right).$$

La función $Q(x)$ es de gran interés en esta asignatura porque se utilizará como ya veremos, para evaluar probabilidades de error en sistemas de comunicaciones digitales. Dada su importancia en la asignatura, a continuación se ilustrarán algunos ejemplos de cómo se puede utilizar esta función para obtener probabilidades de que una variable aleatoria con una media μ y una varianza σ^2 dadas pueda tomar valores en ciertos rangos. A partir de su definición, es evidente cómo se pueden calcular probabilidades de que una variable aleatoria tome valores mayores que un cierto umbral x . En la Figura 2.10 se muestran algunas de las simetrías de la función $Q(x)$ que permiten también obtener probabilidades de que la variable aleatoria gaussiana tome valores menores que un cierto argumento (a la derecha), a partir de un problema equivalente sobre la propia definición de la función $Q(X)$ (a la izquierda). Y en la Figura 2.11, se ilustra un ejemplo de cómo calcular la probabilidad de que la variable aleatoria tome valores en un intervalo entre dos umbrales, problema que se puede resolver replanteándolo como la diferencia entre dos problemas con un único umbral.

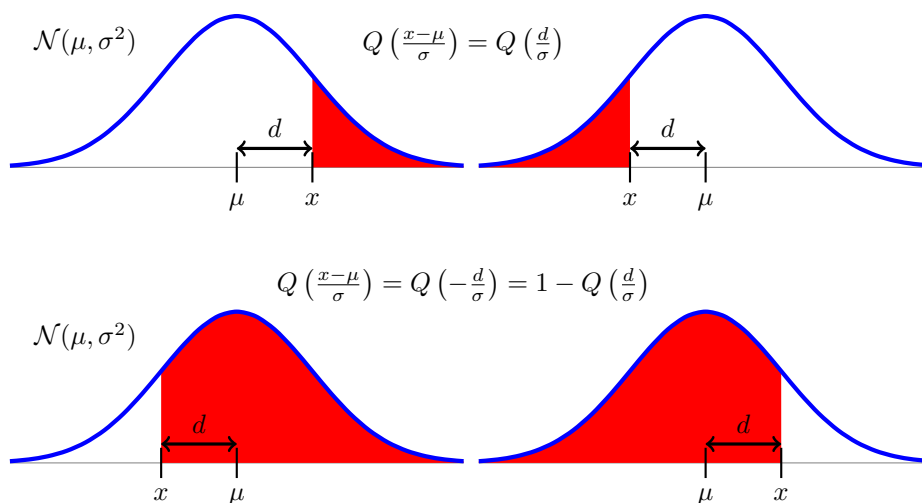


Figura 2.10: Interpretación gráfica de las simetrías de la función $Q(x)$.

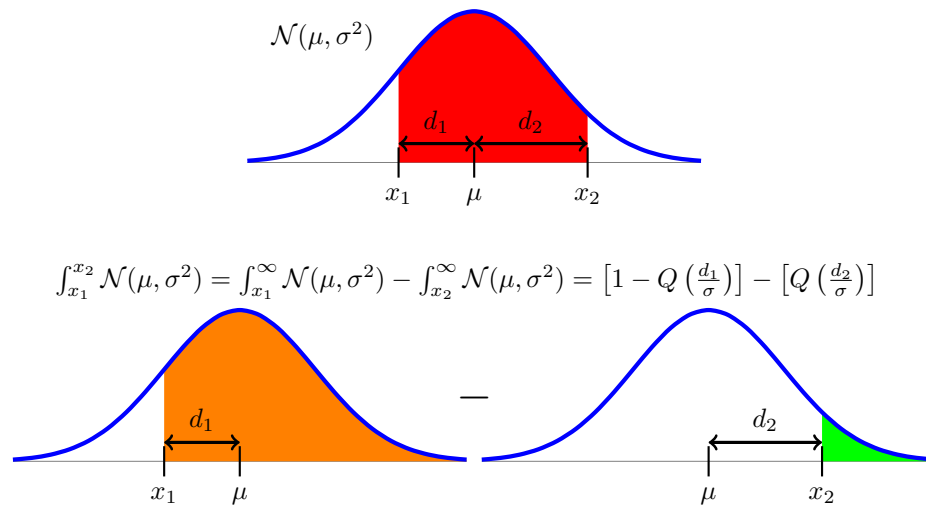


Figura 2.11: Cálculo de la probabilidad de que una variable aleatoria gaussiana tome valores en un intervalo utilizando la función $Q(x)$.

2.2.4. Funciones de una variable aleatoria

Una función de una variable aleatoria $Y = g(X)$ es también ella misma un variable aleatoria. Para encontrar su función de distribución podemos partir de la definición de la misma

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$

Esta probabilidad se reduce a

$$F_Y(y) = P(x \in B_X^g(y)),$$

donde $B_X^g(y)$ es el conjunto de valores de X tales que $g(x) < y$, es decir

$$B_X^g(y) = \{x \in \mathbb{R} : g(x) \leq y\}.$$

Ejemplo

Para la transformación $Y = -2X$, queremos calcular $F_Y(y)$.

En este caso, es sencillo calcular $B_X^g(y)$,

$$B_X^g(y) = \{x \in \mathbb{R} : -2x \leq y\} = \{x \geq -y/2\},$$

de modo que

$$F_Y(y) = P(Y \leq y) = P(X \geq -y/2).$$

Esta probabilidad para una cierta variable aleatoria X se puede calcular conocida $F_X(x)$ o $f_X(x)$.

Por otro lado, la función densidad de probabilidad de la variable aleatoria Y se puede calcular, a partir de $f_X(x)$ y de la transformación $g(x)$, como

$$f_Y(y) = \sum_{i=1}^{N_r} \frac{f_X(x_i)}{|g'(x_i)|},$$

donde N_r y $\{x_i\}_{i=1}^{N_r}$ son el número de soluciones y las propias soluciones de la ecuación $y = g(x)$, respectivamente: la función $g'(x)$ es la derivada de $g(x)$. Para poder obtener esta expresión es preciso que la ecuación tenga un número finito de soluciones, que para todas estas soluciones exista la derivada $g'(x_i)$ y que la derivada en las soluciones no sea nula.

Ejemplo

Tenemos una variable aleatoria X gaussiana con media nula y varianza unidad, es decir $\mu = 0$ y $\sigma = 1$. Queremos encontrar la función densidad de probabilidad de la variable aleatoria

$$Y = aX + b.$$

En este caso $g(x) = ax + b$, y por tanto su derivada es $g'(x) = a$. La ecuación $y = ax + b$ tiene una única solución

$$x_1 = \frac{y - b}{a}$$

Aplicando la expresión para calcular la f.d.p. tenemos

$$f_Y(y) = \frac{f_X\left(\frac{y-b}{a}\right)}{|a|} = \frac{1}{\sqrt{2\pi}|a|} e^{-\frac{(y-b)^2}{2a^2}}.$$

Se puede comprobar que esta distribución es una gaussiana de media b y varianza a^2 , i.e.

$$f_Y(y) = \mathcal{N}(b, a^2).$$

De este ejemplo se puede sacar una conclusión importante: *una función lineal de una variable aleatoria gaussiana es también una variable aleatoria gaussiana.*

2.2.5. Momentos estadísticos

A continuación vamos a ver cómo se calculan algunos momentos estadísticos asociados a una variable aleatoria. Conviene no olvidar que una variable aleatoria representa la salida de un experimento aleatorio. Si se conoce la f.d.p. es posible obtener algunos estadísticos de la misma, lo que equivale a decir estadísticos del experimento aleatorio.

Valor esperado (Media)

El valor esperado (esperanza matemática) de una variable aleatoria es equivalente a su media (aritmética), y a menudo se denota como m_X . El valor esperado mide el valor medio obtenido cuando el número de experimentos es suficientemente grande. Este valor esperado se define como

$$m_X = E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx.$$

Valor esperado de una función de X

El valor esperado de una función de una variable aleatoria, $Y = g(X)$, se obtiene como

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) dx.$$

Momento de orden n

En general, el momento de orden n nos da el valor esperado (la media) de x^n , y se define como

$$m_X^n = \int_{-\infty}^{\infty} x^n \cdot f_X(x) dx.$$

En este caso, la función es $g(x) = x^n$. El valor esperado, la media, es por tanto el momento de orden 1.

Varianza

La varianza se puede ver como el valor esperado para el caso particular de la función

$$g(x) = (x - m_X)^2.$$

Por tanto,

$$\sigma_X^2 = E[(X - m_X)^2] = \int_{-\infty}^{\infty} (x - m_X)^2 \cdot f_X(x) dx.$$

σ_X^2 es la varianza de la v.a. y σ_X es por tanto la desviación típica. Estos parámetros nos dan idea de la variabilidad de la variable aleatoria. Como curiosidad, la media y la varianza tienen la siguiente relación

$$\begin{aligned}\sigma_X^2 &= E[(X - E(X))^2] = E[X^2] - (E[X])^2. \\ \sigma_X^2 &= E[(X - m_X)^2] = E[X^2] - (m_X)^2.\end{aligned}$$

Propiedades

A continuación se presentan algunas de las propiedades de estos estadísticos. En esta lista de propiedades, c denota una constante arbitraria.

1. $E[X + Y] = E[X] + E[Y] = m_X + m_Y$ (Operador lineal)
2. $E[c] = c$
3. $E[c \cdot X] = c \cdot E[X]$
4. $E[X + c] = E[X] + c$
5. $\text{Var}(c) = 0$
6. $\text{Var}(c \cdot X) = c^2 \cdot \text{Var}(X)$
7. $\text{Var}(X + c) = \text{Var}(X)$

2.2.6. Variables aleatorias multidimensionales

Si dos variables aleatorias están definidas sobre el mismo espacio muestral Ω , es posible trabajar con ellas de forma conjunta. Este caso podemos plantearlo como un problema multidimensional, o también como un problema de vectores de variables aleatorias. En este caso seguiremos la primera alternativa.

Funciones de distribución y densidad de probabilidad conjuntas

Para dos variables aleatorias x e Y se define su *función de distribución conjunta* como

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y).$$

La *función de densidad de probabilidad conjunta* se define como

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y).$$

Estas dos funciones tienen las siguientes propiedades (la mayoría extensión de las propiedades para el caso de una única variable aleatoria)

1. $F_X(x) = F_{X,Y}(x, \infty)$.
2. $F_Y(y) = F_{X,Y}(\infty, y)$.
3. $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$.
4. $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$.
5. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$.
6. $P((X, Y) \in A) = \int \int_{(x,y) \in A} f_{X,Y}(x, y) dx dy$.
7. $F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) du dv$.

Función de densidad de probabilidad condicional

Como sucedía para el caso de sucesos, el hecho de conocer el resultado de una variable aleatoria puede condicionar el conocimiento que se tiene sobre la otra. La función de densidad de probabilidad de la variable Y condicionada por $X = x$ se define como

$$f_{Y|X}(y|x) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)}, & f_X(x) \neq 0 \\ 0, & \text{en otro caso} \end{cases}.$$

De aquí surge la definición de *variables aleatorias estadísticamente independientes*. Si el conocimiento de X no aporta nada sobre el conocimiento de Y y viceversa, entonces

$$f_{Y|X}(y|x) = f_Y(y), \text{ y también } f_{X|Y}(x|y) = f_X(x).$$

Por tanto, para este tipo de variables aleatorias se cumple que la distribución conjunta es igual al producto de las distribuciones marginales

$$f_{X,Y}(x, y) = f_X(x) \times f_Y(y).$$

Momentos estadísticos

El valor esperado de una función $g(X, Y)$ de las variables aleatorias X e Y se obtiene como

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f_{X,Y}(x, y) dx dy.$$

Es interesante resaltar los siguientes casos particulares:

- Si $g(X, Y) = X \times Y$, se tiene la esperanza del producto de las dos variables aleatorias, que se denomina la *correlación* entre X e Y .
- En el caso en que $g(X, Y) = (X - m_X) \times (Y - m_Y)$ tenemos la denominada *covarianza* entre ambas variables aleatorias.

La versión normalizada de la covarianza es lo que se conoce como *coeficiente de correlación*, $\rho_{X,Y}$, que se define como

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

Su módulo está limitado entre 0 y 1, es decir $0 \leq |\rho_{X,Y}| \leq 1$, o lo que es lo mismo

$$-1 \leq \rho_{X,Y} \leq +1.$$

Algunos valores particulares de este coeficiente nos aportan una información especial sobre las variables aleatorias implicadas.

- Cuando $\rho_{X,Y} = 0$ se dice que las señales están *incorreladas*. Si dos variables aleatorias son independientes, entonces es fácil comprobar que están incorreladas. Sin embargo, lo recíproco no es cierto: incorrelación no implica independencia.
- Por otro lado, un valor $\rho_{X,Y} = \pm 1$ indica una relación lineal entre las variables aleatorias, es decir $Y = aX + b$. En este caso, $\rho_{X,Y} = +1$ indica un valor positivo de a , mientras que $\rho_{X,Y} = -1$ indica que a es negativo.

Es común utilizar la notación ρ , sin hacer referencia a las variables aleatorias implicadas cuando estas se sobreentienden.

De forma intuitiva, la correlación nos va a indicar el grado de relación estadística entre las dos variables aleatorias. En general, una correlación alta indica una relación alta, y una correlación baja suele indicar una relación baja.

Funciones de variables aleatorias multidimensionales

Sobre variables aleatorias multidimensionales (o múltiples), al igual que para las unidimensionales, se pueden definir funciones sobre las variables X e Y

$$\begin{cases} Z = g(X, Y) \\ W = h(X, Y) \end{cases}.$$

Para obtener $F_{Z,W}(z, w)$ se procede como en el caso unidimensional.

$$F_{Z,W}(z, w) = P(Z \leq z, W \leq w) = P((x, y) \in B_{X,Y}^{g,h}(z, w)),$$

donde en este caso

$$B_{X,Y}^{g,h}(z, w) = \{(x, y) \in \mathbb{R}^2 : g(x, y) \leq z, h(x, y) \leq w\}.$$

Al igual que en el caso de una única v.a., si se conocen las raíces (soluciones) $\{x_i, y_i\}$ del sistema de ecuaciones

$$\begin{cases} z = g(x, y) \\ w = h(x, y) \end{cases},$$

entonces la f.d.p. de las nuevas variables se obtiene mediante la expresión

$$f_{Z,W}(z, w) = \sum_i \frac{f_{X,Y}(x_i, y_i)}{|\det \mathbf{J}(x_i, y_i)|}.$$

donde $\det \mathbf{J}$ denota el determinante de la matriz jacobiano \mathbf{J} . Se necesita que el número de soluciones sea finito y que el jacobiano sea no nulo. El jacobiano se define como

$$\mathbf{J}(x, y) = \begin{bmatrix} \frac{\partial z(x,y)}{\partial x} & \frac{\partial z(x,y)}{\partial y} \\ \frac{\partial w(x,y)}{\partial x} & \frac{\partial w(x,y)}{\partial y} \end{bmatrix}.$$

Para poder calcular la nueva distribución conjunta es necesario que el número de raíces sea finito y que el determinante sea no nulo para todas ellas.

Todo lo que hemos estado viendo aplicado a dos variables aleatorias se puede extender de forma inmediata a un número mayor de variables aleatorias.

Variables aleatorias conjuntamente gaussianas

Las variables aleatorias conjuntamente gaussianas también se denominan en ocasiones variables aleatorias gaussianas multidimensionales. Por su importancia para esta asignatura, a continuación se presentan algunas de sus propiedades. En primer lugar, se definen probabilísticamente. Dos variables aleatorias X e Y conjuntamente gaussianas están caracterizadas por una función densidad de probabilidad conjunta que es una gaussiana bidimensional

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{2\sigma_X^2} + \frac{(y-\mu_Y)^2}{2\sigma_Y^2} - \frac{\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right)}$$

Cuando se tiene este tipo de distribución conjuntamente gaussiana en las variables X e Y , no sólo X e Y van a tener una distribución gaussiana (son v.a. gaussianas) sino que además las probabilidades condicionales también son gaussianas. Esta es la principal diferencia entre dos variables aleatorias que cada una tiene una distribución gaussiana y dos variables aleatorias con una distribución conjuntamente gaussiana. Con una distribución conjuntamente gaussiana, las variables aleatorias individuales son de la forma siguiente: X es gaussiana de media μ_X y varianza σ_X^2 , Y es gaussiana de media μ_Y y varianza σ_Y^2 , y además su coeficiente de correlación es ρ .

Este concepto se puede extender a un número arbitrario n de variables aleatorias, llegándose a la expresión para la distribución de una gaussiana n -dimensional, parametrizada por un vector de medias y una matriz de covarianzas

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})^T}.$$

donde la variable aleatoria $\mathbf{X} = (X_1, X_2, \dots, X_n)$, $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, y el vector de medias es $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_n]^T$. Finalmente, \mathbf{C} es la matriz de covarianzas, que incluye en la fila i y columna j la covarianza entre las variables i -ésima y j -ésima

$$C_{i,j} = \text{Cov}(X_i, X_j) = \rho_{i,j}\sigma_i\sigma_j,$$

es decir,

$$\mathbf{C} = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,n}\sigma_1\sigma_n \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2,n}\sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1,n}\sigma_1\sigma_n & \rho_{2,n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix}.$$

Las propiedades de las variables aleatorias conjuntamente gaussianas son

1. Las variables aleatorias conjuntamente gaussianas están completamente caracterizadas por su vector de medias $\boldsymbol{\mu}$ y su matriz de covarianza \mathbf{C} . A estos dos parámetros se les denomina *estadísticos de segundo orden*, y describen completamente estas variables aleatorias.
2. Si n variables aleatorias son conjuntamente gaussianas, cualquier subconjunto también está distribuido de forma conjuntamente gaussiana. En particular, todas las variables individuales son gaussianas.
3. Cualquier subconjunto de v.a. conjuntamente gaussianas, condicionadas a otro subconjunto de las mismas v.a. conjuntamente gaussianas originales, tiene una distribución conjuntamente gaussiana, aunque los parámetros se modifican en este caso.
4. Cualquier conjunto de variables aleatorias obtenidas como combinaciones de lineales de (X_1, X_2, \dots, X_n)

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

es conjuntamente gaussiano. En particular, individualmente cualquier combinación lineal Y_i es gaussiana.

5. Dos variables aleatorias conjuntamente gaussianas incorreladas son independientes. Por tanto, *para v.a. conjuntamente gaussianas, independencia e incorrelación son equivalentes*. Esto no es cierto en general para otro tipo de v.a.
6. Si las señales están incorreladas, $\rho_{i,j} = 0 \forall i \neq j$, es decir, que \mathbf{C} es una matriz diagonal.

Suma de variables aleatorias

Si tenemos una secuencia de variables aleatorias, (X_1, X_2, \dots, X_n) , que tienen básicamente las mismas propiedades, parece lógico pensar que el comportamiento del promedio de las mismas,

$$Y = \frac{1}{n} \sum_{i=1}^n X_i,$$

sea, por así decirlo, “menos aleatorio”. La *ley de los grandes números* y el *teorema del límite central* plantean de forma rigurosa esta intuición.

Ley de los grandes números (débil) Esta ley plantea que si las variables aleatorias (X_1, X_2, \dots, X_n) están *incorreladas* y todas tienen la misma media m_X y varianza $\sigma_X^2 < \infty$, independientemente de su distribución, para cualquier $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P(|Y - m_X| > \varepsilon) = 0.$$

Esto significa que el promedio (Y) converge, en probabilidad, al valor esperado de las v.a. X_i . Es decir, que cuantas más variables sumemos, más se parece su combinación a la media (menor es su varianza).

Teorema del límite central Este teorema va un poco más allá que la Ley de los grandes números. No sólo dice que el promedio de variables aleatorias converge a la media sino que nos dice como es su distribución. En concreto, el teorema plantea que: si (X_1, X_2, \dots, X_n) son variables n aleatorias *independientes* con medias m_1, m_2, \dots, m_n , y varianzas $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, entonces la distribución de

$$Y = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - m_i}{\sigma_i}$$

converge a una distribución gaussiana de media 0 y varianza 1, $\mathcal{N}(0, 1)$.

En el caso particular de que sean *independientes e idénticamente distribuidas (i.i.d)*, es decir, que todas tengan la misma distribución con la misma media m y la misma varianza σ^2 , el promedio

$$Y = \frac{1}{n} \sum_{i=1}^n X_i,$$

converge a una distribución $\mathcal{N}(m, \frac{\sigma^2}{n})$. Esto es así aunque la distribución original no sea gaussiana.

Nota: Conviene recordar que la ley de los grandes números es válida para variables aleatorias incorreladas mientras que el teorema del límite central exige independencia entre las variables aleatorias, lo que supone una restricción más fuerte.

2.3. Procesos Aleatorios

Un proceso aleatorio, o proceso estocástico, es la extensión natural del concepto de variable aleatoria para trabajar con señales. Los sistemas de comunicaciones trabajan con señales, señales que dependen del tiempo. Como ya se ha mencionado en varias ocasiones, a veces es posible caracterizar estas señales de forma determinista y en otras ocasiones será preciso tratarlas como señales aleatorias: los ejemplos más claros son el ruido térmico en cualquier dispositivo, o las propias señales de información. Estas señales se van a caracterizar como procesos aleatorios.

Tal vez la forma más intuitiva de ver lo que es un proceso aleatorio es verlo como un conjunto de señales temporales que corresponden a cada una de las posibles salidas de un experimento aleatorio. Cada salida de un experimento tiene asociada una función temporal. Una variable aleatoria real asignaba un valor real a cada valor del espacio muestral ($\Omega \rightarrow \mathbb{R}$), es decir, $\lambda_i \in \Omega \rightarrow X(\lambda_i) \in \mathbb{R}$. Un proceso estocástico se puede interpretar como una situación en la que la asignación de valores del espacio muestral a la recta real varía con el tiempo $X(t, \lambda)$. Desde este punto de vista, cada salida del experimento tiene asociada una función temporal que especifica su valor real asignado en un instante t .

Ejemplo

- Definición del experimento aleatorio
 - Cuatro posibles resultados: $\lambda_1, \lambda_2, \lambda_3, \lambda_4$
 - Probabilidades: $P(\lambda_1) = P(\lambda_2) = P(\lambda_3) = P(\lambda_4) = \frac{1}{4}$
- Descripción analítica del proceso

$$X(t, \lambda_1) = \begin{cases} 1, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

$$X(t, \lambda_2) = \begin{cases} 2, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

$$X(t, \lambda_3) = \begin{cases} e^{-t}, & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

$$X(t, \lambda_4) = \begin{cases} \sin(t), & \text{si } t \geq 0 \\ 0, & \text{si } t < 0 \end{cases}$$

Cada una de las señales que forman parte del proceso aleatorio se representan en la Figura 2.12.

Por tanto podemos interpretar que el proceso aleatorio es en realidad un conjunto de señales, cada una de ellas asociada a uno de los posibles valores de salida del experimento aleatorio asociado. En particular, sobre el proceso se pueden identificar los siguientes valores o funciones:

- $X(t_i, \lambda_j) \rightarrow$ una realización individual del experimento.
- $X(t, \lambda_i) \rightarrow$ una señal temporal que indica el número real asignado en cada instante a una posible salida, λ_i , del espacio muestral. Por simplificar la notación, en ocasiones se denotará también como $x_i(t)$.

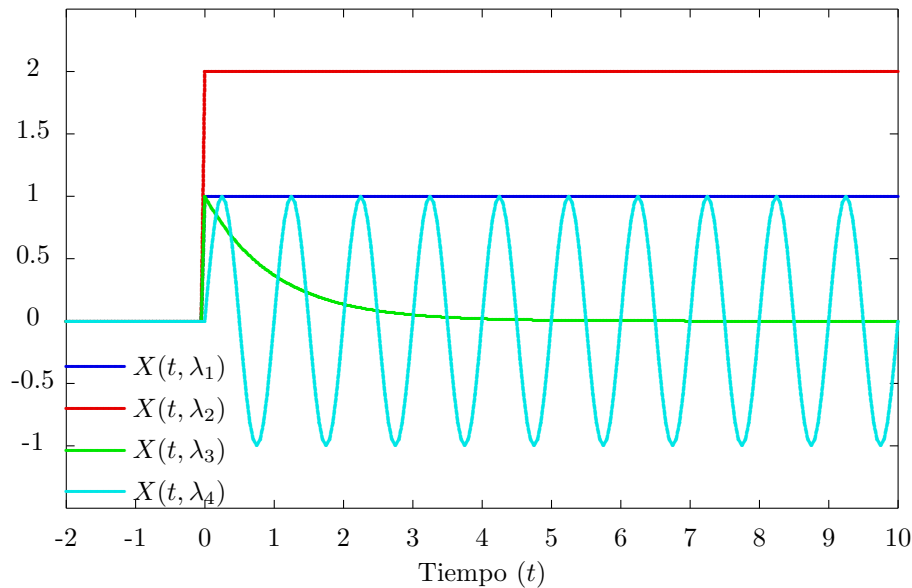


Figura 2.12: Ejemplo de funciones asignadas a cada salida del espacio Ω .

- $X(t_i, \lambda) \rightarrow$ una variable aleatoria (X). Así pues, en cualquier instante de tiempo t_i el valor de un proceso aleatorio es una variable aleatoria.

Ejemplo

(Ω, B, P) es el espacio de probabilidad del experimento aleatorio lanzar un dado. En este caso

$$\Omega = \{1, 2, 3, 4, 5, 6\} = \{\lambda_i = i\}_{i=1,2,\dots,6}.$$

Podemos tener la transformación

$$X(t, \lambda_i) = \lambda_i e^{-t} u(t).$$

En este caso, $X(0, \lambda)$ es una variable aleatoria que toma los valores $\{1, 2, \dots, 6\}$ con una probabilidad $1/6$ para cada uno de esos seis valores. $X(1, \lambda)$ es otra variable aleatoria que toma los valores $\{1e^{-1}, 2e^{-1}, \dots, 6e^{-1}\}$ cada uno con una probabilidad $1/6$. Y en general, para cualquier valor $t \geq 0$, $X(t, \lambda)$ es una variable aleatoria que toma los valores $\{1e^{-t}, 2e^{-t}, \dots, 6e^{-t}\}$ cada uno con una probabilidad $1/6$ (para valores $t < 0$, el proceso aleatorio toma siempre el valor 0).

A la vista de la definición de un proceso aleatorio y del ejemplo anterior, otra forma de interpretar un proceso aleatorio es como un conjunto indexado de variables aleatorias, donde el índice es un índice temporal, bien en tiempo continuo o en tiempo discreto

$$\{X(t_1), X(t_2), \dots\}, \text{ ó } \{X[n_1], X[n_2], \dots\},$$

o en general como

$$\{X(t), t \in \mathbb{R}\}, \text{ ó } \{X[n], n \in \mathbb{Z}\}.$$

Es decir, que un proceso aleatorio es un conjunto de variables aleatorias indexadas por un cierto índice temporal (continuo o discreto).

- Si el índice es t , tomando valores en el conjunto continuo número reales $t \in \mathbb{R}$, el proceso es un *proceso aleatorio en tiempo continuo*

- Si el índice es el conjunto discreto de valores n , con $n \in \mathbb{Z}$, el proceso es un *proceso aleatorio en tiempo discreto*.

Por esta razón, en muchas ocasiones en la notación se obvia la dependencia con la salida del experimento aleatorio λ , pasando a utilizarse la notación $X(t)$ o $X[n]$.

A continuación se van a estudiar los procesos aleatorios en tiempo continuo, y posteriormente se extenderán los principales resultados, de forma trivial, al caso de tiempo discreto.

2.3.1. Descripción de un proceso aleatorio

Un proceso aleatorio se puede describir mediante dos tipos de descripciones:

- Analítica
- Estadística

La *descripción analítica* utiliza una expresión analítica compacta de la transformación, donde por un lado aparece la dependencia temporal, y por otro la dependencia con un conjunto de variables aleatorias θ

$$X(t) = f(t, \theta),$$

El vector $\theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ es un vector de variables aleatorias que modifica algún parámetro de la función $f(\cdot, \cdot)$.

La descripción analítica incluye la expresión analítica involucrando índice temporal y variables aleatorias, y la descripción estadística de las variables aleatorias involucradas (variables incluidas en θ).

Ejemplo

- $X(t) = A \cos(2\pi f_0 t + \theta)$, siendo A y f_0 dos valores constantes y θ una variable aleatoria uniforme en $[0, 2\pi)$. Esta es una descripción analítica de un proceso aleatorio en tiempo continuo que podría por ejemplo utilizarse para modelar la salida de un oscilador en un sistema de comunicaciones (asignando a A y a f_0 los valores de amplitud de tensión y la frecuencia del oscilador, respectivamente).
- $X[n] = A\theta$, siendo A un valor constante y θ una v.a. de Bernoulli con $p = 0,5$. Esta es una descripción analítica de un proceso aleatorio en tiempo discreto que puede por ejemplo utilizarse para representar estadísticamente la transmisión de una secuencia de bits (haciendo la constante $A = 1$) con unos y ceros equiprobables.

La descripción analítica proporciona información intuitiva sobre el proceso aleatorio, porque describe de forma analítica la realización del mismo. Sin embargo, no siempre es posible disponer de este tipo de descripción en aplicaciones reales. En este caso, hay que acudir a la *descripción estadística*. Existen varios tipos de descripción estadística. A continuación se describen los más habituales.

Descripción estadística completa

Una *descripción estadística completa* de un proceso aleatorio $X(t)$ consiste en conocer para cualquier conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$, para cualquier valor de n , la función densidad de probabilidad conjunta de las n variables aleatorias resultantes de evaluar el proceso aleatorio en los n instantes de tiempo especificados, $\{X(t_1), X(t_2), \dots, X(t_n)\}$

$$f_{X(t_1), X(t_2), \dots, X(t_n)}(x_1, x_2, \dots, x_n).$$

Descripción de estadísticos de orden M

Se trata de una descripción similar a la anterior en la que se limita el máximo valor de n . En particular, esta descripción consiste en conocer, para todo $n \leq M$ y todo conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$, la función densidad de probabilidad conjunta de $\{X(t_1), X(t_2), \dots, X(t_n)\}$.

Un caso muy especial para la aplicación a sistemas de comunicaciones es el caso $M = 2$, donde se conoce para cualquier pareja de instantes de tiempo (t_1, t_2)

$$f_{X(t_1), X(t_2)}(x_1, x_2).$$

En este caso se conocen los estadísticos de segundo orden.

Ejemplo

Se dice que un proceso $X(t)$ tiene, para cualquier n y cualquier $(t_1, t_2, \dots, t_n) \in \mathbb{R}^n$, una f.d.p. conjunta de $\{X(t_i)\}_{i=1}^n$ conjuntamente gaussiana con media nula y una matriz de covarianza dada por

$$C_{i,j} = \text{Cov}(X(t_i), X(t_j)) = \sigma^2 \min(t_i, t_j).$$

Esta es una descripción estadística completa de $X(t)$.

En este último ejemplo, aunque se proporciona una descripción estadística completa del proceso, se tiene poca información de como se como es cada realización del mismo. Por eso se recurre en muchos casos a definir una serie de promedios estadísticos que nos dan esa información.

2.3.2. Esperanzas de los procesos (promedios estadísticos)

Al igual que vimos para variables aleatorias, para procesos aleatorios se pueden calcular algunos parámetros estadísticos basados en valores esperados. En particular se van a definir a continuación los dos estadísticos más importantes en el dominio temporal

- Media del proceso aleatorio, $m_X(t)$.
- Función de autocorrelación del proceso aleatorio, $R_X(t_1, t_2)$.

Media de un proceso aleatorio

La *media* o *valor esperado* (o *esperanza matemática*) de un proceso aleatorio $X(t)$ es una función del tiempo, $m_X(t)$, que proporciona para cada instante de tiempo la media de la variable aleatoria $X(t)$, es decir, que $m_X(t) = E[X(t)]$ para todo t .

Si se dispone de la descripción estadística del proceso, y si para cualquier instante t la f.d.p. $f_{X(t)}(x)$ está definida, esta función se puede calcular a partir de la misma como

$$m_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x \cdot f_{X(t)}(x) dx.$$

Función de autocorrelación de un proceso aleatorio

Otro promedio estadístico importante de un proceso aleatorio es la *función de autocorrelación*. Esta función es importante porque, como se verá posteriormente, está relacionada con la descripción en el dominio frecuencial del proceso aleatorio, y con la potencia del proceso. Es una función que depende de dos instantes temporales. En ocasiones se denota como $R_{X,X}(t_1, t_2)$, aunque lo normal es emplear la notación $R_X(t_1, t_2)$.

La función de autocorrelación se define como la esperanza matemática del producto del proceso aleatorio evaluado en los dos instantes que son los argumentos de la función

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)],$$

es decir, la esperanza del producto de las v.a. en los instantes t_1 y t_2 .

De nuevo, si se dispone de la descripción estadística del proceso aleatorio, la función de autocorrelación se puede obtener a través de la función densidad de probabilidad conjunta del proceso en dos instantes como

$$R_X(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdot x_2 \cdot f_{X(t_1), X(t_2)}(x_1, x_2) dx_1 dx_2.$$

En resumen, se han visto dos métodos para describir los procesos aleatorios: descripción analítica y descripción estadística. En este caso se puede tener una descripción estadística completa o de orden M . Lo habitual en sistemas de comunicaciones es tener una descripción de segundo orden. Esta descripción a veces no es práctica en el sentido de que no permite tener una idea clara de como son las realizaciones del proceso. En este caso se utilizan estadísticos que dan promedios sobre realizaciones. En particular son interesantes la media y la función de autocorrelación. Como veremos más tarde, en algunos casos la media y la función de autocorrelación proporcionan una descripción estadística completa del proceso aleatorio (será así para los procesos aleatorios gaussianos).

2.3.3. Estacionariedad y cicloestacionariedad

En una descripción estadística completa de un proceso aleatorio se conoce la f.d.p. conjunta de $\{X(t_1), X(t_2), \dots, X(t_n)\}$

$$f_{X(t_1), X(t_2), \dots, X(t_n)}(x_1, x_2, \dots, x_n).$$

para cualquier conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$ y para cualquier valor de n .

En general, esta distribución conjunta depende de la elección del origen de tiempos. Pero existe una clase muy importante de procesos aleatorios en la que esa función es independiente del origen de tiempos elegido. Es decir, que estos procesos tienen unas propiedades estadísticas que no varían con el tiempo. A estos procesos se les llama *procesos estacionarios*.

Existen dos definiciones de estacionariedad, en sentido estricto y en sentido amplio. Cada una de las definiciones se muestra a continuación.

Estacionariedad en sentido estricto

Un proceso es *estacionario en sentido estricto* si para cualquier conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$, para cualquier valor entero n , y cualquier valor Δ se cumple que

$$f_{X(t_1), X(t_2), \dots, X(t_n)}(x_1, x_2, \dots, x_n) = f_{X(t_1+\Delta), X(t_2+\Delta), \dots, X(t_n+\Delta)}(x_1, x_2, \dots, x_n).$$

Es decir, que la f.d.p. del proceso aleatorio en cualquier conjunto de n instantes de tiempo no depende del origen de coordenadas sino de la diferencia relativa de los n instantes de tiempo $\{t_i\}_{i=1}^n$ que forman parte del conjunto.

Cuando esto sólo se cumple para $n \leq M$, entonces se dice que el proceso es *estacionario de orden M* .

La estacionariedad en sentido amplio es una restricción muy fuerte que muy pocos procesos reales pueden cumplir. Por ello se utiliza a menudo una definición menos restrictiva de estacionariedad.

Estacionariedad en sentido amplio

Un proceso $X(t)$ es *estacionario en sentido amplio* si se cumplen las siguientes condiciones:

1. La media es un valor constante que no depende del tiempo

$$m_X(t) = E[X(t)] = m_X$$

2. La función de autocorrelación no depende de forma explícita de cada uno de los dos valores de tiempo, t_1 y t_2 , sino únicamente de su diferencia

$$R_X(t_1, t_2) = R_X(t_1 - t_2) = R_X(\tau)$$

Al escribir esta condición, para resaltar el hecho de que sólo depende de la diferencia de instantes de tiempo, se suele utilizar la parametrización $t_1 = t + \tau$ y $t_2 = t$, de modo que se puede escribir

$$R_X(t + \tau, t) = R_X(\tau).$$

A partir de ahora, cuando se hable de estacionariedad se hará referencia a estacionariedad en sentido amplio, que se denotará en ocasiones a través del acrónimo WSS (del inglés *Wide-Sense Stationary*).

Función de autocorrelación de procesos estacionarios

La función de autocorrelación de un proceso estacionario $X(t)$, $R_X(\tau)$, tiene las siguientes propiedades:

1. $R_X(-\tau) = R_X(\tau)$. Es una función par.

$$R_X(\tau) = E[X(t)X(t-\tau)] = E[X(t-\tau)X(t)] = R_X(-\tau).$$

2. $|R_X(\tau)| \leq R_X(0)$. El máximo en módulo se obtiene en $\tau = 0$.

$$E[(X(t) \pm X(t-\tau))^2] \geq 0$$

Expandiendo este resultado

$$E[X^2(t)] + E[X^2(t-\tau)] \pm 2E[X(t)X(t-\tau)] \geq 0.$$

$$R_X(0) + R_X(0) \pm 2R_X(\tau) \geq 0 \rightarrow R_X(0) \geq \pm R_X(\tau) \rightarrow R_X(0) \geq |R_X(\tau)|.$$

3. Si para algún T_o se cumple $R_X(T_o) = R_X(0)$, entonces para todo entero k

$$R_X(kT_o) = R_X(0).$$

Se demuestra por inducción.

4. Es una función semidefinida positiva. Formalmente, esto significa que, para cualquier función $g(t)$, se cumple que

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(t) R_X(t-s) g(s) dt ds \geq 0.$$

Esta propiedad tiene una importante implicación práctica, y es que que la transformada de Fourier de la función de autocorrelación es no negativa (sólo toma valores mayores o iguales a cero, pero nunca valores negativos). La demostración de esta propiedad se verá de forma trivial más tarde, cuando se estudie la representación en el dominio frecuencial de un proceso aleatorio.

Procesos cicloestacionarios

Existe una clase específica de procesos no estacionarios muy ligados a los procesos estacionarios, que son los denominados *procesos cicloestacionarios*. En este caso, las propiedades estadísticas no son constantes a lo largo del tiempo, sino que varían a lo largo del mismo, pero estas variaciones son periódicas en el tiempo. La definición de cicloestacionariedad de un proceso aleatorio (en sentido amplio) se reduce de nuevo a condiciones sobre la media y la función de autocorrelación del proceso. En concreto, un proceso aleatorio $X(t)$ será cicloestacionario (en sentido amplio) si su media y su función de autocorrelación son periódicas con un cierto período T_0 , es decir, si se cumple que

1. $m_X(t + T_0) = m_X(t)$.
2. $R_X(t + \tau + T_0, t + T_0) = R_X(t + \tau, t)$, para todo t y τ .

2.3.4. Ergodicidad

Para ver qué es un proceso ergódico se utilizará un ejemplo sencillo. Supongamos que se tiene un proceso $X(t)$ que es estacionario, y se tienen las distintas funciones temporales del proceso, $X(t, \lambda_i) \equiv x_i(t)$.

Es posible calcular la media del proceso en un instante t_0 y en otro instante t_1 . Como el proceso es estacionario ambos valores coinciden

$$m_X(t_0) = m_X(t_1) = m_X.$$

Supongamos que ahora se calcula la media en el tiempo de la realización de índice i , $x_i(t)$, que se denotará como m_i . Si se cumple que ese valor de la media de cualquier realización coincide con el valor m_X , es decir $m_X = m_i \forall i$, estamos ante un *proceso ergódico en la media*. Esta idea puede extenderse a otros estadísticos del proceso aleatorio. Por tanto, un proceso ergódico permite calcular promedios estadísticos sobre realizaciones del proceso aleatorio a partir del promedio temporal de una única realización. Se sustituyen promedios estadísticos por promedios temporales.

A continuación se realiza la definición formal de la ergodicidad..

Proceso ergódico

Para un proceso $X(t)$ estacionario en sentido estricto y para cualquier función $g(x)$ se pueden definir dos tipos de promedios

1. Para un instante t y diferentes realizaciones del proceso tenemos una variable aleatoria $X(t)$ con función de densidad $f_{X(t)}(x)$ independiente de t , al ser el proceso estacionario en sentido amplio. Se puede calcular la *esperanza estadística (promedio estadístico)* de la función $g(X)$ como

$$E[g(X(t))] = \int_{-\infty}^{\infty} g(x) f_{X(t)}(x) dx.$$

Este valor es independiente de t si el proceso es estacionario.

2. Si se toma una realización individual del proceso, se tiene una función temporal determinista $x(t, \lambda_i)$. Se puede calcular la *esperanza temporal (promedio temporal)* de esa función

$$\langle g(x) \rangle_i = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(X(t, \lambda_i)) dt.$$

Este valor $\langle g(x) \rangle_i$ es un valor independiente de t pero que en general dependerá de la realización, de modo que para cada λ_i vamos a tener un $\langle g(x) \rangle_i$ distinto en principio. Si sucede que $\langle g(x) \rangle_i$ es independiente de i , es decir, vale lo mismo para todo i y además

$$\langle g(x) \rangle_i = E[g(X(t))],$$

el proceso es *ergódico*.

Por tanto, un proceso aleatorio estacionario $X(t)$, es ergódico si para cualquier función $g(x)$ y para cualquier $\lambda_i \in \Omega$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(X(t, \lambda_i)) dt = E[g(X(t))].$$

Esto significa que si todos los promedios temporales son iguales a los promedios estadísticos, el proceso estacionario es ergódico.

Por tanto, para estimar los estadísticos, media y autocorrelación, de un proceso estacionario ergódico es suficiente tener una única realización del mismo. A partir de esta realización, es posible obtener los promedios estadísticos de interés a través de los promedios temporales. Por ejemplo, se pueden calcular la media y la autocorrelación de la señal utilizando las funciones $g(x)$ adecuadas.

2.3.5. Potencia y energía

Se habían definido dos tipos de señales deterministas, señales de potencia y señales de energía. Estas definiciones se pueden extender para realizaciones de procesos aleatorios. Si se tiene una realización $X(t, \lambda_i)$, que como ya se ha dicho también se denotará como $x_i(t)$, la energía y la potencia de la realización se definen respectivamente como

$$E_i = \int_{-\infty}^{\infty} |x_i(t)|^2 dt,$$

y

$$P_i = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x_i(t)|^2 dt.$$

Para cada $\lambda_i \in \Omega$ existen un número real E_i y otro P_i que denotan la energía y potencia respectivamente. Consecuentemente, tanto energía como potencia son variables aleatorias que vamos a denotar como \mathcal{E}_X y \mathcal{P}_X respectivamente.

Sobre estas variables aleatorias se pueden definir promedios estadísticos que darán idea del contenido de energía o potencia del proceso. Los promedios que se definen habitualmente son

- Energía del proceso aleatorio $X(t)$, E_X .
- Potencia del proceso aleatorio $X(t)$, P_X .

Estos promedios se definen como

$$E_X = E[\mathcal{E}_X]$$

$$P_X = E[\mathcal{P}_X]$$

donde

$$\mathcal{E}_X = \int_{-\infty}^{\infty} |X(t)|^2 dt,$$

y

$$\mathcal{P}_X = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |X(t)|^2 dt.$$

En este caso, al igual que para señales deterministas

- Un proceso aleatorio es de energía si $E_X < \infty$
- Un proceso aleatorio es de potencia si $0 < P_X < \infty$

Teniendo en cuenta estas definiciones, la energía del proceso se obtiene como

$$\begin{aligned} E_X &= E \left[\int_{-\infty}^{\infty} |X(t)|^2 dt \right] \\ &= \int_{-\infty}^{\infty} E [|X(t)|^2] dt \\ &= \int_{-\infty}^{\infty} R_X(t, t) dt. \end{aligned}$$

y la potencia del proceso viene dada por

$$\begin{aligned}
 P_X &= E \left[\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |X(t)|^2 dt \right] \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} E [|X(t)|^2] dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t, t) dt.
 \end{aligned}$$

Se puede observar que en ambos casos, el resultado depende de la función de autocorrelación del proceso, evaluada en el mismo instante de tiempo, es decir, $R_X(t, t)$.

Para *procesos aleatorios estacionarios* $R_X(t, t) = R_X(0)$, es independiente de t , y por tanto

$$P_X = R_X(0),$$

y

$$E_X = \int_{-\infty}^{\infty} R_X(0) dt.$$

Para que el proceso sea de energía debe cumplirse que $E_X < \infty$. Esto sólo es posible para el caso $R_X(0) = E[X^2(t)] = 0$, lo que significa que para todo t y para cualquier realización $X(t) = 0$. Así pues, para el caso de procesos aleatorios estacionarios, sólo los procesos de potencia son de interés práctico, y en ese caso la potencia del proceso aleatorio se puede obtener evaluando la función de autocorrelación, $R_X(\tau)$ en el origen (en $\tau = 0$).

Finalmente, si el proceso es, además de estacionario, ergódico, entonces \mathcal{P} ya no es realmente una variable aleatoria, pues todas las realizaciones tienen la misma potencia, que es precisamente la potencia del proceso

$$P_i = P_X = R_X(0).$$

2.3.6. Procesos aleatorios multidimensionales (múltiples)

Al igual que sucedía para el caso de variables aleatorias, es posible trabajar con varios procesos aleatorios (estocásticos), definidos sobre el mismo espacio de probabilidad, de forma simultánea. Además, cuando se trabaja con sistemas de comunicaciones, esto aparece de forma natural. Por ejemplo, cuando se modela la entrada de un sistema con un proceso aleatorio $X(t)$, tenemos asociado al sistema la salida del mismo pasando por un sistema LTI. Cada realización $X(t, \lambda_i)$ tiene asociada una salida

$$X(t, \lambda_i) \rightarrow Y(t, \lambda_i) = X(t, \lambda_i) * h(t)$$

Esto se puede interpretar como que para cada $\lambda_i \in \Omega$ se tienen asociadas dos señales temporales $X(t, \lambda_i)$ e $Y(t, \lambda_i)$. Por tanto, Y otro proceso aleatorio $Y(t)$, y se tienen dos procesos aleatorios definidos sobre el mismo espacio de probabilidad.

Independencia e incorrelación de procesos aleatorios

Dos procesos aleatorios $X(t)$ e $Y(t)$ son *independientes* si para cualquier par de instantes de tiempo, t_1 y t_2 , las variables aleatorias $X(t_1)$ e $Y(t_2)$ son independientes.

De forma similar, dos procesos aleatorios $X(t)$ e $Y(t)$ están *incorrelados* si para cualquier par de instantes de tiempo, t_1 y t_2 , las variables aleatorias $X(t_1)$ e $Y(t_2)$ están incorreladas.

Al igual que para variables aleatorias, independencia implica incorrelación, pero lo inverso no es cierto: incorrelación no implica independencia en general.

Correlación cruzada

La *función de correlación cruzada* entre dos procesos aleatorios $X(t)$ e $Y(t)$ se define como

$$R_{X,Y}(t_1, t_2) = E[X(t_1) \cdot Y(t_2)].$$

En general, por la propia definición de la función de correlación cruzada, se tiene la siguiente relación entre las funciones de correlación cruzada

$$R_{X,Y}(t_1, t_2) = R_{Y,X}(t_2, t_1).$$

Estacionariedad conjunta

Dos procesos aleatorios $X(t)$ e $Y(t)$ son *conjuntamente estacionarios* (en sentido amplio) si se cumplen las dos siguientes condiciones:

- ambos son individualmente estacionarios;
- la función de correlación cruzada entre los dos procesos, $R_{X,Y}(t_1, t_2)$, depende sólo de la diferencia entre los dos instantes de tiempo, $\tau = t_1 - t_2$, y se puede por tanto denotar como

$$R_{X,Y}(t_1, t_2) = R_{X,Y}(t_1 - t_2) = R_{X,Y}(\tau).$$

Al igual que en el caso de la estacionariedad, esta condición se escribe en ocasiones utilizando la parametrización $t_1 = t + \tau$ y $t_2 = t$, de forma que

$$R_{X,Y}(t + \tau, t) = R_{X,Y}(\tau).$$

2.3.7. Procesos aleatorios en el dominio de la frecuencia

Es bien conocida la utilidad de la representación frecuencial de señales en el análisis de sistemas, ya que las relaciones entre señales en el dominio de la frecuencia son en muchos casos más sencillas que en el dominio temporal, lo que simplifica el análisis en el dominio frecuencial.

Buscando una representación frecuencial adecuada para procesos aleatorios, se podría pensar inicialmente en tratar de definir la transformada de Fourier para cada función temporal del proceso, $X(t, \lambda_i)$, y definir de nuevo el proceso a través de una función que dependa de la frecuencia ω en lugar de t , dando lugar a un conjunto de transformadas en el dominio frecuencial con una función de ω asociada a cada posible salida del experimento aleatorio, $X(j\omega, \lambda_i)$, donde

$$X(j\omega, \lambda_i) = \mathcal{TF}\{X(t, \lambda_i)\},$$

y $\mathcal{TF}\{\cdot\}$ representa la transformada de Fourier. Uno de los problemas que podría tener esta representación frecuencial del proceso es que es posible que no todas las funciones temporales que forman parte del proceso, $X(t, \lambda_i)$, tengan definida una transformada de Fourier.

En esta sección se verá cómo aplicar las técnicas de análisis en el dominio de la frecuencia para trabajar con procesos aleatorios, para llegar finalmente a una representación en el dominio de la frecuencia adecuada para los procesos aleatorios, la denominada *densidad espectral de potencia*.

Densidad espectral de potencia de procesos aleatorios

La densidad espectral de potencia de un proceso aleatorio es una extensión natural de la definición de densidad espectral de potencia para señales deterministas.

Para definir la densidad espectral de potencia se define primero la densidad espectral de potencia de cada señal del proceso, $X(t, \lambda_i)$ y posteriormente se promedia para todas las señales. Para asegurar la existencia de la transformada de Fourier, se definen las funciones truncadas

$$X^{[T]}(t, \lambda_i) = \begin{cases} X(t, \lambda_i), & |t| < T/2 \\ 0, & \text{en otro caso} \end{cases},$$

o extendiendo el uso de la notación compacta $x_i(t) \equiv X(t, \lambda_i)$,

$$x_i^{[T]}(t) = \begin{cases} x_i(t), & |t| < T/2 \\ 0, & \text{en otro caso} \end{cases}.$$

De esta forma se garantiza que las señales truncadas son señales de energía, al tener una duración limitada y ser por tanto su módulo al cuadrado integrable. Por tanto, tienen definida una transformada de Fourier, que se denotará

$$X_i^{[T]}(j\omega) = \mathcal{TF}\{x_i^{[T]}(t)\} = \int_{-\infty}^{\infty} x_i^{[T]}(t) e^{-j\omega t} dt = \int_{-T/2}^{T/2} x_i(t) e^{-j\omega t} dt.$$

Para señales de energía, la *densidad espectral de energía* era igual a $|X_i^{[T]}(j\omega)|^2$. Se puede entonces definir la *densidad espectral de potencia* como el promedio temporal de la densidad espectral de energía, es decir

$$\frac{|X_i^{[T]}(j\omega)|^2}{T}.$$

Finalmente, llevando T al límite se tiene la densidad espectral de potencia de cada señal que forma parte del proceso aleatorio

$$S_{X_i}(j\omega) = \lim_{T \rightarrow \infty} \frac{|X_i^{[T]}(j\omega)|^2}{T}.$$

Para cada frecuencia ω se tiene una variable aleatoria, pues hay un valor para cada posible valor de salida del experimento aleatorio. Parece lógico por tanto definir la densidad espectral de potencia del proceso como el promedio de estas variables aleatorias para cada frecuencia,

$$S_X(j\omega) \stackrel{def}{=} E \left[\lim_{T \rightarrow \infty} \frac{|X^{[T]}(j\omega)|^2}{T} \right] = \lim_{T \rightarrow \infty} \frac{E \left[|X^{[T]}(j\omega)|^2 \right]}{T}.$$

Esta definición permite tener una noción intuitiva del significado de la representación, es decir, valor medio del módulo al cuadrado las respuestas en frecuencia de las señales del proceso. Sin

embargo, la aplicación literal de de esta definición para el cálculo de la densidad espectral de potencia es relativamente complicado en la mayoría de los casos.

Afortunadamente, existe un teorema que relaciona la densidad espectral de potencia con la función de autocorrelación del proceso aleatorio, lo que simplifica en gran medida la obtención de estas densidades.

Teorema de Wiener-Khinchin

Si para cualquier valor finito τ y cualquier intervalo \mathcal{A} , de longitud $|\tau|$, la autocorrelación del proceso aleatorio cumple

$$\left| \int_{\mathcal{A}} R_X(t + \tau, t) dt \right| < \infty,$$

entonces la densidad espectral de potencia de $X(t)$ es la transformada de Fourier del promedio temporal de la función de autocorrelación, es decir

$$S_X(j\omega) = \mathcal{TF} \{ \langle R_X(t + \tau, t) \rangle \},$$

donde el promedio temporal de la función de autocorrelación es

$$\langle R_X(t + \tau, t) \rangle \stackrel{def}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t + \tau, t) dt.$$

La demostración, que se puede encontrar por ejemplo en [Proakis y Salehi, 2002], pag. 179, se reproduce a continuación.

Teniendo en cuenta que

$$S_X(j\omega) = \lim_{T \rightarrow \infty} \frac{E[|X_T(j\omega)|^2]}{T},$$

y que

$$X^{[T]}(j\omega) = \int_{-T/2}^{T/2} X(t) e^{-j\omega t} dt,$$

introduciendo esta expresión se tiene

$$\begin{aligned} S_X(j\omega) &= \lim_{T \rightarrow \infty} \frac{1}{T} E \left[\int_{-T/2}^{T/2} X(s) e^{-j2\pi f s} ds \int_{-T/2}^{T/2} X(t) e^{+j2\pi f t} dt \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} R_X(s, t) e^{-j2\pi f (s-t)} dt ds. \end{aligned}$$

Lo que se hace ahora es obtener la transformada de Fourier inversa, para demostrar que es precisamente $\langle R_X(t + \tau, t) \rangle$

$$\begin{aligned} \mathcal{TF}^{-1} \{ S_X(j\omega) \} &= \lim_{T \rightarrow \infty} \frac{1}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{+j\omega\tau} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} R_X(s, t) e^{-j\omega(s-t)} dt ds d\omega \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \frac{1}{2\pi} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} R_X(s, t) ds dt \int_{-\infty}^{\infty} e^{j\omega[\tau-(s-t)]} d\omega. \end{aligned}$$

Teniendo en cuenta que la transformada de Fourier inversa de una constante es una delta

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega[\tau-(s-t)]} d\omega = \delta(\tau - s + t),$$

e incluyendo este resultado en la expresión anterior, se tiene

$$\mathcal{TF}^{-1}\{S_X(j\omega)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) \delta(\tau - s + t) ds dt.$$

Teniendo en cuenta que

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(s, t) \delta(\tau - s + t) ds = \begin{cases} R_X(t + \tau, t), & -\frac{T}{2} < t + \tau < \frac{T}{2}, \\ 0, & \text{en otro caso} \end{cases},$$

entonces

$$\mathcal{TF}^{-1}\{S_X(j\omega)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \begin{cases} R_X(t + \tau, t), & -\frac{T}{2} < t + \tau < \frac{T}{2} \\ 0, & \text{en otro caso} \end{cases} dt$$

Esta expresión se puede reescribir como

$$\mathcal{TF}^{-1}\{S_X(j\omega)\} = \begin{cases} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}-\tau}^{\frac{T}{2}-\tau} R_X(t + \tau, t) dt, & \tau > 0 \\ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt, & \tau < 0 \end{cases}$$

o lo que es lo mismo

$$\mathcal{TF}^{-1}\{S_X(f)\} = \begin{cases} \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt - \int_{\frac{T}{2}-\tau}^{\frac{T}{2}} R_X(t + \tau, t) dt \right], & \tau > 0 \\ \lim_{T \rightarrow \infty} \frac{1}{T} \left[\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt - \int_{-\frac{T}{2}}^{\frac{T}{2}-\tau} R_X(t + \tau, t) dt \right], & \tau < 0 \end{cases}.$$

Por tanto, ya que la integral en un segmento de longitud τ está acotada (es la condición del enunciado del teorema), cuando $T \rightarrow \infty$ el segundo término es despreciable, y por tanto

$$S_X(j\omega) = \mathcal{TF} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t + \tau, t) dt \right\},$$

lo que concluye la demostración.

El teorema incluye además un par de corolarios que simplifican el cálculo de $S_X(j\omega)$ para el caso de que el proceso aleatorio sea estacionario o cicloestacionario.

Corolario 1 Si $X(t)$ es un proceso estacionario y $\tau R_X(\tau) < \infty$ para todo $\tau < \infty$, entonces

$$S_X(j\omega) = \mathcal{TF}\{R_X(\tau)\}.$$

La demostración es inmediata, ya que $R_X(\tau)$ sólo depende de τ . En este caso

$$\langle R_X(t + \tau, t) \rangle = R_X(\tau).$$

Corolario 2 Si $X(t)$ es cicloestacionario y se cumple que

$$\left| \int_0^{T_0} R_X(t + \tau, t) dt \right| < \infty,$$

entonces

$$S_X(j\omega) = \mathcal{TF}\{\tilde{R}_X(\tau)\},$$

donde $\tilde{R}_X(\tau)$ es el promedio en un período de la función de autocorrelación

$$\tilde{R}_X(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_X(t + \tau, t) dt.$$

T_0 es el período del proceso cicloestacionario. La demostración es inmediata, ya que al ser la autocorrelación periódica

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t + \tau, t) dt = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} R_X(t + \tau, t) dt.$$

Aparte de estos dos corolarios, del teorema de Wiener-Khinchin se pueden extraer las siguientes propiedades:

- (1) La potencia del proceso aleatorio se puede obtener integrando la densidad espectral de potencia,

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega.$$

Este valor coincide con el valor obtenido con anterioridad

$$P_X = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_X(t, t) dt.$$

Hay que tener en cuenta que la potencia se puede calcular también en el dominio temporal a partir de la función de autocorrelación.

- Para procesos estacionarios

$$P_X = R_X(0).$$

- Para procesos cicloestacionarios

$$P_X = \tilde{R}_X(0).$$

Esto es así por la definición de la transformada de Fourier. Para procesos estacionarios, la transformada de Fourier inversa de la densidad espectral de potencia es la función de autocorrelación

$$R_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) \cdot e^{j\omega\tau} d\omega,$$

que evaluada en $\tau = 0$ queda

$$R_X(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega.$$

Lo mismo se aplica en el caso de procesos cicloestacionarios a $\tilde{R}_X(0)$.

- (2) Para procesos *estacionarios y ergódicos*, la densidad espectral de potencia de cada señal es igual a la densidad espectral de potencia del proceso. En este caso, la densidad espectral de potencia de cada señal es la transformada de Fourier de la autocorrelación de dicha señal, que se define como

$$R_{x_i(t)}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)x_i(t - \tau)dt.$$

Por tanto

$$S_{x_i(t)}(f) = \mathcal{TF}[R_{x_i(t)}(\tau)].$$

Como el proceso es ergódico, el promedio estadístico coincide con el temporal y por tanto

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x_i(t)x_i(t - \tau)dt = R_X(\tau),$$

con lo cual

$$S_X(j\omega) = \mathcal{TF}[R_X(\tau)] = S_{x_i(t)}(j\omega).$$

- (3) La densidad espectral de potencia se definió originalmente como

$$\lim_{T \rightarrow \infty} \frac{E[|X^{[T]}(j\omega)|^2]}{T}.$$

Partiendo de esta definición es obvio que $S_X(j\omega)$ es una función de ω par, real y no negativa. Las dos primeras condiciones se deben a que $R_X(\tau)$ es real y par, como ya sabíamos que ocurría para procesos estacionarios. Pero además tenemos que es no negativa y esto implica que la autocorrelación es *semidefinida positiva*, lo que quiere decir que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t) R_X(t - s) g(s) dt ds \geq 0,$$

para cualquier función $g(x)$. Así pues, estas características de la densidad espectral de potencia provienen de las cuatro propiedades vistas anteriormente para la función de autocorrelación de procesos estacionarios.

2.3.8. Procesos aleatorios estacionarios y sistemas lineales

En la Sección 2.3.6 se ha visto que la salida de un sistema lineal e invariante cuya entrada es un proceso aleatorio es a su vez un proceso aleatorio, tal y como muestra la Figura 2.13

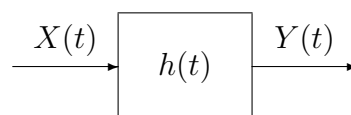


Figura 2.13: Un proceso aleatorio pasando a través de un sistema lineal e invariante.

A continuación, se van a analizar las propiedades del proceso aleatorio de salida, $Y(t)$, basado en el conocimiento que se tiene del proceso aleatorio de entrada, $X(t)$. Se supone que el proceso de entrada es estacionario y que el sistema es un sistema lineal e invariante. En particular, se trata de contestar a las siguientes preguntas:

1. ¿Bajo que condiciones es el proceso de salida estacionario?

2. ¿Bajo que condiciones son los procesos de entrada y salida conjuntamente estacionarios?
3. ¿Cómo se pueden obtener la media y la autocorrelación del proceso de salida y la correlación cruzada entre los procesos de entrada y salida?

El siguiente teorema responde a estas preguntas.

Teorema: Un proceso aleatorio, $X(t)$, es estacionario, de media m_X y función de autocorrelación $R_X(\tau)$. El proceso pasa a través de un sistema lineal e invariante con respuesta al impulso $h(t)$. En este caso, *los procesos de entrada y salida, $X(t)$ e $Y(t)$, son conjuntamente estacionarios*, siendo

$$m_Y = m_X \int_{-\infty}^{\infty} h(t) dt,$$

$$R_Y(\tau) = R_X(\tau) * h(\tau) * h(-\tau) = R_X(\tau) * r_h(\tau)$$

$$R_{X,Y}(\tau) = R_X(\tau) * h(-\tau).$$

En la expresión para la función de autocorrelación del proceso de salida, $r_h(t)$ denota la función de ambigüedad temporal de la respuesta al impulso del canal

$$r_h(t) = h(t) * h(-t).$$

Se puede ver que también se cumple la relación

$$R_Y(\tau) = R_{X,Y}(\tau) * h(\tau).$$

Demostración: Se parte de la expresión de la convolución, que relaciona la entrada y la salida de un sistema lineal

$$Y(t) = \int_{-\infty}^{\infty} X(s)h(t-s) ds$$

Por tanto

$$\begin{aligned} m_Y(t) &= E \left[\int_{-\infty}^{\infty} X(s)h(t-s) ds \right], \\ &= \int_{-\infty}^{\infty} E[X(s)]h(t-s) ds, \\ &= \int_{-\infty}^{\infty} m_X h(t-s) ds, \\ &\stackrel{u=t-s}{=} m_X \int_{-\infty}^{\infty} h(u) du. \end{aligned}$$

Por otro lado,

$$\begin{aligned}
 R_{X,Y}(t_1, t_2) &= E[X(t_1)Y(t_2)], \\
 &= E \left[X(t_1) \int_{-\infty}^{\infty} X(s)h(t_2 - s) ds \right], \\
 &= \int_{-\infty}^{\infty} E[X(t_1)X(s)]h(t_2 - s) ds, \\
 &= \int_{-\infty}^{\infty} R_X(t_1 - s)h(t_2 - s) ds, \\
 &\stackrel{u=s-t_2}{=} \int_{-\infty}^{\infty} R_X(t_1 - t_2 - u)h(-u) du, \\
 &= \int_{-\infty}^{\infty} R_X(\tau - u)h(-u) du, \\
 &= R_X(\tau) * h(-\tau).
 \end{aligned}$$

Finalmente, utilizando el resultado previo

$$\begin{aligned}
 R_Y(t_1, t_2) &= E[Y(t_1)Y(t_2)], \\
 &= E \left[\left(\int_{-\infty}^{\infty} X(s)h(t_1 - s) ds \right) Y(t_2) \right], \\
 &= \int_{-\infty}^{\infty} E[X(s)Y(t_2)]h(t_1 - s) ds, \\
 &= \int_{-\infty}^{\infty} R_{X,Y}(s - t_2)h(t_1 - s) ds, \\
 &\stackrel{u=s-t_2}{=} \int_{-\infty}^{\infty} R_{X,Y}(u)h(t_1 - t_2 - u) du, \\
 &= R_{X,Y}(\tau) * h(\tau), \\
 &= R_X(\tau) * h(-\tau) * h(\tau).
 \end{aligned}$$

Expresiones equivalentes en el dominio frecuencial

Anteriormente se ha visto lo que ocurre para los promedios estadísticos del proceso de salida de un sistema lineal e invariante cuando a su entrada hay un proceso estacionario. A continuación se analizan las relaciones entre los estadísticos del proceso de entrada y de salida en el dominio frecuencial. Para ello únicamente hay que trasladar las expresiones del dominio temporal obtenidas anteriormente al dominio frecuencial. La relación se obtiene inmediatamente si se tiene en cuenta que

$$\mathcal{TF}\{h(-\tau)\} = H^*(j\omega), \text{ y que } \int_{-\infty}^{\infty} h(t) dt = H(0).$$

La segunda expresión es evidente si se tiene en cuenta que

$$H(j\omega) = \int_{-\infty}^{\infty} h(t) \cdot e^{-j\omega t} dt \stackrel{\omega=0}{=} \int_{-\infty}^{\infty} h(t) dt = H(0).$$

Así pues, las relaciones de los estadísticos en el dominio frecuencial son

$$m_Y = m_X \cdot H(0),$$

para la media, y aplicando la transformada de Fourier a la expresión que relaciona autocorrelaciones de los procesos de entrada y de salida, la densidad espectral de potencia del proceso de salida es

$$S_Y(j\omega) = S_X(j\omega) \cdot |H(j\omega)|^2.$$

La primera ecuación dice que la media sólo se ve afectada por la respuesta en continua del sistema, es decir por la componente de $\omega = 0$, $H(0)$. Y la segunda dice que, en cuanto a la densidad espectral de potencia, la fase del sistema es irrelevante y sólo importa su módulo.

También es posible definir una relación en el dominio de la frecuencia para la correlación cruzada. Se define la *densidad espectral de potencia cruzada*, $S_{X,Y}(f)$ como

$$S_{X,Y}(j\omega) \stackrel{def}{=} \mathcal{TF}\{R_{X,Y}(\tau)\}.$$

En este caso,

$$S_{X,Y}(j\omega) = S_X(j\omega) \cdot H^*(j\omega),$$

y como $R_{Y,X}(\tau) = R_{X,Y}(-\tau)$, entonces

$$S_{Y,X}(j\omega) = S_{X,Y}^*(j\omega) = S_X(j\omega) \cdot H(j\omega).$$

Es interesante remarcar que aunque las densidades espectrales de los procesos X e Y , $S_X(j\omega)$ y $S_Y(j\omega)$, son funciones reales no negativas, las densidades cruzadas $S_{X,Y}(j\omega)$ y $S_{Y,X}(j\omega)$ pueden ser, en general, funciones complejas. La Figura 2.14 representa mediante diagramas de bloques la relación entre las distintas representaciones frecuenciales.

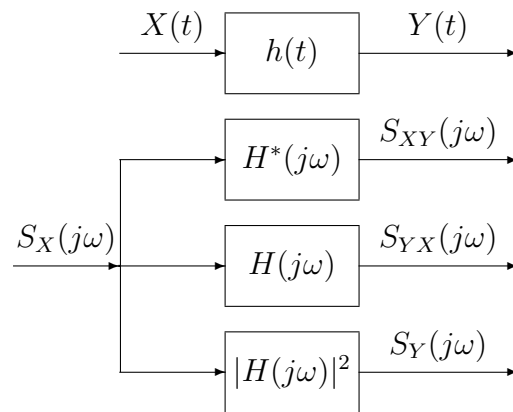


Figura 2.14: Representación esquemática de las relaciones entrada/salida para densidades espectrales de potencia y densidades espectrales cruzadas.

Extensión de los resultados anteriores a procesos cicloestacionarios

Si el proceso aleatorio de entrada a un sistema lineal e invariante es cicloestacionario, se pueden extender de forma sencilla algunos de los resultados anteriores. En particular es trivial hacer la extensión

$$m_Y(t) = m_X(t) \int_{-\infty}^{\infty} h(t) dt = m_X(t) H(0).$$

De la definición de la densidad espectral de potencia, también es trivial obtener la relación

$$S_Y(j\omega) = S_X(j\omega) |H(j\omega)|^2.$$

Dada la relación de la densidad espectral de potencia con el promedio en un período de la función de autocorrelación, también es evidente la relación

$$\tilde{R}_Y(\tau) = \tilde{R}_X(\tau) * h(\tau) * h(-\tau).$$

La relación entre funciones de autocorrelación es un poco más compleja, pero a partir de cálculos sencillos se puede llegar a la expresión

$$R_Y(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_X(s, u) h(t_1 - s) h(t_2 - u) ds du.$$

2.3.9. Suma de procesos aleatorios

En la práctica, a menudo se encuentra el problema de la suma de dos procesos aleatorios. Por ejemplo, en el caso de un sistema de comunicaciones se suma ruido sobre la señal de comunicaciones.

Supongamos que se tiene un proceso aleatorio $Z(t) = X(t) + Y(t)$, donde $X(t)$ e $Y(t)$ son conjuntamente estacionarios. Se van a obtener la media, autocorrelación y densidad espectral de potencia para el proceso $Z(t)$.

La media del proceso $Z(t)$ se obtiene de la forma siguiente

$$m_Z = E[Z(t)] = E[X(t) + Y(t)] = E[X(t)] + E[Y(t)] = m_X + m_Y.$$

La autocorrelación del proceso aleatorio $Z(t)$ es

$$\begin{aligned} R_Z(t + \tau, t) &= E[Z(t + \tau)Z(t)] \\ &= E[(X(t + \tau) + Y(t + \tau))(X(t) + Y(t))] \\ &= E[X(t + \tau)X(t)] + E[X(t + \tau)Y(t)] + E[Y(t + \tau)X(t)] + E[Y(t + \tau)Y(t)] \\ &= R_X(\tau) + R_{X,Y}(\tau) + R_{Y,X}(\tau) + R_Y(\tau). \end{aligned}$$

Como se ve, el proceso es estacionario, ya que la media es constante, y la función de autocorrelación depende únicamente de la diferencia de instantes de tiempo. Si reordenamos esta expresión tenemos

$$R_Z(\tau) = R_X(\tau) + R_Y(\tau) + R_{X,Y}(\tau) + R_{Y,X}(\tau).$$

La densidad espectral de potencia se obtiene mediante la transformada de Fourier de la función de autocorrelación. Tomando la transformada de Fourier en ambos lados de la expresión anterior

$$S_Z(j\omega) = S_X(j\omega) + S_Y(j\omega) + S_{X,Y}(j\omega) + S_{Y,X}(j\omega),$$

que teniendo en cuenta que $S_{Y,X}(j\omega) = S_{X,Y}^*(j\omega)$ lleva a

$$S_Z(j\omega) = S_X(j\omega) + S_Y(j\omega) + 2\text{Re}[S_{X,Y}(j\omega)].$$

Así pues, la densidad espectral de potencia es la suma de las densidades espectrales de potencia de cada proceso individual más un tercer término que depende de la correlación cruzada de los dos procesos.

Cuando los procesos están incorrelados, teniendo en cuenta que la relación entre la covarianza y la autocorrelación es,

$$\text{Cov}(X(t + \tau), Y(t)) = E[(X(t + \tau) - m_X) \cdot (Y(t) - m_Y)] = R_{X,Y}(\tau) - m_X \cdot m_Y$$

y que para procesos incorrelados la covarianza es nula, se cumple que

$$R_{X,Y}(\tau) = m_X m_Y.$$

En este caso, si al menos uno de los dos procesos tiene media nula, $R_{X,Y} = 0$, lo que implica que $S_{X,Y}(j\omega) = 0$ y por tanto

$$S_Z(j\omega) = S_X(j\omega) + S_Y(j\omega).$$

En el caso de la suma de la señal y ruido, lo más habitual es que el ruido esté incorrelado con la señal y tenga media nula. En este caso se puede aplicar esta relación.

2.4. Modelo del ruido térmico: procesos blancos y gaussianos

Un tipo de procesos que juegan un papel muy importante en los sistemas de comunicaciones son los procesos gaussianos y los procesos blancos. Los procesos gaussianos son importantes por dos razones:

1. El ruido térmico, producido por el movimiento aleatorio de los electrones debido a la agitación térmica, presenta una distribución gaussiana. Este tipo de ruido está presente en cualquier dispositivo electrónico y es el más importante en muchos sistemas de comunicaciones.

La explicación de porqué el ruido térmico es gaussiano se puede encontrar en el teorema central del límite. El ruido térmico se debe al movimiento aleatorio de los electrones y la corriente es la suma de múltiples electrones. Si se supone que cada electrón se comporta de forma independiente se tiene la suma de un número de variables aleatorias *i.i.d.*, con lo que al final se llega a que su distribución es gaussiana, $\mathcal{N}(m_X, \frac{\sigma_X^2}{n})$.

2. Los procesos gaussianos proporcionan un buen modelo para algunas fuentes de información, por lo que los procesos gaussianos hacen posible su análisis.

Por tanto, el análisis de procesos gaussianos nos va a permitir analizar el efecto del ruido térmico y las características de algunas fuentes de información.

En cuanto a los procesos blancos, también van a ser importantes en el modelado del ruido en un sistema de comunicaciones, ya que como se verá más tarde las características espectrales del ruido térmico son muy similares a las de un proceso blanco.

A continuación se define lo que son procesos gaussianos y lo que son procesos blancos.

2.4.1. Procesos aleatorios gaussianos

Un proceso aleatorio $X(t)$ es un *proceso gaussiano* si para todo conjunto de n instantes de tiempo $\{t_1, t_2, \dots, t_n\}$ y para cualquier valor de n , las n variables aleatorias resultantes de evaluar el proceso en esos n instantes de tiempo $\{X(t_i)\}_{i=1}^n$, tienen una distribución conjuntamente gaussiana.

Esto implica que para cualquier instante de tiempo t_0 la variable aleatoria $X(t_0)$ es gaussiana y en cada par de puntos t_1 y t_2 las variables aleatorias $(X(t_1), X(t_2))$ tienen una distribución conjuntamente gaussiana.

Una de las ventajas de los procesos conjuntamente gaussianos es que su descripción estadística completa depende sólo del vector de medias $\boldsymbol{\mu}$ y de la matriz de covarianza \mathbf{C} . Debido a esto se puede formular el siguiente teorema.

Teorema: Para procesos gaussianos, el conocimiento de la media, $m_X(t)$, y de la función de autocorrelación, $R_X(t_1, t_2)$, proporciona una descripción estadística completa del proceso.

Este teorema implica que no es necesario conocer el vector $\boldsymbol{\mu}$ o la matriz \mathbf{C} para todo n y todo conjunto de tiempos $\{t_i\}_{i=1}^n$, sino que es suficiente conocer $m_X(t)$ y $R_X(t_1, t_2)$. Siempre es posible construir $\boldsymbol{\mu}$ y \mathbf{C} a partir de estas para cualquier conjunto de instantes de tiempo $\{t_i\}_{i=1}^n$.

- Para $X(t_i)$, $\Rightarrow \mu_i = m_X(t_i)$.
- Para $(X(t_i), X(t_j))$, $\Rightarrow C_{i,j} = \text{Cov}(X(t_i), X(t_j)) = R_X(t_i, t_j) - m_X(t_i)m_X(t_j)$.

Otra de las ventajas de los procesos gaussianos es su comportamiento en sistemas lineales. El siguiente teorema describe este comportamiento que es de gran importancia.

Teorema: Si un proceso gaussiano $X(t)$ atraviesa un sistema lineal e invariante (LTI), el sistema resultante, $Y(t)$, es también un proceso gaussiano.

Para probarlo, utilizamos una de las propiedades que vimos para procesos conjuntamente gaussianos. Se trata de demostrar que para todo n , el vector $(X(t_1), X(t_2), \dots, X(t_n))$ es conjuntamente gaussiano. En general, para un instante de tiempo t_i

$$Y(t_i) = \int_{-\infty}^{\infty} X(\tau)h(t_i - \tau)d\tau.$$

Esta integral se puede interpretar como un sumatorio llevado al límite, donde $X(t)$ es multiplicado por los distintos valores de la respuesta al impulso $h(t)$. En concreto, esta integral es igual a

$$Y(t_i) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_i - k\Delta).$$

Esta expresión se puede interpretar como una combinación lineal de un conjunto de variables aleatorias conjuntamente gaussianas, $\{X(k\Delta)\}_{k=-N}^N$. Se puede interpretar entonces que se tiene un

sistema

$$\left\{ \begin{array}{l} Y(t_1) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_1 - k\Delta) \\ Y(t_2) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_2 - k\Delta) \\ \vdots \\ Y(t_n) = \lim_{N \rightarrow \infty} \lim_{\Delta \rightarrow 0} \sum_{k=-N}^N X(k\Delta)h(t_n - k\Delta) \end{array} \right. .$$

Esta expresión plantea la combinación lineal de las variables aleatorias $\{X(k\Delta)\}_{k=-N}^N$, que forman conjuntamente gaussianas. Y se ha visto en la Sección 2.2.6 que una de las propiedades de las variables aleatorias conjuntamente gaussianas es que cualquier combinación lineal de las mismas forma un conjunto de variables aleatorias conjuntamente gaussianas.

Esta propiedad de los procesos gaussianos es muy importante ya que significa que se conoce el tipo de proceso que hay en la salida de un sistema cuando la entrada es gaussiana. Para procesos en general, puede ser muy difícil saber esto.

Todas las definiciones hechas hasta ahora son para procesos gaussianos en general. A continuación se plantean algunas condiciones para procesos estacionarios.

Teorema: Para procesos gaussianos, estacionariedad en sentido estricto y en sentido amplio son equivalentes.

Esta propiedad se debe a que los procesos gaussianos tienen una descripción estadística completa que sólo depende de la media $m_X(t)$ y de la función de autocorrelación $R_X(t_1, t_2)$.

Teorema: Para procesos gaussianos, estacionarios y de media nula, una condición suficiente para la ergodicidad del proceso $X(t)$ es

$$\int_{-\infty}^{\infty} |R_X(\tau)|d\tau < \infty.$$

Esto es algo que simplifica el análisis de ergodicidad de este tipo de procesos.

Procesos aleatorios conjuntamente gaussianos

Los procesos $X(t)$ e $Y(t)$ son *conjuntamente gaussianos*, si para cualquier valor de n y m , y cualquier par de conjuntos de instantes de tiempo $\{t_1, t_2, \dots, t_n\}$ y $\{\tau_1, \tau_2, \dots, \tau_m\}$, las $n + m$ variables aleatorias

$$\{X(t_1), X(t_2), \dots, X(t_n), Y(\tau_1), Y(\tau_2), \dots, Y(\tau_m)\},$$

tienen una distribución conjuntamente gaussiana (de dimensión $n + m$).

Los procesos aleatorios conjuntamente gaussianos tienen una propiedad muy importante que se expone en el siguiente teorema.

Teorema: Para procesos conjuntamente gaussianos, incorrelación e independencia son equivalentes.

2.4.2. Procesos aleatorios blancos

El término *blanco* asociado a un proceso aleatorio se utiliza para describir a aquellos procesos que tienen igual potencia para todas las frecuencias. El término hace referencia a la analogía con el caso de la luz blanca, que está compuesta por la suma de todos los colores.

A modo de definición, un proceso blanco tiene una densidad espectral de potencia plana, es decir, $S_X(j\omega)$ es constante para todas las frecuencias

$$S_X(j\omega) = C.$$

Por tanto, la función de autocorrelación para un proceso blanco es

$$R_X(\tau) = \mathcal{TF}^{-1}\{C\} = C \cdot \delta(\tau).$$

Esto significa que para cualquier $\tau \neq 0$ $R_X(\tau) = 0$. Y esto implica que las variables aleatorias $X(t_1)$ y $X(t_2)$, $\forall t_1 \neq t_2$, están incorreladas. Si además de blanco el proceso es gaussiano, que como se verá luego es el caso del modelo estadístico habitual para el ruido térmico, esto significa que las variables aleatorias son también independientes.

A partir de la definición de un proceso aleatorio blanco, su potencia es infinita, ya que

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} C d\omega = \infty$$

o si se calcula en el dominio temporal

$$P_X = R_X(0) = C \cdot \delta(0) = \infty.$$

Filtrado de un proceso blanco

Se ha visto que cuando un proceso gaussiano atraviesa un sistema lineal e invariante, el proceso resultante sigue siendo gaussiano. Sin embargo, no sucede lo mismo con la condición de “blanco”. Cuando un proceso blanco pasa por un sistema lineal e invariante, deja de ser blanco. En concreto, la densidad espectral de potencia queda determinada por la respuesta en frecuencia del filtro

$$S_Y(j\omega) = S_X(j\omega) \cdot |H(j\omega)|^2 = C \cdot |H(j\omega)|^2.$$

Salvo en el caso de un filtro trivial paso todo ($h(t) = \alpha\delta(t)$, ó $H(j\omega) = \alpha$), esta respuesta no será constante, por lo que el proceso $Y(t)$ no será blanco.

La función de autocorrelación del proceso filtrado es

$$R_Y(\tau) = R_X(\tau) * h(\tau) * h(-\tau) = R_X(\tau) * r_h(\tau),$$

donde $r_h(\tau)$ es la función de ambigüedad temporal de $h(\tau)$. Teniendo en cuenta la forma de $R_X(\tau)$ para un proceso blanco,

$$R_Y(\tau) = C \cdot r_h(\tau).$$

Esto significa que la potencia del proceso es

$$P_Y = R_Y(0) = C \cdot r_h(0),$$

y hay que tener en cuenta que el valor de la función de ambigüedad temporal de una cierta función evaluada en cero proporciona la energía de dicha función, por lo que

$$P_Y = C \cdot \mathcal{E}\{h(t)\}.$$

Por tanto, la potencia de un proceso blanco filtrado ya no es infinita, sino que está relacionada con la energía de la respuesta el impulso del filtro.

2.4.3. Modelo de ruido térmico

El modelo habitual para el ruido térmico es el de un proceso aleatorio estacionario, ergódico, blanco y gaussiano. Generalmente el ruido se denotará como $n(t)$, por lo que se utilizará la notación $R_n(\tau)$ y $S_n(j\omega)$ para representar la función de autocorrelación y la densidad espectral de potencia del proceso aleatorio que lo modela. En el caso de ruido térmico la constante C que define el valor de la densidad espectral de potencia se denota habitualmente como $N_0/2$, tal y como muestra la Figura 2.4.3.

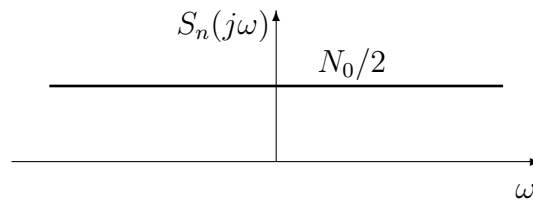


Figura 2.15: Densidad espectral de potencia de un proceso blanco que modela ruido térmico.

Obviamente, ningún proceso físico puede tener potencia infinita. Por tanto, un proceso blanco no existe como tal. La importancia de los procesos blancos en la práctica se debe a que el ruido térmico se puede modelar como un proceso blanco para un gran rango de frecuencias. El análisis mecánico cuántico del ruido térmico dice que la densidad espectral de potencia del ruido blanco es

$$S_n(j\omega) = \frac{h\omega}{4\pi(e^{\frac{h\omega}{2\pi kT}} - 1)},$$

donde $\left\{ \begin{array}{l} h: \text{Constante de Planck } (6.6 \times 10^{-34} \text{ Julios} \times \text{segundo}). \\ k: \text{Constante de Boltzmann } (1.38 \times 10^{-23} \text{ Julios}/^\circ\text{Kelvin}). \\ T: \text{Temperatura en grados Kelvin.} \\ \omega: \text{Frecuencia angular o pulsación en radianes/s } (2\pi \text{ veces la frecuencia}). \end{array} \right.$

Esta densidad espectral de potencia tiene su máximo en $\omega = 0$, donde toma el valor $kT/2$, y tiende a cero cuando f tiende a infinito. Sin embargo, la tendencia es muy lenta. Por ejemplo, a temperatura ambiente, $T = 290$ °K, $S_n(j\omega)$ cae al 90 % de su máximo aproximadamente para $\omega \approx 2\pi \cdot (2 \times 10^{12})$ rad/s, lo que está por encima de las frecuencias comúnmente utilizadas por los sistemas de comunicaciones. La Figura 2.4.3 representa $S_n(j\omega)$.

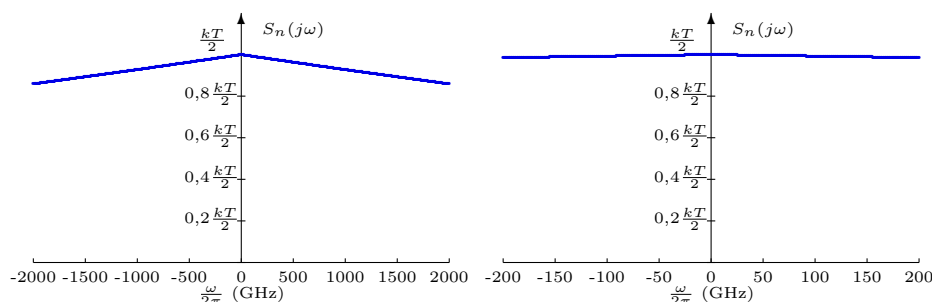


Figura 2.16: Densidad espectral de potencia del ruido térmico

Por esta razón, el ruido térmico se modela como un proceso blanco para el cual $C = \frac{N_0}{2}$, con $N_0 = kT$.

A partir de ahora el modelo para el ruido término es un proceso aleatorio con las siguientes características:

- Estacionario.
- Ergódico.
- Media $m_n = 0$.
- Función de autocorrelación

$$R_n(\tau) = \frac{N_0}{2} \cdot \delta(\tau).$$

- Densidad espectral de potencia

$$S_n(j\omega) = \frac{N_0}{2}.$$

NOTA: es fácil ver que un proceso estacionario gaussiano con esa función de autocorrelación es ergódico, ya que

$$\int_{-\infty}^{\infty} |R_n(\tau)| d\tau = \frac{N_0}{2}.$$

2.4.4. Ruido filtrado y ancho de banda equivalente de ruido

Como se ha visto anteriormente, la potencia de un proceso aleatorio blanco es infinita. En el caso del ruido térmico, aunque no es infinita, es relativamente alta. Para limitarla, lo que se hace habitualmente es filtrarlo. Si se denota como $Z(t)$ el proceso aleatorio resultante al filtrar el proceso de ruido térmico con un filtro lineal e invariante de respuesta al impulso $h(t)$, la densidad espectral de potencia de este proceso es

$$S_Z(j\omega) = S_n(j\omega) |H(j\omega)|^2 = \frac{N_0}{2} |H(j\omega)|^2.$$

La potencia del proceso $Z(t)$ obtenido filtrando el proceso aleatorio que modela el ruido térmico con un filtro de respuesta al impulso $h(t)$ se puede obtener integrando $S_Z(j\omega)$

$$P_Z = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_Z(j\omega) d\omega = \frac{N_0}{2} \cdot \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega}_{\mathcal{E}\{h(t)\}} = \frac{N_0}{2} \cdot \mathcal{E}\{h(t)\}.$$

Si los filtros utilizados son filtros ideales de ancho de banda B Hz (o $W = 2\pi B$ rad/s), bien sean filtros paso bajo o filtros paso banda con frecuencia central f_c Hz (o $\omega_c = 2\pi f_c$ rad/s), sus respuestas en frecuencia son las de la Figura 2.17

Tanto para filtros paso bajo como para filtros ideales, es fácil comprobar que

$$\mathcal{E}\{h(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega = 2B.$$

Por tanto la potencia del ruido filtrado $Z(t)$ es

$$P_Z = N_0 \cdot B.$$

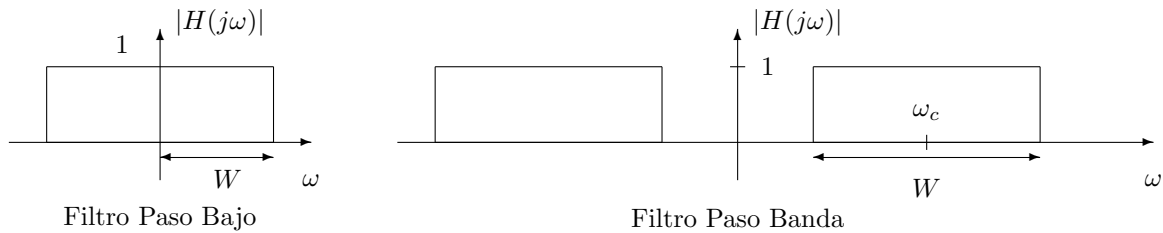


Figura 2.17: Respuesta en frecuencia de filtros ideales, paso bajo y paso banda.

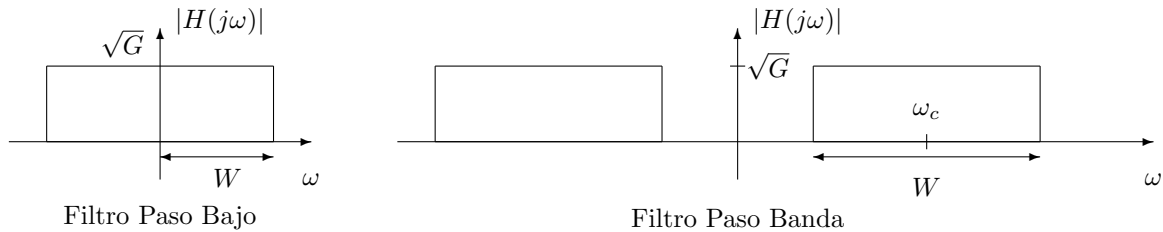


Figura 2.18: Respuesta en frecuencia de filtros ideales, paso bajo y paso banda, con una ganancia en potencia G .

Si los filtros ideales no tienen ganancia unidad, sino ganancia en potencia G , su respuesta en frecuencia es la de la de la Figura 2.18

En ese caso, la energía de los filtros es

$$\mathcal{E}\{h(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega = 2BG,$$

y por tanto la potencia del ruido filtrado es

$$P_Z = N_0 \cdot B \cdot G.$$

Tanto con ganancia unidad, como con una ganancia en potencia arbitraria G , la potencia del ruido filtrado se calcula de forma muy sencilla a partir de las expresiones anteriores.

Cuando se utilizan filtros no ideales, en muchos casos no resulta práctico para los usuarios de un cierto sistema tener que evaluar la integral de la respuesta en frecuencia de un filtro al cuadrado. Sería mucho más conveniente poder aplicar una expresión similar a la que se tiene para los filtros ideales. Por esta razón se utiliza el denominado *ancho de banda equivalente de ruido* de un sistema. El ancho de banda equivalente de ruido, que se denota como B_{eq} , sirve para obtener la potencia del ruido como

$$P_Z = N_0 \cdot B_{eq} \cdot G,$$

donde $G = H_{max}^2$, siendo H_{max} es el máximo de $H(j\omega)$. Por tanto, G denota la ganancia máxima del filtro. Teniendo el ancho de banda equivalente de ruido del filtro, B_{eq} , y su ganancia G , resulta muy sencillo calcular la potencia de ruido a la salida del sistema. Habitualmente, los fabricantes miden estos valores y los publican en las hojas de características para que pueda ser utilizado por los usuarios.

Comparando las expresiones de la potencia del ruido filtrado a partir de B_{eq} y G , e integrando la respuesta al cuadrado del filtro, es evidente que el ancho de banda equivalente de ruido se define

como

$$B_{eq} = \frac{\mathcal{E}\{h(t)\}}{2 \cdot G} = \frac{\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega}{2 \cdot G}.$$

Teniendo en cuenta que la respuesta en frecuencia de sistemas reales es simétrica respecto al origen, se tiene que

$$\int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega = 2 \int_0^{\infty} |H(j\omega)|^2 d\omega,$$

y el ancho de banda equivalente de ruido se puede calcular también como

$$B_{eq} = \frac{\frac{1}{2\pi} \int_0^{\infty} |H(j\omega)|^2 d\omega}{G}.$$

La interpretación del ancho de banda equivalente de ruido sería la del filtro paso bajo ideal, de amplitud H_{max} y frecuencia de corte B_{eq} , que permite pasar el mismo ruido, en términos de potencia, que el sistema caracterizado. La Figura 2.4.4 ilustra esta interpretación.

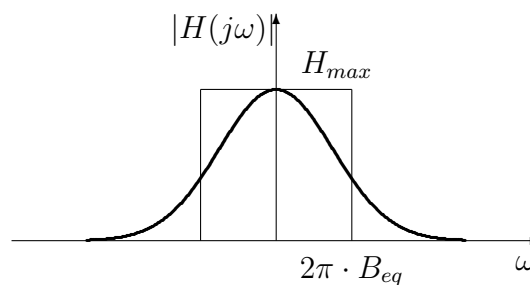


Figura 2.19: Ilustración del significado del ancho de banda equivalente de ruido de un sistema.

Ejemplo

Calcular el ancho de banda equivalente de ruido de un filtro con la siguiente respuesta en frecuencia

$$|H(j\omega)| = \begin{cases} \sqrt{1 + \frac{\omega}{W}}, & \text{si } -W \leq \omega < 0 \\ \sqrt{1 - \frac{\omega}{W}}, & \text{si } 0 \leq \omega \leq W \\ 0, & \text{en otro caso,} \end{cases}$$

donde W es el ancho de banda en rad/s, i.e., $W = 2\pi B$, siendo B el ancho de banda en Hz.

En esta caso, el módulo al cuadrado de la respuesta en frecuencia del filtro es un triángulo entre $-W$ y W ,

$$|H(j\omega)|^2 = \Lambda\left(\frac{\omega}{2W}\right) = \begin{cases} 1 + \frac{\omega}{W}, & \text{si } -W \leq \omega < 0 \\ 1 - \frac{\omega}{W}, & \text{si } 0 \leq \omega \leq W \\ 0, & \text{en otro caso.} \end{cases}$$

Por tanto, se ve claramente que $H_{max} = \max |H(j\omega)|^2 = 1$, valor que se obtiene para $\omega = 0$. Por otro lado, teniendo en cuenta la simetría de la respuesta

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega &= \frac{1}{\pi} \int_0^{\infty} \frac{1}{1 + \omega^2\tau^2} d\omega \\ &= \frac{1}{\pi} \int_0^W 1 - \frac{\omega}{W} d\omega \\ &= \frac{1}{\pi} \times \frac{W}{2} = B. \end{aligned}$$

Por tanto,

$$B_{eq} = \frac{B}{2 \times 1} = \frac{B}{2}.$$

Ejemplo

Calcular el ancho de banda equivalente de ruido de un filtro RC paso bajo.

La respuesta en frecuencia de un filtro RC es

$$H(j\omega) = \frac{1}{1 + j\omega\tau},$$

donde la constante τ vale $\tau = RC$. El módulo de esta respuesta es

$$|H(j\omega)| = \frac{1}{\sqrt{1 + \omega^2\tau^2}}.$$

En este caso es claro que $H_{max} = 1$, para $\omega = 0$. Por otro lado,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 d\omega &= \frac{1}{\pi} \int_0^{\infty} \frac{1}{1 + \omega^2\tau^2} d\omega \\ &\stackrel{u=\omega\tau}{=} \frac{1}{\pi} \int_0^{\infty} \frac{1}{1 + u^2} \frac{du}{\tau} \\ &= \frac{1}{\pi\tau} \underbrace{\int_0^{\infty} \frac{1}{1 + u^2} du}_{\frac{\pi}{2}} = \frac{1}{\pi\tau} \operatorname{arctg}(u) \Big|_0^{\infty} \\ &= \frac{1}{2\tau}. \end{aligned}$$

Por tanto,

$$B_{eq} = \frac{\frac{1}{2\tau}}{2 \times 1} = \frac{1}{4\tau} = \frac{1}{4RC}.$$

2.4.5. Relación señal a ruido

Ya se ha visto como calcular la potencia de los procesos. Un caso particular es el caso en el que se tiene la señal más una componente de ruido y ambas señales se modelan con distintos procesos. En este caso es útil calcular la relación señal a ruido

$$\frac{S}{N} = \frac{\text{Potencia de la señal}}{\text{Potencia del ruido}}.$$

Si se modela la señal con un proceso $X(t)$, y el ruido filtrado con un proceso $Z(t)$, la relación señal a ruido será

$$\frac{S}{N} = \frac{P_X}{P_Z}.$$

La potencia de la señal se puede calcular integrando su densidad espectral de potencia

$$P_X = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(j\omega) d\omega.$$

Para el ruido, si se asume el modelo de proceso blanco y gaussiano, y se tiene en cuenta que para una señal de ancho de banda B , que generalmente se filtrará con un filtro de ese ancho de banda, sólo afecta el ruido en esa banda. Si se utilizan filtros ideales con ganancia unidad

$$P_Z = N_0 \cdot B = K \cdot T \cdot B.$$

Por tanto

$$\frac{S}{N} = \frac{P_X}{N_0 \cdot B}.$$

Si se utilizan filtros ideales con ganancia en potencia G , la potencia de ruido filtrado es

$$P_Z = N_0 \cdot B \cdot G = K \cdot T \cdot B \cdot G.$$

Por tanto

$$\frac{S}{N} = \frac{P_X}{N_0 \cdot B \cdot G}.$$

Si se tiene el ancho de banda equivalente de ruido y la ganancia del filtro receptor, la potencia de ruido es

$$P_Z = N_0 \cdot B_{eq} \cdot G,$$

y la relación señal a ruido

$$\frac{S}{N} = \frac{P_X}{N_0 \cdot B_{eq} \cdot G}.$$

Finalmente, si se conoce la energía de la respuesta al impulso del filtro, la potencia de ruido sería

$$P_Z = \frac{N_0}{2} \cdot \mathcal{E}\{h(t)\},$$

y la relación señal a ruido

$$\frac{S}{N} = \frac{2P_X}{N_0 \cdot \mathcal{E}\{h(t)\}}.$$

2.5. Muestreo de procesos aleatorios limitados en banda

La definición de un proceso aleatorio limitado en banda es la extensión natural de la definición de una señal limitada en banda. Un proceso aleatorio de ancho de banda B tiene la propiedad

$$S_X(j\omega) = 0, \quad \forall |\omega| \geq 2\pi \cdot B.$$

Como ya se ha visto, para señales limitadas en banda existe el teorema del muestreo que permite muestrear la señal en tiempo discreto sin perder ninguna información de la misma. Este teorema dice que para poder reconstruir de forma perfecta la señal original, la frecuencia de muestreo debe ser al menos dos veces la máxima frecuencia de la señal, es decir,

$$f_m \geq 2B \rightarrow T_m = \frac{1}{f_m} \leq \frac{1}{2B}.$$

En este caso a señal se reconstruye como

$$x(t) = 2BT_m \sum_{k=-\infty}^{\infty} x(kT_m) \text{sinc}(2B(t - kT_m)).$$

Para $T_m = \frac{1}{2B}$, esta expresión se simplifica

$$x(t) = \sum_{k=-\infty}^{\infty} x(kT_m) \text{sinc}(2B(t - kT_m)).$$

Parece lógico pensar que el teorema del muestreo se puede extender para procesos aleatorios. El siguiente teorema justifica esta intuición.

Teorema: Si $X(t)$ es un proceso limitado en banda, con $S_X(j\omega) = 0$ para $\omega \geq W = 2\pi B$, tomando un intervalo de muestreo $T_m = \frac{1}{2B} = \frac{\pi}{W}$

$$E \left[\left(X(t) - \sum_{k=-\infty}^{\infty} X(kT_m) \text{sinc}(2B(t - kT_m)) \right)^2 \right] = 0.$$

Para demostrarlo se expande esta expresión

$$\begin{aligned} & E \left[\left(X(t) - \sum_{k=-\infty}^{\infty} X(kT_m) \text{sinc}(2B(t - kT_m)) \right)^2 \right] \\ &= E[X^2(t)] - 2 \sum_{k=-\infty}^{\infty} E[X(t)X(kT_m)] \text{sinc}(2B(t - kT_m)) \\ &\quad + \sum_{k=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} E[X(kT_m)X(uT_m)] \text{sinc}(2B(t - kT_m))\text{sinc}(2B(t - uT_m)) \\ &= R_X(0) - 2 \sum_{k=-\infty}^{\infty} R_X(t - kT_m) \text{sinc}(2B(t - kT_m)) \\ &\quad + \sum_{k=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} R_X((k - u)T_m) \text{sinc}(2B(t - kT_m))\text{sinc}(2B(t - uT_m)). \end{aligned}$$

Para el último término, se puede hacer el cambio de variable $m = u - k$ y queda

$$\begin{aligned} & \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} R_X(mT_m) \text{sinc}(2B(t - kT_m))\text{sinc}(2B(t - kT_m - mT_m)) \\ &= \sum_{k=-\infty}^{\infty} \text{sinc}(2B(t - kT_m)) \sum_{m=-\infty}^{\infty} R_X(mT_m) \text{sinc}(2B(t - kT_m - mT_m)), \end{aligned}$$

donde se ha utilizado la propiedad $R_X(-mT_m) = R_X(mT_m)$.

Como $X(t)$ es limitado en banda, su autocorrelación también lo es, de modo que

$$R_X(t) = \sum_{k=-\infty}^{\infty} R_X(kT_m) \text{sinc}(2B(t - kT_m)),$$

y por tanto

$$\sum_{m=-\infty}^{\infty} R_X(mT_m) \operatorname{sinc}(2B(t - kT_m - mT_m)) = R_X(t - kT_m).$$

Sustituyendo esta expresión se llega a

$$\begin{aligned} E \left[\left(X(t) - \sum_{k=-\infty}^{\infty} X(kT_m) \operatorname{sinc}(2B(t - kT_m)) \right)^2 \right] \\ = R_X(0) - \sum_{k=-\infty}^{\infty} R_X(t - kT_m) \operatorname{sinc}(2B(t - kT_m)). \end{aligned}$$

Y se puede comprobar que el último término es $R_X(0)$, lo que completa la demostración.

Como es la esperanza del error cuadrático lo que se anula, en este caso se dice que el teorema del muestreo se cumple en sentido cuadrático medio, o que $X(t)$ es igual en sentido cuadrático medio a la expresión del teorema del muestreo para $X(kT_m)$

$$X(t) \stackrel{q.m.}{=} \sum_{k=-\infty}^{\infty} X(kT_m) \operatorname{sinc}(2B(t - kT_m)).$$

Otra propiedad interesante es que las muestras del proceso aleatorio sólo están incorreladas si el proceso tiene una densidad espectral de potencia plana en la banda, es decir, si

$$S_X(j\omega) = \begin{cases} K, & |\omega| < W \\ 0, & \text{en otro caso} \end{cases} .$$

2.6. Ejercicios

Ejercicio 2.1 Se tiene una variable aleatoria X .

- a) Si X es una variable aleatoria con una distribución uniforme en el intervalo $[-2, 2]$, calcule las probabilidades
- i) $P(X > 1)$
 - ii) $P(X > -1)$
 - iii) $P(X < -1)$
 - iv) $P(-1 \leq X \leq 1)$
- b) Repita el apartado anterior si X tiene una distribución gaussiana de media 1 y varianza 4.

Ejercicio 2.2 Dos variables aleatorias, X e Y , tienen la función de distribución $F_X(x)$ y la función de probabilidad $f_Y(y)$, respectivamente, que se representan en la Figura 2.20.

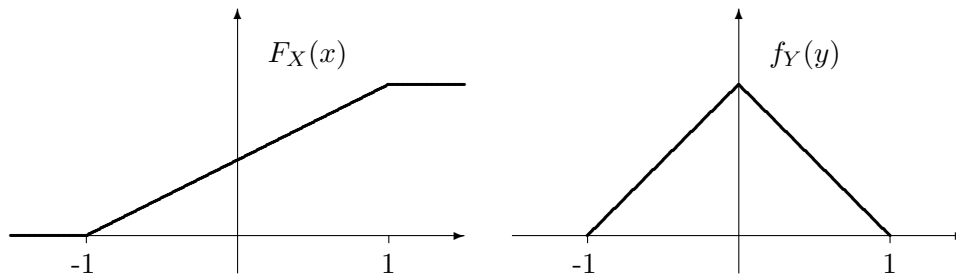


Figura 2.20: Función de distribución $F_X(x)$ y función de probabilidad $f_Y(y)$ del Ejercicio 2.2.

- a) Escriba las expresiones analíticas de las funciones de distribución y de densidad de probabilidad para cada una de las variables aleatorias ($F_X(x), f_X(x), F_Y(y), f_Y(y)$).
- b) Calcule la varianza de las dos variables aleatorias.
- c) Calcule las siguientes probabilidades sobre la variable aleatoria Y : $P(Y > 0)$, $P(Y > -\frac{1}{2})$, $P(Y < -\frac{1}{2})$, $P(Y > \frac{1}{4})$.

Ejercicio 2.3 La variable aleatoria Θ se utiliza para modelar una fase aleatoria. Calcule la función densidad de probabilidad, la media y la varianza de la variable aleatoria

$$X = \tan(\Theta),$$

en los siguientes casos:

- a) La variable aleatoria Θ está uniformemente distribuida en el intervalo $[-\frac{\pi}{2}, \frac{\pi}{2}]$.
- b) La variable aleatoria Θ está uniformemente distribuida en el intervalo $[-\frac{\pi}{4}, \frac{\pi}{4}]$.

Ejercicio 2.4 Un proceso aleatorio $X(t)$ se define como

$$X(t) = A + B \cdot t,$$

donde A y B son dos variables aleatorias independientes, la primera con una función de densidad de probabilidad $f_A(a)$ como la mostrada en la Figura 2.21, y la segunda con una función densidad de probabilidad uniforme en el intervalo $[-1, +1]$, i.e. $f_B(b) = \mathcal{U}(-1, +1)$.

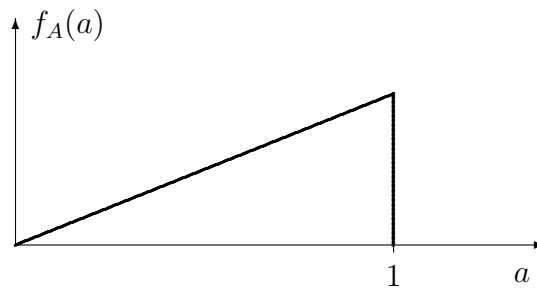


Figura 2.21: Función densidad de probabilidad de la variable aleatoria A .

- a) Calcule la media, $m_X(t)$, del proceso $X(t)$.
- b) Calcule la función de autocorrelación, $R_X(t + \tau, t)$, del proceso $X(t)$.
- c) ¿Es este proceso estacionario en sentido amplio?

Ejercicio 2.5 Se describe un proceso $X(t)$ que para $t \geq 0$ tiene la propiedad de que para todo n y todo conjunto de n instantes de tiempo, $(t_1, t_2, \dots, t_n) \in \mathbb{R}^n$, la función densidad de probabilidad conjunta de $\{X(t_i)\}_{i=1}^n$ es una probabilidad conjuntamente gaussiana de media nula y matriz de covarianza dada por

$$C_{i,j} = \text{Cov}(X(t_i), X(t_j)) = \sigma^2 \min(t_i, t_j)$$

¿Es este proceso estacionario en sentido amplio?

Ejercicio 2.6 Sea $X[n]$ un proceso aleatorio en tiempo discreto, estacionario, de media m_X , y función de autocorrelación $R_X[k]$. Considere el proceso

$$Y[n] = X[n] + a \cdot X[n - 1],$$

donde a es una constante.

- a) Calcule la media, $m_Y[n]$, del proceso $Y[n]$.
- b) Calcule la función de autocorrelación, $R_Y[n + k, n]$, del proceso $Y[n]$.
- c) Obtenga la densidad espectral de potencia, $S_Y(e^{j\omega})$, del proceso $Y[n]$.

Ejercicio 2.7 El proceso aleatorio $Z(t)$ se define como

$$Z(t) = X \cdot \cos(\omega_0 t) + Y \cdot \text{sen}(\omega_0 t),$$

donde X e Y son variables aleatorias gaussianas independientes, de media cero y varianzas σ_X^2 y σ_Y^2 .

- a) Calcule la media $m_Z(t)$.
- b) Calcule la función de autocorrelación $R_Z(t + \tau, t)$.
- c) ¿Es $Z(t)$ estacionario o cicloestacionario?
- d) Calcule la densidad espectral de potencia $S_Z(j\omega)$.
- e) Repita las cuestiones anteriores para $\sigma_X = \sigma_Y$.

Ejercicio 2.8 El proceso aleatorio $Z(t)$ tiene la siguiente descripción analítica

$$Z(t) = X \cdot \cos(\omega_0 t) + Y \cdot \text{sen}(\omega_0 t),$$

donde X es una variable aleatoria uniforme en el intervalo $(-1,1)$, e Y es una variable aleatoria uniforme en el intervalo $(0,1)$. Además, X e Y son independientes.

- Calcule la media $m_Z(t)$.
- Calcule la función de autocorrelación $R_Z(t + \tau, t)$.
- ¿Es $Z(t)$ estacionario o cicloestacionario?
- Calcule la densidad espectral de potencia $S_Z(j\omega)$.

Ejercicio 2.9 Las modulaciones digitales de modulación de pulsos en amplitud (PAM, del inglés *pulse amplitude modulation*), se modelan mediante un proceso aleatorio con la siguiente descripción analítica

$$X(t) = \sum_{n=-\infty}^{\infty} A[n] \cdot g(t - nT).$$

En esta expresión T es una constante que define el tiempo de símbolo, $g(t)$ es una señal determinista que define el pulso conformador del transmisor, con una transformada de Fourier $P(j\omega)$, y $A[n]$ es un proceso aleatorio en tiempo discreto que modela la secuencia de símbolos transmitidos. Este proceso es estacionario, y se conocen su media, y su función de autocorrelación (y por tanto, también su densidad espectral de potencia)

$$m_A = E[A[n]], R_A[k] = E[A[n+k] \cdot A[n]], S_A(e^{j\omega}) = \mathcal{TF}\{R_A[k]\}.$$

- Calcule la media del proceso aleatorio $X(t)$, $m_X(t)$.
- Calcule la función de autocorrelación del proceso aleatorio $X(t)$, $R_X(t + \tau, t)$.
- Demuestre si el proceso es estacionario o cicloestacionario.
- Calcule la densidad espectral de potencia del proceso aleatorio $X(t)$, $S_X(j\omega)$.

Ejercicio 2.10 El proceso de reconstrucción de una señal a partir de sus muestras definido en el Teorema de Nyquist para el muestreo se puede modelar a través del siguiente proceso aleatorio

$$Y(t) = \sum_{n=-\infty}^{\infty} X[n] \cdot \text{sinc}(2B(t - nT))$$

donde para el proceso que modela las muestras, $X[n]$, se va a asumir que cada una de las variables aleatorias que lo forman son variables aleatorias independientes de media zero y varianza σ_X^2 . T y B son dos constantes que modelan en este caso el período de muestreo y el ancho de banda, respectivamente. Tenga en cuenta que la expresión analítica de este proceso es equivalente a la del Ejercicio 2.9, por lo que si lo ha resuelto puede utilizar los resultados obtenidos como ayuda para la resolución de este ejercicio.

- Calcule la densidad espectral de potencia del proceso que modela la señal reconstruida, $Y(t)$.

- b) Para el caso de que $T = \frac{1}{2B}$, relación definida por el Teorema de Nyquist, calcule la potencia del proceso.
- c) Sea ahora $X(t)$ un proceso aleatorio estacionario con densidad espectral de potencia

$$S_X(j\omega) = \frac{N_0}{2} \cdot \Pi\left(\frac{\omega}{2W}\right),$$

donde $W = 2\pi B$ rad/s. La secuencia de muestras $X[n]$ se obtiene ahora muestreando el proceso $X(t)$, es decir, $X[n] = X(nT)$, siendo $T = \frac{1}{2B}$. Determine para este caso la densidad espectral de potencia y la potencia del proceso reconstruido $Y(t)$. ¿Cuál es la relación entre $X(t)$ e $Y(t)$?

Ejercicio 2.11 Calcule la media y autocorrelación del proceso

$$X(t) = A \cdot \cos(\omega t + \theta) \quad (2.1)$$

y determine si es estacionario o cicloestacionario (si es así, determine el periodo) en los siguientes casos:

- a) A es una variable aleatoria gaussiana real de media nula y varianza unidad y ω y θ son constantes reales no nulas.
- b) A es una variable aleatoria gaussiana real de media igual a 1 y varianza unidad y ω y θ son constantes reales no nulas.
- c) ω es una variable aleatoria con distribución uniforme en el intervalo $(0, 2\pi]$ y A y θ son constantes reales no nulas.
- d) θ es una variable aleatoria con distribución uniforme en el intervalo $(0, 2\pi]$ y A y ω son constantes reales no nulas.
- e) θ es una variable aleatoria con distribución uniforme en el intervalo $(0, \pi]$ y A y ω son constantes reales no nulas.
- f) A es una variable aleatoria gaussiana real de media nula y varianza unidad, ω y θ son variables aleatorias con distribución uniforme en el intervalo $(0, 2\pi]$, y las tres estadísticamente independientes.

Ejercicio 2.12 El proceso aleatorio $Y(t)$ viene dado por

$$Y(t) = A_c \cdot \cos(\omega_c t) \cdot X(t),$$

donde A_c y ω_c son dos constantes que determinan la amplitud y frecuencia de la senoide, respectivamente, y $X(t)$ es un proceso aleatorio estacionario, de media nula, función de autocorrelación $R_X(\tau)$, densidad espectral de potencia $S_X(j\omega)$, y potencia P_X .

- a) Calcule la media del proceso $Y(t)$.
- b) Calcule la función de autocorrelación del proceso $Y(t)$.
- c) Calcule la densidad espectral de potencia de $Y(t)$.
- d) Calcule la potencia del proceso $Y(t)$.

Ejercicio 2.13 Una determinada modulación analógica que modula simultáneamente dos señales moduladoras se puede definir mediante el siguiente proceso aleatorio

$$S(t) = M_A(t) \cdot \cos(\omega_c t) + M_B(t) \cdot \text{sen}(\omega_c t),$$

donde $M_A(t)$ y $M_B(t)$ son dos procesos aleatorios que modelan las dos señales moduladoras. Se supone que ambos son procesos aleatorios independientes, estacionarios, de media nula, e idénticas función de autocorrelación $R_{M_A}(\tau) = R_{M_B}(\tau) = R_M(\tau)$, densidad espectral de potencia $S_{M_A}(j\omega) = S_{M_B}(j\omega) = S_M(j\omega)$, y potencia $P_{M_A} = P_{M_B} = P_M$.

- a) Calcule la media del proceso aleatorio $S(t)$, $m_S(t)$.
- b) Calcule la función de autocorrelación del proceso $S(t)$, $R_S(t + \tau, t)$, y diga si el proceso es estacionario o cicloestacionario.
- c) Calcule la densidad espectral de potencia del proceso, $S_S(j\omega)$, y obtenga su potencia, P_S .

Ejercicio 2.14 Una señal de ruido que se modela como un proceso aleatorio estacionario, blanco, gaussiano, de media nula y con densidad espectral de potencia $N_0/2$ pasa por un filtro paso bajo ideal con ancho de banda B Hz.

- a) Calcule la función de autocorrelación del proceso de salida, $Y(t)$.
- b) Asumiendo que $\tau = \frac{1}{2B}$, calcule la función densidad de probabilidad conjunta de las variables aleatorias $Y(t)$ e $Y(t + \tau)$. ¿Son estas variables aleatorias independientes? ¿Y para $\tau = \frac{1}{4B}$?

Ejercicio 2.15 $X(t)$ es un proceso estacionario con densidad espectral de potencia $S_X(j\omega)$. Este proceso atraviesa el sistema que se muestra en la Figura 2.22

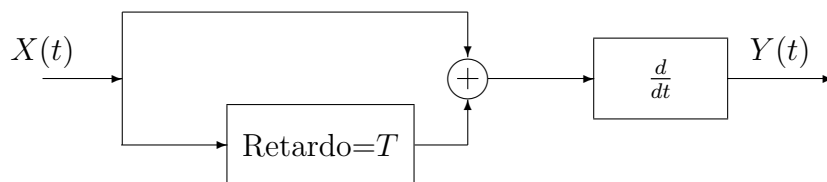


Figura 2.22: Sistema para el Problema 2.15.

- a) ¿Es el proceso $Y(t)$ estacionario? Razone la respuesta.
- b) Calcule la densidad espectral de potencia del proceso $Y(t)$.

Ejercicio 2.16 La señal recibida en un sistema de comunicaciones se obtiene como la suma de dos componentes, uno de señal y otro de ruido

$$r(t) = s(t) + n(t).$$

La señal $s(t)$ se modela con un proceso aleatorio con una densidad espectral de potencia

$$S_s(j\omega) = \frac{P_0}{1 + (\omega/\omega_0)^2},$$

donde $\omega_0 = 2\pi f_0$ y f_0 es el ancho de banda a 3 dB (en Hz) de la señal. El término de ruido, $n(t)$ se modela con un proceso aleatorio blanco que tiene una densidad espectral de potencia $N_0/2$ para todas las frecuencias. La señal recibida se filtra con un filtro paso bajo ideal con ancho de banda B Hz y ganancia unidad. Calcule y dibuje la relación señal a ruido que se tiene a la salida del filtro en función de la relación B/f_0 . ¿Para qué ancho de banda del filtro, B , se obtiene la máxima relación señal a ruido?

Ejercicio 2.17 Calcule el ancho de banda equivalente de ruido de un filtro RC paso bajo, teniendo en cuenta que la respuesta en frecuencia de este tipo de filtros es

$$H(j\omega) = \frac{1}{1 + j\omega\tau}$$

donde la constante de tiempo τ toma el valor $\tau = R \cdot C$.

Ejercicio 2.18 La señal de información que se recibe a la entrada de un receptor de comunicaciones se modela mediante un proceso aleatorio $X(t)$ cuya densidad espectral de potencia es

$$S_X(j\omega) = 10^{-18} \cdot \Pi\left(\frac{\omega}{2W_X}\right),$$

donde W_X es el ancho de banda de la señal en rad/s, que se puede escribir como $W = 2\pi B_X$, y B_X es el ancho de banda en Hz, que en este caso es de $B_X = 5$ MHz. Durante la transmisión, a la señal se le suma ruido blanco y gaussiano con densidad espectral de potencia $N_0/2$, con $N_0 = 4 \cdot 10^{-21}$ Watt/Hz. En el receptor, la señal recibida se filtra con un filtro con respuesta al impulso $h(t)$. Calcule la relación señal a ruido en dB que se tiene a la salida del filtro receptor, si este filtro es:

- a) Un filtro paso bajo ideal con ancho de banda $B = B_X = 5$ MHz.
- b) Un filtro paso bajo ideal con ancho de banda $B = 2B_X = 10$ MHz.
- c) Un filtro paso bajo ideal con ancho de banda $B = B_X/2 = 2,5$ MHz.
- d) Un filtro RC paso bajo con constante de tiempo $\tau = R \cdot C = 3 \cdot 10^{-8}$.

Ejercicio 2.19 La señal de información que se recibe a la entrada de un receptor de comunicaciones se modela mediante un proceso aleatorio $X(t)$ cuya densidad espectral de potencia es

$$S_X(j\omega) = \begin{cases} 10^{-18} \cdot \left[1 + \cos\left(\frac{\omega}{2B_X}\right)\right], & \text{si } |\omega| \leq W_X = 2\pi B_X \\ 0, & \text{si } |\omega| > W_X = 2\pi B_X. \end{cases}$$

donde W_X es el ancho de banda de la señal en rad/s, que se puede escribir como $W_X = 2\pi B_X$, y B_X es el ancho de banda en Hz, que en este caso es de $B_X = 5$ MHz. Durante la transmisión, a la señal se le suma ruido térmico, siendo la temperatura $T = 290$ °K. En el receptor, la señal recibida se filtra con un filtro con respuesta al impulso $h(t)$. Calcule la relación señal a ruido en dB que se tiene a la salida del filtro receptor, cuando este filtro es:

- a) Un filtro paso bajo ideal con ancho de banda $B = B_X = 5$ MHz.
- b) Un filtro paso bajo ideal con ancho de banda $B = 2B_X = 10$ MHz.
- c) Un filtro paso bajo ideal con ancho de banda $B = B_X/2 = 2,5$ MHz.

Ejercicio 2.20 Sea $X(t)$ un proceso aleatorio estacionario, de media $m_X = 1$, y función de autocorrelación $R_X(\tau) = \cos(\omega_A\tau) + \cos(\omega_B\tau)$. A partir del mismo, se obtiene el proceso $Y(t)$ como

$$Y(t) = X(t) + a \cdot X(t - b),$$

donde a y b son dos valores enteros y constantes. Este proceso pasa por un sistema lineal e invariante con la siguiente respuesta en frecuencia

$$H(j\omega) = 2 \cdot \Lambda\left(\frac{\omega}{W}\right) = \begin{cases} 2 \cdot \left(1 - \frac{|\omega|}{W}\right), & \text{si } |\omega| \leq W \\ 0 & \text{si } |\omega| > W \end{cases},$$

dando lugar al proceso $Z(t)$. Considere que $\omega_A = 4\pi$, $\omega_B = 10\pi$ y $W = 8\pi$.

- Calcule la media, $m_Y(t)$, la función de autocorrelación, $R_Y(t + \tau, t)$, la densidad espectral de potencia, $S_Y(j\omega)$, y la potencia, P_Y , del proceso $Y(t)$, y diga si este proceso es o no estacionario y por qué.
- Calcule la media, $m_Z(t)$, la función de autocorrelación, $R_Z(t + \tau, t)$, la densidad espectral de potencia, $S_Z(j\omega)$, y la potencia, P_Z , del proceso $Z(t)$, y diga si este proceso es o no estacionario y por qué, y si los procesos aleatorios $Y(t)$ y $Z(t)$ son conjuntamente estacionarios y por qué.

NOTA: Algunas constantes y relaciones trigonométricas de interés

Constante de Planck: $h = 6,6 \times 10^{-34} \text{ J}\cdot\text{s}$, Constante de Boltzmann: $k = 1,38 \times 10^{-23} \text{ J}/^\circ\text{K}$.

$$\begin{aligned} \cos(a) \cos(b) &= \frac{1}{2} \cos(a - b) + \frac{1}{2} \cos(a + b), & \text{sen}(a) \text{sen}(b) &= \frac{1}{2} \cos(a - b) - \frac{1}{2} \cos(a + b), \\ \text{sen}(a) \cos(b) &= \frac{1}{2} \text{sen}(a - b) + \frac{1}{2} \text{sen}(a + b) \end{aligned}$$

2.7. Soluciones de los ejercicios

Ejercicio 2.1 Solución

- a) La probabilidad se calcula integrando la función densidad de probabilidad de X en el intervalo correspondiente, obteniéndose para la distribución uniforme

$$P(X > 1) = \frac{1}{4}, \quad P(X > -1) = \frac{3}{4}, \quad P(X < -1) = \frac{1}{4}, \quad P(-1 \leq X \leq 1) = \frac{1}{2}.$$

- b) Para la distribución gaussiana

$$P(X > 1) = Q(0) = \frac{1}{2}, \quad P(X > -1) = 1 - Q(1) = 0,8414,$$

$$P(X < -1) = Q(1) = 0,1586, \quad P(-1 \leq X \leq 1) = 1 - Q(1) - Q(0) = 0,3414.$$

Ejercicio 2.2 Solución

- a) Las expresiones analíticas son

$$F_X(x) = \begin{cases} 0, & x < -1 \\ \frac{x+1}{2}, & -1 \leq x \leq 1. \\ 1, & x > 1 \end{cases}$$

$$f_X(x) = \frac{1}{2}, \quad -1 \leq x \leq 1.$$

$$f_Y(y) = \begin{cases} 0, & y < -1 \\ y + 1, & -1 \leq y \leq 0 \\ 1 - y, & 0 \leq y \leq 1 \\ 0, & y > 1 \end{cases}.$$

$$F_Y(y) = \begin{cases} 0, & y < -1 \\ \left[\frac{u^2}{2} + u \right]_{-1}^y = \frac{y^2}{2} + y - \frac{1}{2}, & -1 \leq y \leq 0 \\ \int_{-1}^0 f_Y(y) + \left[u - \frac{u^2}{2} \right]_0^x = \frac{1}{2} + y - \frac{y^2}{2}, & 0 \leq y \leq 1 \\ 1, & y > 1 \end{cases}.$$

- b) Las varianzas resultantes son

$$\sigma_X^2 = \frac{1}{3}, \quad \sigma_Y^2 = \frac{1}{6}.$$

- c) Las probabilidades requeridas valen

$$P(Y > 0) = \frac{1}{2}, \quad P\left(Y > -\frac{1}{2}\right) = \frac{7}{8}, \quad P\left(Y < -\frac{1}{2}\right) = \frac{1}{8}, \quad P\left(Y > -\frac{1}{2}\right) = \frac{9}{32}$$

Ejercicio 2.3 Solución

a) La función de densidad de probabilidad de la nueva variable aleatoria es

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty \leq x \leq \infty.$$

$$m_X = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)} dx = \frac{1}{2\pi} \ln(1+x^2) \Big|_{-\infty}^{\infty} = 0.$$

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x) dx = \int_{-\infty}^{\infty} \frac{x^2}{\pi(1+x^2)} dx = \frac{1}{\pi} (x - \text{atan}(x)) \Big|_{-\infty}^{\infty} = \infty$$

b) En este caso

$$f_X(x) = \frac{2}{\pi(1+x^2)}, \quad -1 \leq x \leq 1.$$

$$m_X = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-1}^1 \frac{2x}{\pi(1+x^2)} dx = \frac{1}{\pi} \ln(1+x^2) \Big|_{-1}^1 = 0.$$

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x) dx = \int_{-1}^1 \frac{2x^2}{\pi(1+x^2)} dx = \frac{2}{\pi} (x - \text{atan}(x)) \Big|_{-1}^1 = 0,2732.$$

Ejercicio 2.4 Solución

a) $m_X(t) = \frac{2}{3}$

b) $R_X(t_1, t_2) = \frac{1}{2} + \frac{t_1 \cdot t_2}{3}$

c) Este proceso no es un proceso estacionario en sentido amplio. Pese a que la media no depende del tiempo t , la función de autocorrelación $R_X(t_1, t_2)$ depende explícitamente del valor de cada uno de los instantes t_1 y t_2 y no únicamente de su diferencia $\tau = t_1 - t_2$.

Ejercicio 2.5 Solución

La respuesta es NO. En primer lugar, porque la definición se limita a instantes temporales mayores que cero, con lo que la función densidad de probabilidad para instantes negativos va a cambiar y ya no se cumple la condición de estacionariedad.

En cualquier caso, aunque no existiera esta restricción, tampoco sería WSS.

$$\text{Cov}(X(t_1)X(t_2)) = E[X(t_1)X(t_2)] - m_X(t_1)m_X(t_2).$$

Teniendo en cuenta que la media es nula,

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \sigma^2 \text{mín}(t_i, t_j),$$

depende de los instantes precisos t_i y t_j y no de su diferencia $\tau = t_i - t_j$.

Ejercicio 2.6 Solución

a) $m_Y[n] = (1 + a) \cdot m_X$

b) $R_Y(n + k, n) = (1 + a^2)R_X[k] + aR_X[k + 1] + aR_X[k - 1]$

c) $S_Y(e^{j\omega}) = S_X(e^{j\omega}) ((1 + a^2) + ae^{j\omega} + ae^{-j\omega})$

Ejercicio 2.7 Solución

a) $m_Z(t) = 0$

 b) La función de autocorrelación $R_X(t + \tau, t)$ vale

$$R_Z(t + \tau, t) = \frac{\sigma_X^2 + \sigma_Y^2}{2} \cos(\omega_0 \tau) + \frac{\sigma_X^2 - \sigma_Y^2}{2} \cos(\omega_0(2t + \tau)).$$

 c) El proceso no es estacionario, pues la autocorrelación depende del valor concreto de t y no sólo de τ . Pero el proceso es cicloestacionario pues $R_Z(t + \tau + T_0, t + T_0) = R_Z(t + \tau, t)$ para $T_0 = \frac{\pi}{\omega_0}$.

d) La densidad espectral de potencia vale

$$S_Z(j\omega) = TF\{\tilde{R}_X(\tau)\} = \pi \frac{\sigma_X^2 + \sigma_Y^2}{2} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)].$$

 e) En el caso $\sigma_X = \sigma_Y = \sigma$ la media no cambia, sigue valiendo $m_Z(t) = 0$. En cambio, la función de autocorrelación vale ahora

$$R_Z(t + \tau, t) = \sigma^2 \cos(\omega_0 \tau) = R_Z(\tau).$$

Lo que significa que el proceso es estacionario en sentido amplio (WSS). Esto implica que la densidad espectral de potencia es

$$S_Z(j\omega) = TF\{R_Z(\tau)\} = \sigma^2 \pi [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$$

Ejercicio 2.8 Solución

a) $m_Z(t) = \frac{1}{2} \text{sen}(\omega_0 t)$

b) $R_X(t + \tau, t) = \frac{1}{3} \cos(\omega_0 \tau)$

 c) El proceso aleatorio no es estacionario, pero sí cicloestacionario, ya que aunque la función de autocorrelación sólo depende de τ , y no de t , la media sí depende de t , y además $m_Z(t + T_0) = m_Z(t)$ y $R_Z(t + \tau + T_0, t + T_0) = R_Z(t + \tau, t)$ para $T_0 = \frac{2\pi}{\omega_0}$.

d) $S_z(j\omega) = \frac{\pi}{3} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$

Ejercicio 2.9 Solución

a) $m_X(t) = m_A \sum_n g(t - nT)$

b) $R_X(t + \tau, t) = \sum_n \sum_k R_A[n - k] g(t + \tau - nT) g(t - kT)$

c) El proceso aleatorio es cicloestacionario, ya que

$$\begin{aligned} m_X(t + T) &= m_A \sum_n g(t + T - nT) = m_A \sum_n g(t - (n - 1)T) \\ &\stackrel{k=n-1}{=} m_A \sum_k g(t - kT) = m_X(t) \end{aligned}$$

$$\begin{aligned}
 R_X(t + \tau + T, t + T) &= \sum_n \sum_k R_A[n - k]g(t + \tau + T - nT)g(t + T - kT) \\
 &= \sum_n \sum_k R_A[n - k]g(t + \tau - (n - 1)T)g(t - (k - 1)T) \\
 &\stackrel{m=n-1, j=k-1}{=} \sum_m \sum_j R_A[m + 1 - (j + 1)]g(t + \tau - mT)g(t - jT) \\
 &= \sum_m \sum_j R_A[m - j]g(t + \tau - mT)g(t - jT) = R_X(t + \tau, t)
 \end{aligned}$$

d) $S_X(j\omega) = \frac{1}{T} \cdot |G(j\omega)|^2 \cdot S_A(e^{j\omega T})$

Ejercicio 2.10 Solución

a) Definiendo $W = 2\pi B$

$$S_Y(j\omega) = \frac{1}{T} \left(\frac{1}{2W} \Pi \left(\frac{\omega}{4\pi W} \right) \right)^2$$

b) $P_Y = 1$

c) La densidad espectral de potencia es

$$S_Y(j\omega) = \frac{N_0}{2} \cdot \Pi \left(\frac{\omega}{2W} \right) = S_X(j\omega)$$

y la potencia es

$$P_Y = P_X = N_0 W$$

El proceso $Y(t)$ es el proceso reconstruido a partir de las muestras tomadas sobre $X(t)$. Como la tasa de muestreo cumple el criterio de Nyquist para el muestreo, ambos procesos son iguales.

Ejercicio 2.11 Solución

a)

$$m_X(t) = 0$$

$$R_X(t + \tau, t) = \frac{\sigma^2}{2} (\cos(\omega(2t + \tau) + 2\theta) + \cos(\omega\tau))$$

El proceso no es estacionario puesto que la autocorrelación depende del origen de tiempos. Sin embargo, es cicloestacionario con período $T = \frac{\pi}{\omega}$.

b)

$$m_X(t) = \cos(\omega t + \theta)$$

$$R_X(t + \tau, t) = (\sigma^2 + 1) \cos(\omega(t + \tau) + \theta) \cos(\omega t + \theta)$$

El proceso es cicloestacionario con período $T = 2\pi\omega$. Tanto la media como la autocorrelación son periódicas con dicho período. No es estacionario en sentido amplio.

c)

$$m_X(t) = A \frac{\text{sen}(2\pi t + \theta) - \text{sen}(\theta)}{2\pi t}$$

$$R_X(t + \tau, t) = \frac{A^2}{2} \left(\frac{\text{sen}(2\pi\tau + 4\pi t + 2\theta) - \text{sen}(2\theta)}{2\pi(\tau + 2t)} + \frac{\text{sen}(2\pi\tau)}{2\pi\tau} \right)$$

No es estacionario ni cicloestacionario.

d)

$$m_X(t) = 0$$

$$R_X(t + \tau, t) = \frac{A^2}{2} \cos(\omega\tau)$$

El proceso es estacionario.

e)

$$m_X(t) = -\frac{2A}{\pi} \text{sen}(\omega t)$$

$$R_X(t + \tau, t) = \frac{A^2}{2} \cos(\omega\tau)$$

El proceso es cicloestacionario con período $T = \frac{2\pi}{\omega}$.

f)

$$m_X(t) = 0$$

$$R_X(t + \tau, t) = \frac{1}{4\pi} \frac{\text{sen}(2\pi\tau)}{\tau} = \frac{1}{2} \text{sinc}(2\tau)$$

El proceso es estacionario.

Ejercicio 2.12 Solución

a) $m_Y(t) = 0$.

b)

$$R_Y(t + \tau, t) = \frac{A_c^2}{2} R_X(\tau) [\cos(\omega_c\tau) + \cos(\omega_c(2t + \tau))].$$

c) $S_Y(j\omega) = \frac{A_c^2}{4} [S_X(j\omega - j\omega_c) + S_X(j\omega + j\omega_c)]$.

d) $P_Y = \frac{A_c^2}{2} \cdot P_X$.

Ejercicio 2.13 Solución

a) $m_S(t) = 0$

b) $R_S(t + \tau, t) = R_M(\tau) \cdot \cos(\omega_c\tau)$. Como la media es una constante, y la función de autocorrelación es una función que no depende del tiempo, sino de la diferencia entre instantes temporales τ , el proceso aleatorio $S(t)$ es un proceso estacionario (en sentido amplio).

c)

$$S_S(j\omega) = \frac{1}{2} \cdot S_M(j\omega - j\omega_c) + \frac{1}{2} \cdot S_M(j\omega + j\omega_c).$$

$$P_S = P_M.$$

Ejercicio 2.14 Solución

a) $R_Y(\tau) = N_0 B \text{sinc}(2B\tau)$.

- b) La función densidad de probabilidad conjunta es una distribución conjuntamente gaussiana, definida por el vector de medias $\boldsymbol{\mu} = [0, 0]$ y por la matriz de covarianza

$$\mathbf{C} = \begin{bmatrix} N_0B & 0 \\ 0 & N_0B \end{bmatrix},$$

Las variables son independientes, al ser nula su covarianza y ser conjuntamente gaussianas.

Para $\tau = \frac{1}{4B}$, la función densidad de probabilidad conjunta está ahora definida por el vector de medias $\boldsymbol{\mu} = [0, 0]$ y por la matriz de covarianza

$$\mathbf{C} = \begin{bmatrix} N_0B & N_0B \operatorname{sinc}\left(\frac{1}{2}\right) \\ N_0B \operatorname{sinc}\left(\frac{1}{2}\right) & N_0B \end{bmatrix}.$$

Ahora, no son independientes, ya que al no ser nula la covarianza entre ambas variables aleatorias, la matriz de covarianza ya no es una matriz diagonal.

Ejercicio 2.15 Solución

- a) La respuesta es SÍ. Si un proceso estacionario $X(t)$ pasa por un sistema lineal e invariante, no sólo el proceso de salida $Y(t)$ es estacionario, sino que además $X(t)$ e $Y(t)$ son conjuntamente estacionarios. Y en este caso, el sistema completo es un sistema lineal e invariante (SLI). Un retardador es lineal, un sumador también, y la operación derivada, que tiene como respuesta impulsiva $h(t) = \delta'(t)$, es un SLI.

b) $S_Y(j\omega) = S_X(j\omega)2\omega^2 [1 + \cos(\omega T)].$

Ejercicio 2.16 Solución

$$\frac{S}{N} = 2 \underbrace{\frac{P_0}{N_0}}_{cte} \frac{\operatorname{tg}^{-1}(B/f_0)}{(B/f_0)}.$$

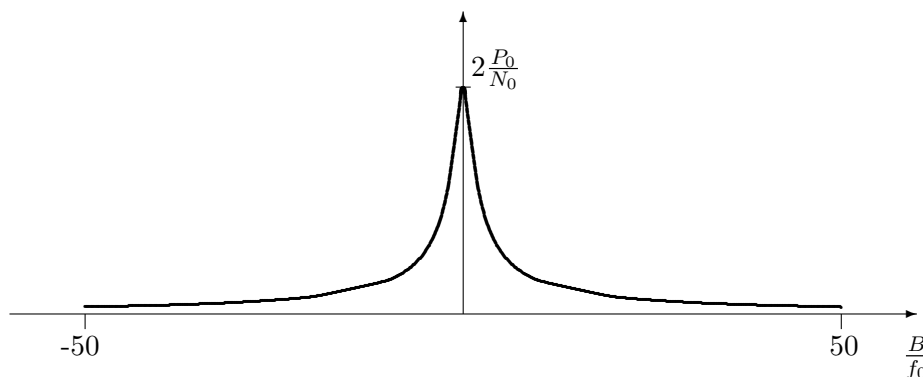


Figura 2.23: Representación de la relación señal a ruido para el Ejercicio 2.16.

El máximo de $\frac{S}{N}$ se obtiene para $\frac{B}{f_0} \rightarrow 0$ (para cero no habría ni señal ni ruido), donde tiende al valor

$$\frac{S}{N} \rightarrow 2 \frac{P_0}{N_0}.$$

Una aspecto a tener en cuenta es que aunque un valor lo más bajo posible de B mejora la relación S/N , cuando B está por debajo de f_0 la señal se degrada al perder componentes frecuenciales

importantes. Por eso, un valor de compromiso apropiado en este caso es $B = f_0$, donde la relación señal a ruido vale

$$\frac{S}{N} = 2 \frac{P_0}{N_0} 0,785 = 1,57 \frac{P_0}{N_0}.$$

Ejercicio 2.17 Solución

$$B_{eq} = \frac{\frac{1}{2\tau}}{2 \times 1} = \frac{1}{4\tau} = \frac{1}{4RC}.$$

Ejercicio 2.18 Solución

a)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{10^{-11}}{2 \cdot 10^{-14}} = 27 \text{ dB}.$$

b)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{10^{-11}}{4 \cdot 10^{-14}} = 24 \text{ dB},$$

c)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{5 \cdot 10^{-12}}{10^{-14}} = 27 \text{ dB}.$$

d)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{8 \cdot 10^{-12}}{3,33 \cdot 10^{-14}} = 23,8 \text{ dB}.$$

-o-O-o-

Ejercicio 2.19 Solución

a)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{10^{-11}}{2,001 \times 10^{-14}} = 26,97 \approx 27 \text{ dB}.$$

b)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{10^{-11}}{4,002 \times 10^{-14}} = 23,97 \approx 24 \text{ dB}.$$

c)

$$\frac{S}{N} \text{ (dB)} = 10 \log_{10} \frac{P_Y}{P_Z} = 10 \log_{10} \frac{0,818 \times 10^{-11}}{1,0005 \times 10^{-14}} = 29,12 \text{ dB}.$$

Ejercicio 2.20 Solución

a)

$$m_Y(t) = 1 + a$$

$$R_Y(\tau) = (1 + a^2) \cdot [\cos(\omega_A \tau) + \cos(\omega_B \tau)] \\ + a \cdot [\cos(\omega_A(\tau - b)) + \cos(\omega_B(\tau - b)) + \cos(\omega_A(\tau + b)) + \cos(\omega_B(\tau + b))].$$

Como la media es una constante, y la función de autocorrelación depende únicamente de la diferencia entre instantes de tiempo, τ , el proceso aleatorio es estacionario.

$$S_Y(j\omega) = \pi \cdot (1 + 2 \cdot a \cdot \cos(\omega b) + a^2) \cdot [\delta(\omega - \omega_A) + \delta(\omega + \omega_A) + \delta(\omega - \omega_B) + \delta(\omega + \omega_B)].$$

NOTA: Como tanto ω_A como ω_B son múltiplos enteros pares de π , al ser b un entero, ωb es un múltiplo par de π tanto para $\omega = \omega_A$ como para $\omega = \omega_B$, por lo que $\cos(\omega(\tau \pm b)) = \cos(\omega\tau)$ y $\cos(\omega b) = 1$ para esas dos frecuencias, por lo que las expresiones anteriores se pueden simplificar, y teniendo también en cuenta que $1 + 2 \cdot a + a^2 = (1 + a)^2$, quedarían como

$$R_Y(\tau) = (1 + a)^2 \cdot [\cos(\omega_A\tau) + \cos(\omega_B\tau)].$$

$$S_Y(j\omega) = \pi \cdot (1 + a)^2 \cdot [\delta(\omega - \omega_A) + \delta(\omega + \omega_A) + \delta(\omega - \omega_B) + \delta(\omega + \omega_B)].$$

La potencia del proceso es

$$P_Y = 2 \cdot (1 + a^2) + 2 \cdot a \cdot [\cos(\omega_A b) + \cos(\omega_B b)] = 2 \cdot (1 + a^2) + 4 \cdot a = 2 \cdot (1 + a)^2 \text{ Watt}$$

b)

$$m_Z = 2 \cdot m_Y = 2 \cdot (1 + a)$$

$$R_Z(\tau) = (1 + a^2) \cdot [\cos(\omega_A\tau)].$$

El proceso es estacionario, al no depender del tiempo ni la media ni la función de autocorrelación.

$$S_Z(j\omega) = \pi \cdot (1 + a^2) \cdot [\delta(\omega - \omega_A) + \delta(\omega + \omega_A)].$$

$$P_Z = (1 + a)^2 \text{ Watt.}$$

NOTA: En caso de no haberse dado cuenta de las simplificaciones al considerar que ω_A y ω_B son múltiplos pares de π , las expresiones quedarían

$$R_Z(\tau) = TF^{-1} [S_Z(j\omega)] = (1 + a^2) \cdot [\cos(\omega_A\tau)] + a \cdot [\cos(\omega_A(\tau - b)) + \cos(\omega_A(\tau + b))].$$

$$S_Z(j\omega) = \pi \cdot (1 + 2 \cdot a \cdot \cos(\omega b) + a^2) \cdot [\delta(\omega - \omega_A) + \delta(\omega + \omega_A)].$$

$$P_Z = 1 + a^2 + 2 \cdot a \cdot \cos(\omega_A b) \text{ Watt.}$$

Capítulo 3

Modulaciones Analógicas

En este capítulo se presentan las modulaciones analógicas más frecuentemente utilizadas, las modulaciones de amplitud y las modulaciones angulares, sus principales características y el efecto del ruido en las mismas.

3.1. Introducción al concepto de modulación

Una señal analógica es una señal en tiempo continuo con un rango continuo de posibles valores (señal continua en tiempo continuo). La mayoría de las fuentes de información tienen este tipo de señales de salida. Voz y vídeo son dos ejemplos significativos de fuentes analógicas. La Figura 3.1 muestra una señal de voz. En este tipo de señales, la información está impresa en la forma de onda de la propia señal. Por tanto, un sistema de comunicaciones analógico debe tratar de transmitir dicha forma de onda de la forma más fiel posible hasta el otro extremo del sistema.

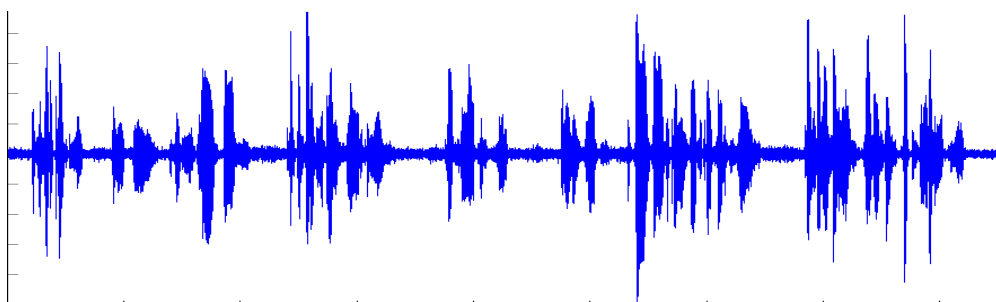


Figura 3.1: Ejemplo de una señal de voz.

La tendencia general consiste en muestrear y cuantificar las señales (conversión analógico/digital, o conversión A/D) y transmitir las por un sistema de comunicaciones digitales, para en el receptor reconstruirlas (conversión digital/analógico, o conversión D/A). Sin embargo, todavía existen hoy en día algunos sistemas de comunicaciones analógicos, especialmente para la radiodifusión de señales de audio y de vídeo. Por eso es necesario analizar estos sistemas.

La transmisión de señales analógicas se puede realizar fundamentalmente de dos modos diferentes:

1. Transmisión en banda base o transmisión sin modular: se transmite directamente, sin modificar, la señal de información.
2. Transmisión de la señal paso banda o modulada: se modifica la señal de información, trasladando el espectro de la señal que pasará a estar centrado en torno a una cierta frecuencia central o de portadora ω_c rad/s. Se puede modificar o no la forma del espectro o el ancho de banda de la señal, tal y como se muestra en la Figura 3.2 en los ejemplos A y B, respectivamente.

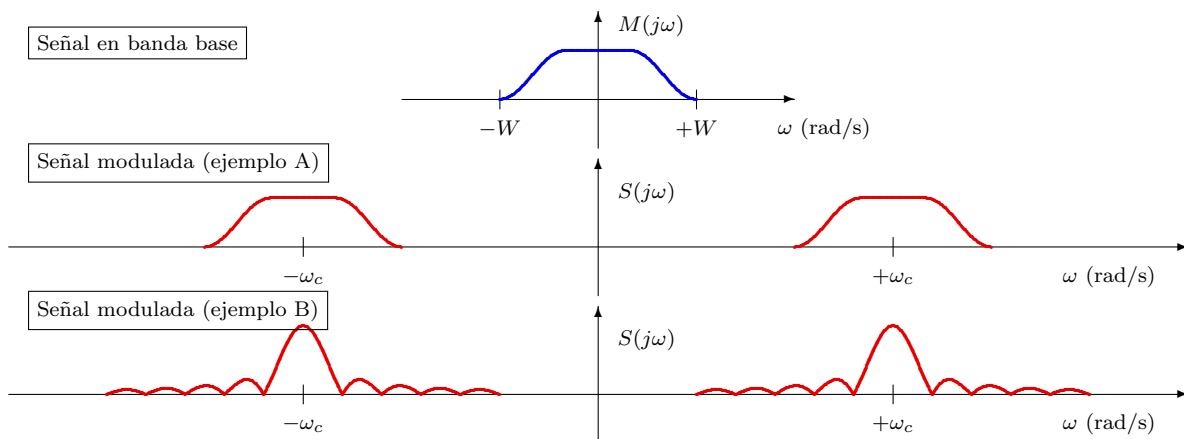


Figura 3.2: Ejemplo sencillo de multiplexación de tres señales por división en frecuencia.

La opción más frecuente es la transmisión de la señal modulada. Al proceso de modificación de la señal de información para generar otra señal con distintas características espectrales pero que contenga impresa la información se le denomina genéricamente modulación. La señal generada que será transmitida se denomina señal modulada, mientras que la señal de información se denomina señal moduladora. Habitualmente, la transmisión se suele realizar mediante la modulación de una señal llamada portadora. En general, *se imprime la señal analógica en la amplitud, frecuencia o fase de una portadora sinusoidal*. Esto da lugar a las denominadas modulaciones de amplitud, frecuencia o fase.

La modulación de una señal analógica persigue uno de los siguientes propósitos:

1. Trasladar en frecuencia la señal original para adecuarla a las características del canal, es decir, llevarla a aquella región de frecuencia donde el canal se comporta de un modo que permita la transmisión de la misma sin distorsión (banda de paso).
2. Acomodar la transmisión simultánea de distintas señales o fuentes de información en un mismo canal, lo que se denomina *multiplexación*. En este caso, se trataría de la denominada multiplexación por división de frecuencia (FDM), cuya idea fundamental se ilustra en la Figura 3.3. En este caso, se tienen tres señales que comparten la misma banda de frecuencias. Si se transmiten las tres señales directamente por el mismo medio, no sería posible separarlas posteriormente. En cambio, si se modifica su rango de frecuencias de cada señal y se mueve el espectro de cada una de ellas a una banda de frecuencias de tal modo que no exista solapamiento entre el espectro de las tres señales moduladas, es posible la transmisión simultánea de las tres señales de forma que en el receptor, tras filtrar cada una de las tres señales, y volver a situarlas en su rango de frecuencias original, se puede recuperar por separado cada una de las tres señales.

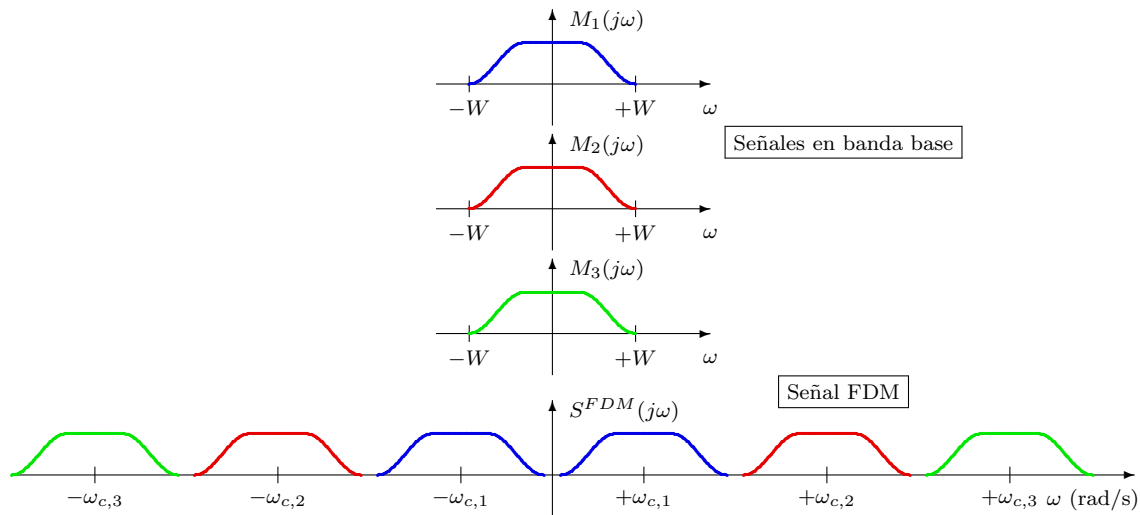


Figura 3.3: Ejemplo sencillo de multiplexación de tres señales por división en frecuencia.

- Expandir el ancho de banda de la señal transmitida para aumentar la inmunidad al ruido de la transmisión.

Los dos primeros objetivos se consiguen con los tres tipos de modulaciones anteriormente mencionadas (amplitud, frecuencia o fase), mientras que el tercero sólo se produce con las denominadas *modulaciones angulares*, que son las modulaciones en frecuencia y en fase.

3.1.1. Notación básica y modelos de señal moduladora

A continuación estableceremos algunos aspectos de la notación que se utilizará durante todo el capítulo. En cuanto a la señal de información analógica que se quiere transmitir, o señal moduladora, se denotará como $m(t)$. A la hora de presentar y analizar los distintos tipos de modulación, en algunos casos se supondrá que se conoce de forma determinista dicha señal; en otros casos se supondrá que no se conoce de forma determinista, sino que será una señal aleatoria de la que se conocen sus parámetros estadísticos.

En el primer caso, cuando se considera la señal moduladora determinista, en general se asumirá que se trata de una señal con las siguientes características:

- Señal paso bajo con ancho de banda B Hz o $W = 2\pi B$ rad/s; es decir, que su transformada de Fourier, $M(j\omega)$, es tal que $M(j\omega) = 0$ para $|\omega| > W$ rad/s.
- Es una señal de potencia. Su potencia es

$$P_m = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{T/2}^{T/2} |m(t)|^2 dt$$

Cuando se considera la señal moduladora aleatoria, para el análisis de estos sistemas, esta señal se caracterizará mediante un proceso estacionario $M(t)$ con las siguientes características:

- Es un proceso aleatorio estacionario en sentido amplio (WSS).

2. Tiene una función de autocorrelación $R_M(\tau)$ conocida.
3. La densidad espectral de potencia es $S_M(j\omega)$ (relacionada con la función de autocorrelación a través de la transformada de Fourier).
4. Es un proceso aleatorio limitado en banda, con ancho de banda B Hz o $W = 2\pi B$ rad/s; es decir, $S_M(j\omega) = 0$ para $|\omega| > W$ rad/s.
5. Potencia del proceso es P_M , que se puede obtener a partir de las funciones anteriores como

$$P_M = R_M(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_M(j\omega) d\omega.$$

La señal se transmite a través de un canal de comunicaciones imprimiéndola en una señal portadora de la forma

$$c(t) = A_c \cos(2\pi f_c t + \phi_c) = A_c \cos(\omega_c t + \phi_c),$$

donde A_c es la amplitud, f_c es la frecuencia y ϕ_c es la fase de la señal portadora. Se dice que la señal $m(t)$ modula la señal portadora en amplitud, frecuencia o fase, si la amplitud frecuencia o fase dependen de $m(t)$. En cualquier caso, el efecto de la modulación es *convertir la señal paso bajo, $m(t)$, en una señal paso banda en el entorno de la frecuencia de la señal portadora, f_c* . Resumiendo, se tienen los siguientes tipos de modulaciones analógicas:

1. Modulaciones de amplitud (AM: *Amplitude Modulation*)

La amplitud de la portadora varía en el tiempo en función de la señal moduladora

$$A_c \rightarrow A_c(t) = f(m(t))$$

2. Modulaciones angulares

El valor angular de la portadora varía en el tiempo en función de la señal moduladora

- a) Modulación de fase (PM: *Phase Modulation*)

La fase de la portadora varía en el tiempo en función de la señal moduladora

$$\phi_c \rightarrow \phi_c(t) = f(m(t))$$

- b) Modulación de frecuencia (FM: *Frequency Modulation*)

La frecuencia instantánea de la portadora varía en el tiempo en función de la señal moduladora

$$f_i(t) = f_c \rightarrow f_i(t) = f(m(t))$$

$f_i(t)$: frecuencia instantánea de la señal portadora en el instante t

3.2. Modulaciones de amplitud (AM)

En una modulación de amplitud o AM, la señal moduladora $m(t)$ se imprime en la amplitud de la señal portadora $c(t)$, es decir, que pasa de tener un valor constante A_c a variar en el tiempo en función de $m(t)$. Por eso se le llama también modulación de doble banda lateral con portadora. Existen diferentes variantes de modulaciones de amplitud. A continuación se van a ver las cuatro siguientes

1. AM: Modulación AM convencional (o modulación AM de doble banda lateral con portadora).
2. DBL: Modulación de Doble Banda Lateral (sin portadora).
3. BLU: Modulación de Banda Lateral Única.
4. BLV: Modulación de Banda Lateral Vestigial.

3.2.1. AM convencional

Una señal AM convencional consiste en la suma de la señal portadora más una señal AM de doble banda lateral, que consiste en el producto entre la señal moduladora y la portadora. La expresión analítica general para la señal modulada es por tanto

$$s(t) = \underbrace{A_c \cos(\omega_c t + \phi_c)}_{\text{Portadora } c(t)} + \underbrace{m(t) \times A_c \cos(\omega_c t + \phi_c)}_{\text{Doble banda lateral (DBL): } m(t) \times c(t)}$$

$$s(t) = A_c \cdot [1 + m(t)] \cdot \cos(\omega_c t + \phi_c). \tag{3.1}$$

La amplitud que tiene ahora la senoide a lo largo del tiempo es $A_c \cdot [1 + m(t)]$. En la mayor parte de los casos resulta útil imponer la restricción $|m(t)| \leq 1$ de modo que esta amplitud $A_c[1 + m(t)]$ sea siempre positiva, ya que en ese caso este será el valor de la envolvente de la señal modulada y se facilitará la demodulación de la señal. Si para algún valor de t se cumple que $m(t) < -1$, se dice que la señal está *sobremodulada*, y el proceso necesario para la demodulación se complica. En la práctica, cuando se desea evitar la sobremodulación $m(t)$ se escala de forma que la amplitud sea siempre menor que la unidad. La forma más habitual de hacerlo es introduciendo una normalización y el denominado índice de modulación.

Teniendo en cuenta el rango dinámico de la señal moduladora, y asumiendo que $-C_M \leq m(t) \leq +C_M$, se define la señal moduladora normalizada $m_n(t)$ como

$$m_n(t) = \frac{m(t)}{\text{máx } |m(t)|} = \frac{m(t)}{C_M}.$$

A partir de esta señal normalizada, se define la señal moduladora con índice de modulación a como

$$m_a(t) = a \times m_n(t).$$

El factor de escala a se denomina *índice de modulación*, y se trata de un valor positivo. Ahora, la señal modulada AM con un índice de modulación a se define reemplazando en la expresión (3.1) la señal modulada por la señal modulada con índice de modulación a , es decir

$$s(t) = c(t) + m_a(t) \times c(t) = A_c \cdot [1 + m_a(t)] \cdot \cos(\omega_c t + \phi_c).$$

Para evita la sobremodulación se debe cumplir que el índice de modulación esté en el rango,

$$0 < a \leq 1.$$

La Figura 3.4 representa un ejemplo de la forma de onda resultante para un índice de modulación $a = \frac{1}{2}$. Se puede ver cómo la señal de información está impresa en la envolvente de la señal, y como la amplitud de la misma varía entre $A_c(1 - a)$ y $A_c(1 + a)$, en este caso entre $\frac{A_c}{2}$ y $\frac{3A_c}{2}$.

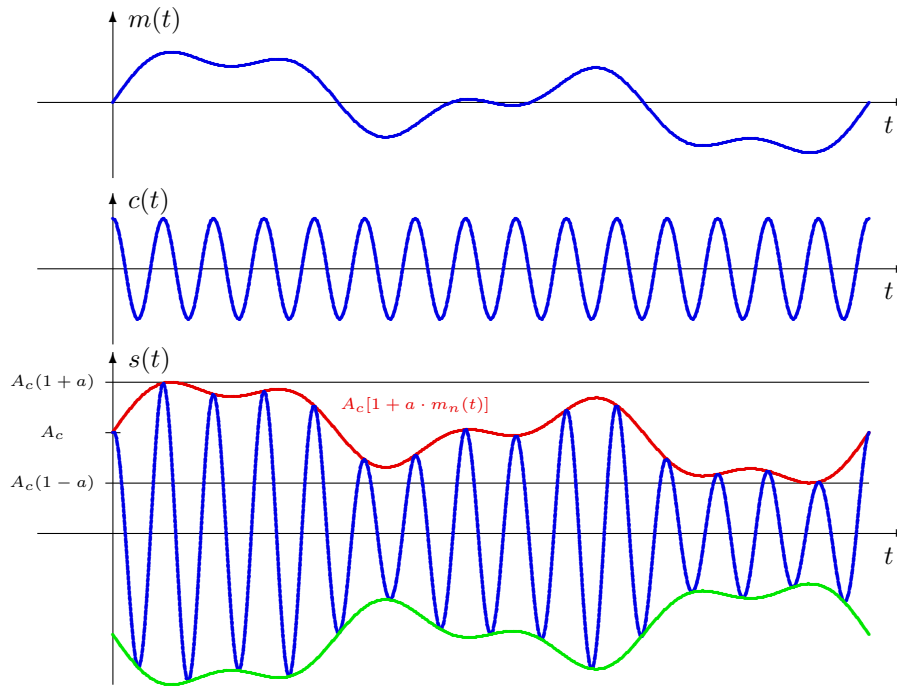


Figura 3.4: Ejemplo de señal AM convencional con índice de modulación $a = \frac{1}{2}$.

Si para la misma señal moduladora se modifica el índice de modulación, y se utiliza $a = \frac{3}{4}$, la señal modulada resultante es la que se muestra en la Figura 3.5. De nuevo la señal moduladora (información) está impresa en la envolvente de la señal, pero ahora la amplitud de la misma varía en un rango mayor, entre $\frac{A_c}{4}$ y $\frac{7A_c}{4}$. Con respecto al caso anterior, ahora los valores más pequeños de la envolvente, están más cerca de cero.

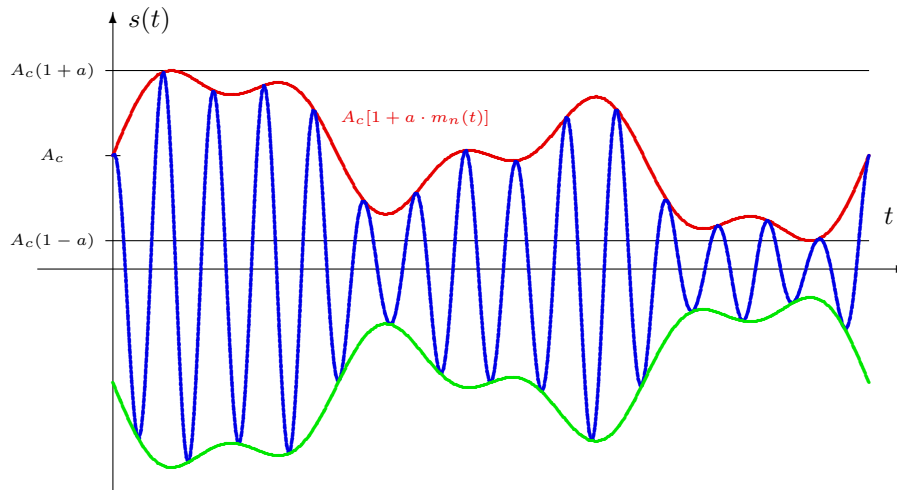


Figura 3.5: Ejemplo de señal AM convencional con índice de modulación $a = \frac{3}{4}$.

Si el índice de modulación tomara un valor mayor que 1, por ejemplo $a = \frac{3}{2}$, ahora el término de amplitud $A_c[1 + m_a(t)]$ podrá tomar valores negativos. la señal modulada resultante es la que se muestra en la Figura 3.6. Se puede ver cómo ahora la envolvente de la señal ya no contiene la forma de la señal moduladora (información), y que donde se produce el paso por cero en el término de amplitud $A_c[1 + m_a(t)]$ hay un salto de 180° en la sinusoide de la señal modulada. Esto, dificultará la demodulación de la señal.

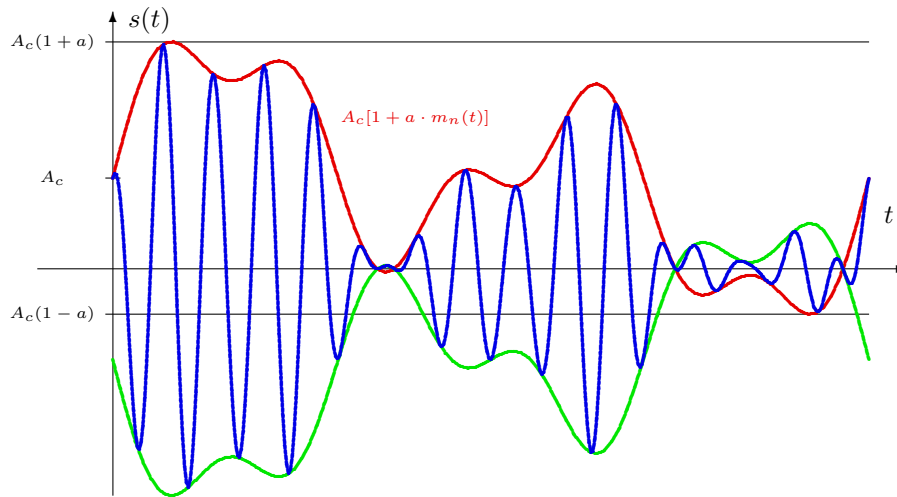


Figura 3.6: Ejemplo de señal AM convencional con índice de modulación $a = \frac{3}{4}$.

Espectro de la señal AM convencional - Caso determinista

Se va a obtener en este apartado la respuesta en frecuencia de la señal modulada cuando se considera una señal moduladora $m(t)$ determinista con transformada de Fourier $M(j\omega)$, con $M(j\omega) = 0$ para $|\omega| > W = 2\pi B$. Resulta conveniente recordar que $m_a(t) = a \cdot m_n(t)$ con $m_n(t) = \frac{1}{C_M} \cdot m(t)$, por lo que en el dominio frecuencial $M_a(j\omega) = a \cdot M_n(j\omega) = \frac{a}{C_M} \cdot M(j\omega)$. Teniendo en cuenta que la señal modulada es

$$s(t) = A_c \cos(\omega_c t + \phi_c) + m_a(t) \times A_c \cos(\omega_c t + \phi_c)$$

aplicando la propiedad básica de la transformada de Fourier de que un producto de señales en el tiempo se convierte en una convolución de sus transformadas de Fourier en el dominio frecuencial, y teniendo en cuenta que la transformada de Fourier de una senoide son dos deltas, y que el término de fase en la senoide implica una exponencial compleja en el dominio frecuencial, la transformada de Fourier de la señal modulada es

$$\begin{aligned} S(j\omega) &= TF[A_c \cos(\omega_c t + \phi_c)] + TF[m_a(t)] * TF[A_c \cos(\omega_c t + \phi_c)] \\ &= A_c \cdot \pi \cdot [\delta(\omega - \omega_c) \cdot e^{j\phi_c} + \delta(\omega + \omega_c) \cdot e^{-j\phi_c}] \\ &\quad + \frac{A_c}{2} \cdot \left[\underbrace{M_a(j\omega - j\omega_c)}_{\frac{a}{C_M} \cdot M(j\omega - j\omega_c)} \cdot e^{j\phi_c} + \underbrace{M_a(j\omega + j\omega_c)}_{\frac{a}{C_M} \cdot M(j\omega + j\omega_c)} \cdot e^{-j\phi_c} \right]. \end{aligned}$$

Si se analiza esta expresión, se llega a las siguientes conclusiones

- Módulo de la transformada de Fourier $S(j\omega)$
 - Dos deltas, en $-\omega_c$ y en $+\omega_c$
 - Amplitud $A_c \pi$
 - Réplicas de la forma de $M(j\omega)$ desplazadas a $-\omega_c$ y $+\omega_c$
 - Factor de escala $\frac{aA_c}{2C_M}$
- Fase de la transformada de Fourier

- La fase de la portadora introduce el término $e^{-j\phi_c}$
 - Término de fase constante
- Ancho de banda de la señal modulada

$$BW_{AM} = 2 \cdot B \text{ Hz}$$

El ancho de banda de la señal modulada es por tanto el doble que el ancho de banda de la señal moduladora transmitida $m(t)$

Las Figuras 3.7 y 3.8 muestran dos ejemplos de transformada de Fourier de la señal modulada, para dos casos particulares de respuesta en frecuencia de la señal moduladora.

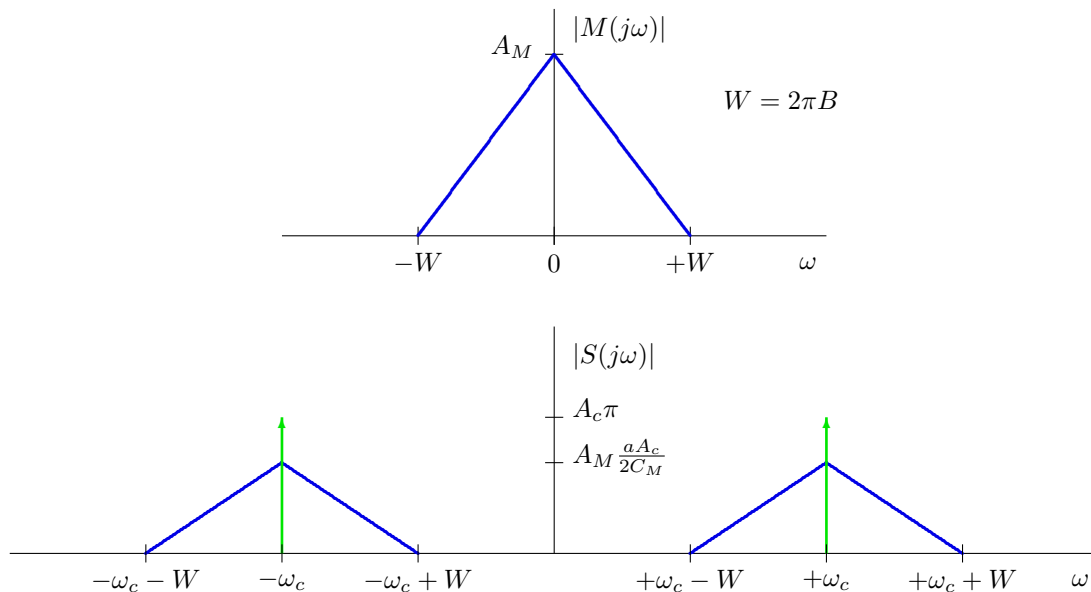


Figura 3.7: Un ejemplo de transformada de Fourier de la señal modulada AM convencional.

Análisis estadístico de la modulación AM convencional

Se considera ahora el caso en el la señal moduladora se considera una señal aleatoria cuya estadística es conocida y modelada mediante un proceso aleatorio $M(t)$ estacionario, con las características definidas en la Sección 3.1.1

$$M(t), \text{ estacionario, con } m_M = 0, R_M(\tau), S_M(j\omega), \text{ y potencia } P_M$$

En primer lugar se obtendrán los parámetros estadísticos del proceso aleatorio que modela la señal modulada, que se define como

$$S(t) = A_c[1 + M_a(t)] \cos(\omega_c t + \phi_c),$$

donde el proceso $M_a(t)$ modela la señal moduladora con índice de modulación a , que viene dada por $M_a(t) = a \cdot M_n(t) = \frac{a}{C_M} \cdot M(t)$. La media de este proceso aleatorio $S(t)$ es

$$m_S(t) = E[S(t)] = A_c[1 + E[M_a(t)]] \cos(\omega_c t + \phi_c) = A_c \cos(\omega_c t + \phi_c)$$

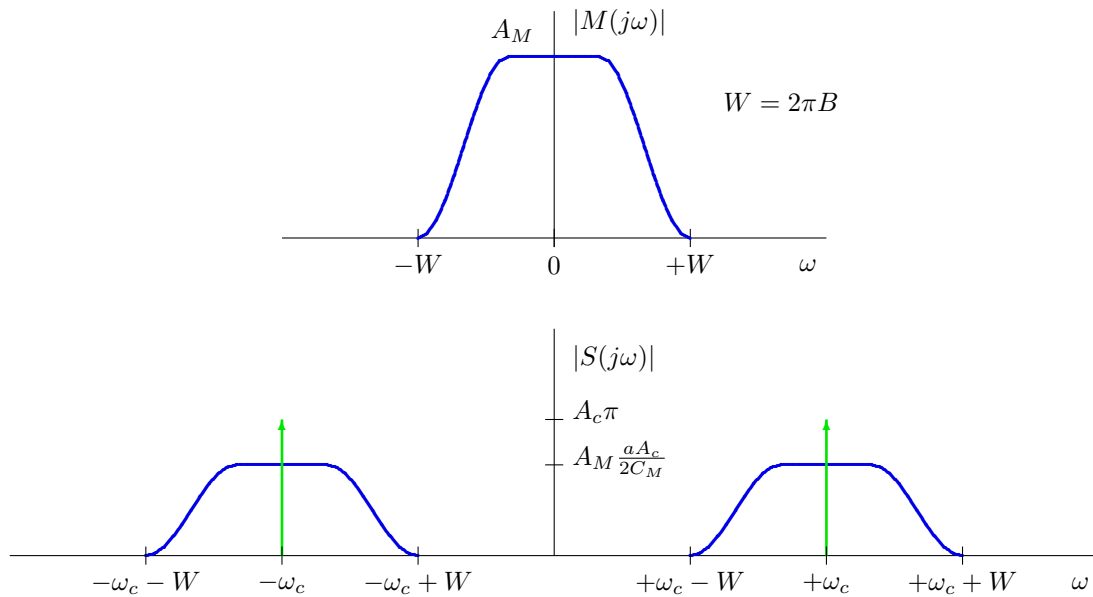


Figura 3.8: Otro ejemplo de transformada de Fourier de la señal modulada AM convencional.

ya que $E[M_a(t)]$ es la definición de la media del proceso $M_a(t)$, y si $M_a(t) = a \cdot M_n(t) = \frac{a}{C_M} \cdot M(t)$, $E[M_a(t)] = \frac{a}{C_M} \cdot E[M(t)] = 0$.

Y la función de autocorrelación se obtiene como

$$\begin{aligned}
 R_S(t, t + \tau) &= E[S(t) \times S(t + \tau)] \\
 &= A_c^2 \cdot E \left[\underbrace{(1 + M_a(t))(1 + M_a(t + \tau))}_{1 + M_a(t) + M_a(t + \tau) + M_a(t) \times M_a(t + \tau)} \right] \cos(\omega_c t + \phi_c) \cos(\omega_c(t + \tau) + \phi_c) \\
 &= \frac{A_c^2}{2} [1 + R_{M_a}(\tau)] [\cos(\omega_c \tau) + \cos(\omega_c(2t + \tau) + 2\phi_c)]
 \end{aligned}$$

Se ha hecho uso de la propiedad de linealidad del operador esperanza matemática, de que $E[1] = 1$, y del hecho de que al ser $M_a(t)$ un proceso aleatorio estacionario, $E[M_a(t)] = E[M_a(t + \tau)] = 0$, y de que $E[M_a(t) \times M_a(t + \tau)]$ es la definición de la función de autocorrelación del proceso aleatorio $M_a(t)$, i.e, $R_{M_a}(\tau)$. Asimismo se ha utilizado la igualdad trigonométrica

$$\cos(a) \times \cos(b) = \frac{1}{2} \cos(a - b) + \frac{1}{2} \cos(a + b).$$

Se puede comprobar que tanto la media como la autocorrelación son funciones periódicas de período $T_m = \frac{2\pi}{\omega_c} = \frac{1}{f_c}$ para la media y $T_R = \frac{2\pi}{2\omega_c} = \frac{1}{2f_c}$ para la función de autocorrelación. El período común para ambas es $T = T_m$ el mayor de ambos. Por tanto el proceso es un proceso *cicloestacionario* con período T . Así pues, para caracterizarlo hace falta calcular la autocorrelación promediada en un período, que vale

$$\begin{aligned}
 \tilde{R}_S(\tau) &= \frac{1}{T} \int_{-T/2}^{T/2} R_S(t, t + \tau) dt = \frac{A_c^2}{2} \cdot [1 + R_{M_a}(\tau)] \cdot \cos(\omega_c \tau) \\
 &= \frac{A_c^2}{2} \cdot \left[1 + \frac{a^2}{C_M^2} R_M(\tau) \right] \cdot \cos(\omega_c \tau)
 \end{aligned}$$

En este caso se ha tenido en cuenta que la integral sobre la variable t en el coseno de frecuencia $2\omega_c$ rad/s se hace sobre 2 períodos completos de la senoide, lo que da lugar a cero. Tenga también en cuenta que si $M_a(t) = a \cdot M_n(t) = \frac{a}{C_M} \cdot M(t)$, entonces $R_{M_a}(\tau) = a^2 \cdot R_{M_n}(\tau) = \frac{a^2}{C_M^2} \cdot R_M(\tau)$; por tanto también se tiene que $S_{M_a}(j\omega) = \frac{a^2}{C_M^2} \cdot S_M(j\omega)$ y $P_{M_a} = \frac{a^2}{C_M^2} \cdot P_M$.

Ahora la densidad espectral de potencia se calcula tomando la transformada de Fourier de este promedio en un período de la función de autocorrelación

$$\begin{aligned} S_S(j\omega) &= TF[\tilde{R}_S(\tau)] = \frac{A_c^2}{2} \cdot \pi \cdot [\delta(\omega - \omega_c) + \delta(\omega + \omega_c)] \\ &\quad + \frac{A_c^2}{4} \cdot [S_{M_a}(j\omega - j\omega_c) + S_{M_a}(j\omega + j\omega_c)] \\ &= \frac{A_c^2}{2} \cdot \pi \cdot [\delta(\omega - \omega_c) + \delta(\omega + \omega_c)] \\ &\quad + \frac{A_c^2}{4} \cdot \left[\frac{a^2}{C_M^2} \cdot S_M(j\omega - j\omega_c) + \frac{a^2}{C_M^2} \cdot S_M(j\omega + j\omega_c) \right] \end{aligned}$$

Y la potencia del proceso se puede calcular como

$$\begin{aligned} P_S = \tilde{R}_S(0) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_S(j\omega) d\omega \\ &= \frac{A_c^2}{2} [1 + R_{M_a}(0)] \\ &= \frac{A_c^2}{2} [1 + P_{M_a}] = \frac{A_c^2}{2} \cdot \left[1 + \frac{a^2}{C_M^2} \cdot P_M \right]. \end{aligned}$$

Analizando estas expresiones, se llega a las siguientes conclusiones:

- Ancho de banda de la señal AM convencional

$$BW_{AM} = 2 \cdot B \text{ Hz}$$

Densidad espectral de potencia consta de

- Dos deltas, en $-\omega_c$ y en $+\omega_c$
 - Amplitud $\frac{A_c^2}{2} \pi$
- Réplicas de la forma de $S_M(j\omega)$ desplazadas a $-\omega_c$ y $+\omega_c$
 - Factor de escala $\left(\frac{aA_c}{2C_M}\right)^2$

- Potencia AM convencional

$$P_S = \tilde{R}_S(0) = \frac{A_c^2}{2} \cdot [1 + R_{M_a}(0)] = \frac{A_c^2}{2} \cdot [1 + P_{M_a}] = \frac{A_c^2}{2} \cdot \left[1 + \frac{a^2}{C_M^2} \cdot P_M \right]$$

- Potencia de la portadora: $\frac{A_c^2}{2}$
- Potencia de la DBL: $\left(\frac{A_c^2}{2} \frac{a^2}{C_M^2}\right) \times P_M$

La potencia de la portadora no es potencia útil desde el punto de vista de que no transporta información (aunque sí es útil porque como veremos permite la utilización de un receptor simple).

Las Figuras 3.9 y 3.10 muestran dos ejemplos de densidad espectral de potencia de la señal modulada, para dos casos particulares de densidad espectral de potencia de la señal moduladora.

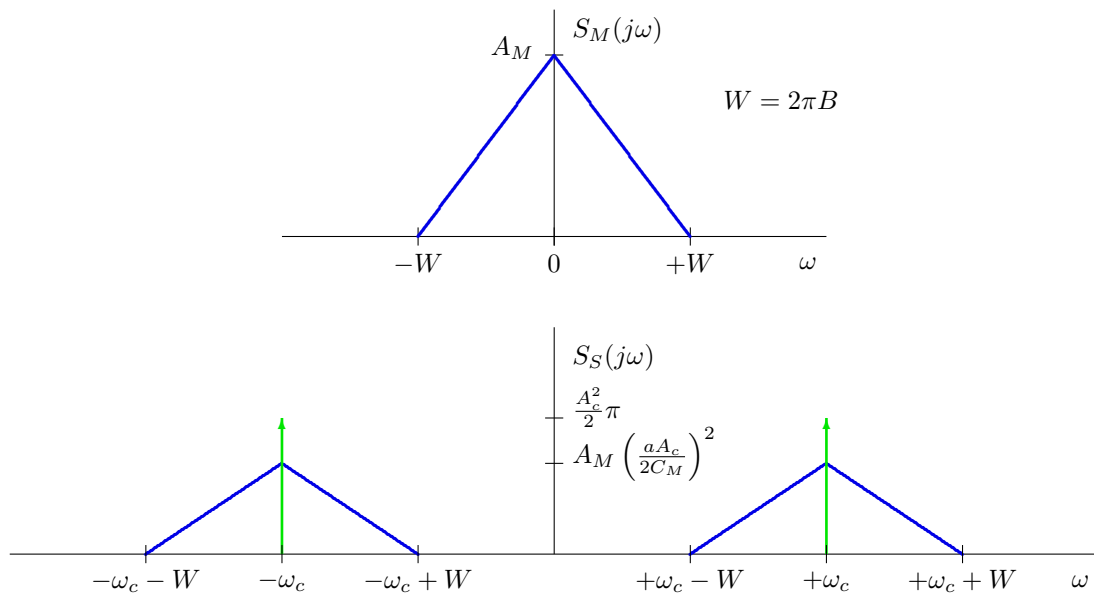


Figura 3.9: Un ejemplo de densidad espectral de potencia de la señal modulada AM convencional.

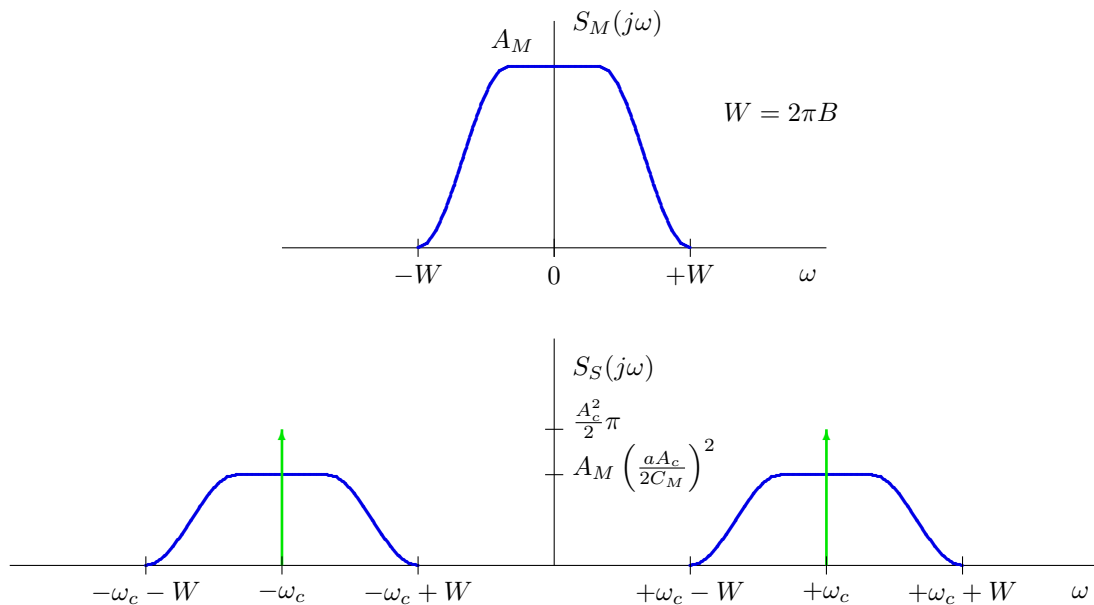


Figura 3.10: Otro ejemplo de densidad espectral de potencia de la señal modulada AM convencional.

Demodulación de la modulación AM convencional

La gran ventaja de esta modulación está en su demodulación:

1. Es muy simple, ya que al estar la señal moduladora presente en la forma de la envolvente, un simple detector de envolvente permite recuperarla.
2. No necesita demodulador síncrono o coherente, aunque se puede utilizar también (este receptor se verá más tarde, al analizar la modulación de doble banda lateral con portadora).

Como se cumple la condición $|m_a(t)| < 1$ la envolvente (amplitud) es $1 + m_a(t) > 0$. Esto permite implementar el receptor mediante un simple detector de envolvente, que se puede implementar mediante un rectificador y un filtro paso bajo con la frecuencia de corte adaptada al ancho de banda de la señal, B .

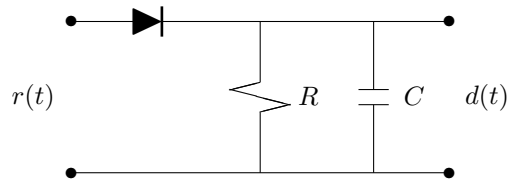


Figura 3.11: Detección de envolvente de una señal AM convencional

Este demodulador recupera la señal con un factor de ganancia y con un término de continua que hay que eliminar.

$$d(t) \approx A_c \left[1 + \frac{a}{C_M} m(t) \right].$$

La simplicidad de este demodulador hace que sea la que se utiliza para la radiodifusión de radio AM. Esto se debe a que

1. Los receptores son muy sencillos, y hay millones de ellos.
2. Aunque no es eficiente en cuanto a la potencia transmitida, hay pocos transmisores, lo que limita este problema en la práctica.

Como resumen de las características de esta modulación se podrían citar los siguientes aspectos:

- Inconvenientes de la modulación AM convencional:
 - Escasa eficiencia en potencia
 - Se gasta potencia en la transmisión de la portadora (que no contiene información)
 - Escasa eficiencia espectral
 - El ancho de banda de la señal modulada es el doble del de la moduladora
- Ventaja fundamental de la modulación AM convencional
 - Si $a \leq 1$, no hay sobremodulación y la envolvente de la señal es proporcional a $1 + m_a(t) \geq 0$, de donde se puede extraer $m(t)$

- Eliminación de la media y escalado
- Receptor simple: detector de envolvente
 - No se necesita un demodulador síncrono

3.2.2. Doble banda lateral sin portadora (DBL)

Esta técnica suprime la portadora de la modulación AM convencional de forma que se elimina el problema de la eficiencia en potencia de la misma. La expresión matemática de la misma es

$$s(t) = m(t) \times c(t) = m(t) \times A_c \cos(2\pi f_c t + \phi_c).$$

La eliminación de la portadora hace que la envolvente de la señal ya no contenga la forma de onda de la señal moduladora, lo que veremos que requerirá de un receptor más complejo para la demodulación. La Figura 3.12 muestra un ejemplo de señal modulada.

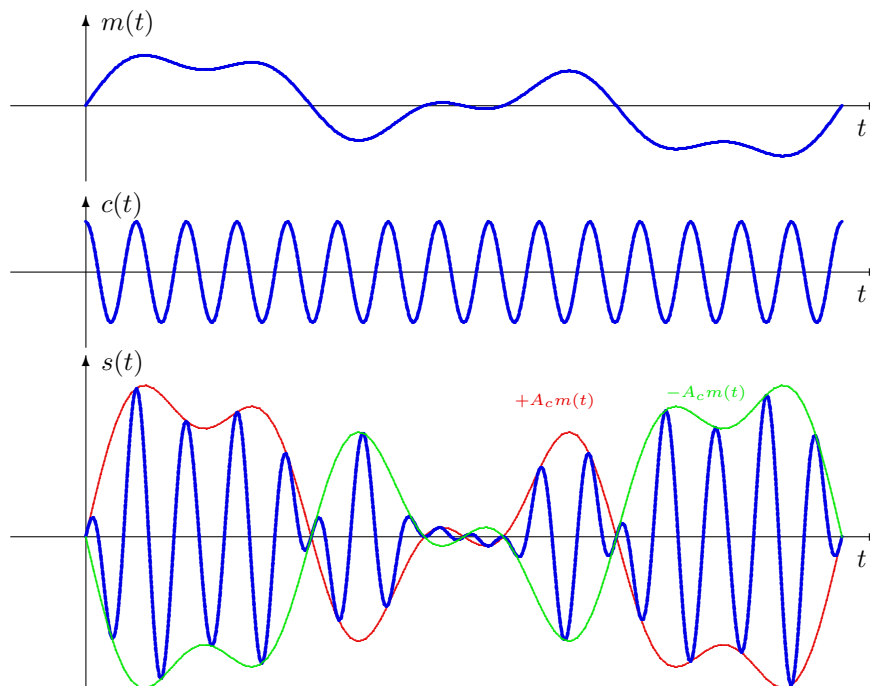


Figura 3.12: Ejemplo de señal de doble banda lateral.

Espectro de la señal DBL - Caso determinista

Se considera una señal $m(t)$ determinista con transformada de Fourier $M(j\omega)$. El espectro de la señal modulada en amplitud es

$$\begin{aligned} S(j\omega) &= TF[m(t)] * TF[A_c \cdot \cos(\omega_c t + \phi_c)] \\ &= \frac{A_c}{2} \cdot [M(j\omega - j\omega_c) \cdot e^{j\phi_c} + M(j\omega + j\omega_c) \cdot e^{-j\phi_c}] \end{aligned}$$

Con respecto a la modulación AM convencional se puede apreciar que:

- Desaparecen las deltas de la modulación AM convencional.

- Cambia el escalado de las réplicas del espectro de $M(t)$ al no haber en este caso normalización, de forma que desaparecen los factores a y C_M .

El nombre de la modulación hace referencia a que en el espectro de la señal aparecen dos bandas laterales, inferior ($|\omega| < \omega_c$) y superior ($|\omega| > \omega_c$), cada una de ellas siendo simétrica con respecto a la otra:

1. $|\omega| > \omega_c$: Banda lateral superior
2. $|\omega| < \omega_c$: Banda lateral inferior

Las Figuras 3.13 y 3.14 muestran dos ejemplos de transformada de Fourier de la señal modulada, para dos casos particulares de respuesta en frecuencia de la señal moduladora.

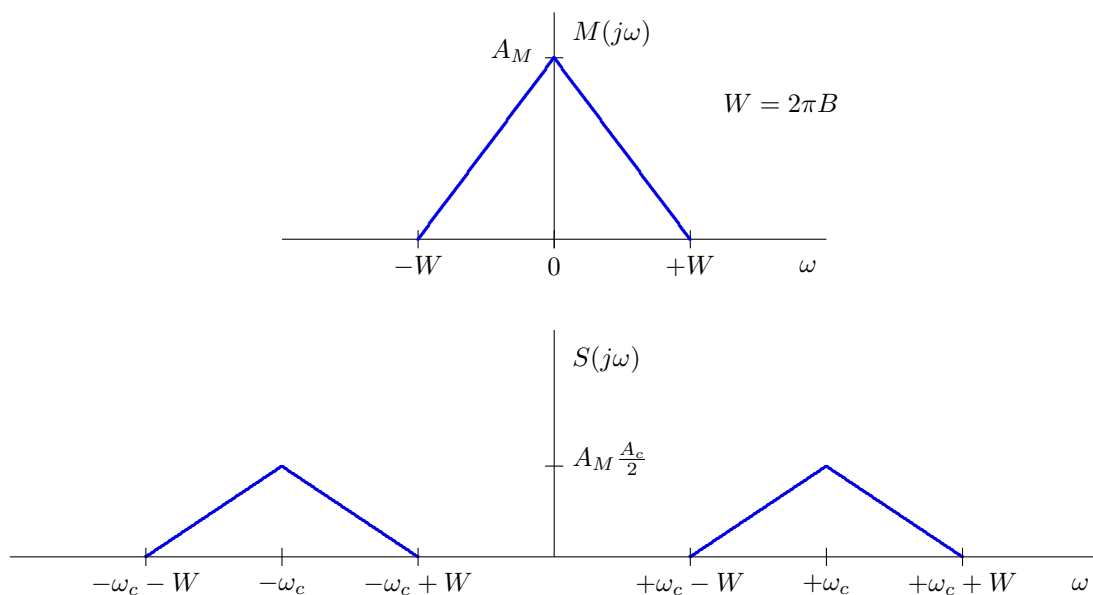


Figura 3.13: Un ejemplo de transformada de Fourier de la señal modulada de doble banda lateral.

Como se ve, el espectro de la señal DBL ocupa también dos veces el ancho de banda de la señal moduladora

$$BW_{DBL} = 2B \text{ Hz},$$

por lo que espectralmente sigue siendo igual de ineficiente que la modulación AM convencional.

La modulación DBL contiene las dos bandas. Hay que remarcar que cada una de las bandas laterales contiene toda la información de la señal, tiene todas las componentes frecuenciales de la misma. Además la portadora se ha suprimido, con lo que no aparecen las deltas en $\pm f_c$.

Análisis estadístico de la modulación de doble banda lateral

Para realizar este análisis se recurre de nuevo a los procesos estocásticos. El mensaje, o señal moduladora, se modela como un proceso estocástico estacionario en sentido amplio $M(t)$, con las características descritas en la Sección 3.1.1. La señal modulada se modela mediante el proceso aleatorio definido por

$$S(t) = M(t) \times c(t) = A_c M(t) \cos(\omega_c t + \phi_c)$$

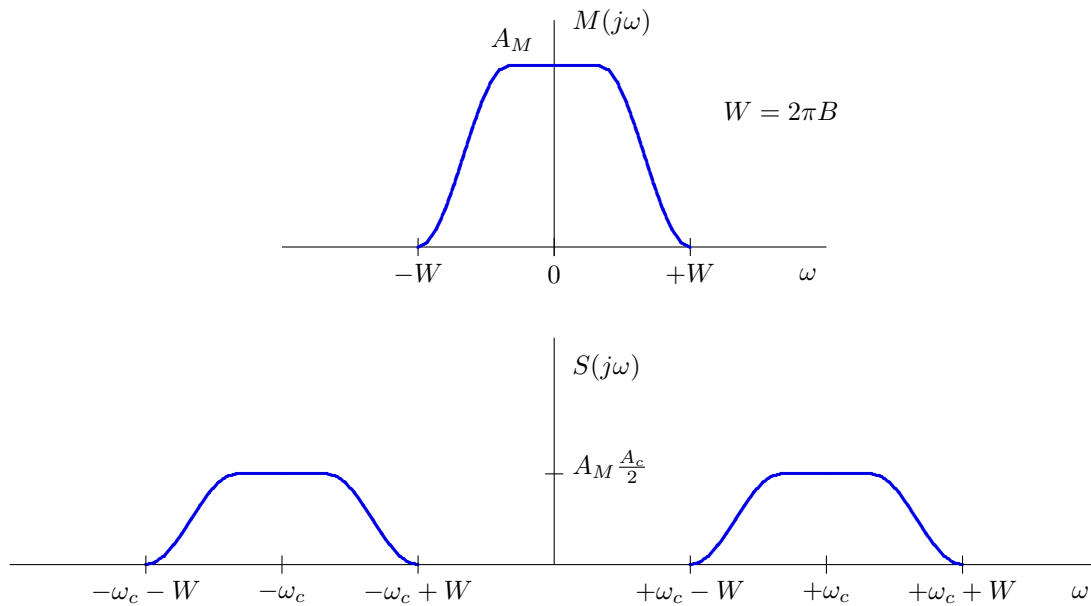


Figura 3.14: Otro ejemplo de transformada de Fourier de la señal modulada de doble banda lateral.

La media de la señal modulada es

$$m_S(t) = E[S(t)] = A_c \cdot E[M(t)] \cdot \cos(\omega_c t + \phi_c) = 0$$

Y la función de autocorrelación

$$\begin{aligned} R_S(t, t + \tau) &= A_c^2 \cdot E[M(t) \cdot M(t + \tau)] \cdot \cos(\omega_c t + \phi_c) \cdot \cos(\omega_c(t + \tau) + \phi_c) \\ &= \frac{A_c^2}{2} \cdot R_M(\tau) \cdot [\cos(\omega_c \tau) + \cos(\omega_c(2t + \tau) + 2\phi_c)] \end{aligned}$$

Se puede comprobar que aunque la media es constante, la función de autocorrelación es periódica de período $T = \frac{1}{2f_c}$. Por tanto el proceso es un proceso *cicloestacionario*. Así pues, para caracterizarlo hace falta calcular la autocorrelación promediada en un período

$$\tilde{R}_S(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} R_S(t, t + \tau) dt = \frac{A_c^2}{2} \cdot R_M(\tau) \cdot \cos(\omega_c \tau).$$

De nuevo se ha tenido en cuenta que la integral en un período de una senoide vale cero. La densidad espectral de potencia se obtiene a través de la transformada de Fourier

$$S_S(j\omega) = TF \{ \tilde{R}_S(\tau) \} = \frac{A_c^2}{4} \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)]$$

La potencia de la señal modulada es

$$P_S = \tilde{R}_S(0) = \frac{A_c^2}{2} R_M(0) = \frac{A_c^2}{2} P_M.$$

Ahora ya no aparece un término de potencia asociado a la portadora, de modo que esta modulación es más eficiente en potencia que la modulación AM convencional.

Las Figuras 3.15 y 3.16 muestran dos ejemplos de densidad espectral de potencia de la señal modulada, para dos casos particulares de densidad espectral de potencia de la señal moduladora.

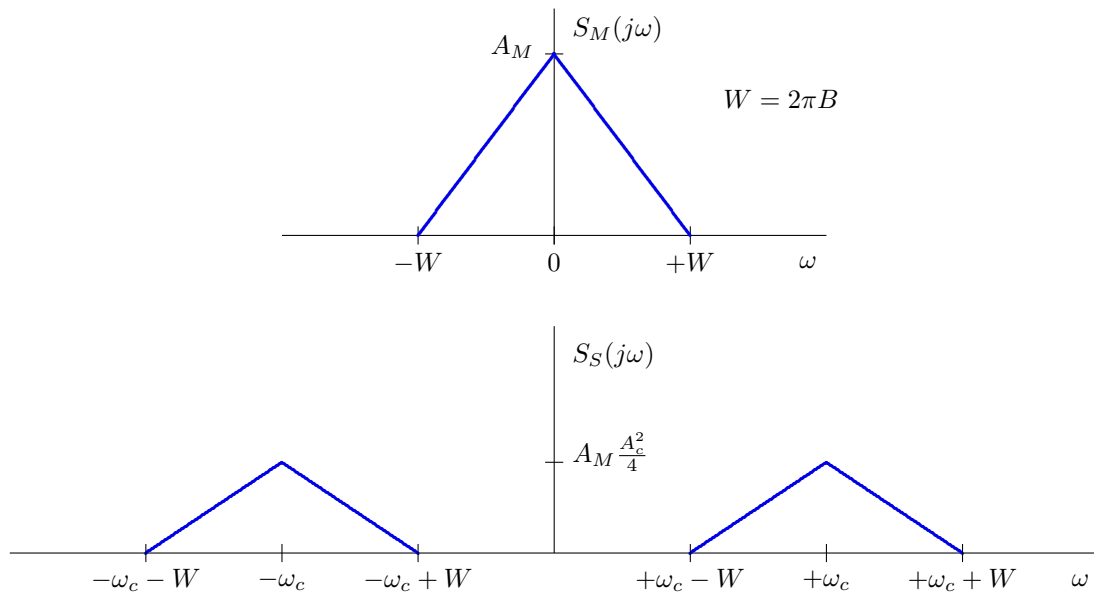


Figura 3.15: Un ejemplo de densidad espectral de potencia de la señal modulada de doble banda lateral.

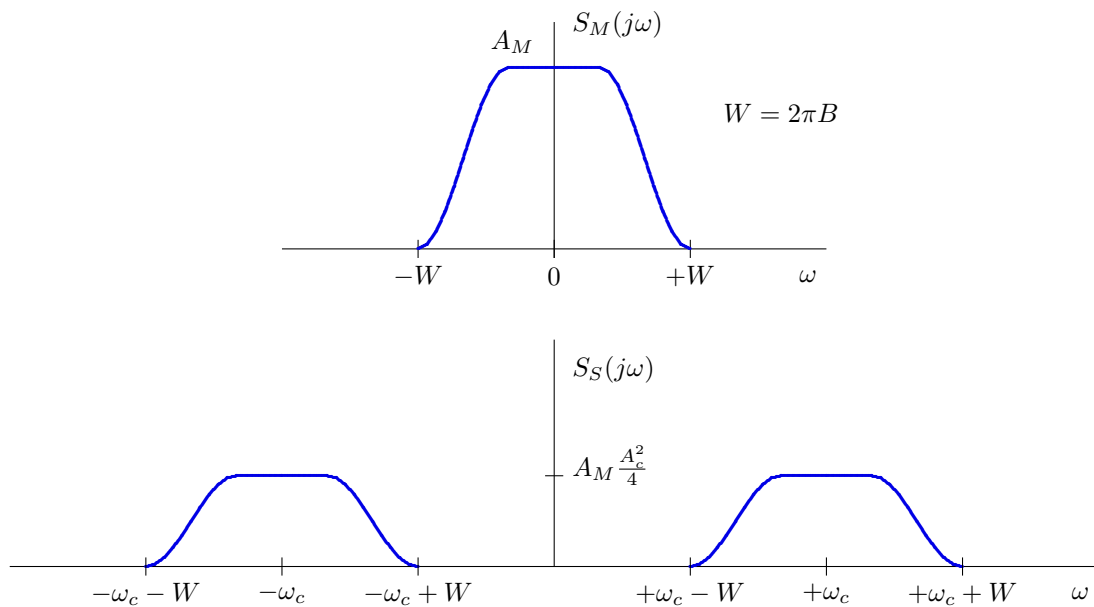


Figura 3.16: Otro ejemplo de densidad espectral de potencia de la señal modulada de doble banda lateral.

Demodulación de señales DBL

La supresión de la portadora, tal y como ya se indicó con anterioridad, y como puede verse por ejemplo en la Figura 3.12, hace que la forma de la envolvente de la señal ya no sea proporcional a la señal moduladora, por lo que no es posible utilizar un receptor coherente para demodulación. En este caso es necesario utilizar un receptor síncrono o receptor coherente. Este tipo de receptor se muestra en la Figura 3.17, donde el acrónimo LPF representa filtro paso bajo (del inglés *Low Pass Filter*) de ancho de banda de B Hz. El rendimiento óptimo se obtiene con un receptor síncrono o

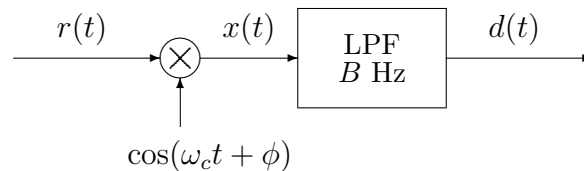


Figura 3.17: Demodulador síncrono para una señal DBL.

coherente, lo que quiere decir que la fase de la portadora del receptor es la misma que la fase de la portadora que se utilizó en el transmisor para la generación de la señal, es decir,

$$\phi = \phi_c.$$

Como veremos a continuación, si esta condición no se cumple (receptor no síncrono o incoherente), se produce una atenuación de la señal recibida y por tanto una pérdida de relación señal a ruido y de prestaciones.

Se procederá a continuación al análisis del funcionamiento del receptor. Inicialmente se asume que la señal no sufre ninguna distorsión durante su transmisión (situación ideal), por lo que la señal recibida es igual a la señal modulada transmitida

$$r(t) = s(t) = A_c m(t) \cos(\omega_c t + \phi_c).$$

La señal demodulada antes de filtrar, $x(t)$, vale por tanto

$$\begin{aligned} x(t) &= r(t) \times \cos(\omega_c t + \phi) \\ &= A_c m(t) \cos(\omega_c t + \phi_c) \cos(\omega_c t + \phi) \\ &= \frac{A_c}{2} \cdot m(t) \cdot [\cos(\phi - \phi_c) + \cos(2\omega_c t + \phi_c + \phi)]. \end{aligned}$$

El filtro paso bajo elimina las componentes de alta frecuencia, con lo que la salida filtrada del demodulador es

$$d(t) = \frac{A_c}{2} \cdot m(t) \cdot \cos(\phi_c - \phi).$$

El valor ideal, con un receptor coherente ($\phi = \phi_c$) es

$$d(t) = \frac{A_c}{2} \cdot m(t).$$

En caso de utilizar un receptor no coherente, aparece un factor $\cos(\phi_c - \phi)$ multiplicando a la señal deseada. Este término es un término de atenuación. Se pueden dar múltiples valores para la diferencia de ángulos. Algunos casos son

1. $\phi = \phi_c$. Este es el caso *ideal*, en el que el coseno vale 1.
2. $\phi_c - \phi = 45^\circ$. La amplitud se reduce por un factor $\sqrt{2}$, lo que significa que la potencia se reduce a la mitad.
3. $\phi_c - \phi = 90^\circ$. La señal desaparece.

Esto indica la necesidad de tener una *demodulación síncrona* o *demodulación coherente en fase* o simplemente coherente.

Para generar una senoide en recepción enganchada en fase con la portadora recibida, existen dos posibilidades

1. Incluir un tono piloto (una portadora de baja amplitud). Este tono se extrae en recepción con un filtro de banda estrecha sintonizado a la frecuencia ω_c . Esta opción tiene la desventaja de que se pierde eficiencia en potencia, ya que se utiliza una parte de la potencia de la señal transmitida en la generación del piloto, que no contiene propiamente información.
2. Introducir un lazo enganchado en fase o PLL (*Phase-Locked Loop*), un dispositivo que permite recuperar la fase de la portadora a partir de la señal recibida. Esta alternativa hace que el receptor sea más complejo y de mayor coste.

La Figura 3.18 muestra una interpretación frecuencial del proceso de demodulación.

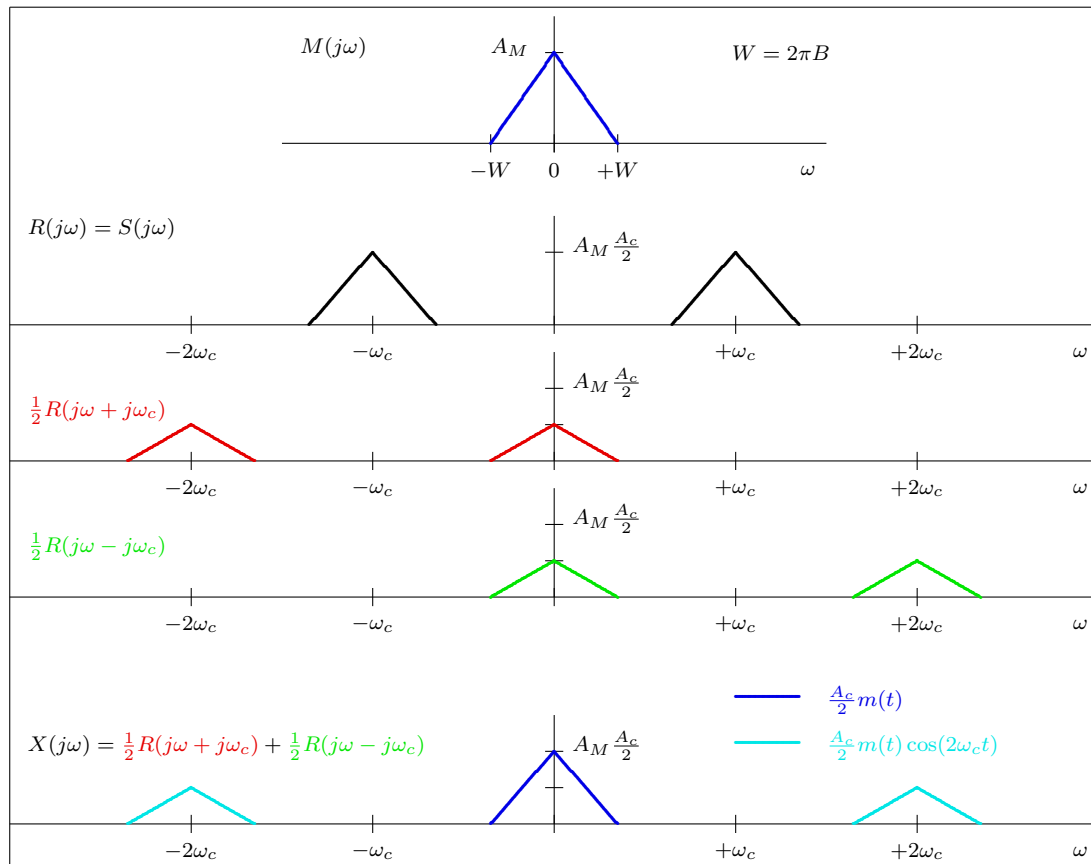


Figura 3.18: Interpretación frecuencial del proceso de demodulación de una señal de doble banda lateral.

Teniendo en cuenta que el producto con una senoide produce dos réplica del espectro (cada una con la mitad de amplitud de la senoide), una desplazada ω_c a la derecha, y otra ω_c a la izquierda

$$X(j\omega) = \frac{1}{2}R(j\omega - j\omega_c) + \frac{1}{2}R(j\omega + j\omega_c),$$

la contribución de ambos términos reconstruye el espectro de la señal y añade una componente de alta frecuencia (centrada en $2\omega_c$ que se elimina con el filtrado paso bajo).

3.2.3. Modulación de banda lateral única (BLU)

La modulación con doble banda lateral y la AM convencional tienen presentes las dos bandas laterales, cada una de las cuales contiene toda la información de la señal transmitida o moduladora por la propiedad de simetría de la respuesta en frecuencia de señales reales. El uso de las dos bandas es redundante y se desperdicia un recurso importante como es el ancho de banda. La modulación de banda lateral única transmite una única banda lateral, de forma que se reduce el ancho de banda a la mitad.

$$BW_{BLU} = B \text{ Hz.}$$

La Figura 3.19 muestra la respuesta en frecuencia de este tipo de modulación.

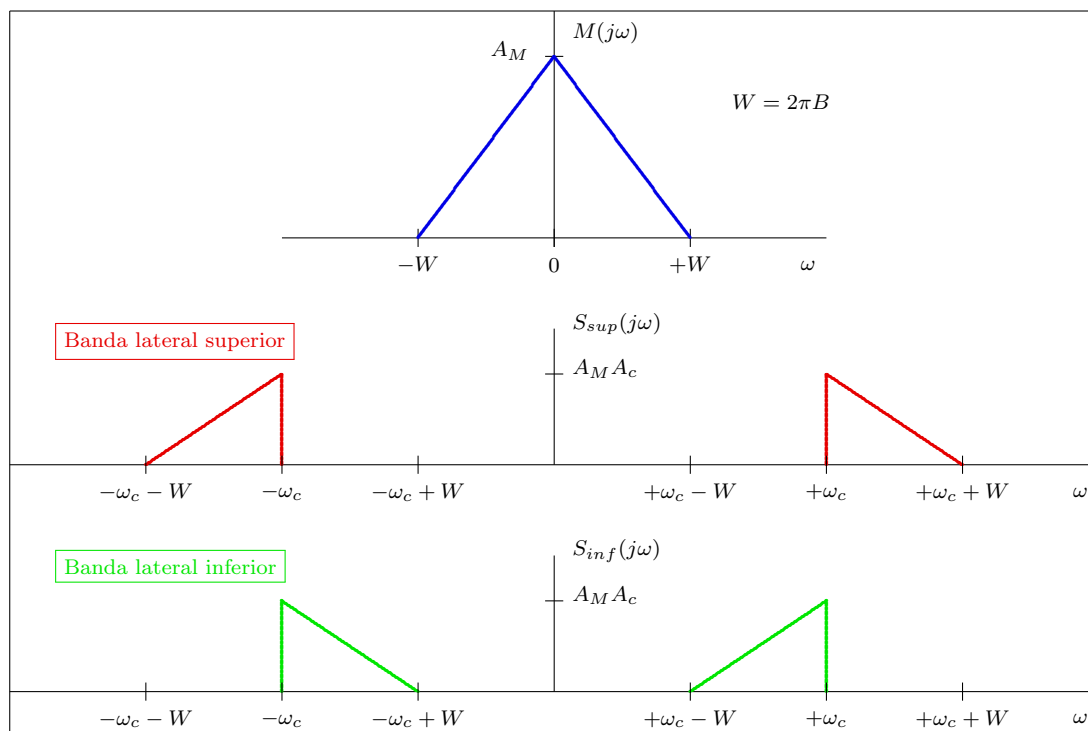


Figura 3.19: Espectro de la señales de banda lateral única, de banda lateral superior y banda lateral inferior.

La opción más simple para la generación de este tipo de señales es mediante filtrado directo; en este caso se genera una señal de doble banda lateral, y posteriormente se elimina mediante filtrado una de las dos bandas laterales:

- BLU de banda lateral superior (BLS): se eliminan las frecuencias $|\omega| < \omega_c$

- BLU de banda lateral inferior (BLI): se eliminan las frecuencias $|\omega| > \omega_c$

El esquema de este método de generación de la señal BLU se muestra en la Figura 3.20. Por conveniencia en la notación, nótese que la señal intermedia $s_D(t)$ es una señal de doble banda lateral pero de amplitud doble (con respecto a la considerada en la sección anterior).

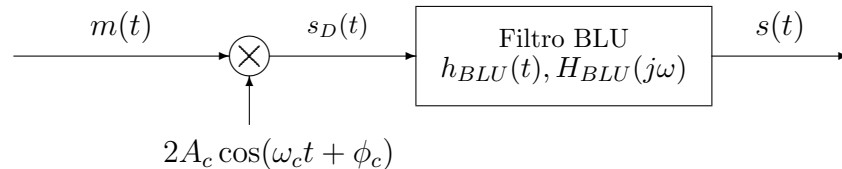


Figura 3.20: Generación de una señal de banda lateral única (BLU) mediante filtrado directo.

La respuesta en frecuencia genérica de los filtros de banda lateral única, para banda lateral superior, y para banda lateral inferior, será

$$H_{BLU}^{sup}(j\omega) = \begin{cases} 1, & \text{si } |\omega| \geq \omega_c \\ 0, & \text{si } |\omega| < \omega_c \end{cases} \text{ y } H_{BLU}^{inf}(j\omega) = \begin{cases} 0, & \text{si } |\omega| > \omega_c \\ 1, & \text{si } |\omega| \leq \omega_c \end{cases}$$

La Figura 3.21 muestra la respuesta en frecuencia de las señales generadas con esta opción.

Existe otra alternativa para la generación de la señal de banda lateral única. Como se demostrará a continuación, la señal BLU tiene la expresión analítica

$$s(t) = A_c \cdot m(t) \cdot \cos(\omega_c t) \mp A_c \cdot \hat{m}(t) \cdot \sin(\omega_c t)$$

donde el signo negativo corresponde a la banda lateral superior, el signo positivo a la banda lateral inferior y $\hat{m}(t)$ es la transformada de Hilbert de la señal $m(t)$. Por simplicidad en la notación se ha asumido una fase nula en las portadoras, $\phi_c = 0$.

La transformada de Hilbert de una señal es la señal que se obtiene al filtrar la señal original con un transformador de Hilbert

$$\hat{m}(t) = m(t) * h_{Hilbert}(t),$$

donde un transformador de Hilbert es un filtro lineal con respuesta al impulso

$$h_{Hilbert}(t) = \frac{1}{\pi t}$$

y respuesta en frecuencia

$$H_{Hilbert}(j\omega) = \begin{cases} -j, & \omega > 0 \\ +j, & \omega < 0 \\ 0, & \omega = 0 \end{cases}$$

Esto permite la generación de la señal utilizando un transformador de Hilbert y dos osciladores en cuadratura (desfasados 90°, un coseno y un seno). Este esquema para la generación, que se muestra en la Figura 3.22, recibe el nombre de modulador de Hartley.

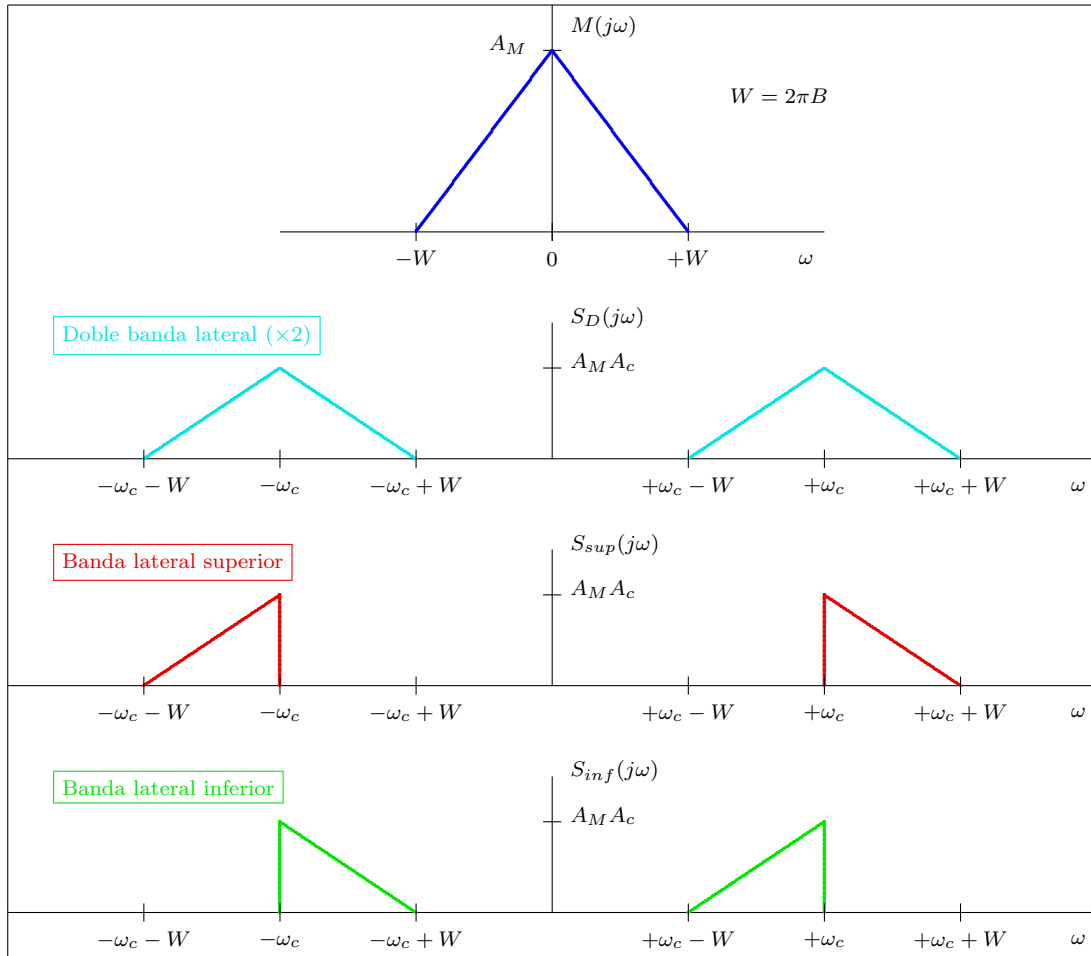


Figura 3.21: Ejemplo de espectro de la señales de banda lateral única, de banda lateral superior y banda lateral inferior, y de la señal de doble banda lateral intermedia que se genera antes del filtrado.

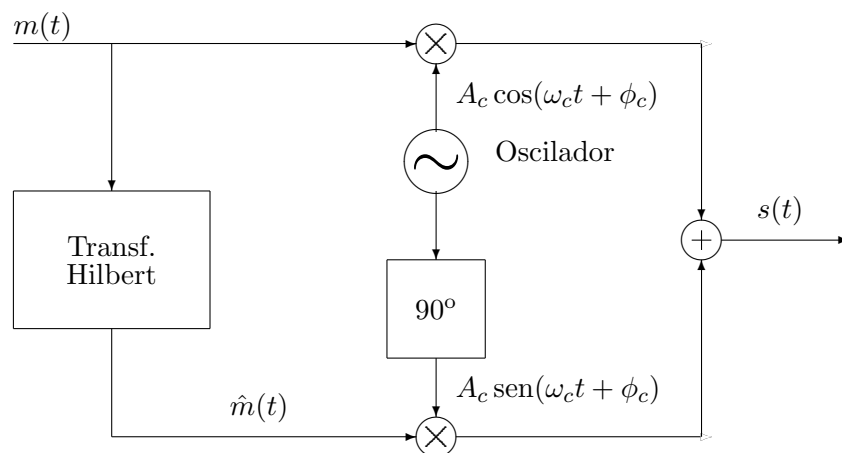


Figura 3.22: Generación de una señal de banda lateral única (BLU) utilizando un transformador de Hilbert y dos portadoras en cuadratura (modulador de Hartley).

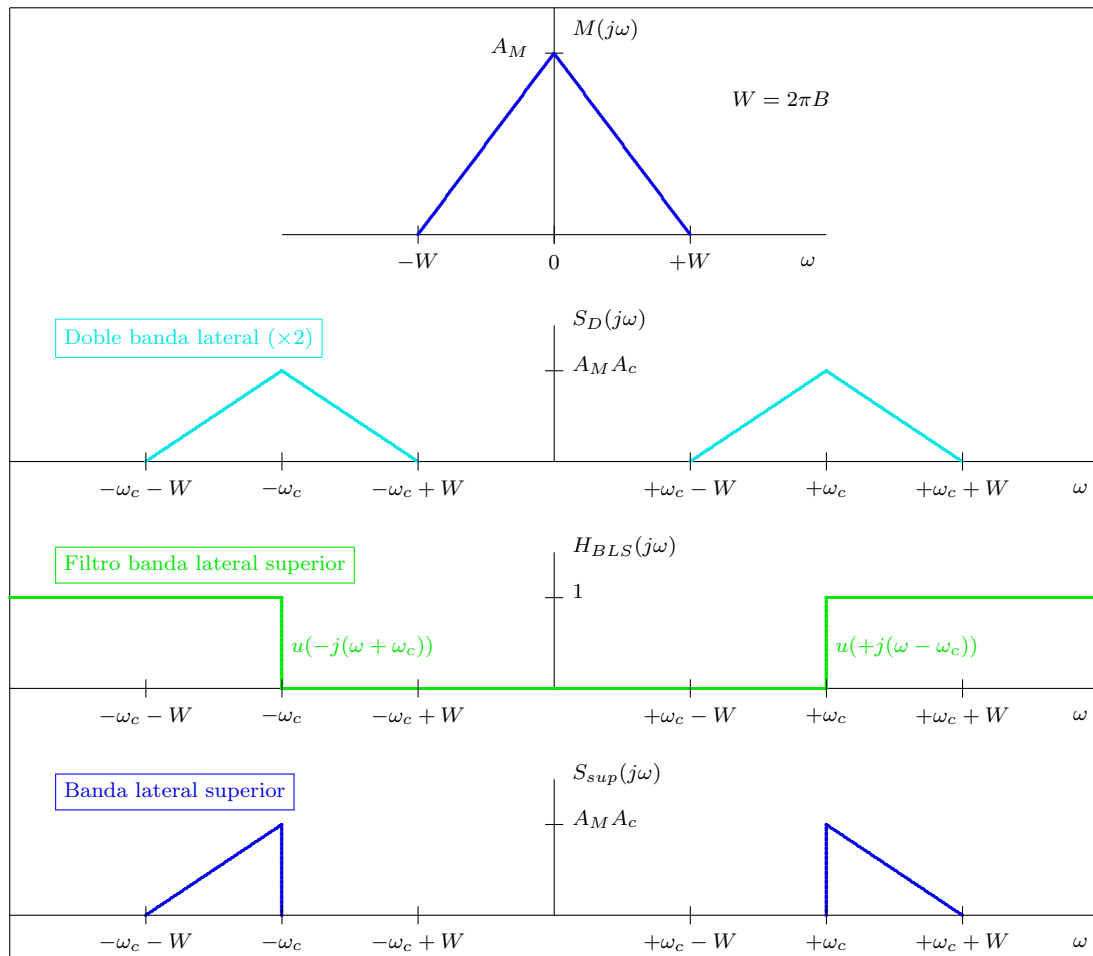


Figura 3.23: Representación de la generación de una señal de banda lateral superior en el dominio de la frecuencia.

Expresión analítica de la señal modulada - Banda lateral superior

La Figura 3.23 muestra la respuesta en frecuencia de las señales generadas con esta opción. En este caso se puede ver que la respuesta en frecuencia del filtro de banda lateral única para banda lateral superior se puede escribir en función de la función escalón $u(x)$, en concreto Respuesta en frecuencia del filtro de banda lateral superior:

$$H_{BLU}^{Sup}(j\omega) = u(\omega - \omega_c) + u(-(\omega + \omega_c))$$

La respuesta en frecuencia de la señal DBL con amplitud doble, $s_D(t)$, es

$$S_D(j\omega) = A_c [M(j\omega - j\omega_c) + M(j\omega + j\omega_c)]$$

A partir de la figura es fácil ver que la respuesta en frecuencia de la señal de banda lateral superior se puede escribir como Espectro de la señal BLU con banda lateral superior

$$\begin{aligned} S(j\omega) &= S_D(j\omega) \cdot H_{BLU}^{Sup}(j\omega) \\ &= A_c M(j\omega) \cdot u(\omega)|_{\omega=\omega-\omega_c} + A_c M(j\omega) \cdot u(-\omega)|_{\omega=\omega+\omega_c}, \end{aligned}$$

es decir, la suma de dos términos: el producto de la respuesta de la moduladora y un escalón desplazados a $+\omega_c$, y el producto de la respuesta de la moduladora y un escalón abatido desplazados a $-\omega_c$, en ambos casos con el factor de escala A_c .

Para obtener la respuesta en el dominio temporal, lo único que hay que hacer es calcular la transformada de Fourier inversa, para lo que resultarán de utilidad las siguientes expresiones (propiedades de la transformada de Fourier y fórmulas de Euler para sinusoides):

$$TF \left[\frac{1}{2} \delta(t) + \frac{j}{2\pi t} \right] = u(\omega),$$

$$TF \left[\frac{1}{2} \delta(t) - \frac{j}{2\pi t} \right] = u(-\omega),$$

$$TF[x(t) \cdot e^{j\omega_c t}] = X(j\omega - j\omega_c)$$

$$\cos(\omega_c t) = \frac{e^{+j\omega_c t} + e^{-j\omega_c t}}{2}$$

$$\text{sen}(\omega_c t) = \frac{e^{+j\omega_c t} - e^{-j\omega_c t}}{2j} = j \frac{e^{-j\omega_c t} - e^{+j\omega_c t}}{2}$$

Teniendo en cuenta estas relaciones, es fácil calcular la transformada de Fourier inversa de $S(j\omega)$

$$\begin{aligned} s(t) &= A_c \cdot m(t) * \left[\frac{1}{2} \delta(t) + \frac{j}{2\pi t} \right] \cdot e^{j\omega_c t} + A_c \cdot m(t) * \left[\frac{1}{2} \delta(t) - \frac{j}{2\pi t} \right] \cdot e^{-j\omega_c t} \\ &= \frac{A_c}{2} \cdot [m(t) + j\hat{m}(t)] \cdot e^{j\omega_c t} + \frac{A_c}{2} \cdot [m(t) - j\hat{m}(t)] \cdot e^{-j\omega_c t} \\ &= A_c \cdot m(t) \cdot \cos(\omega_c t) - A_c \cdot \hat{m}(t) \cdot \text{sen}(\omega_c t). \end{aligned}$$

Expresión analítica de la señal modulada - Banda lateral inferior

Se podría en este caso repetir el procedimiento anterior variando la expresión del filtro de banda lateral única para utilizar el de banda lateral inferior. Pero resulta más sencillo aplicar la propiedad de que la suma de las señales de banda lateral inferior y de banda lateral superior da lugar a la señal de doble banda lateral con amplitud doble

$$s_D(t) = 2 \cdot A_c \cdot m(t) \cdot \cos(\omega_c t) = s_{sup}(t) + s_{inf}(t)$$

tal y como se puede ver en la Figura 3.21. Teniendo esto en cuenta

$$\begin{aligned} s_{inf}(t) &= s_D(t) - s_{sup}(t) \\ &= A_c \cdot m(t) \cdot \cos(\omega_c t) + A_c \cdot \hat{m}(t) \cdot \text{sen}(\omega_c t). \end{aligned}$$

Densidad espectral de potencia

Teniendo en cuenta que las señales se pueden obtener a partir del filtrado de una señal de doble banda lateral con amplitud doble, es trivial obtener las expresiones de la densidad espectral de potencia de una señal de banda lateral única a partir de las expresiones obtenidas en su momento para la modulación de doble banda lateral

- Banda lateral superior

$$S_{S_{sup}}(j\omega) = \begin{cases} A_c^2 \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)], & |\omega| > \omega_c \\ 0, & |\omega| < \omega_c \end{cases}$$

- Banda lateral inferior

$$S_{S_{inf}}(j\omega) = \begin{cases} 0, & |\omega| > \omega_c \\ A_c^2 \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)], & |\omega| < \omega_c \end{cases}$$

La potencia se puede calcular mediante la integral de la densidad espectral de potencia. En ambos casos se cumple que

$$P_S = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_S(j\omega) d\omega = A_c^2 P_M.$$

Demodulación de señales BLU

Para la demodulación óptima de señales de banda lateral única de nuevo es necesaria la utilización de un demodulador síncrono o coherente, como el mostrado en el esquema de la Figura 3.17, en el que $\phi = \phi_c$. Recuerde que por simplicidad en la notación se supuso $\phi_c = 0$. De nuevo se asume una transmisión ideal sin ninguna distorsión, en cuyo caso la señal recibida coincide con la señal modulada que se transmite

$$r(t) = s(t) = A_c \cdot m(t) \cdot \cos(\omega_c t) \mp A_c \cdot \hat{m}(t) \cdot \text{sen}(\omega_c t).$$

La señal demodulada antes de filtrar, $x(t)$, es

$$\begin{aligned} x(t) &= r(t) \times \cos(\omega_c t + \phi) \\ &= [A_c \cdot m(t) \cdot \cos(\omega_c t) \mp A_c \cdot \hat{m}(t) \cdot \text{sen}(\omega_c t)] \times \cos(\omega_c t + \phi) \\ &= \frac{A_c}{2} \cdot m(t) \cdot \cos(\phi) \pm \frac{A_c}{2} \cdot \hat{m}(t) \cdot \text{sen}(\phi) \\ &\quad + \frac{A_c}{2} \cdot m(t) \cdot \cos(2\omega_c t + \phi) \mp \frac{A_c}{2} \cdot \hat{m}(t) \cdot \text{sen}(2\omega_c t + \phi). \end{aligned}$$

El filtrado elimina las componentes de alta frecuencia (términos en $2\omega_c$), con lo que la señal demodulada filtrada queda

$$d(t) = \frac{A_c}{2} \cdot m(t) \cdot \cos(\phi) \pm \frac{A_c}{2} \cdot \hat{m}(t) \cdot \text{sen}(\phi).$$

Si en el generador la fase ϕ_c fuera una fase arbitraria no nula la expresión resultante sería

$$d(t) = \frac{A_c}{2} \cdot m(t) \cdot \cos(\phi - \phi_c) \pm \frac{A_c}{2} \cdot \hat{m}(t) \cdot \text{sen}(\phi - \phi_c).$$

Ahora, el efecto del error de fase es por un lado reducir la amplitud de la señal recibida, que puede incluso desaparecer, tal y como ocurría para una modulación de doble banda lateral, y por otro lado (segundo término) se suma una señal indeseada $\hat{m}(t)$ a la señal recibida, lo que se puede interpretar como un término de distorsión. Por tanto, el uso de un receptor coherente es incluso más necesario que en el caso de la modulación DBL.

La solución es la misma que para la DBL, o bien utilizar un tono piloto o bien utilizar un PLL.

Las Figuras 3.24 y 3.25 muestran una interpretación frecuencial del proceso de demodulación para los casos de banda lateral superior y banda lateral inferior, respectivamente.

Debido a la eficiencia espectral de este método de modulación, se utiliza por ejemplo para la transmisión de canales de voz por cables telefónicos. Con esta modulación se dobla la capacidad, en número de canales, del cable en cuestión.

El principal inconveniente está en su generación. Utilizando la técnica de filtrado directo, para no tener ninguna distorsión en la señal generada se requieren filtros ideales, lo que no es posible en la práctica. Utilizando la técnica basada en el filtro de Hilbert, el inconveniente es que ahora se necesitan dos osciladores en cuadratura, y además una implementación exacta de un transformador de Hilbert, que tampoco es posible. Por tanto, en el propio proceso de generación de la señal habitualmente se generará un cierto grado de distorsión.

3.2.4. Modulación de banda lateral vestigial (BLV)

Para relajar los requerimientos de los filtros necesarios para la modulación BLU, la modulación de banda lateral vestigial permite que una parte, llamada *vestigio*, de la banda no deseada aparezca en la señal modulada. De este modo se facilita la implementación de los filtros al precio de incrementar ligeramente el ancho de banda de la señal.

Para construir esta señal, como en el caso de la señal BLU por filtrado directo, se parte de la una señal de doble banda lateral (de amplitud doble), $s_D(t)$ que se filtra, tal y como se muestra en la Figura 3.26.

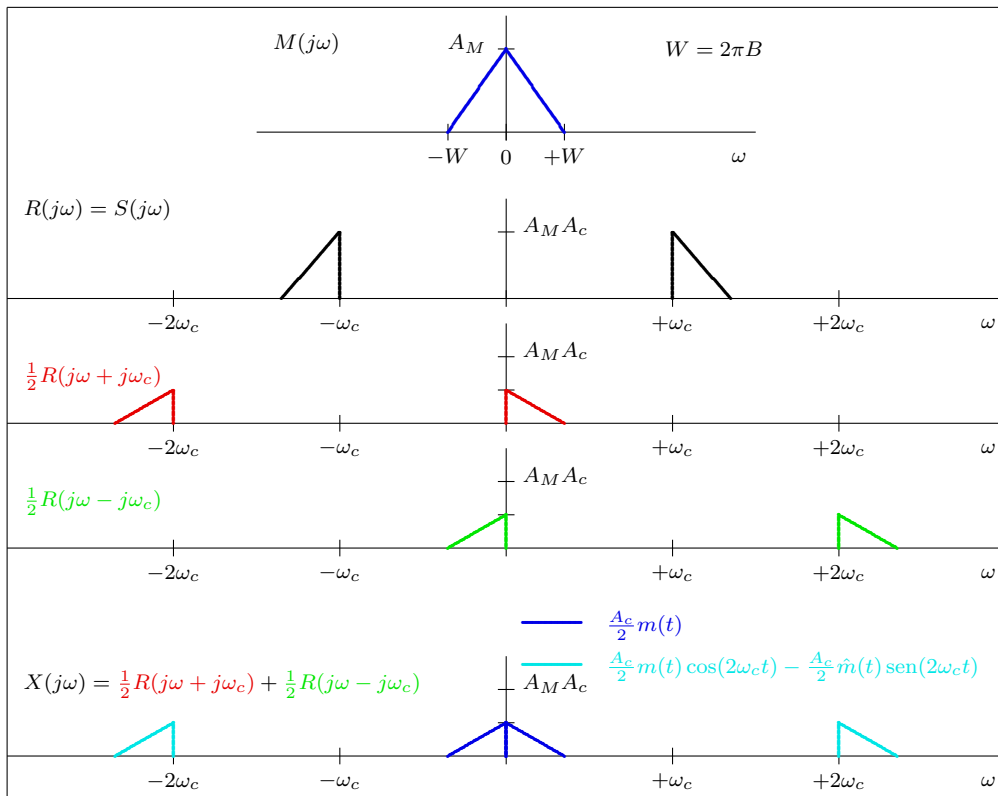


Figura 3.24: Interpretación frecuencial del proceso de demodulación de una señal de banda lateral única (para banda lateral superior).

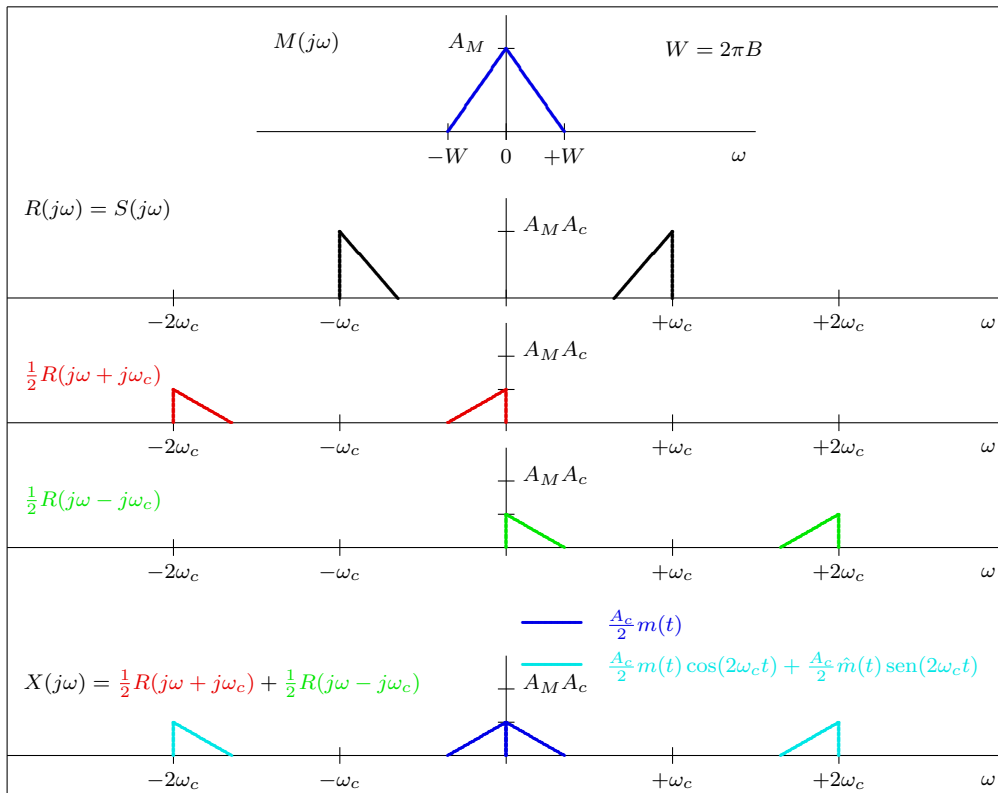


Figura 3.25: Interpretación frecuencial del proceso de demodulación de una señal de banda lateral única (para banda lateral superior).

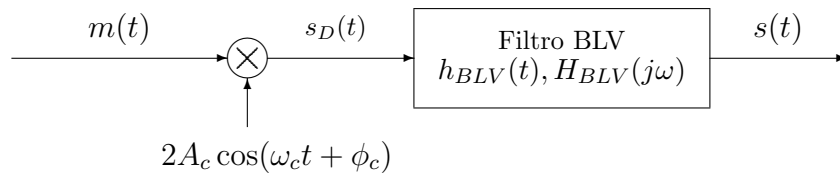


Figura 3.26: Generación de una señal de banda lateral vestigial (BLV) mediante filtrado directo.

La señal modulada de banda lateral vestigial tiene por tanto la expresión analítica

$$s(t) = \left[\underbrace{m(t) \times 2A_c \cos(\omega_c t)}_{s_D(t)} \right] * h_{BLV}(t).$$

En el dominio frecuencial, esta expresión se corresponde con

$$S(j\omega) = A_c \cdot [M(j\omega - j\omega_c) + M(j\omega + \omega_c)] \cdot H_{BLV}(j\omega).$$

Analicemos lo que ocurre en el receptor, si se utiliza un receptor síncrono como el de la Figura 3.17. Se asumirá que la señal se ha transmitido sin distorsión, por lo que la señal recibida será igual a la señal BLV transmitida. En este caso

$$R(j\omega) = S(j\omega) = A_c \cdot [M(j\omega - j\omega_c) + M(j\omega + \omega_c)] \cdot H_{BLV}(j\omega).$$

La señal demodulada (sin filtrar) en el dominio frecuencial es

$$x(t) = r(t) \cdot \cos(\omega_c t) \rightarrow X(j\omega) = \frac{1}{2} \cdot [R(j\omega - j\omega_c) + R(j\omega + j\omega_c)].$$

$$\begin{aligned} X(j\omega) &= \frac{A_c}{2} \cdot [M(j\omega - j2\omega_c) + M(j\omega)] \cdot H_{BLV}(j\omega - \omega_c) \\ &\quad + \frac{A_c}{2} \cdot [M(j\omega) + M(j\omega + j2\omega_c)] \cdot H_{BLV}(j\omega + j\omega_c). \end{aligned}$$

Finalmente, la señal demodulada filtrada en el dominio frecuencial

$$D(j\omega) = \frac{A_c}{2} \cdot M(j\omega) \cdot [H_{BLV}(j\omega - j\omega_c) + H_{BLV}(j\omega + j\omega_c)].$$

Se puede interpretar que esta señal se ha obtenido filtrando la señal modulada con un filtro con respuesta equivalente

$$H_{EQ}(j\omega) = H_{BLV}(j\omega - j\omega_c) + H_{BLV}(j\omega + j\omega_c).$$

Para que no se produzca distorsión, esta respuesta conjunta debe cumplir que en la banda de paso de la señal, para $|\omega| \leq W = 2\pi B$ rad/s, ha de ser ideal; es decir, que su módulo ha de ser constante, y su fase lineal. De aquí se pueden obtener la condición que ha de cumplir el filtro de banda lateral vestigial, que es

$$|H_{BLV}(j\omega - j\omega_c) + H_{BLV}(j\omega + j\omega_c)| = \text{cte}, \text{ en } |\omega| \leq 2\pi B \text{ rad/s}.$$

Idealmente, el valor constante será 1. Esta condición se cumple si la respuesta en frecuencia del filtro de banda lateral vestigial tiene una simetría impar respecto a ω_c en el rango de frecuencias

$\omega_c - \Delta_W < \omega < \omega_c + \Delta_W$, siendo Δ_W el exceso de ancho de banda (vestigio) en radianes/s. En ese caso, el ancho de banda de la señal modulada es

$$BW_{BLV} = B + \Delta_B \text{ Hz, con } \Delta_B = \frac{\Delta_W}{2\pi} \text{ Hz,}$$

donde Δ_B es exceso de ancho de banda vestigial en Hz (habitualmente $\Delta_B \ll B$).

La Figura 3.27 muestra un ejemplo de filtros para banda lateral superior y banda lateral inferior.

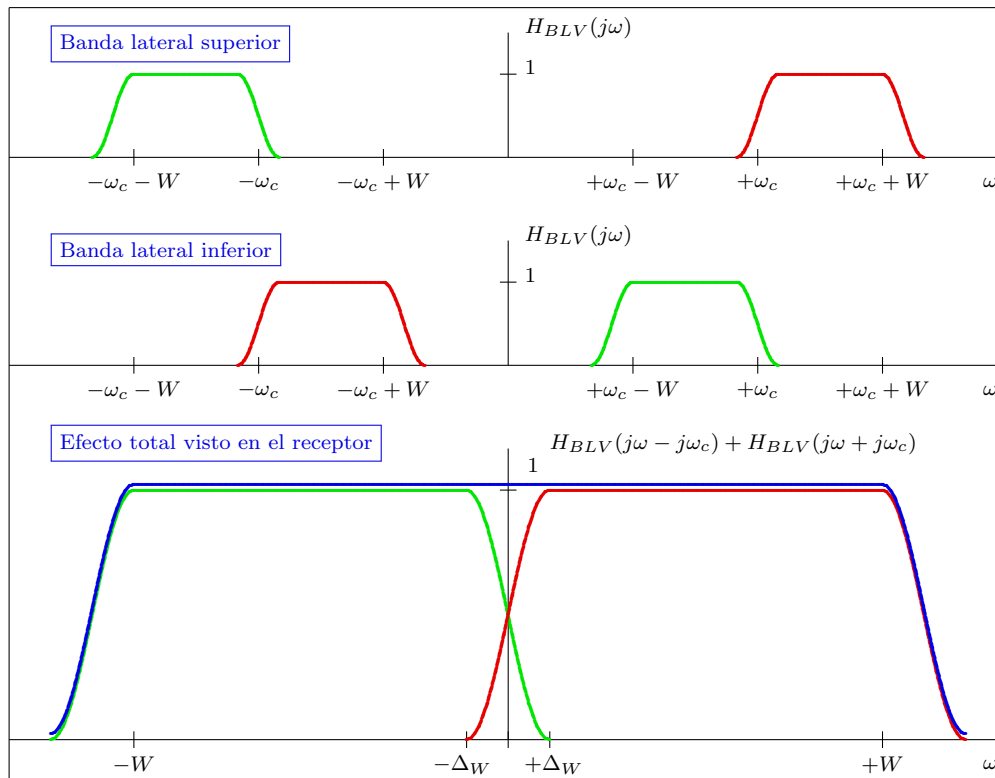


Figura 3.27: Ejemplos de filtros de banda lateral vestigial para banda lateral superior y banda lateral inferior.

Demodulación de señales BLV

Si los filtros cumplen la condición especificada, es posible recuperar la señal sin distorsión utilizando un receptor síncrono. La Figura 3.28 muestra la interpretación frecuencial del proceso de modulación y demodulación para una señal BLV de banda lateral superior.

3.2.5. Resumen de características y comparativa entre las distintas modulaciones de amplitud

La Tabla 3.1 muestra las principales características de las distintas modulaciones de amplitud que se han estudiado. En particular se hace referencia a la potencia de la señal modulada, y cuanta de ella está relacionada con la información transmitida, y al consumo de ancho de banda. Si se considera la eficiencia en potencia y la eficiencia espectral, las conclusiones que se pueden extraer son:

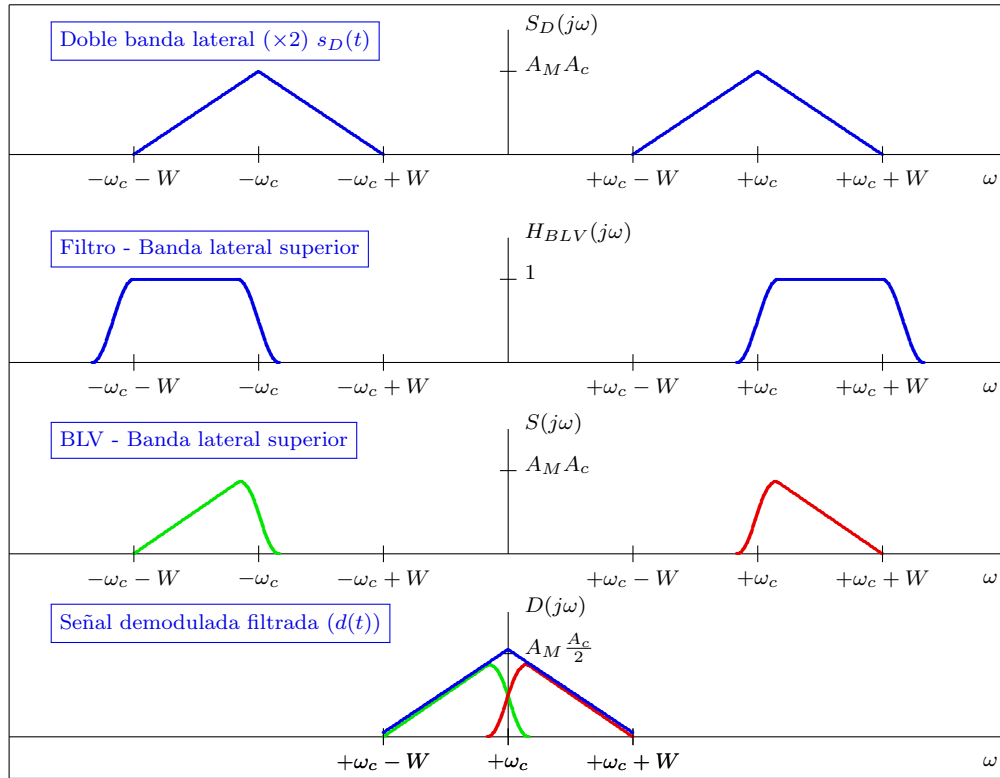


Figura 3.28: Interpretación frecuencial del proceso de modulación y demodulación para una modulación de banda lateral vestigial (banda lateral superior).

Modulación	BW (Hz)	P_S	$P_S(m(t))$	$d(t)$	$P_d(m(t))$
AM convencional	$2B$	$\frac{A_c^2}{2} [1 + P_{M_a}]$	$\frac{A_c^2}{2} P_{M_a}$	$\frac{A_c}{2} [1 + m_a(t)]$	$\frac{A_c^2}{4} P_{M_a}$
DBL	$2B$	$\frac{A_c^2}{2} P_M$	$\frac{A_c^2}{2} P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$
BLU	B	$A_c^2 P_M$	$A_c^2 P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$
BLV	$B + \Delta_B$	$A_c^2 P_M$	$A_c^2 P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$

BW (Hz): ancho de banda de la señal modulada en Hz

P_S : potencia de la señal modulada

$P_S(m(t))$: potencia de la señal modulada relativa al término relacionado con $m(t)$

$d(t)$: señal recuperada con un receptor síncrono o coherente

$P_d(m(t))$: potencia de la señal demodulada relativa al término relacionado con $m(t)$

Tabla 3.1: Comparativa de modulaciones de amplitud.

- Eficiencia en potencia
 - Se entiende como tal que toda la potencia de la señal está relacionada con la transmisión de la información (señal $m(t)$).
 - DBL, BLU y BLV
- Eficiencia espectral
 - Mínimo ancho de banda para la transmisión (mismo ancho de banda que la señal moduladora, B Hz)
 - BLU y BLV (en este caso con un incremento vestigial Δ_B)

En cuanto a los datos correspondientes al demodulador que aparecen en la tabla, se ha considerado siempre que se ha utilizado un receptor síncrono o coherente. Como para el caso de la modulación AM convencional no se había estudiado este receptor, por completitud, se incluye a continuación

Detección síncrona de una modulación AM convencional

En esta sección se estudia la demodulación de una señal AM convencional con un receptor como el de la Figura 3.17. Como siempre, se asume que la señal modulada se transmite sin distorsión, por lo que la señal recibida es igual a la señal modulada que se ha transmitido, que en el caso de una modulación AM convencional es

$$r(t) = s(t) = A_c \cdot [1 + m_a(t)] \cdot \cos(\omega_c t + \phi_c)$$

La señal demodulada sin filtrar $x(t)$ es entonces

$$\begin{aligned} x(t) &= r(t) \times \cos(\omega_c t + \phi) \\ &= A_c [1 + m_a(t)] \cdot \cos(\omega_c t + \phi_c) \times \cos(\omega_c t + \phi) \\ &= \frac{A_c}{2} \cdot [1 + m_a(t)] \cdot \cos(\phi_c - \phi) + \frac{A_c}{2} \cdot [1 + m_a(t)] \cdot \cos(2\omega_c t + \phi_c + \phi) \end{aligned}$$

El filtrado paso bajo elimina los términos de alta frecuencia (términos en $2\omega_c$), con lo que la señal demodulada filtrada es entonces

$$d(t) = \frac{A_c}{2} \cdot [1 + m_a(t)] \cdot \cos(\phi_c - \phi).$$

Por tanto, se revela también para esta modulación la conveniencia de un demodulador síncrono o coherente, con $\phi = \phi_c$, en cuyo caso la señal demodulada vale

$$d(t) = \frac{A_c}{2} \cdot [1 + m_a(t)].$$

3.3. Modulaciones Angulares

En la sección anterior se ha visto la modulación de amplitud de una portadora, en la que la amplitud de la misma se modifica en función de la señal mensaje $m(t)$. A los métodos de

modulación de amplitud se les denominan también *modulaciones lineales*, aunque una señal AM no es lineal en sentido estricto.

Las modulaciones de frecuencia (FM) y de fase (PM) se denominan *modulaciones angulares*. En este caso se modifica la frecuencia o fase de la señal portadora en función de la señal mensaje. Estas modulaciones son claramente no lineales, lo que implica una serie de propiedades:

1. Son más difíciles de implementar (tanto el modulador como el demodulador).
2. Su análisis es en general más complejo. En muchos casos, sólo es posible realizar análisis mediante aproximaciones.

Por otro lado, las modulaciones angulares expanden el espectro de la señal de modo que el ancho de banda de la señal es varias veces mayor que el de la señal. En sentido estricto, estas señales tienen un ancho de banda infinito, pero normalmente se trabaja con el denominado *ancho de banda efectivo*, que es aquel en el que la amplitud del espectro es relevante.

La razón por la que las modulaciones angulares se utilizan, a pesar de estas desventajas, es que tienen la ventaja de que presentan una gran inmunidad frente al ruido. Así pues, lo que hacen estas modulaciones es “cambiar” ancho de banda por inmunidad frente a ruido. Por esta razón, la modulación FM se emplea en sistemas de difusión de música de alta fidelidad o en sistemas de comunicaciones punto a punto en el que la potencia de transmisión está limitada.

3.3.1. Representación de las señales FM y PM

Si se tiene en cuenta que ambas modulaciones modifican, en función de la señal mensaje, el argumento de una portadora sinusoidal, es posible hacer un análisis conjunto de las dos modulaciones y, como se verá posteriormente, existe una clara relación entre ambas.

En general, una modulación angular se puede representar matemáticamente como

$$s(t) = A_c \cdot \cos(\theta(t)),$$

donde teniendo en cuenta que se trata de una señal generada a partir de una portadora de frecuencia f_c Hz ($\omega_c = 2\pi f_c$ rad/s) en general se escribirá

$$\theta(t) = 2\pi f_c t + \phi(t) = \omega_c t + \phi(t).$$

Esta será la representación que se utilizará de forma conjunta para modulaciones PM y FM. El argumento $\theta(t)$ es la fase de la señal en el instante t , y la frecuencia instantánea de la señal, $f_i(t)$, viene dada por su derivada

$$f_i(t) = \frac{1}{2\pi} \frac{d}{dt} \theta(t).$$

Para la expresión particular anterior

$$f_i(t) = f_c + \frac{1}{2\pi} \frac{d}{dt} \phi(t).$$

Si $m(t)$ es la señal moduladora (mensaje a transmitir), la relación de los parámetros de la señal $s(t)$ con la señal moduladora es la siguiente para cada una de los dos tipos de modulación:

- Modulación de fase (PM)

$$\phi(t) = k_p \cdot m(t)$$

k_p : constante de desviación de fase

- Modulación de frecuencia (FM)

$$\Delta f_i(t) = f_i(t) - f_c = \frac{1}{2\pi} \cdot \frac{d}{dt} \phi(t) = k_f \cdot m(t)$$

k_f : constante de desviación de frecuencia

Las constantes k_p y k_f son las *constantes de desviación* de fase y frecuencia, respectivamente.

Hay una relación entre ambos métodos de modulación, que se puede ver fácilmente si se escriben las expresiones de $\phi(t)$ y $\frac{d}{dt} \phi(t)$ en PM y FM

$$\phi(t) = \begin{cases} k_p \cdot m(t), & \text{PM} \\ 2\pi \cdot k_f \cdot \int_{-\infty}^t m(\tau) d\tau, & \text{FM} \end{cases}$$

$$\frac{d}{dt} \phi(t) = \begin{cases} k_p \cdot \frac{d}{dt} m(t), & \text{PM} \\ 2\pi \cdot k_f \cdot m(t), & \text{FM} \end{cases}$$

Se puede ver que modular en FM una portadora con la integral de un mensaje es equivalente a modular en PM la portadora con el mensaje original. Es decir, si se modula en frecuencia (FM) con la derivada de un mensaje, se obtiene la misma señal que si se modula en fase (PM) con el mensaje original. La Figura 3.29 representa la relación entre ambas modulaciones

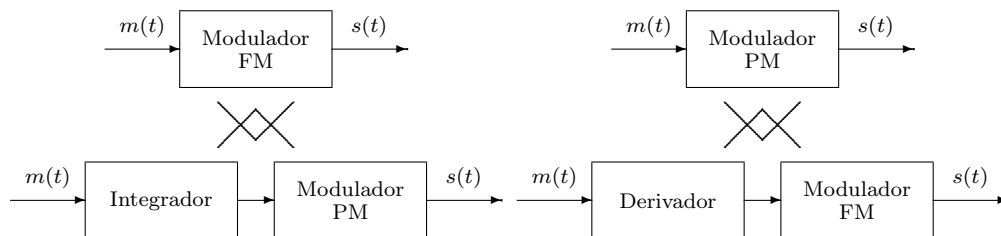


Figura 3.29: Relación entre las modulaciones FM y PM.

Esta relación entre ambos tipos de modulación permite analizarlas en común para luego destacar sus diferencias. La forma de onda de una modulación angular es la de una senoide de amplitud constante, cuya información angular se modifica a lo largo del tiempo, lo que visualmente se traduce en que cambia la longitud entre pasos por cero de la señal. La Figura 3.30 muestra un ejemplo de señal moduladora, la señal portadora, y la señal modulada en FM resultante.

El efecto de las constantes de desviación de frecuencia o fase, es ponderar el grado de desviación de la señal modulada respecto de la señal portadora de frecuencia y fase constantes. Las Figuras 3.31 y 3.32 muestran la señal modulada para una modulación de fase PM para dos valores distintos de constante de desviación de fase, $k_p = 2\pi \times \frac{1}{4}$ y $k_p = 2\pi \times \frac{3}{4}$, respectivamente.

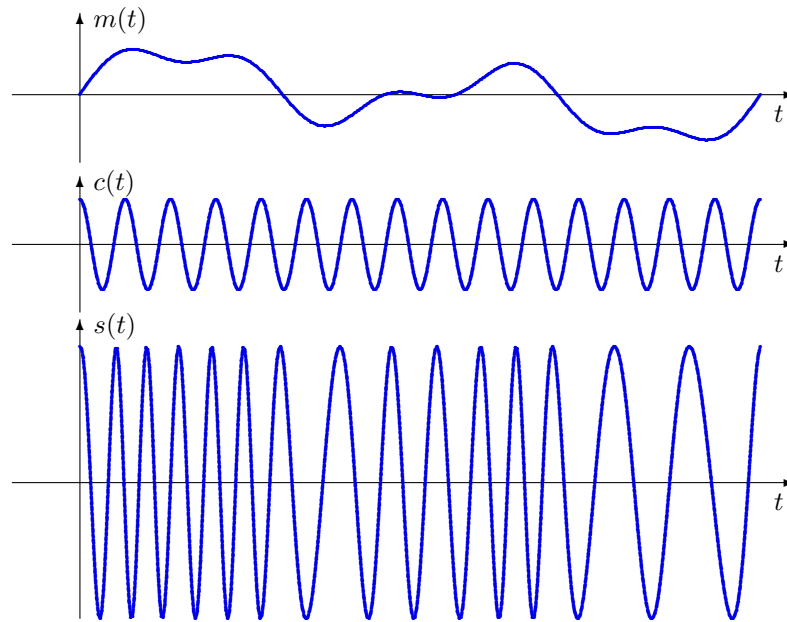


Figura 3.30: Forma de onda de una modulación angular (en este caso FM) para un ejemplo de señal moduladora.

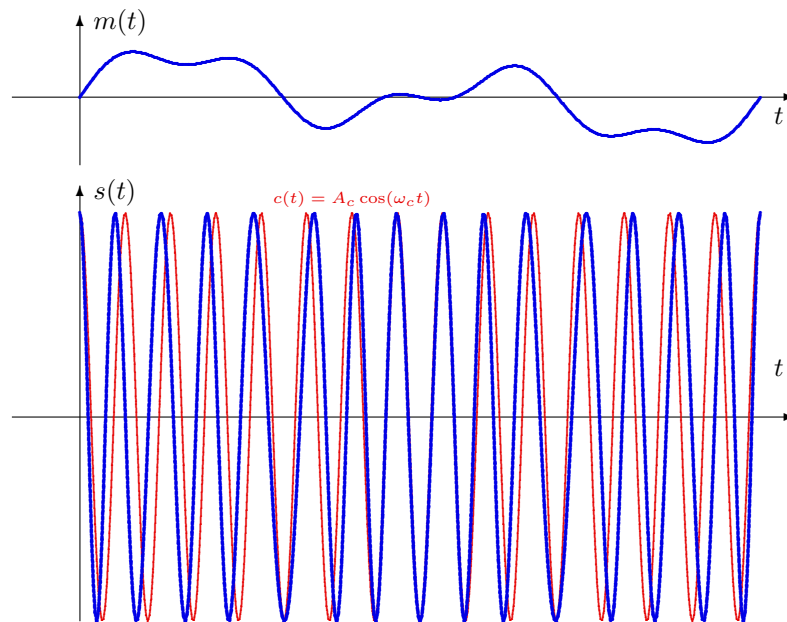


Figura 3.31: Señal modulada para una modulación PM con $k_p = 2\pi \times \frac{1}{4}$ para un ejemplo de señal moduladora.

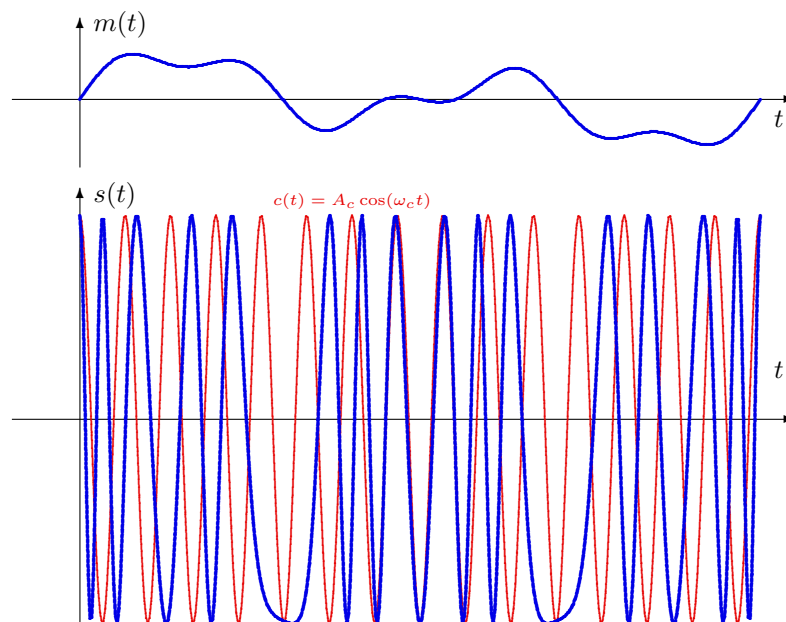


Figura 3.32: Señal modulada para una modulación PM con $k_p = 2\pi \times \frac{3}{4}$ para un ejemplo de señal moduladora.

Se puede ver cómo con respecto a la señal portadora, cuando la señal moduladora toma valores positivos, la señal modulada se adelanta (al incrementarse el término de fase proporcionalmente a $m(t)$), y se retrasa cuando toma valores negativos (al reducirse el término de fase en este caso). La magnitud del adelanto o retardo en un instante dado, es proporcional al parámetro k_p .

Las Figuras 3.33 y 3.34 muestran la señal modulada para una modulación de fase PM para dos valores distintos de constante de desviación de fase, $k_p = 2\pi \times \frac{1}{4}$ y $k_p = 2\pi \times \frac{3}{4}$, respectivamente.

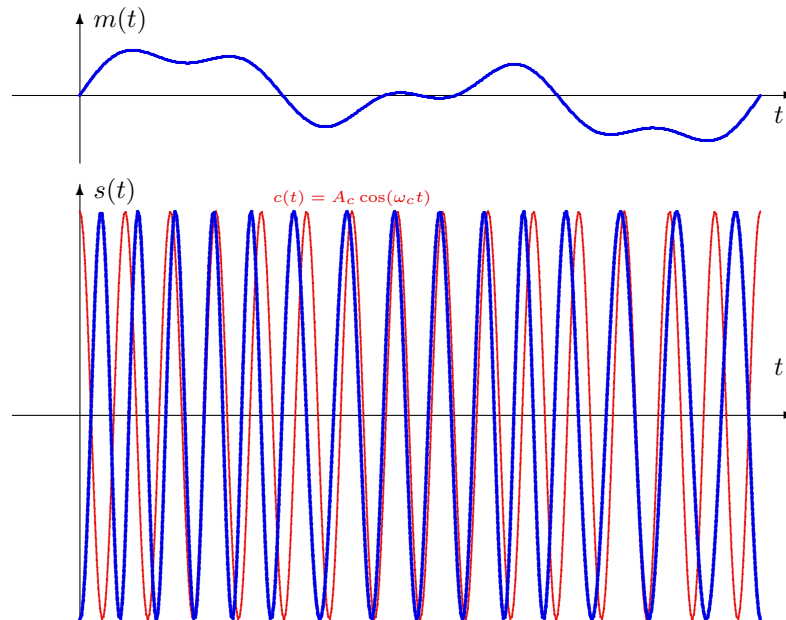


Figura 3.33: Señal modulada para una modulación FM con $k_f = 2\pi \times \frac{1}{4}$ para un ejemplo de señal moduladora.

Ahora, la señal modulada se adelanta frente a la portadora cuando la integral de la señal es positiva, y se retarda cuando es negativa, ya que ahora es la integral de la señal moduladora la que modifica el término de fase. De nuevo, la magnitud de la desviación con respecto a la portadora en un instante, es proporcional al valor de la constante de desviación, en este caso de frecuencia, k_f .

3.3.2. Índices de modulación

Un parámetro importante de una modulación angular es el índice de modulación, ya que varios aspectos de la modulación, como el ancho de banda o la inmunidad frente al ruido, dependen de su valor. Los índices de modulación se definen a partir de las constantes de desviación, de fase o frecuencia, respectivamente. En una señal modulada en fase, la máxima desviación en fase de la señal es

$$\Delta\phi_{\text{máx}} = k_p \cdot \text{máx}(|m(t)|).$$

Del mismo modo, en una señal modulada en frecuencia, la máxima desviación de frecuencia de la señal es

$$\Delta f_{\text{máx}} = k_f \cdot \text{máx}(|m(t)|).$$

A partir de estas desviaciones máximas, se definen los índices de modulación de una modulación PM y de una modulación FM, respectivamente, como

$$\beta_p = \Delta\phi_{\text{máx}} = k_p \cdot \text{máx}(|m(t)|) = k_p \cdot C_M$$

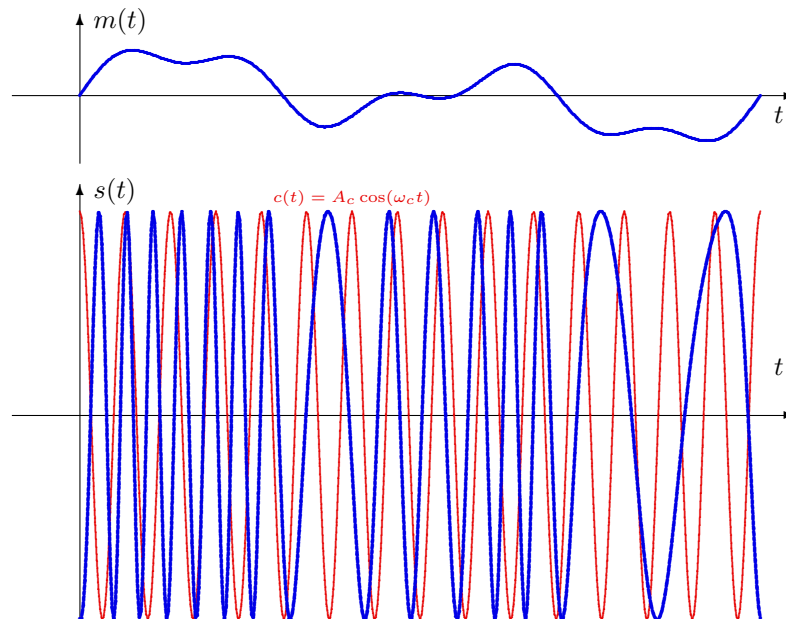


Figura 3.34: Señal modulada para una modulación FM con $k_f = 2\pi \times \frac{3}{4}$ para un ejemplo de señal moduladora.

$$\beta_f = \frac{\Delta f_{\text{máx}}}{B} = \frac{k_f \text{máx}(|m(t)|)}{B} = \frac{k_f \cdot C_M}{B},$$

donde B es el ancho de banda en Hz de la señal moduladora $m(t)$, y C_M es el valor máximo del módulo de la misma, lo que define su rango: $-C_M \leq m(t) \leq +C_M$.

3.3.3. Características espectrales de una modulación angular

Debido a la no linealidad de las modulaciones angulares, en muchos casos la caracterización precisa de su espectro no se puede tratar de forma matemática estricta. Normalmente se estudia para señales moduladoras sencillas y se hacen ciertas aproximaciones. A continuación se ven algunos casos particulares: la modulación de banda estrecha, las modulaciones mediante una señal sinusoidal y mediante una señal periódica, y finalmente se discutirá el caso de una señal moduladora arbitraria no periódica.

Modulación angular de banda estrecha

Una modulación angular se denomina de banda estrecha si las constantes k_p o k_f y la señal $m(t)$ son tales que se cumple que

$$\phi(t) \ll 1.$$

Esto implica valores pequeños de las constantes de desviación. En este caso, dado que la señal modulada tiene la expresión general

$$s(t) = A_c \cdot \cos(\omega_c t + \phi(t)),$$

y teniendo en cuenta la relación trigonométrica

$$\cos(A \pm B) = \cos(A) \cos(B) \mp \sin(A) \sin(B),$$

se puede expandir la expresión de la señal modulada y hacer la siguiente aproximación

$$\begin{aligned} s(t) &= A_c \cdot \cos(\omega_c t) \cdot \cos \phi(t) - A_c \cdot \sin(\omega_c t) \cdot \sin \phi(t) \\ &\approx A_c \cdot \cos(\omega_c t) - A_c \cdot \phi(t) \cdot \sin(\omega_c t). \end{aligned}$$

Se han considerado las siguientes aproximaciones para $\phi(t)$ pequeño

$$\cos(\phi(t)) \approx 1, \quad \sin(\phi(t)) \approx \phi(t).$$

Esta ecuación es muy similar a la expresión de una señal AM convencional, donde aparece un signo negativo y un seno en lugar de un coseno, y en lugar de la señal $m(t)$ tenemos la fase $\phi(t)$, que para una PM es proporcional a $m(t)$ y para una FM es proporcional a la integral de $m(t)$. Recuerde que

$$\text{AM Convencional: } s(t) = A_c \cdot \cos(\omega_c t) + A_c \cdot m_a(t) \cdot \cos(\omega_c t)$$

$$\text{Recuerde también que } \phi(t) = \begin{cases} k_p \cdot m(t) & \text{PM} \\ 2\pi \cdot k_f \cdot \int_{-\infty}^t m(\tau) d\tau & \text{FM} \end{cases}$$

Para este tipo de modulaciones el espectro de la señal es muy similar al de una señal AM, al menos en cuanto a su ancho de banda. Hay que tener en cuenta que el ancho de banda de la integral de una señal es el mismo que el de la propia señal (la respuesta en frecuencia de un derivador es $j\omega$). Por tanto el espectro tendrá las siguientes componentes y características:

- Dos deltas, situadas en $\pm\omega_c$ (espectro portadora)
- Réplicas del espectro de $\phi(t)$ situadas en $\omega = \pm\omega_c$
- Forma del espectro de $\phi(t)$
 - PM: proporcional al espectro de $m(t)$

$$\phi(t) = k_p \cdot m(t) \leftrightarrow \Phi(j\omega) = k_p \cdot M(j\omega)$$

- FM: proporcional al espectro de la integral de $m(t)$

$$\phi(t) = 2\pi \cdot k_f \cdot \int_{-\infty}^t m(\tau) d\tau \leftrightarrow \Phi(j\omega) = 2\pi k_f \cdot \frac{M(j\omega)}{j\omega}$$

Así pues, el ancho de banda de estas señales de banda estrecha es

$$BW_{Angular-BE} \approx 2B.$$

Modulación mediante una señal sinusoidal

En esta sección se analizará el caso en que la señal moduladora es una señal sinusoidal, de amplitud a y frecuencia ω_m rad/s. Por conveniencia, para poder analizar simultáneamente las modulaciones PM y FM, la expresión analítica de la señal será

$$m(t) = \begin{cases} a \cdot \sin(\omega_m t) & \text{para modulación PM} \\ a \cdot \cos(\omega_m t) & \text{para modulación FM} \end{cases}$$

De esta forma, la expresión analítica de la señal moduladora coincide para las dos variantes. Los índices de modulación de una modulación PM y FM son

$$\beta_p = \Delta\phi_{\text{máx}} = k_p \cdot \text{máx}(|m(t)|) = k_p \cdot C_M = k_p \cdot a$$

$$\beta_f = \frac{\Delta f_{\text{máx}}}{B} = \frac{k_f \cdot \text{máx}(|m(t)|)}{B} = \frac{k_f \cdot C_M}{B} = k_f \cdot a \cdot \frac{2\pi}{\omega_m}$$

Por tanto, las expresiones del término de fase $\phi(t)$ son

- Expresiones de $\phi(t)$ para PM

$$\phi(t) = k_p \cdot m(t) = k_p \cdot a \cdot \sin(\omega_m t) = \beta_p \sin(\omega_m t)$$

- Expresiones de $\phi(t)$ para FM

$$\phi(t) = 2\pi \cdot k_f \cdot \int_{-\infty}^t m(\tau) d\tau = 2\pi \cdot k_f \cdot a \cdot \frac{1}{\omega_m} \cdot \sin(\omega_m t) = \beta_f \sin(\omega_m t)$$

Lo que significa que la expresión de la señal modulada es común para ambos tipos de modulación

$$s(t) = A_c \cos(\omega_c t + \phi(t)) = A_c \cdot \cos(\omega_c t + \beta \cdot \sin(\omega_m t))$$

donde β es el índice de modulación de la modulación correspondiente, β_p o β_f . Como un coseno es la parte real de una exponencial compleja, la señal modulada se puede reescribir como

$$s(t) = \text{Re} (A_c \cdot e^{j\omega_c t} \cdot e^{j\beta \cdot \sin(\omega_m t)})$$

La función $e^{j\beta \cdot \sin(\omega_m t)}$ es periódica de frecuencia $f_m = \frac{\omega_m}{2\pi}$ Hz. Esto implica que admite un Desarrollo en serie de Fourier, de la forma

$$e^{j\beta \cdot \sin(\omega_m t)} = \sum_{n=-\infty}^{\infty} J_n(\beta) \cdot e^{j(n \cdot \omega_m)t}.$$

El coeficiente de índice n del desarrollo en serie, que se denota aquí como $J_n(\beta)$, es en este caso la función de Bessel de primera especie de orden n y argumento β . Una expansión en serie de la función de Bessel es la siguiente:

$$J_n(\beta) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{\beta}{2}\right)^{n+2k}}{k!(k+n)!}$$

Para valores pequeños de β , se puede utilizar la aproximación

$$J_n(\beta) \approx \frac{\beta^n}{2^n n!}.$$

Se puede ver que para valores pequeños de β únicamente son significativos los primeros términos (en general se puede limitar a $n = 1$ o $n = 2$).

Las funciones de Bessel cumplen además las siguientes propiedades de simetría

$$J_{-n}(\beta) = \begin{cases} J_n(\beta), & n \text{ par} \\ -J_n(\beta), & n \text{ impar} \end{cases}.$$

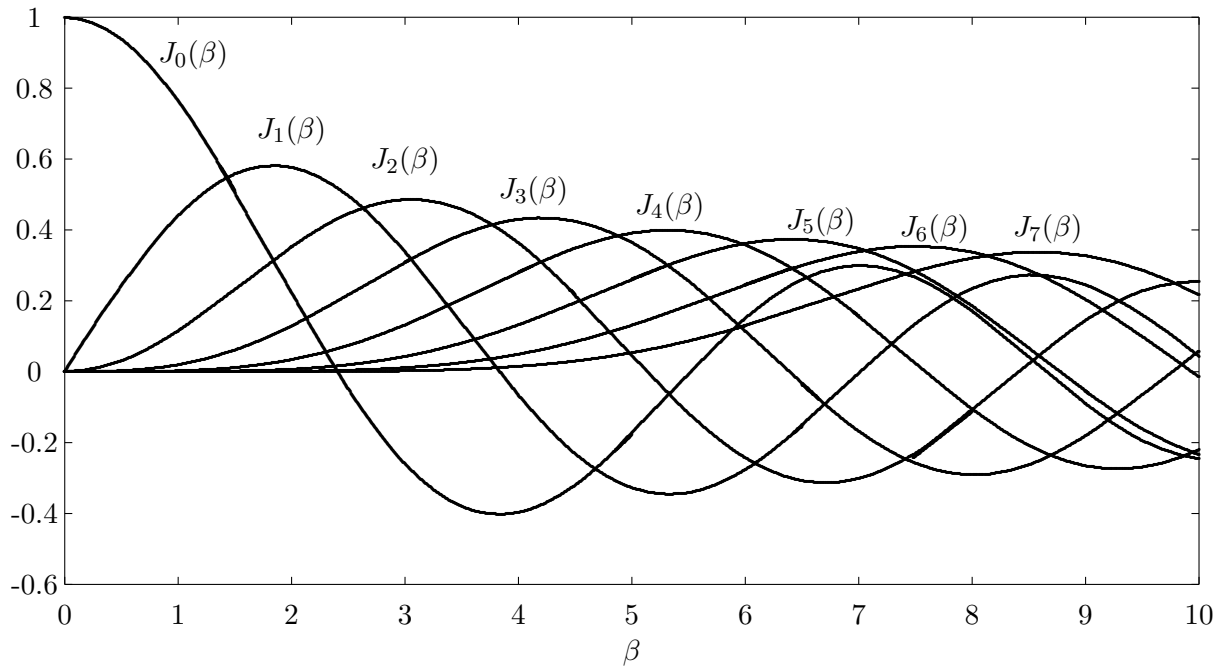


Figura 3.35: Funciones de Bessel, $J_n(\beta)$, para distintos valores de n .

La Figura 3.35 representa las funciones de Bessel, en función de β , para varios valores de n .

Utilizando el desarrollo en serie de Fourier, se puede obtener una expresión alternativa de la señal modulada

$$\begin{aligned}
 s(t) &= \text{Re} \left(A_c \cdot e^{j\omega_c t} \sum_{n=-\infty}^{\infty} J_n(\beta) \cdot e^{j(n\omega_m)t} \right) = \text{Re} \left(\sum_{n=-\infty}^{\infty} A_c \cdot J_n(\beta) \cdot \underbrace{e^{j\omega_c t} \cdot e^{j(n\omega_m)t}}_{e^{j(\omega_c + n\omega_m)t}} \right) \\
 &= \sum_{n=-\infty}^{\infty} A_c \cdot J_n(\beta) \cdot \cos((\omega_c + n \cdot \omega_m) \cdot t)
 \end{aligned}$$

La señal modulada se puede expresar como una suma infinita de sinusoides (cosenos) con las siguientes características (teniendo en cuenta que el índice del sumatorio es n):

- Frecuencias de las sinusoides

$$\text{Frecuencias (Hz)} : f_c + n \cdot f_m, \quad \text{para } n = 0, \pm 1, \pm 2, \dots$$

$$\text{Pulsación (rad/s)} : \omega_c + n \cdot \omega_m, \quad \text{para } n = 0, \pm 1, \pm 2, \dots$$

- Amplitud de la senoide de pulsación $\omega_c + n \cdot \omega_m$

$$A_c \cdot J_n(\beta)$$

Esto quiere decir que el ancho de banda es teóricamente infinito. Sin embargo, el ancho de banda de la señal modulada no es infinito, ya que la amplitud de las componentes para valores altos de n es muy pequeña. Por ello se define un *ancho de banda efectivo* finito. En general, se define el ancho de banda efectivo como aquel que contiene al menos el 98% de la potencia de la señal. En este caso, este ancho de banda efectivo es

$$B_e = 2 \cdot (\beta + 1) \cdot f_m \text{ Hz}$$

Veamos como afecta la señal moduladora al espectro de la señal. En este caso, suponiendo las mismas señales moduladoras sinusoidales de amplitud a y frecuencia f_m

$$B_e = 2(\beta + 1)f_m = \begin{cases} 2(k_p a + 1)f_m, & \text{PM} \\ 2\left(\frac{k_f a}{f_m} + 1\right)f_m, & \text{FM} \end{cases},$$

o lo que es lo mismo

$$B_e = 2(\beta + 1)f_m = \begin{cases} 2(k_p a + 1)f_m, & \text{PM} \\ 2(k_f a + f_m), & \text{FM} \end{cases}.$$

Por otro lado, el número total de armónicos en el ancho de banda efectivo B_e es

$$M_e = 2[\beta] + 3 = \begin{cases} 2[k_p a] + 3, & \text{PM} \\ 2\left[\frac{k_f a}{f_m}\right] + 3, & \text{FM} \end{cases}$$

Estas expresiones muestran que el aumento de la amplitud de la señal, a , tiene prácticamente el mismo efecto en ambas modulaciones, incrementa el ancho de banda. Por otro lado, el incremento de la frecuencia, f_m , también incrementa el ancho de banda, pero el efecto es mucho mayor en la modulación PM que en la FM. En la PM el incremento es multiplicativo mientras que en la FM es aditivo.

Un incremento de la amplitud incrementa el número de armónicos en el ancho de banda de la señal. Sin embargo, el aumento de la frecuencia, para la PM, no varía el número de armónicos, mientras que en la FM lo decreta. Esto explica la relativa insensibilidad del ancho de banda de la FM respecto a la frecuencia de la señal. Por un lado, aumentar f_m aumenta el espacio entre los armónicos, pero por otro lado su número es cada vez menor.

A continuación se verá un ejemplo de la forma del espectro. En este caso, se hará para una modulación con índice de modulación $\beta = 5$. En este caso, la señal modulada es una suma de sinusoides de amplitud $A_c \cdot J_n(\beta)$ y pulsaciones $\omega_c + n \cdot \omega_m$. Los valores de las funciones de Bessel para el valor de β son

$$J_0(5) = -0,18, \quad J_1(5) = -0,32, \quad J_2(5) = 0,05, \quad J_3(5) = 0,37, \quad J_4(5) = 0,39, \quad J_5(5) = 0,26, \dots$$

La Figura 3.36 muestra la forma de la respuesta en frecuencia para frecuencias positivas. Hay que tener en cuenta que la respuesta en frecuencia de un coseno es una delta con amplitud π veces la amplitud del coseno. La propiedad de las funciones de Bessel de que $J_{-n}(\beta)$ sea igual a $J_{+n}(\beta)$ o a dicho valor cambiado de signo, dependiendo de que n sea par o impar, proporciona esa peculiar simetría de la respuesta en frecuencia con respecto a la frecuencia de la portadora ω_c .

Modulación mediante una señal periódica

En este caso se llega, mediante un desarrollo similar a la misma conclusión que en el caso de una señal sinusoidal, también se van a tener únicamente frecuencias de la forma $f_c + n f_m$. Una señal periódica admite un desarrollo en serie de Fourier, de forma que se puede expresar como una suma de sinusoides de frecuencias múltiplo de la que define el período. Por tanto, las frecuencias en el espectro de la señal

$$f_c \pm n \cdot f_m \quad \text{ó} \quad (\omega_c \pm n \cdot \omega_m)$$

Las amplitudes de cada frecuencia vendrán dadas por la suma de las contribuciones de todos los armónicos.

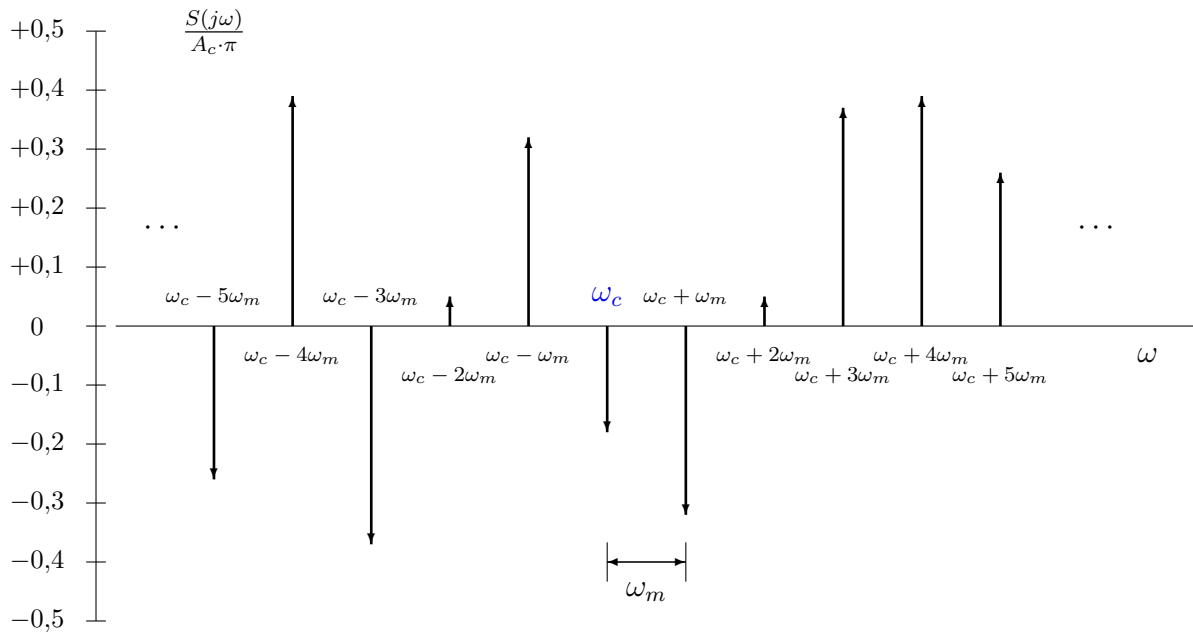


Figura 3.36: Respuesta en frecuencia de una modulación angular para un índice de modulación $\beta = 5$ con una moduladora sinusoidal de frecuencia ω_m rad/s.

Modulación mediante una señal determinista no periódica

Al análisis de este caso es muy complejo debido a la no linealidad de las modulaciones. Sin embargo, existe una regla heurística que proporciona el valor aproximado del ancho de banda de la señal modulada. Es la denominada *regla de Carson*, que dice que el ancho de banda de la señal modulada cuando la señal moduladora tiene un ancho de banda de B Hz, es aproximadamente

$$BW_{Carson} \approx 2 \cdot (\beta + 1) \cdot B \text{ Hz.}$$

Teniendo en cuenta que en modulaciones FM de banda ancha el valor de β está normalmente alrededor de 5 o incluso en valores mayores, este ancho de banda es mucho mayor que el ancho de banda de los esquemas de modulación AM, que era B (BLV o BLU) o $2B$ (DBL y AM convencional).

3.3.4. Modulación de las señales FM y PM

Cualquier proceso de modulación o demodulación, tanto para modulación de amplitud como para modulaciones angulares, implica la generación de frecuencias que no se encuentran en la señal mensaje original. Esto significa que tanto un modulador como un demodulador no se pueden modelar mediante un sistema lineal e invariante, ya que un sistema lineal e invariante no produce nuevas frecuencias, frecuencias no presentes en la señal de entrada del sistema, en la salida del mismo.

Los moduladores angulares son, en general, sistemas no lineales y variantes en el tiempo. Un método para generar una señal de FM de forma directa es diseñar un oscilador cuya tensión varía con una tensión de entrada. Este tipo de osciladores se denominan *osciladores controlados por tensión* y se suelen denotar con las siglas inglesas VCO. Un VCO se puede implementar de varias formas:

1. Mediante un *diodo varactor*: un varactor es un elemento cuya capacidad varía con la tensión que se le aplica. Por tanto, si tal dispositivo se utiliza en el diseño de un oscilador, la frecuencia del mismo varía en función de la capacidad y por tanto de la tensión aplicada.

$$C(t) = C_0 + k_0 m(t).$$

Cuando $m(t) = 0$, la frecuencia del circuito sintonizado es

$$f_c = \frac{1}{2\pi\sqrt{L_0 C_0}}.$$

En general, para $m(t)$ no nulo se tiene

$$\begin{aligned} f_i(t) &= \frac{1}{2\pi\sqrt{L_0(C_0 + k_0 m(t))}} \\ &= \frac{1}{2\pi\sqrt{L_0 C_0}} \frac{1}{\sqrt{1 + \frac{k_0}{C_0} m(t)}} \\ &= f_c \frac{1}{\sqrt{1 + \frac{k_0}{C_0} m(t)}}. \end{aligned}$$

Asumiendo que

$$\varepsilon = \frac{k_0}{C_0} m(t) \ll 1,$$

y utilizando las aproximaciones

$$\sqrt{1 + \varepsilon} \approx 1 + \frac{\varepsilon}{2}, \quad \frac{1}{1 + \varepsilon} \approx 1 - \varepsilon,$$

se obtiene la expresión

$$f_i(t) \approx f_c \left(1 - \frac{k_0}{2C_0} m(t) \right).$$

2. Utilizando un *tubo de reactancia*: un tubo de reactancia es un dispositivo cuya inductancia varía con la tensión aplicada. Se puede hacer un análisis similar al del diodo varactor.

Otra posibilidad para la generación de señales de una modulación angular es el denominado *método indirecto*. En este caso, el proceso se divide en dos partes:

1. Se genera una modulación angular de banda estrecha. Debido a la relación con la modulación AM esto es sencillo de realizar.
2. El segundo paso es generar la señal de banda ancha a partir de la señal de banda estrecha.

Si la señal de banda estrecha es

$$s_{be}(t) = A_c \cos(2\pi f_c t + \phi(t)),$$

la salida del multiplicador de frecuencia es

$$y(t) = A_c \cos(2\pi n f_c t + n\phi(t)).$$

Finalmente, para ajustar la frecuencia portadora deseada, se multiplica por un oscilador local

$$s(t) = A_c \cos(2\pi(n f_c - f_{OL})t + n\phi(t)).$$

3.3.5. Demodulación de las señales FM y PM

La demodulación de una señal FM consiste en encontrar la frecuencia instantánea de la señal modulada $s(t)$, y a continuación restarle la frecuencia de la portadora, ya que

$$m(t) = \frac{f_i(t) - f_c}{k_f}.$$

En cuanto a una señal PM, la demodulación consiste en encontrar la fase de la señal, ya que

$$m(t) = \frac{\phi(t)}{k_p}.$$

En general un demodulador FM se puede implementar mediante un conversor de FM a AM y luego utilizar un demodulador AM, tal y como muestra la Figura 3.37.

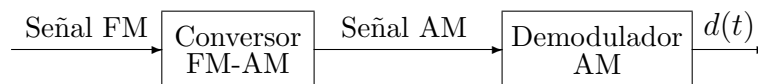


Figura 3.37: Esquema general de un demodulador FM.

La conversión FM-AM se puede hacer de muchas formas:

1. Un derivador

$$|H(j\omega)| = \omega.$$

2. Característica en frecuencia de un circuito sintonizado. Se aprovecha la parte lineal de dicha respuesta. Este sistema es fácil de implementar, aunque la zona lineal puede ser pequeña. Cuando la zona lineal es demasiado pequeña, se pueden utilizar dos circuitos sintonizados a dos frecuencias y combinarlos con el denominado *discriminador balanceado*.

Estos métodos, tienen el inconveniente de que el ancho de banda de la señal AM generada en el paso intermedio es igual al ancho de banda equivalente, B_e , de la modulación FM, con lo que el ruido correspondiente es el ruido contenido en esa banda que es generalmente mayor que B .

Para minimizar el efecto del ruido, y sobre todo del umbral de ruido que veremos posteriormente, se utilizan varias alternativas:

1. Demodulador de FM con realimentación.
2. Demodulador mediante lazo enganchado en fase (PLL).

3.4. Ruido en sistemas de comunicaciones analógicas

Anteriormente se han estudiado las propiedades de potencia y de ancho de banda de las distintas modulaciones analógicas, de amplitud y angulares. En este apartado se analiza el efecto del ruido en estas modulaciones. En todos los casos se van a considerar las siguientes premisas:

- La señal moduladora $m(t)$ tiene un ancho de banda B Hz.
- Se asume una transmisión ideal, o transmisión sobre canal gaussiano, donde la señal transmitida $s(t)$ no sufre ninguna distorsión lineal y el único efecto producido durante la transmisión es la suma de ruido térmico

$$r(t) = s(t) + n(t).$$

La potencia recibida a la entrada del receptor debida al término de señal es por tanto P_S , la potencia de la señal modulada transmitida. En cuanto al término de ruido térmico, se utilizará el modelo estadístico habitual

- Proceso aleatorio $n(t)$ estacionario, ergódico, blanco, gaussiano, con densidad espectral de potencia $S_n(j\omega) = \frac{N_0}{2}$
- El receptor utilizado para modulaciones e amplitud será un receptor coherente
 - Se introducirán filtros para limitar el efecto del ruido antes de proceder a la demodulación. Los filtros se ajustarán al ancho de banda de la señal transmitida, de forma que se minimice el efecto del ruido sin producir ninguna distorsión en la señal de información $s(t)$.
 - Los filtros se considerarán ideales, de forma que se obtengan las prestaciones máximas alcanzables.

El objetivo es calcular la relación señal a ruido (S/N o SNR) de la señal demodulada para los distintos tipos de modulación y se comparará con la relación señal a ruido de la señal en una transmisión en banda base, cuando se transmite la señal moduladora sin modular, y que se denotará como $(\frac{S}{N})_b$.

3.4.1. Relación señal a ruido en una transmisión en banda base

En este apartado se obtiene la relación señal a ruido de referencia, la obtenida cuando se transmite la señal sin modular, de forma que

$$s(t) = m(t) \rightarrow P_S = P_M.$$

La señal a la entrada del receptor será por tanto

$$r(t) = s(t) + n(t)$$

En el receptor el único procesamiento que se realizará será un filtrado para minimizar el efecto del ruido sin distorsionar la señal de información. Para ello se utilizará un filtro paso bajo ideal de ancho de banda B Hz. La Figura 3.38 muestra el esquema del receptor.

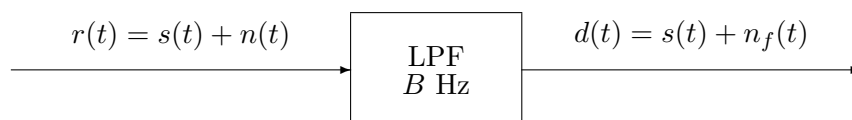


Figura 3.38: Esquema de un receptor en banda base para la recepción de una señal de ancho de banda B Hz.

La potencia del ruido a la salida del filtro se obtiene integrando su densidad espectral de potencia. Teniendo en cuenta que a la salida de un sistema lineal e invariante la densidad espectral de

potencia es la de la entrada multiplicada por el módulo al cuadrado de la respuesta en frecuencia del filtro, la densidad espectral de potencia del ruido filtrado $n_f(t)$ es

$$S_{n_f}(j\omega) = \begin{cases} \frac{N_0}{2} & \text{si } |\omega| \leq W = 2\pi B \text{ rad/s} \\ 0 & \text{si } |\omega| > W = 2\pi B \text{ rad/s} \end{cases}$$

Por tanto, la potencia del ruido filtrado es

$$P_{n_f} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_{n_f}(j\omega) d\omega = \frac{1}{2\pi} \int_{-2\pi B}^{+2\pi B} \frac{N_0}{2} d\omega = N_0 \cdot B \text{ Watt.}$$

Este mismo resultado se podría obtener teniendo en cuenta que la potencia de ruido que atraviesa un filtro ideal de ancho de banda B Hz es $N_0 \times B$ Watt. En cualquier caso, la relación señal a ruido en banda base es

$$\left(\frac{S}{N}\right)_b = \frac{P_S}{N_0 \cdot B}.$$

Este será el valor de referencia con el que se comparará la relación señal a ruido obtenida con las distintas variantes de modulación que se analizarán.

3.4.2. Efecto del ruido en modulaciones de amplitud

En esta sección se determina la relación señal a ruido a la salida del receptor que demodula las señales moduladas en amplitud. Los resultados obtenidos se comparan con el resultado del efecto del ruido en un sistema de comunicaciones banda base.

Receptor coherente y filtrado para minimizar el efecto del ruido

Para todas las modulaciones de amplitud se va a utilizar el receptor que se muestra en la Figura 3.39, donde $\phi = \phi_c$ al tratarse de un receptor coherente. Cuando sea preciso, por cuestión de simplicidad, se considerará $\phi_c = 0$.

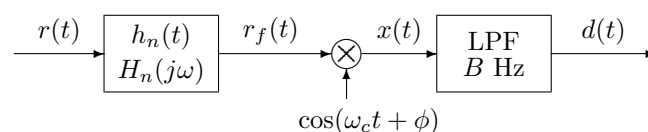


Figura 3.39: Receptor utilizado para modulaciones de amplitud.

El filtro de ruido $h_n(t) / H_n(j\omega)$ se coloca antes del demodulador síncrono para minimizar el efecto del ruido. Para ello, es un filtro paso banda ideal cuya banda de paso, y por tanto el ancho de banda, es igual que la de la señal modulada $s(t)$.

La señal recibida se modela con el modelo de ruido aditivo térmico

$$r(t) = s(n) + n(t).$$

La señal filtrada por el filtro de ruido, teniendo en cuenta que este no modifica la señal de información recibida, es

$$r_f(t) = s(n) + n_f(t), \text{ con } n_f(t) = n(t) * h_n(t).$$

La señal demodulada se obtiene como

$$x(t) = r_f(t) \times \cos(\omega_c t) = s(t) \cdot \cos(\omega_c t) + n_f(t) \cdot \cos(\omega_c t) = x_S(t) + x_n(t)$$

y la señal demodulada filtrada es

$$d(t) = x(t) * h_{LPF-B}(t) = x_S(t) * h_{LPF-B}(t) + x_n(t) * h_{LPF-B}(t) = d_S(t) + d_n(t).$$

Como se puede ver, a la salida del receptor hay dos términos:

- Un término debido a la señal modulada $s(t)$, que es $d_S(t)$.
- Un término debido al ruido térmico $n(t)$, que es $d_n(t)$.

En cuanto al término de señal, no se ve afectado por el filtro de ruido. Para modulaciones de amplitud se calculó en su momento, y el resultado obtenido, junto con la potencia de la señal demodulada, P_{d_S} , se muestra en la Tabla 3.2.

Modulación	P_S	$d_S(t)$	P_{d_S}
AM convencional	$\frac{A_c^2}{2} [1 + P_{M_a}]$	$\frac{A_c}{2} [1 + m_a(t)]$	$\frac{A_c^2}{4} P_{M_a}$
DBL	$\frac{A_c^2}{2} P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$
BLU	$A_c^2 P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$
BLV	$A_c^2 P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$

P_{d_S} : potencia en $d_S(t)$ relativa a $m(t)$

P_{M_a} : potencia de $m_a(t)$, $P_{M_a} = \frac{a^2}{C_M^2} P_M$

Tabla 3.2: Potencia de la señal modulada, señal de salida del demodulador, y potencia de la señal para cada tipo de modulación de amplitud cuando se utiliza un demodulador síncrono.

En cuanto al término de ruido, su potencia depende del filtro de ruido empleado, que a su vez depende del tipo de modulación, y en particular de su ancho de banda. Denotaremos su potencia como P_{d_n} , y se calculará posteriormente su valor para cada tipo de modulación.

Una vez que se haya obtenido la potencia de ruido a la salida del demodulador, la relación señal a ruido tras la demodulación se obtendrá como

$$\left(\frac{S}{N}\right)_d = \frac{P_{d_S}}{P_{d_n}}$$

y se comparará con la relación señal a ruido en banda base

$$\left(\frac{S}{N}\right)_b = \frac{P_S}{N_o \cdot B}$$

Potencia de ruido a la salida del demodulador - Análisis general

La densidad espectral de potencia del ruido filtrado $n_f(t)$ es

$$S_{n_f}(j\omega) = S_n(j\omega) \cdot |H_n(j\omega)|^2 = \frac{N_0}{2} |H_n(j\omega)|^2$$

La densidad espectral de potencia del ruido demodulado $x_n(t)$ es entonces

$$S_{x_n}(j\omega) = \frac{1}{4}S_{n_f}(j\omega - j\omega_c) + \frac{1}{4}S_{n_f}(j\omega + j\omega_c) = \frac{N_0}{8} [|H_n(j\omega - j\omega_c)|^2 + |H_n(j\omega + j\omega_c)|^2]$$

Por tanto, la densidad espectral de potencia tras el filtrado paso bajo $d_n(t)$ es

$$S_{d_n}(j\omega) = S_{x_n}(j\omega) \cdot |H_{LFP-B}(j\omega)|^2 = \begin{cases} S_{x_n}(j\omega), & \text{si } |\omega| \leq W = 2\pi B \\ 0, & \text{si } |\omega| > W = 2\pi B \end{cases}$$

En ese caso, la potencia tras el filtrado paso bajo se calcula como

$$\begin{aligned} P_{d_n} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{d_n}(j\omega) d\omega = \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} S_{x_n}(j\omega) d\omega \\ &= \frac{N_0}{8} \left[\frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega - j\omega_c) d\omega + \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega + j\omega_c) d\omega \right] \end{aligned}$$

Cálculo de la potencia de ruido - AM convencional y DBL

Como el ancho de banda de las modulaciones AM convencional y de doble banda lateral es idéntico, el filtro de ruido y por tanto la potencia de ruido es también la misma. Para ambas modulaciones el filtro de ruido es idéntico

$$H_n(j\omega) = \begin{cases} 1, & \text{si } \omega_c - W \leq |\omega| \leq \omega_c + W \\ 0, & \text{en otro caso} \end{cases}$$

donde W es el ancho de banda en rad/s ($W = 2\pi B$). La Figura 3.40 representa el módulo de la respuesta en frecuencia del filtro de ruido utilizado.

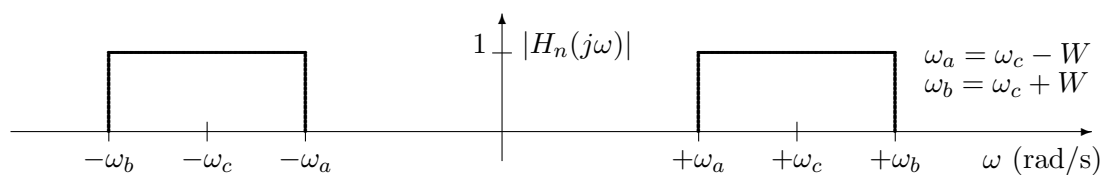


Figura 3.40: Respuesta en frecuencia del filtro de ruido utilizado en la modulación AM convencional y en la modulación de doble banda lateral.

Para este filtro de ruido se tiene que

$$\frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega - j\omega_c) d\omega + \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega + j\omega_c) d\omega = 4B$$

y la potencia de ruido es por tanto

$$P_{d_n} = \frac{1}{2}N_0B.$$

La Figura 3.41 muestra la interpretación frecuencial del proceso que sufre el ruido en el receptor.

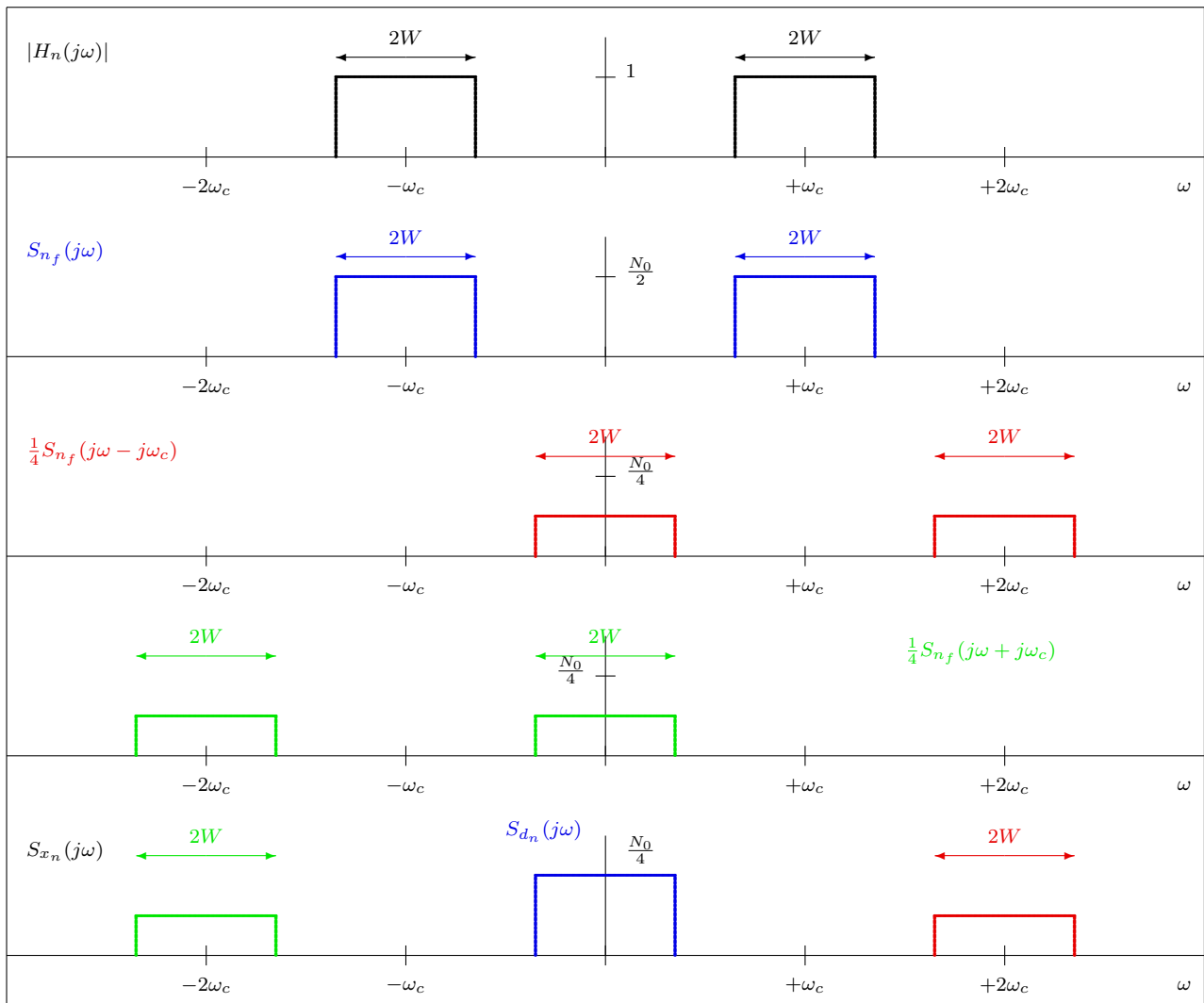


Figura 3.41: Interpretación frecuencial del proceso que sufre el ruido en el receptor para las modulaciones AM convencional y DBL.

Cálculo de la potencia de ruido - Banda lateral única (BLU)

Se va a considerar el caso de una modulación BLU de banda lateral superior. Para banda lateral inferior el resultado es el mismo. Para esta modulación el filtro de ruido tiene la siguiente respuesta en frecuencia

$$H_n(j\omega) = \begin{cases} 1, & \text{si } \omega_c \leq |\omega| \leq \omega_c + W \\ 0, & \text{en otro caso} \end{cases}$$

donde W es el ancho de banda en rad/s ($W = 2\pi B$). La Figura 3.42 representa el módulo de la respuesta en frecuencia del filtro de ruido utilizado.

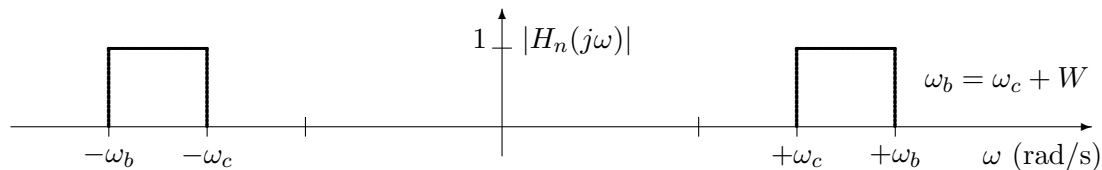


Figura 3.42: Respuesta en frecuencia del filtro de ruido utilizado en una modulación de banda lateral única (banda lateral superior).

Para este filtro de ruido se tiene que

$$\frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega - j\omega_c) d\omega + \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega + j\omega_c) d\omega = 2B$$

y la potencia de ruido es por tanto

$$P_{d_n} = \frac{1}{4} N_0 B.$$

La Figura 3.43 muestra la interpretación frecuencial del proceso que sufre el ruido en el receptor.

Cálculo de la potencia de ruido - Banda lateral vestigial (BLV)

Se va a considerar el caso de una modulación BLV de banda lateral superior. Para banda lateral inferior el resultado es el mismo. Para esta modulación el filtro de ruido tiene la siguiente respuesta en frecuencia

$$H_n(j\omega) = \begin{cases} 1, & \text{si } \omega_c - \Delta_W \leq |\omega| \leq \omega_c + W \\ 0, & \text{en otro caso} \end{cases}$$

donde W es el ancho de banda en rad/s ($W = 2\pi B$), y Δ_W es el exceso de ancho de banda vestigial en rad/s ($\Delta_W = 2\pi \Delta_B$). La Figura 3.42 representa el módulo de la respuesta en frecuencia del filtro de ruido utilizado.

Para este filtro de ruido se tiene que

$$\frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega - j\omega_c) d\omega + \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} H_n(j\omega + j\omega_c) d\omega = 2(B + \Delta_B)$$

y la potencia de ruido es por tanto

$$P_{d_n} = \frac{1}{4} N_0 (B + \Delta_B).$$

La Figura 3.45 muestra la interpretación frecuencial del proceso que sufre el ruido en el receptor.

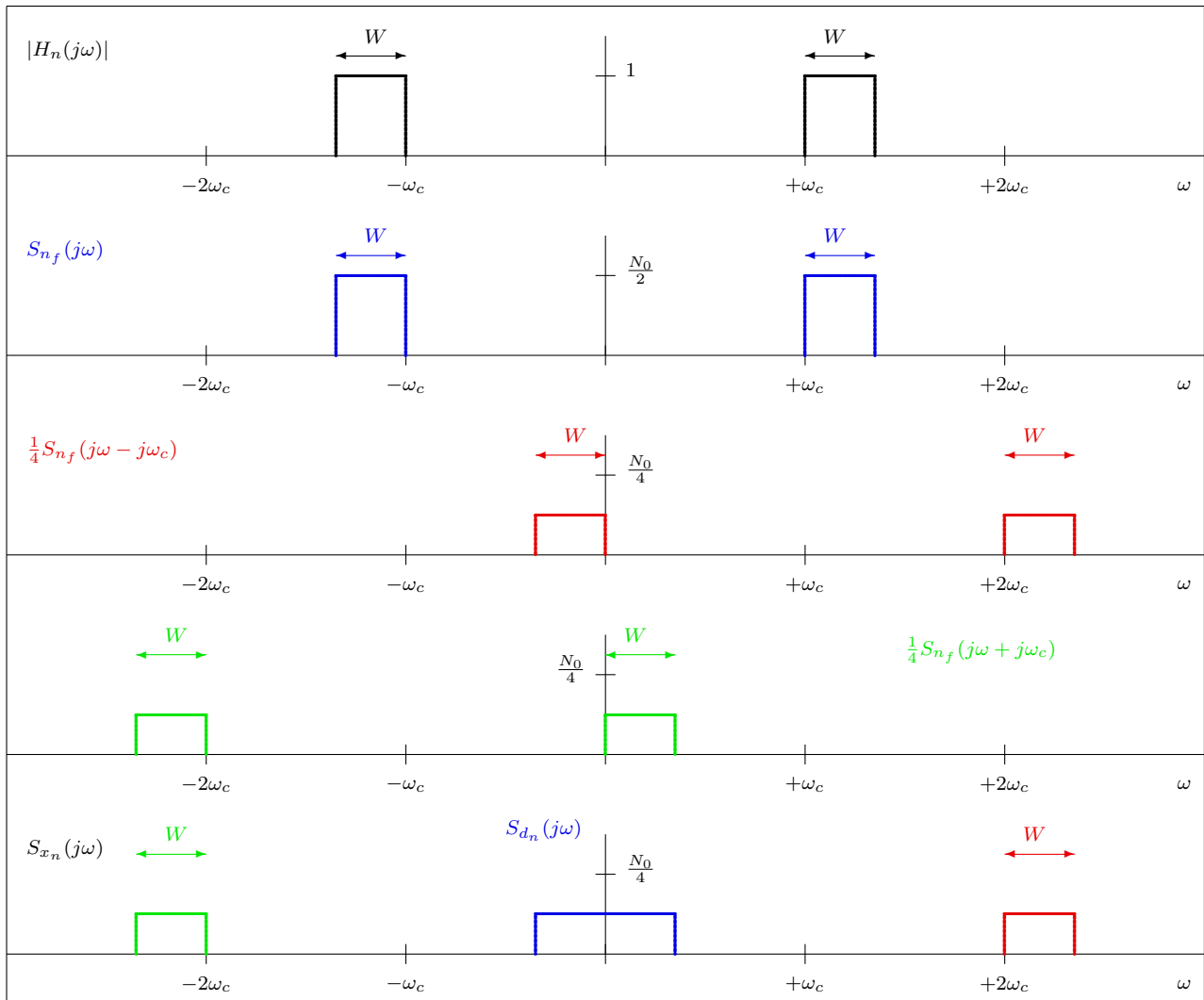


Figura 3.43: Interpretación frecuencial del proceso que sufre el ruido en el receptor para una modulación de banda lateral única (banda lateral superior).

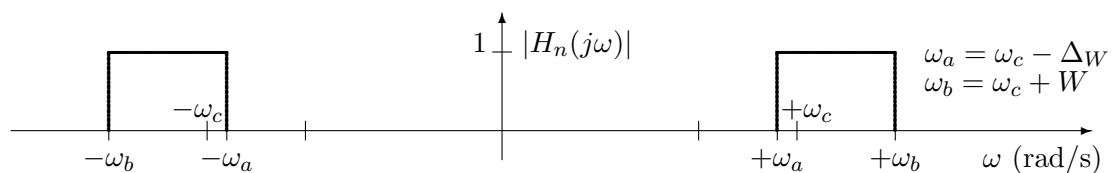


Figura 3.44: Respuesta en frecuencia del filtro de ruido utilizado en una modulación de banda lateral vestigial (banda lateral superior).

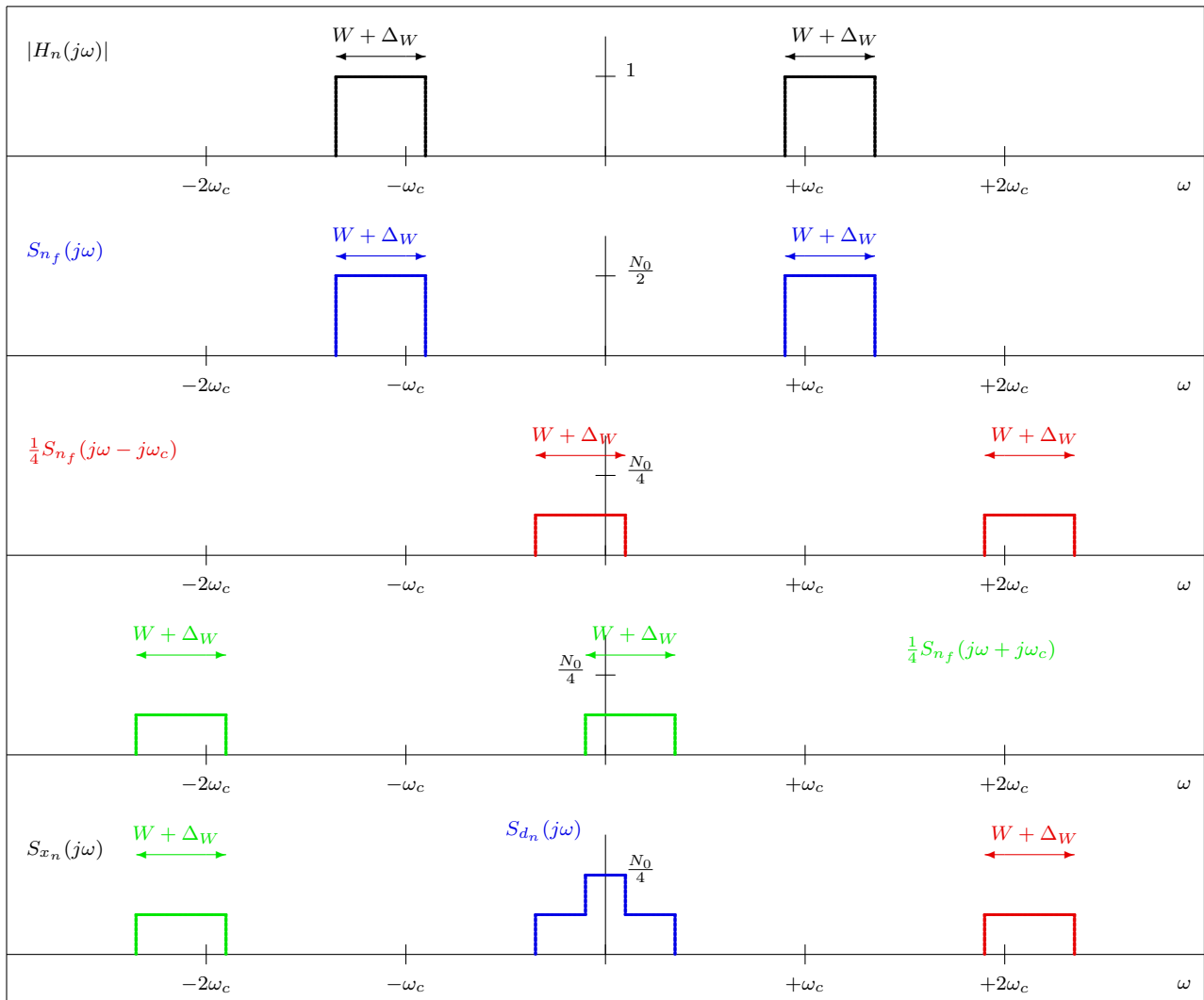


Figura 3.45: Interpretación frecuencial del proceso que sufre el ruido en el receptor para una modulación de banda lateral vestigial (banda lateral superior).

Cálculo de las relaciones señal a ruido

Una vez conocidas la potencia de la señal y del ruido a la salida del demodulador (que se resumen en la Tabla 3.3), sólo falta calcular la relación señal a ruido y compararla con la obtenida en una transmisión en banda base.

Modulación	P_S	$d_S(t)$	P_{d_S}	P_{d_n}
AM convencional	$\frac{A_c^2}{2} [1 + P_{M_a}]$	$\frac{A_c}{2} [1 + m_a(t)]$	$\frac{A_c^2}{4} P_{M_a}$	$\frac{1}{2} N_0 B$
DBL	$\frac{A_c^2}{2} P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$	$\frac{1}{2} N_0 B$
BLU	$A_c^2 P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$	$\frac{1}{4} N_0 B$
BLV	$A_c^2 P_M$	$\frac{A_c}{2} m(t)$	$\frac{A_c^2}{4} P_M$	$\frac{1}{4} N_0 (B + \Delta_B)$

Tabla 3.3: Potencias de señal y de ruido a la salida del demodulador para las modulaciones de amplitud.

Modulación de doble banda lateral (DBL) La relación señal a ruido para este tipo de modulación es

$$\left(\frac{S}{N}\right)_{DBL} = \frac{P_{d_S}}{P_{d_n}} = \frac{\frac{A_c^2}{4} P_M}{\frac{1}{2} N_0 B} = \frac{\frac{A_c^2}{2} P_M}{N_0 B} = \frac{P_S}{N_0 B} = \left(\frac{S}{N}\right)_b$$

Se puede ver que para una modulación DBL se tiene exactamente la misma relación señal a ruido que transmitiendo en banda base.

Modulación de banda lateral única (BLU) La relación señal a ruido para una modulación BLU es

$$\left(\frac{S}{N}\right)_{BLU} = \frac{P_{d_S}}{P_{d_n}} = \frac{\frac{A_c^2}{4} P_M}{\frac{1}{4} N_0 B} = \frac{A_c^2 P_M}{N_0 B} = \frac{P_S}{N_0 B} = \left(\frac{S}{N}\right)_b$$

De nuevo, la misma relación señal a ruido que transmitiendo en banda base.

Modulación AM convencional La relación señal a ruido para una modulación AM convencional es

$$\begin{aligned} \left(\frac{S}{N}\right)_{AM} &= \frac{P_{d_S}}{P_{d_n}} = \frac{\frac{A_c^2}{4} P_{M_a}}{\frac{1}{2} N_0 B} = \frac{\frac{A_c^2}{2} P_{M_a}}{N_0 B} = \frac{P_{M_a}}{1 + P_{M_a}} \frac{\frac{A_c^2}{2} [1 + P_{M_a}]}{N_0 B} \\ &= \underbrace{\frac{P_{M_a}}{1 + P_{M_a}}}_{\eta_{AM}} \frac{P_S}{N_0 B} = \eta_{AM} \left(\frac{S}{N}\right)_b \end{aligned}$$

En este caso se tiene peor relación señal a ruido que transmitiendo en banda base. Esto es debido a la transmisión de la portadora, que no contiene información, y que hace que este tipo de modulación sea poco eficiente en potencia. El factor de eficiencia $\eta_{AM} < 1$ es en este caso

$$\eta_{AM} = \frac{P_{M_a}}{1 + P_{M_a}} = \frac{\frac{a^2}{C_M^2} P_M}{1 + \frac{a^2}{C_M^2} P_M} = \frac{P_M}{\frac{C_M^2}{a^2} + P_M}$$

Como se puede ver, la eficiencia depende del índice de modulación a , habiendo peor eficiencia para valores bajos de a .

Modulación de banda lateral vestigial (BLV) La relación señal a ruido para una modulación BLV es

$$\begin{aligned} \left(\frac{S}{N}\right)_{BLV} &= \frac{P_{dS}}{P_{dn}} = \frac{\frac{A_c^2}{4} P_M}{\frac{1}{4} N_0 (B + \Delta_B)} = \frac{A_c^2 P_M}{N_0 (B + \Delta_B)} = \frac{B}{B + \Delta_B} \frac{A_c^2 P_M}{N_0 B} \\ &= \underbrace{\frac{B}{B + \Delta_B}}_{\eta_{BLV}} \frac{P_S}{N_0 B} = \eta_{BLV} \left(\frac{S}{N}\right)_b \end{aligned}$$

En este caso la relación es también peor relación señal a ruido que transmitiendo en banda base. El factor de eficiencia $\eta_{BLV} < 1$ es ahora

$$\eta_{BLV} = \frac{B}{B + \Delta_B}.$$

La eficiencia depende por tanto del exceso de ancho de banda Δ_B vestigial. Si el vestigio es pequeño en relación al ancho de banda, $\Delta_B \ll B$, en ese caso $\eta_{BLV} \approx 1$, i.e., la relación señal a ruido es prácticamente igual que en banda base.

3.4.3. Efecto del ruido en modulaciones angulares

El análisis es relativamente complicado debido a la no linealidad presente al estar la dependencia de la señal modulada con respecto a la señal moduladora en el argumento de una senoide.

En general, por resumir las principales características sin entrar en un desarrollo analítico riguroso, la señal demodulada se puede escribir como

$$d(t) = \begin{cases} k_p m(t) + Y_n(t), & \text{PM} \\ k_f m(t) + \frac{1}{2\pi} \frac{d}{dt} Y_n(t), & \text{FM} \end{cases}.$$

El término de ruido, $Y_n(t)$, es de tal naturaleza que lleva a obtener las siguientes expresiones para la densidad espectral del ruido a la salida del demodulador

$$S_{n_d}(j\omega) = \begin{cases} \frac{N_0}{A_c^2}, & \text{PM} \\ \frac{N_0}{A_c^2} \omega^2, & \text{FM} \end{cases}.$$

Esto implica que la potencia de ruido a la salida del demodulador es

$$\begin{aligned} P_{n_d} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{n_d}(j\omega) d\omega \\ &= \begin{cases} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{N_0}{A_c^2} d\omega, & \text{PM} \\ \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{N_0}{A_c^2} \omega^2 d\omega, & \text{FM} \end{cases} \\ &= \begin{cases} \frac{2N_0 B}{A_c^2}, & \text{PM} \\ \frac{2N_0 B^3}{3A_c^2} \omega^2, & \text{FM} \end{cases}. \end{aligned}$$

Y la relación señal a ruido,

$$\left(\frac{S}{N}\right)_d = \frac{P_R}{P_{n_d}} = \frac{P_S}{P_{n_d}},$$

es

$$\left(\frac{S}{N}\right)_d = \begin{cases} \frac{k_p^2 A_c^2 P_M}{2 N_0 B}, & \text{PM} \\ \frac{3k_f^2 A_c^2 P_M}{2B^2 N_0 B}, & \text{FM} \end{cases}.$$

Si se denota la potencia recibida como $P_R = \frac{A_c^2}{2}$ y se consideran los índices de modulación

$$\left(\frac{S}{N}\right)_d = \begin{cases} P_R \left(\frac{\beta_p}{\max |m(t)|}\right)^2 \frac{P_M}{N_0 B}, & \text{PM} \\ 3P_R \left(\frac{\beta_f}{\max |m(t)|}\right)^2 \frac{P_M}{N_0 B}, & \text{FM} \end{cases}.$$

Teniendo en cuenta que

$$\left(\frac{S}{N}\right)_b = \frac{P_R}{N_0 B},$$

la relación señal a ruido se puede escribir en función de la relación señal a ruido transmitiendo en banda base.

$$\left(\frac{S}{N}\right)_d = \begin{cases} P_M \left(\frac{\beta_p}{\max |m(t)|}\right)^2 \left(\frac{S}{N}\right)_b, & \text{PM} \\ 3P_M \left(\frac{\beta_f}{\max |m(t)|}\right)^2 \left(\frac{S}{N}\right)_b, & \text{FM} \end{cases}.$$

Se puede observar que en las modulaciones angulares se tiene una ganancia en relación señal a ruido proporcional al índice de modulación al cuadrado.

Efecto umbral en modulaciones angulares

Este efecto de ganancia sólo se produce en la práctica si la relación señal a ruido banda base es mayor que un umbral que viene dado por

$$\left(\frac{S}{N}\right)_{\text{umbral}} = 20 \cdot (\beta + 1).$$

Esto implica en la práctica que hay un nivel umbral de potencia recibida a partir del que se obtiene esa ganancia en relación señal a ruido

$$P_{R_{\text{umbral}}} = (N_0 B) \times \left(\frac{S}{N}\right)_{\text{umbral}} \rightarrow A_{c,\text{umbral}} = \sqrt{2P_{R_{\text{umbral}}}}.$$

3.5. Ejercicios

Ejercicio 3.1 Una misma señal de entrada se aplica a 4 moduladores analógicos diferentes. Se monitoriza la respuesta en frecuencia a la salida de los cuatro moduladores, dando lugar a los cuatro espectros representados en la Figura 3.46. Tenga en cuenta que en el eje de abscisas se representa frecuencia en kHz, y que cada subfigura tiene distintas escalas.

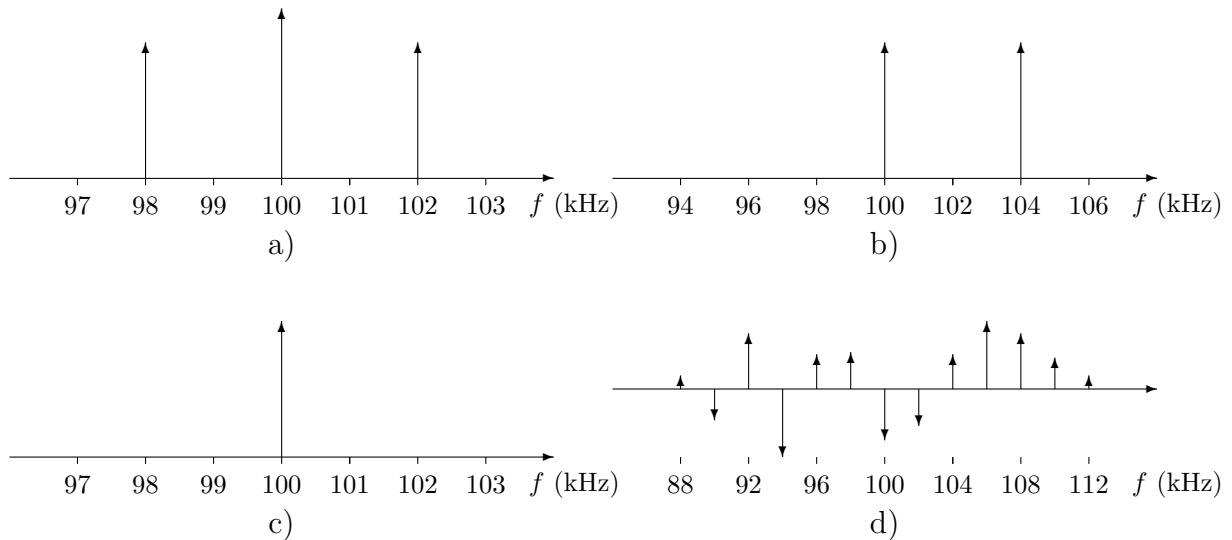


Figura 3.46: Espectros de la señal de salida de los moduladores del Ejercicio 3.1.

- a) Especifique a cuál o cuáles de las posibles variantes de modulaciones analógicas pertenece cada uno de los espectros, y cuál es la frecuencia de la señal portadora de cada modulador.
- b) Proporcione la expresión analítica completa (incluidos los posibles valores numéricos que se puedan extraer a la vista de las figuras) de la señal de entrada a los moduladores (señal moduladora $m(t)$).

Ejercicio 3.2 En un modulador FM con índice de modulación 5 y frecuencia de portadora $f_c = 100$ MHz se introduce una señal moduladora $m(t) = \cos(2\pi f_m t)$, con $f_m = 2$ MHz. Explique cómo es el espectro de la señal y dibújelo indicando claramente las frecuencias y amplitudes de las componentes más relevantes.

Ejercicio 3.3 La señal moduladora de un sistema de comunicaciones analógico, $m(t)$, es una señal de ancho de banda B Hz (ó $W = 2\pi B$ rad/s) con densidad espectral de potencia

$$S_M(j\omega) = \begin{cases} \frac{|\omega|}{W}, & \text{si } |\omega| \leq W = 2\pi B \\ 0, & \text{si } |\omega| > W = 2\pi B \end{cases}$$

La portadora tiene una frecuencia $f_c \gg B$ Hz.

- a) Indique cuál es el ancho de banda de la señal modulada mediante AM convencional con índice de modulación a , y dibuje su densidad espectral de potencia, $S_S(j\omega)$.
- b) Indique cuál es el ancho de banda de la señal modulada mediante modulación de amplitud de doble banda lateral y dibuje su densidad espectral de potencia, $S_S(j\omega)$.

- c) Indique cuál es el ancho de banda de la señal modulada mediante modulación de amplitud de banda lateral única y dibuje su densidad espectral de potencia, $S_S(j\omega)$, para el caso de banda lateral inferior.
- d) Indique cuál es el ancho de banda de la señal modulada mediante una modulación FM con índice de modulación $\beta = 5$.

Ejercicio 3.4 La señal moduladora de un sistema de comunicaciones analógico es $m(t) = \cos(2\pi f_m t)$, con $f_m = 2$ MHz. La portadora $c(t) = A_c \cos(\omega_c t)$, siendo $\omega_c = 2\pi f_c$ con una frecuencia $f_c = 100$ MHz.

- a) Dibuje la densidad espectral de potencia e indique la potencia de la señal modulada $s(t)$ si se emplea una modulación de amplitud con índice de modulación $a = 0,5$.
- b) Dibuje la densidad espectral de potencia e indique la potencia de la señal modulada $s(t)$ si se emplea una modulación AM de doble banda lateral.
- c) Dibuje la densidad espectral de potencia de la señal modulada $s(t)$ si se emplea una modulación AM de banda lateral única (banda lateral inferior).

Ejercicio 3.5 En un sistema de comunicaciones analógicas la señal a transmitir (señal moduladora) es una señal en banda base de ancho de banda B Hz.

- a) Proporcione la eficiencia espectral (dando el ancho de banda) de una modulación AM convencional (con portadora), y diga cómo se comporta frente al ruido comparándola con una transmisión sin modular.
- b) Proporcione la eficiencia espectral (dando el ancho de banda) de una modulación AM de doble banda lateral sin portadora, y diga cómo se comporta frente al ruido comparándola con una transmisión sin modular.
- c) Proporcione la eficiencia espectral (dando el ancho de banda) de una modulación AM de banda lateral única, y diga cómo se comporta frente al ruido comparándola con una transmisión sin modular.
- d) Proporcione la eficiencia espectral (dando el ancho de banda) de una modulación AM de banda lateral vestigial, y diga cómo se comporta frente al ruido comparándola con una transmisión sin modular.
- e) Proporcione la eficiencia espectral (dando el ancho de banda) de una modulación FM con índice de modulación $\beta = 3$, y diga cómo se comporta frente al ruido comparándola con una transmisión sin modular.

Ejercicio 3.6 Las siguientes ecuaciones presentan las expresiones analíticas de la señal modulada $s(t)$ para una señal moduladora $m(t)$ en distintas modulaciones analógicas.

- a) $s_1(t) = A_c \cdot m(t) \cdot \cos(\omega_c t + \phi_c)$
- b) $s_2(t) = A_c \cdot \cos(\omega_c t + a \cdot m(t))$
- c) $s_3(t) = A_c \cdot [1 + m(t)] \cdot \cos(\omega_c t + \phi_c)$
- d) $s_4(t) = A_c \cdot m(t) \cdot \cos(\omega_c t + \phi_c) - A_c \cdot \hat{m}(t) \cdot \sin(\omega_c t + \phi_c)$

e) $s_5(t) = A_c \cdot \cos(\omega_c t + b \cdot \int_{-\infty}^t m(\tau) d\tau)$

f) $s_6(t) = A_c \cdot m(t) \cdot \cos(\omega_c t + \phi_c) + A_c \cdot \hat{m}(t) \cdot \sin(\omega_c t + \phi_c)$

donde $\hat{m}(t)$ denota la transformada de Hilbert de la señal moduladora.

Ejercicio 3.7 Diga a qué variante de modulación analógica corresponde cada una de las siguientes expresiones analíticas de la densidad espectral de potencia, $S_S(j\omega)$, o transformada de Fourier, $S(j\omega)$, de la señal modulada $s(t)$.

a) $S_S(j\omega) = \begin{cases} 0, & |\omega| > \omega_c \\ A_c^2 \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)], & |\omega| \leq \omega_c \end{cases}$

b) $S_S(j\omega) = \frac{A_c^2}{4} \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)]$

c) $S(j\omega) = A_c \cdot \pi \cdot [\delta(\omega - \omega_c) \cdot e^{j\phi_c} + \delta(\omega + \omega_c) \cdot e^{-j\phi_c}] + \frac{A_c}{2} \cdot [M_a(j\omega - j\omega_c) \cdot e^{j\phi_c} + M_a(j\omega + j\omega_c) \cdot e^{-j\phi_c}]$

d) $S_S(j\omega) = \begin{cases} A_c^2 \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)], & |\omega| \geq \omega_c \\ 0, & |\omega| < \omega_c \end{cases}$

e) $S(j\omega) = \sum_{n=-\infty}^{\infty} A_c \cdot \pi \cdot J_n(3) \cdot [\delta(\omega - \omega_c - n \cdot \omega_m) + \delta(\omega + \omega_c + n \cdot \omega_m)]$

f) $S_S(j\omega) = \frac{A_c^2}{2} \cdot \pi \cdot [\delta(\omega - \omega_c) + \delta(\omega + \omega_c)] + \frac{A_c^2}{4} \cdot [S_{M_a}(j\omega - j\omega_c) + S_{M_a}(j\omega + j\omega_c)]$

Ejercicio 3.8 Para modulaciones de amplitud (AM), y en particular para sus receptores

- a) Explique qué es un receptor coherente y dibuje un diagrama de bloques que explique los distintos módulos que lo forman.
- b) Diga cuáles son las variantes de AM que requieren la utilización de un receptor coherente y cuáles pueden utilizar un receptor más simple sin pérdida de prestaciones (en este caso diga qué receptor simple se puede utilizar).
- c) Indique, de forma específica para cada una de las variantes AM que requieren de su utilización, cuál es el efecto de no utilizar un receptor coherente.

Ejercicio 3.9 La señal moduladora de un sistema de comunicaciones se modela mediante un proceso aleatorio estacionario de media nula y cuya densidad espectral de potencia es la que se muestra en la Figura 3.47.

El sistema de comunicaciones utiliza una modulación de banda lateral única (BLU) con banda lateral inferior y frecuencia de portadora $\omega_c \gg W$ rad/s.

- a) Represente la densidad espectral de potencia de la señal modulada.
- b) Compare, en términos de eficiencia espectral, la modulación BLU con las modulaciones AM convencional, AM de doble banda lateral, y modulación FM con índice de modulación $\beta > 1$.

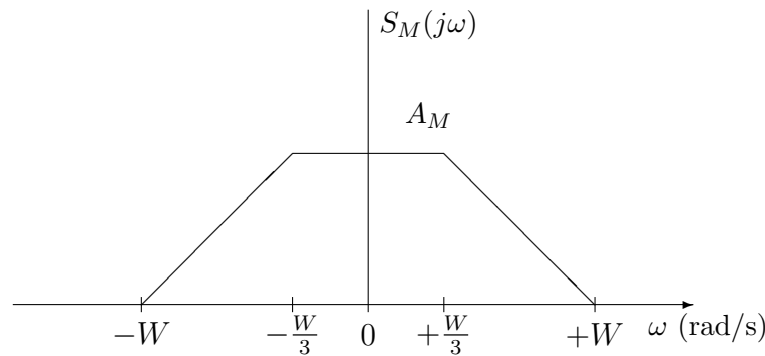


Figura 3.47: Densidad espectral de potencia de la señal moduladora para el Ejercicio 3.9.

Ejercicio 3.10 En un sistema de transmisión analógica, cuya frecuencia de portadora es de 50 kHz y amplitud $A_c = 1$, se va a transmitir como señal moduladora un tono de 5 kHz, es decir

$$m(t) = \cos(2\pi f_m t), \text{ con } f_m = 5 \times 10^3 \text{ Hz.}$$

- Si se utiliza modulación AM convencional, calcule el ancho de banda de la señal modulada, y dibuje, de forma aproximada, la señal modulada en el dominio del tiempo si el índice de modulación es $a = \frac{1}{2}$ (etiquete adecuadamente ambos ejes, especialmente el de amplitud de la señal modulada).
- Si se considera una modulación de amplitud, pero se desea eliminar la ineficiencia que supone la transmisión de la portadora, ¿qué modulaciones serían posibles? Indique el ancho de banda de la señal modulada en todos los casos posibles.
- En el caso de utilizar una modulación FM con índice de modulación $\beta_f = 7$, indique cuál es el ancho de banda teórico (de forma exacta), defina lo que es el ancho de banda efectivo de la señal modulada, y diga cuál sería su valor para este caso.
- Cuando se comparan las modulaciones angulares frente a las modulaciones lineales o de amplitud, explique la ventaja fundamental y la principal desventaja de las primeras respecto a las segundas.

Ejercicio 3.11 Considere las señales de la Figura 3.48.

- Las figuras A, B, C y D muestran la señal modulada cuando se transmite la señal moduladora $m(t)$ utilizando cuatro tipos de modulaciones analógicas: AM convencional (en este caso con índice de modulación $a = \frac{1}{2}$), modulación de doble banda lateral (DBL), modulación de fase (PM) y modulación de frecuencia (FM). Identifique a qué tipo de modulación corresponde cada una de las figuras, indicando los rasgos distintivos de la misma (sin esta indicación, la respuesta no será considerada como válida).

NOTA: para facilitar la identificación, se ha superpuesto en las 4 figuras en trazo fino la portadora $c(t)$.

- Diga qué tipo de modulación analógica, de entre todas las posibles (no se limite a las 4 indicadas en el apartado anterior), utilizaría en los casos que se van a describir a continuación. Para cada caso, debe indicar de forma precisa qué característica de dicha modulación es la que la hace apropiada, y qué ventaja proporciona respecto al resto de modulaciones (sin esta indicación, la respuesta no será considerada como válida).

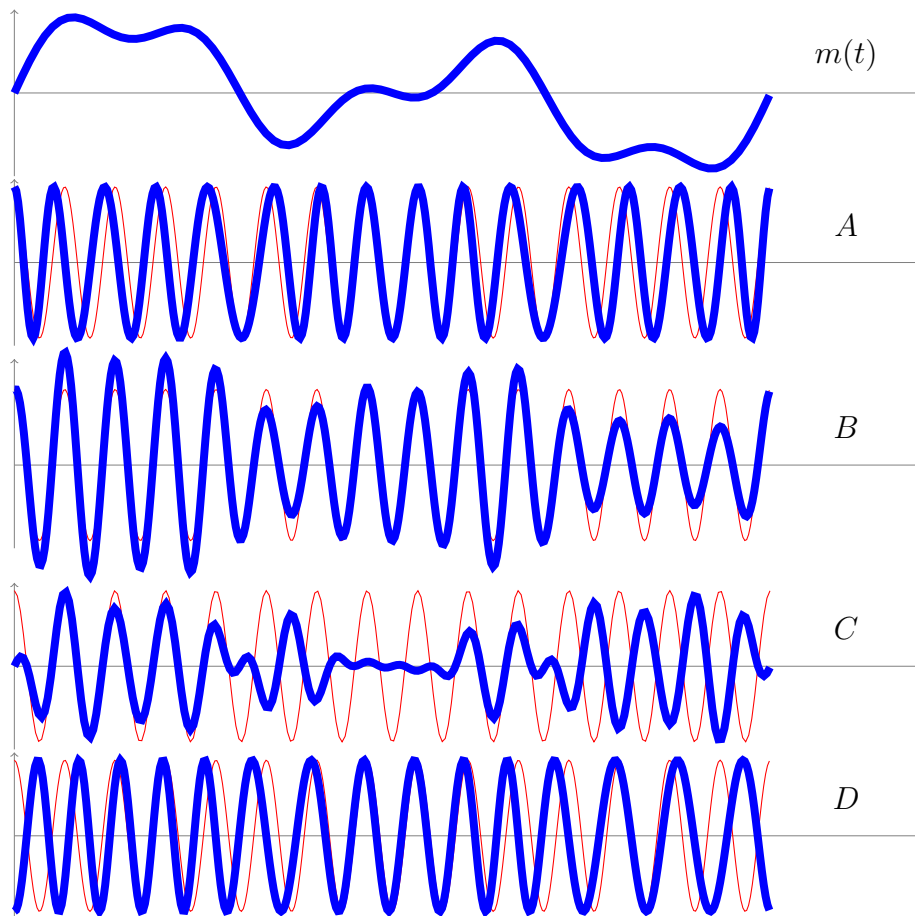


Figura 3.48: Señales para el Ejercicio 3.11.

- I) Sistema que transmite señales de alta fidelidad, donde lo más importante es la calidad de la señal recibida.
- II) Sistema en que la principal prioridad es que los receptores sean lo más económicos que sea posible.
- III) Un sistema de multiplexación de señales de voz por un mismo cable, con un ancho de banda de 400 MHz, en el que se desea poder multiplexar el mayor número posible de señales de voz.

Ejercicio 3.12 Para una modulación de amplitud de banda lateral vestigial, con una señal moduladora de ancho de banda W rad/s

- a) Explique cómo se genera la señal modulada, y dibuje un diagrama de bloques con los distintos elementos que forman el transmisor.
- b) Indique qué condición tiene que cumplir el filtro de banda lateral vestigial para este tipo de modulación, y para cada uno de los filtros de la Figura 3.49, demuestre si se cumple o no dicha condición, y en caso de que así sea, diga si se utilizaría para una modulación de banda lateral superior o inferior.

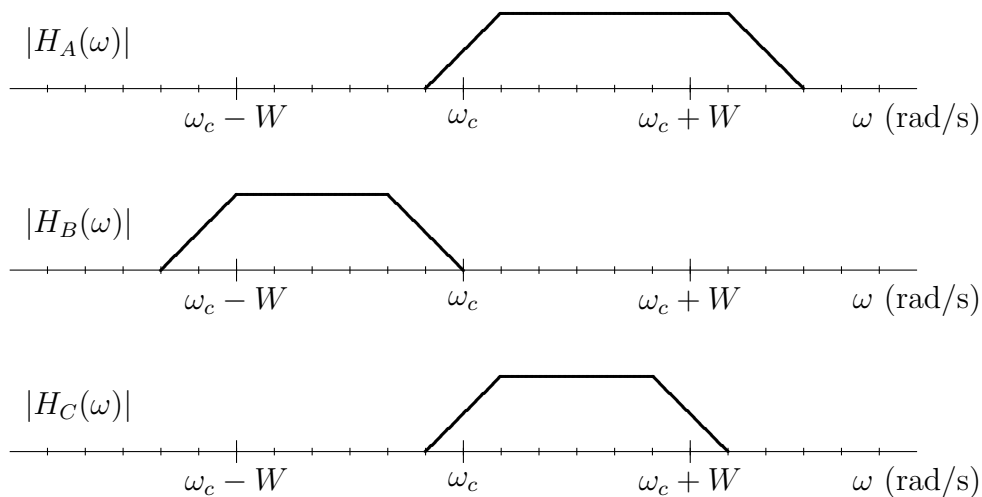


Figura 3.49: Respuestas en frecuencia de los filtros para el Ejercicio 3.12.

3.6. Soluciones de los ejercicios

Ejercicio 3.1 Solución

a) Las modulaciones y frecuencia de portadora son

- Figura (a): modulación AM convencional, con frecuencia de portadora $f_c = 100$ kHz.
- Figura (b): modulación de doble banda lateral, con frecuencia de portadora $f_c = 102$ kHz.
- Figura (c): modulación AM de banda lateral única. Aquí hay dos posibilidades
 - Banda lateral superior con frecuencia de portadora $f_c = 98$ kHz.
 - Banda lateral inferior con frecuencia de portadora $f_c = 102$ kHz.
- Figura (d): modulación angular, con frecuencia de portadora $f_c = 100$ kHz.

b) La señal moduladora es una senoide de frecuencia 2 kHz, es decir

$$m(t) = A_c \cos(\omega_c t + \phi)$$

con $\omega_c = 2\pi f_c$ y $f_c = 2$ kHz.

Ejercicio 3.2 Solución

El espectro de la señal modulada son deltas en las frecuencias

$$\omega_c + m \times \omega_m$$

para todo m entero, donde ω_c es la frecuencia de la portadora y ω_m es la frecuencia de la señal moduladora (en ambos casos son frecuencias angulares en rad/s). En este caso $\omega_c = 2\pi f_c$ y $\omega_m = 2\pi f_m$. La amplitud de cada delta se π veces el coeficiente $J_m(\beta)$, en este caso $J_m(5)$. Teniendo en cuenta que

$$J_0(5) = -0,18, J_1(5) = -0,32, J_2(5) = 0,05$$

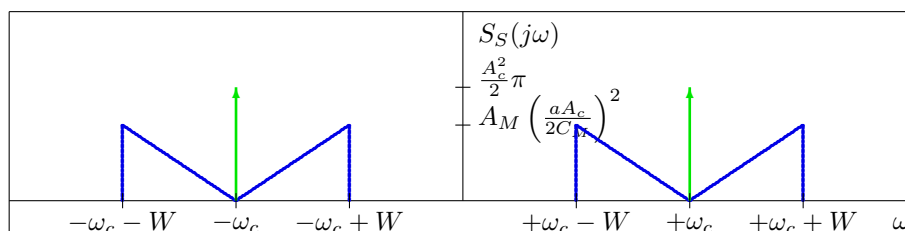
$$J_3(5) = 0,37, J_4(5) = 0,39, J_5(5) = 0,26, \dots$$

y que $J_{-m}(\beta) = (-1)^m J_m(\beta)$, la transformada de Fourier de la señal modulada tendrá la forma de la Figura 3.50 (se representan sólo las 11 frecuencias centrales, para el resto, se seguiría el mismo procedimiento).

Ejercicio 3.3 Solución

En este caso se define para todos los apartados $A_M = 1$.

a) El ancho de banda de $2B$ Hz. La densidad espectral de potencia, para un índice de modulación genérico a , es la de la figura



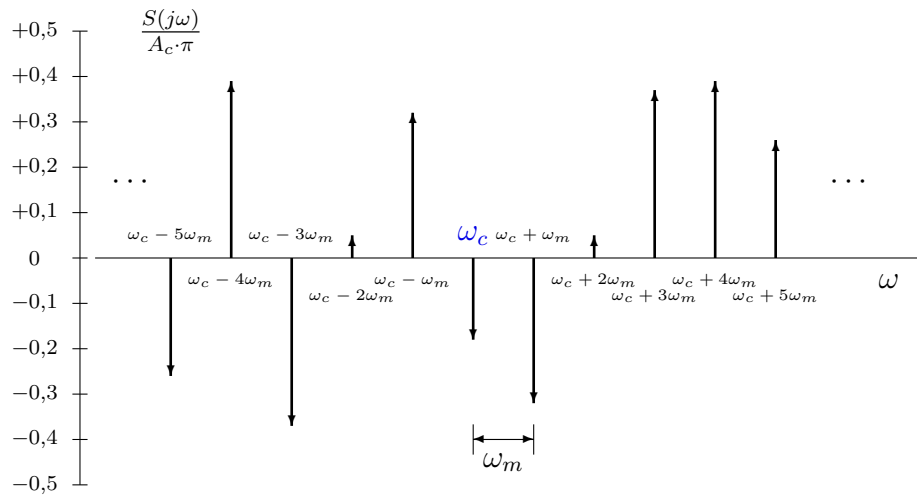
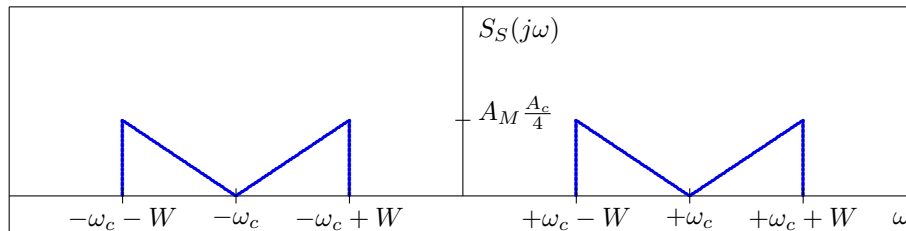
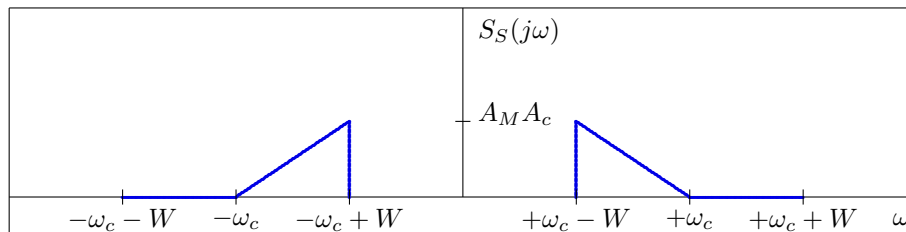


Figura 3.50: Respuesta en frecuencia de la señal modulada para el Ejercicio 3.2.

b) El ancho de banda de $2B$ Hz. La densidad espectral de potencia es la de la figura



c) El ancho de banda de B Hz. La densidad espectral de potencia es la de la figura

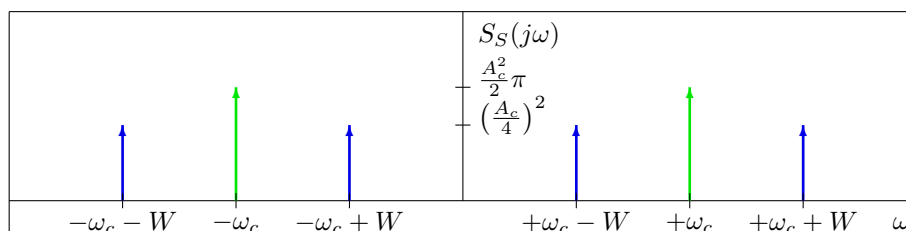


d) El ancho de banda aproximado es $12B$ Hz.

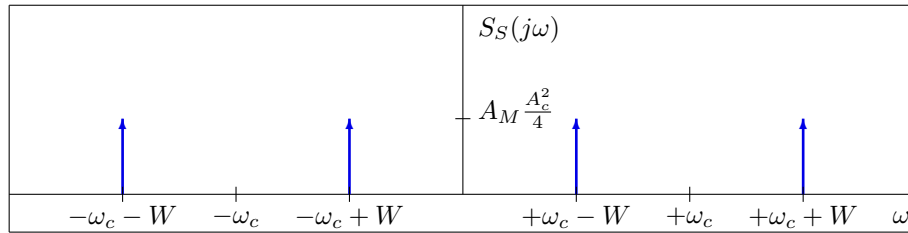
Ejercicio 3.4 Solución

En todos los casos se considera $P_M = \frac{1}{2}$, que es la potencia de la señal moduladora, y $W = 2\pi f_m$.

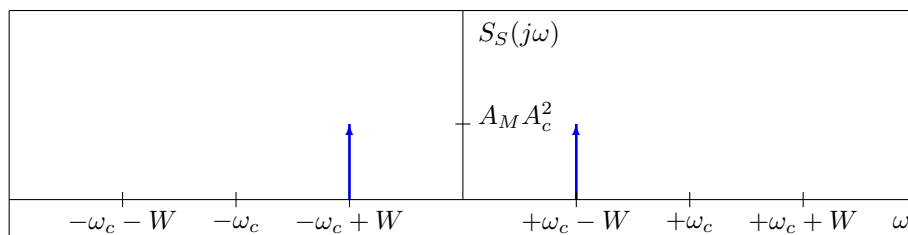
a) La potencia de la señal modulada es $P_S = \frac{A_c^2}{2} [1 + \frac{P_M}{4}]$. La densidad espectral de potencia, para un índice de modulación genérico a , es la de la figura



- b) La potencia de la señal modulada es $P_S = \frac{A_c^2}{2} P_M$. La densidad espectral de potencia es la de la figura



- c) La potencia de la señal modulada es $P_S = A_c^2 P_M$. La densidad espectral de potencia es la de la figura



Ejercicio 3.5 Solución

- a) El ancho de banda es $2B$ Hz, y la relación señal a ruido es peor que la obtenida en una transmisión en banda base

$$\left(\frac{S}{N}\right)_{AM} = \eta_{AM} \left(\frac{S}{N}\right)_b$$

con

$$\eta_{AM} = \frac{P_{M_a}}{1 + P_{M_a}} = \frac{\frac{a^2}{C_M^2} P_M}{1 + \frac{a^2}{C_M^2} P_M} = \frac{P_M}{\frac{C_M^2}{a^2} + P_M}$$

- b) El ancho de banda es $2B$ Hz, y la relación señal a ruido es igual que la obtenida en una transmisión en banda base

$$\left(\frac{S}{N}\right)_{DBL} = \eta \left(\frac{S}{N}\right)_b$$

- c) El ancho de banda es B Hz, y la relación señal a ruido es igual que la obtenida en una transmisión en banda base

$$\left(\frac{S}{N}\right)_{BLU} = \eta \left(\frac{S}{N}\right)_b$$

- d) El ancho de banda es $B + \Delta_B$ Hz, donde Δ_B es el exceso de ancho de banda debido al vestigio, y la relación señal a ruido es ligeramente que la obtenida en una transmisión en banda base

$$\left(\frac{S}{N}\right)_{BLV} = \eta_{BLV} \left(\frac{S}{N}\right)_b$$

con

$$\eta_{BLV} = \frac{B}{B + \Delta_B}$$

- e) El ancho de banda es aproximadamente $8B$ Hz, y la relación señal a ruido es mejor que la obtenida en una transmisión en banda base, siendo la ganancia proporcional al índice de modulación al cuadrado

$$\left(\frac{S}{N}\right)_{FM} \approx 3 \frac{P_M}{C_M^2} \beta_f^2 \cdot \left(\frac{S}{N}\right)_b$$

Ejercicio 3.6 Solución

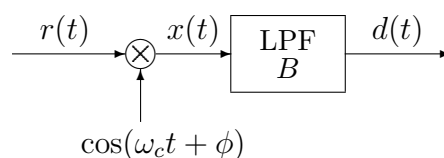
- a) Modulación de amplitud de doble banda lateral (DBL).
- b) Modulación angular de fase (PM).
- c) Modulación de amplitud AM convencional (o doble banda lateral con portadora).
- d) Modulación de banda lateral única (BLU), de banda lateral superior.
- e) Modulación angular de frecuencia (FM).
- f) Modulación de banda lateral única (BLU), de banda lateral inferior.

Ejercicio 3.7 Solución

- a) Modulación de banda lateral única (BLU), de banda lateral inferior.
- b) Modulación de amplitud de doble banda lateral (DBL).
- c) Modulación de amplitud AM convencional (o doble banda lateral con portadora).
- d) Modulación de banda lateral única (BLU), de banda lateral superior.
- e) Modulación angular, de fase (PM) o frecuencia (FM), cuando la señal moduladora es una senoide.
- f) Modulación de amplitud AM convencional (o doble banda lateral con portadora).

Ejercicio 3.8 Solución

- a) Un receptor coherente es un receptor compuesto por un multiplicador por una senoide de la frecuencia de la portadora (a veces denominado demodulador, por su función de devolver el espectro de la señal a banda base) seguido de un filtro paso bajo con el ancho de banda de la señal moduladora, B Hz, tal y como se muestra en la figura

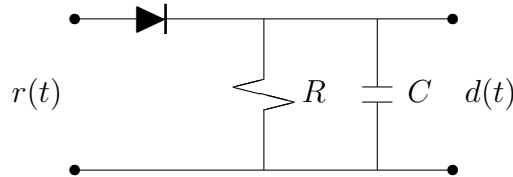


En un receptor coherente, la fase de la portadora en el receptor ha de coincidir con la fase de la portadora que se utilizó en el modulador para generar la señal, $c(t) = A_c \cos(\omega_c t + \phi_c)$, es decir, que debe cumplirse que $\phi = \phi_c$. Normalmente, para lograr esta sincronización de las fases se puede transmitir una señal piloto a la frecuencia y fase de la portadora, o bien se utilizará un lazo enganchado en fase, o PLL, en el receptor.

b) Las variantes de modulaciones AM que requieren un receptor coherente son

- Modulación AM de doble banda lateral (DBL)
- Modulación AM de banda lateral única (BLU)
- Modulación AM de banda lateral vestigial (BLV)

La modulación AM convencional, que también puede hacer uso de un receptor coherente, puede utilizar un receptor más simple, en este caso un detector de envolvente, que se puede implementar con un diodo rectificador y un filtro RC paso bajo, como se muestra en la figura



c) En el caso de la modulación de doble banda lateral, el efecto de tener una fase diferente en las portadoras del transmisor (ϕ_c) y del receptor (ϕ), va a ser que la señal demodulada idealmente (sin ruido ni distorsiones) tendrá la expresión

$$d(t) = \frac{A_c}{2} \cdot m(t) \cdot \cos(\phi_c - \phi),$$

con lo que el desfase entre portadoras generará una atenuación en la señal recibida dependiente de la diferencia de fases. Se puede comprobar, por ejemplo, que un desfase de 90° sería crítico, ya que en este caso se anularía totalmente la señal recibida.

En el caso de la modulación de banda lateral única (y de la de banda lateral vestigial, que es muy similar), el problema es aún mayor, ya que la expresión de la señal a la salida del demodulador es

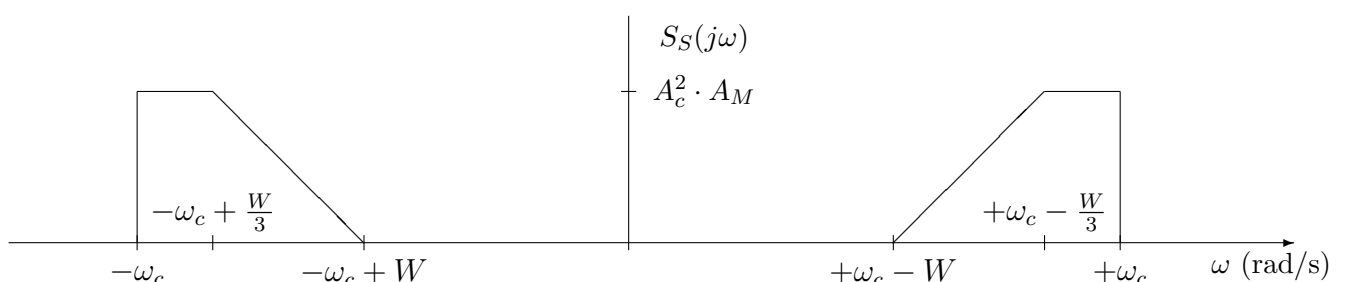
$$d(t) = \frac{A_c}{2} \cdot m(t) \cdot \cos(\phi_c - \phi) \pm \frac{A_c}{2} \cdot \hat{m}(t) \cdot \sin(\phi_c - \phi),$$

con lo que una diferencia de fases no sólo atenuará el término asociado a la señal $m(t)$, sino que hará no nulo el término proporcional a la transformada de Hilbert de la señal moduladora, $\hat{m}(t)$, que es una interferencia sobre la señal deseada $m(t)$.

Ejercicio 3.9 Solución

a)

$$S_S(j\omega) = \begin{cases} 0, & |\omega| > \omega_c \\ A_c^2 \cdot [S_M(j\omega - j\omega_c) + S_M(j\omega + j\omega_c)], & |\omega| < \omega_c \end{cases}$$



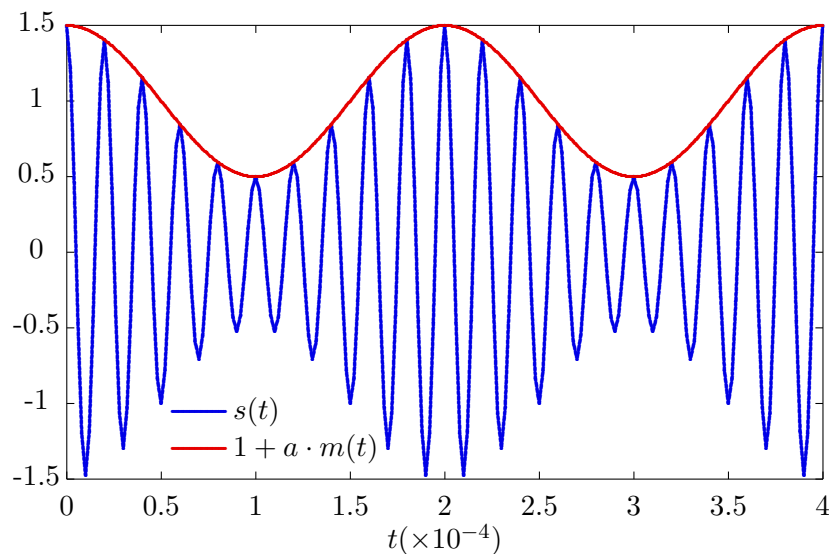
- b) En términos de eficiencia espectral, la modulación BLU es la más eficiente de todas, ya que el ancho de banda requerido para transmitir una señal moduladora de ancho de banda B Hz es B Hz, mientras que tanto la AM convencional como la AM de doble banda lateral requieren el doble de ancho de banda, $2B$ Hz, mientras que la modulación FM requiere un ancho de banda mayor, dependiente del índice de modulación, en concreto, según la regla de Carson, $2(\beta + 1)B$ Hz.

Ejercicio 3.10 Solución

- a) El ancho de banda es

$$BW_{AM} = 2B = 10 \text{ kHz.}$$

La señal modulada en el dominio temporal tendrá la forma



- b) Las modulaciones de amplitud que no incluyen la transmisión de la portadora son
- Modulación de doble banda lateral (DBL).
 - Modulación de banda lateral única (BLU).
 - Modulación de banda lateral vestigial (BLV).

El ancho de banda de la DBL es igual que el de la AM convencional, mientras que el de la BLU es la mitad, y el de la BLV es aproximadamente la mitad (despreciando el vestigio de la banda lateral eliminada)

$$BW_{DBL} = 2B = 10 \text{ kHz, } BW_{BLU} = B = 5 \text{ kHz, } BW_{BLV} \approx B = 5 \text{ kHz.}$$

- c) El ancho de banda teórico es infinito, ya que la transformada de Fourier se puede escribir como una suma infinita de sinusoides en las frecuencias $\omega_c \pm n \cdot \omega_m$ rad/s. El ancho de banda efectivo es el ancho de banda que contiene el 98 % de la potencia de la señal, y en este caso es

$$BW_{FM} = 2 \cdot (\beta_f + 1) \cdot f_m = 80 \text{ kHz.}$$

- d) La principal ventaja de las modulaciones angulares es su mayor relación señal a ruido (inmunidad frente al ruido), que aumenta respecto a la relación en banda base de forma proporcional al cuadrado del índice de modulación. El principal inconveniente es que el ancho de banda se incrementa de forma proporcional a dos veces el índice de modulación más uno, lo que es sensiblemente superior al requerido por las modulaciones lineales.

Ejercicio 3.11 Solución

- a) Las modulaciones de amplitud se distinguen de las angulares en que en las primeras lo que varía con la señal de información o moduladora es la amplitud de la señal, mientras que en las últimas la amplitud permanece constante, pero se modifica la información angular (frecuencia instantánea o término de fase). Por tanto, las modulaciones de amplitud son las de las figuras B y C. Las expresiones analíticas para la modulación AM convencional y la modulación de doble banda lateral son, respectivamente

$$s_{AM}(t) = [1 + a \times m_a(t)] \times c(t), \quad s_{DBL}(t) = m(t) \times c(t).$$

En la primera, la envolvente de la señal es $[1 + a \times m_a(t)]$, que para $a < 1$, como es el caso, es siempre positiva, mientras que en la segunda, simplemente se multiplica la señal moduladora por la portadora, lo que hace que la señal se invierta respecto a la portadora para valores negativos de $m(t)$. Es trivial ver que

- Señal B: modulación AM convencional
- Señal C: modulación de doble banda lateral.

En cuanto a las modulaciones angulares, las modulaciones PM y FM tienen la siguiente expresión analítica

$$s(t) = A_c \cos(\omega_c t + \phi(t)) \text{ con } \phi(t) = \begin{cases} k_p \times m(t), & \text{para PM} \\ 2\pi k_f \times \int_{-\infty}^t m(\tau) d\tau, & \text{para FM} \end{cases}$$

Por tanto, el término de fase $\phi(t)$ es proporcional en un caso a la señal moduladora, y en otro caso a su integral. En cualquier caso, cuando $\phi(t)$ toma valores positivos, se incrementa el argumento de la senoide, lo que significa que la señal modulada se “adelanta” respecto a la portadora, mientras que cuando toma valores negativos, se reduce el argumento de la senoide, lo que significa que la señal modulada se “retarda” respecto de la portadora. Por tanto, en una modulación PM, la señal se adelantará respecto a la portadora para valores positivos de $m(t)$, y se retardará para valores negativos, lo que sucede en la señal A. En una modulación FM, que se adelante o retrase respecto a la portadora está relacionado con el signo de la integral de $m(t)$, lo que sucede en la figura D. Por tanto

- Señal A: modulación PM.
- Señal D: modulación FM.

- b) La modulación más conveniente para cada caso es:

- I) En este caso lo mejor es utilizar una modulación angular con un índice de modulación alto, ya que estas tienen una mejor relación señal a ruido que las modulaciones angulares, relación que es proporcional a β^2 . De entre las dos variantes, para un mismo valor de β (mismo ancho de banda), tiene mejor relación señal a ruido la modulación FM (3 veces la de la PM).
- II) El receptor más simple que se puede utilizar en una modulación analógica es un detector de envolvente, que se puede implementar con un diodo, y un filtro RC. La modulación que puede usar este receptor es la modulación AM convencional. El resto de modulaciones de amplitud debe usar un receptor coherente, y las modulaciones angulares usan receptores más complejos.

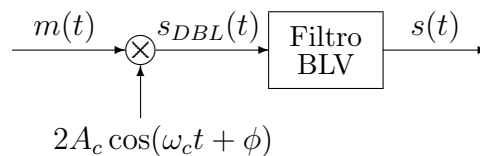
- III) En este caso, se podrán multiplexar más señales cuanto menor sea el ancho de banda de la señal modulada, y la modulación que tiene un menor ancho de banda es la modulación de banda lateral única, que necesita el mismo ancho de banda que la señal moduladora. El resto de modulaciones usa un mayor ancho de banda.

Ejercicio 3.12 Solución

a) La señal modulada de una modulación AM de banda lateral vestigial se genera de la siguiente forma

- Se genera una señal de doble banda lateral (de amplitud doble, en la notación que se siguió en la asignatura, pero este factor de amplitud no es en realidad relevante) multiplicando la moduladora por la señal portadora.
- Se realiza un filtrado de esta señal de doble banda lateral con un filtro de banda lateral vestigial, un filtro que tiene que cumplir unas ciertas condiciones (siguiente apartado).

La figura muestra el diagrama de bloques de un transmisor de AM de banda lateral vestigial



b) La condición que debe cumplir un filtro de banda lateral vestigial es que su respuesta en frecuencia tenga una simetría impar respecto de la frecuencia de portadora, ω_c , de modo que la contribución de la respuesta del filtro desplazada ω_c a la izquierda más la misma respuesta desplazada ω_c a la derecha (efecto de un demodulador) sea constante en el ancho de banda de la señal

$$|H_{BLV}(j\omega - j\omega_c) + H_{BLV}(j\omega + j\omega_c)| = \text{cte, en } |\omega| \leq W = 2\pi B.$$

El único filtro que cumple la condición es el filtro A. En este caso, se trataría de un filtro para una modulación de banda lateral vestigial de banda lateral superior, ya que se deja pasar la banda de frecuencias por encima de la frecuencia de portadora (más el correspondiente vestigio de la banda lateral inferior).

Capítulo 4

Modulación y detección en canales gaussianos

La función básica de un sistema de comunicaciones digitales consiste en el envío y la correspondiente recuperación de manera fiable de una secuencia digital de información enviada a través de un canal de naturaleza analógica. Este es el problema que se aborda en este capítulo.

El capítulo comienza introduciendo la conveniencia de una representación geométrica de las señales para el diseño y análisis de un sistema de comunicaciones digitales, y se presentan las características básicas de dicha representación. A continuación se estudia el problema de la transmisión en canales de ruido aditivo blanco y gaussiano, lo que se denomina comúnmente como *canal gaussiano*; en este modelo se supone que la única distorsión que sufre la señal de comunicaciones durante su transmisión es la suma de ruido térmico, que se modela mediante un proceso aleatorio estacionario, ergódico, blanco, gaussiano, con densidad espectral de potencia $N_0/2$ W/Hz. Se analizan los elementos funcionales del transmisor y del receptor, abordando el diseño óptimo de cada elemento, y se presenta la metodología para obtener las prestaciones que se obtienen con un cierto sistema.

4.1. Introducción

Este capítulo se centra en el análisis de sistemas de comunicaciones digitales. Conviene por tanto recordar cuáles son las características de un sistema de comunicaciones digitales y cuál es el modelo de un sistema de comunicaciones digitales.

4.1.1. Ventajas de los sistemas de comunicaciones digitales

Los sistemas de comunicaciones digitales predominan claramente sobre los sistemas digitales porque tienen múltiples ventajas sobre estos. Las más importantes son las siguientes:

- Capacidad de “*regeneración*”. Esta es sin duda la principal ventaja. Bajo ciertas circunstancias es posible transmitir sin errores, o en general, con una probabilidad de error arbitrariamente baja. Esto es posible debido a que en un cierto instante de tiempo, el símbolo transmitido es uno entre un conjunto finito, lo que supone que la señal puede tener una

forma entre un conjunto finito, y si la distorsión no es muy severa se puede buscar la más parecida, y “regenerar” la señal, como se ilustra en la Figura 4.1. Como se ve en la figura, en cada intervalo se transmite una entre dos posibles formas de onda, nivel alto y nivel, bajo, por lo que si la distorsión es moderada en el receptor se puede recuperar la señal en cada intervalo y recuperar de forma exacta la señal transmitida.

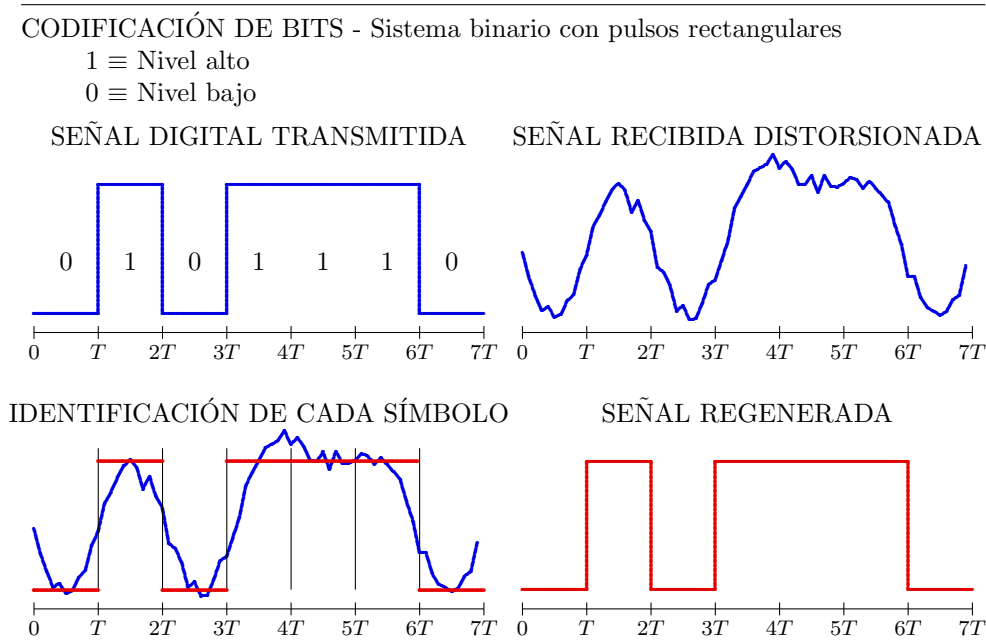


Figura 4.1: Ilustración de la capacidad de regeneración en un sistema de comunicaciones digitales.

- Existen técnicas de detección y corrección de errores. En el ejemplo de la Figura 4.1 la distorsión que sufre la señal transmitida es moderada, por lo que se recupera la señal transmitida (secuencia binaria) sin errores. En casos reales se producen errores, pero existen técnicas que permiten detectar y corregir errores.
- La información se puede encriptar (proteger) de forma relativamente sencilla. Basta con aplicar operaciones aritméticas sobre la secuencia binaria de información.
- Permite corregir la distorsión introducida por el canal (igualación de canal) de una forma mucho más sencilla que en sistemas analógicos. El hecho de saber qué posibles valores puede tomar la señal en un cierto intervalo, hace más fácil estimar el efecto del canal sobre la señal transmitida para invertirlo.
- El formato de la información transmitida es independiente del tipo de información (voz, datos, TV, etc.). En cualquier caso, la información está contenida en una secuencia de símbolos binarios (bits), por lo que al diseñar un sistema, hay que tratar de transmitir bits de forma eficiente, independientemente del tipo de información que contengan.
- Permite utilizar TDM/TDMA y CDM/CDMA (además de FDM/FDMA) como mecanismo de multiplexación/acceso al medio.
- Los circuitos son, en general
 - Más fiables
 - De menor coste

- Más flexibles (programables)

Evidentemente no todo son ventajas en la comparación entre sistemas analógicos y sistemas digitales. Aunque las ventajas claramente predominan sobre las desventajas, es conveniente conocer también estas. Las más importantes son las siguientes:

- La necesidad de sincronismo en el receptor. En el ejemplo de la Figura 4.1, están identificados los instantes donde comienza y finaliza la transmisión de cada bit, lo que permite comparar en cada intervalo la señal recibida con un nivel alto o bajo para realizar la regeneración de la señal (o lo que es lo mismo, la identificación de la secuencia binaria transmitida). En un sistema real, el receptor digital debe generar la señal de sincronismo que permite identificar estos instantes. Esto no es necesario en sistemas analógicos.
- En general, una señal de información para una fuente del mismo tipo (por ejemplo una señal de voz) requiere un mayor ancho de banda para su transmisión en formato digital que en formato analógico. Esta desventaja se puede compensar utilizando técnicas de compresión de las señales en formato digital, lo que permite transmitir las con un menor ancho de banda.
- Muchas fuentes de información son de naturaleza analógica. En realidad, este es un inconveniente menor, ya que como ya se ha visto en el Capítulo 1, la conversión de señales analógicas a formato digital (y la correspondiente reconversión de formato digital a analógico) permite la transmisión de fuentes de naturaleza analógica a través de sistemas de comunicaciones digitales. La tecnología de conversión analógico/digital (A/D) y digital/analógico (D/A) es a día de hoy una tecnología madura y económica.

4.1.2. Descripción general de un sistema de comunicaciones digitales

La principal característica de un sistema de comunicaciones digitales es que la información a transmitir está contenida en una secuencia de símbolos pertenecientes a un alfabeto finito de símbolos, en oposición a los sistemas de comunicaciones analógicos, donde la información está contenida en la forma de una determinada señal moduladora $m(t)$. Esta característica dota a este tipo de sistemas de una serie de ventajas que hacen que hayan desbancado a los sistemas de comunicaciones analógicos, como se ha visto en el capítulo introductorio de la asignatura.

Aunque en general se puede codificar la información en distintos alfabetos, este capítulo se va a centrar en el caso más frecuente, en el que la información está contenida en una secuencia binaria, con lo que el objetivo del sistema será la transmisión eficiente de bits entre la fuente y el destino de la transmisión.

Las Figuras 4.2 y 4.3 muestran los diagramas de bloques de un transmisor y de un receptor digital, respectivamente. Estos diagramas incluyen los elementos funcionales básicos que aparecen en cualquier sistema.

El primer elemento del transmisor, el codificador de fuente, se encarga de reducir la redundancia de la fuente de información (compresión de la información) para reducir la tasa binaria de la información a transmitir. El codificador de canal se encarga de introducir redundancia de forma controlada para permitir la detección y corrección de un cierto número de errores de bit. El ejemplo más sencillo de codificadores de canal son los codificadores de repetición. Estos códigos repiten varias veces cada bit a transmitir. Un código de repetición de orden 1 cuando transmite un cierto

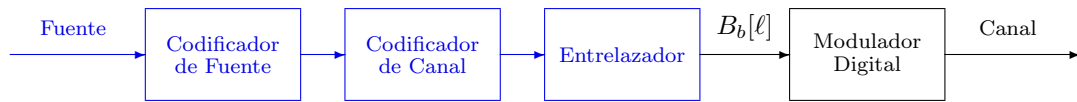


Figura 4.2: Diagrama de bloques funcionales del transmisor de un sistema de comunicaciones digitales.

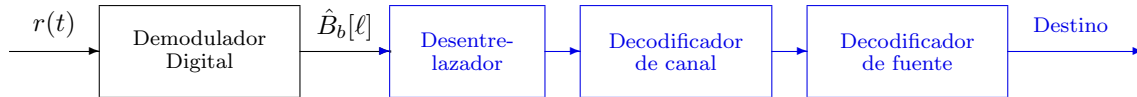


Figura 4.3: Diagrama de bloques funcionales del receptor de un sistema de comunicaciones digitales.

bit lo repite otra vez, de forma que se envía cada bit dos veces. Un código de repetición de orden 2 repite cada bit dos veces, con lo que al final se transmite cada bit tres veces, realizando un proceso de codificación

$$0 \rightarrow 000$$

$$1 \rightarrow 111$$

En cada bloque de tres bits codificados, en el receptor es posible la detección de hasta dos errores; adoptando la estrategia de decisión por mayoría, en el receptor se podrá corregir un error sobre cada bloque de tres bits transmitidos. Si hay dos o tres errores, con este esquema de decisión por mayoría se tomaría una decisión errónea sobre el bloque. Así, cada código de canal tiene una cierta capacidad de detección y corrección.

En muchas ocasiones los errores de bit en la secuencia transmitida en un sistema no se producen de forma aislada, sino que se producen en ráfagas de errores. La aplicación directa de las técnicas de codificación de canal sobre ráfagas de errores no es útil, ya que funcionan bien cuando sobre cada bloque de bits codificados (3 en el caso del ejemplo anterior) hay un número limitado de errores. El papel del entrelazador es convertir ráfagas de errores de bit en errores aislados. Para ello se reordenan los bits antes de la transmisión (entrelazado) y se devuelven al orden original en el receptor (desentrelazado), de forma que errores de ráfaga antes del desentrelazado se convierten en errores aislados a la salida del desentrelazador, tal y como se muestra en el ejemplo de la Figura 4.4.

En cuanto al transmisor, a partir de la fuente del sistema de comunicaciones, y después de realizar los procesos de codificación de fuente, codificación de canal, y del entrelazado, se genera una secuencia binaria $B_b[\ell]$ que contiene la información que se quiere transmitir. En este capítulo se va a estudiar el último elemento del transmisor, el modulador digital. La principal función de este modulador digital es la transmisión de la secuencia de bits, $B_b[\ell]$, a través de un canal de comunicaciones analógico, para lo que debe convertir dicha secuencia en una señal electromagnética $s(t)$ adecuada para su transmisión por el medio analógico utilizado para la transmisión (un cable, fibra óptica, el espectro radioeléctrico, etc.).

En cuanto al receptor, el demodulador digital se encargará de la recuperación de la secuencia de bits que se ha transmitido a partir de la señal recibida través del canal de comunicaciones, $r(t)$. Idealmente se debería recuperar de forma exacta la secuencia transmitida, pero la distorsión sufrida por la señal durante su transmisión hará que esto no sea en general posible, por lo que el demodulador digital deberá proporcionar una estima de la secuencia transmitida, $\hat{B}_b[\ell]$, intentando minimizar el número de errores de la estima.

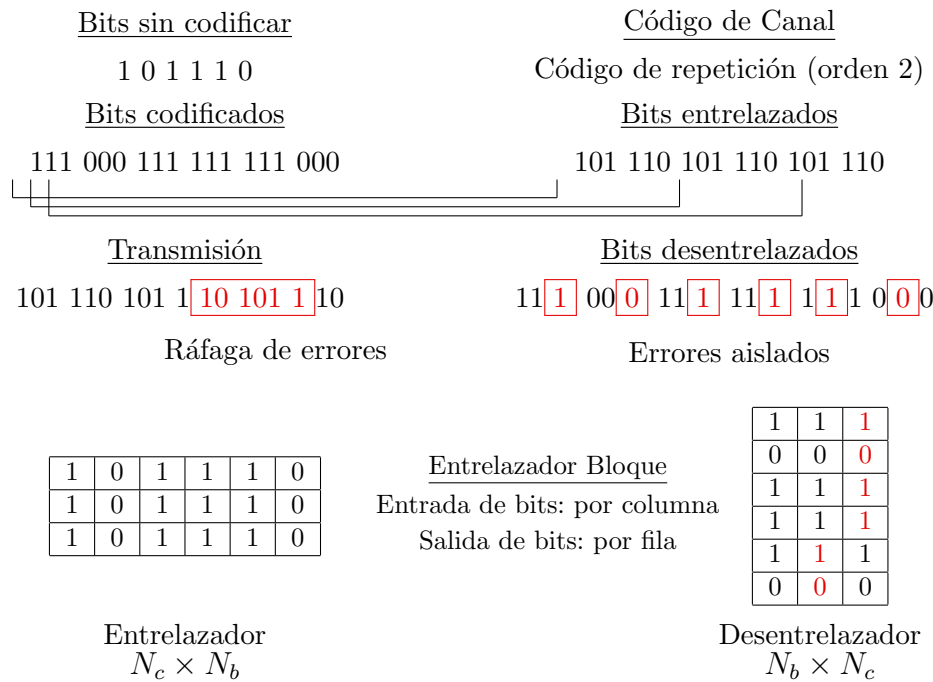


Figura 4.4: Ejemplo de entrelazador matricial, en el que se reordenan los bits introduciéndolos en una estructura matricial por columnas y sacándolos leyendo por filas.

En este capítulo se estudiarán el modulador digital del transmisor y el demodulador digital del receptor. Es decir, que los problemas básicos que se abordan son la transmisión de una secuencia de bits a través de un canal analógico, y la recuperación de dicha secuencia a partir de la señal recibida a la salida del canal, que habrá sufrido cierto tipo de distorsiones durante su transmisión a través del canal analógico.

4.1.3. Diseño genérico de un modulador digital y notación básica

Como se acaba de ver, la función del modulador digital es la transmisión de bits (secuencia $B_b[\ell]$) a una tasa binaria $R_b = \frac{1}{T_b}$ bits/s a través de un medio de transmisión (canal) de naturaleza analógica, para lo que deberá convertir la secuencia de bits en una señal electromagnética $s(t)$.

En general, la transmisión de la secuencia de bits no se realizará bit a bit, sino que se realizará por bloques de m bits. Se segmentará la secuencia de bits $B_b[\ell]$ en bloques de m bits, y se enviará cada uno de estos bloques. A cada bloque de m bits se le denominará *símbolo*, de modo que el alfabeto de los símbolos tendrá $M = 2^m$ posibles valores: las M posibles combinaciones de m bits. Los elementos de este alfabeto se denotarán como b_i , con $i \in \{0, 1, \dots, M - 1\}$, es decir

$$B[n] \in \{b_i\}_{i=0}^{M-1}.$$

Así pues, lo primero que se realizará en el modulador digital será la conversión de la secuencia de símbolos $B_b[\ell]$ a una secuencia de símbolos $B[n]$. Un aspecto importante en esta conversión es el cambio de velocidades, ya que se pasará de la velocidad binaria o de bit R_b bits/s a la velocidad de símbolo R_s símbolos/s o baudios. La relación entre estas dos velocidades o tasas de transmisión es evidente. Dado que cada símbolo contiene m bits, la relación es

$$R_b = m \times R_s.$$

Asociada a cada velocidad hay una duración temporal, el tiempo de bit, T_b y el tiempo de símbolo, T , respectivamente, que indicará cada cuantos segundos se transmite un bit o un símbolo, respectivamente. La relación de estos tiempos con las velocidades de transmisión y entre ellos es también evidente

$$T_b = \frac{1}{R_b}, T = \frac{1}{R_s}, T = m \times T_b.$$

En este capítulo se presentará la forma más sencilla de transmisión de la información, con el objetivo de introducir las técnicas básicas de modulación y de demodulación. La secuencia de información digital ha de convertirse en una señal eléctrica. Esta conversión se realizará por tramos. Se va a realizar la transmisión de un símbolo cada T segundos, de forma que se asociará un intervalo temporal a la transmisión de cada símbolo. El primer símbolo de la secuencia, $B[0]$, se comenzará a transmitir en el instante $t = 0$, y durará hasta $t = T$. A continuación se transmitirá el segundo símbolo de la secuencia, $B[1]$, entre los instantes $t = T$ y $t = 2T$, y así sucesivamente. En general, el símbolo $B[n]$ tendrá asociado el intervalo $nT \leq t < (n + 1)T$. De este modo, la forma de la señal en cada intervalo de símbolo estará asociada al símbolo que se transmite en dicho intervalo.

Como se ha explicado anteriormente, la secuencia de símbolos puede tomar uno entre M posible valores para cada instante discreto n

$$B[n] \in \{b_0, b_1, \dots, b_{M-1}\}.$$

La forma más simple de realizar la conversión de símbolos a señal $s(t)$ es definiendo un conjunto de M señales de duración T segundos

$$\{s_0(t), s_1(t), \dots, s_{M-1}(t)\}, \text{ definidas en } 0 \leq t < T$$

y realizando una asociación símbolo / forma de onda:

$$b_i \leftrightarrow s_i(t).$$

Cuando en el instante n la secuencia de símbolos toma un cierto valor, por ejemplo $B[n] = b_j$, la forma de la señal $s(t)$ en el intervalo asociado a este símbolo, $nT \leq t < (n + 1)T$, será la forma de onda asociada a b_j , que es $s_j(t)$. Obviamente, hay que trasladar la forma de onda $s_j(t)$ al intervalo correspondiente, por lo que se traslada la forma de onda asociada a b_j al intervalo, quedando

$$s(t) = s_j(t - nT), \text{ en } nT \leq t < (n + 1)T.$$

A continuación se ilustra este procedimiento con un ejemplo sencillo. En este caso se van a transmitir los bits en bloques de tamaño $m = 2$ bits, por lo que el sistema tiene un alfabeto de $M = 2^2 = 4$ símbolos. Por ejemplo se puede hacer la siguiente asociación

$$b_0 \equiv 00, b_1 \equiv 01, b_2 \equiv 10, b_3 \equiv 11,$$

aunque se podría haber hecho otra diferente. Hay que elegir un conjunto de $M = 4$ señales, y asociar una señal a cada símbolo. En este ejemplo se han elegido las cuatro señales de la Figura 4.5. La asociación realizada está implícita en los subíndices ($s_i(t)$ está asociada a b_i).

Se quiere realizar la transmisión de la siguiente secuencia binaria

$$B_b[\ell] = 011110001101 \dots$$

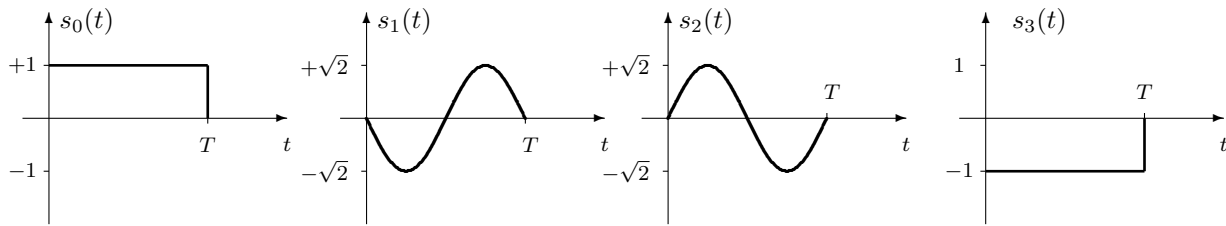


Figura 4.5: Conjunto de 4 señales elegidas para el ejemplo.

Se segmenta esta secuencia en bloques de $m = 2$ bits

$$B_b[\ell] : 01 \mid 11 \mid 10 \mid 00 \mid 11 \mid 01 \mid \dots$$

y se realiza la conversión de secuencia binaria a secuencia de símbolos identificando cada bloque de 2 bits de acuerdo a la asignación especificada anteriormente

$$B[n] = b_1 \mid b_3 \mid b_2 \mid b_0 \mid b_3 \mid b_1 \mid \dots$$

Ahora la señal transmitida se generará por intervalos de símbolo. Como el primer símbolo es b_1 , en el primer intervalo de símbolo ($0 \leq t < T$) se colocará la forma de onda asociada, $s_1(t)$. Como el segundo símbolo es b_3 , en el segundo intervalo de símbolo ($T \leq t < 2T$) se colocará la forma de onda asociada, $s_3(t)$, que se retardará T segundos para colocarla en ese intervalo. Como el tercer símbolo es b_2 , en el tercer intervalo de símbolo ($2T \leq t < 3T$) se colocará la forma de onda asociada, $s_2(t)$, que se retardará $2T$ segundos para colocarla en ese intervalo. De este modo, la señal modulada $s(t)$ se generará por intervalos de la forma siguiente

$$s(t) = \{s_1(t) \mid s_3(t - T) \mid s_2(t - 2T) \mid s_0(t - 3T) \mid s_3(t - 4T) \mid s_1(t - 5T) \mid \dots\}$$

La Figura 4.6 muestra la señal resultante al aplicar este proceso de generación por intervalos de símbolo.

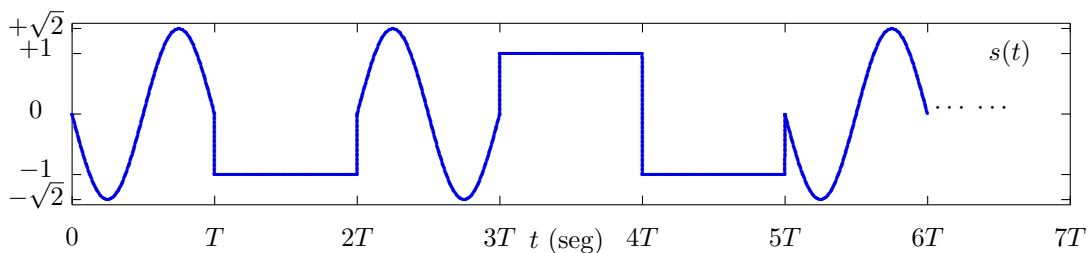


Figura 4.6: Señal generada por la secuencia binaria del ejemplo.

A la luz de este ejemplo, se puede ver claramente que el diseño del modulador digital consiste en la elección de M señales para asociarlas a cada uno de los M posibles valores de la secuencia de símbolo. En este ejemplo la elección ha sido arbitraria. ¿Es la elección del ejemplo una buena elección? ¿Qué criterios hay que tener en cuenta a la hora de elegir el conjunto de M señales? Esta es la pregunta a la que se tratará de dar respuesta a lo largo de este capítulo.

4.1.4. Transmisión a través de un canal de comunicaciones

Una vez generada la señal modulada $s(t)$ que contiene la información digital a enviar, esta se transmite a través de un medio físico o canal de comunicaciones, como podría ser un cable, fibra

óptica, el espectro radioeléctrico o cualquier otro medio que permita la transmisión de una señal. La señal sufre distorsiones durante la transmisión de forma que la señal recibida a la entrada del receptor, $r(t)$, no coincide con la señal transmitida: $r(t) \neq s(t)$.

Existen varios tipos de distorsión que afectan a una señal durante su transmisión por un medio físico. En esta asignatura se considerarán únicamente los dos más importantes: distorsión lineal y ruido. La forma habitual de modelar la distorsión lineal es mediante un modelo de sistema lineal e invariante, caracterizado por una cierta respuesta impulsiva $h(t)$ y su correspondiente respuesta en frecuencia $H(j\omega)$, respuestas relacionadas a través de la transformada de Fourier. El ruido se modela como un proceso aditivo cuyas características estadísticas se asocian en la mayoría de los casos al modelo habitual de ruido térmico: un proceso aleatorio estacionario, ergódico, blanco, gaussiano, de media nula y densidad espectral de potencia $N_0/2$ W/Hz, donde la constante N_0 se obtiene mediante el producto de la constante de Boltzmann y la temperatura expresada en grados Kelvin. Por tanto, el modelo de canal de comunicaciones que se utilizará en este capítulo será el definido por la relación

$$r(t) = s(t) * h(t) + n(t)$$

que se ilustra en la Figura 4.7.

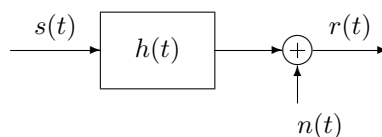


Figura 4.7: Modelo del canal de comunicaciones.

4.1.5. Diseño genérico de un demodulador digital y notación básica

La función del demodulador digital es la recuperación de la secuencia de bits $B_b[\ell]$ a partir de la señal recibida a través del canal, $r(t)$. La señal sufre distorsiones durante su transmisión a través del canal de comunicaciones, de modo que la señal recibida será diferente de la señal transmitida

$$r(t) \neq s(t).$$

El procedimiento básico para recuperar la señal sería el siguiente. Se procesará la señal recibida por tramos, en concreto por intervalos de símbolo. En el intervalo de símbolo asociado al índice discreto n , intervalo $nT \leq t < (n + 1)T$, se observará la señal recibida y se comparará con las M formas de onda del sistema. Se elegirá la señal con la que tenga un mayor “parecido”, y si esta es $s_j(t)$, la estima para el símbolo asociado a ese intervalo será el símbolo asociado a esa forma de onda $\hat{B}[n] = b_j$.

Volviendo al ejemplo anterior, tras la transmisión de la señal, esta sufre una cierta distorsión, dando lugar a la señal de la Figura 4.8.

Tras segmentar la señal en intervalos de símbolo, se buscará para cada intervalo de símbolo cuál de las M posibles señales del sistema (señales de la Figura 4.5 en este ejemplo) es más parecida a la señal recibida en el intervalo. En este caso, una sencilla inspección visual permite identificar la señal correspondiente a cada intervalo

- $n = 0$, intervalo $0 \leq t < T$ - Señal “más parecida”: $s_1(t) \rightarrow \hat{B}[0] = b_1$

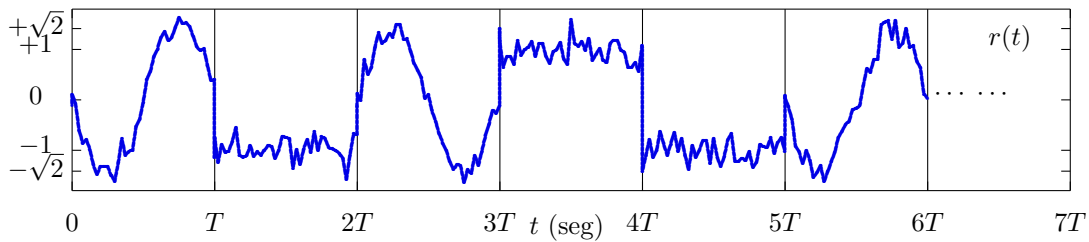


Figura 4.8: Señal recibida a la entrada del receptor en el ejemplo.

- $n = 1$, intervalo $T \leq t < 2T$ - Señal “*más parecida*”: $s_3(t) \rightarrow \hat{B}[1] = b_3$
- Siguiendo el mismo proceso: $\hat{B}[2] = b_2$, $\hat{B}[3] = b_0$, $\hat{B}[4] = b_3$, $\hat{B}[5] = b_1$

De este modo, la secuencia estimada sería

$$\hat{B}[n] = b_1 | b_3 | b_2 | b_0 | b_3 | b_1 | \dots$$

que tras la conversión de símbolos a bits proporcionaría

$$\hat{B}_b[\ell] : 01 | 11 | 10 | 00 | 11 | 01 | \dots$$

En el ejemplo la distorsión de la señal es moderada, de forma que mediante una inspección visual se recupera la secuencia original sin problemas. Pero el sistema real deberá medir de algún modo el “*parecido*” o “*diferencia*” entre señales. Es necesario definir una medida que cuantifique la diferencia de la señal recibida en el intervalo asociado al instante discreto n con la forma de onda $s_i(t)$, que podríamos denotar como

$$\text{diferencia}(n, i).$$

De este modo, el criterio para realizar una decisión sería decidir el símbolo asociado a la señal con una menor diferencia (mayor parecido) respecto a la señal recibida en el intervalo

$$\hat{B}[n] = b_i \text{ si } \text{diferencia}(n, i) < \text{diferencia}(n, j) \text{ para todo } j \neq i.$$

Por ejemplo, se nos podría ocurrir una medida que sume la diferencia en módulo entre la señal observada en el intervalo de símbolo y cada una de las posibles señales de referencia a lo largo del intervalo de símbolo (la suma sobre cada instante de tiempo se convierte en una integral)

$$\text{diferencia}_A(n, i) = \int_{nT}^{(n+1)T} |s(t) - s_i(t - nT)| dt.$$

Pero también se nos podría ocurrir medir directamente la energía de la señal diferencia, que vendría dada por

$$\text{diferencia}_B(n, i) = \int_{nT}^{(n+1)T} |s(t) - s_i(t - nT)|^2 dt.$$

Intuitivamente ambas medidas cuantifican de algún modo la diferencia entre las señales, dando valores más bajos cuanto más se parezcan (darían 0 sólo si las señales son idénticas). Y se nos podrían ocurrir muchas más medidas que podrían servir para cuantificar el parecido o diferencia entre señales. ¿Cuál sería la mejor medida para minimizar la probabilidad de equivocarse? El diseño del demodulador digital consistirá precisamente en eso, en buscar la mejor forma de cuantificar ese parecido para que la secuencia estimada sea la que minimice el número de estimas erróneas.

4.1.6. Factores a considerar a la hora de elegir las M formas de onda

Acabamos de ver qué función básica han de realizar el modulador y el demodulador digital en un sistema de comunicaciones digitales, qué procedimiento se utilizará para realizar dicha función, y en qué consiste su diseño. Resumiendo:

- Modulador digital
 - Función: Conversión de una secuencia de bits $B_b[\ell]$ en una señal electromagnética $s(t)$.
 - Procedimiento: agrupar bits en bloques de m bits (símbolos), asociar una forma de onda a cada uno de los M posibles valores de un bloque (símbolo), y transmitir en el intervalo asociado a cada símbolo la forma de onda que lo representa.
 - Diseño: Elección de un conjunto apropiado de M señales para transportar cada uno de los M símbolos posibles (bloques de m bits).

- Demodulador digital
 - Función: Estima de la secuencia de bits enviada, $\hat{B}_b[\ell]$ a partir de la señal recibida $r(t)$.
 - Procedimiento: Decisión para cada intervalo de símbolo, de cuál de las M señales se transmitió en cada intervalo de símbolo a la vista de la señal recibida en dicho intervalo, comparando la forma de la señal en el intervalo con las M posibles señales y eligiendo una de ellas.
 - Diseño: Determinar la medida óptima para realizar la comparación y la elección, de forma que se minimicen los errores cometidos teniendo en cuenta la distorsión que sufre la señal en su transmisión a través del canal.

A la hora de realizar el diseño del sistema, lo primero es tener claros los factores que habrá que considerar. En particular, a la hora de diseñar el modulador y el demodulador digital habrá que tener en cuenta principalmente tres factores:

1. Prestaciones: probabilidad de equivocarse en el receptor (P_e)
 - Si en el receptor para cada intervalo se decide la señal más “parecida”, obviamente la probabilidad de error dependerá del “parecido” entre señales, por lo que habrá que buscar señales que sean lo más diferentes que sea posible.
 - Otro problema es determinar la medida de “parecido” o “diferencia” que lleva a minimizar la probabilidad de error teniendo en cuenta el tipo de distorsión que sufre la señal en su transmisión.

2. Energía/potencia de la señal transmitida
 - La potencia de la señal transmitida está limitada en la práctica, ya que los amplificadores utilizados en el transmisor tienen una potencia nominal que no se puede exceder por varios tipos de razones. Como la señal transmitida está compuesta por “trozos” en cada intervalo de símbolo, y en cada intervalo aparece una de las M formas de onda $s_i(t)$ que definen el modulador, en la práctica una medida razonable de la energía de la señal es el promedio de la energía de cada una de las señales que la componen, que se denominará energía media por símbolo transmitido (E_s), y que en la práctica estará acotada. Esta energía media por símbolo se obtiene ponderando la energía de cada símbolo (más propiamente, de la señal asociada a cada símbolo) teniendo en cuenta la probabilidad con la que se transmite dicho símbolo:

- Probabilidad de cada símbolo: $p_B(b_i) = P(B[n] = b_i)$
- Energía del símbolo $b_i \equiv$ energía de la señal $s_i(t)$
- Energía media por símbolo: promedio de la energía de los M símbolos

$$E_s = \sum_{i=0}^{M-1} p_B(b_i) \cdot \mathcal{E}\{s_i(t)\}, \text{ siendo } \mathcal{E}\{s_i(t)\} = \int_{-\infty}^{\infty} |s_i(t)|^2 dt.$$

3. Características del canal de comunicaciones ($h(t)$)

- A la hora de diseñar el sistema habrá que tratar de minimizar la distorsión que sufre la señal en la transmisión. Como se vió en la Sección 4.1.4, el efecto del canal habitualmente se modela a través de la relación

$$r(t) = s(t) * h(t) + n(t),$$

es decir, que la señal transmitida sufre fundamentalmente dos efectos: una distorsión lineal dada por la respuesta del canal (descrita por la respuesta al impulso $h(t)$ o respuesta en frecuencia $H(j\omega)$) y la adición de ruido (normalmente ruido térmico). El ruido siempre va a estar presente, y lo único que se puede hacer es filtrar a la entrada del receptor en el ancho de banda de la señal para minimizar su efecto sin distorsionar la señal transmitida. Pero la elección de las señales $s_i(t)$, que determinan las características de la señal transmitida, sí es relevante para minimizar la distorsión lineal; teniendo en cuenta la respuesta del canal, habrá señales que sufran una mayor distorsión, y señales que tengan una distorsión menor. La situación ideal sería que distorsión lineal introducida por el canal fuera nula, lo que teóricamente se podría conseguir si se elige un conjunto apropiado de M señales. Ese conjunto apropiado sería uno que cumpliera que

$$s_i(t) * h(t) = s_i(t) \text{ para } i \in \{0, 1, \dots, M - 1\},$$

o lo que es lo mismo, tal vez sea más fácil de interpretar esta condición en el dominio frecuencial, si se cumple

$$S_i(j\omega) \times H(j\omega) = S_i(j\omega) \text{ para } i \in \{0, 1, \dots, M - 1\}.$$

Si se logra el cumplimiento de esta condición para las M señales, la única distorsión que sufrirá la señal transmitida será la suma de ruido, lo que se denomina modulo de ruido aditivo blanco y gaussiano o modelo de canal gaussiano para abreviar

$$r(t) = s(t) + n(t).$$

Seleccionar M señales atendiendo de forma conjunta a estos tres factores trabajando con las señales en el dominio temporal es complicado, ya que algunos de los factores van incluso en oposición. Por ejemplo, es fácil aumentar la diferencia entre señales con señales de mayor amplitud, pero estas tienen mayor energía. Buscar M señales de una energía limitada, lo más diferentes posibles, y que al mismo tiempo tengan una transformada de Fourier adecuada a la respuesta en frecuencia del canal, es un problema complicado de resolver trabajando con la representación temporal de las señales. Además, está también el problema de determinar la medida de “*diferencia*” que lleva a tener la menor posibilidad de equivocarse en el demodulador a la hora de estimar la señal transmitida en cada intervalo de símbolo. Este problema puede tratarse de una forma más sencilla y eficiente manejando una representación geométrica de las señales a través de un espacio vectorial. Esta representación se presentará en la siguiente sección.

4.2. Representación geométrica de las señales

A la hora de diseñar el modulador, este debe asociar una forma de onda a cada símbolo del alfabeto de $B[n]$, lo que indirectamente supone que a cada b_i se le asocia una señal $s_i(t)$. Si el objetivo de diseño es tener la menor probabilidad de error, intuitivamente es lógico pensar que esas señales han de ser lo más diferentes que sea posible. Pero no es tan fácil buscar una medida de diferencia entre señales, cuando estas están definidas en el dominio temporal, que sea fácil de manejar analíticamente y que sea apropiada para tomar decisiones en un sistema de comunicaciones.

Esta es la razón para emplear una representación geométrica de las señales. Supongamos que cada señal se puede representar dentro de un espacio vectorial, es decir, que cada señal se puede representar por un vector, como se muestra en la Figura 4.9, en la que se representan en línea continua 4 señales en formato vectorial. La señal $s_i(t)$ estará representada por el vector \mathbf{a}_i (asociación basada en los subíndices). Por un lado, si se utiliza una medida de distancia entre vectores, es fácil determinar lo lejos o cerca que están los vectores asociados a dos señales. Y esa medida de distancia puede ser una buena medida del parecido entre señales. Por otro lado, el ruido será también representado como un vector, que se sumará a la señal, y tendremos otro vector resultante (en este caso el vector \mathbf{q} de la figura). Con una medida de distancia entre vectores es fácil decir de qué señal está más cerca el nuevo vector de la señal recibida (en este ejemplo, de la señal $s_0(t)$, al estar el vector \mathbf{q} más próximo al vector \mathbf{a}_0 que a los otros tres vectores).

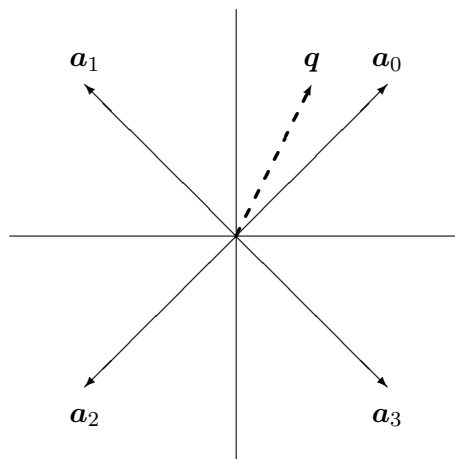


Figura 4.9: Representación geométrica de señales (ejemplo en un espacio bidimensional).

En esta sección se presentará una representación vectorial de las señales que resulta apropiada para el propósito que se persigue en el capítulo, que no es otro que el diseño y análisis de sistemas de comunicaciones digitales. En primer lugar se revisarán las principales características de los espacios vectoriales, y de los espacios de Hilbert, se verá cómo obtener una base ortonormal para la representación vectorial de un conjunto de M señales, y cómo calcular energías y medidas de distancia sobre las representaciones vectoriales de las señales.

4.2.1. Espacios vectoriales

Las señales admiten una representación como vectores de un espacio vectorial. Como vamos a ver, esta representación permite aplicar todas las herramientas de análisis y síntesis de este tipo

de espacios para trabajar con las señales.

Analicemos en primer lugar si una señal puede ser considerada como un vector dentro de un espacio vectorial. Un espacio vectorial \mathbb{V} es un conjunto de elementos, que denominamos vectores, que poseen las siguientes propiedades:

1. Existe una ley de composición interna, que se denomina suma y se representa con el símbolo $+$ que, aplicada a dos vectores $(\mathbf{x}, \mathbf{y}, \in \mathbb{V})$ de la forma $\mathbf{x} + \mathbf{y}$ da como resultado otro vector del espacio $(\mathbf{x} + \mathbf{y} \in \mathbb{V})$, cumpliendo además las siguientes propiedades
 - a) Conmutativa: $\forall \mathbf{x}, \mathbf{y} \in \mathbb{V}; \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
 - b) Asociativa: $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{V}; \mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.
 - c) Existencia de elemento neutro: $\exists \mathbf{0} \in \mathbb{V} : \forall \mathbf{x} \in \mathbb{V}; \mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$.
 - d) Existencia de elemento inverso: $\forall \mathbf{x} \in \mathbb{V} \exists (-\mathbf{x}) : \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

2. Existe una ley de composición externa que denominamos producto con un conjunto \mathbb{C} de elementos denominados escalares (que deben tener la estructura de cuerpo) que, aplicada a un escalar α ($\alpha \in \mathbb{C}$) y a un vector \mathbf{x} ($\mathbf{x} \in \mathbb{V}$) de la forma $\alpha\mathbf{x}$ da como resultado otro vector del espacio ($\alpha\mathbf{x} \in \mathbb{V}$), cumpliendo además las siguientes propiedades:
 - a) Asociativa: $\forall \alpha, \beta \in \mathbb{C}; \forall \mathbf{x} \in \mathbb{V}; \alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$.
 - b) Existencia de elemento neutro: $\exists \mathbf{1} \in \mathbb{C} : \forall \mathbf{x} \in \mathbb{V}; \mathbf{1}\mathbf{x} = \mathbf{x}$.
 - c) Distributiva respecto a la suma: $\forall \alpha \in \mathbb{C}; \forall \mathbf{x}, \mathbf{y} \in \mathbb{V}; \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.
 - d) Distributiva respecto al producto por un escalar:

$$\forall \alpha, \beta \in \mathbb{C}; \forall \mathbf{x} \in \mathbb{V}; (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}.$$

Si se considera el caso general de una señal compleja (tanto en tiempo continuo como en tiempo discreto), esta señal cumple todas las condiciones de un espacio vectorial. La operación suma del espacio vectorial es la suma punto a punto de la señal, que es conmutativa, asociativa, tiene un elemento neutro ($x(t) = 0$) y un elemento inverso, la propia señal invertida (multiplicada por -1).

Si los escalares son el cuerpo de los números complejos, la ley de composición externa es la multiplicación de la señal por un complejo, y es fácil verificar que se cumplen todas las propiedades.

El caso de las señales reales y los números reales también lo cumplen, ya que de hecho no dejan de ser un caso particular de las señales y números complejos.

Sin embargo esta estructura de espacio vectorial genérico es muy simple como para ser de utilidad. Una estructura más elaborada la encontramos en los espacios vectoriales de Hilbert o simplemente espacios de Hilbert.

4.2.2. Espacios de Hilbert para señales de energía finita

Un espacio de Hilbert es, básicamente, un espacio vectorial con producto escalar¹. El producto escalar es una aplicación de pares de vectores en el campo de los escalares (complejos o reales en

¹Estrictamente, es un espacio vectorial con producto escalar que cumple la propiedad de completitud. La completitud se cumple cuando toda sucesión de Cauchy es convergente en la métrica inducida por el producto escalar. Si no posee esta propiedad el espacio vectorial recibe el nombre de espacio de pre-Hilbert.

nuestro caso),

$$f : (\mathbb{V}, \mathbb{V}) \rightarrow \mathbb{C},$$

que se denota como $\langle \mathbf{x}, \mathbf{y} \rangle$, y que cumple las siguientes propiedades:

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^*$
2. $\langle (\alpha \mathbf{x} + \beta \mathbf{y}), \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$
3. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$
4. $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ (vector elemento neutro)

A partir del producto escalar se define una norma para el espacio vectorial como

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle},$$

y a partir de la norma una medida de distancia entre vectores

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

El ángulo entre dos vectores se mide como

$$\theta = \cos^{-1} \left(\frac{\operatorname{Re}\{\langle \mathbf{x}, \mathbf{y} \rangle\}}{\|\mathbf{x}\| \|\mathbf{y}\|} \right).$$

Para las señales y, en general, para cualquier espacio vectorial genérico no existe una única posibilidad de definición de producto escalar: se puede escoger como producto escalar cualquier función que cumpla los requisitos anteriormente establecidos. Cada definición de producto escalar da lugar a un espacio de Hilbert distinto, con métrica y norma distinta. A continuación se presenta la estructura de dos espacios de Hilbert para señales de energía, que son los más apropiados para su aplicación en sistemas de comunicaciones

1. L_2 : espacio de Hilbert para señales de energía en tiempo continuo.
2. ℓ_2 : espacio de Hilbert para señales de energía en tiempo discreto.

Espacio de Hilbert para señales de energía en tiempo continuo

El espacio L_2 se define mediante el siguiente producto escalar

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_{-\infty}^{\infty} x(t) \cdot y^*(t) dt.$$

La norma inducida por este producto escalar es

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\int_{-\infty}^{\infty} |x(t)|^2 dt} = \sqrt{\mathcal{E}\{x(t)\}},$$

es decir, que es la raíz cuadrada de la energía de la señal. Esto es interesante porque la norma de un vector define la distancia de la representación vectorial al origen de coordenadas, por lo que

la energía de una señal se podrá evaluar de forma sencilla como la distancia al cuadrado de su representación vectorial con respecto al origen de coordenadas.

En cuanto a la distancia entre la representación vectorial de dos señales, viene dada por

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\int_{-\infty}^{\infty} |x(t) - y(t)|^2 dt} = \sqrt{\mathcal{E}\{x(t) - y(t)\}}.$$

En este caso se tiene la raíz cuadrada de la energía de la señal diferencia, que intuitivamente parece una medida cuantitativa razonable de la diferencia entre dos señales (posteriormente veremos que en ciertos casos es la medida óptima para decidir con la mínima probabilidad de equivocarse en un demodulador digital).

Espacio de Hilbert para señales de energía en tiempo discreto

El espacio ℓ_2 se define mediante el siguiente producto escalar

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=-\infty}^{\infty} x[n] \cdot y^*[n].$$

La norma inducida por este producto escalar es

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{n=-\infty}^{\infty} |x[n]|^2} = \sqrt{\mathcal{E}\{x[n]\}},$$

y la distancia es la bien conocida distancia euclídea

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{n=-\infty}^{\infty} |x[n] - y[n]|^2} = \sqrt{\mathcal{E}\{x[n] - y[n]\}}.$$

Como se puede ver, las definiciones son equivalentes en ambos espacios, con la única diferencia de la naturaleza continua o discreta del índice temporal (t o n , respectivamente).

Aspectos de interés

En cualquier caso, el producto escalar proporciona una medida del parecido o similitud entre dos señales. El producto escalar de dos señales cuya variación con el tiempo sea similar será “grande”, mientras que el de dos señales cuya variación con el tiempo sea distinta será “pequeño”.

Cuando el producto escalar de dos señales es igual a cero, se dice que esas señales son *ortogonales*. Esto significa que las correspondientes representaciones vectoriales tienen un ángulo de 90° entre ellas.

Hay una relación importante entre las normas de los vectores y el módulo del producto escalar. La desigualdad de Cauchy-Schwarz hace uso de esta relación y establece que

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

La igualdad sólo se cumple si el vector \mathbf{y} es una versión escalada del vector \mathbf{x} , lo que en el caso de señales significa que $y(t) = \alpha \cdot x(t)$ o $y[n] = \alpha \cdot x[n]$. Esta desigualdad toma la forma siguiente para los espacios L_2 y ℓ_2 , respectivamente

$$\left| \int_{-\infty}^{\infty} x(t) \cdot y^*(t) dt \right| \leq \sqrt{\int_{-\infty}^{\infty} |x(t)|^2 dt} \cdot \sqrt{\int_{-\infty}^{\infty} |y(t)|^2 dt}.$$

$$\left| \sum_{n=-\infty}^{\infty} x[n] \cdot y^*[n] \right| \leq \sqrt{\sum_{n=-\infty}^{\infty} |x[n]|^2} \cdot \sqrt{\sum_{n=-\infty}^{\infty} |y[n]|^2}.$$

4.2.3. Representación de una señal en una base del espacio vectorial

El producto escalar también permite encontrar de forma sencilla la representación de una señal en una base del espacio vectorial.

Una base para un espacio de Hilbert \mathbb{H} de dimensión D es un subconjunto de D elementos $\{\mathbf{b}_n\} \in \mathbb{H}$, que determinan un conjunto de D coeficientes únicos $\{c_n(\mathbf{x})\}$, $n \in \{0, 1, \dots, D-1\}$, para cualquier elemento \mathbf{x} del espacio, de modo que el elemento se puede representar de forma unívoca como

$$\mathbf{x} = \sum_{n=0}^{D-1} c_n(\mathbf{x}) \cdot \mathbf{b}_n.$$

A los D coeficientes que definen unívocamente al vector \mathbf{x} , $\{c_n(\mathbf{x})\}_{n=0}^{D-1}$, se les suelen denominar coordenadas del vector en la base.

Se dice que una base es una *base ortogonal* cuando se cumple que los elementos de la base son ortogonales entre ellos, lo que significa que su producto escalar es nulo

$$\langle \mathbf{b}_n, \mathbf{b}_m \rangle = 0, \quad \forall n \neq m.$$

Se dice que una base es una *base ortonormal* cuando es una base ortogonal, es decir, se cumple que los elementos de la base son ortogonales, y además se cumple que cada uno de ellos tiene norma unidad, que teniendo en cuenta la definición de norma a través del producto escalar equivale en este caso a decir que el producto escalar de cada elemento de la base consigo mismo es uno

$$\|\mathbf{b}_n\| = 1 \rightarrow \langle \mathbf{b}_n, \mathbf{b}_n \rangle = 1.$$

En espacios de Hilbert, una de las ventajas de trabajar con una base ortonormal es que los coeficientes de la expansión en términos de la base (coordenadas) se obtienen mediante el producto escalar del vector con los distintos elementos que forman la base ortonormal

$$c_n(\mathbf{x}) = \langle \mathbf{x}, \mathbf{b}_n \rangle.$$

En el caso del espacio L_2 , un ejemplo de base ortonormal trivial lo constituye el conjunto de funciones delta de Dirac $\{\delta(t - \tau), \tau \in \mathbb{R}\}$. Si se denota $\mathbf{b}_i = \delta(t - \tau_i)$,

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \int_{-\infty}^{\infty} \delta(t - \tau_i) \delta(t - \tau_j) dt = \begin{cases} 1, & \tau_i = \tau_j \\ 0, & \tau_i \neq \tau_j \end{cases}.$$

La representación de la señal en términos de esta base es

$$\mathbf{x} \equiv x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau.$$

Los coeficientes de la expansión se obtienen como

$$c_i(\mathbf{x}) = \langle \mathbf{x}, \delta(t - \tau_i) \rangle = \int_{-\infty}^{\infty} x(t) \cdot \delta(t - \tau_i) dt = x(\tau_i).$$

Lo mismo se aplica para el espacio ℓ_2 , donde el conjunto $\{\delta[n - k], k \in \mathbb{Z}\}$, forma una base ortonormal. En este caso se puede denotar $\mathbf{b}_k = \delta[n - k]$

$$\langle \mathbf{b}_k, \mathbf{b}_i \rangle = \sum_{n=-\infty}^{\infty} \delta[n - k] \delta[n - i] = \delta[k - i] = \begin{cases} 1, & k = i \\ 0, & k \neq i \end{cases}.$$

La representación de la señal en términos de esta base es

$$\mathbf{x} \equiv x[n] = \sum_{k=-\infty}^{\infty} x_k \delta[n - k].$$

Los coeficientes de la expansión se obtienen como

$$c_k(\mathbf{x}) = \langle \mathbf{x}, \mathbf{b}_k \rangle = \sum_{n=-\infty}^{\infty} x[n] \cdot \delta[n - k] = x[k].$$

4.2.4. Procedimiento de ortogonalización de Gram-Schmidt

Las dos bases ortonormales presentadas con anterioridad no son prácticas ya que involucran infinitos coeficientes, lo que supone en la práctica conocer la señal completa. Es posible obtener una base ortonormal de dimensión finita que permita representar un conjunto finito de elementos del espacio vectorial (en el caso particular que afecta al diseño y análisis de sistemas de comunicaciones, una base de dimensión finita que permita representar M señales). Una base adecuada para este tipo de representación se puede obtener mediante el denominado procedimiento de *ortogonalización de Gram-Schmidt*. Este procedimiento permite, a partir de M elementos del espacio vectorial que se quieren representar, obtener una base ortonormal de $N \leq M$ elementos a través de la que representar cada uno de los M elementos como una combinación lineal de los elementos de la base, donde los coeficientes (coordenadas) de la combinación lineal determinan la representación vectorial de cada elemento. De este modo es posible representar cada elemento como un punto, de dimensión N , en el espacio vectorial definido por esa base.

En el caso que nos interesa, los M elementos son señales de duración T segundos, y la correspondiente representación vectorial es una representación compacta del conjunto de señales a transmitir por un sistema de comunicaciones que simplifica el diseño y el análisis de su funcionamiento. A continuación se formaliza este problema y se particulariza el procedimiento de Gram-Schmidt para su aplicación en dicho problema.

En un sistema de comunicaciones se tiene un conjunto de M señales $\{s_i(t), i = 0, \dots, M - 1\}$, de duración T segundos y definidas en el intervalo $0 \leq t < T$, que se van a emplear para transmitir información a través del canal de comunicaciones. A partir de este conjunto se va a construir otro

conjunto de N señales ortonormales $\{\phi_j(t), j = 0, \dots, N-1\}$, con $N \leq M$, que permitirá obtener cada una de las señales $s_i(t)$ como una combinación lineal de las mismas.

$$s_i(t) = \sum_{j=0}^{N-1} a_{i,j} \cdot \phi_j(t).$$

Según la definición de una base ortonormal, cualquier señal $s_i(t)$ se puede expresar como combinación lineal de los elementos de la base donde los coeficientes de la expansión, o coordenadas de la representación vectorial para la señal $s_i(t)$, $a_{i,j}$, se obtienen a través del producto escalar

$$a_{i,j} = \langle s_i(t), \phi_j(t) \rangle = \int_{-\infty}^{\infty} s_i(t) \cdot \phi_j^*(t) \cdot dt.$$

Esto da lugar a una representación vectorial (N dimensional) de cada señal a través del vector de coordenadas

$$s_i(t) \rightarrow \mathbf{a}_i = \begin{bmatrix} a_{i,0} \\ a_{i,1} \\ \vdots \\ a_{i,N-1} \end{bmatrix}.$$

Aunque a simple vista puede parecer que este cambio de representación no supone un gran avance, hay que tener en cuenta que de este modo N es la dimensión del espacio de señales, y cada señal se representa con N coordenadas, lo que es más cómodo que manejar la expresión de la señal en tiempo continuo. Además, también se simplifica el cálculo de la energía de las señales $s_i(t)$ y de su diferencia. Es importante recordar que como se trata de una base ortonormal se cumplirá que el producto escalar de dos elementos distintos de la base será nulo, y que el producto escalar de un elemento de la base consigo mismo valdrá uno, es decir

$$\langle \phi_i(t), \phi_j(t) \rangle = \int_{-\infty}^{\infty} \phi_i(t) \cdot \phi_j^*(t) \cdot dt = \delta[i - j].$$

El producto escalar entre dos señales $s_i(t)$ y $s_k(t)$ se calcula como

$$\begin{aligned} \langle s_i(t), s_k(t) \rangle &= \int_{-\infty}^{\infty} s_i(t) \cdot s_k^*(t) dt \\ &= \int_{-\infty}^{\infty} \left(\sum_{j=0}^{N-1} a_{ij} \cdot \phi_j(t) \right) \cdot \left(\sum_{\ell=0}^{N-1} a_{k\ell}^* \cdot \phi_{\ell}^*(t) \right) dt \\ &= \int_{-\infty}^{\infty} \left(\sum_{j=0}^{N-1} \sum_{\ell=0}^{N-1} a_{ij} \cdot a_{k\ell}^* \cdot \phi_j(t) \cdot \phi_{\ell}^*(t) \right) dt \\ &= \sum_{j=0}^{N-1} a_{ij} \cdot a_{kj}^* \int_{-\infty}^{\infty} \phi_j^2(t) dt + \sum_{j=0}^{N-1} \sum_{\substack{\ell=0 \\ \ell \neq j}}^{N-1} a_{ij} a_{k\ell}^* \int_{-\infty}^{\infty} \phi_j(t) \phi_{\ell}^*(t)^* dt, \end{aligned}$$

que aplicando la propiedad de ortonormalidad de la base se reduce a

$$\langle \mathbf{s}_i, \mathbf{s}_k \rangle = \langle s_i(t), s_k(t) \rangle = \sum_{j=0}^{N-1} a_{ij} \cdot a_{kj}^*.$$

La energía de una señal, que en el dominio temporal involucra una integral

$$\mathcal{E}_i = \mathcal{E} \{s_i(t)\} = \int_{-\infty}^{\infty} |s_i(t)|^2 dt$$

se puede calcular por tanto a partir de la representación vectorial de la señal de una forma más sencilla

$$\mathcal{E}_i = \mathcal{E} \{s_i(t)\} = \langle \mathbf{s}_i, \mathbf{s}_i \rangle = \|s_i(t)\|^2 = \sum_{j=0}^{N-1} |a_{ij}|^2.$$

El cálculo a partir de integrales de las señales en el dominio temporal ha sido reemplazado por una suma de N coordenadas al cuadrado, mucho más fácil de realizar.

Lo mismo se puede aplicar al cálculo de la diferencia entre señales. Antes se dijo que una medida intuitivamente razonable (luego veremos además que en muchos casos es la medida óptima) es la energía de la señal diferencia, que en el dominio temporal de nuevo requiere de una integral

$$\int_{-\infty}^{\infty} |s_i(t) - s_k(t)|^2 dt.$$

Con la definición de producto escalar que se ha utilizado, la raíz cuadrada de esta cantidad corresponde a la norma del vector diferencia, es decir, a la distancia entre la representación vectorial de las señales. Resumido, la distancia entre señales se puede obtener a partir de la definición de norma como la norma del vector diferencia (de la diferencia de señales)

$$d(s_i(t), s_k(t)) = \|s_i(t) - s_k(t)\| = \sqrt{\int_{-\infty}^{\infty} |s_i(t) - s_k(t)|^2 dt}.$$

Teniendo en cuenta la expansión lineal que tienen ambas señales, las coordenadas de la señal diferencia son la diferencia de coordenadas de cada una de las señales, y por tanto, la norma de la misma se obtiene como la raíz cuadrada del producto escalar de la señal diferencia consigo misma

$$d(s_i(t), s_k(t)) = \sqrt{\langle \mathbf{s}_i - \mathbf{s}_k, \mathbf{s}_i - \mathbf{s}_k \rangle} = \sqrt{\sum_{j=0}^{N-1} |a_{ij} - a_{kj}|^2}.$$

De nuevo se han sustituido integrales sobre señales por sumas sobre N coordenadas. Por eso resulta conveniente poder representar las señales como vectores obtenidos a partir de combinaciones lineales de un conjunto de señales que forman una base ortonormal.

Para la obtención de una base ortonormal que permita representar M señales, el proceso de ortogonalización de Gram-Schmidt comienza seleccionando una forma de onda de energía no nula. Si $s_0(t)$ es de energía nula, se ha de reordenar el conjunto de señales para tener una nueva señal $s_0(t)$ de energía $\mathcal{E}_0 > 0$. El primer elemento de la base se calcula simplemente normalizando la primera señal

$$\phi_0(t) = \frac{s_0(t)}{\sqrt{\mathcal{E}_0}}.$$

Esto asegura que la energía de $\phi_0(t)$, $\mathcal{E}\{\phi_0(t)\} = 1$, o lo que es lo mismo, que tenemos el primer elemento de una base ortonormal.

Para obtener el segundo elemento de la base, primero se calcula la proyección de $s_1(t)$ sobre $\phi_0(t)$, es decir, la coordenada de dicha señal sobre el primer elemento de la base, que es

$$a_{1,0} = \int_{-\infty}^{\infty} s_1(t) \cdot \phi_0^*(t) dt.$$

Entonces se sustrae esta proyección de la segunda función

$$d_1(t) = s_1(t) - a_{1,0}\phi_0(t).$$

Ahora $d_1(t)$ es ortogonal a $\phi_0(t)$, pero no tiene energía unidad. Si \mathcal{E}_1 denota la energía de $d_1(t)$, es decir

$$\mathcal{E}_1 = \mathcal{E}\{d_1(t)\} = \int_{-\infty}^{\infty} |d_1(t)|^2 dt,$$

el segundo elemento de la base ortonormal es

$$\phi_1(t) = \frac{d_1(t)}{\sqrt{\mathcal{E}_1}},$$

que también tiene energía unidad.

En general, la obtención del elemento $k + 1$ de la base, $\phi_k(t)$ se obtiene como

$$\phi_k(t) = \frac{d_k(t)}{\sqrt{\mathcal{E}_k}},$$

donde

$$d_k(t) = s_k(t) - \sum_{j=0}^{k-1} a_{k,j} \cdot \phi_j(t),$$

$$\mathcal{E}_k = \mathcal{E}\{d_k(t)\} = \int_{-\infty}^{\infty} |d_k(t)|^2 dt,$$

y

$$a_{k,j} = \int_{-\infty}^{\infty} s_k(t) \cdot \phi_j^*(t) dt.$$

A continuación se presenta un ejemplo para el cálculo de una base que permita representar 4 señales.

Ejemplo

Se aplica el procedimiento de Gram-Schmidt al conjunto de señales de la Figura 4.10.

Para obtener el primer elemento de la base calculamos la energía de $s_0(t)$. Es fácil comprobar que $\mathcal{E}_0 = 2$. Por tanto

$$\phi_0(t) = \frac{s_0(t)}{\sqrt{2}}.$$

Luego se calcula la proyección de $s_1(t)$ sobre $\phi_0(t)$.

$$a_{1,0} = \int_{-\infty}^{\infty} s_1(t) \cdot \phi_0^*(t) dt = 0.$$

Por tanto, en este caso $d_1(t) = s_1(t)$, que es ortogonal a $\phi_0(t)$ pero no es ortonormal. Hay que dividir por la energía

$$\mathcal{E}_1 = 2.$$

$$\phi_1(t) = \frac{s_1(t)}{\sqrt{2}}.$$

Para calcular $\phi_2(t)$ hay que calcular las proyecciones de $s_2(t)$ sobre $\phi_0(t)$ y $\phi_1(t)$.

$$a_{2,0} = \int_{-\infty}^{\infty} s_2(t) \cdot \phi_0^*(t) dt = 0.$$

$$a_{2,1} = \int_{-\infty}^{\infty} s_2(t) \cdot \phi_1^*(t) dt = -\sqrt{2}.$$

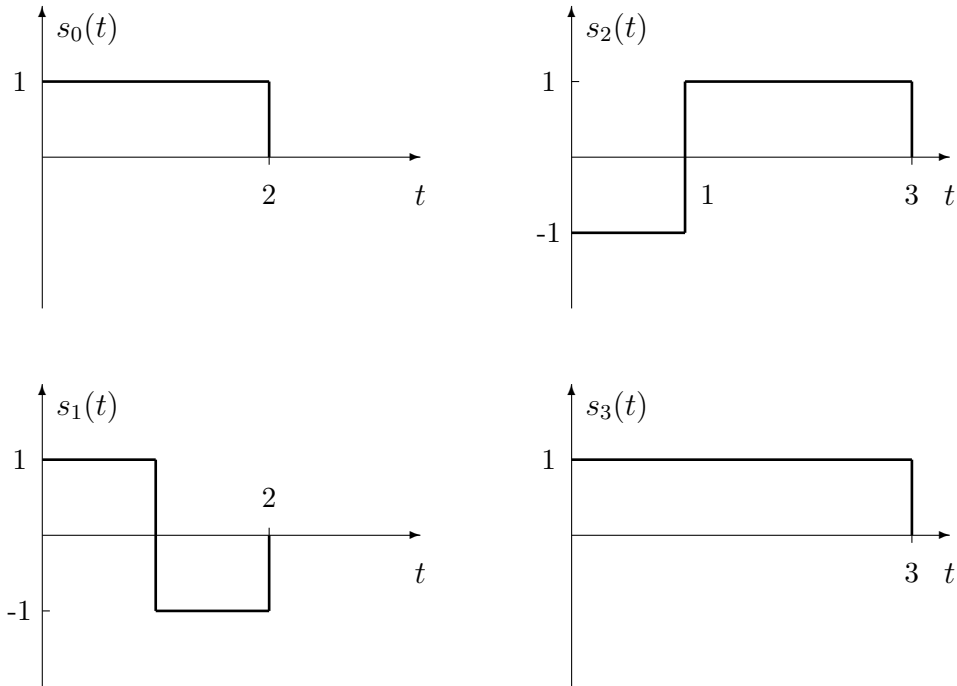


Figura 4.10: Conjunto original de señales.

Por tanto

$$\begin{aligned} d_2(t) &= s_2(t) - a_{20} \cdot \phi_0(t) - a_{21} \cdot \phi_1(t) \\ &= s_2(t) + \sqrt{2}\phi_1(t). \end{aligned}$$

La energía de $d_2(t)$ es

$$\mathcal{E}_2 = 1.$$

$$\phi_2(t) = d_2(t) = s_2(t) + \sqrt{2}\phi_1(t).$$

Ahora aplicamos el procedimiento sobre la última de las señales, $s_3(t)$. En primer lugar vamos a calcular las proyecciones de esta señal sobre los 3 elementos de la base ya disponibles:

$$a_{3,0} = \int_{-\infty}^{\infty} s_3(t) \cdot \phi_0^*(t) dt = \sqrt{2}.$$

$$a_{3,1} = \int_{-\infty}^{\infty} s_3(t) \cdot \phi_1^*(t) dt = 0.$$

$$a_{3,2} = \int_{-\infty}^{\infty} s_3(t) \cdot \phi_2^*(t) dt = 1.$$

Por tanto, ahora

$$d_3(t) = s_3(t) - a_{3,0} \cdot \phi_0(t) - a_{3,1} \cdot \phi_1(t) - a_{3,2} \cdot \phi_2(t) = 0.$$

Esto quiere decir que no es necesario ningún elemento adicional en la base para poder representar $s_3(t)$, ya que esta puede obtenerse como una combinación lineal de los 3 elementos de la base ya obtenidos anteriormente. Así pues, en este caso, la dimensión de este espacio de señales es 3, y las 3 funciones base se muestran en la Figura 4.11.

En resumen, una vez calculadas las N señales de la base ortonormal

$$\{\phi_j(t), j = 0, \dots, N - 1\}$$

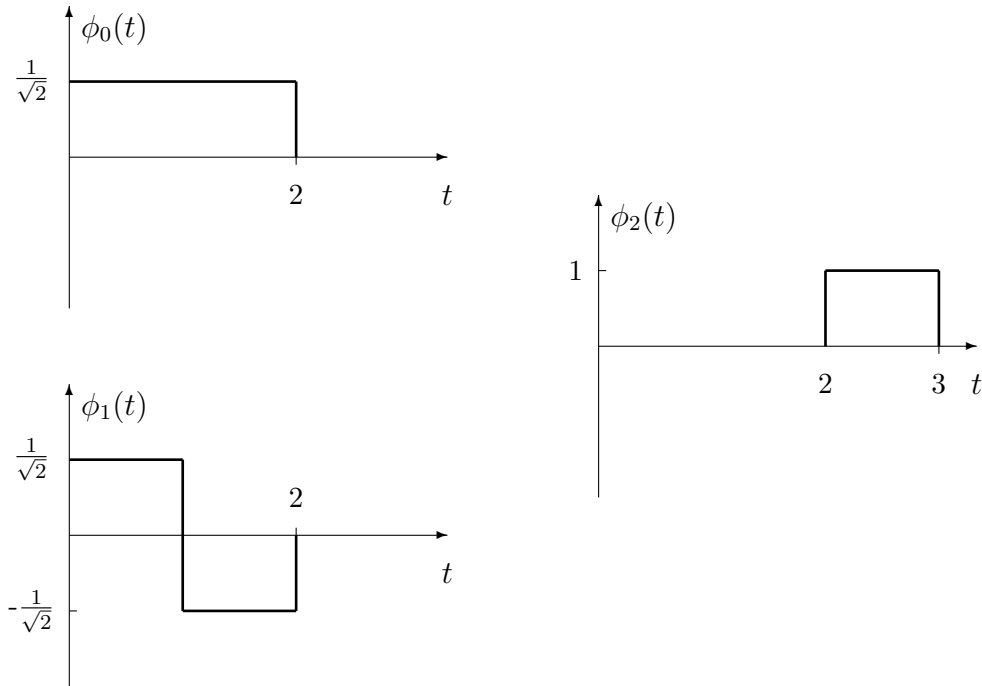


Figura 4.11: Conjunto de señales ortonormales.

cada señal se puede obtener como una combinación lineal de los N elementos de la base

$$s_i(t) = \sum_{j=0}^{N-1} a_{ij} \phi_j(t), \quad i = 0, \dots, M - 1.$$

Por tanto, cada señal puede ser representada mediante un vector

$$\mathbf{a}_i = \begin{bmatrix} a_{i,0} \\ a_{i,1} \\ \vdots \\ a_{i,N-1} \end{bmatrix}.$$

o equivalentemente, como un punto en un espacio de señales N -dimensional con coordenadas $(a_{i,0}, a_{i,1}, \dots, a_{i,N-1})$.

Por otro lado, la potencia de la señal $s_i(t)$, \mathcal{E}_i , se puede calcular de forma sencilla a partir de dichas coordenadas

$$\begin{aligned} \mathcal{E}_i &= \int_{-\infty}^{\infty} |s_i(t)|^2 dt, \\ \mathcal{E}_i &= \langle \mathbf{s}_i, \mathbf{s}_i \rangle = \sum_{j=0}^{N-1} |a_{i,j}|^2. \end{aligned}$$

Ejemplo

Las coordenadas de las 4 señales del ejemplo anterior (que se han ido calculando durante el desarrollo del procedimiento de ortogonalización) son

$$\left. \begin{array}{l} a_{0,0} = \sqrt{2} \\ a_{0,1} = 0 \\ a_{0,2} = 0 \end{array} \right\} \rightarrow \mathbf{a}_0 = [\sqrt{2}, 0, 0]^T.$$

$$\left. \begin{aligned} a_{1,0} &= 0 \\ a_{1,1} &= \sqrt{2} \\ a_{1,2} &= 0 \end{aligned} \right\} \rightarrow \mathbf{a}_1 = [0, \sqrt{2}, 0]^T.$$

$$\left. \begin{aligned} a_{2,0} &= 0 \\ a_{2,1} &= -\sqrt{2} \\ a_{2,2} &= 1 \end{aligned} \right\} \rightarrow \mathbf{a}_2 = [0, -\sqrt{2}, 1]^T.$$

$$\left. \begin{aligned} a_{3,0} &= \sqrt{2} \\ a_{3,1} &= 0 \\ a_{3,2} &= 1 \end{aligned} \right\} \rightarrow \mathbf{a}_3 = [\sqrt{2}, 0, 1]^T.$$

Estas señales se pueden representar como puntos en un espacio 3D, tal y como se muestra en la Figura 4.12

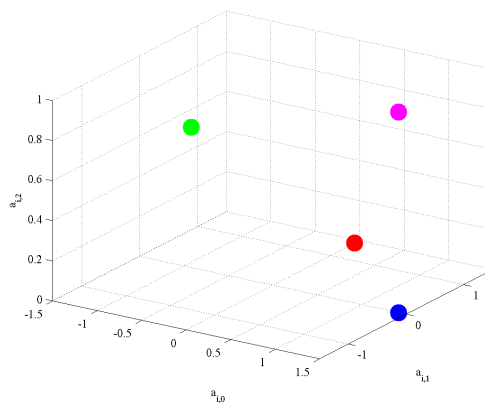


Figura 4.12: Señales representadas en la base ortonormal.

Es fácil comprobar que

$$\mathcal{E}_0 = 2, \mathcal{E}_1 = 2, \mathcal{E}_2 = 3, \mathcal{E}_3 = 3.$$

También se pueden calcular de forma sencilla las distancias entre señales, que en este caso son

$$d(\mathbf{s}_0, \mathbf{s}_1) = 2, \quad d(\mathbf{s}_0, \mathbf{s}_2) = \sqrt{5}, \quad d(\mathbf{s}_0, \mathbf{s}_3) = 1$$

$$d(\mathbf{s}_1, \mathbf{s}_2) = \sqrt{9}, \quad d(\mathbf{s}_1, \mathbf{s}_3) = \sqrt{5}, \quad d(\mathbf{s}_2, \mathbf{s}_3) = 2$$

Es importante hacer notar que la base para representar un conjunto de M señales no es única. Si se realiza el procedimiento de Gram-Schmidt cambiando el orden de las señales, la base obtenida, y las correspondientes coordenadas son otras, pero las energías y distancias entre señales se mantienen. Es decir, el resultado, si se realiza una representación de las señales, es que los nuevos puntos estarán rotados.

También es posible ver para un conjunto de señales una base adecuada para su representación. Por ejemplo, para estas señales se podrían utilizar como bases las siguientes señales:

$$\phi_0(t) = \begin{cases} 1, & \text{si } 0 \leq t < 1 \\ 0, & \text{en otro caso} \end{cases}.$$

$$\phi_1(t) = \begin{cases} 1, & \text{si } 1 \leq t < 2 \\ 0, & \text{en otro caso} \end{cases}.$$

$$\phi_2(t) = \begin{cases} 1, & \text{si } 2 \leq t < 3 \\ 0, & \text{en otro caso} \end{cases}$$

En este caso, es trivial ver que

$$\mathbf{a}'_0 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{a}'_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \mathbf{a}'_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{a}'_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Es fácil comprobar como se mantienen energías (que se relacionan con la distancia de cada punto al origen) y distancias entre las representaciones vectoriales utilizando esta nueva representación.

Por tanto, la representación vectorial correspondiente supone una rotación de la representación utilizando la otra base ortonormal, tal y como se puede ver a en en la Figura 4.13, que muestra la representación vectorial de las 4 señales en la nueva base ortonormal.

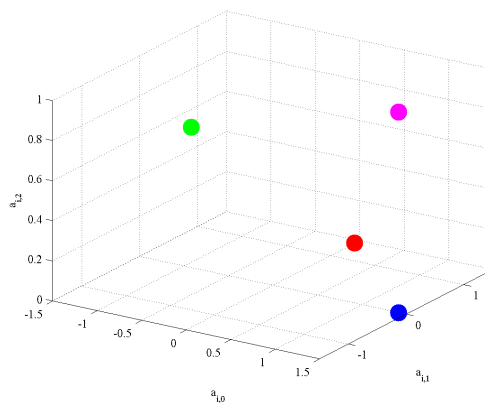


Figura 4.13: Señales representadas en la base ortonormal alternativa.

4.3. Modelo de comunicación digital

La representación vectorial de las señales a través de una base ortonormal de dimensión N permite utilizar un modelo de comunicación digital basado en esta representación que simplifica notablemente el diseño y análisis del sistema. El esquema de este modelo consta de 4 elementos funcionales (además del canal de comunicaciones), y se muestra en la Figura 4.14.

En este modelo se dividen tanto el modulador como el demodulador digital en dos módulos:

- División del modulador digital en dos módulos
 - Codificador
 - Modulador
- División del demodulador digital en dos módulos
 - Demodulador
 - Decisor

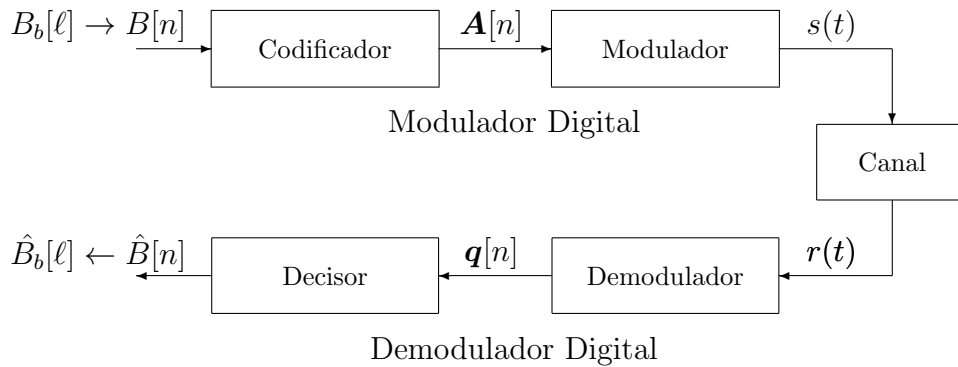


Figura 4.14: Modelo básico de un sistema de comunicaciones digitales.

Hay que recordar que la tarea del modulador y demodulador digital era convertir bits (agrupados en símbolos) en señales y hacer la conversión inversa, respectivamente. Con esta división en módulos se ha introducido en ambos casos una representación vectorial de las señales intermedia, $\mathbf{A}[n]$ en el transmisor y $\mathbf{q}[n]$ en el receptor. Van a ser precisamente estas representaciones vectoriales intermedias las que faciliten notablemente el diseño y análisis del sistema.

En el transmisor, ahora la asignación de una señal $s_i(t)$ a cada valor del alfabeto de símbolos b_i , se realiza ahora en dos pasos:

- Se hace una asignación de la representación vectorial de $s_i(t)$, \mathbf{a}_i , para b_i (codificador).
- Se convierte la representación vectorial de la señal, \mathbf{a}_i , a señal $s_i(t)$ (modulador). Para ello hay que definir una base ortonormal de dimensión N , ya que

$$s_i(t) = \sum_{j=0}^{N-1} a_{i,j} \times \phi_j(t).$$

Esta división simplifica notablemente el problema de la elección de las M señales para tener en cuenta los 3 factores que se especificaron en la Sección 4.1.6: prestaciones (parecido/diferencia entre señales), energía y características del canal. Esto es así porque los dos primeros factores se desacoplan del tercero. Como se ha visto, la energía y el parecido entre señales se pueden evaluar a partir de la representación vectorial de las señales, independientemente de la base ortonormal elegida. De este modo, el diseño del codificador, que consiste en elegir la representación vectoriales de las M señales,

$$\{\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_{M-1}\},$$

que se denominará habitualmente constelación, se realizará teniendo en cuenta estos dos factores: prestaciones (medida de diferencia entre las señales) y energía.

En cuanto a las características del canal, hay que hacer que las señales tengan una respuesta en frecuencia adecuada a la respuesta en frecuencia del canal. La forma de la respuesta en frecuencia de la señal $s_i(t)$, dado que

$$s_i(t) = \sum_{j=0}^{N-1} a_{i,j} \times \phi_j(t) \leftrightarrow S_i(j\omega) = \sum_{j=0}^{N-1} a_{i,j} \times \Phi_j(j\omega),$$

depende de la elección de la base ortonormal. Elegir una base ortonormal u otra no tendrá ninguna repercusión en la energía y medida de diferencia (distancia) entre señales, ya que estos parámetros

quedan completamente definidos por la representación vectorial de las señales. Lo que se modificará será la forma de las señales, tanto en el dominio temporal como frecuencial. De este modo, el diseño del modulador, que consistirá en la elección de una base ortonormal apropiada, se realizará teniendo en cuenta sólo el tercer factor, es decir, para que las señales resulten adecuadas a las características del canal.

En cuanto al receptor, la división en dos módulos también simplifica el procedimiento. En primer lugar, el demodulador obtiene la representación vectorial de la señal recibida en cada intervalo de símbolo, $\mathbf{q}[n]$ para el intervalo $nT \leq t < T$. Y a partir de esa representación, será mucho más sencillo buscar la regla óptima para decidir cuál de las M señales se transmitió en ese intervalo (la regla que produzca menos errores) que trabajando sobre la representación en tiempo continuo de las señales.

A continuación se resume brevemente la función y principales características de cada uno de los 4 elementos funcionales del sistema:

■ Codificador

- Define la representación vectorial de la señal asociada a cada símbolo (constelación)
 - Intervalo de índice n : vector $\mathbf{A}[n]$ representando $s(t)$ en $nT \leq t < (n+1)T$
- Criterios de diseño (para seleccionar la constelación)
 - Energía
 - Distancia (“*parecido*”) entre señales (prestaciones)

■ Modulador

- Define la base ortonormal del espacio de señales
- Criterios de diseño (para seleccionar las N funciones $\phi_j(t)$, para $i = 0, 1, \dots, N-1$)
 - Adecuación al canal de comunicaciones

■ Demodulador

- Convierte la señal recibida, por intervalos de símbolo, en vectores en el espacio de señales definido por la base $\{\phi_j(t)\}_{j=0}^{N-1}$
 - Intervalo de índice n : vector $\mathbf{q}[n]$ representando $r(t)$ en $nT \leq t < (n+1)T$

■ Decisor

- Compara el “*parecido*” entre la señal recibida y las M posibles señales $s_i(t)$ para decidir símbolos
 - Medida de distancia sobre las representaciones vectoriales
 - Compara las distancias de:
 - Vector de la señal recibida en el intervalo de símbolo: $\mathbf{q}[n]$
 - Vectores de los M posibles símbolos: \mathbf{a}_i , para $i \in \{0, 1, \dots, M-1\}$

4.3.1. Ejemplo para ilustrar la ventaja de la representación vectorial en el diseño de un sistema

Empezaremos considerando un sistema de comunicación binaria en el que cada símbolo corresponde a 1 bit ($m = 1, M = 2^m = 2$) y los dos bits se transmiten con la misma probabilidad. Como señales para la transmisión de cada símbolo, se puede escoger de entre cuatro posibles conjuntos de señales $\{s_0(t), s_1(t)\}$ para realizar la transmisión. Estos conjuntos se representan en las Figuras 4.15, 4.16, 4.17 y 4.18 respectivamente.

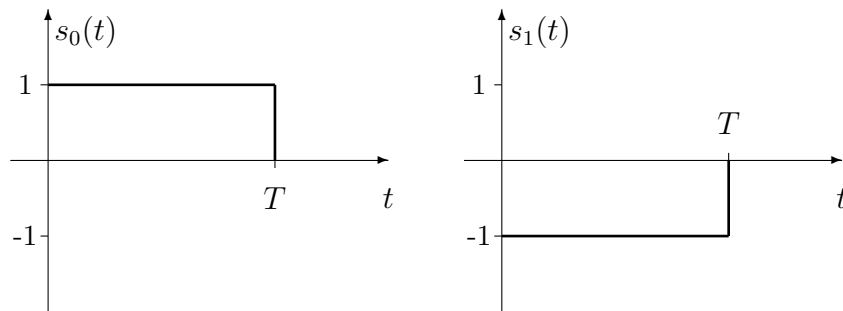


Figura 4.15: Primer conjunto de señales.

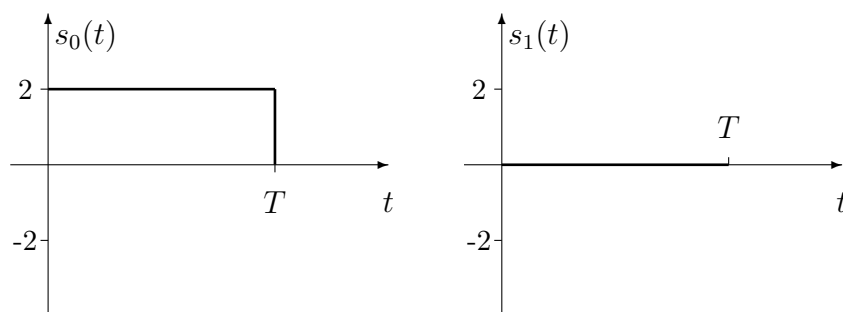


Figura 4.16: Segundo conjunto de señales.

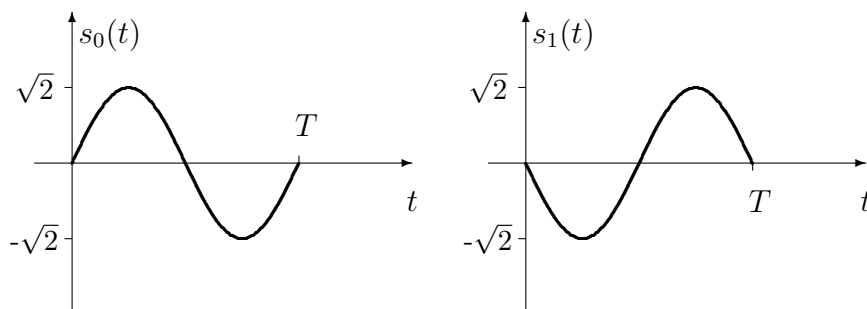


Figura 4.17: Tercer conjunto de señales

Para decidir qué conjunto de señales es el más apropiado hay que recordar que el objetivo es poder recuperar el símbolo enviado con la menor probabilidad de error teniendo en cuenta las características del canal; en este caso se suma ruido blanco gaussiano. Parece lógico pensar que si mayor es la “separación” o “diferencia” entre $s_0(t)$ y $s_1(t)$ más fácil será discernir entre ambas en el receptor pese al ruido sumado. Para medir esta separación se necesita una medida de distancia entre señales.

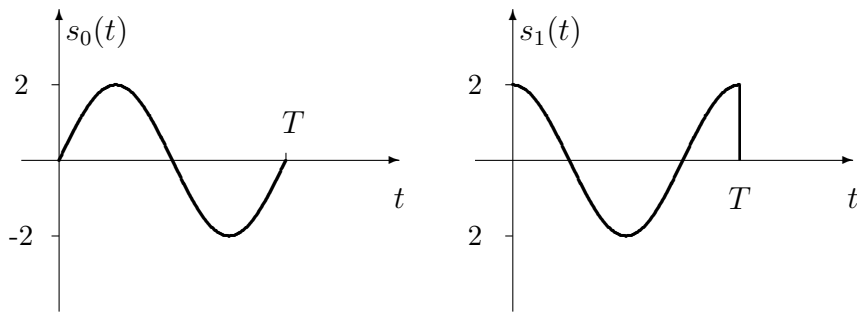


Figura 4.18: Cuarto conjunto de señales

Se va a utilizar la medida de distancia definida en el espacio de Hilbert de señales de energía finita, L_2 . En este espacio, las señales $s_i(t)$ pasan a ser vectores \mathbf{a}_i y la distancia se define como

$$d(s_i(t), s_j(t)) = \|\mathbf{a}_i - \mathbf{a}_j\| = \sqrt{\int_{-\infty}^{\infty} |s_i(t) - s_j(t)|^2 dt}.$$

Aplicando esta medida a las dos señales del primer conjunto

$$d(s_0(t), s_1(t)) = \sqrt{\int_0^T |1 - (-1)|^2 dt} = 2\sqrt{T}.$$

Para el segundo conjunto de señales tenemos

$$d(s_0(t), s_1(t)) = \sqrt{\int_0^T |2 - 0|^2 dt} = 2\sqrt{T}.$$

Para el tercer conjunto de señales

$$\begin{aligned} d(s_0(t), s_1(t)) &= \sqrt{\int_0^T \left| \sqrt{2} \sin\left(\frac{2\pi t}{T}\right) - \left(-\sqrt{2} \sin\left(\frac{2\pi t}{T}\right)\right) \right|^2 dt} \\ &= \sqrt{\int_0^T 8 \sin^2\left(\frac{2\pi t}{T}\right) dt} \\ &= \sqrt{4 \left[t - \frac{T}{2\pi} \sin\left(\frac{2\pi t}{T}\right) \cos\left(\frac{2\pi t}{T}\right) \right]_0^T} \\ &= 2\sqrt{T}. \end{aligned}$$

Y para el último conjunto de señales

$$\begin{aligned} d(s_0(t), s_1(t)) &= \sqrt{\int_0^T \left| 2 \sin\left(\frac{2\pi t}{T}\right) - \left(2 \cos\left(\frac{2\pi t}{T}\right)\right) \right|^2 dt} \\ &= \sqrt{\int_0^T 4 - 8 \sin\left(\frac{2\pi t}{T}\right) \cos\left(\frac{2\pi t}{T}\right) dt} \\ &= 2\sqrt{T}, \end{aligned}$$

ya que

$$\int_0^T 8 \sin\left(\frac{2\pi t}{T}\right) \cos\left(\frac{2\pi t}{T}\right) dt = \left[\frac{2T}{\pi} \sin^2\left(\frac{2\pi t}{T}\right) \right]_0^T = 0.$$

Así pues, los cuatro conjuntos tienen la misma separación y en principio las señales se comportarían igual ante la perturbación que introduce el canal. Pero también debemos preguntarnos si les suponen el mismo esfuerzo al transmisor enviarlos a través del canal. Y la forma de medir el esfuerzo es la *energía media por símbolo* que se define como

$$\begin{aligned} E_s &= E[\mathcal{E}\{s(t)\}] \\ &= E\left[\int_{-\infty}^{\infty} |s(t)|^2 dt\right] \\ &= \sum_{i=0}^{M-1} P(s(t) = s_i(t)) \times \mathcal{E}\{s_i(t)\} \\ &= \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \int_{-\infty}^{\infty} |s_i(t)|^2 dt. \end{aligned}$$

Si los símbolos son equiprobables, $p_A(\mathbf{a}_i) = \frac{1}{M}$, por lo que

$$E_s = \sum_{i=0}^{M-1} \frac{1}{M} \int_{-\infty}^{\infty} |s_i(t)|^2 dt.$$

Aplicando esta definición al primer conjunto de señales

$$E_s = \frac{1}{2} \int_0^T |1|^2 dt + \frac{1}{2} \int_0^T |-1|^2 dt = \frac{1}{2}T + \frac{1}{2}T = T.$$

Si se calcula esta energía para el resto de conjuntos se obtiene, respectivamente, $E_s = 2T$, $E_s = T$ y $E_s = 2T$. Se puede decir por tanto que los mejores conjuntos son el primero y el tercero, pues tienen la misma distancia con la menor energía.

Cuando se manejan conjuntos de señales con diferente distancia entre señales, es necesario emplear algún tipo de normalización que nos permita establecer su comparación. La más común de estas normalizaciones consiste en expresar la distancia entre señales en función de la energía media por símbolo. Sobre el primer conjunto de símbolos, esta distancia normalizada es

$$d(s_0(t), s_1(t)) = 2\sqrt{T} = 2\sqrt{E_s}.$$

Para el resto de conjuntos, respectivamente

$$d(s_0(t), s_1(t)) = 2\sqrt{T} = \sqrt{2E_s},$$

$$d(s_0(t), s_1(t)) = 2\sqrt{T} = 2\sqrt{E_s},$$

y

$$d(s_0(t), s_1(t)) = 2\sqrt{T} = \sqrt{2E_s}.$$

Se observa que los conjuntos de señales 1 y 3 proporcionan una mayor distancia de separación entre señales que los conjuntos 2 y 4 para un mismo valor de E_s .

Otra cuestión a plantearse es si es posible encontrar una forma de representar las señales que nos permita ver de manera más simple las prestaciones que vamos a obtener con un conjunto determinado. La respuesta la encontramos en la misma estructura de espacio vectorial que nos ha permitido medir la distancia entre señales: si somos capaces de encontrar una base ortonormal que nos permita representar un determinado conjunto de señales podremos trabajar directamente con las coordenadas en esa base y evitar cualquier tipo de cálculo sobre las formas de onda.

Expresado formalmente, se busca una representación para cada una de las señales del conjunto $\{s_i(t) : i = 0, \dots, M - 1\}$ de la forma

$$s_i(t) = \sum_{j=0}^{N-1} a_{i,j} \phi_j(t),$$

donde $\{\phi_j(t) : j = 0, \dots, N-1\}$ son los elementos de la base ortonormal, y $\{a_{i,j} : j = 0, \dots, N-1\}$ son las coordenadas de la señal $s_i(t)$ en dicha base, que se agrupan en el vector \mathbf{a}_i . Así pues, los símbolos \mathbf{a}_i corresponden a las coordenadas de la señal de índice i .

Si se obtiene la base para el primer conjunto de señales, se puede comprobar que

$$s_0(t) = \sqrt{T} \phi_0(t), \quad s_1(t) = -\sqrt{T} \phi_0(t),$$

donde $\phi_0(t) = s_0(t)/\sqrt{T}$, que se representa en la Figura 4.19.

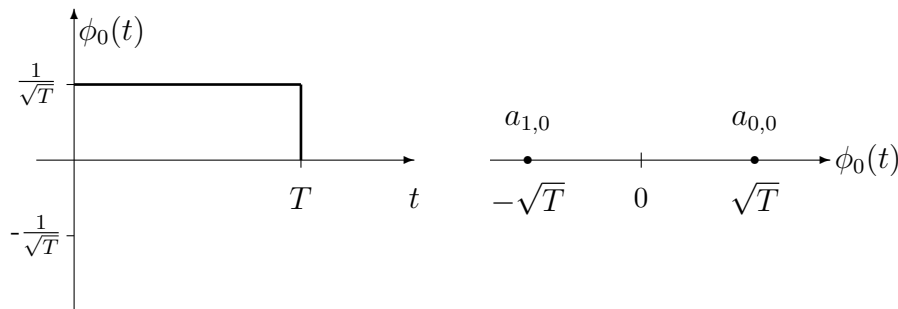


Figura 4.19: Ortogonalización para el primer conjunto de señales.

El hecho de obtener un único elemento de la base, $\phi_0(t)$, tiene una justificación muy sencilla: como $s_1(t) = -s_0(t)$, un cambio de signo en la coordenada de $s_0(t)$ nos basta para generar $s_1(t)$. La amplitud de $\phi_0(t)$ se toma para cumplir la condición de ortonormalidad, que no es otra cosa que la normalización de la energía ($\int_{-\infty}^{\infty} \phi_0^2(t) dt = 1$). Ahora es fácil calcular la distancia entre las señales y su energía

$$d(s_0(t), s_1(t)) = d(\mathbf{a}_0, \mathbf{a}_1) = \sqrt{(a_{0,0} - a_{1,0})^2} = \sqrt{(\sqrt{T} - (-\sqrt{T}))^2} = 2\sqrt{T}.$$

La energía por símbolo es

$$E_s = E[\mathcal{E}\{s(t)\}] = E[\mathcal{E}\{\mathbf{a}_i\}] = p_A(\mathbf{a}_0)a_{0,0}^2 + p_A(\mathbf{a}_1)a_{1,0}^2 = \frac{1}{2}T + \frac{1}{2}T = T.$$

El resultado, como no podía ser de otra forma, es igual al que hemos obtenido calculando sobre las formas de onda.

Para el segundo conjunto de señales se obtiene la base y coordenadas mostradas en la Figura 4.20.

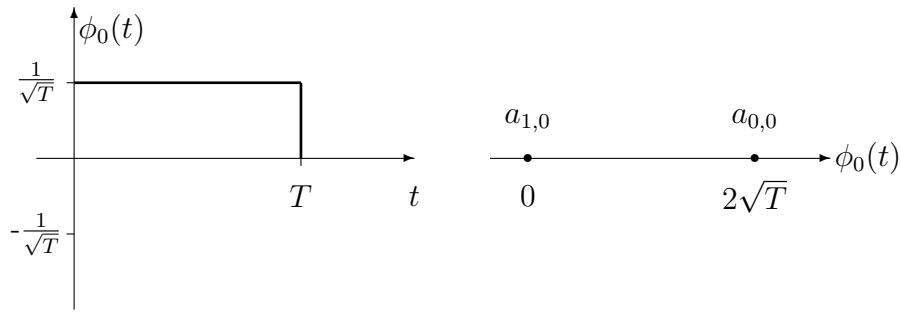


Figura 4.20: Ortogonalización para el segundo conjunto de señales.

Si se compara con el resultado obtenido para el primer conjunto, se observa que el elemento único de la base es la misma señal y que las coordenadas simplemente se han desplazado un valor $\sqrt{2}$. Este desplazamiento hace que se mantenga la distancia entre señales, de modo que frente al ruido se comportaría igual que el conjunto primero. Sin embargo, aumenta la energía media de la constelación

$$E_s = E[\mathcal{E}\{\mathbf{a}_i\}] = p_A(\mathbf{a}_0)a_{0,0}^2 + p_A(\mathbf{a}_1)a_{1,0}^2 = \frac{1}{2}4T = 2T.$$

La orthogonalización del tercer conjunto se muestra en la Figura 4.21.

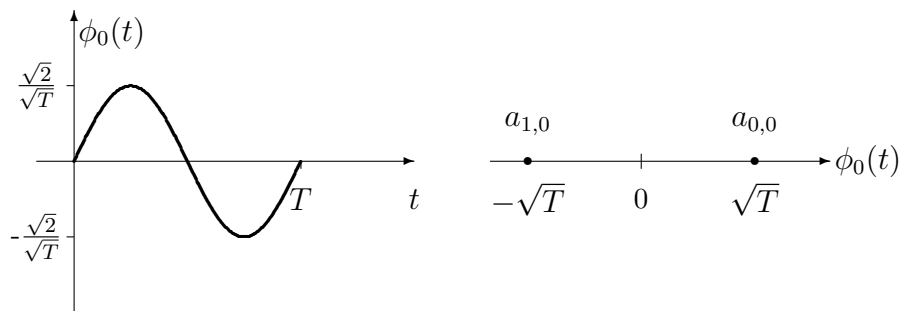


Figura 4.21: Ortogonalización para el tercer conjunto de señales.

La orthogonalización del cuarto conjunto se muestra en la Figura 4.22.

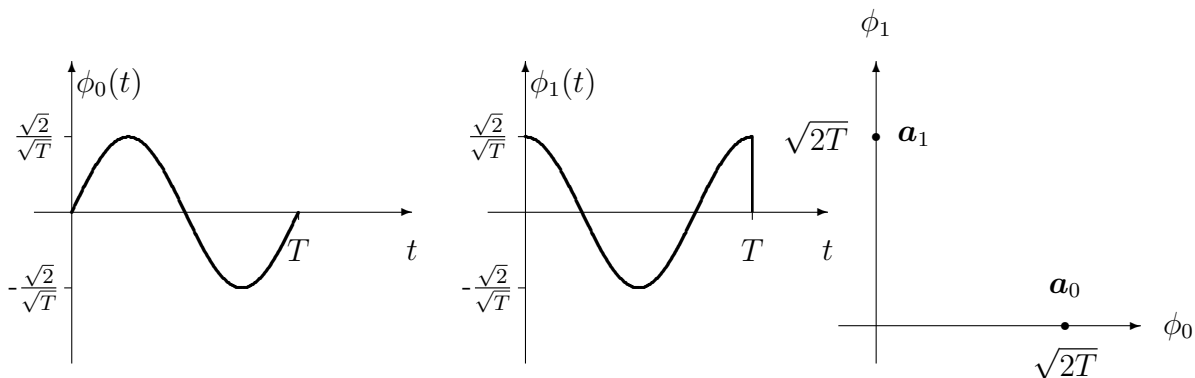


Figura 4.22: Ortogonalización para el cuarto conjunto de señales.

A diferencia de los conjuntos anteriores, aparecen dos elementos en la base del espacio vectorial. Si anteriormente los símbolos de la constelación \mathbf{a}_0 y \mathbf{a}_1 podían ser interpretados como puntos sobre

una recta, ahora se interpretan como puntos sobre un plano cuyos ejes representan las coordenadas sobre $\phi_0(t)$ y $\phi_1(t)$. Los símbolos \mathbf{a}_0 y \mathbf{a}_1 se expresan ahora como vectores de dimensión 2 cuyos valores son

$$\mathbf{a}_0 = \begin{bmatrix} a_{0,0} \\ a_{0,1} \end{bmatrix} = \begin{bmatrix} \sqrt{2T} \\ 0 \end{bmatrix} \quad \mathbf{a}_1 = \begin{bmatrix} a_{1,0} \\ a_{1,1} \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2T} \end{bmatrix}.$$

Ahora la distancia se calcula como la distancia sobre el plano, esto es

$$\begin{aligned} d(s_0(t), s_1(t)) &= d(\mathbf{a}_0, \mathbf{a}_1) = \|\mathbf{a}_0 - \mathbf{a}_1\| \\ &= \sqrt{(a_{0,0} - a_{1,0})^2 + (a_{0,1} - a_{1,1})^2} \\ &= \sqrt{(\sqrt{2T})^2 + (-\sqrt{2T})^2} = 2\sqrt{T}. \end{aligned}$$

Y la energía media por símbolo es

$$\begin{aligned} E_s &= E[\mathcal{E}\{\mathbf{a}_i\}] = E[\|\mathbf{a}_i\|^2] \\ &= p_A(\mathbf{a}_0)\|\mathbf{a}_0\|^2 + p_A(\mathbf{a}_1)\|\mathbf{a}_1\|^2 \\ &= p_A(\mathbf{a}_0)(a_{0,0}^2 + a_{0,1}^2) + p_A(\mathbf{a}_1)(a_{1,0}^2 + a_{1,1}^2) \\ &= \frac{1}{2}2T + \frac{1}{2}2T = 2T. \end{aligned}$$

Si se comparan el primer conjunto y el tercero, se observa que el elemento de la base de señales es distinto pero la constelación es la misma. Si la constelación es la misma, son idénticas la separación entre señales y la energía media de la constelación, o energía media por símbolo. A partir de este hecho podemos extraer una conclusión: la elección del conjunto de elementos de la base $\{\phi_j(t), j = 0, \dots, N - 1\}$ no va a afectar a la fiabilidad de la comunicación ni a la energía de la señal; la fiabilidad y la energía quedarán fijadas por la constelación.

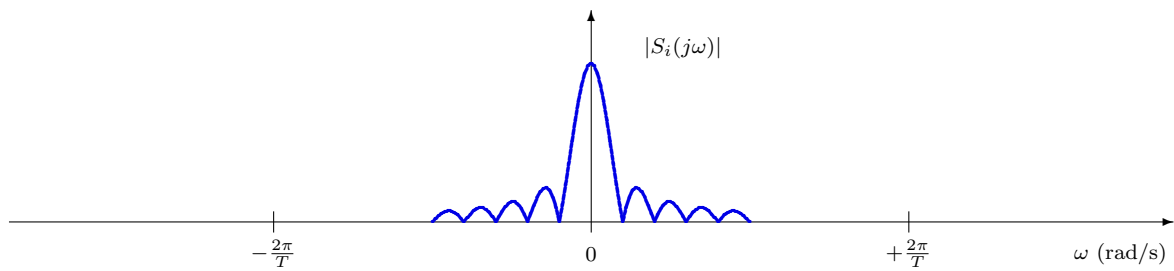
Para este ejemplo, el mejor compromiso entre prestaciones y energía se obtiene con los conjuntos 1 y 3, ya que comparten constelación, por lo que la diferencia (distancia) entre señales y la energía de cada señal es la misma. ¿Cuándo será más adecuado utilizar el conjunto 1, y cuando el conjunto 3? ¿De qué depende la elección de una u otra base ortonormal? Para responder a esta pregunta hay que considerar qué cambia al elegir entre una u otra base ortonormal. Lo que se modifica es la forma de las señales, en el dominio temporal (visto en las Figuras 4.15 y 4.17) y en el dominio frecuencial. La respuesta en este dominio para los cuatro conjuntos, que depende de la transformada de Fourier de las funciones base, se muestra en la Figura 4.23.

Se puede ver que las señales de los conjuntos 1 y 2 tienen su respuesta en frecuencia centrada en 0 Hz, con lo que son señales apropiadas para transmisión en canales con “buena respuesta” en bajas frecuencias. Las señales de los conjuntos 3 y 4 tienen su respuesta en frecuencia centrada en $\omega = \frac{2\pi}{T}$ rad/s, por lo que son señales apropiadas para transmisión en canales con “buena respuesta” en torno a la frecuencia $\frac{2\pi}{T}$ radianes/s. Es decir, que la elección de una u otra base ortonormal, dependerá de las características del canal por el que se ha de transmitir.

Esta conclusión no es específica de los conjuntos de señales que estamos evaluando, sino que es una conclusión general para el diseño de transmisores. En una situación en la que el objetivo sea diseñar el transmisor desde el inicio, la constelación y la base se diseñan prácticamente de forma independiente:

- La constelación atendiendo a criterios de fiabilidad de la comunicación y de energía de la señal, que estará limitada.

- Conjunto 1 y Conjunto 2



- Conjunto 3 y Conjunto 4

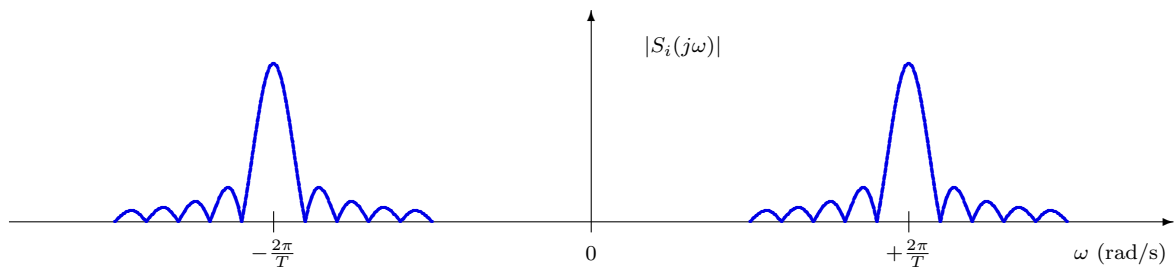


Figura 4.23: Respuesta en frecuencia de las señales de los cuatro conjuntos propuestos en el ejemplo.

- Los elementos de la base atendiendo al comportamiento del canal de comunicaciones (intendiendo que el canal físico se comporte como el modelo de ruido aditivo que hemos considerado).

La interpretación geométrica de las señales viene derivada de la estructura de espacio vectorial que poseen las señales y más concretamente de la definición de producto escalar en el espacio vectorial. Continuando con esta interpretación geométrica, la energía de un símbolo concreto es el cuadrado de la distancia del símbolo al origen de coordenadas, o lo que es lo mismo, el cuadrado de la norma del vector $E_i = \|\mathbf{a}_i\|^2$. También se puede medir el ángulo que forman los distintos vectores a través del producto escalar.

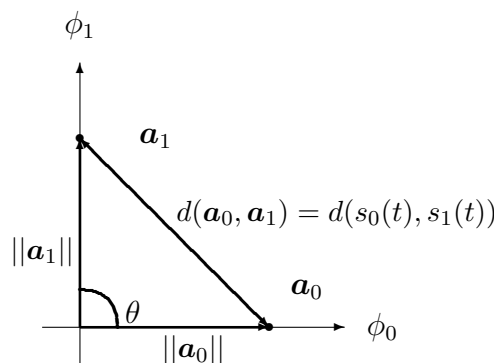


Figura 4.24: Interpretación geométrica de la constelación del cuarto conjunto de señales

La validez de todo lo planteado y discutido no se restringe a este conjunto de señales, sino que se puede aplicar al caso general de una constelación de M símbolos $\{\mathbf{a}_i, i = 0, \dots, M - 1\}$ y

una base ortonormal de N señales $\{\phi_j(t), j = 0, \dots, N - 1\}$ que generan el conjunto de señales $\{s_i(t), i = 0, \dots, M - 1\}$.

La constelación se diseñará de tal forma que se consiga una separación máxima entre los símbolos que la componen intentando al mismo tiempo que la energía media por símbolo, o energía de la constelación, esté dentro de los límites del sistema. La elección del conjunto de elementos de la base, de acuerdo con el modelo de canal que estamos considerando, es irrelevante en cuanto a prestaciones y energía. El conjunto de señales de la base se escoge para que sean adecuadas para el canal concreto sobre el que se realizará la transmisión. Y la elección de una u otra constelación no tendrá ninguna relevancia sobre la forma de las señales, en el dominio temporal y frecuencial, y por tanto en la adecuación de las señales al canal. Esto permite desacoplar el diseño de la constelación del diseño de la base ortonormal que define el espacio de señales (que se corresponden con el diseño del codificador y el diseño del modulador, respectivamente).

A continuación se analizará de forma individual cada uno de los 4 elementos funcionales del modelo de comunicación ilustrado en la Figura 4.14. Se empezará por el receptor para acabar con los módulos del transmisor.

4.4. Demodulador

El demodulador es el primer elemento funcional del receptor. Como se ilustra en la Figura 4.25, este elemento convierte la señal recibida $r(t)$ en una secuencia discreta $\mathbf{q}[n]$ de vectores de dimensión N .

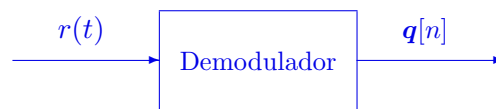


Figura 4.25: Demodulador de un sistema de comunicaciones digital.

La secuencia $\mathbf{q}[n]$ contiene la representación vectorial de la señal recibida en el espacio de señales del sistema, definido por la base ortonormal de dimensión N

$$\{\phi_0(t), \phi_1(t), \dots, \phi_{N-1}(t)\}.$$

Se trata pues de una secuencia de vectores de dimensión N

$$\mathbf{q}[n] = \begin{bmatrix} q_0[n] \\ q_1[n] \\ \vdots \\ q_{N-1}[n] \end{bmatrix}$$

En concreto, la señal recibida se procesa por intervalos de símbolo, y en el instante discreto n , la secuencia $\mathbf{q}[n]$ contiene la representación vectorial del fragmento de la señal recibida en el intervalo de símbolo asociado a ese instante, es decir, en $nT \leq t < (n + 1)T$

$$\mathbf{q}[n] \equiv \text{proyección de } r(t) \text{ en } nT \leq t < (n + 1)T \text{ sobre } \{\phi_j(t)\}_{j=0}^{N-1}$$

La proyección de una señal sobre un elemento de la base, $\phi_k(t)$, se obtiene mediante el producto escalar. Como las señales que forman la base tiene por definición su soporte en el intervalo $0 \leq$

$t < T$, hacer el producto escalar del fragmento de la señal $r(t)$ en el intervalo del símbolo de índice n , en $nT \leq t < (n+1)T$, es equivalente al producto escalar de $r(t)$ con la señal $\phi_k(t)$ retardada nT segundos, es decir, que la coordenada de índice k del vector $\mathbf{q}[n]$ se puede obtener como

$$q_k[n] = \langle r(t), \phi_k(t - nT) \rangle = \int_{-\infty}^{\infty} r(t) \cdot \phi_k^*(t - nT) dt = \int_{nT}^{(n+1)T} r(t) \cdot \phi_k^*(t) dt.$$

El operador conjugado es relevante para una representación compleja de señales. En esta asignatura sólo se considerarán señales reales, aunque por completitud en la notación se incluirá la posibilidad de trabajar con señales complejas². Por tanto, para la implementación de un demodulador se necesita algún tipo de estructura capaz de obtener N productos escalares. Aquí se van a estudiar dos estructuras equivalentes para llevar a cabo esta labor:

1. Basada en operaciones de correlación.
2. Basada en operaciones de filtrado.

Por simplicidad en la notación, en los desarrollos inicialmente se considerará la obtención del primer vector de la secuencia, $\mathbf{q}[0]$, que contiene la representación vectorial de la señal $r(t)$ en el primer intervalo de símbolo, en $0 \leq t < T$. Por simplicidad, en este caso se obviaré el índice temporal discreto $[0]$, de forma que se utilizará la notación

$$\mathbf{q} = \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_{N-1} \end{bmatrix} \equiv \mathbf{q}[0] = \begin{bmatrix} q_0[0] \\ q_1[0] \\ \vdots \\ q_{N-1}[0] \end{bmatrix}$$

La extensión para un instante discreto n arbitrario se hará de forma inmediata.

4.4.1. Demodulador por correlación

La idea del demodulador por correlación es utilizar la estructura que sale de la implementación directa de la operación producto escalar. Para el primer intervalo de símbolo

$$q_k = \langle r(t), \phi_k(t) \rangle = \int_0^T r(t) \phi_k^*(t) dt.$$

Por tanto, el demodulador tiene la estructura que se presenta en la Figura 4.26.

La notación vectorial que estamos utilizando puede dar la impresión de una estructura del demodulador excesivamente compleja para su implementación en hardware. Sin embargo, cuando particularizamos la estructura en sistemas prácticos, la estructura no suele ser en general muy compleja. Por ejemplo, para los cuatro conjuntos de señales del ejemplo de la Sección 4.3.1, la estructura es muy sencilla. Para los dos primeros conjuntos, en los que la función base toma un valor constante en todo el intervalo de símbolo, el demodulador se reduce a un escalado de la señal recibida y una integración en el intervalo de símbolo, como se muestra en la Figura 4.27 a); para el tercer conjunto, el demodulador multiplica la señal recibida por una señal sinusoidal (en hardware, un oscilador) e integra el producto en el intervalo de símbolo (Figura 4.27 b)); para el cuarto conjunto de señales se ha de multiplicar la señal recibida por dos componentes sinusoidales

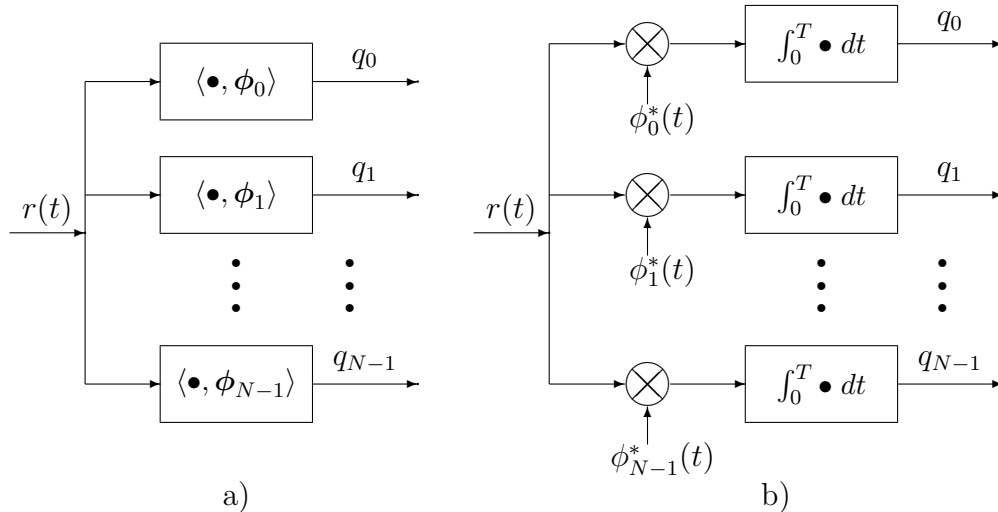


Figura 4.26: Estructura del demodulador por correlación; a) genérica, b) realización.

en cuadratura (un seno y un coseno, que en hardware se pueden obtener con un oscilador y un desfasador de 90°) y dos integradores (Figura 4.27 c)).

4.4.2. El filtro adaptado

Otra posibilidad para el diseño del receptor es la utilización de filtros. Esta opción permite conservar y observar la característica temporal y frecuencial de las señales, y en algunos casos ofrece una alternativa más eficiente para la implementación en hardware del demodulador, dependiendo de la forma de las funciones base que definen el espacio de señales del sistema.

Si se filtra la señal recibida a la entrada del receptor, $r(t)$, con un cierto filtro con respuesta al impulso $h_k(t)$, su salida es

$$y_k(t) = r(t) * h_k(t) = \int_{-\infty}^{\infty} r(\tau) \cdot h_k(t - \tau) d\tau$$

Vamos a ver que mediante un proceso de filtrado es posible implementar la operación producto escalar que requiere el receptor. Recuerde que se busca una estructura para obtener

$$q_k = \langle r(t), \phi_k(t) \rangle = \int_0^T r(t) \phi_k^*(t) dt.$$

Si se comparan las dos expresiones anteriores, para $y_k(t)$ y q_k , respectivamente, se observa que en la segunda aparece $\phi_k^*(t)$ en lugar de $h_k(t)$, un complejo conjugado, que en la primera se integra sobre τ mientras en la segunda se integra sobre t , y que el signo sobre la variable de integración es negativo en la respuesta del filtro. Si se elige $h_k(t) = \phi_k^*(-t)$, la expresión de $y_k(t)$ se convierte en

$$y_k(t) = \int_{-\infty}^{\infty} r(\tau) \cdot \phi_k^*(-(t - \tau)) d\tau$$

Si se compara con la expresión analítica para q_k , y tenemos en cuenta que $-(t - \tau)$ se puede escribir como $(\tau - t)$ vemos que las dos expresiones son iguales si se toma el valor $t = 0$. Por tanto

²No existen señales electromagnéticas complejas, pero en algunos casos se utiliza la notación compleja para el manejo simultáneo de dos señales, una contenida en la parte real y otra en la parte imaginaria de la señal analítica.

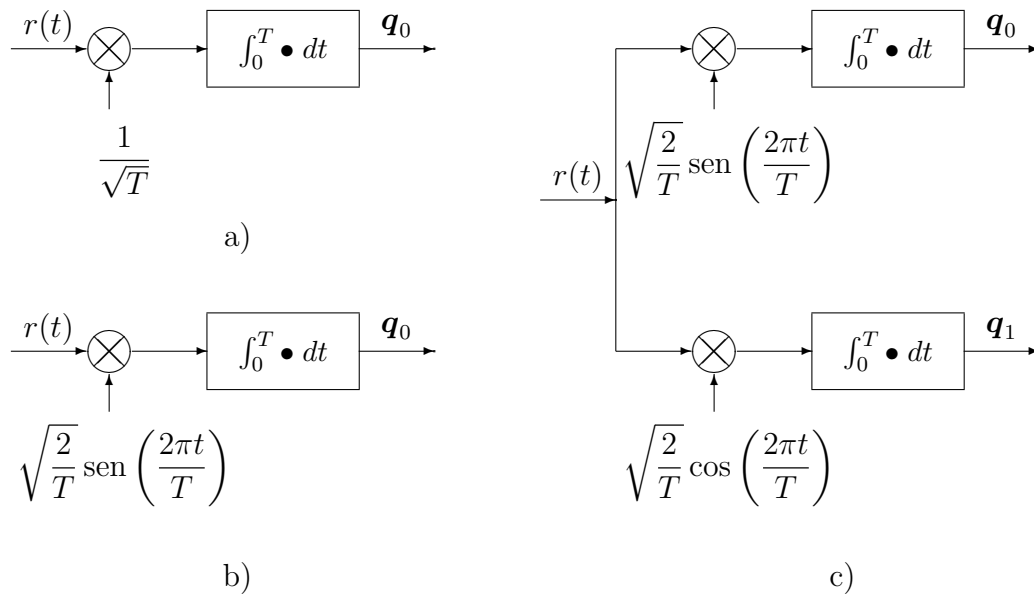


Figura 4.27: Estructura del demodulador para los cuatro conjuntos de señales analizados; a) dos primeros conjuntos, b) tercer conjunto, c) cuarto conjunto.

si se muestra la señal de salida del filtro $y_k(t)$ en $t = 0$, tenemos

$$y_k(0) = \int_{-\infty}^{\infty} r(\tau) \cdot h_k(-\tau) d\tau = \int_{-\infty}^{\infty} r(\tau) \cdot \phi_k^*(\tau) d\tau = q_k,$$

es decir, el producto escalar de la señal recibida con $\phi_k(t)$.

En general, un filtro con respuesta al impulso $h(t) = x^*(-t)$ se dice que es un *filtro adaptado* a la señal $x(t)$. Por lo que el demodulador se puede implementar utilizando un banco de N filtros adaptados a los elementos de la base $\phi_k(t)$, con respuesta $h_k(t) = \phi_k^*(-t)$, y muestreando en $t = 0$ (al principio del intervalo de símbolo), tal y como se muestra en la Figura 4.28

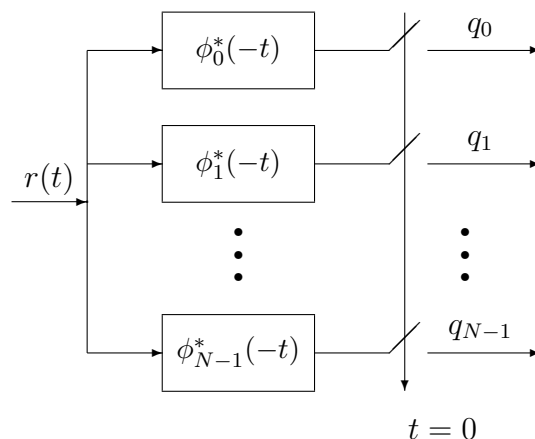


Figura 4.28: Estructura del demodulador basada en filtros adaptados.

Para la implementación práctica de esta estructura surge el problema de que los filtros son anticausales: como $\phi_k(t)$ tiene por definición como soporte $0 \leq t < T$, $\phi_i^*(-t)$ tiene como soporte $-T < t \leq 0$. Para poder tener una implementación causal, se puede retardar la respuesta $\phi^*(-t)$ T segundos, con lo que la salida del filtro se retarda la misma cantidad y es necesario muestrear

en $t = T$ en lugar de en cero, tal y como se muestra en la Figura 4.29. De este modo se tiene lo mismo que con el filtro anticausal muestreando en cero.

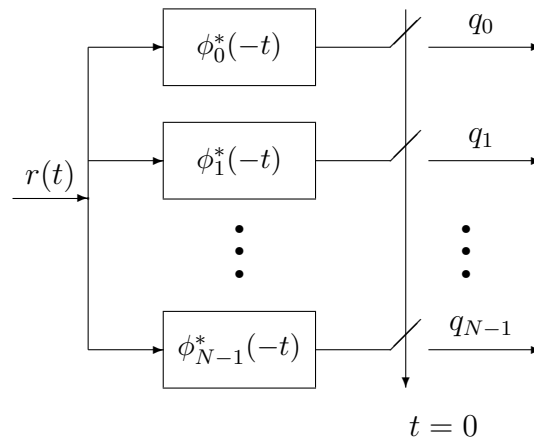


Figura 4.29: Estructura del demodulador basada en filtros adaptados causales.

Analíticamente, desplazar $\phi_k^*(-t)$ un tiempo T significa tener $\phi_k^*(-(t - T)) = \phi_k^*(T - t)$. La salida de ese filtro desplazado T , denotada $y_k^T(t)$ es entonces

$$y_k^T(t) = \int_{-\infty}^{\infty} r(\tau) \times \phi_k^*(T - (t - \tau)) d\tau.$$

Muestreando en $t = T$ se tiene

$$y_k^T(T) = \int_{-\infty}^{\infty} r(\tau) \cdot \phi_k^*(\tau) d\tau = y_k(0) = q_k.$$

Se observa por tanto que esta estructura proporciona exactamente el mismo resultado que la que utiliza filtros anticausales. Para la implementación, evidentemente hay que utilizar la estructura con filtros causales, pero como ambas son analíticamente equivalentes, para hacer cálculos se puede utilizar la expresión con filtros anticausales que es más sencilla.

Propiedades del filtro adaptado

El filtro adaptado no aporta por sí mismo ninguna novedad más que una estructura alternativa para la implementación de las operaciones de producto escalar que debe llevar a cabo el demodulador. Sin embargo, esta estructura permite comprobar de una forma sencilla algunas de las propiedades del demodulador.

Máxima relación señal a ruido Esta es seguramente la propiedad más importante. Para demostrar esta propiedad se propone el caso genérico en el que se tiene una señal conocida $s(t)$ a la que se suma ruido y que se filtra con un filtro con respuesta al impulso $h(t)$. El modelo para el término de ruido es el modelo estadístico habitual para el ruido térmico: proceso aleatorio estacionario, ergódico, blanco y gaussiano, con media nula y densidad espectral de potencia y función de autocorrelación

$$S_n(j\omega) = \frac{N_0}{2}, R_n(\tau) = \frac{N_0}{2} \delta(\tau).$$

Este esquema general se muestra en la Figura 4.30.

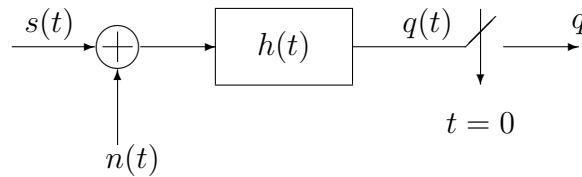


Figura 4.30: Modelo para la derivación de la propiedad de máxima relación señal a ruido del filtro adaptado.

Se va a comprobar en primer lugar que el filtro que proporciona la máxima relación señal a ruido a su salida, definida como

$$\left(\frac{S}{N}\right)_q \equiv \frac{\text{Energía en } q \text{ debida a } s(t)}{\text{Energía en } q \text{ debida a } n(t)},$$

es un filtro adaptado a la señal transmitida $s(t)$. Para simplificar la notación se considera el caso de señales reales, aunque los resultados se extienden trivialmente al caso de señales complejas. La salida del filtro, $q(t)$, se define como

$$\begin{aligned} q(t) &= (s(t) + n(t)) * h(t) = s(t) * h(t) + n(t) * h(t) \\ &= \int_{-\infty}^{\infty} s(\tau) \cdot h(t - \tau) d\tau + \int_{-\infty}^{\infty} n(\tau) \cdot h(t - \tau) d\tau. \end{aligned}$$

Su valor en $t = 0$ es

$$q = q(0) = \int_{-\infty}^{\infty} s(\tau) \cdot h(-\tau) d\tau + \int_{-\infty}^{\infty} n(\tau) \cdot h(-\tau) d\tau = s + n.$$

La relación señal a ruido de la salida se define como

$$\left(\frac{S}{N}\right)_q = \frac{E[|s|^2]}{E[|n|^2]} = \frac{|s|^2}{E[n^2]},$$

ya que el término s es determinista si la señal $s(t)$ es conocida

$$E[|s|^2] = |s|^2 = \left| \int_{-\infty}^{\infty} s(\tau) \cdot h(-\tau) d\tau \right|^2 \quad (\text{valor determinista})$$

El valor de $E[|n|^2]$ se obtiene como

$$\begin{aligned} E[|n|^2] &= E \left[\left(\int_{-\infty}^{+\infty} n(\tau) \cdot h(-\tau) d\tau \right) \left(\int_{-\infty}^{+\infty} n(\theta) \cdot h(-\theta) d\theta \right) \right] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \underbrace{E[n(\tau) \cdot n(\theta)]}_{R_n(\tau-\theta)} \cdot h(-\tau) \cdot h(-\theta) d\tau d\theta \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \underbrace{\frac{N_0}{2} \delta(\tau - \theta)}_{R_n(\tau-\theta)} \cdot h(-\tau) \cdot h(-\theta) d\tau d\theta \\ &= \frac{N_0}{2} \int_{-\infty}^{+\infty} |h(-\tau)|^2 d\tau = \frac{N_0}{2} \int_{-\infty}^{+\infty} |h(\tau)|^2 d\tau = \frac{N_0}{2} \cdot \mathcal{E}\{h(t)\} \end{aligned}$$

En el desarrollo se ha aplicado la propiedad de la integral del producto de una delta y una función

$$\int_{-\infty}^{+\infty} f(x) \cdot \delta(x - x_0) dx = f(x_0).$$

Se puede ver que la energía de la componente de ruido únicamente depende de la energía de la respuesta en frecuencia del filtro. Sustituyendo este resultado en la expresión de la relación señal a ruido se tiene

$$\left(\frac{S}{N}\right)_q = \frac{|s|^2}{E[|n|^2]} = \frac{\left|\int_{-\infty}^{\infty} s(\tau) \cdot h(-\tau) d\tau\right|^2}{\frac{N_0}{2} \mathcal{E}\{h(t)\}}.$$

Para buscar el máximo de esta relación señal a ruido respecto a la respuesta al impulso del filtro, $h(t)$, se va a hacer el siguiente razonamiento: se supone que la energía de $h(t)$ es constante, y se busca el máximo de la relación $(S/N)_q$ respecto a $h(t)$; si ese máximo no depende de $\mathcal{E}\{h(t)\}$ tendremos el máximo para cualquier valor de $\mathcal{E}\{h(t)\}$ y, por consiguiente, el máximo que estamos buscando.

Si se supone $\mathcal{E}\{h(t)\}$ constante, el máximo de la relación señal a ruido se limita a calcular el máximo del numerador. Si se particulariza la desigualdad de Cauchy-Schwarz (ver Sección 4.2.2) para las señales $s(t)$ y $h(-t)$, y elevando al cuadrado cada uno de los términos de la desigualdad se tiene

$$\left|\int_{-\infty}^{\infty} s(\tau) \cdot h(-\tau) d\tau\right|^2 \leq \left(\int_{-\infty}^{\infty} |s(\tau)|^2 d\tau\right) \left(\int_{-\infty}^{\infty} |h(-\tau)|^2 d\tau\right).$$

La igualdad, y por tanto el máximo, se produce cuando $h(-t) = \alpha \times s(t)$ para algún valor de la constante α . Introduciendo este resultado se obtiene

$$\begin{aligned} \max_{h(t)} \left(\frac{S}{N}\right)_q &= \frac{\left(\int_{-\infty}^{\infty} s(\tau)h(-\tau)d\tau\right)^2}{\frac{N_0}{2} \mathcal{E}\{h(t)\}} \Bigg|_{h(-t)=\alpha s(t)} \\ &= \frac{\left(\int_{-\infty}^{\infty} |s(\tau)|^2 d\tau\right) \left(\alpha^2 \int_{-\infty}^{\infty} |s(\tau)|^2 d\tau\right)}{\frac{N_0}{2} \alpha^2 \mathcal{E}\{s(t)\}} \\ &= \frac{2}{N_0} \mathcal{E}\{s(t)\}. \end{aligned}$$

De este resultado se obtienen dos conclusiones:

1. La relación señal a ruido a la salida se hace máxima cuando $h(t) = \alpha s(-t)$ para cualquier valor de α (exceptuando $\alpha = 0$) y, particularmente, para el filtro adaptado

$$h(t) = s(-t),$$

Para señales complejas se llega de forma trivial a la misma conclusión para

$$h(t) = s^*(-t).$$

2. La relación señal a ruido a la salida del filtro adaptado no depende de la forma específica de $s(t)$, sino únicamente de su energía y de la densidad espectral de potencia de ruido a la entrada del filtro.

Esta demostración realizada para una señal genérica se puede aplicar al conjunto de señales del transmisor, $\{s_i(t), i = 0, \dots, M - 1\}$ y al conjunto de señales de la base, $\{\phi_i(t), i = 0, \dots, M - 1\}$, con lo que se demuestra que la estructura de demodulador basado en filtros adaptados (o su equivalente de correlación) es la estructura que nos permite obtener la máxima relación señal a ruido en cada componente del vector \mathbf{q} .

4.4.3. Extensión de las estructuras del demodulador para cualquier instante n

Se ha estudiado en las secciones anteriores el demodulador para la obtención de la primera de las observaciones de la secuencia de vectores, $\mathbf{q} \equiv \mathbf{q}[0]$. La extensión para la obtención del vector de la secuencia en cualquier instante arbitrario, $\mathbf{q}[n]$, es trivial, ya que sólo hay que procesar $r(t)$ en el intervalo de símbolo correspondiente, $nT \leq t < (n+1)T$. En el caso de la estructura con correladores, sólo hay que retardar las funciones $\phi_k(t)$ nT segundos e integrar en el intervalo de símbolo, y en el caso de los filtros adaptados sólo hay que modificar el instante de muestreo, para que coincida con el inicio del intervalo de símbolo, $t = nT$, si se utilizan filtros anticausales, o con el final del intervalo de símbolo, $(n+1)T$, si se utilizan filtros causales, como se muestra en la Figura 4.31.

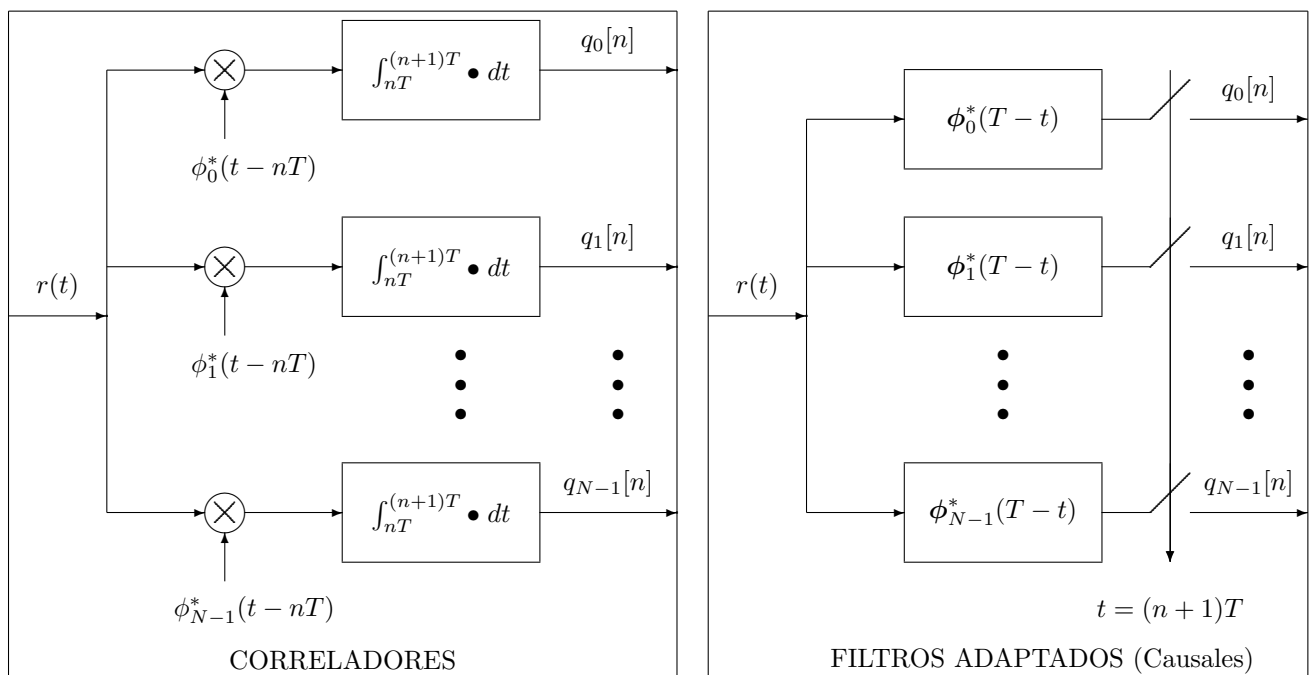


Figura 4.31: Estructura del demodulador para un flujo indefinido de símbolos.

Es interesante remarcar que ambas estructuras proporcionan el mismo resultado, ya que se trata de dos estructuras que realizan la misma operación analítica: el producto escalar de la señal recibida con cada una de las N funciones que forman la base ortonormal que define el espacio de señales del sistema.

4.4.4. Caracterización estadística de la salida del demodulador en el caso de transmisión sobre un canal gaussiano

Una vez que se han visto las estructuras que pueden utilizarse para la implementación del demodulador, a continuación se realizará la caracterización estadística de la secuencia de observaciones $\mathbf{q}[n]$. En particular se considerará el caso de transmisión de la señal sobre un canal gaussiano, cuando la señal transmitida no sufre ninguna distorsión lineal (lo que implica suponer que la base ortonormal se adapta perfectamente a las características del canal), y que la única distorsión se

debe al efecto del ruido, con lo que la señal recibida será

$$r(t) = s(t) + n(t).$$

De nuevo, para simplificar la notación se considerará el primer instante discreto, $n = 0$, y se obviará la dependencia con el índice temporal discreto n . El objetivo es obtener el modelo estadístico de la observación \mathbf{q} cuando se conoce el símbolo que se ha transmitido, que con carácter general se supondrá que es el símbolo de índice i , i.e. $\mathbf{A} \equiv \mathbf{A}[0] = \mathbf{a}_i$. Esto significa que en el intervalo de símbolo de interés, el primero, $0 \leq t < T$, la señal transmitida es

$$s(t) = s_i(t) = \sum_{j=0}^{N-1} a_{i,j} \times \phi_j(t).$$

A continuación se analiza la salida del demodulador bajo esta situación. Introduciendo la expresión analítica de $r(t)$ en la obtención de la coordenada de índice k , q_k , se tiene

$$\begin{aligned} q_k &= \langle r(t), \phi_k(t) \rangle = \int_0^T r(t) \cdot \phi_k^*(t) dt = \int_0^T (s(t) + n(t)) \cdot \phi_k^*(t) dt \\ &= \int_0^T \left(\sum_{j=0}^{N-1} a_{i,j} \cdot \phi_j(t) \right) \phi_k^*(t) dt + \underbrace{\int_0^T n(t) \cdot \phi_k^*(t) dt}_{z_k} \\ &= \sum_{j=0}^{N-1} a_{i,j} \int_0^T \phi_j(t) \cdot \phi_k^*(t) dt + z_k = \sum_{j=0}^{N-1} a_{i,j} \cdot \delta[j - k] + z_k = a_{i,k} + z_k. \end{aligned}$$

Se ha tenido en cuenta en el desarrollo que los elementos de la base son ortonormales, lo que analíticamente implica que

$$\langle \phi_j(t), \phi_k(t) \rangle = \int_0^T \phi_j(t) \cdot \phi_k^*(t) dt = \begin{cases} 0, & \text{si } k \neq j \\ 1, & \text{si } k = j \end{cases} \equiv \delta[j - k].$$

En este caso $a_{i,k}$ es la coordenada de índice k del símbolo \mathbf{a}_i y z_k es la contribución del ruido gaussiano en esa coordenada. Todas las componentes (coordenadas) de \mathbf{q} pueden expresarse de forma conjunta empleando notación vectorial como

$$\mathbf{q} = \begin{bmatrix} a_{i,0} \\ a_{i,1} \\ \vdots \\ a_{i,N-1} \end{bmatrix} + \begin{bmatrix} z_0 \\ z_1 \\ \vdots \\ z_{N-1} \end{bmatrix} = \mathbf{a}_i + \mathbf{z}$$

Los valores del vector \mathbf{a}_i son deterministas (se parte de la asunción de que el símbolo transmitido es conocido), por lo que sólo queda caracterizar el vector de ruido \mathbf{z} . Cada una de las componentes del vector de ruido se considera como una variable aleatoria. Dado que el proceso de que proviene, $n(t)$, es gaussiano, cada componente tendrá una función de densidad de probabilidad gaussiana. El vector con las coordenadas de ruido está por tanto compuesto por N variables aleatorias conjuntamente gaussianas. Para caracterizarlas, hay que encontrar su función densidad de probabilidad conjunta, lo que en el caso de v.a. gaussianas se reduce al cálculo de su vector de medias y su matriz de covarianzas. Comenzaremos calculando la media de cada una de las coordenadas del vector de ruido. Aplicando la definición de media de una variable aleatoria, la media de la coordenada de índice k es

$$m_{z_k} = E[z_k] = E \left[\int_0^T n(t) \phi_j^*(t) dt \right] = \int_0^T E[n(t)] \phi_j^*(t) dt = 0,$$

ya que $n(t)$ tiene media nula. A continuación se calcula la matriz de covarianzas. Teniendo en cuenta que todas las componentes tienen media nula, la covarianza entre dos de las coordenadas es

$$\begin{aligned} \text{Cov}(z_j, z_k) &= E[z_j \cdot z_k^*] = E \left[\left(\int_0^T n(t) \cdot \phi_j^*(t) dt \right) \left(\int_0^T n^*(\tau) \cdot \phi_k(\tau) d\tau \right) \right] \\ &= \int_0^T \int_0^T \underbrace{E[n(t) \cdot n^*(\tau)]}_{R_n(t-\tau) = \frac{N_0}{2} \cdot \delta(t-\tau)} \phi_j^*(t) \cdot \phi_k(\tau) dt d\tau \\ &= \int_0^T \int_0^T \frac{N_0}{2} \cdot \delta(t - \tau) \cdot \phi_j^*(t) \cdot \phi_k(\tau) dt d\tau \\ &= \frac{N_0}{2} \int_0^T \phi_j^*(t) \cdot \phi_k(t) dt = \frac{N_0}{2} \cdot \delta[j - k]. \end{aligned}$$

Por tanto, la covarianza entre dos coordenadas distintas es nula, y la covarianza de una coordenada consigo misma (varianza) es

$$\sigma_{z_k}^2 = \frac{N_0}{2}.$$

Esto significa que las variables aleatorias que modelan cada uno de los elementos del vector de ruido están incorreladas, lo que para variables aleatorias gaussianas es equivalente a decir que son independientes. Por tanto, la matriz de covarianzas es una matriz diagonal con la varianza de cada componente en la diagonal principal. A la luz de estos resultados se puede concluir que las N componentes de ruido son variables aleatorias gaussianas incorreladas (independientes) de media nula y varianza $N_0/2$. Como una variable aleatoria gaussiana queda unívocamente determinada a partir de su media y su varianza, la función densidad de probabilidad de cada componente z_k es

$$f_{z_k}(z_k) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{z_k^2}{N_0}},$$

y como bajo estadística gaussiana incorrelación implica independencia, la función densidad de probabilidad conjunta de \mathbf{z} se obtiene mediante el producto de las distribuciones de cada una de las componentes del vector

$$f_{\mathbf{z}}(\mathbf{z}) = \prod_{k=0}^{N-1} f_{z_k}(z_k) = \frac{1}{(\pi N_0)^{N/2}} e^{-\sum_{k=0}^{N-1} \frac{z_k^2}{N_0}} = \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{z}\|^2}{N_0}}.$$

Una vez obtenida la función de densidad de probabilidad del vector de ruido, es fácil obtener la función de densidad de probabilidad condicional de la observación \mathbf{q} . La función densidad de probabilidad de \mathbf{q} cuando se conoce el símbolo transmitido, $\mathbf{A} = \mathbf{a}_i$, es una distribución condicional: distribución de q_k condicionada a que el símbolo transmitido es $\mathbf{A} = \mathbf{a}_i$. Para cada una de las componentes del vector \mathbf{q} , q_k , se tiene también una distribución condicional: distribución de q_k condicionada a que el símbolo transmitido es $\mathbf{A} = \mathbf{a}_i$. En ese caso, $q_k = a_{i,k} + z_k$, es una variable aleatoria formada por la suma de una constante determinista $a_{i,k}$ y una variable aleatoria, z_k , por lo que tiene el mismo tipo de distribución que z_k pero con la media modificada sumando el valor constante, $a_{i,k}$, a la media de z_k , que en este caso era cero. Por tanto, la distribución es una distribución gaussiana, de media la coordenada de índice k del símbolo transmitido, $a_{i,k}$, y varianza $N_0/2$

$$f_{q_k|\mathbf{A}}(q_k|\mathbf{a}_i) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(q_k - a_{i,k})^2}{N_0}}.$$

Debido a la independencia, la función densidad de probabilidad condicional conjunta del vector \mathbf{q} cuando se conoce que $\mathbf{A} = \mathbf{a}_i$ se obtiene mediante el producto de las distribuciones condicionales de cada una de las componentes del vector

$$\begin{aligned} f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i) &= \prod_{k=0}^{N-1} f_{q_k|A_k}(q_k|a_{i,k}) = \frac{1}{(\pi N_0)^{N/2}} e^{-\sum_{k=0}^{N-1} \frac{(q_k - a_{i,k})^2}{N_0}} \\ &= \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{q} - \mathbf{a}_i\|^2}{N_0}}. \end{aligned}$$

Como se puede ver, la distribución del vector de observación cuando se transmite un cierto símbolo es una distribución gaussiana N dimensional, de componentes independientes, media el símbolo transmitido y varianza $N_0/2$ en cada una de las N direcciones del espacio

$$f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i) = \mathcal{N}^N \left(\mathbf{a}_i, \frac{N_0}{2} \right).$$

En el desarrollo del demodulador hemos partido de la idea de recuperar las coordenadas de la señal recibida en la base del espacio de señales del transmisor, pero no nos hemos preguntado si la información que recuperamos de la señal recibida es toda la información relevante a la hora de tomar la decisión del símbolo que fue transmitido o existe información adicional que pueda ayudar a tomar la decisión. La respuesta a esta pregunta es que tal y como se ha definido el vector \mathbf{q} , contiene toda la información relevante para decidir qué símbolo fue el transmitido. Estrictamente, se dice que el vector \mathbf{q} es un *estadístico suficiente para la detección*. La demostración se puede encontrar por ejemplo en [Artés-Rodríguez et al., 2007].

4.4.5. Canal discreto equivalente

Algunas de las conclusiones obtenidas del análisis del modelo de comunicación digital son las siguientes:

1. La fiabilidad de un esquema de comunicación, por lo que respecta al proceso de modulación, viene dada por las características de la constelación de señales y no por los elementos de la base de señales que empleemos en el proceso de modulación.
2. La relación señal a ruido no depende de la forma concreta que toman cada uno de los elementos de la base de señales, sino de la energía de las señales y de la densidad espectral de potencia del ruido (la potencia del ruido).
3. La salida del demodulador es un vector \mathbf{q} que toma la forma $\mathbf{q} = \mathbf{A} + \mathbf{z}$, donde \mathbf{A} es el símbolo transmitido y \mathbf{z} es una componente de ruido con una función densidad de probabilidad conjuntamente gaussiana.

Estas conclusiones permiten obtener una simplificación del modelo general de un sistema de comunicación, englobando modulador, canal y demodulador en un único elemento que denominamos *canal discreto equivalente*. Este modelo, que se representa en la Figura 4.32, es de utilidad en el análisis del decisor porque nos permite “ocultar” la naturaleza analógica del canal y centrarnos en los aspectos que van a influir en la fiabilidad de la comunicación.

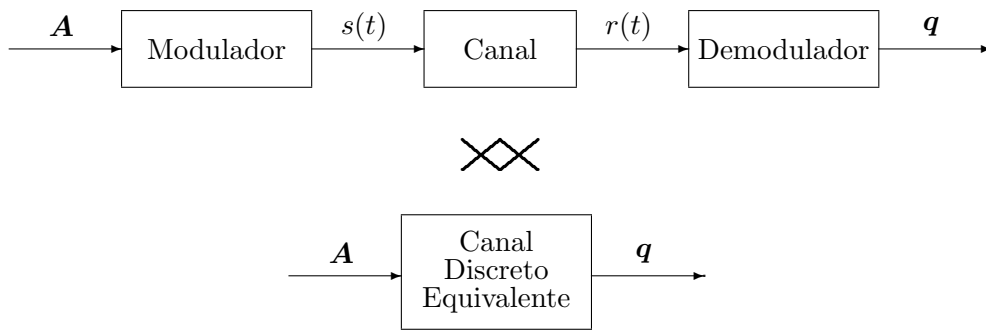


Figura 4.32: Definición de canal discreto equivalente.

El canal discreto equivalente es en general vectorial, con la dimensión dada por el espacio de señales del sistema, N . La relación entrada salida, para un modelo de transmisión sobre canal gaussiano, viene dada por

$$\mathbf{q} = \mathbf{A} + \mathbf{z},$$

tal y como se representa en la Figura 4.33

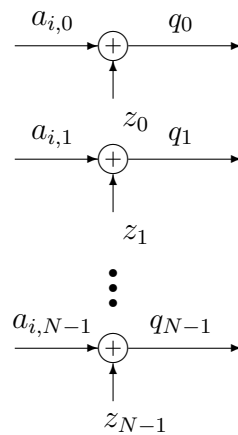


Figura 4.33: Relaciones entre representaciones vectoriales de transmisor y receptor dadas por el canal discreto equivalente, cuando se transmite el símbolo $\mathbf{A} = \mathbf{a}_i$ sobre un canal gaussiano.

4.5. Decisor

El decisor es el elemento final de un demodulador digital, y su función es proporcionar la estima del símbolo que se ha transmitido en un instante n , $\hat{B}[n]$, a partir de la representación vectorial de la señal recibida en el intervalo de símbolo asociado a dicho instante, $\mathbf{q}[n]$, tal y como se ilustra en la Figura 4.34. Con esta estima, de forma implícita se tiene también la estima de los bits transmitidos, $\hat{B}_b[\ell]$, debido a la identificación de cada posible valor del símbolo, b_i , con una m -tupla de bits específica.



Figura 4.34: Decisor de un sistema de comunicaciones.

Dada su función, es un elemento determinante para las prestaciones del sistema, parametrizadas por la probabilidad de error, definida a nivel de símbolo o a nivel de bit. Por ello, el decisor se diseña de acuerdo con el siguiente criterio: hacer mínima la probabilidad de error de símbolo.

4.5.1. Diseño del decisor - Regiones de decisión

Debido a que $\mathbf{q}[n]$ es un estadístico suficiente para la detección de $\mathbf{A}[n]$ (y por tanto de $B[n]$), la decisión se realizará símbolo a símbolo sin memoria; es decir, para cada instante temporal n , se decidirá $\hat{B}[n]$ a partir de $\mathbf{q}[n]$, aplicando siempre la misma regla, ya que la estadística de $\mathbf{q}[n]$ no varía con el tiempo.

Teniendo en cuenta que el alfabeto de símbolos tiene M posibles valores

$$B[n] \in \{b_0, b_1, \dots, b_{M-1}\},$$

el diseño del decisor consiste en establecer para cada posible valor de $\mathbf{q}[n]$, qué valor del alfabeto, de los M posibles, se decide. O visto de otra forma, establecer cuáles son los posibles valores de $\mathbf{q}[n]$ que darán lugar a la decisión de cada uno de los M posibles valores de $B[n]$. La forma de establecer esos valores es definiendo lo que se denominan “*regiones de decisión*”. Se dividirá el dominio de $\mathbf{q}[n]$ en M regiones disjuntas que formen una partición del espacio de $\mathbf{q}[n]$

$$\{I_0, I_1, \dots, I_{M-1}\}.$$

Cada región se asocia a uno de los valores del alfabeto de símbolos, identificándose la asociación a través del subíndice (I_k se asocia a b_k). Una vez definidas las regiones, a la vista de un valor de observación $\mathbf{q}[n]$, se decidirá $\hat{B}[n] = b_k$ cuando el valor de observación esté en la región de decisión de b_k , i.e. cuando $\mathbf{q}[n] \in I_k$. Por eso se denominan regiones de decisión.

Por tanto, el diseño del decisor se plantea como el problema de establecer las M regiones de decisión que minimicen la probabilidad de error de símbolo.

4.5.2. Obtención del decisor óptimo

A continuación se obtendrán las reglas a seguir para establecer de forma óptima (para minimizar la probabilidad de error de símbolo) las regiones de decisión que definen el decisor. De nuevo, para simplificar la notación se tratará analíticamente el caso del primer símbolo de la secuencia (índice discreto $n = 0$) y se obviará la dependencia con el índice temporal, para lo que se utilizará como notación

$$\hat{B} \equiv \hat{B}[n], \quad \mathbf{q} \equiv \mathbf{q}[n].$$

Cuando \mathbf{q} toma un valor determinado \mathbf{q}_0 , el decisor toma una decisión sobre el símbolo transmitido, por ejemplo $\hat{B} = b_i$; la probabilidad de error en esa decisión, que denotamos como $P_e^{(\mathbf{q}=\mathbf{q}_0 \rightarrow \hat{B}=b_i)}$ es

$$P_e^{(\mathbf{q}=\mathbf{q}_0 \rightarrow \hat{B}=b_i)} = P(B \neq b_i | \mathbf{q} = \mathbf{q}_0) = 1 - P(B = b_i | \mathbf{q} = \mathbf{q}_0) = 1 - p_{B|\mathbf{q}}(b_i | \mathbf{q}_0).$$

La probabilidad de error en una decisión ($\hat{B} = b_i$) tomada a partir de un valor de observación determinado ($\mathbf{q} = \mathbf{q}_0$) será igual a la probabilidad de que, siendo la entrada igual a \mathbf{q}_0 el símbolo transmitido no sea igual al decidido. El valor $p_{B|\mathbf{q}}(b_i | \mathbf{q}_0)$ es, obviamente, la probabilidad de acertar cuando ante un valor de entrada \mathbf{q}_0 el decisor asigna el símbolo b_i ($\hat{B} = b_i$). Esta es una probabilidad condicional, que se denomina *probabilidad a posteriori* del símbolo b_i .

Si el decisor tomara siempre la misma decisión, $\hat{B} = b_i$, independientemente del valor de observación \mathbf{q} , la probabilidad media de error, que se denotará como $P_e^{(\hat{B}=b_i, \forall \mathbf{q})}$, se obtiene calculando el promedio sobre el conjunto de todos los posibles valores de la observación \mathbf{q} , que estadísticamente es la esperanza matemática bajo la distribución de \mathbf{q}

$$P_e^{(\hat{B}=b_i, \forall \mathbf{q})} = E_{f_{\mathbf{q}}(\mathbf{q}_0)} \left[P_e^{(\mathbf{q}=\mathbf{q}_0 \rightarrow \hat{B}=b_i)} \right] = \int_{-\infty}^{\infty} [1 - p_{B|\mathbf{q}}(b_i|\mathbf{q}_0)] \cdot f_{\mathbf{q}}(\mathbf{q}_0) d\mathbf{q}_0 \quad (4.1)$$

$$= \int_{-\infty}^{\infty} f_{\mathbf{q}}(\mathbf{q}_0) d\mathbf{q}_0 - \int_{-\infty}^{\infty} p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) \cdot f_{\mathbf{q}}(\mathbf{q}_0) d\mathbf{q}_0 \quad (4.2)$$

$$= 1 - \int_{-\infty}^{\infty} p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) \cdot f_{\mathbf{q}}(\mathbf{q}_0) d\mathbf{q}_0 \quad (4.3)$$

Cuando el decisor toma decisiones distintas cuando la observación \mathbf{q} toma valores en cada una de las M regiones de decisión, la probabilidad de error del sistema, teniendo en cuenta que se trata de regiones disjuntas que forman una partición del dominio de \mathbf{q} , es

$$P_e = 1 - \sum_{i=0}^{M-1} \int_{I_i} p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) \cdot f_{\mathbf{q}}(\mathbf{q}_0) d\mathbf{q}_0$$

Para encontrar el mínimo de esta expresión hay que tener en cuenta lo siguiente:

- El mínimo se obtiene cuando se maximiza el segundo término.
- La función dentro de las integrales de este término, $p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) f_{\mathbf{q}}(\mathbf{q}_0)$, es siempre mayor o igual que cero por ser el producto de funciones no negativas ($0 \leq p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) \leq 1$, $0 \leq f_{\mathbf{q}}(\mathbf{q}_0) \leq 1$). Esto implica que la probabilidad de error se minimiza cuando se maximiza este argumento.
- $f_{\mathbf{q}}(\mathbf{q}_0)$ es independiente tanto del símbolo transmitido como del recibido. Por tanto, el máximo del sumatorio se obtiene cuando se maximiza el valor de $p_{B|\mathbf{q}}(b_i|\mathbf{q}_0)$ en cada una de las regiones de decisión.
- Por consiguiente, la región de decisión I_i , para la cual se asigna como salida del decisor el valor $\hat{B} = b_i$, es aquella para la que se cumple que

$$p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) > p_{B|\mathbf{q}}(b_j|\mathbf{q}_0), \quad \forall j \neq i.$$

Es decir, que ante un valor de entrada $\mathbf{q} = \mathbf{q}_0$, el decisor ha de calcular el conjunto de probabilidades $\{p_{B|\mathbf{q}}(b_j|\mathbf{q}_0), j = 0, \dots, M-1\}$ y la salida del decisor será el símbolo b_i que cumple la relación anterior. En el caso en que dos símbolos distintos b_i, b_k , obtengan el máximo valor, es decir

$$p_{B|\mathbf{q}}(b_i|\mathbf{q}_0) = p_{B|\mathbf{q}}(b_k|\mathbf{q}_0) > p_{B|\mathbf{q}}(b_j|\mathbf{q}_0), \quad \forall j \neq i, k,$$

el decisor podrá optar por cualquiera de ellos de forma arbitraria sin que esta decisión tenga efecto alguno sobre la probabilidad de error.

Este criterio se conoce con el nombre de *máximo a posteriori* o MAP (de “*maximum a posteriori*”). Su nombre viene de la denominación de $p_{B|\mathbf{q}}(b_i|\mathbf{q}_0)$ como probabilidades *a posteriori* de B dado \mathbf{q} , ya que representa la probabilidad de los símbolos una vez realizada la transmisión, en lugar de las probabilidades *a priori* $p_B(b_i)$. La Figura 4.35 muestra un ejemplo de como podrían

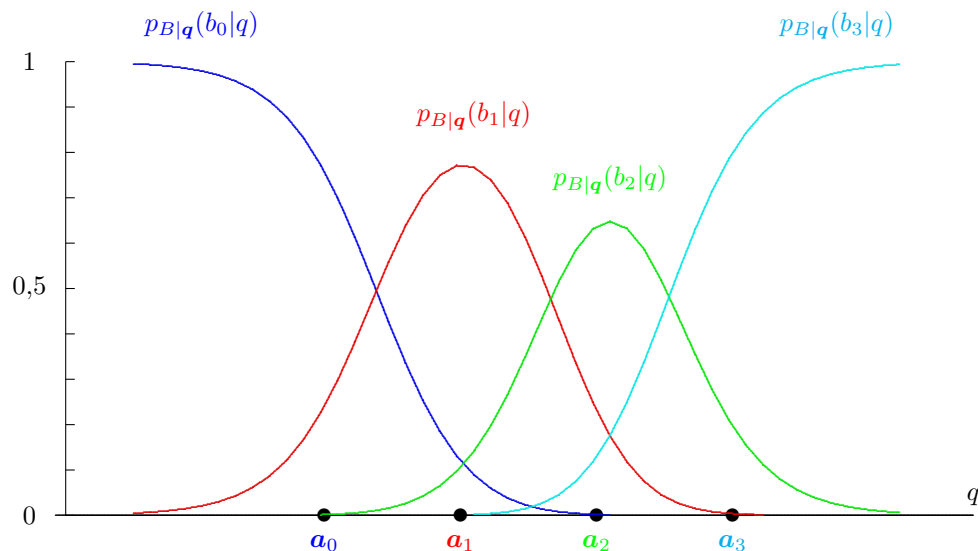


Figura 4.35: Un ejemplo de las probabilidades a posteriori para una constelación de 4 símbolos en un espacio unidimensional.

ser esas probabilidades a posteriori para una constelación de 4 símbolos en un espacio unidimensional. Se puede comprobar que para cada posible valor de observación q , la suma de las cuatro probabilidades a posteriori es uno.

Para obtener las regiones de decisión, hay que buscar cuál de los 4 símbolos tiene asociado el valor máximo de probabilidad a posteriori para cada valor de q . El máximo valor de probabilidad a posteriori, para cada posible valor de observación, se realiza en negro en la Figura 4.36. La región de decisión de un símbolo estará formada por el conjunto de valores de observación para los que la probabilidad a posteriori de ese símbolo es la mayor de todas. La región de decisión de b_0 (o de \mathbf{a}_0), es el conjunto de puntos para los que $p_{B|q}(b_0|q)$ es mayor que las probabilidades a posteriori de los otros 3 símbolos; la región de decisión de b_1 (o de \mathbf{a}_1), es el conjunto de puntos para los que $p_{B|q}(b_1|q)$ es mayor que las probabilidades a posteriori de los otros 3 símbolos, y así sucesivamente. En la figura se ilustra esta forma de obtener las regiones de decisión para la constelación del ejemplo.

La expresión analítica de las probabilidades *a posteriori* puede obtenerse mediante la regla de Bayes

$$p_{B|q}(b_j|\mathbf{q}_0) = \frac{p_B(b_j) \cdot f_{q|B}(\mathbf{q}_0|b_j)}{f_q(\mathbf{q}_0)}$$

Teniendo en cuenta que $B = b_j$ implica que $\mathbf{A} = \mathbf{a}_j$ y viceversa,

$$f_{q|B}(\mathbf{q}_0|b_j) = f_{q|\mathbf{A}}(\mathbf{q}_0|\mathbf{a}_j)$$

Sustituyendo esta distribución en la expresión de las probabilidades *a posteriori*, el criterio MAP se reduce a encontrar el símbolo b_i que cumple

$$\frac{p_B(b_i) \cdot f_{q|\mathbf{A}}(\mathbf{q}_0|\mathbf{a}_i)}{f_q(\mathbf{q}_0)} > \frac{p_B(b_j) \cdot f_{q|\mathbf{A}}(\mathbf{q}_0|\mathbf{a}_j)}{f_q(\mathbf{q}_0)} \quad j = 0, \dots, M - 1, j \neq i.$$

Como $f_q(\mathbf{q}_0)$ es una cantidad no negativa independiente del símbolo transmitido, esta condición es equivalente a la condición

$$p_B(b_i) \cdot f_{q|\mathbf{A}}(\mathbf{q}_0|\mathbf{a}_i) > p_B(b_j) \cdot f_{q|\mathbf{A}}(\mathbf{q}_0|\mathbf{a}_j) \quad j = 0, \dots, M - 1, j \neq i,$$

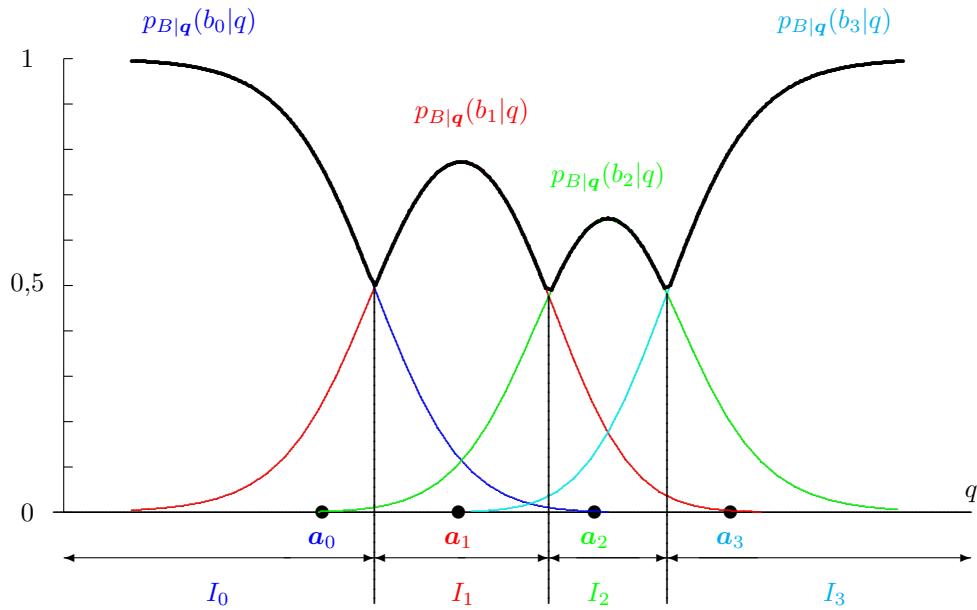


Figura 4.36: Obtención de las regiones de decisión mediante el criterio de máximo a posteriori para un ejemplo con una constelación de 4 símbolos en un espacio unidimensional.

o la expresión equivalente

$$p_A(\mathbf{a}_i) \cdot f_{q|A}(\mathbf{q}_0|\mathbf{a}_i) > p_A(\mathbf{a}_j) \cdot f_{q|A}(\mathbf{q}_0|\mathbf{a}_j) \quad j = 0, \dots, M - 1, j \neq i.$$

A continuación se muestra un ejemplo que ilustra la aplicación de este criterio bajo las condiciones analizadas en la Sección 4.4.4, i.e. transmisión sobre un canal gaussiano. Recuerde que en ese caso las distribuciones condicionales de la observación eran gaussianas de la dimensión del espacio de señales, N , de media el símbolo transmitido y varianza $N_0/2$ en todas las direcciones del espacio

$$f_{q|A}(\mathbf{q}|\mathbf{a}_i) = \mathcal{N}^N \left(\mathbf{a}_i, \frac{N_0}{2} \right).$$

Ejemplo

Se considera un sistema binario ($M = 2$) que emplea un espacio de señal unidimensional ($N = 1$). En este caso tanto \mathbf{A} como \mathbf{q} son escalares, y las probabilidades de símbolos son $p_B(b_1) = \frac{2}{3}$ y $p_B(b_0) = \frac{1}{3}$. Es decir, el símbolo b_1 es el doble de probable. La figura 4.37 muestra las curvas de $p_B(b_1) \cdot f_{q|A}(\mathbf{q}|\mathbf{a}_1)$ y $p_B(b_0) \cdot f_{q|A}(\mathbf{q}|\mathbf{a}_0)$, respectivamente cuando la distribución de \mathbf{q} condicionada a cada símbolo es gaussiana de media el símbolo transmitido y varianza $N_0/2$.

Si el valor $\mathbf{q} = \mathbf{q}_0$ en el receptor es mayor que \mathbf{q}_2 , el punto en que se cortan las curvas, entonces se cumple

$$p_B(b_0) \cdot f_{q|A}(\mathbf{q}_0|\mathbf{a}_0) > p_B(b_1) \cdot f_{q|A}(\mathbf{q}_0|\mathbf{a}_1),$$

por lo que tales puntos formarán parte de la región de decisión I_0 , mientras que si es menor que \mathbf{q}_2

$$p_B(b_1) \cdot f_{q|A}(\mathbf{q}_0|\mathbf{a}_1) > p_B(b_0) \cdot f_{q|A}(\mathbf{q}_0|\mathbf{a}_0),$$

por lo que tales puntos formarán parte de la región de decisión I_1 . Por tanto, las regiones de decisión son

$$I_0 = (\mathbf{q}_2, \infty)$$

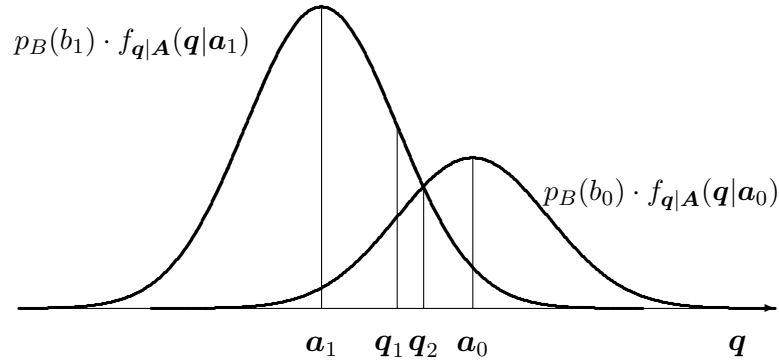


Figura 4.37: Aplicación del criterio MAP con símbolos no equiprobables.

$$I_1 = (-\infty, q_2)$$

El valor que define el cambio de región de decisión es el punto $q = q_2$ en que se cortan ambas curvas, es decir

$$p_B(b_1) \cdot f_{q|A}(q_2|a_1) = p_B(b_0) \cdot f_{q|A}(q_2|a_0).$$

Naturalmente, el decisor queda unívocamente determinado mediante la relación

$$q \in I_i \Leftrightarrow \hat{B} = b_i.$$

En un gran número de sistemas, todos los símbolos tienen la misma probabilidad de ser transmitidos. Bajo la hipótesis de símbolos equiprobables, $p_B(b_j) = 1/M$, las probabilidades *a priori* de los símbolos pasan a ser irrelevantes en la comparación, por lo que la regla de decisión para un valor de observación $q = q_0$ se reduce a encontrar el símbolo b_i que cumple

$$f_{q|A}(q_0|a_i) > f_{q|A}(q_0|a_j) \quad j = 0, \dots, M - 1, j \neq i,$$

es decir, el símbolo que maximiza la función $f_{q|A}(q_0|a_i)$. Esta particularización del criterio MAP para símbolos equiprobables recibe el nombre de *criterio de máxima verosimilitud* o criterio ML (del inglés “*maximum likelihood*”). El nombre viene de la propia función $f_{q|A}(q_0|a_i)$, denominada función de verosimilitud ya que representa una medida de certidumbre o verosimilitud de que la hipótesis cierta sea que el símbolo transmitido fuese a_i cuando la entrada del decisor es $q = q_0$.

Ejemplo

Se considera el mismo sistema binario ($M = 2$) en un espacio de señal unidimensional ($N = 1$) del ejemplo anterior, pero en este caso con símbolos equiprobables. La figura 4.38 muestra las curvas de $p_B(b_1) \cdot f_{q|A}(q|a_1)$ y $p_B(b_0) \cdot f_{q|A}(q|a_0)$ respectivamente. Ahora, eliminar las probabilidades *a priori* de cada símbolo en la comparación no modifica la representación más que en un factor de escala, ya que al ser los símbolos equiprobables $p_B(b_0) = p_B(b_1)$.

Si el valor de $q = q_0$ en el receptor es mayor que q_1 , entonces se cumple

$$p_B(b_0) \cdot f_{q|A}(q_0|a_0) > p_B(b_1) \cdot f_{q|A}(q_0|a_1),$$

o lo que es lo mismo

$$f_{q|A}(q_0|a_0) > f_{q|A}(q_0|a_1),$$

por lo que tales puntos formarán parte de la región de decisión I_0 , mientras que si es menor que q_1

$$p_B(b_1) \cdot f_{q|A}(q_0|a_1) > p_B(b_0) \cdot f_{q|A}(q_0|a_0),$$

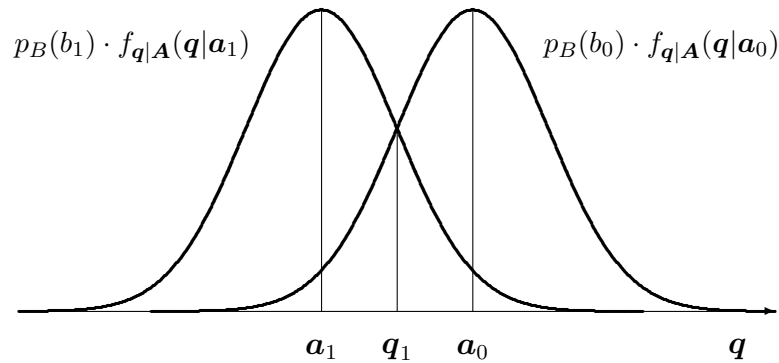


Figura 4.38: Aplicación del criterio MAP con símbolos equiprobables, en cuyo caso se particulariza en el criterio de máxima verosimilitud.

o lo que es lo mismo

$$f_{q|A}(q_0|a_1) > f_{q|A}(q_0|a_0),$$

por lo que tales puntos formarán parte de la región de decisión I_1 . Esto nos lleva a ver que las regiones de decisión son

$$I_0 = (q_1, \infty)$$

$$I_1 = (-\infty, q_1).$$

Intuitivamente se ve que en este caso que como ambas funciones gaussianas tienen la misma varianza y factor de escala, el umbral es el punto medio entre sus centros, $q_1 = \frac{a_0 + a_1}{2}$.

Cuando el espacio de señales es unidimensional, como en los dos ejemplos vistos anteriormente, los valores que separan las regiones de decisión se suelen denominar *umbrales de decisión*. Si el espacio de señales es multidimensional, las curvas ($N = 2$) o superficies que separan las regiones de decisión se denominan *fronteras de decisión*. En los dos ejemplos anteriores el umbral de decisión es q_2 para el primer ejemplo, y q_1 , el punto medio entre los símbolos a_0 y a_1 , en el segundo ejemplo, cuando los símbolos son equiprobables. Se puede ver que cuando los símbolos no se transmiten con la misma probabilidad, el decisor óptimo tiende a aumentar la región de decisión de los símbolos que se transmiten con mayor probabilidad, lo que parece bastante razonable de forma intuitiva (además de verse refrendado por los desarrollos analíticos pertinentes). Al modificar las probabilidades *a priori* para dar más probabilidad al símbolo b_1 , el umbral se ha desplazado de q_1 a q_2 , aumentando la región de decisión del símbolo más probable con respecto al caso de símbolos igualmente probables.

Para finalizar, se particularizará la regla de decisión para el criterio ML (o MAP con símbolos equiprobables) para el caso en que las distribuciones condicionales de la observación son gaussianas, como en el caso visto en la Sección 4.4.4. Introduciendo la función densidad de probabilidad a la entrada del receptor en la formulación del decisor ML se obtiene que la decisión para una observación $\mathbf{q} = \mathbf{q}_0$ será el símbolo b_i (o equivalentemente \mathbf{q}_0 pertenece a I_i) si se cumple

$$\frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{q}_0 - \mathbf{a}_i\|^2}{N_0}} > \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{q}_0 - \mathbf{a}_j\|^2}{N_0}} \quad j = 0, \dots, M - 1, j \neq i.$$

Multiplicando ambos términos de la desigualdad por $(\pi N_0)^{N/2}$ se obtiene

$$e^{-\frac{\|\mathbf{q}_0 - \mathbf{a}_i\|^2}{N_0}} > e^{-\frac{\|\mathbf{q}_0 - \mathbf{a}_j\|^2}{N_0}} \quad j = 0, \dots, M - 1, j \neq i.$$

Teniendo en cuenta que la función exponencial es una función monótona creciente y, por tanto, cumple

$$e^a > e^b \Leftrightarrow a > b,$$

la expresión anterior es equivalente a decir

$$-\frac{\|\mathbf{q}_0 - \mathbf{a}_i\|^2}{N_0} > -\frac{\|\mathbf{q}_0 - \mathbf{a}_j\|^2}{N_0} \quad j = 0, \dots, M-1, j \neq i,$$

y multiplicando por N_0 y teniendo en cuenta el signo negativo se llega a la condición

$$\|\mathbf{q}_0 - \mathbf{a}_i\|^2 < \|\mathbf{q}_0 - \mathbf{a}_j\|^2 \quad j = 0, \dots, M-1, j \neq i.$$

Aplicando la definición de la norma de un vector

$$\|\mathbf{q}_0 - \mathbf{a}_i\|^2 = \sum_{k=0}^{N-1} |q_{0,k} - a_{i,k}|^2 = |d(\mathbf{q}_0, \mathbf{a}_i)|^2.$$

Por tanto, finalmente se establece que la regla de decisión se limita a escoger el símbolo más cercano al vector de observación $\mathbf{q} = \mathbf{q}_0$. Alternativamente, se puede decir que la región de decisión de un símbolo, I_i , estará formada por todos los puntos del espacio de \mathbf{q} que se encuentran más cerca del símbolo, \mathbf{a}_i , que de cualquier otro símbolo de la constelación. Un esquema que ilustra el decisor resultante de este criterio, que se denomina *criterio de mínima distancia euclídea*, se muestra en la Figura 4.39.

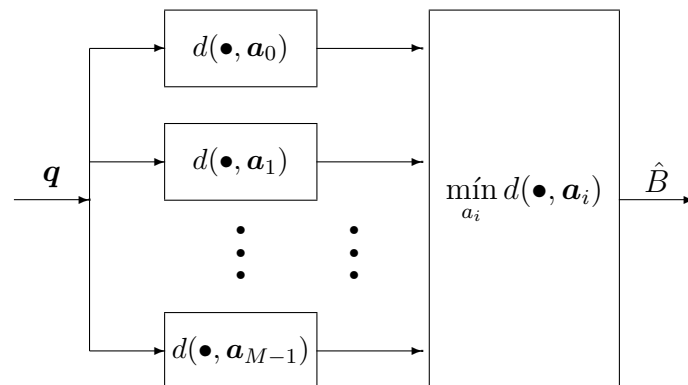


Figura 4.39: Estructura del decisor de mínima distancia euclídea.

En este punto es necesario realizar las siguientes precisiones:

- Este último desarrollo que nos ha llevado a formular el decisor ML como un decisor de mínima distancia euclídea está basado en una función densidad de probabilidad gaussiana a la entrada del decisor, y esta función densidad de probabilidad gaussiana viene dada por la naturaleza del ruido térmico presente habitualmente en canales de comunicaciones. Cuando el ruido que aparece en el canal no tenga una distribución gaussiana, como ocurre por ejemplo en algunos sistemas de comunicaciones basados en fibra óptica, el desarrollo del decisor óptimo conducirá a reglas de decisión diferentes.
- La definición de producto escalar que se adoptó en la Sección 4.2.2, no es la única posible, y se adoptó en su momento sin una justificación estricta. Esta definición da lugar a que la medida

de distancia sobre el espacio de Hilbert resultante sea la de distancia euclídea. Ahora es posible justificar que para el caso de estadística gaussiana para el ruido, esta definición resulta conveniente al permitir simplificar notablemente la regla de decisión para el caso de símbolos equiprobables, que por otra parte es el caso más frecuente en sistemas de comunicaciones digitales.

4.5.3. Cálculo de las probabilidades de error

En el apartado anterior se han obtenido las reglas de diseño de un decisor que permiten obtener la menor probabilidad de error. En esta sección se estudiará como evaluar dicha probabilidad de error en diferentes tipos de sistemas. Como se ha visto con anterioridad, en un sistema de comunicaciones las prestaciones vienen determinadas por la constelación transmitida (junto con la probabilidad con la que se transmite cada símbolo de la constelación), y son independientes de la base ortonormal que define el espacio de señales del sistema, siempre y cuando esta se elija de forma apropiada a las características del canal. Por ello, en esta sección tras plantear de forma general el problema de evaluar la probabilidad de error para un sistema, se particularizará el estudio para distintos tipos de constelaciones.

Cálculo exacto de la probabilidad de error de símbolo

La probabilidad de error de símbolo de un sistema de comunicaciones digitales se define cómo la probabilidad de que la decisión de un cierto símbolo en un instante n no sea correcta

$$P_e = P(\hat{B}[n] \neq b_i | B[n] = b_i).$$

El cálculo de esta probabilidad de error se obtiene promediando las probabilidades de error condicionales, es decir, las probabilidades de error condicionadas a que se ha transmitido cada uno de los símbolos de la constelación. Por notación, y teniendo en cuenta la relación unívoca entre un símbolo y la representación vectorial de su señal asociada, se representarán estas probabilidades condicionales como

$$P_{e|B[n]=b_i} = P_{e|A[n]=\mathbf{a}_i} \equiv P_{e|\mathbf{a}_i}.$$

La probabilidad de error de símbolo se obtendrá promediando estas probabilidades condicionales teniendo en cuenta la probabilidad con la que se transmitirá cada símbolo

$$P_e = \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \cdot P_{e|\mathbf{a}_i}.$$

Por tanto, el problema se reduce al cálculo de las probabilidades de error condicionales. Para ello sólo hay que analizar bajo qué circunstancias se produce un error cuando se ha transmitido un símbolo, $\mathbf{A}[n] = \mathbf{a}_i$, y evaluar la probabilidad de que se produzcan tales circunstancias. De nuevo, por simplicidad en la notación se obviará el índice temporal n (lo que es posible debido a la independencia de los símbolos y observaciones en distintos instantes temporales). Cuando se ha transmitido el símbolo $\mathbf{A} = \mathbf{a}_i$, se produce una decisión errónea cuando se decide $\hat{B} = b_j \neq b_i$, y esto sucede cuando el valor de observación no está en la región de decisión del símbolo, I_i . Esto significa que la probabilidad de error condicional para \mathbf{a}_i es la probabilidad de que habiendo transmitido dicho símbolo, el valor de observación no esté en la región de decisión I_i . Como la

distribución de la observación cuando se transmite \mathbf{a}_i es $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i)$, esta probabilidad se obtiene como

$$P_{e|\mathbf{a}_i} = \int_{\mathbf{q} \notin I_i} f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i) d\mathbf{q}.$$

Comentar finalmente que para poder obtener la probabilidad de error de símbolo que tiene un sistema, como se deduce claramente de las expresiones anteriores, hay que conocer los siguientes parámetros:

- probabilidades a priori de cada símbolo, $p_B(b_i) = p_A(\mathbf{a}_i)$;
- regiones de decisión de cada símbolo, I_i ;
- distribuciones condicionales de la observación para cada símbolo, $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i)$;

en todos los casos para $i \in \{0, 1, \dots, M-1\}$. A continuación se realizará el cálculo de la probabilidad de error para distintos tipos de constelaciones. Por defecto, si no se indica de forma explícita lo contrario, se asumirán probabilidades a priori idénticas (símbolos equiprobables), y distribuciones condicionales de la observación gaussianas, como las obtenidas en la transmisión sobre un canal gaussiano (ver Sección 4.4.4).

Constelación binaria ($M = 2$) en espacio unidimensional ($N = 1$)

Como ejemplo inicial se considera una constelación de dos símbolos en un espacio unidimensional, el símbolo \mathbf{a}_0 tiene como coordenada $+A$ y el símbolo \mathbf{a}_1 , $-A$. Bajo la asunción de símbolos equiprobables y transmisión sobre canal gaussiano, la región de decisión I_1 estará formada por todos los valores de \mathbf{q} más cercanos a \mathbf{a}_1 que a \mathbf{a}_0 y al revés para la región de decisión I_0 . El umbral de decisión es, por tanto, cero, y las regiones de decisión

$$\text{Umbral } q_u = 0 \rightarrow I_0 = [0, \infty), I_1 = (-\infty, 0),$$

tal y como se muestra en la Figura 4.40.

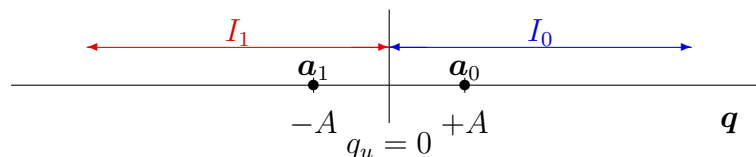


Figura 4.40: Constelación y regiones de decisión para el ejemplo de constelación binaria en espacio unidimensional.

Calculemos ahora las probabilidades de error condicionales, y empezemos por el símbolo \mathbf{a}_0 . En este caso, la distribución condicional de la observación para este símbolo es la que se muestra en la Figura 4.41.

Se trata de una distribución $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0)$ gaussiana de media $+A$, la representación vectorial del símbolo transmitido, y varianza $N_0/2$. La probabilidad de error condicional para \mathbf{a}_0 es la integral

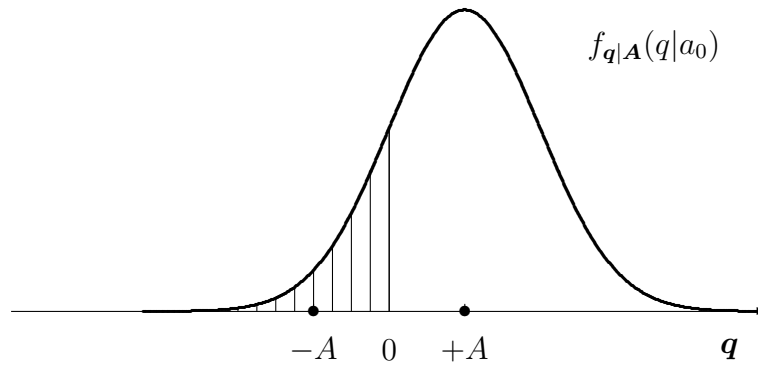


Figura 4.41: Distribución condicional de la observación para el símbolo \mathbf{a}_0 .

de esta distribución condicional fuera de la región de decisión de \mathbf{a}_0 , que en este caso es el área resaltada en la figura sobre la distribución

$$P_{e|\mathbf{a}_0} = \int_{q \notin I_0} f_{q|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0) dq = Q\left(\frac{A}{\sqrt{N_0/2}}\right).$$

Como cuando se transmite el símbolo $\mathbf{A} = \mathbf{a}_i$ la distribución de la observación es $f_{q|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0)$, y el error en la decisión se produce cuando la observación \mathbf{q} toma valores que caen fuera de I_0 (en este caso binario eso significa que cae en I_1), la probabilidad de que esto ocurra se calcula integrando la distribución de \mathbf{q} fuera de I_0 .

Si se calcula la probabilidad de error condicional para el símbolo \mathbf{a}_1 , en este caso la distribución condicional de la observación para este símbolo es la que se muestra en la Figura 4.42.

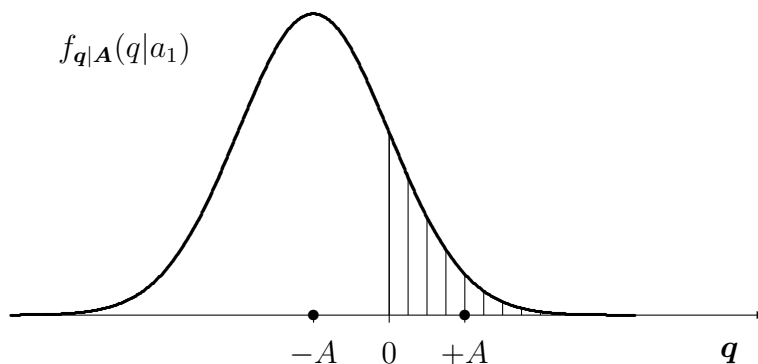


Figura 4.42: Distribución condicional de la observación para el símbolo \mathbf{a}_1 .

Se trata de una distribución $f_{q|\mathbf{A}}(\mathbf{q}|\mathbf{a}_1)$ gaussiana de media $-A$ y varianza $N_0/2$. La probabilidad de error condicional es la integral de esta distribución condicional fuera de su región de decisión, que en este caso es el área resaltada en la figura sobre la distribución

$$P_{e|\mathbf{a}_1} = \int_{q \notin I_1} f_{q|\mathbf{A}}(\mathbf{q}|\mathbf{a}_1) dq = Q\left(\frac{A}{\sqrt{N_0/2}}\right).$$

Ahora, como cuando se transmite el símbolo $\mathbf{A} = \mathbf{a}_i$ la distribución de la observación es $f_{q|\mathbf{A}}(\mathbf{q}|\mathbf{a}_1)$, y el error en la decisión se produce cuando la observación \mathbf{q} toma valores que caen fuera de I_1

(en este caso binario eso significa que cae en I_0), la probabilidad de que esto ocurra se calcula integrando la distribución de \mathbf{q} fuera de I_1 .

Una vez calculadas las probabilidades de error condicionales, la probabilidad de error de símbolo se obtiene promediandolas, lo que en este caso significa

$$P_e = \frac{1}{2}P_{e|a_0} + \frac{1}{2}P_{e|a_1} = \frac{1}{2} \int_{\mathbf{q} \notin I_0} f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0) d\mathbf{q} + \frac{1}{2} \int_{\mathbf{q} \notin I_1} f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_1) d\mathbf{q}$$

La Figura 4.43 muestra la interpretación gráfica del significado de esta probabilidad de error, y cómo la modificación del umbral de decisión haría aumentar dicha probabilidad de error.

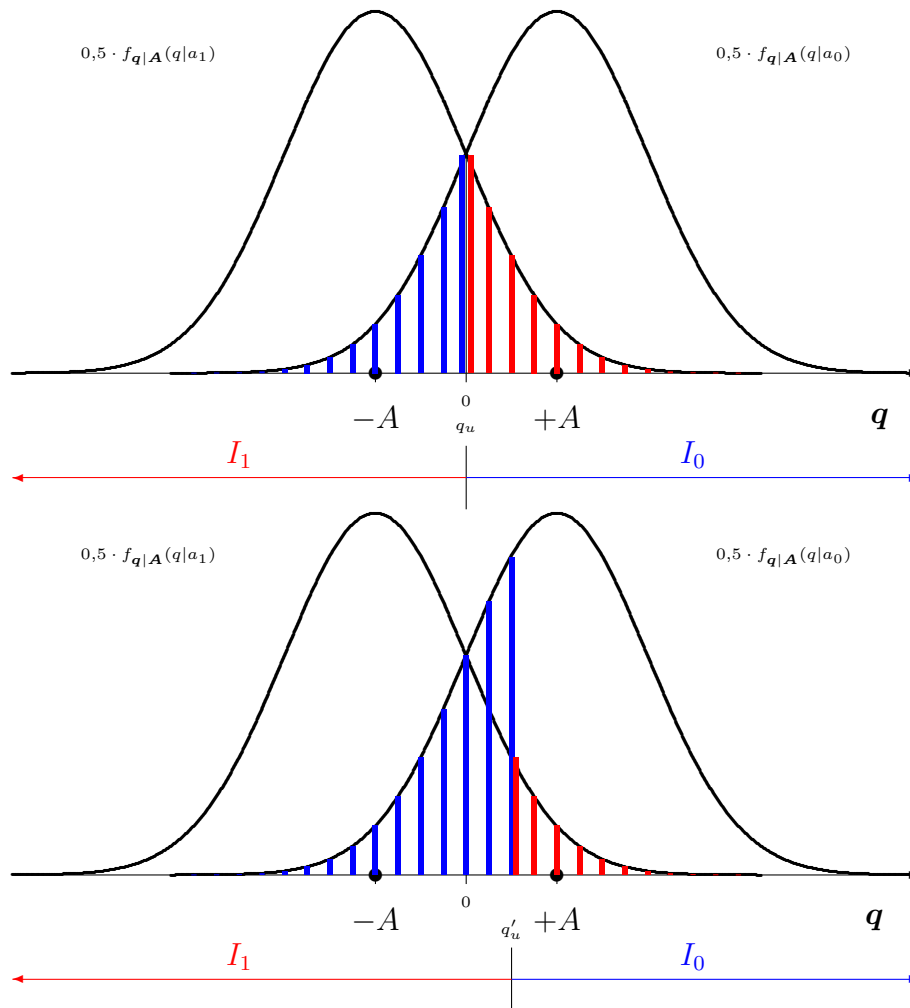


Figura 4.43: Interpretación gráfica de la probabilidad de error de símbolo para el decisor óptimo y para otro decisor.

De este ejemplo se pueden extrapolar de forma inmediata el mismo resultado cuando los valores de la representación vectorial de los dos símbolos es otra arbitraria, \mathbf{a}_0 y \mathbf{a}_1 . Independientemente de los valores de los símbolos, para símbolos equiprobables y distribuciones condicionales gaussianas el umbral está en el punto medio entre ambos

$$\mathbf{q}_u = \frac{\mathbf{a}_0 + \mathbf{a}_1}{2},$$

de modo que la distancia de cada símbolo al umbral, que define el argumento de la función $Q(x)$ para evaluar la integral de la distribución gaussiana, es la mitad de la distancia entre los dos

símbolos

$$d(\mathbf{a}_0, \mathbf{q}_u) = d(\mathbf{a}_1, \mathbf{q}_u) = \frac{d(\mathbf{a}_0, \mathbf{a}_1)}{2}$$

por lo que las probabilidades de error condicionales son iguales y la probabilidad de error total es

$$P_e = Q\left(\frac{d(\mathbf{a}_0, \mathbf{a}_1)}{2\sqrt{N_0/2}}\right).$$

Decisor M -ario en espacio unidimensional

Para ilustrar este caso se considera un ejemplo de un sistema con una constelación de $M = 4$ símbolos que se transmiten con la misma probabilidad sobre un canal gaussiano. Las coordenadas de la representación vectorial de los símbolos son

$$\mathbf{a}_0 = -3, \mathbf{a}_1 = -1, \mathbf{a}_2 = +1, \mathbf{a}_3 = +3$$

En este caso, al poder aplicarse el criterio de mínima distancia euclídea para el diseño del decisor, los umbrales del decisor aparecen en los puntos medios entre cada dos símbolos de la constelación

$$q_{u1} = -2, q_{u2} = 0, q_{u3} = +2,$$

con lo que las regiones de decisión son

$$I_0 = (-\infty, -2], I_1 = (-2, 0], I_2 = (0, +2], I_3 = (+2, +\infty)$$

tal y como se ilustra en la Figura 4.44.

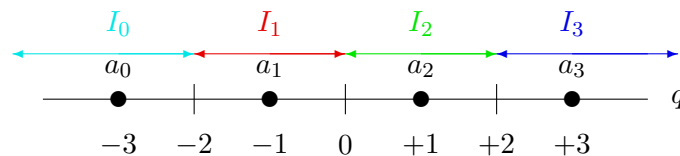


Figura 4.44: Constelación y regiones de decisión para el ejemplo de una constelación de 4 símbolos en espacio unidimensional.

La probabilidad de error condicional para el símbolo \mathbf{a}_0 , requiere la utilización de la distribución condicional de la observación para este símbolo, que es la que se muestra en la Figura 4.45.

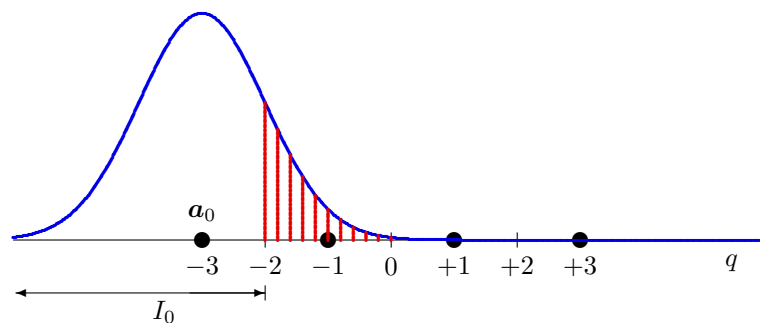


Figura 4.45: Distribución condicional de la observación para el símbolo \mathbf{a}_0 .

Se trata de una distribución $f_{q|A}(q|a_0)$ gaussiana de media -3 , la representación vectorial del símbolo transmitido, y varianza $N_0/2$. La probabilidad de error condicional es la integral de esta distribución condicional fuera de su región de decisión, que en este caso es el área resaltada en la figura sobre la distribución

$$P_{e|a_0} = \int_{q \notin I_0} f_{q|A}(q|a_0) dq = Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

La probabilidad de error condicional para el símbolo a_1 , se basa en la distribución condicional de la observación para este símbolo, que es la que se muestra en la Figura 4.46.

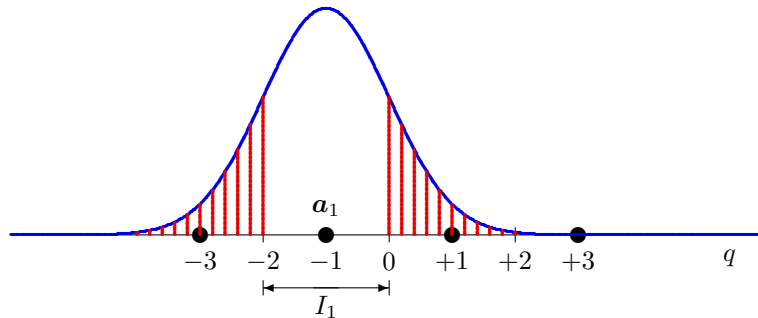


Figura 4.46: Distribución condicional de la observación para el símbolo a_1 .

Se trata de una distribución $f_{q|A}(q|a_1)$ gaussiana de media -1 , la representación vectorial del símbolo transmitido, y varianza $N_0/2$. La probabilidad de error condicional es la integral de esta distribución condicional fuera de su región de decisión, que en este caso es el área resaltada en la figura sobre la distribución

$$P_{e|a_1} = \int_{q \notin I_1} f_{q|A}(q|a_1) dq = 2Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

La distribución condicional de la observación para el símbolo a_2 se muestra en la Figura 4.47.

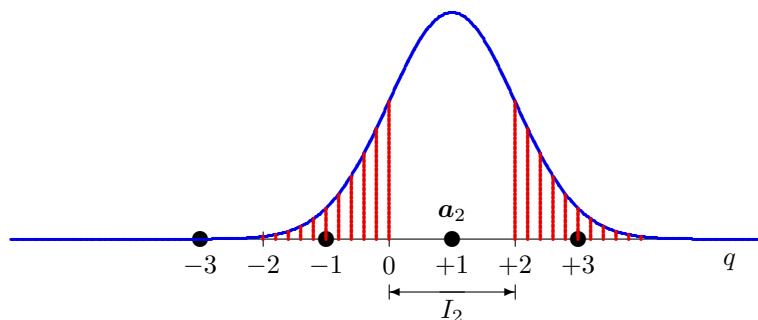


Figura 4.47: Distribución condicional de la observación para el símbolo a_2 .

Se trata de una distribución $f_{q|A}(q|a_2)$ gaussiana de media $+1$, la representación vectorial del símbolo transmitido, y varianza $N_0/2$. La probabilidad de error condicional es la integral de esta distribución condicional fuera de su región de decisión, que en este caso es el área resaltada en la figura sobre la distribución

$$P_{e|a_2} = \int_{q \notin I_2} f_{q|A}(q|a_2) dq = 2Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

Finalmente, para el símbolo \mathbf{a}_3 , la distribución condicional de la observación para este símbolo, que es la que se muestra en la Figura 4.48.

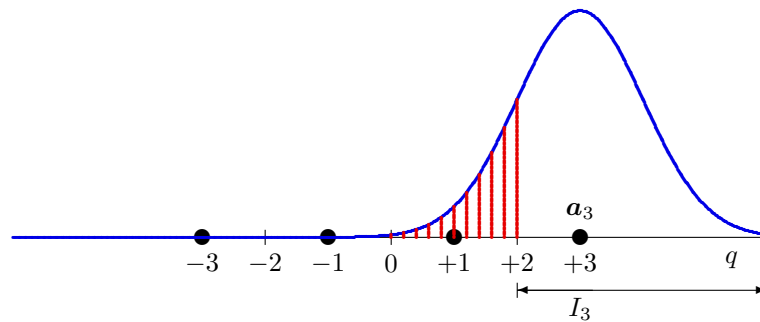


Figura 4.48: Distribución condicional de la observación para el símbolo \mathbf{a}_3 .

Se trata de una distribución $f_{q|\mathbf{A}}(\mathbf{q}|\mathbf{a}_3)$ gaussiana de media +3, la representación vectorial del símbolo transmitido, y varianza $N_0/2$. La probabilidad de error condicional es la integral de esta distribución condicional fuera de su región de decisión, que en este caso es el área resaltada en la figura sobre la distribución

$$P_{e|\mathbf{a}_3} = \int_{q \notin I_3} f_{q|\mathbf{A}}(q|\mathbf{a}_3) dq = Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

En este ejemplo, dada la simetría de las regiones de decisión, claramente se puede ver que

$$P_{e|\mathbf{a}_0} = P_{e|\mathbf{a}_3} \text{ y } P_{e|\mathbf{a}_1} = P_{e|\mathbf{a}_2},$$

lo que se podría haber aprovechado para simplificar el cálculo.

Una vez calculadas las probabilidades de error condicionales, la probabilidad de error de símbolo se obtiene promediándolas

$$P_e = \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) P_{e|\mathbf{a}_i} = \frac{1}{4} \sum_{i=0}^{M-1} P_{e|\mathbf{a}_i} = \frac{3}{2} Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

Sistema binario ($M = 2$) en espacio multidimensional ($N > 1$)

Para ilustrar este caso se considera la siguiente constelación

$$\mathbf{a}_0 = \begin{bmatrix} A \\ 0 \end{bmatrix}, \quad \mathbf{a}_1 = \begin{bmatrix} 0 \\ A \end{bmatrix}.$$

En ella, los símbolos son vectores de dimensión 2. Para el caso por defecto de símbolos equiprobables y bajo transmisión sobre un canal gaussiano, se puede aplicar de nuevo el criterio de mínima distancia euclídea, con lo que la región de decisión de cada símbolo está formada por los puntos del espacio de \mathbf{q} (en este caso el plano bidimensional formado por las dos coordenadas, q_0 y q_1) que están más cerca de ese símbolo que del otro. La frontera de decisión que separa los puntos más cercanos de \mathbf{a}_0 de los de \mathbf{a}_1 es la recta $q_0 = q_1$. La región de decisión I_0 es el semiplano de la derecha, y la región de decisión I_1 es el semiplano de la izquierda

$$I_0 = \left\{ \mathbf{q} = \begin{bmatrix} q_0 \\ q_1 \end{bmatrix} \mid q_0 \geq q_1 \right\} \quad I_1 = \left\{ \mathbf{q} = \begin{bmatrix} q_0 \\ q_1 \end{bmatrix} \mid q_0 < q_1 \right\}$$

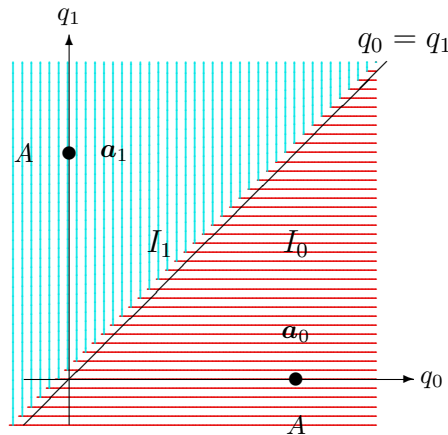


Figura 4.49: Frontera de decisión y regiones de decisión para el ejemplo de sistema binario en espacio bidimensional.

tal y como se representa en la Figura 4.49.

Si el símbolo transmitido es \mathbf{a}_0 la función de densidad de probabilidad de la entrada al receptor $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0)$ es una gaussiana de media $\mathbf{a}_0 = \begin{bmatrix} A \\ 0 \end{bmatrix}$, componentes independientes y varianza de cada una de esas componentes igual a $N_0/2$, y para el símbolo \mathbf{a}_1 $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_1)$ es una gaussiana de media $\mathbf{a}_1 = \begin{bmatrix} 0 \\ A \end{bmatrix}$, de nuevo con componentes independientes y varianza $N_0/2$

$$f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i) = \frac{1}{\pi N_0} e^{-\frac{\|\mathbf{q}-\mathbf{a}_i\|^2}{N_0}} = \frac{1}{\pi N_0} e^{-\frac{(q_0 - \sqrt{2T})^2 + q_1^2}{N_0}},$$

tal y como se muestra en la Figura 4.51 (en este caso se representan para $A = 1$, por un lado cada una por separado, y en la figura inferior ambas juntas, cada una dentro de su región de decisión).

Para calcular las probabilidades de error condicionales, $P_{e|\mathbf{a}_i}$, hay que integrar la función densidad de probabilidad de cada símbolo \mathbf{a}_i fuera de su región de decisión, I_i , en este caso el semiplano definido por $q_0 > q_1$ ó $q_0 < q_1$, respectivamente para \mathbf{a}_0 y \mathbf{a}_1 , tal y como se muestra en la Figura 4.51.

Como es bien sabido, la integral de una función gaussiana no tiene una expresión analítica definida. En el caso unidimensional se recurre a la función tabulada $Q(x)$, que se puede calcular de forma numérica, para obtener el valor de la integral de una distribución gaussiana unidimensional. Pero ahora, para espacios bidimensionales, no hay ninguna función tabulada que calcule integrales de distribuciones gaussianas sobre semiplanos. En lugar de resolver el cálculo de forma directa, aprovechando que si se tienen dos símbolos siempre se puede hacer pasar una recta por esos puntos, se va a hacer una transformación del sistema de coordenadas para convertir el problema en otro problema unidimensional equivalente. Para ello se realiza el siguiente cambio de variables:

$$q'_0 = \frac{1}{\sqrt{2}}(q_0 - q_1),$$

$$q'_1 = \frac{1}{\sqrt{2}}(q_0 + q_1 - A).$$

La transformación gira la constelación 45° y desplaza el sistema resultante $\frac{A}{\sqrt{2}}$ hacia abajo, tal y como se ilustra en el Figura 4.52, de forma que los puntos de la constelación pasan ahora a

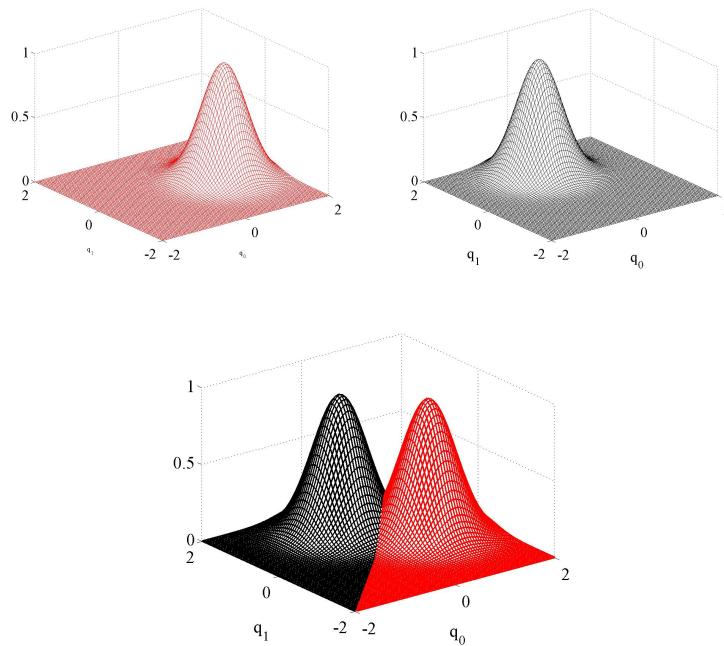


Figura 4.50: Funciones densidad de probabilidad condicionales para el ejemplo de sistema binario en espacio bidimensional. Cada función por separado, y ambas representadas cada una dentro de su región de decisión.

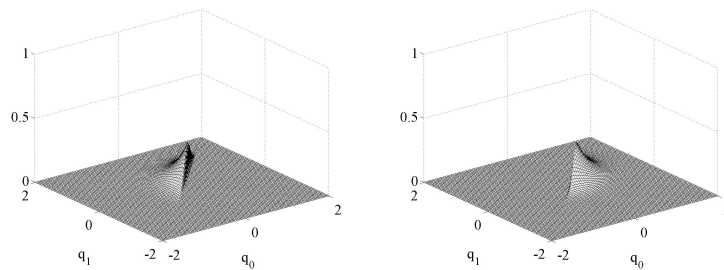


Figura 4.51: Interpretación gráfica de las probabilidades de error condicionales para el ejemplo de sistema binario en espacio bidimensional.

estar sobre un espacio unidimensional, ya que su coordenada sobre el segundo eje, q'_1 , es nula para ambos símbolos,

$$\mathbf{a}'_0 = \begin{bmatrix} +\frac{A}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad \mathbf{a}'_1 = \begin{bmatrix} -\frac{A}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

Podríamos plantearnos eliminar por tanto la segunda coordenada y resolver el problema para la nueva constelación unidimensional resultante, tal y como se muestra en la Figura 4.53. Para poder hacer esto, es necesario que las componentes de ruido sobre los nuevos ejes sean independientes, y hay que comprobar la distribución del ruido en la coordenada q'_0 al realizar la transformación, y ver si sigue siendo una distribución gaussiana de media nula y varianza $N_0/2$. Vamos a ver qué estas condiciones sí se cumplen.

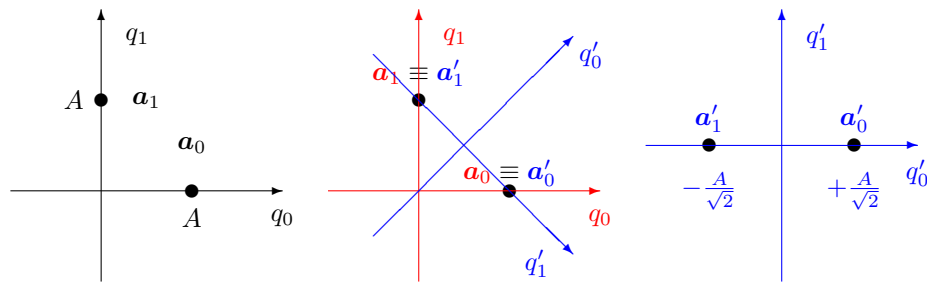


Figura 4.52: Ilustración de la transformación geométrica implícita en el cambio de coordenadas planteado.

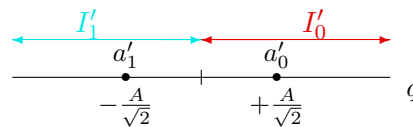


Figura 4.53: Regiones de decisión sobre el nuevo sistema de referencia tras el cambio de coordenadas.

Empezemos por recordar que cuando se transmite el símbolo $\mathbf{A} = \mathbf{a}_i$, se tiene que

$$q_0 = a_{i,0} + z_0, \quad q_1 = a_{i,1} + z_1.$$

Denotaremos la observación recibida como

$$q'_0 = a'_{i,0} + z'_0, \quad q'_1 = a'_{i,1} + z'_1.$$

Si el símbolo transmitido es \mathbf{a}_0 , q'_0 toma la forma

$$\begin{aligned} q'_0|_{\mathbf{A}=\mathbf{a}_0} &= \frac{1}{\sqrt{2}} ((a_{0,0} + z_0) - (a_{0,1} + z_1)) \\ &= \frac{1}{\sqrt{2}} ((A + z_0) - (0 + z_1)) \\ &= \underbrace{\frac{A}{\sqrt{2}}}_{a'_{0,0}} + \underbrace{\frac{1}{\sqrt{2}}(z_0 - z_1)}_{z'_0} \end{aligned}$$

y q'_1

$$\begin{aligned} q'_1|_{\mathbf{A}=\mathbf{a}_0} &= \frac{1}{\sqrt{2}} ((a_{0,0} + z_0) + (a_{0,1} + z_1) - A) \\ &= \frac{1}{\sqrt{2}} ((A + z_0) + (0 + z_1) - A) \\ &= \underbrace{\frac{1}{\sqrt{2}}(z_0 + z_1)}_{z'_1} \end{aligned}$$

De forma análoga, si el símbolo transmitido es \mathbf{a}_1 , q'_0 y q'_1 toman la forma

$$\begin{aligned} q'_0|_{\mathbf{A}=\mathbf{a}_1} &= \frac{1}{\sqrt{2}} ((a_{1,0} + z_0) - (a_{1,1} + z_1)) \\ &= \frac{1}{\sqrt{2}} ((0 + z_0) - (A + z_1)) \\ &= \underbrace{-\frac{A}{\sqrt{2}}}_{a'_{1,0}} + \underbrace{\frac{1}{\sqrt{2}}(z_0 - z_1)}_{z'_0} \end{aligned}$$

y

$$\begin{aligned} q'_1|_{\mathbf{A}=\mathbf{a}_1} &= \frac{1}{\sqrt{2}} ((a_{1,0} + z_0) + (a_{1,1} + z_1) - A) \\ &= \frac{1}{\sqrt{2}} ((0 + z_0) + (A + z_1) - A) \\ &= \underbrace{\frac{1}{\sqrt{2}}(z_0 + z_1)}_{z'_1} \end{aligned}$$

De las dos componentes, sólo q'_0 contiene información sobre el símbolo transmitido; q'_1 contiene ruido. Esto implica que se ha reducido la dimensión del espacio de señal de 2 a 1. Las coordenadas de los nuevos elementos de la constelación respecto a q'_0 , que denotaremos a'_0 y a'_1 , son

$$\begin{aligned} a'_0 &= +\frac{A}{\sqrt{2}}, \\ a'_1 &= -\frac{A}{\sqrt{2}}, \end{aligned}$$

es decir, exactamente las mismas que en la constelación unidimensional considerada anteriormente.

Respecto al ruido, es fácil demostrar que los términos z'_0 y z'_1 son independientes. El primero es proporcional a $z_0 - z_1$ y el segundo a $z_0 + z_1$, ya que la suma (o resta) de dos variables aleatorias con función densidad de probabilidad gaussiana es otra variable aleatoria gaussiana y

$$E[(z_0 - z_1)(z_0 + z_1)] = E[z_0^2 - z_1^2] = \frac{N_0}{2} - \frac{N_0}{2} = 0.$$

Esto implica que el valor de q'_1 es irrelevante para la decisión y que q'_0 es un estadístico suficiente para la detección.

Para determinar la probabilidad de error debemos conocer los estadísticos de la componente de ruido, $\frac{1}{\sqrt{2}}(z_0 - z_1)$, ya que por ahora sólo sabemos que es gaussiana. El valor medio es

$$E[z'_0] = E\left[\frac{1}{\sqrt{2}}(z_0 - z_1)\right] = \frac{1}{\sqrt{2}}E[z_0] - \frac{1}{\sqrt{2}}E[z_1] = 0.$$

Y la varianza vale

$$\begin{aligned} \text{Var}(z'_0) &= E\left[\left(\frac{1}{\sqrt{2}}(z_0 - z_1)\right)^2\right] = \frac{1}{2}E[z_0^2] + \frac{1}{2}E[z_1^2] - E[z_0z_1] \\ &= \frac{1}{2}\frac{N_0}{2} + \frac{1}{2}\frac{N_0}{2} - 0 \\ &= \frac{N_0}{2}. \end{aligned}$$

Finalmente, si la nueva constelación es la misma que en el caso unidimensional y los estadísticos del ruido también (media cero y varianza $N_0/2$), también lo será la probabilidad de error,

$$P_e = Q\left(\frac{d(\mathbf{a}'_0, \mathbf{a}'_1)}{2\sqrt{N_0/2}}\right).$$

La distancia entre los símbolos es la misma que en la constelación original, $d(\mathbf{a}'_0, \mathbf{a}'_1) = d(\mathbf{a}_0, \mathbf{a}_1)$, por lo que es indistinto realizar la medida en el espacio original que en el transformado. De hecho, el transformado ni siquiera hay que calcularlo para evaluar la probabilidad media de error en este caso.

El desarrollo realizado puede extenderse a cualquier constelación binaria en un espacio de dimensión arbitraria, ya que siempre es posible definir una dirección del espacio a través de la recta que pasa por los dos puntos. Así pues, de forma general, la probabilidad de error para sistemas binarios con símbolos equiprobables y con transmisión sobre un canal gaussiano es

$$P_e = Q\left(\frac{d(\mathbf{a}_0, \mathbf{a}_1)}{2\sqrt{N_0/2}}\right),$$

independientemente de la dimensión del espacio de señales.

Decisor M -ario en espacio multidimensional

El caso general de cálculo de la probabilidad media de error en constelaciones multidimensionales es un problema complejo por las formas que pueden tomar las regiones de decisión. Para constelaciones binarias siempre es posible transformar el problema en unidimensional simplificando así la resolución de las integrales necesarias en el cálculo de la probabilidad de error de cada símbolo. Sin embargo, para constelaciones de más de dos símbolos esto no es siempre posible. Como ejemplo, baste considerar las constelaciones de la Figura 4.54, que presenta dos ejemplos de constelaciones bidimensionales de uso común en sistemas de comunicaciones y sus regiones de decisión asociadas. En algunos casos será posible calcular la probabilidad de error exacta, y en otros casos esto no será posible de forma analítica, y habrá que recurrir a cálculos numéricos, a aproximaciones o a cotas de la probabilidad de error.

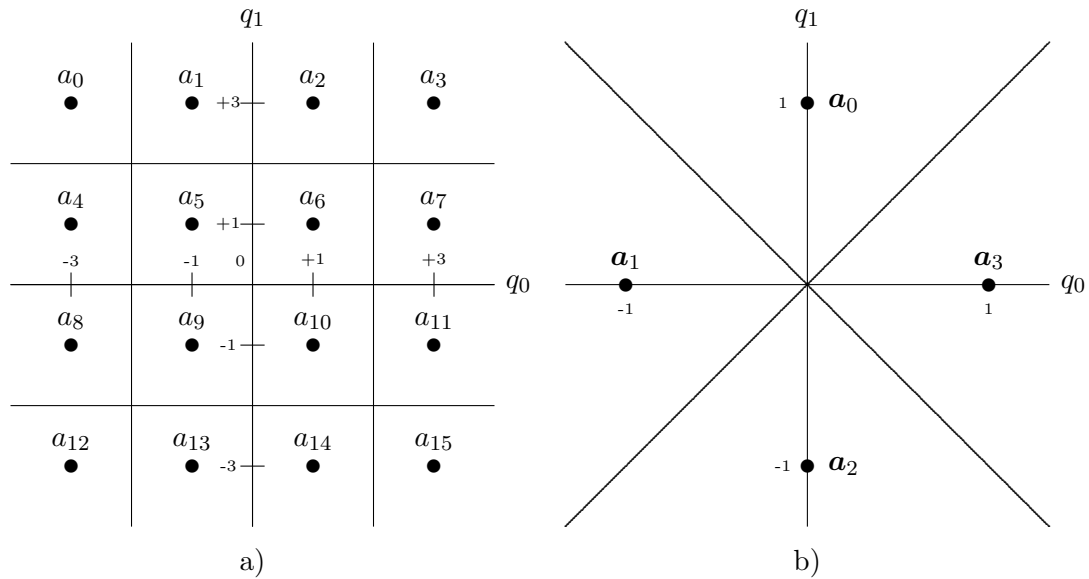


Figura 4.54: Ejemplos de constelaciones bidimensionales y sus fronteras de decisión.

Para sistemas con constelaciones como las de la figura es aún posible calcular de manera sencilla la probabilidad media de error. Veremos que esto es posible cuando las fronteras de las regiones de decisión forman una retícula sobre el espacio, como ocurre con las dos constelaciones de la figura.

Por ejemplo, primero se considera la constelación de la Figura 4.54 a). Veamos cómo se calcularía la probabilidad de error condicional para uno de los símbolos, por ejemplo \mathbf{a}_6 . Este símbolo tiene coordenadas

$$\mathbf{a}_6 = \begin{bmatrix} +1 \\ +1 \end{bmatrix},$$

por lo que la distribución condicional de la observación, $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_6)$, es una distribución gaussiana bidimensional de media \mathbf{a}_6 y varianzas $\sigma^2 = N_0/2$, tal y como se muestra en la Figura 4.55.

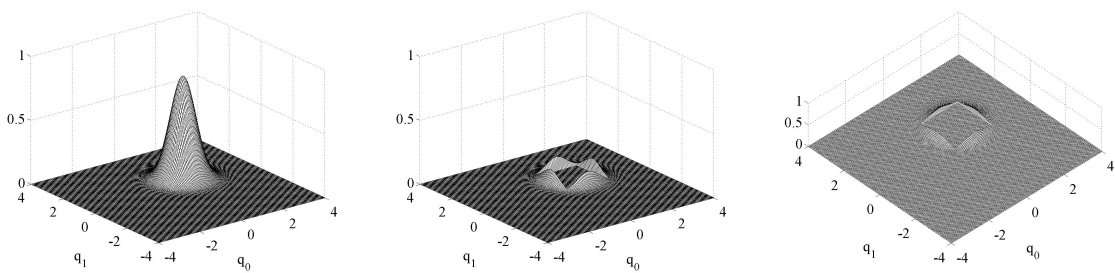


Figura 4.55: Distribución condicional de la observación para el símbolo \mathbf{a}_6 , $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_6)$: distribución completa, y distribución fuera de su región de decisión (dos vistas desde diferentes perspectivas).

La región de decisión de este símbolo es una cuadrícula, que en este caso, al estar alineada con los dos ejes de la representación se puede escribir como dos condiciones independientes (una por cada eje) que deben darse simultáneamente; en este caso

$$0 \leq q_0 < 2 \text{ y } 0 \leq q_1 < 2.$$

No hay expresiones analíticas para calcular directamente la integral de una gaussiana fuera de un cuadrado o un rectángulo, pero se puede transformar este problema 2D en dos problemas

unidimensionales acoplados. Cuando la región de decisión se puede parametrizar en dos condiciones independientes, una por cada dimensión del espacio, que han de cumplirse simultáneamente, la probabilidad de error condicional de un símbolo se puede escribir como

$$P_{e|\mathbf{a}_i} = 1 - P_{a|\mathbf{a}_i} = 1 - P_{a|a_{i,0}} \times P_{a|a_{i,1}} = 1 - [(1 - P_{e|a_{i,0}}) \times (1 - P_{e|a_{i,1}})],$$

donde se han seguido los siguientes pasos:

- La probabilidad de error condicional se puede escribir como 1 menos la probabilidad de acierto condicional, $P_{a|\mathbf{a}_i}$.
- La probabilidad condicional de acierto $P_{a|\mathbf{a}_i}$ se puede escribir como el producto de la probabilidad de acertar en cada una de las dos direcciones del espacio $P_{a|a_{i,0}} \times P_{a|a_{i,1}}$ (ya que la región de decisión se establece con dos condiciones independientes, una sobre cada dirección del espacio).
- La probabilidad de acierto en una de las direcciones del espacio ($P_{a|a_{i,0}}$ ó $P_{a|a_{i,1}}$) se puede escribir como 1 menos la probabilidad de error en esa dirección ($P_{e|a_{i,0}}$ ó $P_{e|a_{i,1}}$). Y esa probabilidad es igual a las probabilidades de error en espacios unidimensionales que ya se ha aprendido a calcular.

Esto permite calcular la probabilidad condicional de error para los símbolos si sus regiones de decisión forman una retícula alineada con los ejes del espacio de observación \mathbf{q} .

Veamos ahora con ejemplos numéricos cómo se calcularía la probabilidad de error total para la constelación anterior. Se pueden agrupar las regiones de decisión en tres tipos, dependiendo del número de regiones de decisión con las que limitan.

- Tipo 1: $\{I_0, I_3, I_{12}, I_{15}\}$
 - Una única frontera de decisión en cada dirección del espacio
- Tipo 2: $\{I_5, I_6, I_9, I_{10}\}$
 - Dos fronteras de decisión en cada dirección del espacio
- Tipo 3: $\{I_1, I_2, I_4, I_7, I_8, I_{11}, I_{13}, I_{14}\}$
 - Una frontera en una de las direcciones del espacio
 - Dos fronteras en la otra dirección

Todos los símbolos de un mismo tipo tienen la misma probabilidad condicional de error, ya que hay que integrar la distribución gaussiana centrada en el símbolo fuera de una región de las mismas dimensiones. Por tanto, es posible tomar un ejemplo de cada tipo, y extrapolar los resultados. Lo ejemplos elegidos pueden ser, por ejemplo: \mathbf{a}_0 (Tipo 1), \mathbf{a}_5 (Tipo 2), \mathbf{a}_7 (Tipo 3). De este modo, la probabilidad de error total será

$$\begin{aligned} P_e &= \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \cdot P_{e|\mathbf{a}_i} \\ &= 4 \times \frac{1}{16} P_{e|\mathbf{a}_0} + 4 \times \frac{1}{16} P_{e|\mathbf{a}_5} + 8 \times \frac{1}{16} P_{e|\mathbf{a}_7}, \end{aligned}$$

que ahora se comprobará que vale

$$P_e = 3Q\left(\frac{1}{\sqrt{N_0/2}}\right) - \frac{9}{4}Q^2\left(\frac{1}{\sqrt{N_0/2}}\right).$$

Empezaremos el cálculo con el símbolo que ejemplifica las regiones de tipo 1, que es \mathbf{a}_0 . El procedimiento de cálculo para este caso, que se ilustra en la Figura 4.56, es el siguiente:

▪ Dirección del espacio q_0

- Media de la distribución gaussiana 1-D: $a_{0,0} = -3$
- Región de decisión : $-\infty < q_0 < -2$

$$P_{a|a_{0,0}} = 1 - P_{e|a_{0,0}} = 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

▪ Dirección del espacio q_1

- Media de la distribución gaussiana 1-D: $a_{0,1} = +3$
- Región de decisión : $+2 \leq q_1 < +\infty$

$$P_{a|a_{0,1}} = 1 - P_{e|a_{0,1}} = 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

▪ Probabilidad de error condicional

$$P_{e|a_0} = 1 - \left[1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)\right]^2 = 2Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q^2\left(\frac{1}{\sqrt{N_0/2}}\right)$$

A continuación realizaremos el cálculo para el símbolo que ejemplifica las regiones de tipo 2, que es \mathbf{a}_5 . El procedimiento de cálculo para este caso, que se ilustra en la Figura 4.57, es el siguiente:

▪ Dirección del espacio q_0

- Media de la distribución gaussiana 1-D: $a_{5,0} = -1$
- Región de decisión : $-2 \leq q_0 < 0$

$$P_{a|a_{5,0}} = 1 - P_{e|a_{5,0}} = 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

▪ Dirección del espacio q_1

- Media de la distribución gaussiana 1-D: $a_{5,1} = +1$
- Región de decisión : $0 \leq q_1 < +2$

$$P_{a|a_{5,1}} = 1 - P_{e|a_{5,1}} = 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

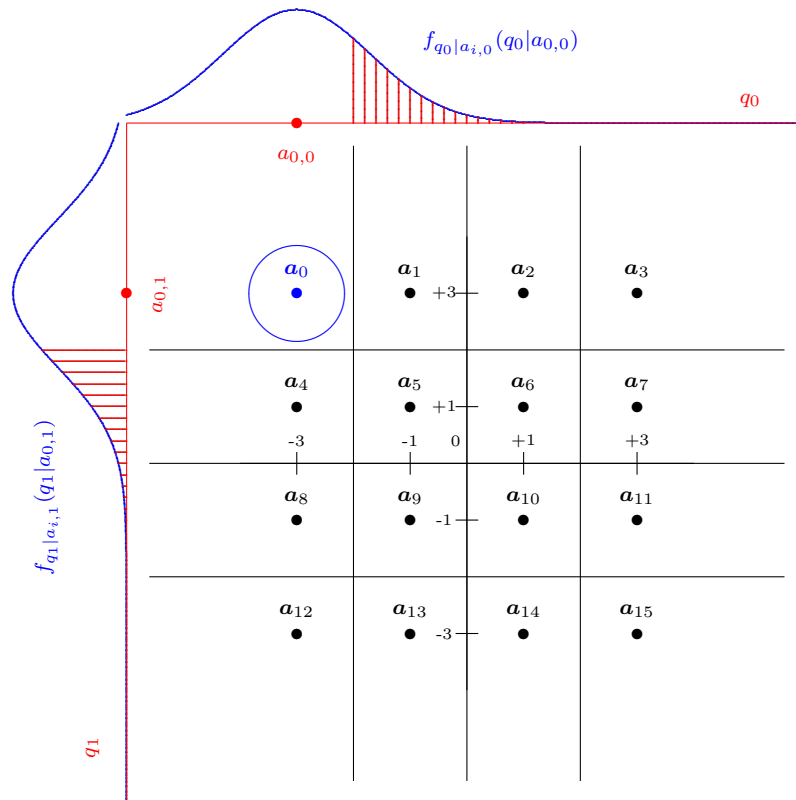


Figura 4.56: Distribuciones condicionales marginales para el símbolo a_0 .

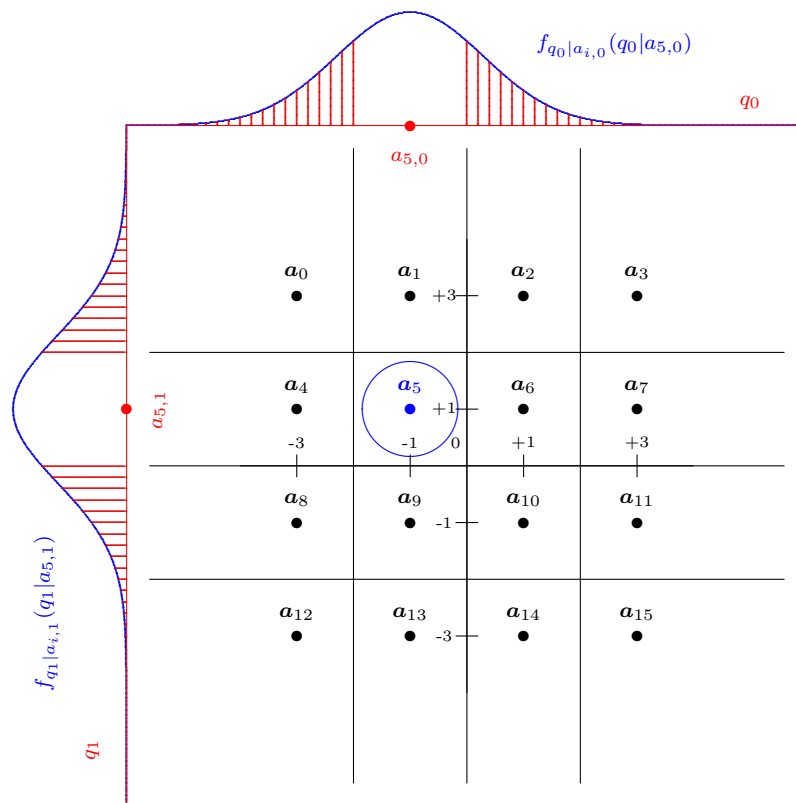


Figura 4.57: Distribuciones condicionales marginales para el símbolo a_5 .

- Probabilidad de error condicional

$$P_{e|a_5} = 1 - \left[1 - 2Q \left(\frac{1}{\sqrt{N_0/2}} \right) \right]^2 = 4Q \left(\frac{1}{\sqrt{N_0/2}} \right) - 4Q^2 \left(\frac{1}{\sqrt{N_0/2}} \right)$$

Finalizaremos con el cálculo para el símbolo que ejemplifica las regiones de tipo 3, que es a_7 . El procedimiento de cálculo para este caso, que se ilustra en la Figura 4.56, es el siguiente:

- Dirección del espacio q_0

- Media de la distribución gaussiana 1-D: $a_{7,0} = +3$
- Región de decisión : $+2 \leq q_0 < +\infty$

$$P_{a|a_{7,0}} = 1 - P_{e|a_{7,0}} = 1 - Q \left(\frac{1}{\sqrt{N_0/2}} \right)$$

- Dirección del espacio q_1

- Media de la distribución gaussiana 1-D: $a_{7,1} = +1$
- Región de decisión : $0 \leq q_1 < +2$

$$P_{a|a_{7,1}} = 1 - P_{e|a_{7,1}} = 1 - 2Q \left(\frac{1}{\sqrt{N_0/2}} \right)$$

- Probabilidad de error condicional

$$P_{e|a_7} = 1 - \left[1 - Q \left(\frac{1}{\sqrt{N_0/2}} \right) \right] \cdot \left[1 - 2Q \left(\frac{1}{\sqrt{N_0/2}} \right) \right] = 3Q \left(\frac{1}{\sqrt{N_0/2}} \right) - 2Q^2 \left(\frac{1}{\sqrt{N_0/2}} \right)$$

Finalmente, promediando las distribuciones condicionales para los 16 símbolos se obtiene el resultado final que se mostró anteriormente.

Se ha visto cómo se puede calcular la probabilidad de error para regiones de decisión que formen una retícula alineada con los ejes del espacio \mathbf{q} . El problema 2D puede convertirse en dos problemas 1D acoplados que pueden resolverse de forma independiente. Para la constelación de la Figura 4.54 b), las regiones de decisión forman una retícula, pero esta no está alineada con los ejes del espacio de \mathbf{q} , q_0 y q_1 en este ejemplo bidimensional. Sin embargo, es fácil ver que si se aplica un cambio de variables que produzca un giro de 45° , tal y como se hizo en el ejemplo de constelación binaria en espacio bidimensional, las regiones de decisión pasan a estar alineadas con los nuevos ejes y se puede aplicar el mismo procedimiento. En este caso, todas las regiones son del tipo 1, y como la distancia entre símbolos en este caso es $\sqrt{2}$, es fácil comprobar que la probabilidad de error de cada símbolo es la misma, con lo que su valor coincide con la probabilidad media de error y vale

$$P_e = 1 - \left(1 - Q \left(\frac{1}{\sqrt{N_0}} \right) \right)^2 = 2Q \left(\frac{1}{\sqrt{N_0}} \right) - Q^2 \left(\frac{1}{\sqrt{N_0}} \right).$$

En las constelaciones representadas en la Figura 4.54 se pueden calcular las probabilidades de error de forma sencilla debido a que o bien se tienen regiones de decisión que forman una retícula

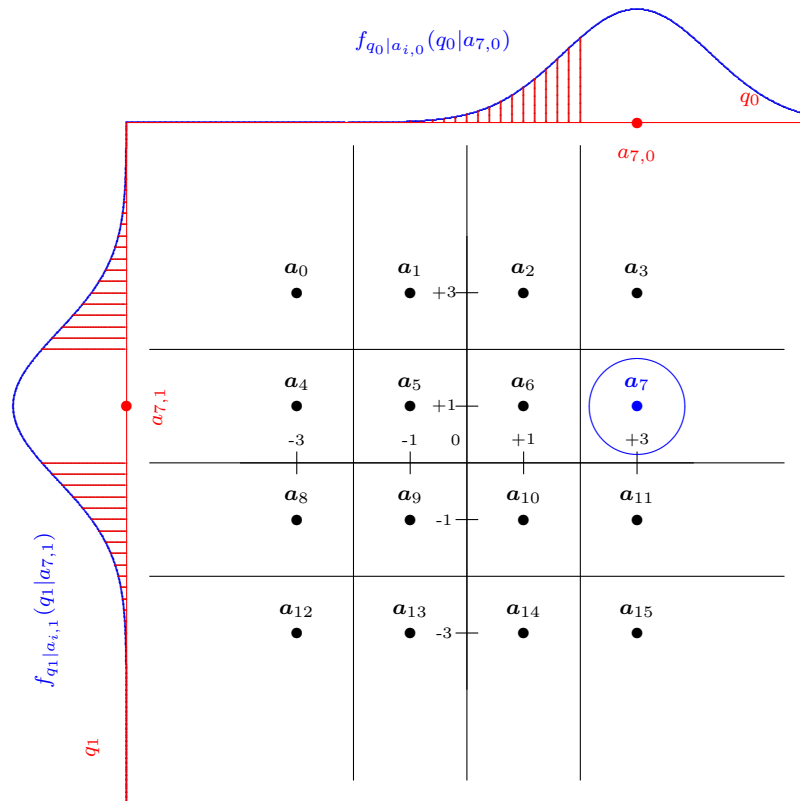


Figura 4.58: Distribuciones condicionales marginales para el símbolo a_7 .

alineada con los ejes del espacio, o bien existe una transformación sencilla que las convierte en ese tipo de regiones. Pero en otros casos esto puede no ser posible, como en el caso de un sistema con la constelación de la Figura 4.59, donde se muestran las ocho regiones de decisión resultantes. Y no resulta posible evaluar de forma analítica la integral de una distribución gaussiana fuera de ese tipo de regiones de decisión. En este caso, para calcular la probabilidad media de error es preciso resolver la integral de la función densidad de probabilidad condicional de la observación fuera de la región de decisión mediante otros procedimientos. En algunos casos será posible calcularla de forma analítica, pero en otros, como el de este ejemplo, no será posible. En casos como este, si se desea calcular la probabilidad de error de forma exacta, habrá que recurrir a cálculos numéricos, con programas como MATLAB que permiten evaluar numéricamente este tipo de resultados. Si se desea utilizar expresiones analíticas, habrá que recurrir a aproximaciones o cotas de la probabilidad de error. A continuación se verán estas últimas herramientas analíticas.

4.5.4. Aproximación y cotas para la probabilidad de error

Las aproximaciones y cotas son útiles cuando no es posible evaluar de forma exacta la probabilidad de error de forma analítica, o cuando se requiere una idea de la magnitud de la probabilidad de error, sin necesidad de conocer el valor exacto.

Aproximación de la probabilidad de error

La probabilidad de error entre símbolos depende de la distancia ente los mismos. Lo más probable, en caso de error, es decidir de forma errónea un símbolo que está a mínima distancia del

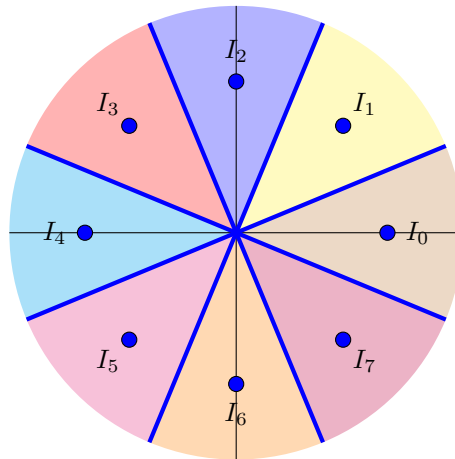


Figura 4.59: Ejemplo de una constelación en un espacio de señales bidimensional en la que no es posible calcular la probabilidad de error de forma analítica.

transmitido, mientras que la probabilidad de que se produzca un error con símbolos más alejados es considerablemente mayor. La aproximación más utilizada asume que sólo se cometerán errores con símbolos que están a mínima distancia, y que todos los símbolos tienen el mismo número de símbolos a mínima distancia, siendo este número el mayor posible para la constelación. Estas consideraciones llevan a la aproximación

$$P_e \approx k \cdot Q \left(\frac{d_{min}}{2\sqrt{N_0/2}} \right),$$

donde los dos parámetros que aparecen en la expresión son:

- d_{min} : mínima distancia entre dos símbolos de la constelación;
- k : máximo número de símbolos a mínima distancia de un símbolo de la constelación.

Estos dos parámetros son muy fáciles de calcular para cualquier constelación. Por ejemplo, para la constelación de 16 símbolos de la Figura 4.54, figura (a), estos parámetros serían

$$d_{min} = 2, k = 4.$$

Para el parámetro k , en la constelación hay símbolos que tienen 2 símbolos a mínima distancia (los de las esquinas), símbolos que tienen 4 símbolos a mínima distancia (los 4 del centro) y otros que tienen a 3 símbolos a mínima distancia (el resto). El valor que hay que incluir en el parámetro es el máximo número de símbolos que hay a mínima distancia de un cierto símbolo, en este caso son 4. Para la constelación de la figura (b), los parámetros serían

$$d_{min} = \sqrt{2}, k = 2.$$

Cota de la unión

La aproximación de la probabilidad de error es útil cuando se quiere tener una idea de la magnitud de la probabilidad de error, y no es relevante que la probabilidad de error total sea

mayor o menor que el valor aproximado. Sin embargo, en ciertas ocasiones es necesario tener una idea aproximada de la probabilidad de error, pero con la certeza de que la probabilidad de error está por debajo del valor especificado (que lo acota). En este caso se recurre a cotas de la probabilidad de error, valores que cumplen que

$$P_e \leq \text{Cota.}$$

Aquí se verán dos cotas: la *cota de la unión*, que acota la probabilidad de error de forma relativamente ajustada, es decir, que la probabilidad de error no es lejana al valor de la cota, especialmente para relaciones señal a ruido altas, aunque su cálculo se complica para constelaciones con muchos símbolos; y la *cota holgada*, una cota con una expresión analítica mucho más sencilla, pero que proporciona valores algo más alejados de la probabilidad de error exacta que la cota de la unión.

Comenzaremos con la cota de la unión. Se trata de una cota que se expresa como la suma de probabilidades de error de sistemas binarios. La idea es que acota la probabilidad condicional de error de un símbolo por la suma de las probabilidades de error de los $M - 1$ sistemas binarios resultantes de utilizar siempre ese símbolo y cada uno de los $M - 1$ símbolos restantes de la constelación, es decir

$$P_{e|\mathbf{a}_i} \leq \sum_{\substack{j=0 \\ j \neq i}}^{M-1} Q \left(\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{2\sqrt{N_0/2}} \right),$$

lo que lleva a acotar la probabilidad de error total como

$$P_e \leq \sum_{i=0}^{M-1} p_{\mathbf{A}}(\mathbf{a}_i) \sum_{\substack{j=0 \\ j \neq i}}^{M-1} Q \left(\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{2\sqrt{N_0/2}} \right).$$

Para ilustrar el procedimiento, se utilizará como ejemplo la constelación de la Figura 4.54 b). En esta constelación, la probabilidad de error cuando el símbolo transmitido es \mathbf{a}_0 , $P_{e|\mathbf{a}_0}$, se obtiene integrando $f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0)$ fuera de la región de decisión I_0 , lo que corresponde al área rayada de la Figura 4.60 a).

En lugar de calcular esta integral, se calcula el error que se obtendría empleando tres decisores binarios para decidir entre el símbolo \mathbf{a}_0 y cada uno de los otros tres símbolos, es decir, considerar los casos binarios

- \mathbf{a}_0 y \mathbf{a}_1 ;
- \mathbf{a}_0 y \mathbf{a}_2 ;
- \mathbf{a}_0 y \mathbf{a}_3 ;

Esto supone integrar la función densidad de probabilidad en cada una de las tres regiones mostradas en las Figuras 4.60 b), c) y d), respectivamente. Comenzando con el símbolo \mathbf{a}_1 , denotamos la probabilidad de error del decisor binario entre \mathbf{a}_0 y \mathbf{a}_1 como $P_e(\mathbf{a}_0, \mathbf{a}_1)$, y se obtiene integrando el área rayada en la Figura 4.60 b). Para un decisor binario

$$P_e(\mathbf{a}_0, \mathbf{a}_1) = Q \left(\frac{d(\mathbf{a}_0, \mathbf{a}_1)}{2\sqrt{N_0/2}} \right) = Q \left(\frac{\sqrt{2}}{2\sqrt{N_0/2}} \right) = Q \left(\frac{1}{\sqrt{N_0}} \right).$$

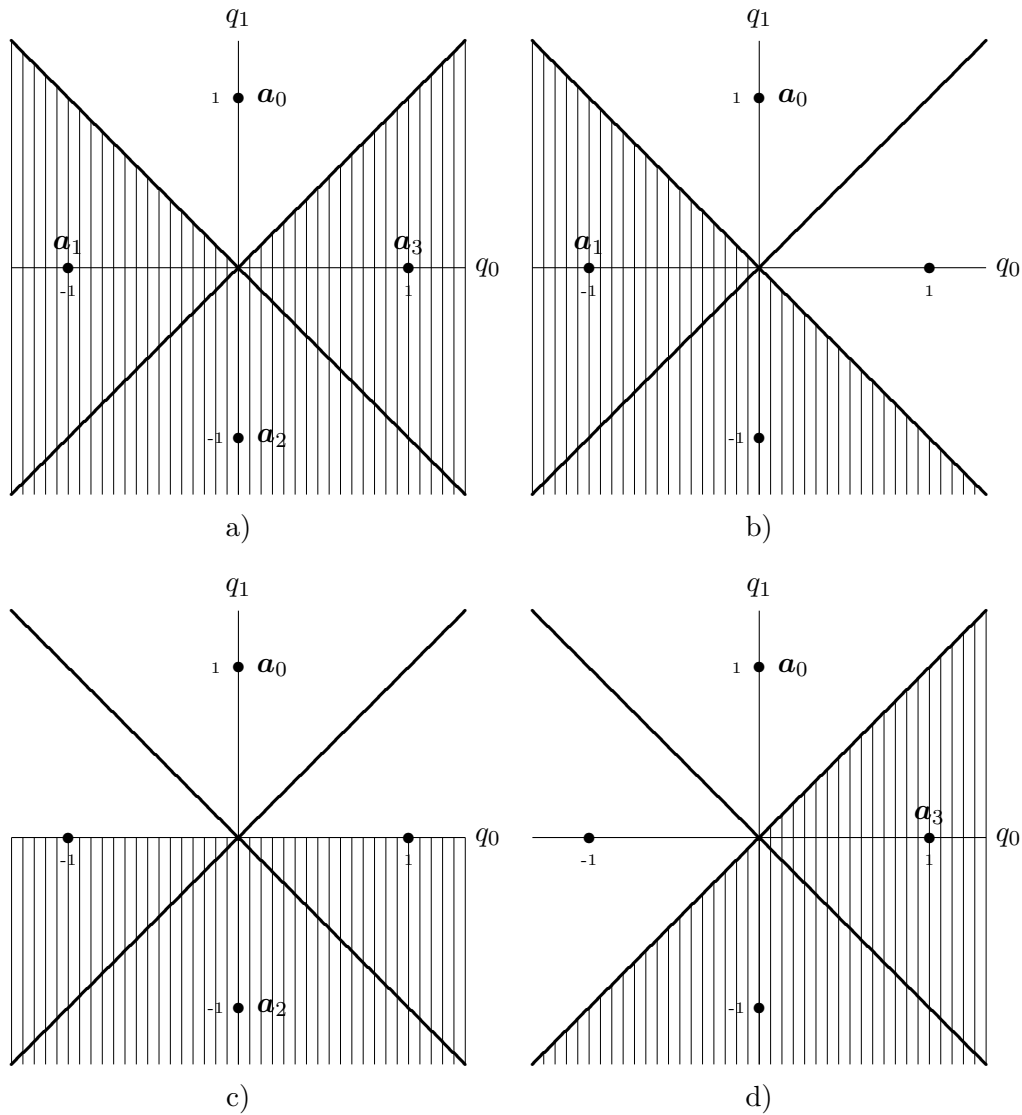


Figura 4.60: Cota de la unión.

Procediendo de igual modo con los otros dos símbolos

$$P_e(\mathbf{a}_0, \mathbf{a}_2) = Q\left(\frac{d(\mathbf{a}_0, \mathbf{a}_2)}{2\sqrt{N_0/2}}\right) = Q\left(\frac{2}{2\sqrt{N_0/2}}\right) = Q\left(\frac{\sqrt{2}}{\sqrt{N_0}}\right).$$

$$P_e(\mathbf{a}_0, \mathbf{a}_3) = Q\left(\frac{d(\mathbf{a}_0, \mathbf{a}_3)}{2\sqrt{N_0/2}}\right) = Q\left(\frac{\sqrt{2}}{2\sqrt{N_0/2}}\right) = Q\left(\frac{1}{\sqrt{N_0}}\right).$$

A partir de estas probabilidades de error con decisores binarios se define la cota de la unión para

$P_{e|a_0}$ como

$$\begin{aligned} P_{e|a_0} &\leq \sum_{j=1}^{M-1} P_e(\mathbf{a}_0, \mathbf{a}_j) \\ &= P_e(\mathbf{a}_0, \mathbf{a}_1) + P_e(\mathbf{a}_0, \mathbf{a}_2) + P_e(\mathbf{a}_0, \mathbf{a}_3) \\ &= 2Q\left(\frac{1}{\sqrt{N_0}}\right) + Q\left(\frac{2}{2\sqrt{N_0/2}}\right). \end{aligned}$$

Procediendo de igual modo con el resto de símbolos de la constelación obtenemos la denominada **cota de la unión para la probabilidad media de error**, que para una constelación genérica de M símbolos se escribe como

$$P_e \leq \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \sum_{\substack{j=0 \\ j \neq i}}^{M-1} P_e(\mathbf{a}_i, \mathbf{a}_j) = \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \sum_{\substack{j=0 \\ j \neq i}}^{M-1} Q\left(\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{2\sqrt{N_0/2}}\right)$$

Para constelaciones con símbolos equiprobables

$$P_e \leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ j \neq i}}^{M-1} Q\left(\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{2\sqrt{N_0/2}}\right)$$

Particularizando para el ejemplo que nos ocupa, dada la simetría de la constelación, da como resultado

$$P_e \leq \frac{1}{4} \left(2Q\left(\frac{1}{\sqrt{N_0}}\right) + Q\left(\frac{2}{2\sqrt{N_0/2}}\right) \right)$$

Para comprender por qué la cota de la unión es una cota superior hay que tener en cuenta que $P_{e|a_0}$ se obtiene integrando $f_{a|A}$ en la región sombreada de la Figura 4.60 a), esto es, fuera de I_0 . En lugar de proceder a calcular esa integral, se han sumado tres integrales de la misma función sobre las regiones mostradas por las Figuras 4.60 b), c) y d). La superposición de estas áreas comprende todo el espacio fuera de I_0 , con lo que se asegura que como mínimo se obtiene esa probabilidad de error buscada. Sin embargo, en algunas regiones se está integrando varias veces. Este aspecto se muestra en la Figura 4.61. En las zonas etiquetadas como I sólo se ha integrado una vez, en las zonas etiquetadas como II se ha integrado dos veces y en la zona etiquetada como III se ha integrado tres veces.

Se ha comprobado que la cota de la unión es una cota superior de la probabilidad de error, pero no sabemos si la cota está cerca o lejos del valor verdadero de probabilidad media de error. Para la constelación anterior se han calculado tanto la probabilidad exacta como la cota de la unión. La Figura 4.62 compara ambos valores como una función de la relación E_s/N_0 . Se puede ver que la cota de la unión proporciona un valor superior a la probabilidad de error exacta que además es bastante ajustado al valor exacto, especialmente para valores altos de relación señal a ruido (parametrizada por E_s/N_0).

Cota holgada

En constelaciones con un número de símbolos elevado, la evaluación de la cota de error puede ser engorrosa debido al número de términos a evaluar, que crece aproximadamente con el cuadrado

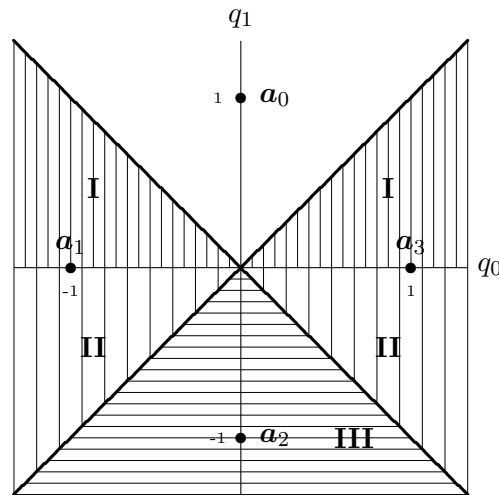


Figura 4.61: Cota de la unión: Cota superior

del número de símbolos de la constelación. Para una constelación de 4 símbolos hay que evaluar 6 distancias (aprovechando que $d(\mathbf{a}_i, \mathbf{a}_j) = d(\mathbf{a}_j, \mathbf{a}_i)$ y hacer 6 evaluaciones de la función $Q(x)$. En general, el número de distancias mínimo a considerar en la expresión, la simetría $d(\mathbf{a}_i, \mathbf{a}_j) = d(\mathbf{a}_j, \mathbf{a}_i)$, es

$$N_{\text{distancias}} = \sum_{k=1}^{M-1} k$$

Esto, para algunos ejemplos numéricos supone

- $M = 4$ significa $N_{\text{distancias}} = 6$.
- $M = 8$ significa $N_{\text{distancias}} = 28$.
- $M = 16$ significa $N_{\text{distancias}} = 120$.
- $M = 64$ significa $N_{\text{distancias}} = 2016$.

Para constelaciones con un número elevado de símbolos, el cálculo analítico sin utilizar un computador es relativamente costoso, por lo que cabe buscar una cota más compacta analíticamente. Una cota con una expresión analítica compacta y que es comúnmente utilizada es la denominada *cota holgada* (del inglés “*looser bound*”). Esta aproximación utiliza como base la cota de la unión, pero asume que todos los símbolos están a una distancia d_{\min} del resto, con lo que es una aproximación pesimista que para constelaciones con un número grande de símbolos puede estar muy alejada de la probabilidad de error exacta. Con esta asunción, la expresión de la probabilidad de error queda acotada por

$$P_e \leq (M - 1)Q \left(\frac{d_{\min}}{2\sqrt{N_0/2}} \right).$$

La Figura 4.63 compara la probabilidad de error exacta con la cota holgada para la constelación de cuatro símbolos de la Figura 4.54 (b), como una función de la relación E_s/N_0 . Se puede ver que la cota de la unión proporciona un valor superior a la probabilidad de error exacta que ahora es menos ajustado al valor exacto que en el caso de la cota de la unión.

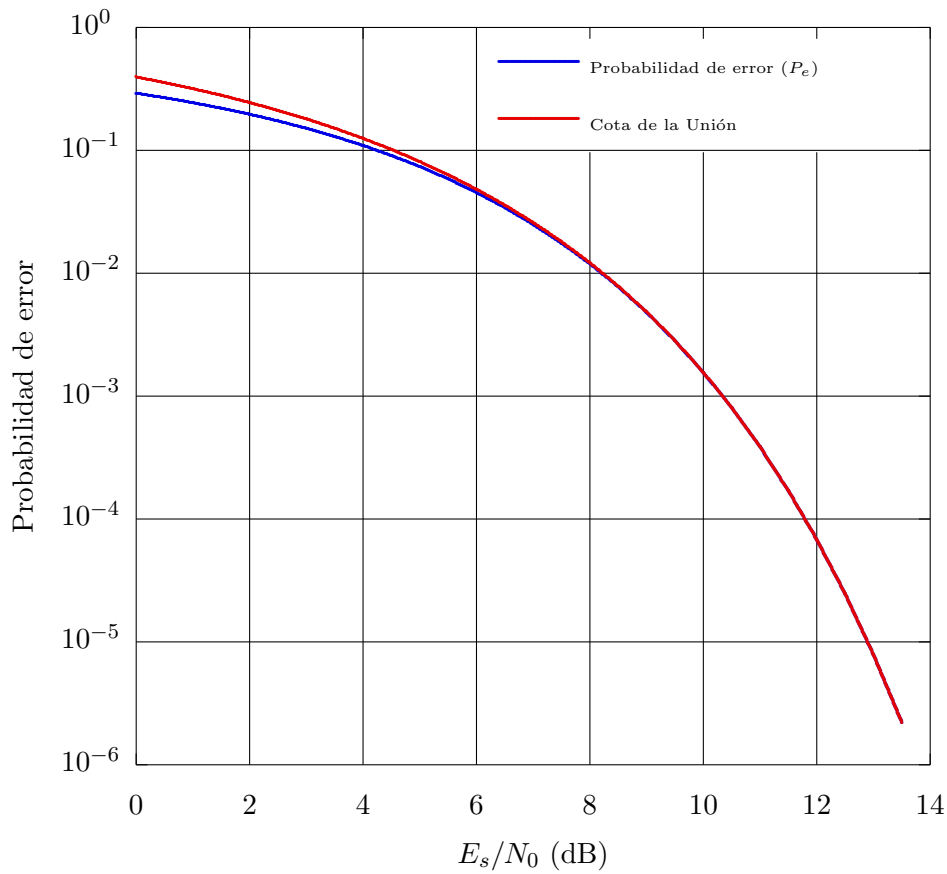


Figura 4.62: Probabilidades de error exacta y cota de la unión para la constelación de la Figura 4.54 (b), como una función de la relación E_s/N_0 .

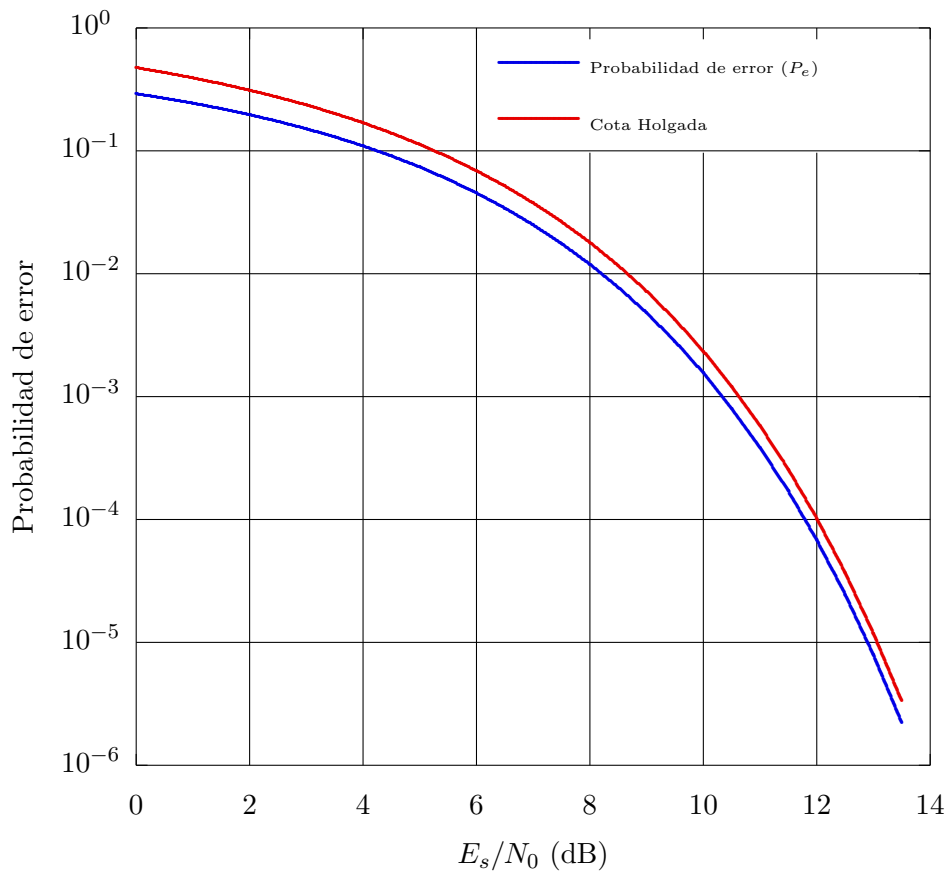


Figura 4.63: Probabilidades de error exacta y cota de la holgada para la constelación de la Figura 4.54 (b), como una función de la relación E_s/N_0 .

4.5.5. Expresiones de la probabilidad de error en función de E_s/N_0

En muchas ocasiones, como ocurre en las Figuras 4.62 o 4.63, resulta interesante expresar la probabilidad de error (bien el valor exacto, una aproximación o una cota), como una función de la relación señal a ruido del sistema, parametrizada por la relación entre energía media por símbolo de la señal transmitida, y densidad espectral de potencia del ruido, i.e. E_s/N_0 . Obtener expresiones de la probabilidad de error en función de esta relación es muy sencillo en la mayoría de los casos. Los términos que aparecen habitualmente en las probabilidades de error se suelen escribir como

$$Q\left(\frac{d_{min}}{2\sqrt{N_0/2}}\right) \text{ ó } Q\left(\frac{A}{\sqrt{N_0/2}}\right).$$

En cualquier caso, multiplicando y dividiendo por $\sqrt{E_s}$, se pueden reescribir como

$$Q\left(v \cdot \sqrt{\frac{E_s}{N_0}}\right), \text{ donde } Q\left(\underbrace{\frac{d_{min}}{\sqrt{2}\sqrt{E_s}}}_v \sqrt{\frac{E_s}{N_0}}\right) \text{ ó } Q\left(\underbrace{\frac{A\sqrt{2}}{\sqrt{E_s}}}_v \sqrt{\frac{E_s}{N_0}}\right)$$

El factor v es un valor constante, que depende de la constelación y que puede evaluarse de forma sencilla: basta con conocer el valor numérico de E_s y sustituir en la expresión anterior. Este factor se puede ver como una medida de eficiencia de la constelación, ya que cuanto más alto es v , menor es el valor de la función $Q(x)$ para ese argumento, y más eficiente es la constelación. A continuación se introduce un ejemplo en que se calcula la probabilidad de error en función de E_s/N_0 para dos constelaciones binarias diferentes.

Ejemplo

Se ha visto que para cualquier dimensión N , cualquier sistema binario con símbolos equiprobables que transmita sobre un canal gaussiano tiene una probabilidad de error

$$P_e = Q\left(\frac{d(\mathbf{a}_0, \mathbf{a}_1)}{2\sqrt{N_0/2}}\right)$$

Se compararán dos casos

- Caso (a): Constelación binaria simétrica ($N = 1$)

$$\mathbf{a}_0 = +A, \mathbf{a}_1 = -A$$

- Caso (b): Constelación ortogonal ($N = 2$)

$$\mathbf{a}_0 = \begin{bmatrix} A \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} 0 \\ A \end{bmatrix}$$

Sólo hay que calcular las distancias y energías medias por símbolo para cada constelación

- Caso (a): $E_s = A^2$, $d(\mathbf{a}_0, \mathbf{a}_1) = 2 \cdot A$
- Caso (b): $E_s = A^2$, $d(\mathbf{a}_0, \mathbf{a}_1) = \sqrt{2} \cdot A$

A partir de estas expresiones, el factor de eficiencia v se calcula como

- Caso (a): Constelación binaria simétrica

$$v = \frac{d(\mathbf{a}_0, \mathbf{a}_1)}{\sqrt{2}\sqrt{E_s}} = \sqrt{2} \rightarrow P_e = Q\left(\sqrt{2\frac{E_s}{N_0}}\right)$$

- Caso (b): Constelación ortogonal (se extiende para $N > 2$)

$$v = \frac{d(\mathbf{a}_0, \mathbf{a}_1)}{\sqrt{2}\sqrt{E_s}} = 1 \rightarrow P_e = Q\left(\sqrt{\frac{E_s}{N_0}}\right)$$

Se puede ver que la constelación binaria simétrica es más eficiente. Para un valor de E_s dado, tiene mejores prestaciones que la constelación ortogonal.

4.6. Codificador

Hasta ahora, la descripción que se ha hecho del codificador se reduce a decir que realiza una transformación de una secuencia de símbolos, $B[n]$, a una representación vectorial de las señales que estarán asociadas a cada símbolo de la secuencia, $\mathbf{A}[n]$. Esta función se ilustra en la Figura 4.64. Cada uno de los símbolos estará formado por un bloque de m bits.



Figura 4.64: Codificador en el transmisor de un sistema de comunicaciones digital.

Es ahora cuando, una vez analizado el receptor y a la vista de las expresiones obtenidas en el cálculo de las prestaciones que se pueden obtener con un sistema de comunicaciones, se establecerán las condiciones que debe cumplir el codificador para conseguir una comunicación de la forma más eficiente que sea posible.

4.6.1. Diseño del codificador

El diseño del codificador consiste en dos partes:

1. Diseño o elección de la constelación a transmitir.
2. Asignación binaria para cada uno de los símbolos.

La primera parte consiste en elegir la constelación de M puntos en un espacio de dimensión N que define la representación vectorial de las M señales que se utilizan en el transmisor para transportar cada uno de los $M = 2^m$ valores del alfabeto de símbolos (bloques de m bits). De modo que cada posible combinación de m bits, símbolo b_i , tendrá asociada una señal $s_i(t)$, y el codificador lo que hace es definir la representación vectorial de la señal, \mathbf{a}_i . Luego, el modulador convertirá esta representación vectorial en una señal continua en el tiempo $s_i(t)$. Es importante recordar que la representación vectorial de un conjunto de M señales determina dos factores importantes

1. La energía de cada señal, y por tanto, la energía media de las M señales.
2. La distancia entre cada par de señales, que se relaciona con la energía de la señal diferencia, y que hemos visto que determina las prestaciones del sistema.

Por tanto, la elección de la constelación determina estos dos factores: energía y prestaciones. Y de hecho, el diseño de la constelación se hará en base a estos dos factores: se buscará el mejor compromiso entre energía y prestaciones.

La segunda parte del diseño del codificador es la asignación binaria. Esta asignación consiste en adjudicar a cada uno de los posibles valores del alfabeto de \mathbf{A} (o equivalentemente de $B[n]$, ya que hay una asignación biunívoca $b_i \leftrightarrow \mathbf{a}_i$) una de las M posibles combinaciones de m bits. Como veremos en su momento, la asignación determina la probabilidad de error a nivel de bit, es decir, las prestaciones del sistema. Por tanto, los criterios que han de guiar el diseño del codificador son

1. Prestaciones (probabilidades de error de símbolo y de bit).
2. Energía.

A continuación se analizará cómo se realiza el diseño de la constelación para el caso de transmisión de símbolos equiprobables cuando el ruido es gaussiano, que es el caso más frecuente en sistemas de comunicaciones.

4.6.2. Diseño de la constelación

El diseño de la constelación depende de dos factores: prestaciones y energía. Y las prestaciones están relacionadas con la distancia entre símbolos, dependiendo fundamentalmente de la mínima distancia entre dos símbolos de la constelación. Recuerde la aproximación de la probabilidad de error

$$P_e \approx k \cdot Q\left(\frac{d_{min}}{2\sqrt{N_0/2}}\right).$$

Así que dependiendo del tipo de limitaciones que a priori se tengan en el sistema, el problema de diseño se puede plantear de dos formas. Si existe una limitación en energía, para conseguir la menor probabilidad de error posible, el codificador ha de generar una constelación que tenga la mayor distancia entre símbolos, pero con la limitación que ponga el valor máximo de energía de símbolo o de energía media por símbolo, E_s que sea admisible teniendo en cuenta las limitaciones en energía que tenga el sistema. Si lo que se limita es la probabilidad de error máxima admisible, habrá que buscar una constelación que la consiga utilizando la menor cantidad de energía, de modo que se minimize el coste del sistema transmisor.

En cualquier caso, los dos factores que determinan la elección de la constelación van en sentidos contrapuestos: si se aumenta la distancia entre símbolos de una constelación para disminuir la probabilidad de error, en general aumentará la energía necesaria para transmitir cada símbolo. Es pues necesario encontrar un compromiso entre ambos requerimientos. Conviene también recordar que la energía de una señal, en cuanto a la representación vectorial de la misma, está relacionada con la distancia que tiene respecto del origen dicha representación vectorial, ya que

$$\mathcal{E}\{\mathbf{a}_i\} = \|\mathbf{a}_i\|^2 = \sum_{k=0}^{N-1} |a_{i,k}|^2,$$

es decir, que la energía de una señal es la distancia al cuadrado de su representación vectorial al origen de coordenadas. De este modo, se puede plantear el problema del diseño óptimo de un codificador como el problema de colocar M puntos en un espacio de dimensión N de forma que tengan la mayor separación mínima entre ellos que sea posible, al mismo tiempo que estén todos lo más cerca posible del origen de coordenadas, buscando un compromiso entre las dos cosas.

Relacionada con la idea de colocar los puntos lo más cerca posible del origen, está la siguiente propiedad, que determina una de las condiciones que debe cumplir un codificador óptimo desde el punto de vista de compromiso prestaciones/energía: Para unas distancias entre símbolos dadas, se minimiza la energía media por símbolo cuando la media de la constelación es nula

$$E[\mathbf{a}_i] = \begin{bmatrix} E[a_{i,0}] \\ E[a_{i,1}] \\ \vdots \\ E[a_{i,N-1}] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}$$

Esta propiedad se ilustrará con un sencillo ejemplo de sistema binario en un espacio unidimensional, que es fácilmente extensible a cualquier otra situación. Se consideran dos símbolos en un espacio 1D, en concreto los símbolos

$$\mathbf{a}_0 = B - A, \quad \mathbf{a}_1 = B + A$$

que se representan en la Figura 4.65

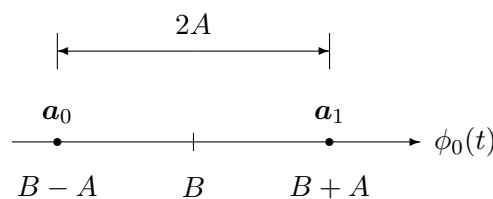


Figura 4.65: Ejemplo para un caso binario y unidimensional de la relación de la energía de una constelación con su media y con la distancia entre símbolos de la constelación.

Los dos parámetros de los símbolos, A y B , parametrizan por separado la distancia entre símbolos y la media de los mismos:

- Media: B
- Distancia entre símbolos: $2A$

Si se calcula la energía media por símbolo (asumiendo símbolos equiprobables) para esta constelación, se tiene

$$\begin{aligned} E_s &= \frac{1}{2} \cdot \mathcal{E}\{\mathbf{a}_0\} + \frac{1}{2} \cdot \mathcal{E}\{\mathbf{a}_1\} = \frac{1}{2} \cdot (B - A)^2 + \frac{1}{2} \cdot (B + A)^2 \\ &= \frac{1}{2} \cdot (B^2 + A^2 - 2AB) + \frac{1}{2} \cdot (B^2 + A^2 + 2AB) = B^2 + A^2. \end{aligned}$$

Se puede ver que aparecen dos términos independientes, uno relacionado con la media de la constelación, y otro con la distancia entre símbolos.

- Contribución de la media: B^2
- Contribución de la distancia entre símbolos: A^2

Por tanto, para cualquier distancia entre símbolos, la mínima energía por símbolo se obtiene cuando la media es nula

$$\text{Media nula } (B = 0) \rightarrow E_s = A^2.$$

Resulta trivial extender este desarrollo para espacios de mayor dimensión o para un mayor número de símbolos. Así que un diseño óptimo siempre contará con esta característica: la media de los símbolos de la constelación será nula (siendo la media de dimensión N).

Técnica de empaquetado de esferas

Teniendo en cuenta todas estas consideraciones, el problema de diseñar una constelación óptima, entendiéndolo por óptima el que tenga el mejor compromiso posible entre prestaciones y consumo de energía, se puede plantear como la búsqueda de constelaciones que para una mínima distancia entre símbolos dada, requieran la mínima energía, que implica que los símbolos estén lo más cerca posible del origen (teniendo una media nula). Atendiendo a esta descripción, el problema de diseño de la constelación se puede resolver mediante la denominada *técnica de empaquetado de esferas*: modelar cada símbolo como una esfera de diámetro d_{min} , de modo que dos esferas que se toquen estén a distancia d_{min} , como se muestra en la Figura 4.66, y buscar como empaquetarlas en un espacio N -dimensional ocupando el menor volumen posible.

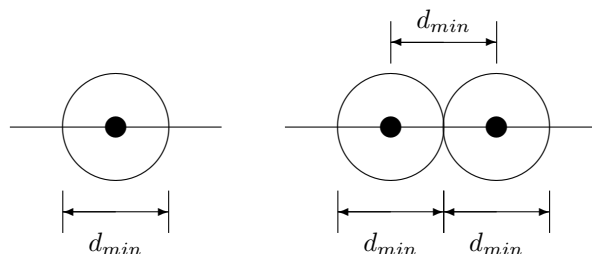


Figura 4.66: Modelo de un símbolo en la técnica de empaquetado de esferas.

Técnica de empaquetado de esferas - Espacio 1D

En un espacio unidimensional, la aplicación de esta metodología es muy simple. Se busca colocar M símbolos sobre una recta, separados como mínimo d_{min} , y que estén lo más cercanos que sea posible del origen. Esto se consigue colocando los puntos equiespaciados, centrados en el origen, con la mitad a cada lado, para conseguir que la media sea nula, con lo que las coordenadas de los M símbolos serán

$$\mathbf{a}_i \in \left\{ \pm \frac{d_{min}}{2}, \pm 3 \cdot \frac{d_{min}}{2}, \dots, \pm (M - 1) \cdot \frac{d_{min}}{2} \right\},$$

es decir, símbolos con coordenadas \pm números impares por la mitad de la distancia mínima entre símbolos. La Figura 4.67 muestra un ejemplo para $M = 4$.

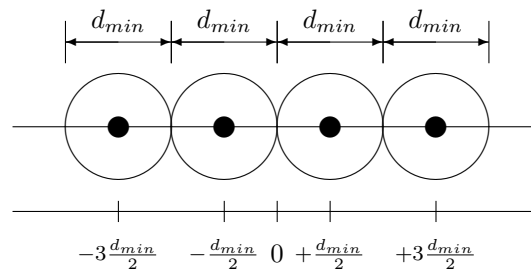


Figura 4.67: Ejemplo de aplicación de la técnica de empaquetado de esferas para un espacio 1D.

Técnica de empaquetado de esferas - Espacio 2D

En un espacio 2D, la mejor forma de empaquetar esferas, tal y como se ilustra en la Figura 4.68, es siguiendo una configuración hexagonal, más eficiente que una configuración en una retícula.

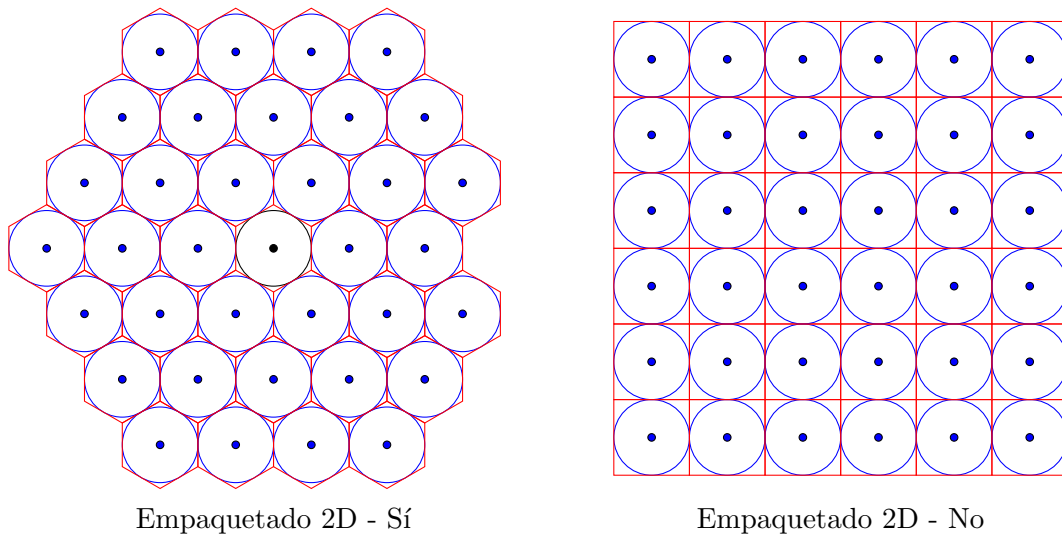


Figura 4.68: Ejemplo de aplicación de la técnica de empaquetado de esferas para un espacio 2D, en la que se compara la forma más eficiente de colocar las esferas, utilizando una configuración hexagonal, frente a un modo menos eficiente en una configuración reticular (se compara el empaquetado de 36 esferas con ambas configuraciones).

La opción de colocar una esfera en el centro, y colocar un círculo de esferas tocándola a su alrededor, provoca una configuración hexagonal mucho más eficiente que una configuración rectangular con los puntos formando una retícula. Si hay más símbolos, se colocará un segundo anillo sobre el primero, teniendo también una configuración hexagonal, y se seguirá así hasta completar M esferas. Como M es una potencia de 2, es posible que el último anillo no se complete, como ocurre en el ejemplo de la constelación de 32 símbolos ($m = 5$ bits por símbolo) de la Figura 4.69. En ese caso, si la esfera inicial se mantiene en el origen, la media de la constelación no será nula, por lo que habrá que “recolocar la constelación” para que tenga media nula, como se ha hecho ya para la constelación de la figura.

La forma en que ha de modificarse una constelación con media no nula para que manteniendo la relación de distancias entre símbolos (lo que geoméricamente equivale a hacer un desplazamiento sin escalado) pase a tener media nula es restando a cada símbolo de la constelación la media de

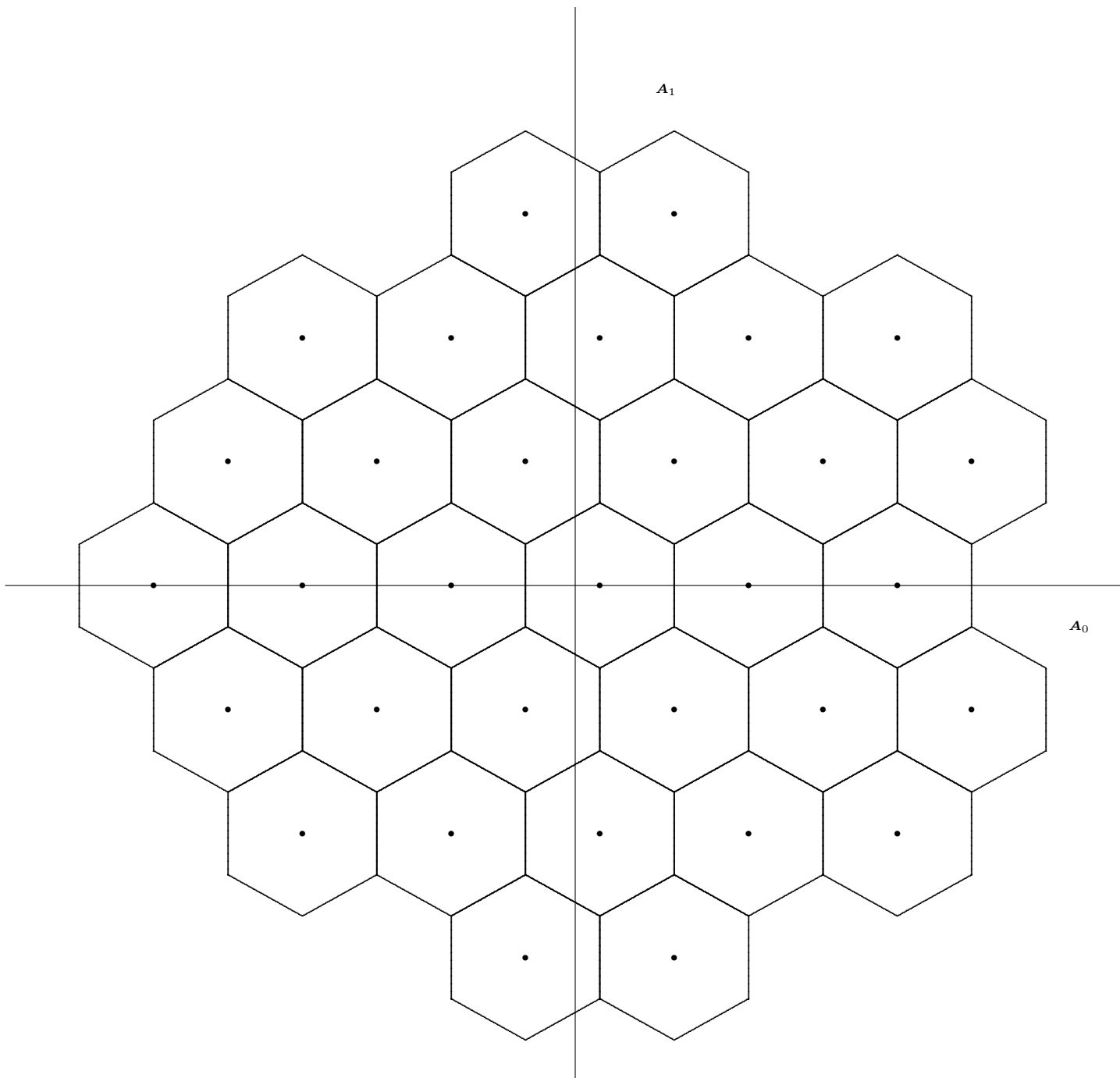


Figura 4.69: Ejemplo de aplicación de la técnica de empaquetado de esferas para un espacio 2D, para una constelación de 32 puntos, en la que se han sustituido las esferas por hexágonos para resaltar la estructura hexagonal resultante.

la misma, es decir, hay que calcular la media de la constelación

$$E[\mathbf{a}_i] = \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \cdot \mathbf{a}_i.$$

y la nueva constelación modificada para tener media nula tendrá símbolos con coordenadas

$$\mathbf{a}'_i = \mathbf{a}_i - E[\mathbf{a}_i]$$

de forma que la media de la nueva constelación es

$$E[\mathbf{a}'_i] = \sum_{i=0}^{M-1} \underbrace{p_{A'}(\mathbf{a}'_i)}_{p_A(\mathbf{a}_i)} \cdot \mathbf{a}'_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}$$

A continuación se muestra un ejemplo de como modificar una constelación para, manteniendo la relación de distancias entre símbolos, la media de la nueva constelación sea nula, y por tanto, la energía media por símbolo sea mínima.

Ejemplo

En concreto, se considera una constelación con $M = 4$ símbolos, $p_A(\mathbf{a}_i) = \frac{1}{4}$, $\forall i$, con coordenadas

$$\mathbf{a}_0 = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} +\frac{3}{2} \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -\frac{1}{2} \\ +2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} +\frac{3}{2} \\ +2 \end{bmatrix}$$

La constelación se representa en la Figura 4.70 muestra esta constelación.

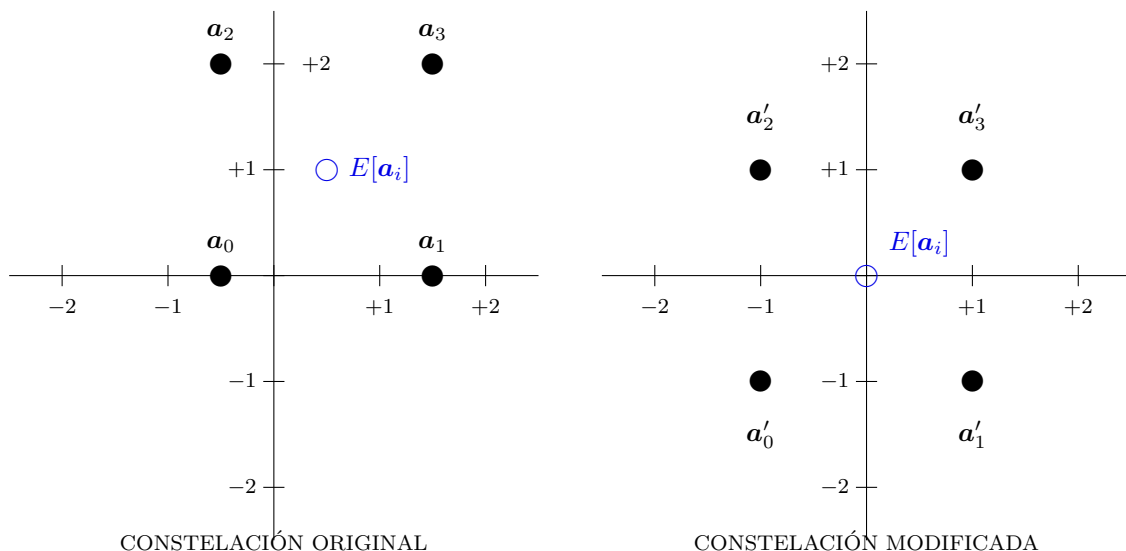


Figura 4.70: Ejemplo de modificación de una constelación para minimizar la energía media por símbolo manteniendo las prestaciones (distancias relativas entre símbolos).

La energía media por símbolo de la constelación en este caso es

$$E_s = \frac{1}{4} \cdot \left[\left(-\frac{1}{2} \right)^2 + 0^2 \right] + \frac{1}{4} \cdot \left[\left(\frac{3}{2} \right)^2 + 0^2 \right] + \frac{1}{4} \cdot \left[\left(-\frac{1}{2} \right)^2 + 2^2 \right] + \frac{1}{4} \cdot \left[\left(\frac{3}{2} \right)^2 + 2^2 \right] = \frac{13}{4} = 3,25$$

Para minimizar la energía media por símbolo, primero se calcula la media de la constelación, que en este caso vale

$$E[\mathbf{a}_i] = \frac{1}{4} \cdot \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix} + \frac{1}{4} \cdot \begin{bmatrix} +\frac{3}{2} \\ 0 \end{bmatrix} + \frac{1}{4} \cdot \begin{bmatrix} -\frac{1}{2} \\ +2 \end{bmatrix} + \frac{1}{4} \cdot \begin{bmatrix} +\frac{3}{2} \\ +2 \end{bmatrix} = \begin{bmatrix} +\frac{1}{2} \\ +1 \end{bmatrix}.$$

Esta media se ilustra con un círculo en la figura. Una vez calculada la media de la constelación, a cada símbolo se le resta dicha media, de modo que se obtiene la siguiente constelación modificada

$$\mathbf{a}'_0 = \mathbf{a}_0 - E[\mathbf{a}_i] = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mathbf{a}'_1 = \mathbf{a}_1 - E[\mathbf{a}_i] = \begin{bmatrix} +1 \\ -1 \end{bmatrix}, \mathbf{a}'_2 = \begin{bmatrix} -1 \\ +1 \end{bmatrix}, \mathbf{a}'_3 = \begin{bmatrix} +1 \\ +1 \end{bmatrix},$$

que se representa también en la Figura 4.70, y que tiene media nula. Si se calcula ahora la energía media por símbolo de la constelación modificada, se tiene

$$E'_s = \frac{1}{4} \cdot [(-1)^2 + (-1)^2] + \frac{1}{4} \cdot [(+1)^2 + (-1)^2] + \frac{1}{4} \cdot [(-1)^2 + (+1)^2] + \frac{1}{4} \cdot [(+1)^2 + (+1)^2] = 2.$$

Se puede ver que es menor que en la constelación original, tal y como era de esperar.

4.6.3. Constelaciones utilizadas en sistemas de comunicaciones

Aunque esta es la forma óptima de las constelaciones desde el punto de vista del mejor compromiso entre prestaciones y consumo de energía, existen otros factores a tener en cuenta a la hora de escoger una constelación que hacen que no siempre se elija esta estrategia. De hecho, las constelaciones hexagonales no son las constelaciones que se usan con más frecuencia en los sistemas de comunicaciones prácticos. Entre estos factores que llevan a elegir otro tipo de constelaciones destacan los siguientes:

Facilidad de realización del transmisor Las coordenadas de los distintos símbolos deben poder expresarse con números de precisión finita (no irracionales) pues es necesario almacenarlos en el hardware o software del transmisor. En un empaquetamiento hexagonal aparecen coordenadas de símbolos del tipo irracional. Si se trunca el valor de las coordenadas la posición de los símbolos se modifica, teniendo un cambio en las prestaciones.

Relación energía de pico/energía media Existen tecnologías de amplificación que funcionan bien cuando la energía de las señales no varía entre intervalos de símbolo, pero que generan serias distorsiones cuando hay cambios bruscos en la energía de un intervalo a otro. Constelaciones como la de la Figura 4.69 tienen símbolos con energías relativamente diferentes (muy baja para los símbolos próximos al origen, y varios órdenes de magnitud mayor para los símbolos más alejados). Constelaciones en las que la relación entre la energía de pico (energía máxima de un símbolo), y la energía media por símbolo es alta, no son adecuadas para su uso con este tipo de amplificadores.

Facilidad de realización del receptor Existen aplicaciones en las que el coste del receptor es determinante, hasta tal punto que es preferible sacrificar una menor probabilidad de error en recepción para conseguir una simplificación de la circuitería necesaria en el receptor. Dependiendo de la tecnología empleada, este puede ser un factor determinante para utilizar constelaciones que aunque no tengan el mejor compromiso prestaciones/energía, permitan implementaciones más económicas para el receptor. Un ejemplo de problema que puede aparecer en algunas tecnologías es la recuperación del signo de las coordenadas del símbolo

recibido (implementar la función signo en un ASIC es relativamente complicado) y esto fuerza a utilizar constelaciones donde las coordenadas no sean negativas, denominadas comúnmente *constelaciones unipolares*. Otro ejemplo es la implementación del decisor. En constelaciones hexagonales, las regiones de decisión son hexagonales, y determinar si un punto está dentro de una determinada región de decisión, implica comparar dicho punto con 6 rectas y comprobar que está simultáneamente en el lado correcto de las seis rectas. Constelaciones con regiones de decisión más simples, pueden ser en algunos casos más convenientes.

Por razones de orden práctico, especialmente las mencionadas con anterioridad, las constelaciones utilizadas con mayor frecuencia son las que se citan a continuación.

Constelaciones QAM

Estas constelaciones, cuyo nombre viene del acrónimo inglés para “*Quadrature Amplitude Modulation*”, tienen como principal característica que los puntos de la constelación forman una retícula sobre un espacio 2D, tal y como los ejemplos que se muestran en la Figura 4.71. Para ello, tienen la particularidad de que el número de bits por símbolo m es par.

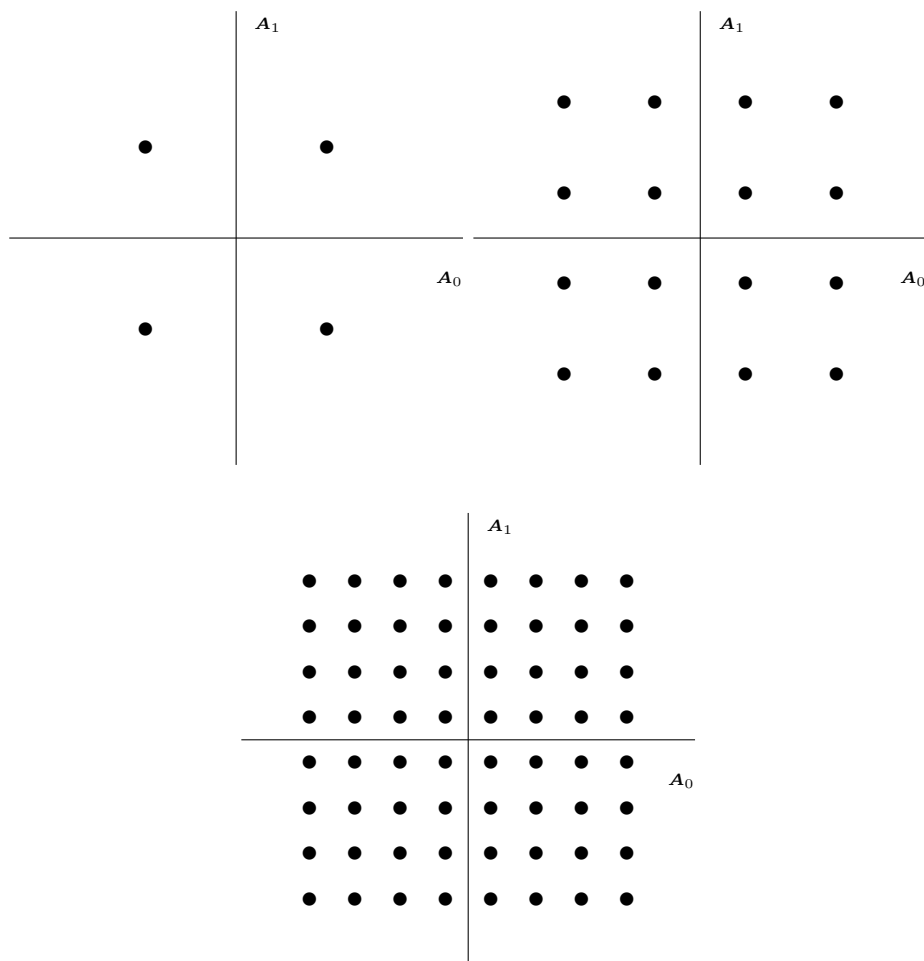


Figura 4.71: Ejemplo de constelaciones QAM de 4, 16 y 64 símbolos.

Este tipo de constelaciones tienen la ventaja de que la implementación de transmisor y receptor es relativamente simple. Por ejemplo, los valores de coordenadas toman los mismos valores en

cada dirección del espacio, y las regiones de decisión de cada símbolo se pueden expresar como condiciones independientes para cada una de las direcciones del espacio, por lo que el procesamiento se puede hacer de forma independiente para cada dirección del espacio tanto en el transmisor como en el receptor.

Constelaciones PSK

Estas constelaciones, cuyo nombre viene del acrónimo inglés para “*Phase Shift Keying*”, tienen como principal característica que los puntos de la constelación forman una circunferencia sobre un espacio 2D, tal y como los ejemplos que se muestran en la Figura 4.72.

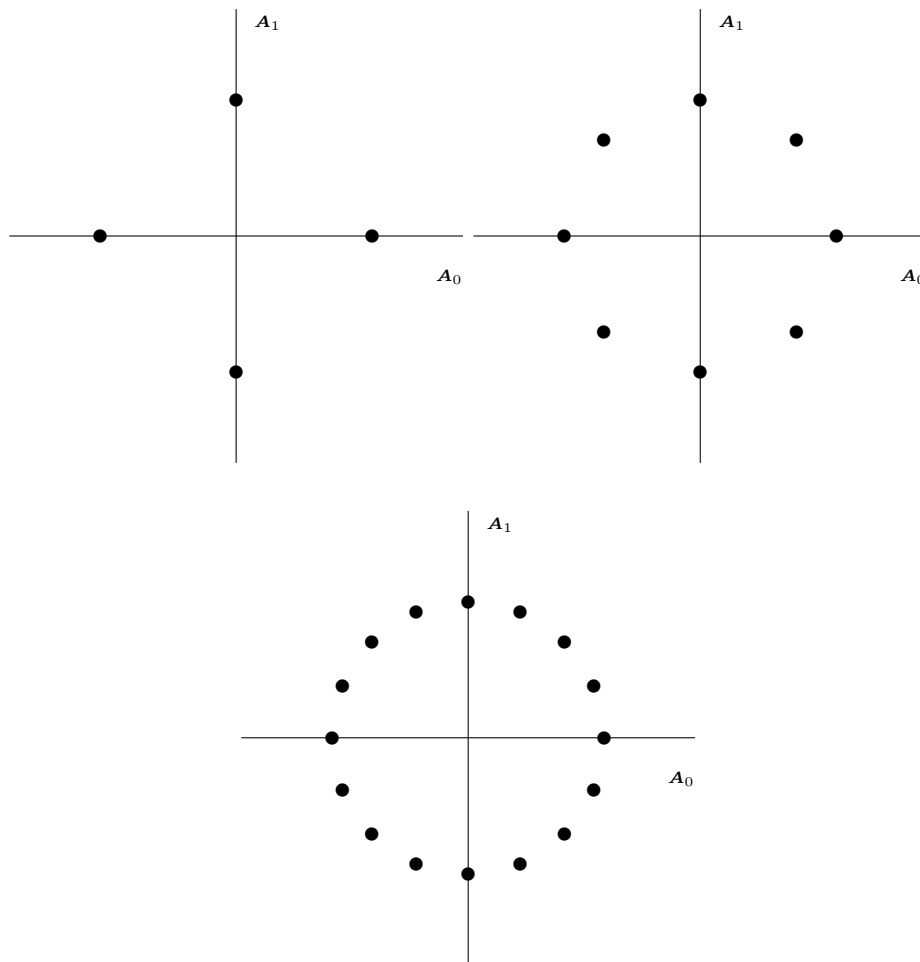


Figura 4.72: Ejemplo de constelaciones QAM de 4, 8 y 16 símbolos.

Su principal ventaja es que como los símbolos tienen todos la misma distancia al origen, todos tienen la misma energía, por lo que la energía permanece constante en todos los intervalos de símbolo durante la transmisión. La implementación de transmisor y receptor tampoco es compleja, ya que los símbolos se pueden codificar como

$$A[n] = \sqrt{E_s} \times e^{j\phi[n]},$$

es decir, que la información de cada símbolo está en una fase. En el decisor, las regiones de decisión se definen a partir de umbrales sobre las fases, por lo que lo único que hay que estimar es la fase de la observación recibida y compararla con los valores que definen las regiones de decisión.

Constelaciones ortogonales unipolares

Estas constelaciones son constelaciones en las que la dimensión del espacio de señales coincide con el número de señales, de modo que se transmiten M señales ortogonales cuya representación vectorial tiene una única componente no nula idéntica, es decir

$$\mathbf{a}_0 = \begin{bmatrix} \sqrt{E_s} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} 0 \\ \sqrt{E_s} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 0 \\ \sqrt{E_s} \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{a}_{M-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \sqrt{E_s} \end{bmatrix}.$$

En este caso, todas las coordenadas son positivas, la energía de todos los símbolos es la misma, y el receptor es muy sencillo, ya que basta con buscar la posición con la mayor componente.

4.6.4. Asignación binaria - Codificación de Gray y cálculo de la BER

Una vez seleccionada una constelación, es necesario realizar la asignación binaria: asignar a cada uno de los M símbolos de la constelación, una de las M posibles combinaciones de m bits. El objetivo al hacer la asignación será minimizar la probabilidad de error de bit, que denotaremos como BER (del inglés “*Bit Error Rate*”), y que se define como

$$BER = P(\hat{B}_b[\ell] \neq B_b[\ell]).$$

Para obtener una regla que permita minimizar esta probabilidad de error, lo primero es saber de qué depende, y como le afecta el tener una cierta asignación binaria en el codificador. El cálculo de esta probabilidad de error se hace de forma similar al de la probabilidad de error de símbolo. Hay que calcular las probabilidades de error de bit condicionales, BER_{a_i} , y promediarlas teniendo en cuenta la probabilidad con la que se transmite cada símbolo de la constelación

$$BER = \sum_{i=0}^{M-1} p_A(a_i) \cdot BER_{a_i}.$$

En cuanto a las probabilidades condicionales, hay que promediar la probabilidad de error que se tiene cuando se comete una equivocación de símbolo decidido teniendo en cuenta con qué probabilidad se producen estos errores, lo que significa evaluar la siguiente expresión

$$BER_{a_i} = \sum_{\substack{j=0 \\ j \neq i}}^{M-1} P_{e|a_i \rightarrow a_j} \times \frac{m_{e|a_i \rightarrow a_j}}{m},$$

donde los parámetros involucrados se definen de la forma siguiente:

- $P_{e|a_i \rightarrow a_j}$: probabilidad de transmitiendo $\mathbf{A} = \mathbf{a}_i$, decidir $\hat{\mathbf{A}} = \mathbf{a}_j$

$$P_{e|a_i \rightarrow a_j} = \int_{\mathbf{q}_0 \in I_j} f_{\mathbf{q}|A}(\mathbf{q}_0 | \mathbf{a}_i) d\mathbf{q}_0$$

- $m_{e|a_i \rightarrow a_j}$: número de errores de bit que conlleva esa decisión
- m : número de bits por símbolo de la constelación

Un ejemplo de cálculo de BER para un sistema unidimensional

A continuación se calcula la BER para una constelación unidimensional de 4 símbolos equiprobables con coordenadas

$$\mathbf{a}_0 = -3, \mathbf{a}_1 = -1, \mathbf{a}_2 = +1, \mathbf{a}_3 = +3.$$

Las regiones de decisión están definidas por los umbrales $q_{u1} = -2, q_{u2} = 0, q_{u3} = +2$

$$I_0 = (-\infty, -2], I_1 = (-2, 0], I_2 = (0, +2], I_3 = (+2, +\infty),$$

y la asignación binaria elegida (de momento de forma arbitraria es)

$$\mathbf{a}_0 \equiv 01, \mathbf{a}_1 \equiv 00, \mathbf{a}_2 \equiv 10, \mathbf{a}_3 \equiv 11.$$

La Figura 4.73 ilustra estas características del sistema bajo evaluación.

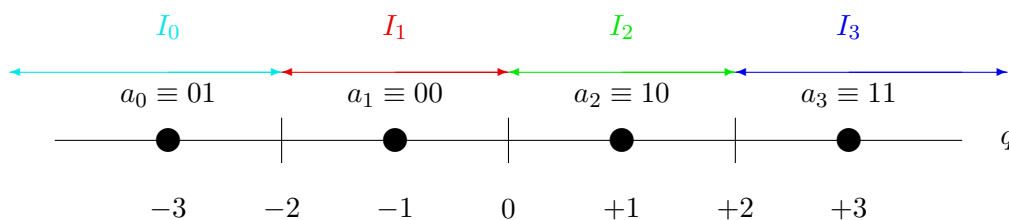


Figura 4.73: Constelación, regiones de decisión, y asignación binaria para el cálculo de la BER.

A continuación se evaluarán las BER condicionales para cada uno de los 4 símbolos de la constelación. Empezamos con el primer símbolo, \mathbf{a}_0 . Para este símbolo, la distribución $f_{q|A}(q|\mathbf{a}_0)$ es gaussiana de media \mathbf{a}_0 y varianza $N_0/2$, por lo que la BER condicional viene dada por

$$BER_{\mathbf{a}_0} = \underbrace{\left[Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_1}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_1}}{m}} + \underbrace{\left[Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_2}} \times \underbrace{\frac{2}{2}}_{\frac{m_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_2}}{m}} + \underbrace{\left[Q\left(\frac{5}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_3}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_3}}{m}}$$

La Figura 4.74 representa graficamente el significado de las probabilidades $P_{e|\mathbf{a}_0 \rightarrow \mathbf{a}_j}$ en distintos colores.

Para el segundo símbolo, \mathbf{a}_1 , la distribución $f_{q|A}(q|\mathbf{a}_1)$ es gaussiana de media \mathbf{a}_1 y varianza $N_0/2$, por lo que la probabilidad de error de bit condicional es

$$BER_{\mathbf{a}_1} = \underbrace{\left[Q\left(\frac{1}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_0}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_0}}{m}} + \underbrace{\left[Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_2}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_2}}{m}} + \underbrace{\left[Q\left(\frac{3}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_3}} \times \underbrace{\frac{2}{2}}_{\frac{m_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_3}}{m}}$$

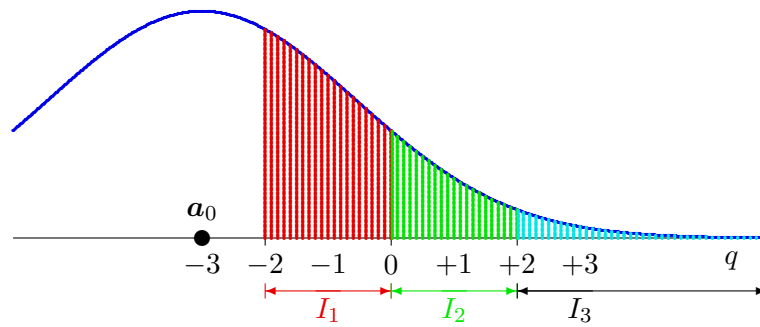


Figura 4.74: Ilustración gráfica de las probabilidades de error entre símbolos, $P_{e|a_0 \rightarrow a_j}$, para $A = a_0$.

La Figura 4.75 representa gráficamente el significado de las probabilidades $P_{e|a_1 \rightarrow a_j}$ en distintos colores.

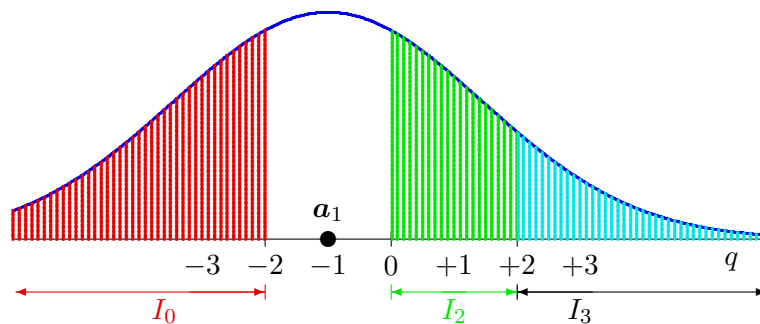


Figura 4.75: Ilustración gráfica de las probabilidades de error entre símbolos, $P_{e|a_1 \rightarrow a_j}$, para $A = a_1$.

Para el tercer símbolo, la distribución $f_{q|A}(q|a_2)$ es gaussiana de media a_2 y varianza $N_0/2$, por lo que

$$\begin{aligned}
 BER_{a_2} = & \underbrace{\left[Q \left(\frac{3}{\sqrt{N_0/2}} \right) \right]}_{P_{e|a_2 \rightarrow a_0}} \times \underbrace{\frac{2}{2}}_{\frac{m_{e|a_2 \rightarrow a_0}}{m}} + \underbrace{\left[Q \left(\frac{1}{\sqrt{N_0/2}} \right) - Q \left(\frac{3}{\sqrt{N_0/2}} \right) \right]}_{P_{e|a_2 \rightarrow a_1}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|a_2 \rightarrow a_1}}{m}} \\
 & + \underbrace{\left[Q \left(\frac{1}{\sqrt{N_0/2}} \right) \right]}_{P_{e|a_2 \rightarrow a_3}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|a_2 \rightarrow a_3}}{m}}
 \end{aligned}$$

La Figura 4.76 representa gráficamente el significado de las probabilidades $P_{e|a_2 \rightarrow a_j}$ en distintos colores.

Finalmente, para el último símbolo de la constelación, a_3 , la distribución condicional de la

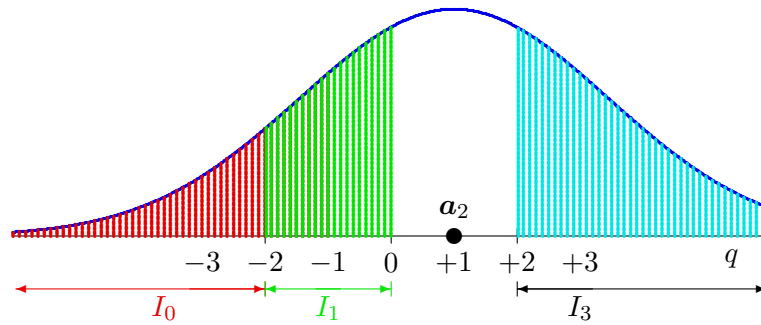


Figura 4.76: Ilustración gráfica de las probabilidades de error entre símbolos, $P_{e|a_2 \rightarrow a_j}$, para $\mathbf{A} = \mathbf{a}_2$.

observación, $f_{q|\mathbf{A}}(q|\mathbf{a}_3)$, es gaussiana de media \mathbf{a}_3 y varianza $N_0/2$, lo que significa que

$$\begin{aligned}
 BER_{\mathbf{a}_3} &= \underbrace{\left[Q\left(\frac{5}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_0}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_0}}{m}} + \underbrace{\left[Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_1}} \times \underbrace{\frac{2}{2}}_{\frac{m_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_1}}{m}} \\
 &+ \underbrace{\left[Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) \right]}_{P_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_2}} \times \underbrace{\frac{1}{2}}_{\frac{m_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_2}}{m}}
 \end{aligned}$$

La Figura 4.77 representa gráficamente el significado de las probabilidades $P_{e|\mathbf{a}_3 \rightarrow a_j}$ en distintos colores.

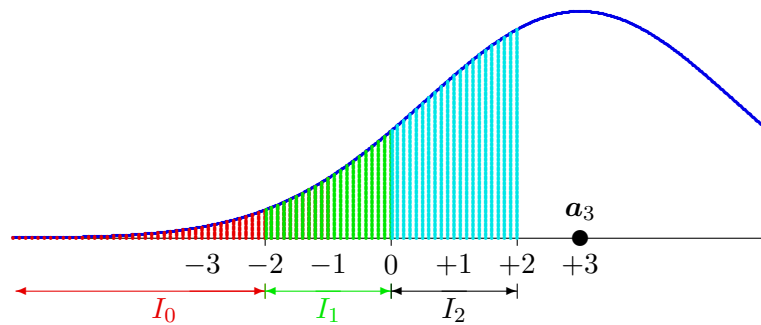


Figura 4.77: Ilustración gráfica de las probabilidades de error entre símbolos, $P_{e|\mathbf{a}_3 \rightarrow a_j}$, para $\mathbf{A} = \mathbf{a}_3$.

Una vez calculadas las BER condicionales, la BER total se obtiene promediándolas, lo que para este ejemplo lleva a

$$\begin{aligned}
 BER &= \frac{1}{4} \times BER_{\mathbf{a}_0} + \frac{1}{4} \times BER_{\mathbf{a}_1} + \frac{1}{4} \times BER_{\mathbf{a}_2} + \frac{1}{4} \times BER_{\mathbf{a}_3} \\
 &= \frac{3}{4} Q\left(\frac{1}{\sqrt{N_0/2}}\right) + \frac{1}{2} Q\left(\frac{3}{\sqrt{N_0/2}}\right) - \frac{1}{4} Q\left(\frac{5}{\sqrt{N_0/2}}\right).
 \end{aligned}$$

Veamos qué ocurre si se modifica la asignación binaria. Si se utilizara ahora, por ejemplo, la asignación binaria

$$\mathbf{a}_0 \equiv 11, \mathbf{a}_1 \equiv 00, \mathbf{a}_2 \equiv 10, \mathbf{a}_3 \equiv 01,$$

si se repitiera el proceso anterior para calcular la BER, se llegaría a otro resultado distinto, en concreto

$$BER = \frac{5}{4}Q\left(\frac{1}{\sqrt{N_0/2}}\right) - \frac{1}{4}Q\left(\frac{3}{\sqrt{N_0/2}}\right).$$

Se puede ver que este valor es mayor (el término dominante es el de menor argumento de la función $Q(x)$, en este caso $Q\left(\frac{1}{\sqrt{N_0/2}}\right)$, que en el primer caso está multiplicado por $\frac{3}{4}$, y en el segundo por $\frac{5}{4}$. ¿Qué es lo que hace que una asignación sea mejor o peor respecto a la BER que produce? Si se observa el procedimiento de cálculo, veremos que modificar la asignación binaria no modifica los términos de probabilidad de error entre un símbolo y otro, es decir

No se modifican los términos $P_{e|a_i \rightarrow a_j}$.

Sin embargo, si que cambia la probabilidad de error de bit asignada a cada error de símbolo

Sí se modifican los términos $m_{e|a_i \rightarrow a_j}$.

Esto significa, que dependiendo de la asignación binaria elegida, los términos

$$\frac{m_{e|a_i \rightarrow a_j}}{m}$$

pueden tomar uno de dos valores: $\frac{1}{2}$ ó $\frac{2}{2} = 1$. Cambiar la asignación binaria implica cambiar los términos que tienen asociado uno u otro valor. Lo que interesa aquí es asociar esos valores de forma que se minimize la probabilidad de error de bit. Eso implica tratar de asignar el menor valor, $1/2$, a aquellos términos que más pesen al ser ponderados por $P_{e|a_i \rightarrow a_j}$. Y los términos que tienen un mayor valor de probabilidad de error de símbolo $P_{e|a_i \rightarrow a_j}$ son los asociados a símbolos que están a mínima distancia en la constelación. De aquí surge la regla que optimiza la asignación binaria, y que se denomina *codificación de Gray*.

Codificación de Gray La asignación binaria de símbolos que están a mínima distancia, debe diferir únicamente en 1 bit, de modo que los términos que más pesan en la BER (para valores de $P_{e|a_i \rightarrow a_j}$ más elevados) se multipliquen por el valor $\frac{1}{m}$.

Dicho de otra forma. Cuando se produce un error de símbolo, lo más probable es que sea un error con un símbolo que está a mínima distancia del transmitido. Por eso, lo mejor es hacer en que el tipo de error más frecuente, sólo se produzca un error de bit sobre los m bits codificados. De esta forma, utilizando un codificación de Gray, la probabilidad de error de bit se puede aproximar para relaciones señal a ruido razonablemente altas como

$$BER \approx \frac{1}{m} \cdot P_e,$$

siendo $m = \log_2(M)$ el número de bits por símbolo del sistema.

En todo codificador deber utilizarse una codificación de Gray, o si no es posible, una codificación pseudo-Gray, donde la regla se haga cumplir para tantos símbolos a mínima distancia como sea posible. Para los tipos de constelaciones más utilizadas en la práctica, QAM y PSK, es siempre posible encontrar una codificación de Gray. En el caso de QAM esto es sencillo porque la codificación bidimensional se puede hacer estableciendo una codificación de Gray unidimensional con la mitad de bits asignados de forma independiente en cada dirección del espacio, tal y como se muestra en la Figura 4.78.

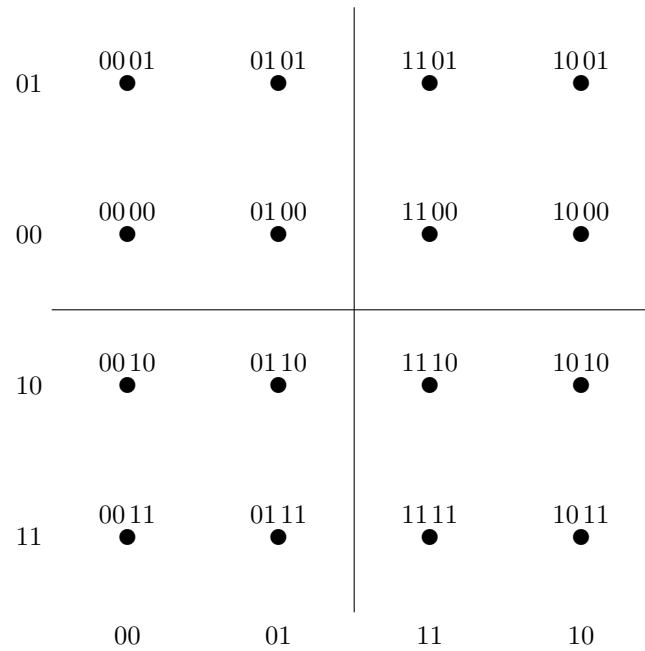


Figura 4.78: Ejemplo de codificación de Gray para una constelación 16-QAM.

Para las constelaciones PSK, al estar los puntos en un círculo, también es relativamente sencillo encontrar una codificación de Gray. Basta plantear una codificación de Gray unidimensional en la que los símbolos de los extremos difieran en un único bit, y convertirla en una asignación circular en el espacio 2D, como en el ejemplo de la Figura 4.79.

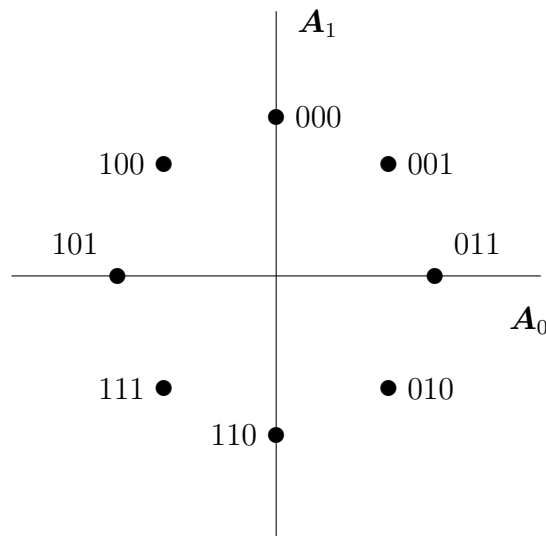


Figura 4.79: Ejemplo de codificación de Gray para una constelación 8-PSK.

4.6.5. Relación entre tasa de bit y tasa de símbolo

Para finalizar con el codificador, recordar únicamente la relación entre tasa de símbolo y tasa de bit. Dado que cada símbolo transporta m bits, la relación entre velocidades de transmisión o tasas de símbolo y de bit es obvia

$$R_b = m \times R_s \text{ bits/s}$$

M (símbolos)	m (bits/símbolo)	E_s con niveles normalizados ($d_{min} = 2$)	d_{min} con $E_s = 2$
4	2	2	2
16	4	10	0,8944
64	8	42	0,4364
256	16	170	0,2169

Tabla 4.1: Transmisión con constelaciones M -QAM.

ó

$$R_s = \frac{R_b}{m} \text{ baudios (símbolos/s).}$$

Esto significa que un sistema con un mayor número de bits por símbolo (o lo que es lo mismo, con un mayor número de símbolos, $M = 2^m$), tendrá para la misma velocidad de símbolo una mayor capacidad de transmisión de bits. Lo que podría invitar a considerar la idea de transmitir constelaciones muy densas (con muchos símbolos). Sin embargo, hay que tener en cuenta otro factor a la hora de elegir una cierta densidad de la constelación. Si el sistema tiene una limitación en energía, constelaciones más densas implicarán menores distancias entre símbolos, lo que llevará a su vez a peores prestaciones. Para ilustrarlo con un ejemplo, analicemos un sistema con constelaciones M -QAM. La Tabla 4.1 muestra para distintos tamaños de la constelación, como aumenta la energía necesaria si se mantiene la distancia mínima (niveles normalizados) o como equivalentemente se reduce la distancia mínima si se mantiene constante la energía media por símbolo de la constelación. En la Figura 4.80, se muestran a escala las constelaciones para ese nivel de energía media por símbolo constante, $E_s = 2$ en este caso. Se puede ver que al elegir constelaciones más densas, aunque el número de bits por símbolo se incrementa, disminuye la relación entre prestaciones y energía, por lo que en la práctica de nuevo hay que buscar un compromiso a la hora de establecer la densidad óptima de la constelación.

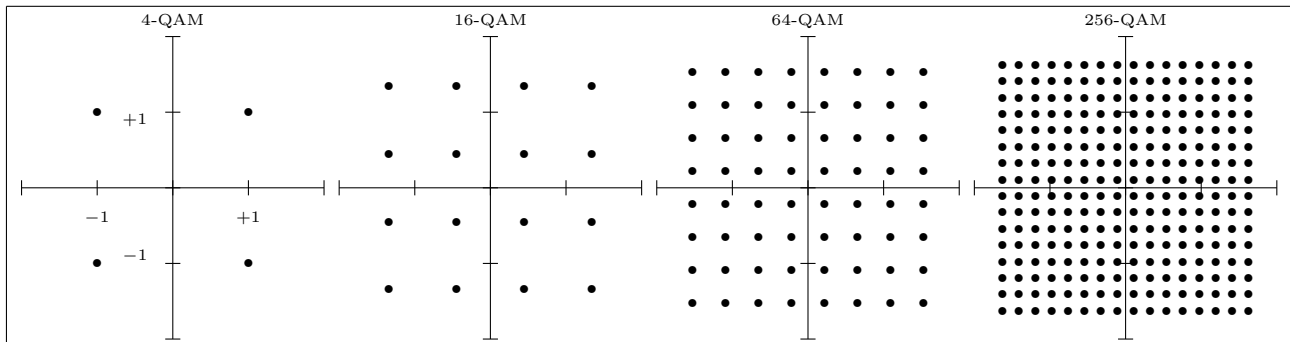


Figura 4.80: Distintas constelaciones M -QAM con la misma energía media por símbolo ($E_s = 2$).

4.7. Modulador

El modulador es el segundo elemento funcional del sistema de transmisión, tras el codificador. Su función consiste en transformar la secuencia de representaciones vectoriales de las señales a transmitir, $\mathbf{A}[n]$, en la señal modulada $s(t)$ que contiene la información a transmitir. Esta función se ilustra en la Figura 4.81. La forma de hacerlo será transformando cada vector de la secuencia $\mathbf{A}[n]$ en una señal de duración T que definirá la forma de señal en el intervalo de símbolo corres-

pendiente, $nT \leq t < (n + 1)T$, de tal modo que si $\mathbf{A}[n] = \mathbf{a}_k$ entonces $s(t) = s_k(t - nT)$ en $nT \leq (n + 1)T$, formando así por intervalos de símbolo la señal modulada $s(t)$.



Figura 4.81: Codificador en el transmisor de un sistema de comunicaciones digital.

4.7.1. Diseño del modulador

La función básica del modulador, como se ha dicho, es generar una señal a partir de su representación discreta. Como la relación entre señal y representación discreta viene dada por la expresión

$$s_i(t) = \sum_{j=0}^{N-1} a_{i,j} \times \phi_j(t),$$

lo que hace falta para realizar esta conversión es definir una base ortonormal de N elementos (dimensión del espacio de señales). Y en eso consiste el diseño del modulador, en la elección de esa base ortonormal de dimensión N

$$\text{Diseño del modulador: } \{\phi_0(t), \phi_1(t), \dots, \phi_{N-1}(t)\}.$$

Así, si $\mathbf{A}[n] = \mathbf{a}_k$ entonces $s(t) = s_k(t - nT)$ en $nT \leq (n + 1)T$, con lo que la expresión analítica de la señal completa es

$$s(t) \sum_n \sum_{j=0}^{N-1} A_j[n] \times \phi_j(t - nT),$$

donde $A_j[n]$ denota la coordenada de índice j del símbolo $\mathbf{A}[n]$, es decir

$$\mathbf{A}[n] = \begin{bmatrix} A_0[n] \\ A_1[n] \\ \vdots \\ A_{N-1}[n] \end{bmatrix}.$$

Así como el diseño del modulador se hacía teniendo en cuenta dos factores, prestaciones y energía, el diseño del modulador se hace en base a un único factor: las características del canal. Se debe buscar una base ortonormal cuyos elementos minimizen la distorsión lineal que sufre la señal durante su transmisión a través del canal. Si la respuesta del canal es $h(t)$ en el dominio temporal, o equivalentemente su transformada de Fourier $H(j\omega)$ en el dominio frecuencial, idealmente se desea que la base sea tal que no exista ninguna distorsión lineal, para lo que debe cumplirse la siguiente condición, expresada en el dominio temporal y en el frecuencial

$$\phi_i(t) * h(t) = \phi_i(t) \Leftrightarrow \Phi_i(j\omega) \times H(j\omega) = \Phi_i(j\omega).$$

Aunque en muchos casos en esta asignatura se considerará que se puede lograr una adecuación perfecta a las condiciones del canal, en la práctica esto no será posible, con lo que habrá que conformarse con seleccionar señales cuya respuesta en frecuencia esté en la banda de paso del canal de transmisiones, lo que da lugar a una distinción entre canales en banda base y canales paso banda.

4.7.2. Algunos ejemplos de moduladores y señales moduladas que generan

En esta sección se muestran algunos ejemplos de moduladores, y del tipo de señales moduladas que se generan con los mismos cuando se desea transmitir una cierta secuencia de bits de información.

El primer ejemplo es un sistema cuya constelación (CODIFICADOR) y base ortonormal (MODULADOR), son los mostrados en la Figura 4.82

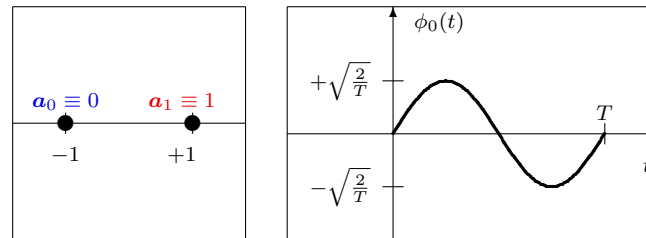


Figura 4.82: Constelación (CODIFICADOR) y base ortonormal (MODULADOR) de un sistema de comunicaciones. Ejemplo A.

Teniendo en cuenta que se trata de un sistema unidimensional, las señales asociadas a cada uno de los símbolos son

$$s_0(t) = -1 \times \phi_0(t), \quad s_1(t) = +1 \times \phi_0(t).$$

Al tratarse de un espacio de dimensión $N = 1$, las dos señales son réplicas escaladas, con distintos factores de escala, de la misma señal, la base ortonormal que define el modulador, $\phi_0(t)$. Las dos señales se muestran en la Figura 4.83

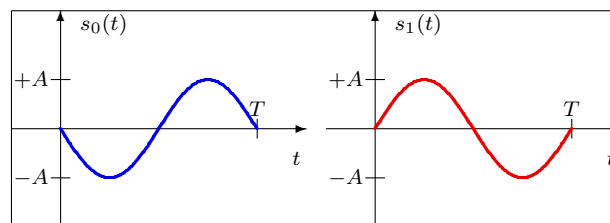


Figura 4.83: Señales asociadas a cada símbolo en el Ejemplo A.

Para la siguiente secuencia binaria de información

$$B_b[\ell] = 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ \dots$$

la señal modulada correspondiente es la que muestra la Figura 4.84

En cada intervalo de símbolo se transmite un bit, que se transporta en una de las dos formas de onda, $s_0(t)$ (en azul) transporta el bit 0, y $s_1(t)$ (en rojo) transporta el bit 1.

En el segundo ejemplo, se considera un sistema binario pero ahora en un espacio bidimensional. La constelación (CODIFICADOR) y base ortonormal (MODULADOR) empleados en este caso son los mostrados en la Figura 4.85

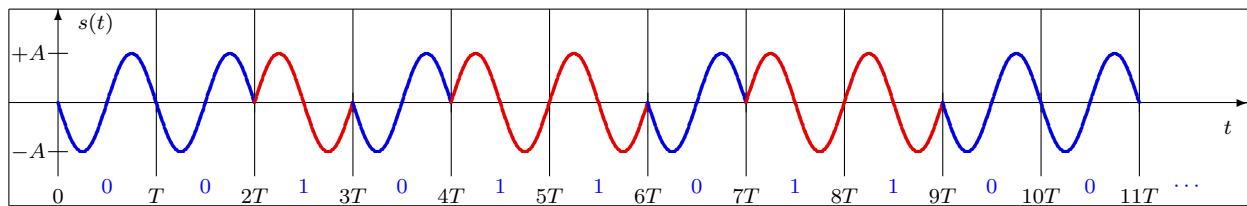


Figura 4.84: Señal modulada para la secuencia de información transmitida en el Ejemplo A.

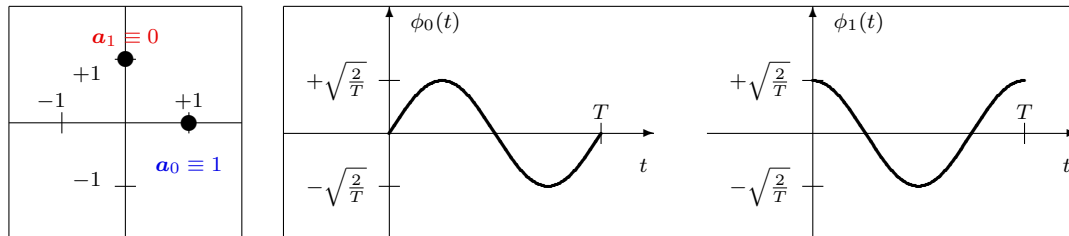


Figura 4.85: Constelación (CODIFICADOR) y base ortonormal (MODULADOR) de un sistema de comunicaciones. Ejemplo B.

Teniendo en cuenta que se trata de un sistema bidimensional, las señales asociadas a cada uno de los símbolos son

$$s_0(t) = +1 \times \phi_0(t) + 0 \times \phi_1(t), \quad s_1(t) = 0 \times \phi_0(t) + 1 \times \phi_1(t),$$

que a diferencia del caso anterior ahora se forman a partir de la combinación lineal de dos señales, los dos elementos de la base ortonormal que define el modulador. Las dos señales se muestran en la Figura 4.86

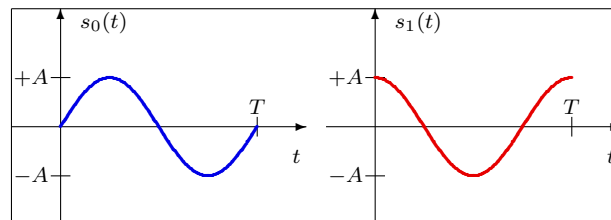


Figura 4.86: Señales asociadas a cada símbolo en el Ejemplo B.

Para la misma secuencia binaria de información que en el caso anterior

$$B_b[\ell] = 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ \dots ,$$

la señal modulada correspondiente es la que muestra la Figura 4.87

Al igual que en el ejemplo anterior, en cada intervalo de símbolo se transmite un bit, que se transporta en una de las dos formas de onda, $s_0(t)$ (en azul) transporta el bit 0, y $s_1(t)$ (en rojo) transporta el bit 1. Lo que cambia con respecto al ejemplo anterior es la forma que tienen ahora las señales que se utilizan para transmitir cada símbolo.

En el tercer ejemplo, utilizando el mismo modulador que en el caso anterior, ahora se utilizará una constelación de $M = 4$ símbolos. La constelación (CODIFICADOR) y base ortonormal (MODULADOR) empleados en este caso son los mostrados en la Figura 4.88

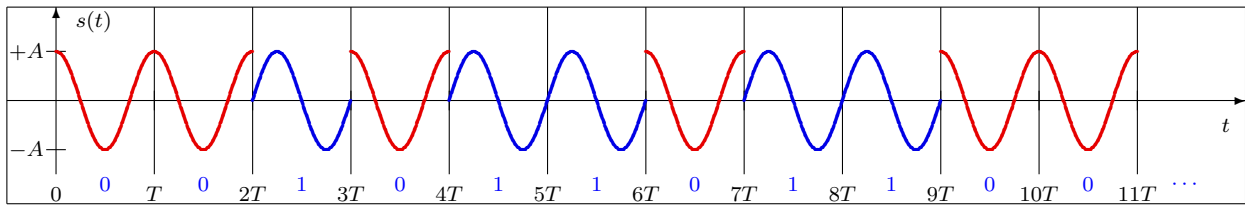


Figura 4.87: Señal modulada para la secuencia de información transmitida en el Ejemplo B.

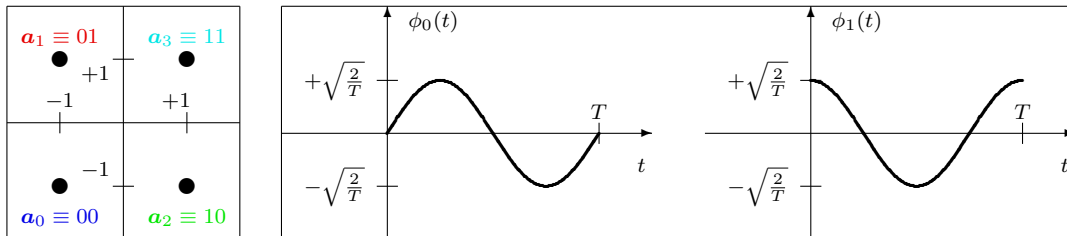


Figura 4.88: Constelación (CODIFICADOR) y base ortonormal (MODULADOR) de un sistema de comunicaciones. Ejemplo C.

De nuevo se trata de un sistema bidimensional, donde las señales asociadas a cada uno de los símbolos son

$$\begin{aligned}
 s_0(t) &= -1 \times \phi_0(t) - 1 \times \phi_1(t), \\
 s_1(t) &= -1 \times \phi_0(t) + 1 \times \phi_1(t), \\
 s_2(t) &= +1 \times \phi_0(t) - 1 \times \phi_1(t),
 \end{aligned}$$

y

$$s_3(t) = +1 \times \phi_0(t) + 1 \times \phi_1(t),$$

que de nuevo se forman a partir de la combinación lineal de dos señales, los dos elementos de la base ortonormal que define el modulador. Las cuatro señales se muestran en la Figura 4.89

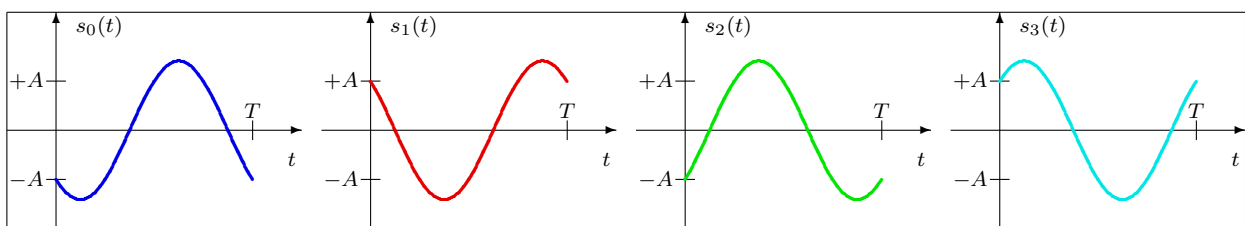


Figura 4.89: Señales asociadas a cada símbolo en el Ejemplo C.

Si la secuencia binaria de información que se desea transmitir es ahora

$$B_b[\ell] = 00\ 10\ 11\ 01\ 10\ 00\ 11\ 00\ 01\ 00\ 10\ \dots,$$

donde por conveniencia se han separado los bits en bloques de tamaño $m = 2$ bits, la señal modulada correspondiente es la que muestra la Figura 4.90

A diferencia de los dos ejemplos anteriores, que eran binarios ($M = 2$), ahora en cada intervalo de símbolo se transmiten no uno sino dos bits, que se transportan en una de las cuatro formas de onda, $s_0(t)$ (en azul) transporta la dupla 00, $s_1(t)$ (en rojo) transporta la dupla 01, $s_2(t)$ (en

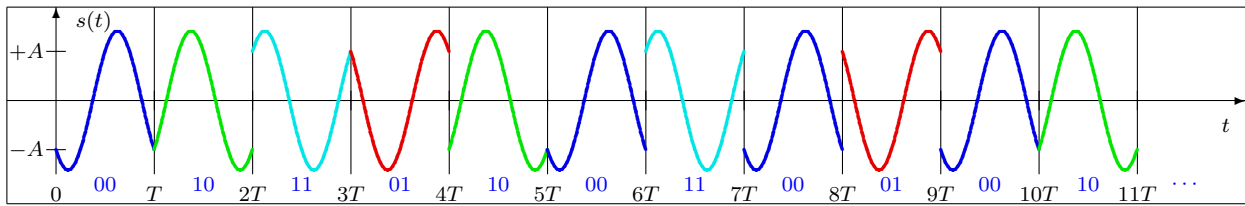


Figura 4.90: Señal modulada para la secuencia de información transmitida en el Ejemplo C.

verde) transporta la dupla 10, y $s_3(t)$ (en cyan) transporta la dupla 11. Ahora, al tratarse de una constelación de $M = 4$ símbolos, cada símbolo transporta $m = \log_2(M) = 2$ bits de información.

Finalmente, el cuarto ejemplo utiliza la misma constelación de cuatro símbolos que el ejemplo anterior, con idéntica asignación binaria, pero con otro modulador de dimensión $N = 2$, tal y como se muestra en la Figura 4.91

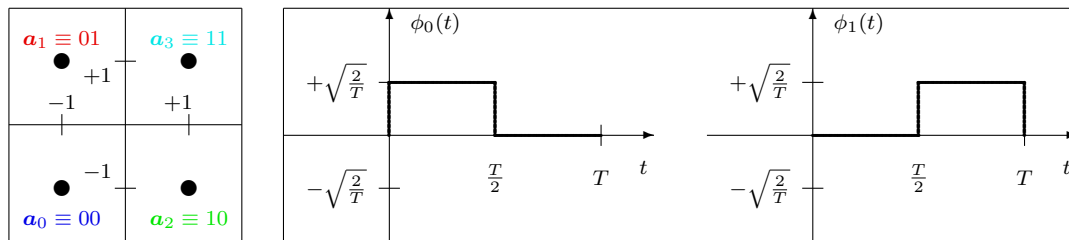


Figura 4.91: Constelación (CODIFICADOR) y base ortonormal (MODULADOR) de un sistema de comunicaciones. Ejemplo D.

La expresión analítica de las cuatro señales en función de los elementos de la base, al no cambiar la constelación, es la misma que en el ejemplo anterior

$$s_0(t) = -1 \times \phi_0(t) - 1 \times \phi_1(t),$$

$$s_1(t) = -1 \times \phi_0(t) + 1 \times \phi_1(t),$$

$$s_2(t) = +1 \times \phi_0(t) - 1 \times \phi_1(t),$$

y

$$s_3(t) = +1 \times \phi_0(t) + 1 \times \phi_1(t),$$

pero con distintos valores para las dos funciones que forman la base ortonormal, por lo que las cuatro señales resultantes son ahora las que se muestran en la Figura 4.92

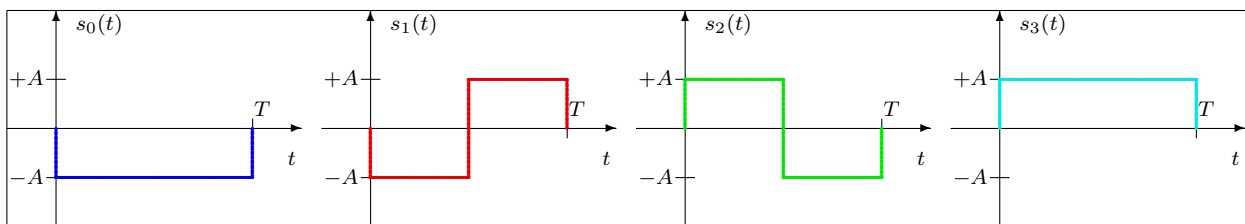


Figura 4.92: Señales asociadas a cada símbolo en el Ejemplo D.

Si se transmite la misma secuencia binaria de información que en el ejemplo anterior

$$B_b[\ell] = 00\ 10\ 11\ 01\ 10\ 00\ 11\ 00\ 01\ 00\ 10\ \dots ,$$

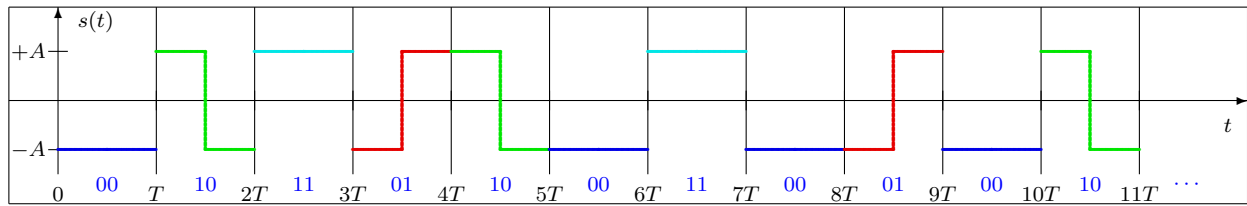


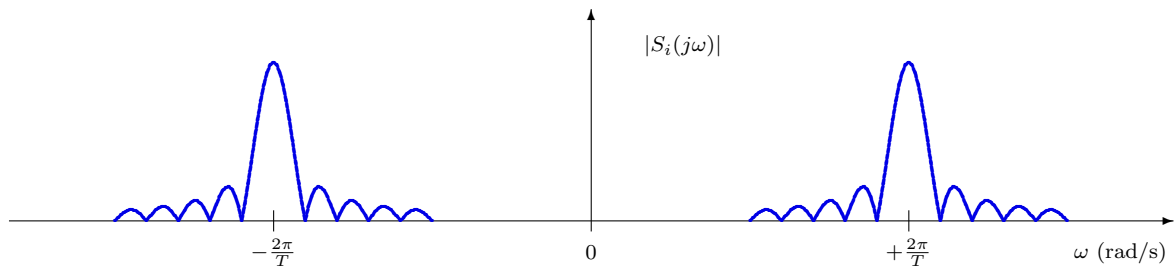
Figura 4.93: Señal modulada para la secuencia de información transmitida en el Ejemplo D.

la señal modulada correspondiente es la que muestra la Figura 4.93

De nuevo, en cada intervalo de símbolo se transmiten dos bits, que se transportan en una de las cuatro formas de onda, $s_0(t)$ (en azul) transporta la dupla 00, $s_1(t)$ (en rojo) transporta la dupla 01, $s_2(t)$ (en verde) transporta la dupla 10, y $s_3(t)$ (en cyan) transporta la dupla 11. Al tratarse de la misma constelación de $M = 4$ símbolos, cada símbolo transporta de nuevo $m = \log_2(M) = 2$ bits de información. Lo que cambia es la forma de onda que transporta cada dupla de bits, al haber cambiado la base ortonormal que define el modulador.

Si comparamos los sistemas de los dos últimos ejemplos, que comparten codificador, las prestaciones si se transmite sobre un canal gaussiano son idénticas, y las energías de los cuatro símbolos también lo son, ya que estas vienen determinadas por la elección del codificador (que contiene la representación discreta de las señales). Lo que cambia del tercer ejemplo al cuarto es la forma de onda de las señales y por tanto su respuesta en frecuencia. En el cuarto ejemplo tenemos señales cuadradas, cuya respuesta en frecuencia es banda base (centrada en 0 Hz), mientras que en el tercer ejemplo tenemos señales cuya respuesta en frecuencia está centrada en la frecuencia de las sinusoides que forman la base, por lo que es una respuesta paso banda, tal y como se muestra en la Figura 4.94. La elección de una base ortonormal (modulador) y otra, dependerá de las características del canal por el que queramos transmitir, y en particular en este caso, de la banda en la que se desee realizar la transmisión. Si el canal tiene respuesta paso bajo, será mejor emplear el modulador del cuarto ejemplo, que genera señales adecuadas para este tipo de canales. Si el canal tiene una respuesta paso banda, será mejor emplear el modulador del tercer ejemplo, que genera señales más adecuadas para este tipo de respuesta.

■ Ejemplo C



■ Ejemplo D

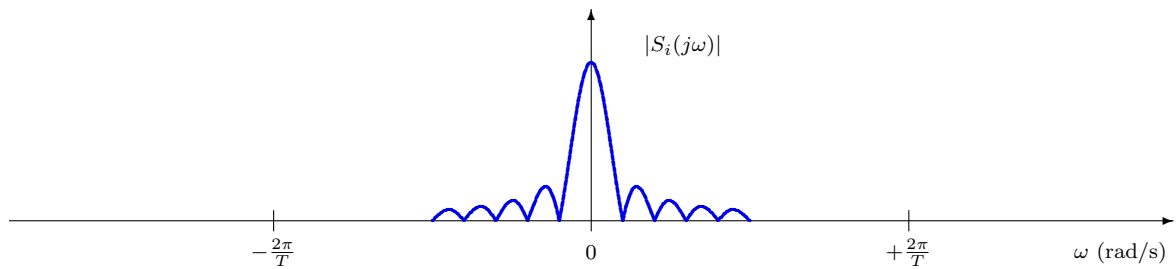


Figura 4.94: Respuesta en frecuencia de las señales de los ejemplos C y D.

4.8. Ejercicios

Ejercicio 4.1 Para el conjunto de símbolos de la Figura 4.95

- a) Aplique el procedimiento de Gram-Schmidt para obtener una base ortonormal que permita la representación vectorial de las señales
 - I) Obtenga la base ortonormal, y diga cuál es la dimension del espacio de señales.
 - II) Obtenga la representación vectorial de las señales.
- b) Calcule la energía de cada señal a partir de su representación vectorial (compárela con la obtenida a partir de la definición temporal de la señal).
- c) Calcule la energía de la diferencia entre cada una las seis señales y la señal $s_0(t)$ a partir de la representación vectorial de las señales (puede compararla con la obtenida en el dominio temporal).

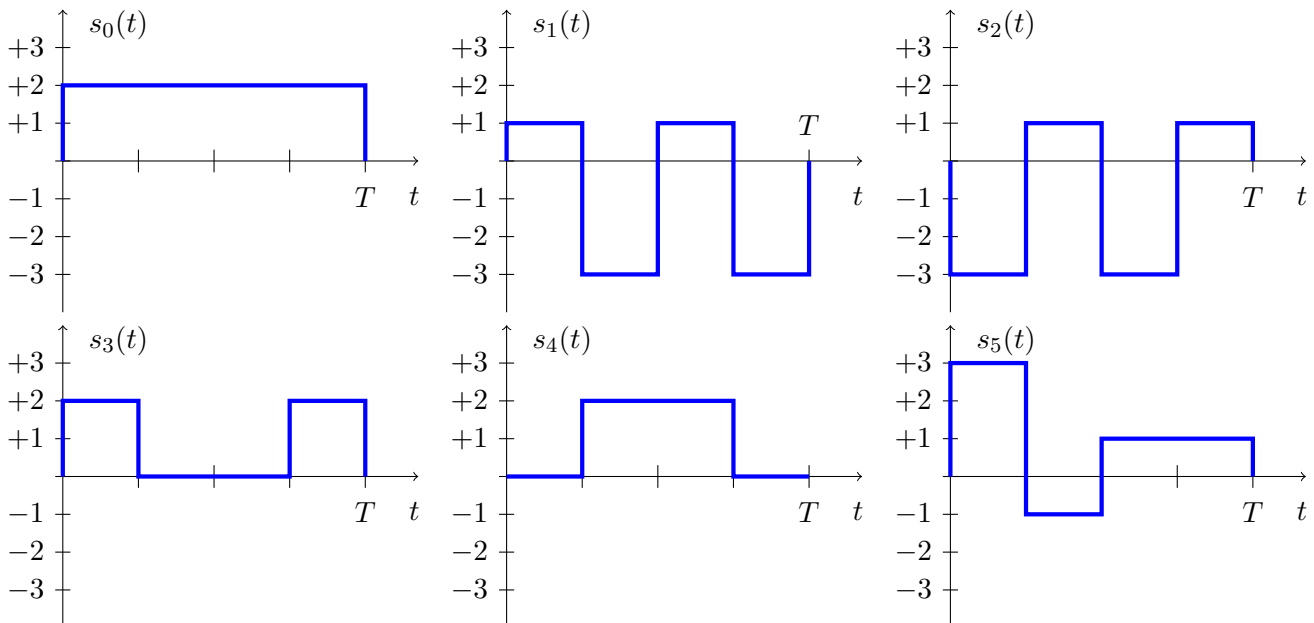


Figura 4.95: Señales para el Ejercicio 4.1.

Ejercicio 4.2 Un sistema de comunicaciones binario emplea pulsos rectangulares causales de duración T y amplitudes $\pm A$ para transmitir información a una velocidad de 100 kbits/s. Si la densidad espectral de potencia del ruido aditivo gaussiano es $N_0/2$, con $N_0 = 10^{-4}$ W/Hz, y suponiendo que ambos símbolos se transmiten con la misma probabilidad, determine el valor de A necesario para:

- a) Obtener una probabilidad de error de símbolo aproximada $P_e \approx 10^{-4}$.
- b) Obtener una probabilidad de error de bit aproximada $BER \approx 10^{-6}$.

Ejercicio 4.3 Un sistema de comunicaciones binario utiliza una constelación con dos símbolos, $\mathbf{a}_0 = +A$ y $\mathbf{a}_1 = 0$. Si se transmite sobre un canal aditivo gaussiano, con densidad espectral de potencia $N_0/2$, y las probabilidades de los símbolos son $p_A(\mathbf{a}_0) = 1/3$, y $p_A(\mathbf{a}_1) = 2/3$:

- a) Calcule el decisor óptimo (es decir, calcule el umbral de decisión).
- b) Calcule la probabilidad de error.

Ejercicio 4.4 Se tiene la constelación de la Figura 4.96. Se supone que los símbolos tienen la misma probabilidad. Tomando $T = 1$

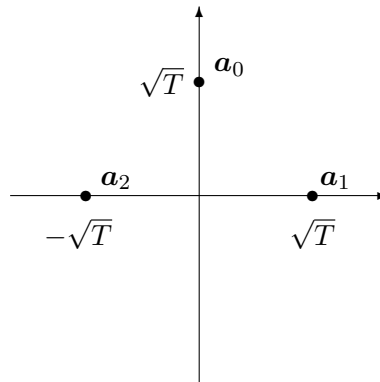


Figura 4.96: Constelación Ejercicio 4.4.

- a) Calcule al energía media por símbolo, E_s .
- b) Diseñe una constelación alternativa que, con la misma probabilidad de error, tenga la mínima energía media por símbolo. Calcule el nuevo valor de E_s .
- c) Para cualquiera de las dos constelaciones, obtenga un límite de la probabilidad de error mediante la cota de la unión y la cota holgada.

Ejercicio 4.5 Para la implementación de un sistema de comunicaciones de tres símbolos equiprobables, el modulador transmite las tres señales mostradas en la Figura 4.97.

- a) Calcule la energía media por símbolo.
- b) Calcule alguna cota de la probabilidad de error.

Ejercicio 4.6 Un sistema de comunicaciones emplea la constelación unidimensional

$$\mathbf{a}_0 = 0, \mathbf{a}_1 = 1, \mathbf{a}_2 = 2, \mathbf{a}_3 = 3,$$

Las observaciones proporcionadas por el demodulador están caracterizadas por el canal discreto equivalente

$$q[n] = A[n] + z[n],$$

donde cada muestra del término de ruido, $z[n]$, tiene la función densidad de probabilidad representada en la Figura 4.98.

- a) Obtenga la expresión analítica de la función densidad de probabilidad de la observación condicionada a la transmisión de cada símbolo (es decir, $f_{q|A}(q|a_i)$, para todo i).
- b) Diseñe el decisor óptimo para símbolos equiprobables.
- c) Calcule la probabilidad de error para el decisor anterior.

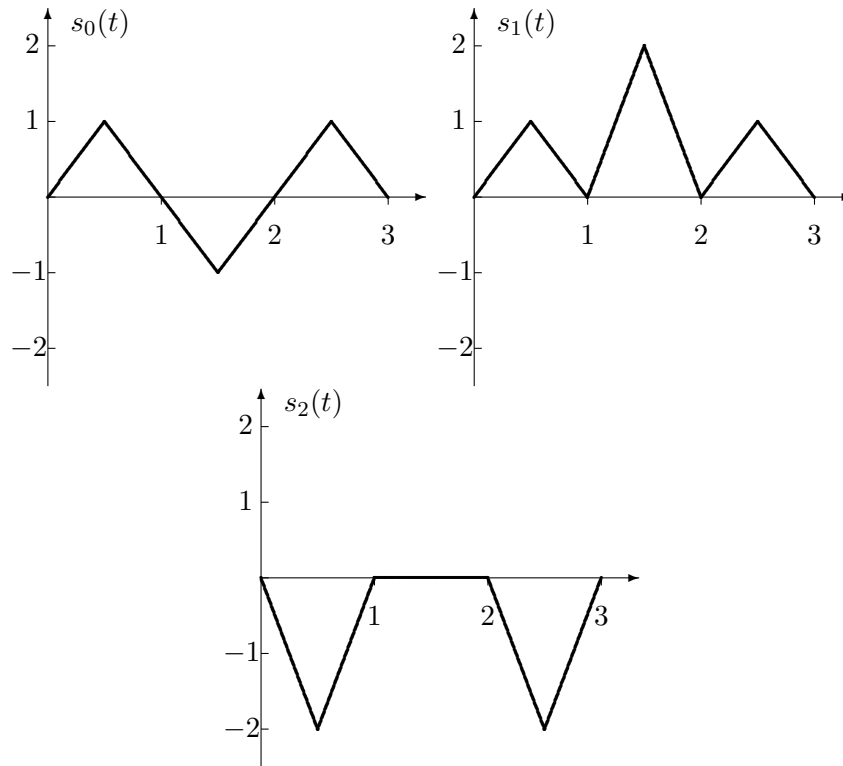


Figura 4.97: Señales Ejercicio 4.5.

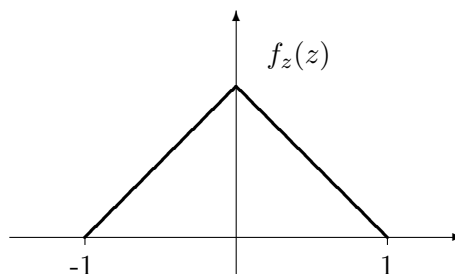


Figura 4.98: Constelación y función densidad de probabilidad del término de ruido para el Ejercicio 4.6.

- d) Diseñe el decisor óptimo para las siguientes probabilidades de símbolo: $p_A(a_0) = p_A(a_3) = 1/6$, $p_A(a_1) = p_A(a_2) = 1/3$.
- e) Calcule la probabilidad de error para el decisor anterior.

Ejercicio 4.7 Un sistema de comunicaciones binario utiliza las siguientes señales para la transmisión de los dos símbolos:

$$s_0(t) = -s_1(t) = \begin{cases} \sqrt{\frac{E_b}{N_0}}, & 0 \leq t \leq T \\ 0, & \text{en otro caso} \end{cases}$$

El receptor se implementa como se muestra en la Figura 4.99 (representación para la recepción del primer símbolo de la secuencia).

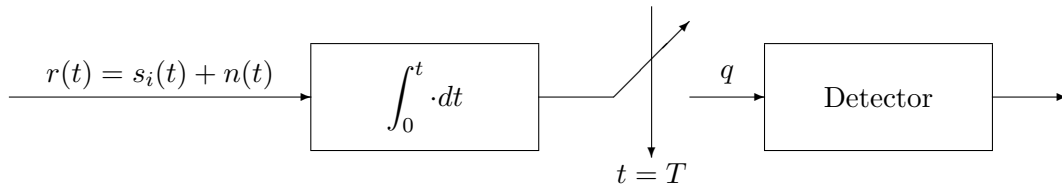


Figura 4.99: Receptor para el Ejercicio 4.7.

Determine la relación señal a ruido a la salida del demodulador (en q), suponiendo que el ruido aditivo es de media nula, gaussiano y con densidad espectral de potencia $N_0/2$ W/Hz.

Ejercicio 4.8 El código de línea de Manchester utiliza las señales de la Figura 4.100 para transmitir los dos símbolos del código.

- a) Determine la probabilidad de error si ambos símbolos se transmiten con igual probabilidad.
- b) Determine la probabilidad de error si la probabilidad de los símbolos es $p_A(a_0) = p$ y $p_A(a_1) = 1 - p$.

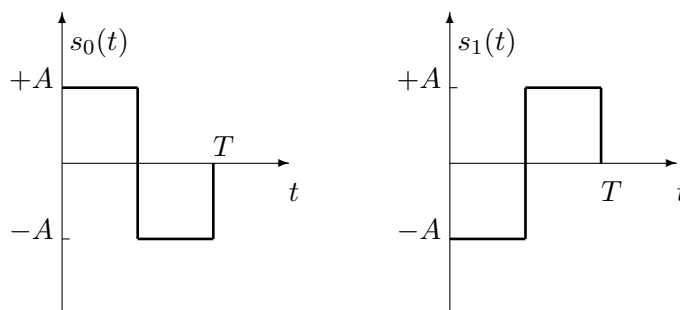


Figura 4.100: Señales para el Ejercicio 4.8.

Ejercicio 4.9 Considere un sistema digital de comunicaciones que transmite la información empleando una constelación QAM a una tasa de símbolo de 2400 baudios (símbolos/s). Se asume ruido aditivo blanco y gaussiano.

- a) Determine la relación E_b/N_0 requerida para conseguir una probabilidad de error de símbolo aproximada de 10^{-5} para una tasa binaria de 4800 bits/s.
- b) Repita el cálculo para una tasa binaria de 9600 bits/s.
- c) Repita el cálculo para una tasa binaria de 19200 bits/s.
- d) Exponga las conclusiones obtenidas a partir de estos resultados.

Ejercicio 4.10 Un sistema de comunicaciones transmite dos símbolos mediante las señales $s_0(t)$ y $s_1(t)$ que se muestran a continuación en la Figura 4.101

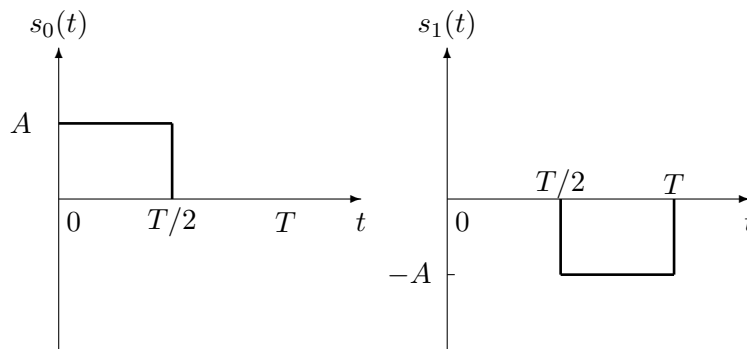


Figura 4.101: Señales para el Ejercicio 4.10.

- a) Se tiene un canal aditivo gaussiano (ruido blanco con densidad espectral de potencia $N_0/2$). Dibuje la constelación, diseñe el receptor (demodulador + decisor) óptimo, y calcule la probabilidad de error.
- b) Si el canal es tal que su salida a una entrada $s(t)$ es $\alpha s(t) + \alpha A$, de nuevo con el mismo tipo de ruido aditivo, rediseñe el decisor óptimo y calcule la probabilidad de error.
- c) Si en la situación del apartado a) se utiliza el demodulador de la Figura 4.102, diseñe el decisor óptimo y calcule la probabilidad de error en este caso. Comente los resultados comparándolos con los obtenidos en el apartado a).

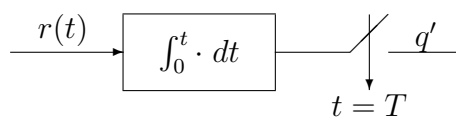


Figura 4.102: Demodulador para el Ejercicio 4.10.

NOTA: Tenga en cuenta que si no se emplea un demodulador normalizado, la varianza de ruido discreto ya no es $N_0/2$.

Ejercicio 4.11 Las cuatro señales de la Figura 4.103 se utilizan para transmitir 4 símbolos igualmente probables en un sistema de comunicaciones. Se considera que dichas señales se transmiten a través de un canal gaussiano con densidad espectral de potencia $N_0/2$.

- a) Diseñe el transmisor: codificador (constelación) y modulador ($\{\phi_i(t)\}$, $i = 0, 1, \dots, N$).

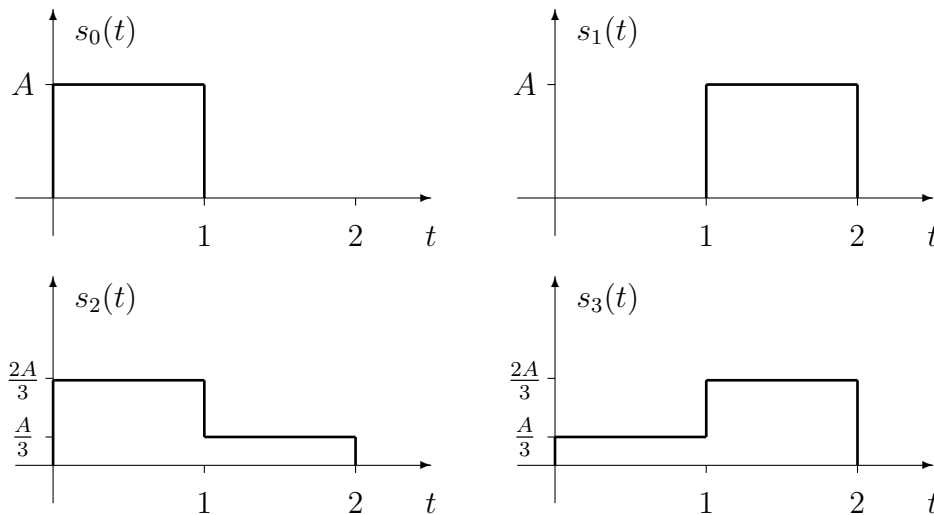


Figura 4.103: Señales para el Ejercicio 4.11.

- b) Calcule la energía media por símbolo del sistema, y realice una asignación óptima de bits a cada símbolo justificando dicha asignación (sin la justificación adecuada, la asignación no será valorada).
- c) Diseñe el receptor óptimo (demodulador + decisor) utilizando filtros adaptados causales (hay que proporcionar la expresión analítica o dibujarlos), obtenga las expresiones de la función densidad de probabilidad de la observación a la salida del demodulador condicionada a la transmisión de cada símbolo ($f_{q|A}(q|\mathbf{a}_i)$, $i = 0, 1, 2, 3$), y calcule la probabilidad de error.
- d) Si se utiliza el demodulador de la Figura 4.104, diseñe el decisor óptimo, obtenga las expresiones de la función densidad de probabilidad de la observación a la salida del demodulador condicionada a la transmisión de cada símbolo ($f_{q|A}(q|\mathbf{a}_i)$, $i = 0, 1, 2, 3$), y calcule la probabilidad de error. Compare este valor con el obtenido en el apartado anterior y explique los resultados obtenidos.

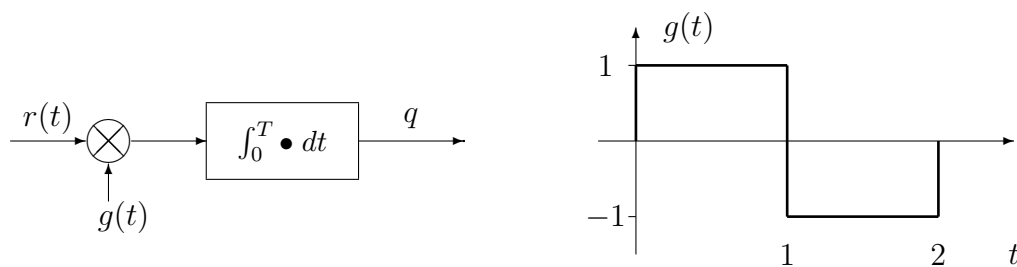


Figura 4.104: Demodulador para el Ejercicio 4.11.

Ejercicio 4.12 Se tiene un sistema de comunicaciones con un transmisor con una tasa de símbolo $R_s = 10^3$ baudios. Se asume ruido aditivo blanco, gaussiano, y densidad espectral de potencia $N_0/2$, con $N_0 = 2 \times 10^{-2}$.

- a) Una aproximación comúnmente empleada en sistemas de comunicaciones digitales para la pro-

babilidad de error de símbolo es

$$P_e \approx k \cdot Q \left(\frac{d_{min}}{2\sqrt{N_o/2}} \right), \quad (4.4)$$

donde d_{min} es la mínima distancia entre dos puntos de la constelación y k es el máximo número de símbolos que se encuentran a d_{min} de un símbolo de la constelación.

- I) Utilizando la aproximación (4.4), diseñe el codificador unidimensional óptimo del sistema de comunicaciones, con la menor energía media por símbolo, para obtener una probabilidad de error de símbolo aproximada $P_e \approx 2 \cdot 10^{-4}$ transmitiendo a una velocidad binaria $R_b = 2 \times 10^3$ bits/s.
 - II) Realice una asignación óptima de bits a cada símbolo, explicando la razón de dicha asignación, y calcule la tasa de error binaria aproximada.
- b) Si el sistema de comunicaciones utiliza el codificador y el modulador definidos en la Figura 4.105:

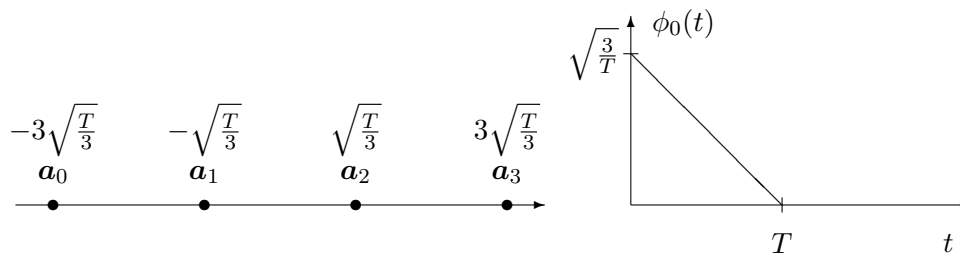


Figura 4.105: Codificador y modulador para el Ejercicio 4.12.

- I) Diseñe el demodulador óptimo utilizando un correlador y utilizando un filtro adaptado causal (en este último caso proporcione la expresión analítica de la respuesta al impulso del filtro o bien dibújela).
- II) Si por simplicidad, en lugar del demodulador óptimo se emplea un demodulador que realiza la siguiente operación sobre la señal recibida ($r(t)$)

$$q = 2 \cdot \int_0^T r(t) dt,$$

diseñe el decisor óptimo y calcule la probabilidad de error de símbolo asumiendo símbolos equiprobables. Discuta si esta probabilidad de error será mayor o menor que la obtenida con el demodulador del apartado anterior.

Ejercicio 4.13 Se va a diseñar un sistema de comunicaciones que utilizará las ocho señales de la Figura 4.106 para transmitir ocho símbolos con igual probabilidad. El canal únicamente introduce ruido, que se considerará blanco, gaussiano, estacionario y con densidad espectral de potencia $N_0/2$. Por simplicidad en los cálculos, considere $T = 2$.

- a) Realice la asignación de los bits que transporta cada una de las señales para obtener la mínima probabilidad de error de bit, y justifique la razón para elegir dicha asignación.

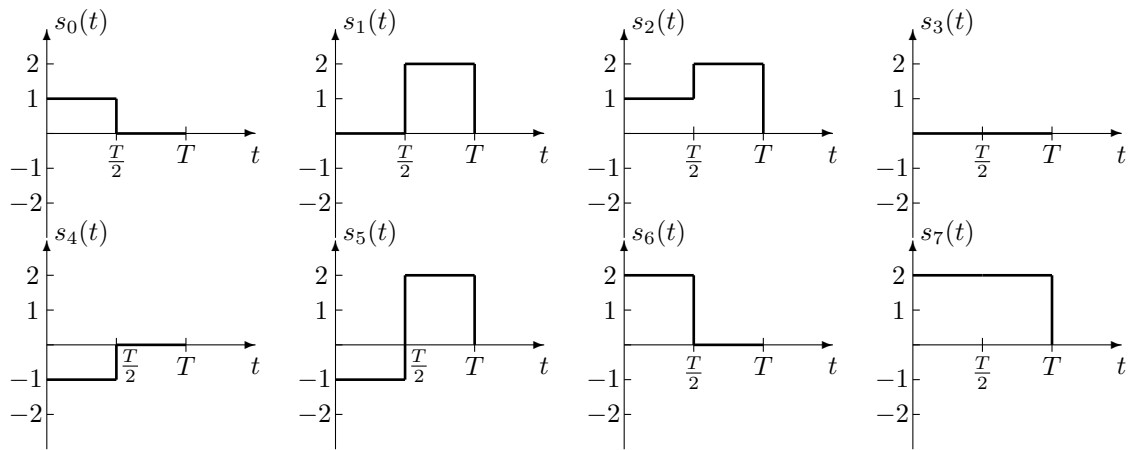


Figura 4.106: Señales para el Ejercicio 4.13.

- b) Obtenga otras 8 señales alternativas (dibuje las ocho señales o proporcione sus expresiones analíticas) con las que se pueda obtener la misma probabilidad de error que con el conjunto original, pero que requieran una mínima energía media por símbolo, y calcule esa energía media por símbolo mínima.
- c) Diseñe el receptor óptimo (demodulador + decisor), y calcule la probabilidad de error obtenida (puede hacerlo utilizando tanto el conjunto original de señales como el obtenido en la sección b), ya que en ambos casos la probabilidad de error ha de ser la misma).

Ejercicio 4.14 Un sistema de comunicaciones utiliza una constelación formada por los siguientes 4 símbolos,

$$\mathbf{a}_0 = \begin{bmatrix} +1 \\ +1 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} -1 \\ +1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} +1 \\ -1 \end{bmatrix},$$

que se transmiten con igual probabilidad, y un modulador dado por las funciones base de la Figura 4.107.

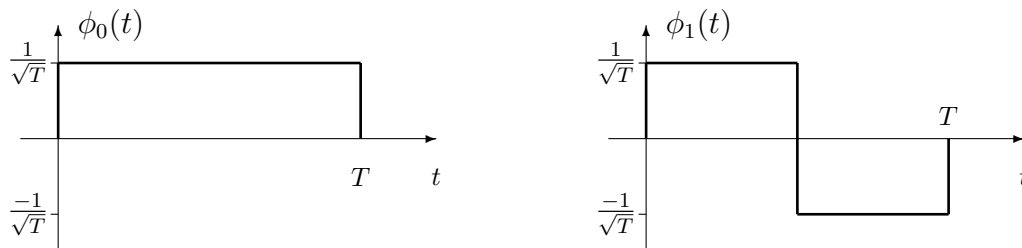


Figura 4.107: Demodulador para el Ejercicio 4.14.

Por simplicidad en los cálculos, en lo sucesivo considere $T = 1$. Considere también la transmisión sobre un canal gaussiano con densidad espectral de potencia de ruido blanco $N_0/2$.

- a) Realice la asignación binaria que minimiza la probabilidad de error de bit, justificando dicha asignación (sin la justificación adecuada, la asignación no será valorada), y calcule la velocidad de transmisión de símbolos, R_s , y la velocidad de transmisión binaria, R_b .
- b) Represente las cuatro señales utilizadas para la transmisión de cada símbolo, $s_i(t)$, $i = \{0, 1, 2, 3\}$, y la señal resultante de la transmisión de la siguiente secuencia de símbolos

$$\mathbf{A}[0] = \mathbf{a}_1, \mathbf{A}[1] = \mathbf{a}_0, \mathbf{A}[2] = \mathbf{a}_3, \mathbf{A}[3] = \mathbf{a}_0, \mathbf{A}[4] = \mathbf{a}_2.$$

- c) Diseñe el receptor óptimo (demodulador + decisor), y calcule la probabilidad de error obtenida con este receptor.
- d) Si en lugar del demodulador óptimo se utiliza el demodulador de la Figura 4.108, diseñe el decisor óptimo para ese demodulador y calcule la probabilidad de error resultante.

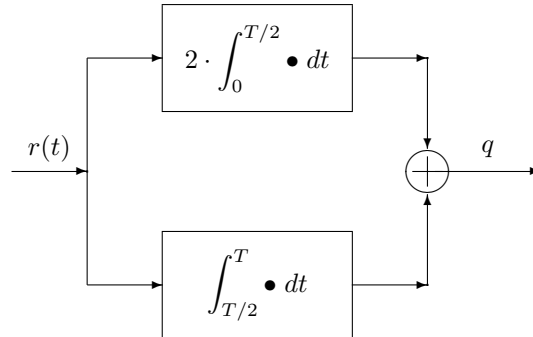
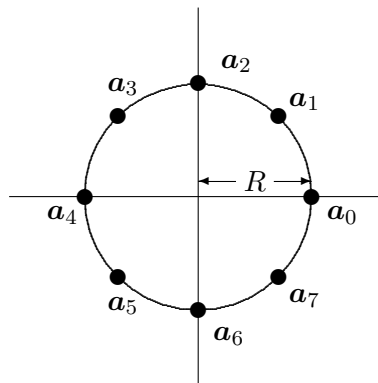


Figura 4.108: Demodulador para el Ejercicio 4.14.

Ejercicio 4.15 Un sistema digital de comunicaciones usa una constelación de 8 símbolos equiprobables como la que se muestra en la figura. Se supone que el ruido asociado a la transmisión se modela como un proceso de ruido térmico habitual, estacionario, blanco, gaussiano y con densidad espectral de potencia $S_n(j\omega) = \frac{N_0}{2}$, siendo $N_0 = 0,03$ para este caso.



- a) Si se desea que la probabilidad de error de símbolo sea aproximadamente $P_e \approx 2 \times 10^{-6}$:
- I) determine el valor del radio R requerido;
 - II) calcule la energía media por símbolo, E_s , de la constelación resultante;
 - III) realice la asignación binaria óptima justificándola (sin justificación no se considerará válida), y estime la tasa de error de bit (BER) aproximada usando dicha asignación si se asume que la relación señal a ruido es alta.
- b) Si se desea transmitir a una tasa binaria $R_b = 1500$ bits por segundo a través de un canal en banda base (inicialmente sin limitación en el ancho de banda):
- I) elija una base ortonormal apropiada para la transmisión justificando su elección (debe proporcionar las expresiones analíticas de cada elemento de la base o un dibujo de la misma, pero en cualquier caso han de estar claramente indicados los valores numéricos de amplitudes y duración de las señales, T seg.);

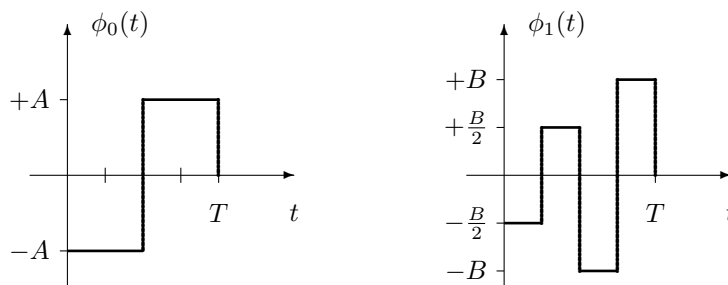
- II) diseñe un demodulador óptimo para el sistema de comunicaciones.
- c) En cuanto al decisor
 - I) diseñe el decisor óptimo para el sistema;
 - II) calcule la cota de la unión (sin despreciar ningún término ni hacer ninguna aproximación);
 - III) obtenga la cota holgada y escriba la cota resultante para la probabilidad de error en función del cociente o relación $\frac{E_s}{N_0}$.

NOTA: la distancia entre dos puntos de una circunferencia de radio R está dada por $d = 2R \sin\left(\frac{\theta}{2}\right)$, donde θ es el ángulo que forman los radios que unen dichos puntos con el centro.

Ejercicio 4.16 Un sistema de comunicaciones utiliza una constelación formada por los siguientes 4 símbolos,

$$\mathbf{a}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

que se transmiten con igual probabilidad, y un modulador dado por las funciones base de la figura



Por simplicidad en los cálculos, en lo sucesivo considere $T = 1$. Considere también la transmisión sobre un canal gaussiano con densidad espectral de potencia de ruido blanco $N_0/2$.

- a) Determine los valores de A y B , demuestre que las funciones forman una base ortonormal, y calcule la velocidad de transmisión o tasa de símbolo y la velocidad de transmisión o tasa de bit.
- b) Represente las señales $s_i(t)$ asociadas a la transmisión de cada símbolo \mathbf{a}_i , y el fragmento de la señal resultante de la transmisión de los siguientes primeros 5 símbolos de la secuencia

$$\mathbf{A}[0] = \mathbf{a}_1, \mathbf{A}[1] = \mathbf{a}_0, \mathbf{A}[2] = \mathbf{a}_3, \mathbf{A}[3] = \mathbf{a}_0, \mathbf{A}[4] = \mathbf{a}_2.$$

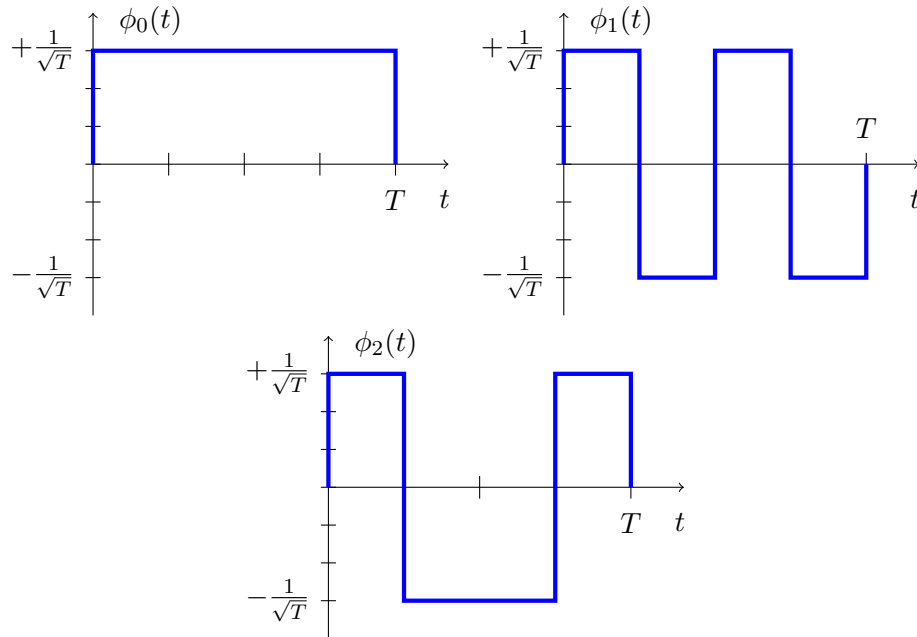
- c) Realice la asignación binaria para cada símbolo, justificando dicha asignación (sin esta justificación no se considerará la asignación como válida), y calcule la energía media por símbolo del sistema. ¿Es la constelación utilizada la más apropiada en términos del compromiso entre uso de energía y prestaciones (explique claramente por qué sí o por qué no)?
- d) Diseñe el demodulador óptimo, el decisor óptimo, y calcule la probabilidad de error de símbolo exacta del sistema.
- e) Obtenga la aproximación de la probabilidad de error de símbolo y acote dicha probabilidad de error mediante la cota de la unión y mediante la cota holgada.

4.9. Soluciones de los ejercicios

Ejercicio 4.1 Solución

a) Resultado de la aplicación del Procedimiento de Gram-Schmidt siguiendo el orden dado por los índices de las señales

1) La dimensión del espacio de señales es $N = 3$, y los tres elementos que forman la base ortonormal son los de la figura



II) La representación vectorial de las señales es

$$\mathbf{a}_0 = \begin{bmatrix} 2\sqrt{T} \\ 0 \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} -\sqrt{T} \\ 2\sqrt{T} \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -\sqrt{T} \\ -2\sqrt{T} \\ 0 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} \sqrt{T} \\ 0 \\ \sqrt{T} \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} \sqrt{T} \\ 0 \\ -\sqrt{T} \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} \sqrt{T} \\ \sqrt{T} \\ \sqrt{T} \end{bmatrix}$$

b)

$$\mathcal{E}\{\mathbf{a}_0\} = 4T, \mathcal{E}\{\mathbf{a}_1\} = 5T, \mathcal{E}\{\mathbf{a}_2\} = 5T, \mathcal{E}\{\mathbf{a}_3\} = 2T, \mathcal{E}\{\mathbf{a}_4\} = 2T, \mathcal{E}\{\mathbf{a}_5\} = 3T$$

c) La energía de la diferencia entre cada par de señales

$$\mathcal{E}\{s_0(t) - s_0(t)\} = 0, \mathcal{E}\{s_0(t) - s_1(t)\} = 13T, \mathcal{E}\{s_0(t) - s_2(t)\} = 13T$$

$$\mathcal{E}\{s_0(t) - s_3(t)\} = 2T, \mathcal{E}\{s_0(t) - s_4(t)\} = 2T, \mathcal{E}\{s_0(t) - s_5(t)\} = 3T$$

Ejercicio 4.2 Solución

En este caso, al ser el sistema binario, la probabilidad de error de símbolo y de bit coinciden

a) $A = 2,6163$

b) $A = 3,3588$

Ejercicio 4.3 Solución

a) El umbral que define el decisor óptimo y las regiones de decisión son

$$q_u = \frac{A^2 - N_0 \ln 2}{2A}, I_0 = \{q : q \geq q_u\}, I_1 = \{q : q < q_u\}$$

b)

$$P_e = \frac{1}{3}Q\left(\frac{A - q_u}{\sqrt{N_0/2}}\right) + \frac{2}{3}Q\left(\frac{q_u}{\sqrt{N_0/2}}\right)$$

Ejercicio 4.4 Solución

a) $E_s = 1$ J.

b) La constelación alternativa, que tiene energía $E'_s = \frac{8}{9}$ J, es

$$\mathbf{a}'_0 = \begin{bmatrix} 0 \\ \frac{2}{3} \end{bmatrix}, \mathbf{a}'_1 = \begin{bmatrix} 1 \\ -\frac{1}{3} \end{bmatrix}, \mathbf{a}'_2 = \begin{bmatrix} -1 \\ -\frac{1}{3} \end{bmatrix}$$

c) La cota de la unión es

$$P_e \leq \frac{4}{3}Q\left(\frac{1}{\sqrt{N_0}}\right) + \frac{2}{3}Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

La cota hogada es

$$P_e \leq 2Q\left(\frac{1}{\sqrt{N_0}}\right)$$

Ejercicio 4.5 Solución

a) $E_s = \frac{17}{9}$ J.

b) La cota holgada vale

$$P_e \leq 2Q\left(\frac{\sqrt{3}}{2\sqrt{N_0/2}}\right)$$

Ejercicio 4.6 Solución

a) Si $\Lambda(x)$ denota la función triangular, con soporte entre -1 y $+1$ y altura unidad

$$\Lambda(x) = \begin{cases} 1 + x, & \text{si } -1 \leq x < 0 \\ 1 - x, & \text{si } 0 \leq x \leq 1 \end{cases}$$

las expresiones de las funciones densidad de probabilidad condicionales son las siguientes

$$f_{q|A}(q|a_0) = \Lambda(x), f_{q|A}(q|a_1) = \Lambda(x - 1), f_{q|A}(q|a_2) = \Lambda(x - 2), f_{q|A}(q|a_3) = \Lambda(x - 3)$$

b) Regiones de decisión para símbolos equiprobables

$$I_0 = \left\{q : q < \frac{1}{2}\right\}, I_1 = \left\{q : \frac{1}{2} \leq q < \frac{3}{2}\right\}, I_2 = \left\{q : \frac{3}{2} \leq q < \frac{8}{3}\right\}, I_3 = \left\{q : q \geq \frac{8}{3}\right\}$$

c) Probabilidad de error para símbolos equiprobables

$$P_e = \frac{3}{16}.$$

d) Regiones de decisión para símbolos no equiprobables

$$I_0 = \left\{ q : q < \frac{1}{3} \right\}, I_1 = \left\{ q : \frac{1}{3} \leq q < \frac{3}{2} \right\}, I_2 = \left\{ q : \frac{3}{2} \leq q < \frac{5}{2} \right\}, I_3 = \left\{ q : q \geq \frac{5}{2} \right\}$$

e) Probabilidad de error para símbolos no equiprobables

$$P_e = \frac{7}{36}.$$

Ejercicio 4.7 Solución

$$\frac{S}{N} \Big|_q = \frac{2E_b}{N_0}$$

Ejercicio 4.8 Solución

a) $P_e = Q \left(\frac{A\sqrt{T}}{\sqrt{N_0/2v}} \right)$

b)

$$P_e = pQ \left(\frac{A\sqrt{T} - q_u}{\sqrt{N_0/2}} \right) + (1 - p)Q \left(\frac{A\sqrt{T} + q_u}{\sqrt{N_0/2}} \right)$$

donde

$$q_u = \frac{N_0 \ln \frac{1-p}{p}}{4A\sqrt{T}}$$

Ejercicio 4.9 Solución

a) $\frac{E_b}{N_0} \approx 9,68$ ($\frac{E_b}{N_0}$ (dB) $\approx 9,85$ dB)

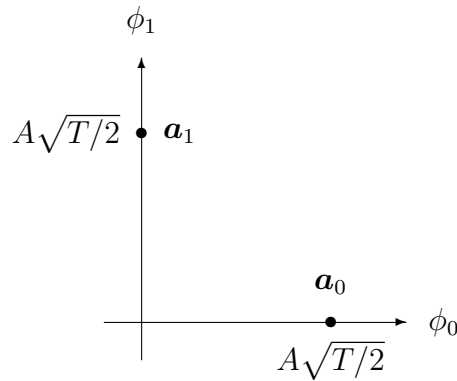
b) $\frac{E_b}{N_0} \approx 25,31$ ($\frac{E_b}{N_0}$ (dB) $\approx 14,03$ dB)

c) $\frac{E_b}{N_0} \approx 439,928$ ($\frac{E_b}{N_0}$ (dB) $\approx 26,43$ dB)

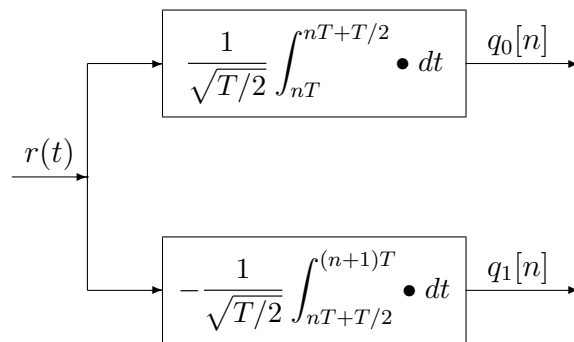
d) Transmitir a una mayor velocidad binaria para la misma tasa de símbolo supone utilizar constelaciones más densas, y con constelaciones más densas es necesaria una mayor energía para obtener las mismas prestaciones.

Ejercicio 4.10 Solución

a) La constelación es la de la figura



El demodulador, mediante correladores es equivalente al dado en la figura



El decisor sigue la regla

$$q_0[n] \underset{\hat{A}[n]=a_1}{\overset{\hat{A}[n]=a_0}{\geq}} q_1[n].$$

O lo que es lo mismo, las regiones de decisión son

$$I_0 = \{\mathbf{q} : q_0 > q_1\}, I_1 = \{\mathbf{q} : q_0 \leq q_1\}$$

La probabilidad de error es

$$P_e = Q\left(\frac{d(\mathbf{a}_0, \mathbf{a}_1)}{2\sqrt{N_0/2}}\right) = Q\left(\frac{A\sqrt{T}}{2\sqrt{N_0/2}}\right)$$

b) El decisor viene dado por las siguientes regiones de decisión

$$I_0 = \{\mathbf{q} : q_0 + 2\alpha A\sqrt{T/2} > q_1\}, I_1 = \{\mathbf{q} : q_0 + 2\alpha A\sqrt{T/2} \leq q_1\}$$

La probabilidad de error es ahora

$$P_e = Q\left(\frac{\alpha A\sqrt{T}}{2\sqrt{N_0/2}}\right),$$

c) El decisor óptimo es en este caso sigue la regla

$$q' \underset{a_1}{\overset{a_0}{\geq}} 0$$

o lo que es lo mismo, las regiones de decisión son

$$I_0 = \{q' : q' > 0\}, I_1 = \{q' : q' \leq 0\}$$

La probabilidad de error es

$$P_e = Q\left(\frac{A\sqrt{T}}{2\sqrt{N_0/2}}\right).$$

Esta probabilidad de error coincide con la del apartado a). Esto es así porque el nuevo demodulador lo que hace es girar la constelación para alinearla sobre el nuevo eje q' , y tiene un factor de escala \sqrt{T} que afecta tanto a la señal como al ruido.

Ejercicio 4.11 Solución

a) El modulador está dado por las dos funciones base

$$\phi_0(t) = \frac{s_0(t)}{A}, \phi_1(t) = \frac{s_1(t)}{A}$$

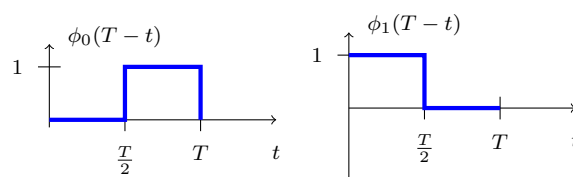
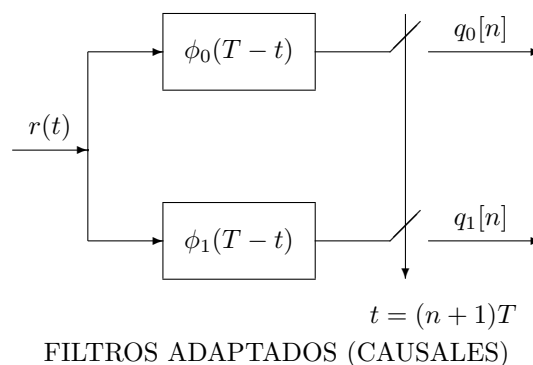
y la constelación es

$$\mathbf{a}_0 = \begin{bmatrix} A \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} 0 \\ A \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} \frac{2A}{3} \\ \frac{A}{3} \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} \frac{A}{3} \\ \frac{2A}{3} \end{bmatrix}$$

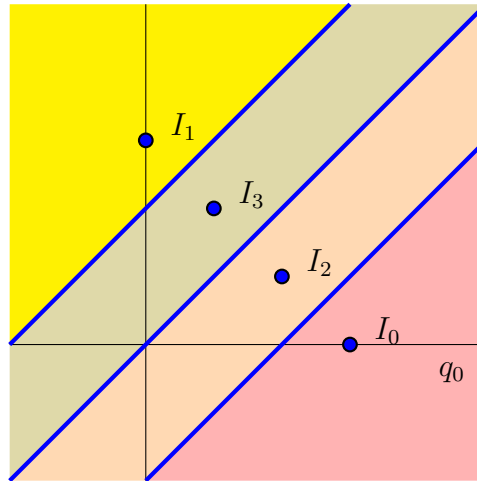
b) $E_s = \frac{7}{9}A^2$ J. La asignación de bits debe seguir la regla de codificación de Gray. En ejemplo (no es el único posible) sería

$$\mathbf{a}_0 \equiv 10, \mathbf{a}_1 \equiv 00, \mathbf{a}_2 \equiv 11, \mathbf{a}_3 \equiv 01$$

c) Demodulador



Decisor



Expresiones de la distribución condicional de la observación (general)

$$f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_i) = \mathcal{N}^N \left(\mathbf{a}_i, \frac{N_0}{2} \right) = \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{q}-\mathbf{a}_i\|^2}{N_0}}$$

Expresiones particularizadas

$$f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_0) = \frac{1}{\pi N_0} e^{-\frac{(q_0-A)^2+q_1^2}{N_0}}, \quad f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_1) = \frac{1}{\pi N_0} e^{-\frac{q_0^2+(q_1-A)^2}{N_0}}$$

$$f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_2) = \frac{1}{\pi N_0} e^{-\frac{(q_0-2A/3)^2+(q_1-A/3)^2}{N_0}}, \quad f_{\mathbf{q}|\mathbf{A}}(\mathbf{q}|\mathbf{a}_3) = \frac{1}{\pi N_0} e^{-\frac{(q_0-A/3)^2+(q_1-2A/3)^2}{N_0}}$$

Probabilidad de error

$$P_e = \frac{3}{2} Q \left(\frac{A}{3\sqrt{N_0}} \right)$$

d) Decisor

$$I_0 = \left\{ q : q > A \frac{\sqrt{2}}{3} \right\}, \quad I_1 = \left\{ q : q < -A \frac{\sqrt{2}}{3} \right\}$$

$$I_2 = \left\{ q : 0 < q \leq A \frac{\sqrt{2}}{3} \right\}, \quad I_3 = \left\{ q : -A \frac{\sqrt{2}}{3} \leq q < 0 \right\}$$

Distribuciones condicionales de la observación

$$f_{q|\mathbf{A}}(q|\mathbf{a}_0) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(q-A/\sqrt{2})^2}{N_0}}, \quad f_{q|\mathbf{A}}(q|\mathbf{a}_1) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(q+A/\sqrt{2})^2}{N_0}}$$

$$f_{q|\mathbf{A}}(q|\mathbf{a}_2) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(q-A/(3\sqrt{2}))^2}{N_0}}, \quad f_{q|\mathbf{A}}(q|\mathbf{a}_3) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(q+A/(3\sqrt{2}))^2}{N_0}}$$

Ejercicio 4.12 Solución

a) En el primer apartado:

1) El codificador está dado por cuatro símbolos con coordenadas

$$\mathbf{a}_0 = -1,05, \quad \mathbf{a}_1 = -0,35, \quad \mathbf{a}_2 = +0,35, \quad \mathbf{a}_3 = +1,05.$$

II) La asignación tiene que seguir la regla de la codificación de Gray, es decir, entre símbolos a mínima distancia la asignación binaria debe diferir únicamente en un bit. En ejemplo (no el único posible) de tal asignación sería

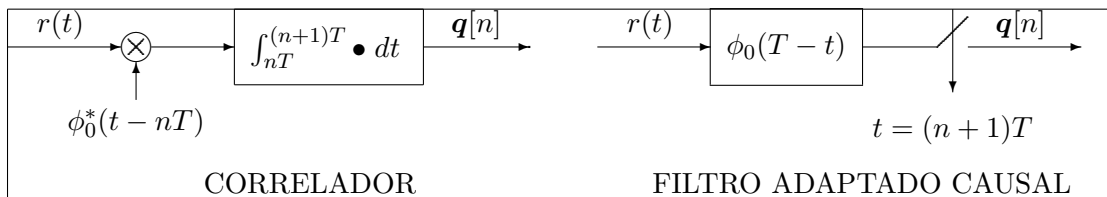
$$\mathbf{a}_0 \equiv 00, \mathbf{a}_1 \equiv 01, \mathbf{a}_2 \equiv 11, \mathbf{a}_3 \equiv 10.$$

La probabilidad de error de bit aproximada, para relaciones señal a ruido suficientemente altas es

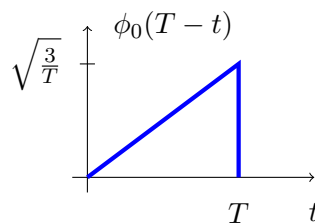
$$BER \approx 10^{-4}.$$

b) En el segundo apartado

i) Estructura de los demoduladores por correlación y mediante filtros adaptados



La respuesta del filtro adaptado es



II) El decisor viene dado por las regiones

$$I_0 = \{q : q < -2T\}, I_1 = \{q : -2T \leq q < 0\}, I_2 = \{q : 0 \leq q < +2T\}, I_3 = \{q : q \geq +2T\}$$

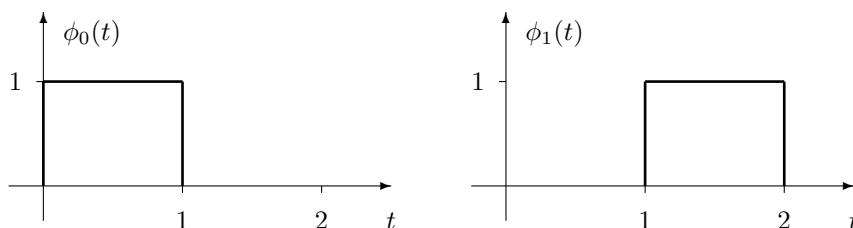
La probabilidad de error es

$$P_e = \frac{3}{2}Q\left(\sqrt{\frac{T}{2N_0}}\right).$$

Las prestaciones obtenidas con este receptor nunca podrán ser mejores que las del receptor óptimo, que por definición es el que proporciona la mínima probabilidad de error de símbolo. Por tanto, si se utiliza un demodulador diferente de los habituales basados en correladores o filtros adaptados, como en este caso, las prestaciones serán en el mejor de los casos iguales (si el receptor hace una operación equivalente) y en general serán peores (si la operación que realiza no es la equivalente a un demodulador óptimo).

Ejercicio 4.13 Solución

La base y la constelación del sistema son

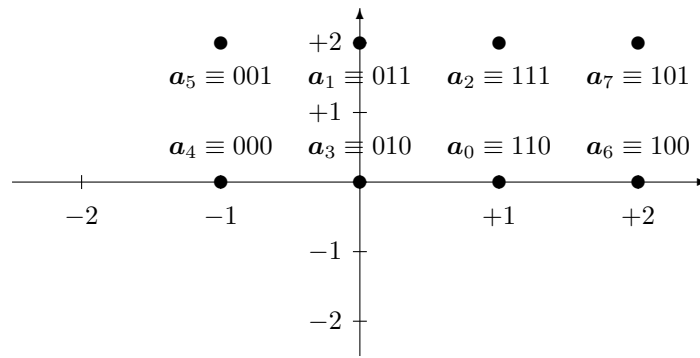


$$\mathbf{a}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{a}_4 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{a}_6 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \mathbf{a}_7 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

a) La asignación binaria tiene que seguir una codificación de Gray, ya que esta es la que minimiza la probabilidad de error de bit para una constelación. Un ejemplo de este tipo de codificación, en la que símbolos a mínima distancia difieren en un único bit en la asignación, sería

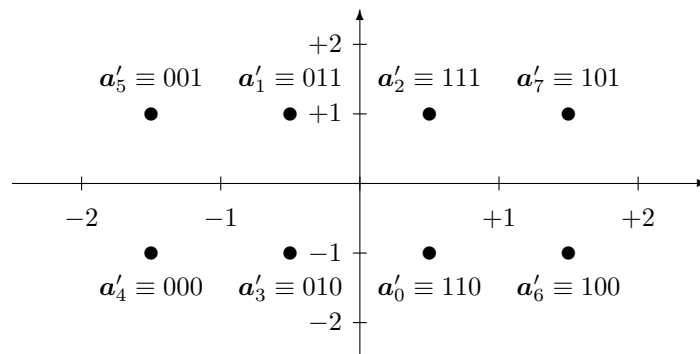
$$\mathbf{a}_0 \equiv 110, \mathbf{a}_1 \equiv 011, \mathbf{a}_2 \equiv 111, \mathbf{a}_3 \equiv 010, \mathbf{a}_4 \equiv 000, \mathbf{a}_5 \equiv 001, \mathbf{a}_6 \equiv 100, \mathbf{a}_7 \equiv 101$$



b) Los símbolos de la constelación modificada serían:

$$\mathbf{a}'_0 = \begin{bmatrix} +1/2 \\ -1 \end{bmatrix}, \mathbf{a}'_1 = \begin{bmatrix} -1/2 \\ +1 \end{bmatrix}, \mathbf{a}'_2 = \begin{bmatrix} +1/2 \\ +1 \end{bmatrix}, \mathbf{a}'_3 = \begin{bmatrix} -1/2 \\ -1 \end{bmatrix}$$

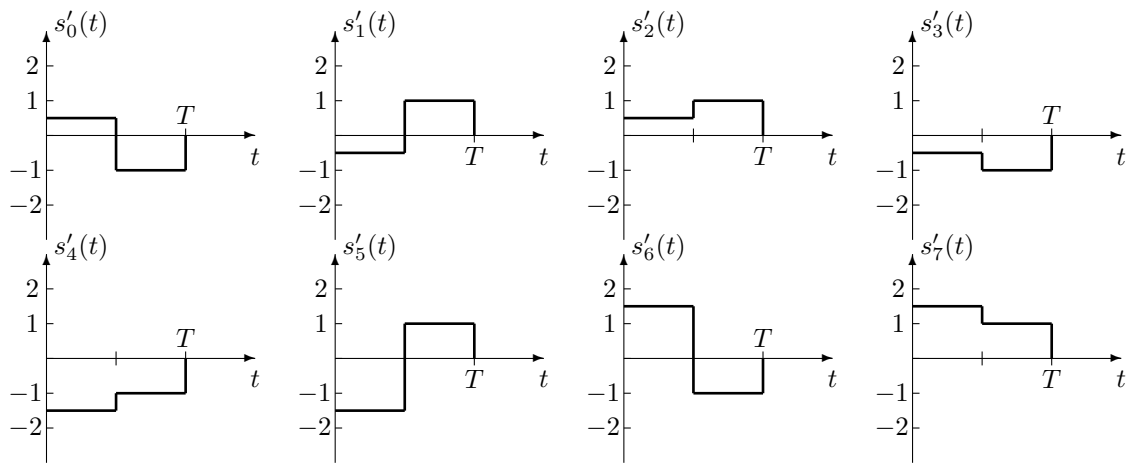
$$\mathbf{a}'_4 = \begin{bmatrix} -3/2 \\ -1 \end{bmatrix}, \mathbf{a}'_5 = \begin{bmatrix} -3/2 \\ +1 \end{bmatrix}, \mathbf{a}'_6 = \begin{bmatrix} +3/2 \\ -1 \end{bmatrix}, \mathbf{a}'_7 = \begin{bmatrix} +3/2 \\ +1 \end{bmatrix}$$



Las nuevas señales tendrían como expresión analítica

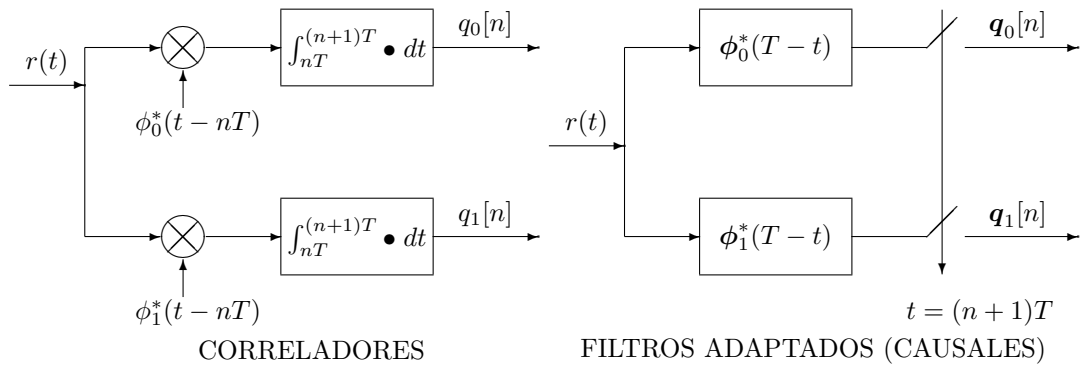
$$s'_i(t) = a'_{i,0} \cdot \phi_0(t) + a'_{i,1} \cdot \phi_1(t),$$

y se pueden ver representadas en la figura

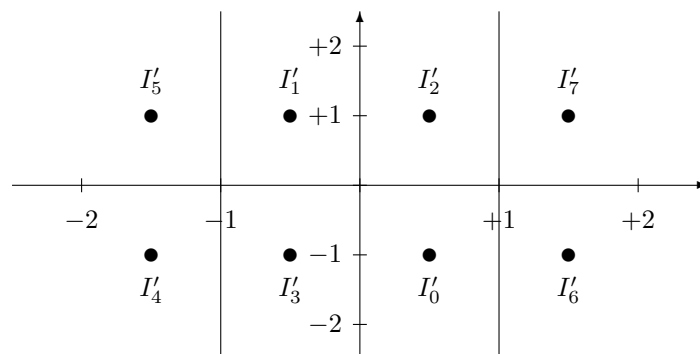


$$E_s = \frac{9}{4}$$

c) El demodulador óptimo está basado en filtros adaptados o en correladores con las funciones que forman la base ortonormal, cualquiera de las dos opciones es válida.



El decisor óptimo es

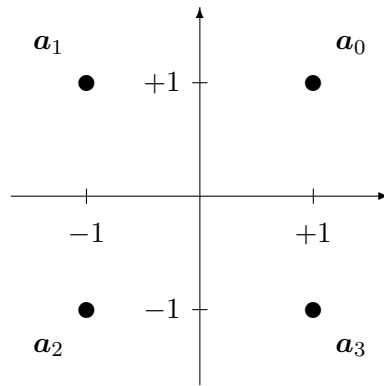


$$P_e = \frac{3}{2} \cdot Q\left(\frac{1}{2\sqrt{N_0/2}}\right) + Q\left(\frac{1}{\sqrt{N_0/2}}\right) - \frac{3}{2} \cdot Q\left(\frac{1}{2\sqrt{N_0/2}}\right) \cdot Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

Ejercicio 4.14 Solución

a) La asignación binaria, para minimizar la probabilidad de error de bit, ha de ser una codificación de Gray, ya que de esta forma los errores más probables a nivel de símbolo suponen un único error a nivel de bit.

En este caso la constelación es la que presenta la figura



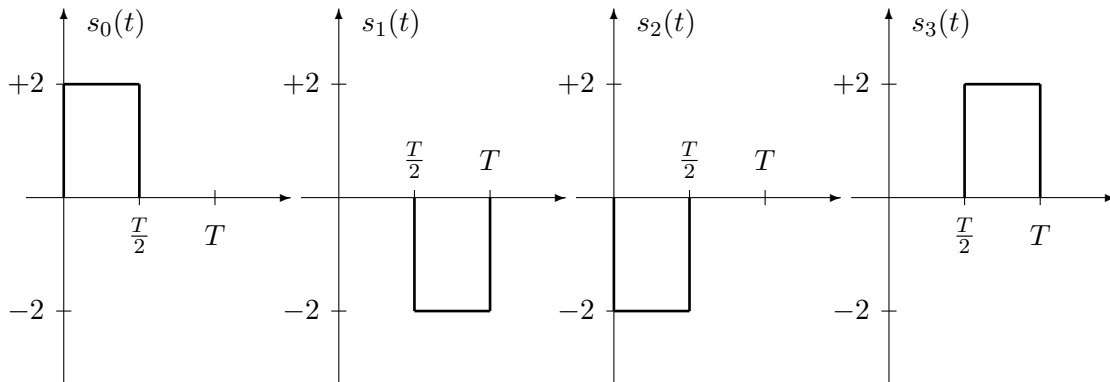
Una posible asignación de Gray (no es la única posible) sería

$$\mathbf{a}_0 = 00, \mathbf{a}_1 = 01, \mathbf{a}_2 = 11, \mathbf{a}_3 = 10.$$

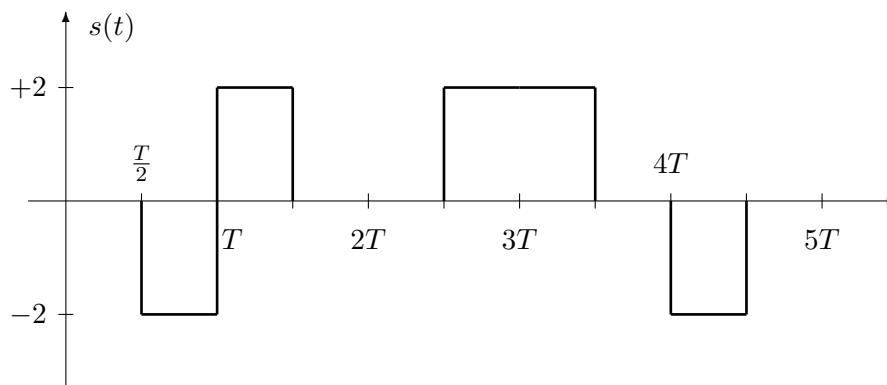
Las velocidades de símbolo y binaria son

$$R_s = \frac{1}{T} = 1 \text{ símbolos/s (baudios)}, \text{ y } R_b = 2 \text{ bits/s.}$$

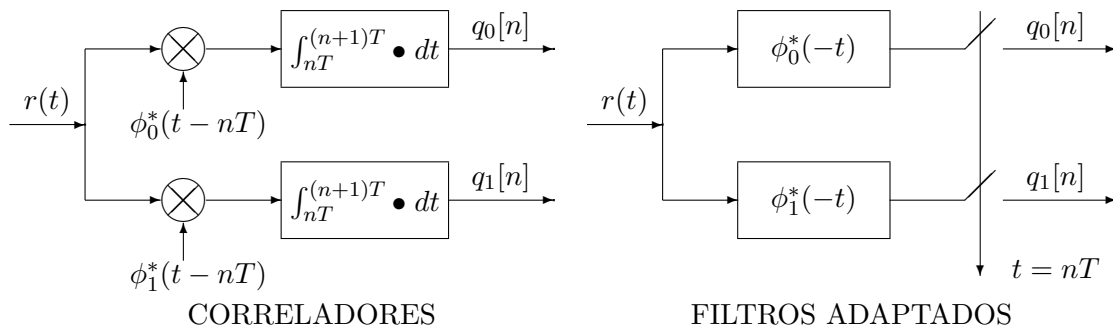
b) Las cuatro señales se muestran en la figura



La señal resultante de transmitir la secuencia dada es



c) Para el demodulador óptimo se puede utilizar cualquiera de las dos estructuras típicas, mediante correladores, o mediante filtros adaptados.



Decisor: dado por las siguientes regiones de decisión

$$\mathbf{q} \in I_0 \text{ si } \begin{cases} q_0 \geq 0 \\ q_1 \geq 0 \end{cases}, \mathbf{q} \in I_1 \text{ si } \begin{cases} q_0 < 0 \\ q_1 \geq 0 \end{cases}, \mathbf{q} \in I_2 \text{ si } \begin{cases} q_0 < 0 \\ q_1 < 0 \end{cases}, \mathbf{q} \in I_3 \text{ si } \begin{cases} q_0 \geq 0 \\ q_1 < 0 \end{cases}.$$

Probabilidad de error

$$P_e = 2 \cdot Q\left(\frac{1}{\sqrt{N_0/2}}\right) - \left(Q\left(\frac{1}{\sqrt{N_0/2}}\right)\right)^2.$$

d) El decisor óptimo es

$$q \in I_0 \text{ si } q \geq \frac{3}{2}, q \in I_1 \text{ si } -\frac{3}{2} < q \leq 0, q \in I_2 \text{ si } q \leq \frac{3}{2}, q \in I_3 \text{ si } 0 < q < \frac{3}{2},$$

Finalmente, la probabilidad de error es

$$P_e = Q\left(\frac{1}{\sqrt{5N_0}}\right) + \frac{1}{2} \cdot Q\left(\frac{2}{\sqrt{5N_0}}\right).$$

Ejercicio 4.15 Solución

a) En este caso se busca una probabilidad de error aproximada de $P_e \approx 2 \times 10^{-6}$

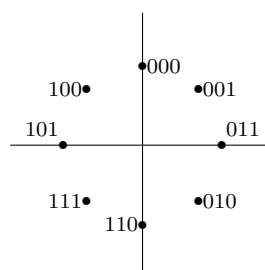
i) El radio vale

$$R = 1,5202.$$

ii) La energía media por símbolo

$$E_s = R^2 = 2,311 \text{ J.}$$

iii) Para realizar la asignación binaria ($m = \log_2 M = 3$ bits/símbolo) hay que utilizar una codificación de Gray, de modo que símbolos que están a mínima distancia tengan una asignación que difiera únicamente en un bit. Así, como la mayor parte de errores se produce con símbolos a mínima distancia, se minimiza la probabilidad de error de bit.



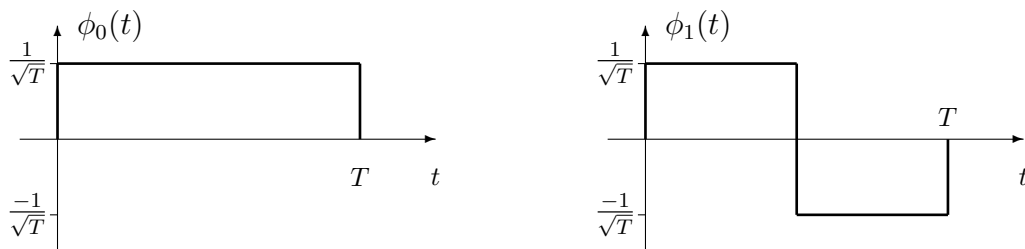
$$\mathbf{a}_0 \equiv 011, \mathbf{a}_1 \equiv 001, \mathbf{a}_2 \equiv 000, \mathbf{a}_3 \equiv 100, \mathbf{a}_4 \equiv 101, \mathbf{a}_5 \equiv 111, \mathbf{a}_6 \equiv 110, \mathbf{a}_7 \equiv 010.$$

Cuando se usa una codificación de Gray, si la relación señal a ruido es alta

$$BER \approx \frac{1}{m} \cdot P_e = \frac{1}{3} \cdot P_e = \frac{2}{3} \times 10^{-6}.$$

b) En este caso se busca una transmisión a la velocidad binaria $R_b = 1500$ bits/s.

i) La dimensión del espacio de señales, dada esta constelación, es $N = 2$, por lo que hay que encontrar dos funciones que formen una base ortonormal (su producto escalar es nulo y cada una tiene energía unidad). Para una transmisión en banda base, formas de onda de tipo pulsos cuadrados son apropiadas, ya que su espectro está en banda base. Hay varias opciones. Una posibilidad (no la única) sería la base dada en la figura



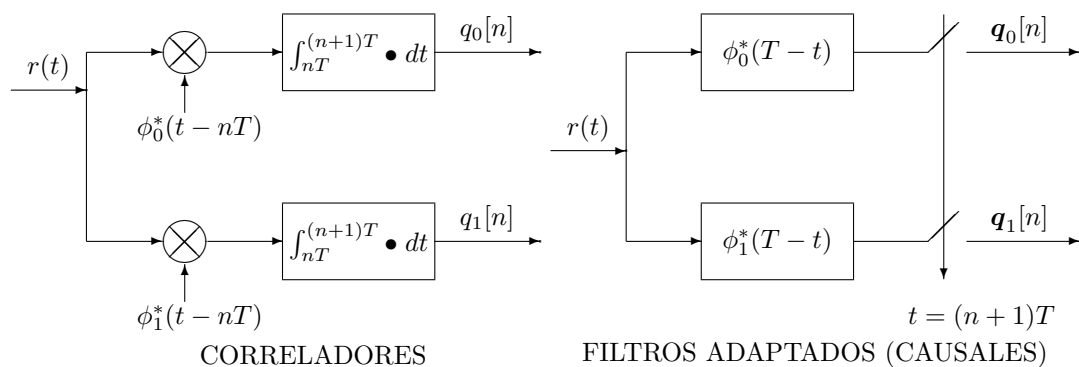
Para tenerlas completamente definidas, hay que determinar el valor de T , y este está relacionado con la velocidad de transmisión, ya que la velocidad de símbolo es $R_s = \frac{1}{T}$ baudios. La velocidad de símbolo está relacionada con la tasa binaria a través del número de bits por símbolo

$$R_s = \frac{R_b}{m} = \frac{1500}{3} = 500 \text{ baudios.}$$

Esto implica que el tiempo de símbolo es

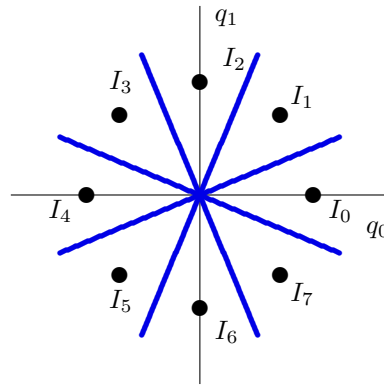
$$T = \frac{1}{R_s} = \frac{1}{500} = 0,002 \text{ segundos} = 2 \text{ ms.}$$

ii) El demodulador óptimo puede ser el basado en filtros adaptados causales o en correladores con las funciones que forman la base ortonormal, cualquiera de las dos opciones es válida.



En este caso, como las funciones base son reales, los conjugados no son relevantes.

c) En esta sección se trabaja sobre el decisor



i) Para el diseño del decisor, al ser los símbolos equiprobables y el ruido gaussiano, se aplica el criterio de mínima distancia euclídea: la región de decisión de un símbolo son aquellos puntos del espacio más cercanos a ese símbolo que al resto de símbolos de la constelación, que en este caso serían las que se muestra en la figura anterior.

ii) La cota de la unión se obtiene como

$$P_e \leq \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \sum_{\substack{j=0 \\ j \neq i}}^{M-1} P_e(\mathbf{a}_i, \mathbf{a}_j) = \sum_{i=0}^{M-1} p_A(\mathbf{a}_i) \sum_{\substack{j=0 \\ j \neq i}}^{M-1} Q \left(\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{2\sqrt{N_o/2}} \right).$$

Para esta constelación, cada símbolo tiene al resto de símbolos a las siguientes distancias:

- Dos símbolos a distancia $d_{min} = 1,1635$ (ángulo $\theta = 45^\circ$)
- Dos símbolos a distancia $d_a = 2,15$ (ángulo $\theta = 90^\circ$)
- Dos símbolos a distancia $d_b = 2,8$ (ángulo $\theta = 135^\circ$)
- Un símbolo a distancia $d_c = 2R = 3,04$ (ángulo $\theta = 180^\circ$)

Por tanto, como la cota para la probabilidad de error condicional para cada símbolo es la misma

$$P_e \leq 2 \cdot Q \left(\frac{d_{min}}{2\sqrt{N_o/2}} \right) + 2 \cdot Q \left(\frac{d_a}{2\sqrt{N_o/2}} \right) + 2 \cdot Q \left(\frac{d_b}{2\sqrt{N_o/2}} \right) + Q \left(\frac{d_c}{2\sqrt{N_o/2}} \right) = 2,0345 \times 10^{-6}.$$

d) La cota holgada es

$$P_e \leq 7 \cdot Q \left(\frac{1,1635}{2\sqrt{N_o/2}} \right) = 7,12 \times 10^{-6}.$$

Expresada en función de E_s/N_0 vale

$$P_e \leq 7 \cdot Q \left(0,5412 \cdot \sqrt{\frac{E_s}{N_0}} \right).$$

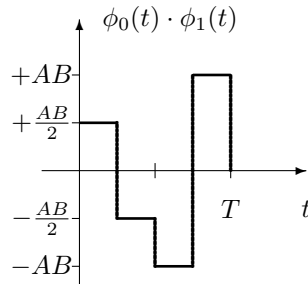
Ejercicio 4.16 Solución

a) $A = 1, B = \sqrt{8/5} = 1,2649$.

Dado que con estos valores la energía de ambas señales es uno, lo único que falta para demostrar que forman una base ortonormal es que el producto escalar entre ambas señales vale cero, que para señales reales se define como

$$\langle \phi_0(t), \phi_1(t) \rangle = \int_{-\infty}^{+\infty} \phi_0(t) \cdot \phi_1(t) dt = 0$$

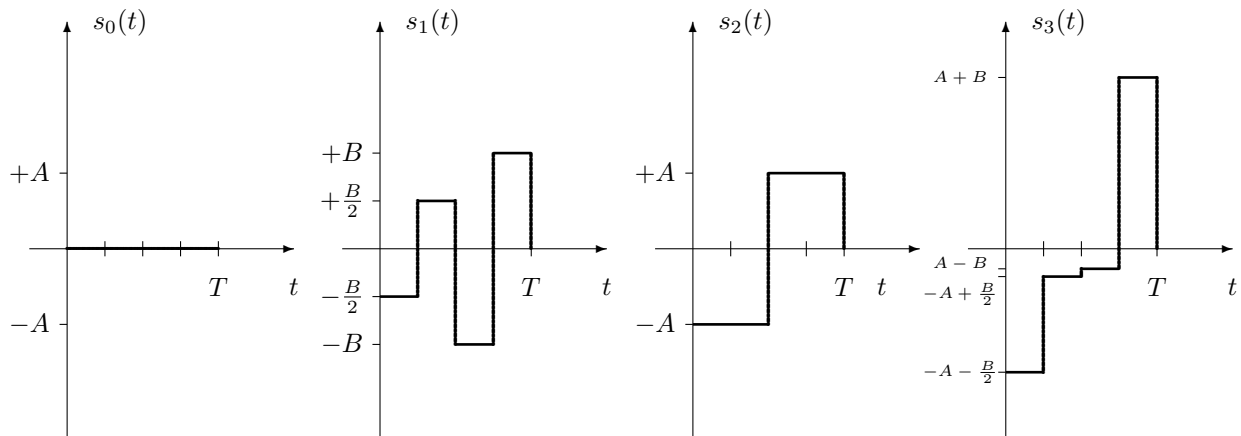
El integrando, la señal resultante del producto entre $\phi_0(t) \cdot \phi_1(t)$, es el que se muestra en la figura, y se puede ver claramente que su integral es cero.



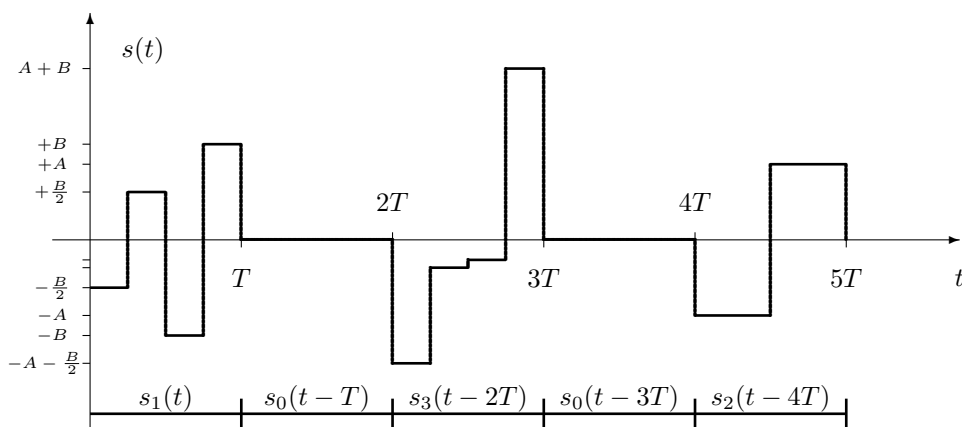
b)

$$s_0(t) = 0, \quad s_1(t) = \phi_1(t), \quad s_2(t) = \phi_0(t), \quad s_3(t) = \phi_0(t) + \phi_1(t),$$

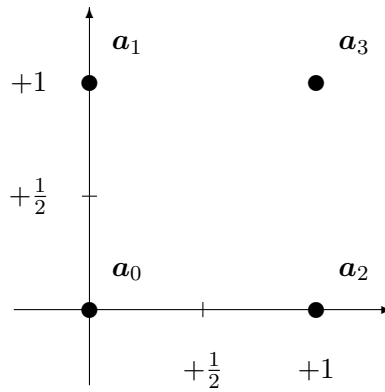
lo que da lugar a las cuatro señales de la figura



La señal asociada a la secuencia es



c) Para realizar la asignación binaria hay que utilizar una codificación de Gray, de modo que símbolos que están a mínima distancia tengan una asignación que difiera únicamente en un bit. Esta es la asignación que minimiza la probabilidad de error de bit (BER).



Una posible asignación de Gray es

$$\mathbf{a}_0 = 00, \mathbf{a}_1 = 01, \mathbf{a}_2 = 10, \mathbf{a}_3 = 11.$$

Velocidades de transmisión de símbolo y de bit $R_s = \frac{1}{T} = 1$ símbolos/s (baudios) y $R_b = 2$ bits/s.

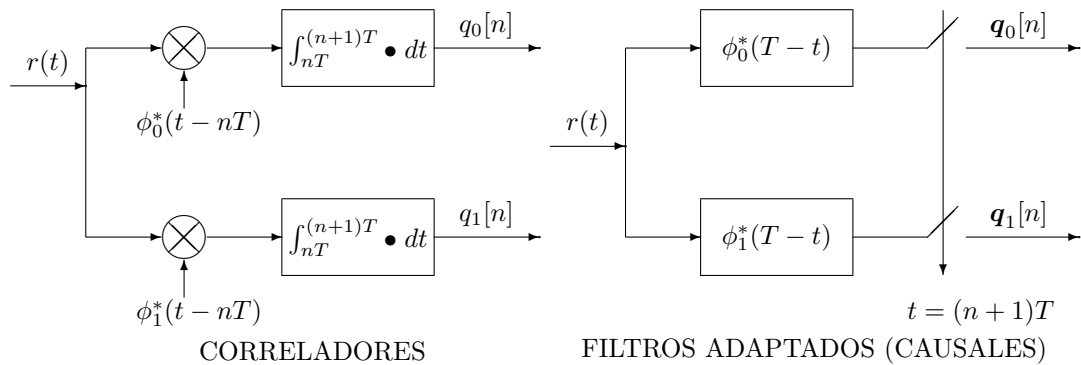
La energía media por símbolo

$$E_s = 1.$$

Esta constelación no es la más apropiada, ya que no cumple la condición de que la media de la constelación sea nula, que es la que garantiza una mínima energía media por símbolo para las distancias relativas dadas por la constelación, y en este caso la media es

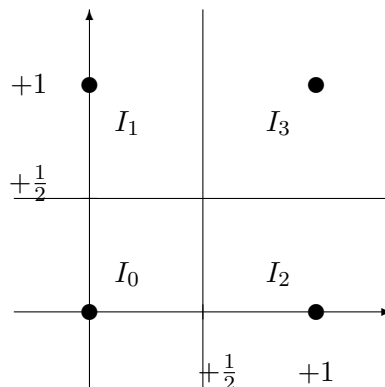
$$E[\mathbf{a}_i] = \begin{bmatrix} +\frac{1}{2} \\ +\frac{1}{2} \end{bmatrix}.$$

- d) El demodulador óptimo está basado en filtros adaptados o en correladores con las funciones que forman la base ortonormal, cualquiera de las dos opciones es válida.



En este caso, como las funciones base son reales, los conjugados no son relevantes.

El decisor óptimo es el que se muestra en la figura



La probabilidad de error vale

$$P_e = 2 \cdot Q\left(\frac{1}{2\sqrt{N_0/2}}\right) - \left[Q\left(\frac{1}{2\sqrt{N_0/2}}\right)\right]^2$$

e) La aproximación habitual para la probabilidad de error es

$$P_e \approx 2 \cdot Q\left(\frac{1}{2\sqrt{N_0/2}}\right).$$

En cuanto a las cotas, la cota de la unión es

$$P_e \leq 2 \cdot Q\left(\frac{1}{2\sqrt{N_0/2}}\right) + Q\left(\frac{\sqrt{2}}{2\sqrt{N_0/2}}\right) = 2 \cdot Q\left(\frac{1}{\sqrt{2N_0}}\right) + Q\left(\frac{1}{\sqrt{N_0}}\right).$$

Y la cota holgada vale

$$P_e \leq 3 \cdot Q\left(\frac{1}{2\sqrt{N_0/2}}\right).$$

Capítulo 5

Límites fundamentales en los sistemas de comunicaciones digitales

El principal objetivo de un sistema de comunicaciones es la transmisión fiable de información. Una fuente produce la información, y el propósito del sistema de comunicaciones es transmitir la salida de la fuente al destino de la información, como se ilustra en la Figura 5.1.

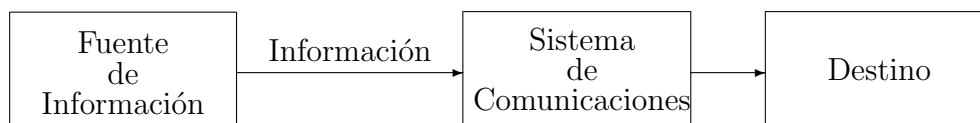


Figura 5.1: Esquema funcional simplificado de un sistema de comunicaciones.

Existe una gran variedad de fuentes de información, y cada una de ellas produce información de naturaleza distinta. Algunos tipos de fuentes de información, con el correspondiente tipo de información que generan, podrían ser, por ejemplo

- Radiodifusión de radio: fuente de voz o de audio.
- Radiodifusión de TV: fuente de vídeo.
- Transmisión de FAX: imagen fija.
- Comunicación entre PC's: fuente binaria o ASCII.
- Almacenamiento de datos: fuente binaria.

A la hora de analizar un sistema de comunicaciones y medir sus prestaciones o sus límites, es posible plantearse cuantificar la cantidad de información transmitida, o la fiabilidad en la transmisión de información. En los capítulos anteriores se han estudiado dos tipos de sistemas de comunicaciones: sistemas de comunicaciones analógicos y sistemas de comunicaciones digitales. En cuanto a sistemas de comunicaciones analógicos, ya se ha estudiado el compromiso entre prestaciones y consumo de recursos de las distintas variantes de modulaciones, especialmente potencia y ancho de banda. En este capítulo se considerarán únicamente los sistemas de comunicaciones digitales, y el objetivo será estudiar los límites fundamentales que se pueden alcanzar en la transmisión fiable

de información con este tipo de sistemas. Conviene recordar que la utilización de un sistema de comunicaciones digital no implica la exclusión de fuentes analógicas. La transmisión digital permite en general una mayor inmunidad frente a ruido, mayor flexibilidad, la aplicación de encriptado y facilita la implementación de los equipos, lo que ha hecho que los sistemas digitales predominen sobre los analógicos. Pero es posible la transmisión de información de naturaleza analógica a través de un sistema digital, realizando una conversión analógico/digital en el lado del transmisor y la correspondiente conversión digital/analógico en el lado del receptor.

Al estudiar las prestaciones de sistemas de comunicaciones digitales en el capítulo anterior, se ha visto que al aumentar la velocidad binaria de transmisión mediante el aumento del número de bits por símbolo de la constelación, para una cierta energía media por símbolo fija, la probabilidad de error aumenta y por tanto las prestaciones se degradan. Este hecho se consideraba algo ineludible hasta que en los años cuarenta Claude Shannon demostró que es posible transmitir con una probabilidad de error tan baja como se desee a una velocidad binaria arbitraria siempre que esa velocidad esté por debajo de la denominada *capacidad del canal*. Esta demostración se considera el principio de la llamada *teoría de la información*, y establece un límite a los sistemas de comunicaciones en cuanto a la máxima cantidad de información que se puede transmitir de forma fiable. Es preciso hacer notar que esta demostración establece el límite, pero no especifica cómo se puede alcanzar dicho límite. De momento, la teoría de la información no tiene respuesta a esta pregunta.

En este capítulo se pretende analizar cuál es la máxima cantidad de información que se puede transmitir de forma fiable utilizando un sistema digital de comunicaciones, y de qué factores depende el valor de ese límite. Aunque casi todo el mundo tiene una noción intuitiva de lo que es información, para realizar el análisis de un sistema de comunicaciones no basta con esa noción intuitiva, sino que es necesario disponer de medidas cuantitativas de información. Hartley, Nyquist o Shannon fueron pioneros en el desarrollo de definiciones de medidas para la información que resultan de utilidad para el propósito de este capítulo. Estas medidas de información están relacionadas con las distribuciones de probabilidad de los elementos cuya información se trata de cuantificar. Para poder hacer un uso de ellas en el análisis de un sistema digital de comunicaciones es preciso definir en primer lugar modelos probabilísticos que puedan utilizarse para representar el comportamiento de fuentes de información digitales. Dado que la información se transmitirá a través de un canal haciendo uso del sistema de comunicaciones, será preciso también disponer de modelos probabilísticos que permitan representar el comportamiento del canal y el comportamiento del sistema de comunicaciones a distintos niveles. A partir de esos modelos probabilísticos, y utilizando medidas cuantitativas de información, se podrán finalmente obtener los límites que puede alcanzar un cierto sistema de comunicaciones.

Este capítulo se dividirá en cuatro partes. En primer lugar se presentarán modelos probabilísticos adecuados para representar el comportamiento de fuentes de información. A continuación se definirán modelos probabilísticos que permitan representar el comportamiento del canal de comunicaciones y de los elementos del sistema de comunicaciones a distintos niveles. Genéricamente se denominarán todos estos modelos como modelos de canal, ya que los distintos elementos del sistema de comunicaciones pueden en cierto modo considerarse como parte del canal a través del que se transmite la información. La tercera parte presentará distintas medidas cuantitativas de información. Finalmente, en la última parte, haciendo uso de los modelos probabilísticos de fuente y de canal, y de las medidas cuantitativas de información presentadas en la tercera parte, se obtendrán los límites fundamentales alcanzables por un sistema digital de comunicaciones.

5.1. Modelado de las fuentes de información

Una fuente de información produce como salida algo, que se denominará genéricamente información, que es de interés para el receptor de la misma, que no la conoce de antemano. La misión del sistema de comunicaciones es asegurar que la información se transmite correctamente.

Como la salida de la fuente de información es una función variante en el tiempo e impredecible (si es predecible no tiene demasiado interés transmitirla), esta salida puede ser modelada mediante un proceso aleatorio. Las características de ese proceso aleatorio dependerán de las características y naturaleza de la fuente en cuestión. Por ejemplo, podrá ser un proceso aleatorio en tiempo continuo o en tiempo discreto dependiendo de si la fuente genera una información de naturaleza analógica o digital. Aunque este capítulo se centrará en el estudio de sistemas digitales, y los modelos más relevantes serán por tanto los modelos para fuentes digitales (las fuentes analógicas serán convertidas a formato digital antes de su transmisión), por completitud se estudiarán también de forma breve los modelos utilizados para fuentes analógicas.

5.1.1. Fuentes analógicas

El modelo utilizado para caracterizar una fuente analógica será habitualmente un proceso aleatorio en tiempo continuo, $X(t)$, cuyas propiedades estadísticas dependerán de la naturaleza de la fuente. Se puede tomar como ejemplo el modelado de una fuente de voz. Es conocido que la señal de voz tiene la mayor parte de su potencia distribuida esencialmente en la banda de frecuencias entre 300 y 4000 Hz. Por tanto, esta fuente se puede modelar mediante un proceso aleatorio cuya densidad espectral de potencia se ajuste a estas características, tal y como se ilustra en la Figura 5.2.

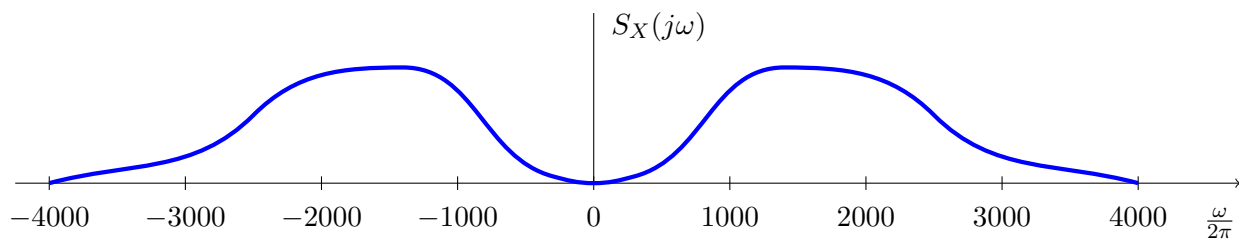


Figura 5.2: Ejemplo de una densidad espectral de potencia que representa un comportamiento típico de una señal de voz.

Habitualmente se considera que el comportamiento medio no cambia con el tiempo, lo que permite asumir que el proceso estacionario. Como además la media de la señal es nula, el proceso aleatorio utilizado para modelar la señal de voz podría ser un proceso aleatorio estacionario, de media nula, y función densidad espectral de potencia como la de la figura. Al ser un proceso estacionario, su función de autocorrelación vendrá dada por la transformada de Fourier inversa de esta densidad espectral de potencia.

El mismo procedimiento es aplicable a distintas fuentes analógicas. Por ejemplo para señales de TV, dependiendo del sistema (PAL, SECAM o NTSC), la banda se encuentra entre 0-6.5 MHz o entre 0-4.5 MHz. Se utilizaría para modelarlas un proceso aleatorio estacionario cuya densidad espectral de potencia representa el comportamiento medio de la respuesta en frecuencia de la señal al cuadrado.

Aunque cada señal analógica tendrá unas características espectrales diferentes, hay algunos aspectos comunes en la mayoría de los modelos utilizados:

1. Se consideran procesos limitados en banda.
2. Esto permite que se puedan muestrear siguiendo el criterio de Nyquist y posteriormente pueden volver a reconstruirse.

5.1.2. Fuentes digitales

En el caso de una fuente digital, su salida se puede modelar mediante un proceso aleatorio discreto en el tiempo. En ese caso, el modelo matemático para una fuente de información es el que muestra la Figura 5.3.



Figura 5.3: Modelo matemático de una fuente de información discreta en el tiempo

La fuente se modela como un proceso aleatorio discreto en el tiempo, $X[n]$. El alfabeto de la fuente puede ser:

- Discreto. Por ejemplo para modelar una fuente de datos digitales o señales analógicas muestreadas y cuantificadas.
- Continuo. Por ejemplo para representar fuentes analógicas muestreadas (como una señal de voz) antes de la cuantificación.

En función del tipo de fuente los parámetros estadísticos del proceso aleatorio serán diferentes. En esta sección se va a estudiar el modelo más sencillo de fuente, que permite realizar todo el desarrollo posterior del capítulo. Este modelo es el *modelo discreto sin memoria*.

Modelo discreto sin memoria

El modelo discreto sin memoria, o DMS (del inglés *discrete memoryless source*), de una fuente de información es un proceso aleatorio discreto en el tiempo y en amplitud, en el que todas las variables aleatorias que forman el proceso aleatorio $X[n]$ son independientes y tienen la misma distribución. Por tanto, una fuente DMS genera una secuencia de variables aleatorias i.i.d. que toman valores de un alfabeto discreto. Para describir de forma completa este tipo de fuente basta con conocer su alfabeto y su distribución, es decir

1. $\mathcal{A}_X = \{x_0, x_1, \dots, x_{M-1}\}$.
2. $p_X(x_i) = P(X = x_i)$ para $i = 0, 1, \dots, M - 1$.

Ejemplo

Una fuente de información binaria utiliza como modelo un modelo discreto sin memoria que viene descrito por el alfabeto $\mathcal{A}_X = \{0, 1\}$ y las probabilidades de cada símbolo binario, $P(X = 1) = p$, y $P(X = 0) = 1 - p$. En el caso particular en que $p = 1/2$, este tipo de fuente se denomina *fente binaria simétrica* o BSS (del inglés *binary symmetric source*).

5.2. Modelos probabilísticos de canal

En el Capítulo 4 se ha presentado el modelo general de un sistema de comunicaciones digitales, cuyos elementos funcionales se muestran en la Figura 5.4.

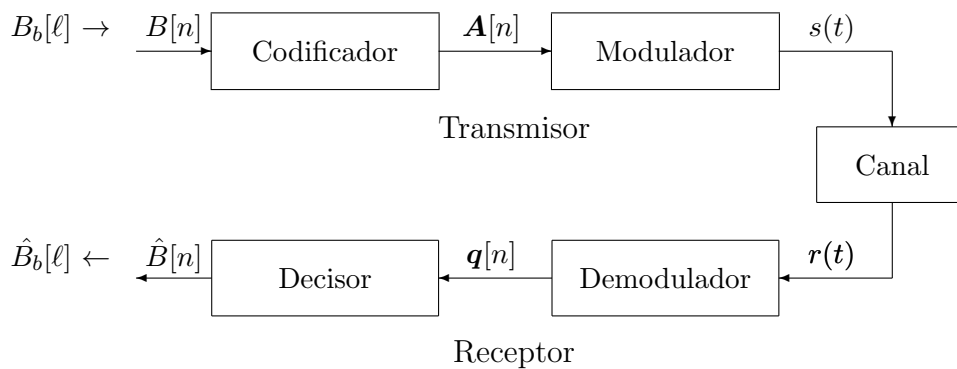


Figura 5.4: Modelo general de un sistema digital de comunicaciones.

En este esquema se denominaba canal a la abstracción del medio físico de transmisión, y se utilizaba como modelo simplificado para representarlo un modelo de canal aditivo gaussiano. En esta sección el término canal tendrá un significado más amplio que el considerado en el esquema de la Figura 5.4. Se van a definir varios modelos probabilísticos, que se denominarán genéricamente modelos de canal; estos modelos establecerán la relación probabilística entre la información recibida y la información transmitida entre distintos puntos del sistema de comunicaciones, lo que puede interpretarse como la definición de varios canales sobre el sistema, cada uno de ellos en un diferente nivel de abstracción. La caracterización de estos modelos vendrá dada por la distribución condicional de la salida dada la entrada. En todos los casos se considerará independencia condicional entre entradas y salidas para instantes de tiempo diferentes, por lo que se podrá eliminar la dependencia temporal. Esto significa que el valor de entrada en un cierto instante se modelará mediante una variable aleatoria X , y el valor de la correspondiente salida en el mismo instante se modelará con otra variable aleatoria, Y . En este caso el modelo probabilístico estará caracterizado por la distribución condicional

$$f_{Y|X}(y|x).$$

La diferencia entre los distintos modelos que se van a presentar está en la definición de lo que en cada caso se considera entrada y salida para cada *canal*. Se van a presentar cuatro modelos probabilísticos (o *canales*) diferentes, cada uno considerando lo que es el canal con un distinto nivel de abstracción sobre el modelo general del sistema de comunicaciones. En concreto, los cuatro modelos son los que se muestran en la Figura 5.5, y que se enumeran a continuación:

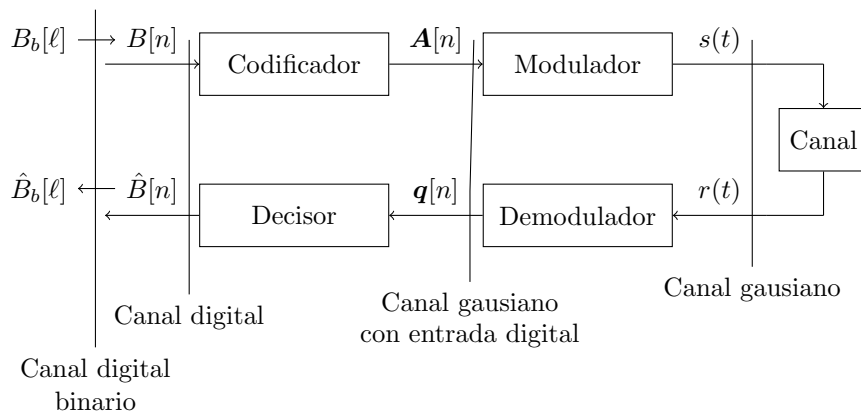


Figura 5.5: Definición de los distintos modelos de canal sobre un sistema de comunicaciones digital.

1. **Canal gaussiano.** Es el modelo para el representar el canal físico propiamente dicho, que transforma la señal transmitida mediante la adición de ruido aditivo, modelado mediante un proceso aleatorio estacionario, ergódico, blanco, gaussiano, de media nula y densidad espectral de potencia $N_0/2$, de forma que

$$r(t) = s(t) + n(t).$$

En este modelo, como entrada y salida se consideran los valores de la señal transmitida y la señal recibida, respectivamente, en el mismo instante temporal, es decir

$$X \equiv s(t), Y \equiv r(t).$$

En ambos casos, las variables aleatorias son variables aleatorias continuas, ya que las señales transmitida y recibida pueden tomar valores en un rango continuo de amplitudes.

2. **Canal gaussiano con entrada digital.** Este modelo coincide con lo que en el capítulo anterior se denominó canal discreto equivalente, que presenta como entrada una secuencia de símbolos de un alfabeto discreto (constelación) de M símbolos y como salida tiene una observación con un dominio continuo, la salida del demodulador. En el capítulo anterior se vió que la relación entre ambas era una relación aditiva con el término de ruido a la salida del demodulador

$$\mathbf{q}[n] = \mathbf{A}[n] + \mathbf{z}[n].$$

Las características estadísticas del vector de ruido $\mathbf{z}[n]$ se estudiaron en el capítulo anterior. Por tanto, en este modelo la entrada es el símbolo transmitido en un instante discreto n y la salida el valor de observación a la salida del demodulador en ese mismo instante. Se trata por tanto de símbolos y observaciones vectoriales de dimensión N , por lo que para representarlas se utilizarán variables aleatorias multidimensionales (vectoriales) de la misma dimensión

$$\mathbf{X} \equiv \mathbf{A}[n], \mathbf{Y} \equiv \mathbf{q}[n].$$

3. **Canal digital.** Este modelo considera el conjunto formado por el codificador, modulador, canal, demodulador y el decisor de símbolos. Como entrada tiene un conjunto de símbolos de un alfabeto de M valores y como salida un alfabeto de los mismo M símbolos: $B[n]$ y $\hat{B}[n]$ (o lo que es lo mismo $\mathbf{A}[n]$ y $\hat{\mathbf{A}}[n]$, ya que hay una equivalencia unívoca entre símbolos como bloques de M bits y vectores N -dimensionales de la constelación que los transportan). Por tanto, en este modelo la entrada y la salida vendrán definidos como

$$X \equiv B[n], Y \equiv \hat{B}[n] \text{ ó equivalentemente } X \equiv \mathbf{A}[n], Y \equiv \hat{\mathbf{A}}[n].$$

4. **Canal digital binario.** Este es el último nivel de abstracción, en el que se considera como canal todo el sistema de comunicaciones, cuyas entradas y salidas son los símbolos binarios. Representa el mayor nivel de abstracción que se puede realizar sobre un sistema de comunicaciones: el sistema completo visto como vehículo o canal para la transmisión de bits. En este caso por tanto, la entrada y salida se definirán como los bits transmitido y recibido en el mismo instante de tiempo

$$X \equiv B_b[\ell], Y \equiv \hat{B}_b[\ell].$$

Una vez que se han presentado estos cuatro canales, a continuación se obtendrá el modelo probabilístico (distribución condicional de la salida dada la entrada) que se utilizará para representar cada uno de ellos.

5.2.1. Canal gaussiano

La relación entre la entrada y la salida del canal gaussiano es

$$r(t) = s(t) + n(t),$$

donde el término de ruido $n(t)$ se modela como un proceso aleatorio estacionario, ergódico, blanco, gaussiano, de media nula y densidad espectral de potencia $S_n(j\omega) = N_0/2$. La función de autocorrelación del ruido es por tanto

$$R_n(\tau) = \frac{N_0}{2} \delta(\tau).$$

Para obtener el modelo probabilístico dado por la probabilidad condicional de la salida dada la entrada, lo primero que hay que tener en cuenta es que estrictamente la potencia del proceso de ruido, al ser blanco, es infinita. Esto implica que en la práctica para minimizar el efecto del ruido en el receptor se utilizará un filtro selectivo en frecuencia. Idealmente, este filtro no introducirá ninguna distorsión en la señal $s(t)$, que se considerará limitada en banda con un ancho de banda B Hz, y al mismo tiempo debe minimizar la potencia del ruido a su salida; el filtro que cumple estas condiciones es un filtro paso bajo ideal, con un ancho de banda igual al de la señal recibida. De este modo el canal gaussiano en realidad modelará la relación entre la señal transmitida y la señal recibida después de este filtrado, tal y como se ilustra en la Figura 5.6, donde la respuesta $h_n(t)$ (o su equivalente en frecuencia $H_n(j\omega)$) es la del filtro ideal utilizado para limitar el ruido.

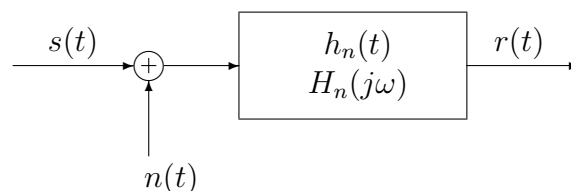


Figura 5.6: Canal gaussiano con filtrado en el receptor para la limitación de la potencia de ruido.

Si la señal tiene un ancho de banda B Hz, el ancho de banda del filtro será B Hz, con lo que su respuesta en frecuencia es

$$H(j\omega) = \Pi\left(\frac{\omega}{2W}\right),$$

donde $W = 2\pi B$ denota el ancho de banda en radianes por segundo, y su respuesta al impulso

$$h(t) = 2B \text{sinc}(2Bt).$$

La potencia del término de ruido a la salida de este filtro, como se vió en el Capítulo 2, se obtiene integrando la densidad espectral de potencia del proceso de ruido filtrado, que para filtros ideales supone multiplicar N_0 por el ancho de banda en Hz del filtro

$$\sigma^2 = \int_{-\infty}^{\infty} S_n(j\omega) |H_n(j\omega)|^2 d\omega = \frac{N_0}{2} \times 2B = N_0B.$$

Teniendo esto en cuenta, el *canal probabilístico gaussiano* se define como aquel que relaciona dos variables aleatorias X e Y con funciones de densidad de probabilidad continuas sobre \mathbb{R} que representan el valor de $s(t)$ y $r(t)$ en un cierto instante de tiempo. Dado que el valor de la salida en un instante será el de la entrada más el valor del ruido filtrado en ese instante, y que ese ruido tiene una distribución gaussiana con potencia N_0B , este modelo probabilístico está caracterizado por la función de densidad de probabilidad condicional

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-x)^2}{2\sigma^2}},$$

donde $\sigma^2 = N_0B$.

5.2.2. Canal gaussiano con entrada digital

La Figura 5.7 ilustra conceptualmente el modelo de canal gaussiano con entrada digital, que modela la relación entre los símbolos de la constelación que se transmiten y la observación a la salida del demodulador en un momento dado.

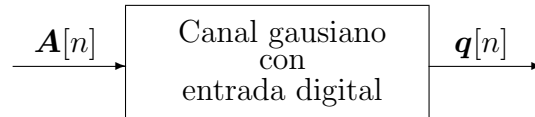


Figura 5.7: Canal gaussiano con entrada digital.

Como se vió en el Capítulo 4, la salida $\mathbf{q}[n]$ toma la forma

$$\mathbf{q}[n] = \mathbf{A}[n] + \mathbf{z}[n],$$

donde $\mathbf{z}[n]$ es la componente de ruido de ruido de la observación, que es un proceso estocástico discreto en el tiempo, multidimensional (de la dimensión del espacio de señales del sistema, N), formado por N variables aleatorias conjuntamente gaussianas, e independiente de $\mathbf{A}[n]$. Además, sus componentes son a su vez estadísticamente independientes las unas de las otras. Esto hace que se pueda eliminar la dependencia temporal y utilizar la representación

$$\mathbf{q} = \mathbf{A} + \mathbf{z},$$

donde \mathbf{z} tiene una función densidad de probabilidad gaussiana N -dimensional, de media nula y varianza $N_0/2$ en todas las direcciones del espacio

$$f_{\mathbf{z}}(\mathbf{z}) = \mathcal{N}^N \left(\mathbf{0}, \frac{N_0}{2} \right) = \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{z}\|^2}{N_0}}.$$

A partir de aquí, la obtención del modelo probabilístico es simple. Formalmente se define como aquel que relaciona las variables aleatorias N -dimensionales \mathbf{X} e \mathbf{Y} , la primera con alfabeto

discreto $\{\mathbf{x}_i\}$, con $i = 0, \dots, M - 1$, donde cada valor del alfabeto estará identificado con uno de los vectores N -dimensionales que forman la constelación del sistema, y la segunda con una función densidad de probabilidad N -dimensional continua sobre \mathbb{R} . Bajo esa premisa, el modelo probabilístico viene dado por la función densidad de probabilidad condicional

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}_i) = \frac{1}{(\pi N_0)^{N/2}} e^{-\frac{\|\mathbf{y}-\mathbf{x}_i\|^2}{N_0}},$$

es decir, una distribución gaussiana N dimensional con media el símbolo transmitido y varianza $\sigma^2 = N_0/2$ en cada dirección del espacio. Como se ve, coincide con el modelo probabilístico que define el canal discreto equivalente, utilizando ahora una notación ligeramente diferente en términos de las variables aleatorias \mathbf{X} e \mathbf{Y} para denotar, respectivamente, \mathbf{A} y \mathbf{q} .

5.2.3. Canal digital

La Figura 5.8 representa el canal digital. Habitualmente se trabaja con los símbolos de la constelación $\mathbf{A}[n]$ en lugar de con los símbolos $B[n]$, aunque dada la relación unívoca entre ambos, se puede aplicar sobre los últimos de igual modo, incluyendo del codificador y decodificador dentro de este modelo.

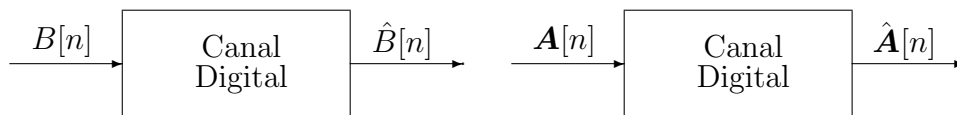


Figura 5.8: Representación conceptual del canal digital.

En este modelo se considera que cada símbolo de la secuencia $B[n]$ es estadísticamente independiente del resto de símbolos de la secuencia. Por otra parte, ya se ha visto en el apartado dedicado a la transmisión indefinida de símbolos, que dado que las señales sólo ocupan el intervalo de duración T dedicado al símbolo y que el ruido es independiente en cada momento, la recepción de cada símbolo es independiente del resto de símbolos. Bajo estas condiciones, la probabilidad de tener un símbolo determinado de la secuencia de salida del canal digital, $\hat{B}[n]$, depende únicamente del símbolo emitido en ese mismo instante $B[n]$. Por tanto, es posible eliminar la dependencia temporal y analizar únicamente el caso de la transmisión de un símbolo aislado, entendiéndose que cada vez que se use el canal para transmitir un símbolo, el canal no va a modificar su comportamiento.

Cuando se transmite el símbolo b_i a la entrada del canal, a la salida del mismo vamos a tener un símbolo, b_j , con una probabilidad determinada que es la probabilidad condicional

$$p_{\hat{B}|B}(b_j|b_i).$$

Debido a la asignación unívoca entre un símbolo y la representación vectorial de la señal que lo transmite, esta probabilidad cumple que

$$p_{\hat{B}|B}(b_j|b_i) = p_{\hat{\mathbf{A}}|\mathbf{A}}(\mathbf{a}_j|\mathbf{a}_i).$$

En el capítulo anterior se estudio cómo se calculan estas probabilidades integrando la distribución condicional de la observación para el símbolo transmitido en la región de decisión de cada uno de los símbolos de la constelación. Como en este modelo la entrada y la salida eran precisamente $X \equiv B$ e $Y \equiv \hat{B}$, si se conocen estas probabilidades para todas las combinaciones posibles de símbolos transmitido y recibido, el canal estará completamente caracterizado.

Existe un modelo probabilístico ampliamente utilizado que contempla como caso particular el canal digital y se denomina *canal discreto sin memoria* o DMC (del inglés *Discrete Memoryless Channel*). El DMC es un modelo estadístico que relaciona una variable aleatoria X con función densidad de probabilidad discreta que denominamos entrada y otra variable aleatoria Y con función de probabilidad discreta que denominamos salida. En el caso particular del DMC, los alfabetos de entrada y salida pueden ser distintos. En nuestra aplicación comparten el mismo alfabeto, por lo que este es un caso particular del DMC.

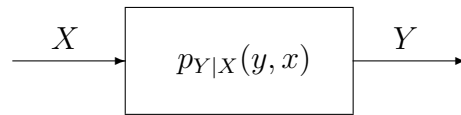


Figura 5.9: Modelo probabilístico del canal discreto sin memoria (DMC).

El canal discreto sin memoria se puede representar mediante el diagrama de la Figura 5.9. Se observa que dentro del bloque que representa el canal no aparecen respuestas al impulso ni respuestas en frecuencia, sino las probabilidades condicionales de la salida respecto a la entrada. El calificativo de *discreto* proviene de la naturaleza de las probabilidades de X e Y , que tienen un alfabeto discreto, y no de *tiempo discreto*, por lo que no hay que confundirlo con el canal discreto equivalente. El calificativo sin memoria proviene del modelo probabilístico de la entrada y la salida, variables aleatorias y no procesos aleatorios (no hay dependencia temporal de los estadísticos). Formalmente, un canal discreto sin memoria se define a través de los siguientes elementos:

1. El *alfabeto de entrada* (de M_X posibles valores)

$$\{x_i\}, i = 0, \dots, M_X - 1.$$

2. El *alfabeto de salida* (de M_Y posibles valores)

$$\{y_i\}, i = 0, \dots, M_Y - 1.$$

3. El conjunto de probabilidades condicionales

$$p_{Y|X}(y_j|x_i).$$

Estas probabilidades se denominan *probabilidades de transición*, y se suelen agrupar en la denominada *matriz de canal*, que es una matriz de M_X filas y M_Y columnas que ordena las probabilidades de transición de la forma siguiente

$$P = \begin{bmatrix} p_{Y|X}(y_0|x_0) & p_{Y|X}(y_1|x_0) & \cdots & p_{Y|X}(y_{M_Y-1}|x_0) \\ p_{Y|X}(y_0|x_1) & p_{Y|X}(y_1|x_1) & \cdots & p_{Y|X}(y_{M_Y-1}|x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p_{Y|X}(y_0|x_{M_X-1}) & p_{Y|X}(y_1|x_{M_X-1}) & \cdots & p_{Y|X}(y_{M_Y-1}|x_{M_X-1}) \end{bmatrix}$$

Tal y como está definida la matriz, una fila está asociada a un cierto valor del alfabeto de entrada, mientras que una columna está asociada a un cierto valor del alfabeto de salida. Por tanto, la suma de los elementos de una fila da como resultado 1 (se tiene la suma de la distribución condicional de probabilidad sobre todo el espacio de observación). Además, si se concatenan dos canales discretos sin memoria, la matriz de canal de la concatenación se obtendrá mediante el producto de las matrices de canal de cada uno de los canales.

En ocasiones, en lugar de la matriz de canal se utiliza una representación gráfica para especificar las probabilidades de transición, utilizando un diagrama de flechas o diagrama de rejilla, como el que muestra la Figura 5.10. En este caso, las probabilidades de transición se incluyen en los pesos asociados a las distintas flechas que forman el diagrama uniendo elementos de la entrada con la salida. Dada la definición de estas probabilidades, las probabilidades asociadas a las flechas que salen de un mismo nodo suman la unidad.

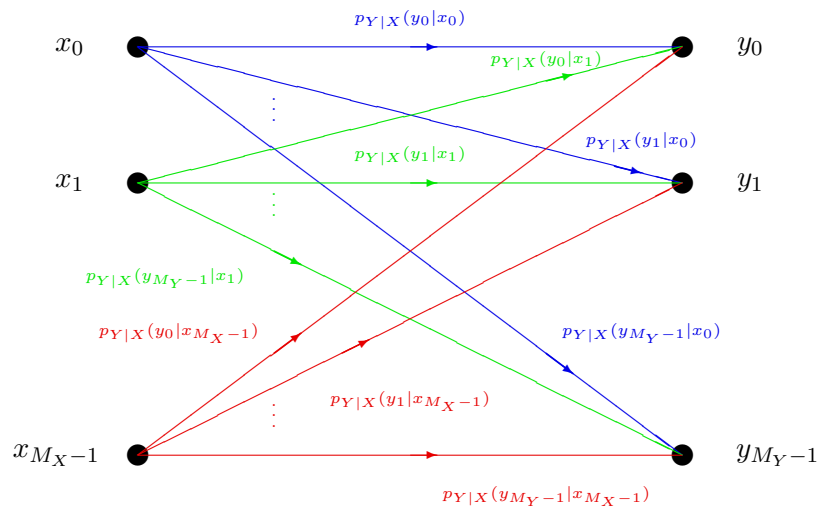


Figura 5.10: Representación de un DMC mediante un diagrama de rejilla.

Hay que resaltar que en la definición de un canal discreto sin memoria aparecen el alfabeto de entrada y el de salida, pero no las distribuciones de probabilidades de la entrada, $p_X(x_i)$ ni de la salida, $p_Y(y_j)$, ya que estas probabilidades no forman parte de la naturaleza del canal.

A continuación se muestra cómo obtener un DMC que represente al canal digital para un cierto sistema de comunicaciones. El modelo se obtiene a partir de las probabilidades de error de símbolo definidas en el capítulo anterior. En primer lugar, para un DMC que modela un sistema de comunicaciones tanto el alfabeto de entrada como el de salida se corresponden con el alfabeto de símbolos del sistema, bien de $\mathbf{A}[n]$ o de $\mathbf{B}[n]$, dada su equivalencia. Por conveniencia utilizaremos la representación vectorial de los símbolos, es decir

$$x_i \equiv \mathbf{a}_i,$$

$$y_j \equiv \mathbf{a}_j,$$

donde ahora $M_X = M_Y = M$. De esta forma, la asociación del alfabeto con los símbolos de la constelación está implícita en los subíndices. En cuanto a las probabilidades de transición, es evidente que

$$p_{Y|X}(y_j|x_i) \equiv p_{\hat{\mathbf{A}}|\mathbf{A}}(\mathbf{a}_j|\mathbf{a}_i).$$

Es decir, las probabilidades de transición $p_{Y|X}(y_j|x_i)$ indica la probabilidad de recibir el símbolo \mathbf{a}_j cuando se ha transmitido el símbolo \mathbf{a}_i . En este caso, los elementos de la diagonal principal de la matriz de canal, para los que $j = i$, se corresponden con las probabilidades condicionales de acierto para cada símbolo

$$p_{Y|X}(y_i|x_i) = p_{\hat{\mathbf{A}}|\mathbf{A}}(\mathbf{a}_i|\mathbf{a}_i) = P_{a|\mathbf{a}_i} = 1 - P_{e|\mathbf{a}_i}.$$

Los elementos fuera de la diagonal en cambio se corresponden con las probabilidades de error entre distintos símbolos

$$p_{Y|X}(y_j|x_i) = p_{\hat{A}|A}(\mathbf{a}_j|\mathbf{a}_i) = P_{e|\mathbf{a}_i \rightarrow \mathbf{a}_j}$$

Consecuentemente, la suma de los elementos de cada fila fuera de la diagonal principal equivale a la probabilidad de error condicional para el símbolo asociado a dicha fila

$$\sum_{\substack{j=0 \\ j \neq i}}^{M-1} p_{Y|X}(y_j|x_i) = \sum_{\substack{j=0 \\ j \neq i}}^{M-1} P_{e|\mathbf{a}_i \rightarrow \mathbf{a}_j} = P_{e|\mathbf{a}_i}$$

Esto significa que en un sistema ideal, la matriz de canal o el diagrama de rejilla deberían ser como se muestra en la Figura 5.11: una matriz diagonal, o un diagrama de rejilla con una única flecha partiendo de cada símbolo de entrada al mismo símbolo de salida.

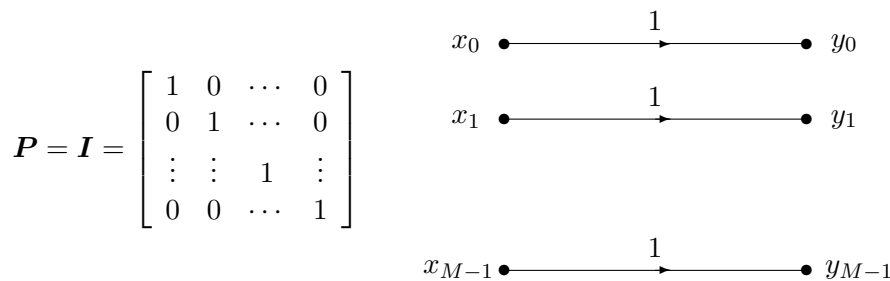


Figura 5.11: Valores ideales de un DMC que modela un canal gaussiano.

Para ilustrar el procedimiento de obtención de la matriz de canal para un sistema de comunicaciones se utilizará como ejemplo un sistema con una constelación de cuatro símbolos, $M = 4$, en un espacio unidimensional, $N = 1$, con coordenadas $\mathbf{a}_0 = -3$, $\mathbf{a}_1 = -1$, $\mathbf{a}_2 = +1$, $\mathbf{a}_3 = +3$ y equiprobables, con lo que las regiones de decisión vienen dadas por los umbrales $q_{u1} = -2$, $q_{u2} = 0$, $q_{u3} = +2$

$$I_0 = (-\infty, -2], I_1 = (-2, 0], I_2 = (0, +2], I_3 = (+2, +\infty)$$

tal y como muestra la Figura 5.12.

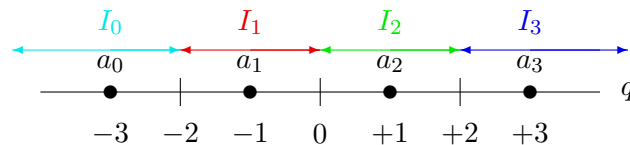


Figura 5.12: Constelación unidimensional de cuatro símbolos equiprobables y sus correspondientes regiones de decisión.

En este caso se tiene la siguiente asociación

$$X \equiv \mathbf{A}[n], Y \equiv \hat{\mathbf{A}}[n],$$

por lo que los alfabetos de entrada y salida coinciden

$$x_i \equiv \mathbf{a}_i, y_j \equiv \mathbf{a}_j \text{ para } i, j \in \{0, 1, \dots, M - 1\}.$$

Las probabilidades de transición $p_{Y|X}(y_j|x_i)$ que definen probabilísticamente el sistema se corresponden en este caso con la probabilidad de recibir el símbolo de índice j cuando se ha transmitido el símbolo de índice i . Estos valores definen la probabilidad de acertar en la transmisión de un símbolo, si $j = i$, o la probabilidad de error entre dos símbolos, si $j \neq i$. Utilizando la notación del capítulo anterior, se tiene

$$p_{Y|X}(y_i|x_i) = P_{a|a_i} = 1 - P_{e|a_i}$$

y

$$p_{Y|X}(y_j|x_i) = P_{e|a_i \rightarrow a_j} \text{ para } j \neq i.$$

En el capítulo anterior se explicó cómo se obtenían estos valores, que se volverán a obtener para esta constelación. En primer lugar se calcularán las probabilidades de transición que aparecen en la primera fila de la matriz de canal, $p_{Y|X}(y_j|x_0)$, $\forall j$; es decir, que en este caso se tienen las probabilidades de recibir cada uno de los 4 símbolos cuando se transmite \mathbf{a}_0 (el símbolo asociado a x_0). La distribución de la observación cuando se transmite el símbolo \mathbf{a}_0 es gaussiana, de media \mathbf{a}_0 y varianza $N_0/2$. Así que para obtener las probabilidades condicionales, sólo hay que integrar esa distribución gaussiana en cada una de las 4 regiones de decisión, tal y como se ilustra en la Figura 5.13.

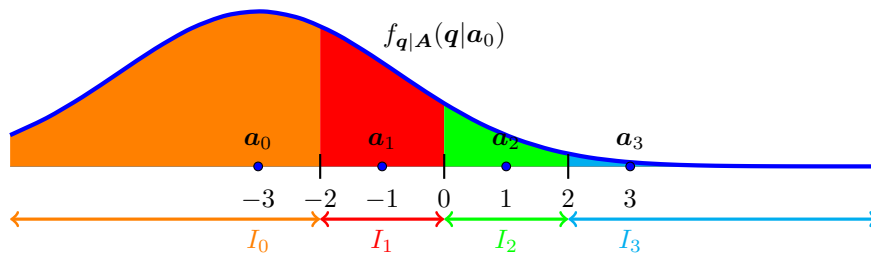


Figura 5.13: Cálculo de las probabilidades de transición asociadas a la primera fila de la matriz de canal.

A continuación se calculan estas probabilidades

- Distribución $f_{q|A}(\mathbf{q}|\mathbf{a}_0)$: gaussiana de media $\mathbf{a}_0 = -3$ y varianza $N_0/2$

$$p_{Y|X}(y_0|x_0) = 1 - P_{e|a_0} = 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_1|x_0) = P_{e|a_0 \rightarrow a_1} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_2|x_0) = P_{e|a_0 \rightarrow a_2} = Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_3|x_0) = P_{e|a_0 \rightarrow a_3} = Q\left(\frac{5}{\sqrt{N_0/2}}\right)$$

Se puede comprobar que estas cuatro probabilidades suman la unidad, como era de esperar.

A continuación se calcularán las probabilidades de transición que aparecen en la segunda fila de la matriz de canal, $p_{Y|X}(y_j|x_1)$, $\forall j$; es decir, que en este caso se tienen las probabilidades de recibir

cada uno de los 4 símbolos cuando se transmite \mathbf{a}_1 (el símbolo asociado a x_1). La distribución de la observación cuando se transmite el símbolo \mathbf{a}_1 es gaussiana, de media \mathbf{a}_1 y varianza $N_0/2$. Así que para obtener las probabilidades condicionales, sólo hay que integrar esa distribución gaussiana en cada una de las 4 regiones de decisión, tal y como se muestra en la Figura 5.14.

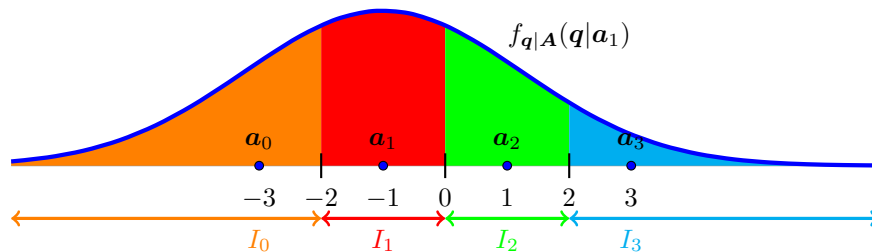


Figura 5.14: Cálculo de las probabilidades de transición asociadas a la primera segunda de la matriz de canal.

A continuación se calculan estas probabilidades

- Distribución $f_{q|A}(\mathbf{q}|\mathbf{a}_1)$: gaussiana de media \mathbf{a}_1 y varianza $N_0/2$

$$p_{Y|X}(y_0|x_1) = P_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_0} = Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_1|x_1) = 1 - P_{e|\mathbf{a}_1} = 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_2|x_1) = P_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_2} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_3|x_1) = P_{e|\mathbf{a}_1 \rightarrow \mathbf{a}_3} = Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

En la tercera fila de la matriz de canal se tienen las probabilidades de transición $p_{Y|X}(y_j|x_2), \forall j$; es decir, que en este caso se tienen las probabilidades de recibir cada uno de los 4 símbolos cuando se transmite \mathbf{a}_2 (el símbolo asociado a x_2). La distribución de la observación cuando se transmite el símbolo \mathbf{a}_2 es en este caso gaussiana de media $\mathbf{a}_2 = +1$ y varianza $N_0/2$. Al igual que en los casos anteriores, para obtener las probabilidades condicionales hay que integrar esa distribución gaussiana en cada una de las 4 regiones de decisión, tal y como se ilustra en la Figura 5.15.

A continuación se calculan estas probabilidades

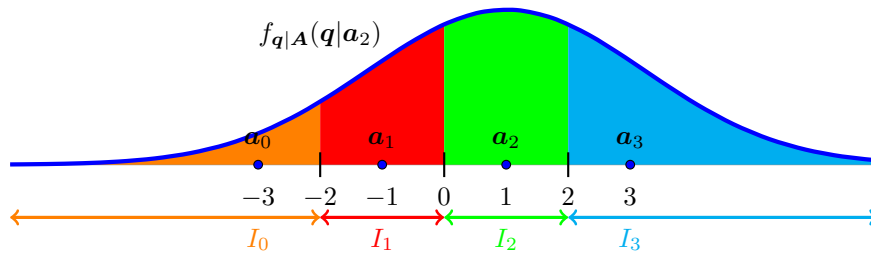


Figura 5.15: Cálculo de las probabilidades de transición asociadas a la tercera fila de la matriz de canal.

- Distribución $f_{q|A}(q|a_2)$: gaussiana de media a_2 y varianza $N_0/2$

$$p_{Y|X}(y_0|x_2) = P_{e|a_2 \rightarrow a_0} = Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_1|x_2) = P_{e|a_2 \rightarrow a_1} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_2|x_2) = 1 - P_{e|a_2} = 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

$$p_{Y|X}(y_3|x_2) = P_{e|a_2 \rightarrow a_3} = Q\left(\frac{1}{\sqrt{N_0/2}}\right)$$

Finalmente, en la cuarta y última fila de la matriz de canal se tienen las probabilidades de transición $p_{Y|X}(y_j|x_2), \forall j$; es decir, que en este caso se tienen las probabilidades de recibir cada uno de los 4 símbolos cuando se transmite a_3 (el símbolo asociado a x_3). La distribución de la observación cuando se transmite el símbolo a_3 es en este caso gaussiana de media $a_3 = +3$ y varianza $N_0/2$. Las probabilidades condicionales se obtienen integrando esta distribución condicional en cada una de las 4 regiones de decisión, tal y como se muestra en la Figura 5.16.

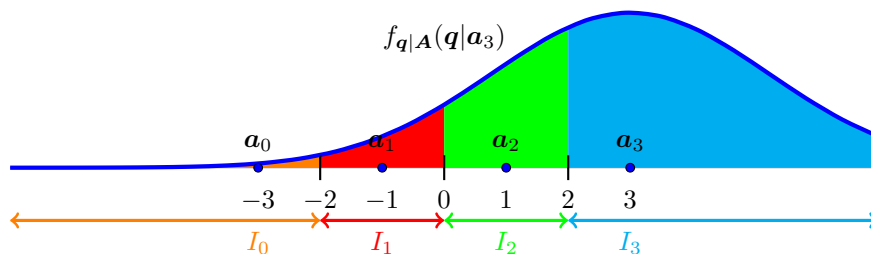


Figura 5.16: Cálculo de las probabilidades de transición asociadas a la cuarta fila de la matriz de canal.

A continuación se calculan estas probabilidades

- Distribución $f_{q|A}(\mathbf{q}|\mathbf{a}_3)$: gaussiana de media \mathbf{a}_3 y varianza $N_0/2$

$$\begin{aligned}
 p_{Y|X}(y_0|x_3) &= P_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_0} = Q\left(\frac{5}{\sqrt{N_0/2}}\right) \\
 p_{Y|X}(y_1|x_3) &= P_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_1} = Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) \\
 p_{Y|X}(y_2|x_3) &= P_{e|\mathbf{a}_3 \rightarrow \mathbf{a}_2} = Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) \\
 p_{Y|X}(y_3|x_3) &= 1 - P_{e|\mathbf{a}_3} = 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)
 \end{aligned}$$

Para este ejemplo, dada la simetría de la constelación, una vez obtenidos los valores de las probabilidades de transición para las dos primeras filas, las de las dos siguientes se pueden obtener de forma inmediata. En cualquier caso, agrupando todas las probabilidades de transición, el DMC que representa al sistema de comunicaciones que utiliza la constelación de 4 símbolos del ejemplo tiene la siguiente matriz de canal

$$\mathbf{P} = \begin{bmatrix}
 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right) \\
 Q\left(\frac{1}{\sqrt{N_0/2}}\right) & 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) \\
 Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) \\
 Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right)
 \end{bmatrix}$$

Para distintos valores de N_0 , la matriz tendrá distintos valores. A medida que N_0 disminuye, la matriz tiende a la matriz identidad, que es la matriz de canal ideal.

5.2.4. Canal digital binario

El canal digital binario es el modelo probabilístico que supone una mayor abstracción al considerar todo el sistema de comunicaciones como un canal que transmite y recibe bits.



Figura 5.17: Representación conceptual del canal digital binario.

Se trata por tanto de un modelo en el que la descripción probabilística vendrá dada por la probabilidad de recibir cada uno de los posibles valores de la secuencia binaria $\hat{B}_b[\ell]$ dada la secuencia transmitida $B_b[\ell]$. En el caso en que se asume independencia temporal sobre la secuencia de bits transmitida y recibida, se puede obviar la dependencia temporal, y el modelo en este caso

está definido por cuatro probabilidades de transición que, si se utiliza la notación $x_0 \equiv 0$, $x_1 \equiv 1$, $y_0 \equiv 0$ e $y_1 \equiv 1$ y se tiene en cuenta que por definición

$$p_{Y|X}(y_0|x_i) = 1 - p_{Y|X}(y_1|x_i),$$

quedan reducidas únicamente a dos probabilidades de transición relevantes: la probabilidad de error (o de acierto) condicional para cada bit. En ese caso, puede utilizarse un DMC particularizado para el caso $M_X = M_Y = 2$ como modelo del canal digital binario, que tendrá la forma

$$\mathbf{P} = \begin{bmatrix} p_{Y|X}(y_0|x_0) & p_{Y|X}(y_1|x_0) \\ p_{Y|X}(y_0|x_1) & p_{Y|X}(y_1|x_1) \end{bmatrix} = \begin{bmatrix} 1 - p_{e|0} & p_{e|0} \\ p_{e|1} & 1 - p_{e|1} \end{bmatrix}.$$

Las probabilidades $p_{e|0}$ y $p_{e|1}$ denotan la probabilidad de error de bit cuando se transmite un cero y la probabilidad de error de bit cuando de transmite un uno, respectivamente. En la mayor parte de los sistemas de comunicaciones estas dos probabilidades son iguales

$$p_{e|0} = p_{e|1} = \varepsilon,$$

en cuyo caso la matriz de canal es simétrica

$$\mathbf{P} = \begin{bmatrix} p_{Y|X}(y_0|x_0) & p_{Y|X}(y_1|x_0) \\ p_{Y|X}(y_0|x_1) & p_{Y|X}(y_1|x_1) \end{bmatrix} = \begin{bmatrix} 1 - \varepsilon & \varepsilon \\ \varepsilon & 1 - \varepsilon \end{bmatrix}.$$

Este caso se conoce como *canal binario simétrico* o BSC (del inglés “*Binary Symmetric Channel*”). La representación en diagrama de rejilla para este modelo se muestra en la Figura 5.18.

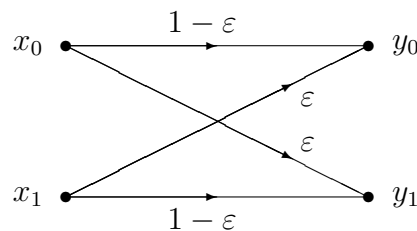


Figura 5.18: Representación en diagrama de rejilla de un canal binario simétrico o BSC.

El BSC es un modelo probabilístico que puede resultar apropiado para representar el canal digital binario. Sin embargo, antes de dar por buena la equivalencia entre canal digital binario y BSC, conviene realizar las siguientes puntualizaciones:

1. Para sistemas binarios, $M = 2$, un símbolo de la secuencia $B[n]$ transporta un único bit, y si los símbolos se transmiten de forma independiente en $B[n]$ eso implica una transmisión independiente a nivel de bit en la secuencia $B_b[\ell]$. En este caso, el modelo BSC representa de forma exacta el canal digital binario.
2. Para sistemas M -arios con $M > 2$, el BSC representa el comportamiento medio a lo largo del tiempo del canal digital binario, porque en el sistema real la transmisión se realiza por bloques de $m = \log_2 M$ bits, símbolo a símbolo (secuencia $B[n]$ y su representación vectorial $\mathbf{A}[n]$); se puede por tanto decir que mientras que el canal digital no posee memoria (porque la transmisión se realiza símbolo a símbolo de manera independiente) y encaja perfectamente en el modelo del DMC, el canal digital binario tiene la memoria introducida por el codificador que transforma la secuencia $B_b[\ell]$ en la secuencia $B[n]$. Eso significa que no puede considerarse de forma estricta un canal sin memoria. Desde este punto de vista, el BSC es una aproximación al canal digital binario que representa su comportamiento medio a lo largo del tiempo.

3. El valor de la probabilidad de error ε definida para el BSC es en ambos casos es la probabilidad de error de bit (BER) del sistema.

A pesar de esta discordancia entre las hipótesis del modelo BSC y la naturaleza del canal digital binario, se suele aceptar en la práctica la equivalencia de ambos, asignando a p el valor de la BER del sistema

$$\mathbf{P} = \begin{bmatrix} 1 - BER & BER \\ BER & 1 - BER \end{bmatrix}.$$

5.3. Medidas cuantitativas de información

Una vez establecidos los modelos probabilísticos para fuentes de información, el modelo discreto sin memoria o DMS, y para los canales digitales (canal digital y canal digital binario), en este caso el modelo sin memoria o DMC, en esta sección se analizan varios tipos de medidas cuantitativas de información. Por un lado medidas que se pueden aplicar sobre una variable aleatoria, que puede ser por ejemplo la que modela la entrada o salida de un canal, y otras que se aplican simultáneamente sobre dos variables aleatorias, que en nuestro problema tendrían en cuenta la relación entre la entrada y la salida de un canal. A través de estas medidas se podrá posteriormente calcular la máxima cantidad de información que se puede transmitir de forma fiable con un sistema de comunicaciones digital.

5.3.1. Información y entropía

En primer lugar se va a cuantificar la información que tiene una cierta variable aleatoria discreta. En la aplicación para el estudio de sistemas de comunicaciones, esta variable aleatoria puede representar tanto la salida de una fuente de información como la entrada o salida de un canal digital.

Para obtener una medida cuantitativa de la información que contiene una variable aleatoria, en primer lugar se buscará una medida para la información que contiene un suceso de esa variable aleatoria, es decir, el hecho de que la variable aleatoria tome un cierto valor dentro de su alfabeto, como $X = x_i$. Antes de establecer una medida cuantitativa de información para este caso, se van a presentar de forma intuitiva algunas de las propiedades básicas que esta medida debe cumplir, con el objetivo de encontrar luego alguna función que cumpla dichas propiedades y que pueda utilizarse como medida cuantitativa de la información que tiene un suceso de una variable aleatoria.

- Una noción intuitiva de información indica que la cantidad de información de un cierto suceso está relacionada con la probabilidad con la que este se produce. El conocimiento de que se ha producido un suceso poco probable aporta en general más información que el de un suceso más probable. Por poner un ejemplo sencillo, el número medio de días de lluvia al año en algunos puntos del desierto del Sahara es de menos de un día al año, mientras que en la ciudad de Santander es de 151. El hecho de conocer que hoy ha llovido en el desierto del Sahara es más informativo que el hecho de conocer que hoy ha llovido en Santander (el primer caso puede aparecer en algún informativo, mientras que es poco probable que el segundo se considere algo noticioso o informativo).
- Por otro lado, cambios pequeños en la probabilidad del suceso deberían dar lugar a pequeños cambios en la información del mismo. Por ejemplo, el suceso de que llueva en Santander

tendrá más o menos la misma cantidad de información que el suceso de que llueva en Gijón, mientras que el suceso de que llueva en el desierto del Sahara tendrá una información similar al suceso de que llueva en el desierto del Gobi o el de Atacama. Esta idea intuitiva se puede interpretar como que la dependencia de la información con respecto de la probabilidad del suceso debe ser continua.

- Finalmente, parece también intuitivo pensar que si se tienen varios sucesos independientes entre sí, la información que tienen en conjunto deberá ser la suma de la información de cada uno de ellos por separado.

Partiendo de estas nociones intuitivas sobre algunas características que debe tener una medida de información sobre un cierto suceso, se llega a la denominada *auto-información*, que es una medida cuantitativa de la información que contiene un suceso de una variable aleatoria.

Auto-información

La auto-información del suceso $X = x_i$ se denota como $I_X(x_i)$. Para obtener la expresión analítica de esta función se han traducido a notación matemática las nociones intuitivas sobre dicha medida que se han comentado anteriormente, lo que da lugar a las cuatro condiciones que ha de cumplir dicha función

1. La medida de información de un suceso deberá depender de la probabilidad de dicho suceso, y no del valor del propio suceso, es decir

$$I_X(x_i) = f(p_X(x_i)).$$

2. Debe además ser una función decreciente de la probabilidad del suceso

$$p_X(x_i) > p_X(x_j) \text{ implicará que } I_X(x_i) < I_X(x_j),$$

lo que significa que la función $f(\cdot)$ deberá ser una función decreciente

$$f(a) < f(b) \text{ para todo } a > b.$$

3. La función $f(\cdot)$ utilizada para la autoinformación deberá también ser una función continua de su argumento, de forma que la variación de información será continua sobre la probabilidad de los sucesos.
4. Finalmente, si dos variables aleatorias son independientes, y se define un suceso conjunto $X = x_i$ e $Y = y_j$, la información del suceso conjunto deberá ser la suma de la información de cada suceso

$$I_{X,Y}(x_i, y_j) = I_X(x_i) + I_Y(y_j).$$

Dado que para variables aleatorias independientes la probabilidad conjunta se puede escribir como el producto de las probabilidades marginales de cada variable

$$p_{X,Y}(x_i, y_j) = p_X(x_i) \times p_Y(y_j),$$

esto significa que la función $f(\cdot)$ elegida para la medida de información deberá ser aditiva sobre el producto de los argumentos, es decir

$$f(a \times b) = f(a) + f(b).$$

Se puede demostrar que la única función que cumple estas propiedades es la función logarítmica. Por tanto, la auto-información se define como

$$I_X(x_i) = -\log(p_X(x_i)).$$

Teniendo en cuenta las propiedades de la función logarítmica, la auto-información se puede escribir alternativamente como

$$I_X(x_i) = \log\left(\frac{1}{p_X(x_i)}\right).$$

La base del logaritmo no es determinante para las características de la medida de información. Lo único que implica son las unidades en que se expresa la información. Las bases utilizadas con mayor frecuencia son 2 y la base natural o número e (logaritmo neperiano). Si la base es 2, las unidades son *bits*, y si se usa el logaritmo natural o neperiano, las unidades son *nats*. En lo sucesivo, cuando no se especifique la base, se asumirán logaritmos de base 2 y por tanto bits como unidades de información. En cualquier caso, el cambio de la base, y por tanto de unidades, no supone más que un escalado, ya que en general los logaritmos en una cierta base se relacionan con los logaritmos en base natural a través de la siguiente relación

$$\log_b x = \frac{\log_e x}{\log_e b} = \frac{\ln x}{\ln b},$$

lo que directamente supone una relación lineal entre los logaritmos en dos bases distintas.

Entropía

La auto-información proporciona una medida cuantitativa de la información de un suceso aislado. Si se quiere cuantificar la información de una variable aleatoria (por ejemplo para modelar una fuente de información) han de tenerse en cuenta todos los posibles sucesos. Una opción razonable consiste en promediar la información de cada suceso considerando su probabilidad. El contenido de información de una variable aleatoria así calculado se denomina *entropía* y se denota mediante $H(X)$. Por tanto, la entropía de la variable aleatoria X se obtiene promediando la auto-información de cada uno de los sucesos que forman parte del alfabeto de la variable aleatoria

$$H(X) = -\sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log p_X(x_i) = \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log\left(\frac{1}{p_X(x_i)}\right).$$

Las unidades serán bits o nats por símbolo, dependiendo de la base empleada.

A efectos del cómputo hay que tener en cuenta que se considerará que $0 \log(0) = 0$. Como se ve, la entropía es una función de la probabilidad de cada suceso, o lo que es lo mismo, de la función densidad de probabilidad discreta, y proporciona un número que representa el contenido de información de esa fuente. No se debe confundir con una función de variable aleatoria que es otra variable aleatoria, tal y como la notación puede hacer parecer.

La entropía puede interpretarse como una cantidad que representa la incertidumbre sobre el valor concreto que toma una variable aleatoria X , que puede modelar, por ejemplo, la salida de una fuente de información. Si X toma siempre el mismo valor x_i , es decir, si $p_X(x_i) = 1$, no se tiene ninguna incertidumbre acerca del valor de la variable aleatoria y la entropía vale 0. Si X deja de tomar siempre el mismo valor, entonces la incertidumbre aumenta y con ella la entropía.

Dos propiedades importantes de la entropía de una variable aleatoria discreta son:

1. La entropía de una variable aleatoria discreta es una función no negativa, es decir

$$H(X) \geq 0.$$

Esto es evidente ya que el rango de valores de una probabilidad es $0 \leq p_X(x_i) \leq 1$ y $\log(x) \leq 0$ para $0 < x \leq 1$. El valor $H(X) = 0$ sólo se produce si uno de los elementos del alfabeto tiene probabilidad uno y por tanto el resto tiene probabilidad nula.

2. El valor máximo que puede tomar la entropía de una variable aleatoria discreta es el logaritmo del número de elementos de su alfabeto

$$H(X) \leq \log(M_X).$$

Ese valor máximo se produce únicamente si los símbolos son equiprobables, $p_X(x_i) = 1/M_X$.

Estas dos propiedades establecen los límites para los valores mínimo y máximo que puede tomar la entropía de una variable aleatoria X , e indican bajo que distribuciones se obtienen dichos valores mínimo y máximo, respectivamente. Para ilustrar estas propiedades, a continuación se calcula la entropía en un caso muy sencillo: una variable aleatoria binaria, $M_X = 2$, donde las probabilidades de cada símbolo se parametrizan con la probabilidad de uno de los dos elementos del alfabeto, $p_X(x_0) = p$, $p_X(x_1) = 1 - p$. En este caso la entropía es igual a

$$H(X) = -p \cdot \log(p) - (1 - p) \cdot \log(1 - p) = p \cdot \log \frac{1}{p} + (1 - p) \log \frac{1}{1 - p} \equiv H_b(p)$$

Esta función, denotada como $H_b(p)$, o en ocasiones como $\Omega(p)$, se denomina *función de entropía binaria*, y se representa en la Figura 5.19 en función de su argumento. Recuerde que dicho argumento representa la probabilidad de uno de los dos elementos del alfabeto de la variable aleatoria binaria.

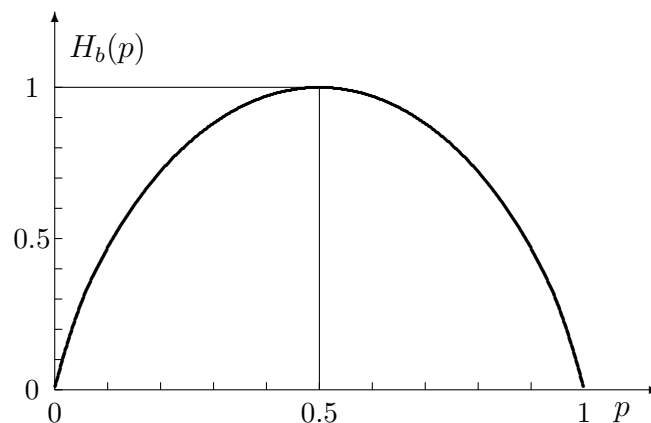


Figura 5.19: Función de entropía binaria, $H_b(p)$, expresada en bits por símbolo.

Se comprueba que en los casos en los que no existe incertidumbre, $p = 0$ o $p = 1$, la entropía es nula. Fuera de estos casos el valor es siempre mayor que cero, tomando el valor máximo, 1 bit por símbolo, cuando los símbolos son equiprobables, lo que representa la situación de máxima incertidumbre posible. Además, dado que la entropía depende únicamente de los valores de la distribución de probabilidad y no de a qué valores está asignado cada uno de esos valores, es una función simétrica respecto a $p = \frac{1}{2}$, lo que es evidente dado que por definición

$$H_b(p) = H_b(1 - p).$$

La función entropía binaria puede servir de referencia para definir un bit de información: un bit es la información que se tiene cuando se transmiten dos símbolos con igual probabilidad.

A continuación se muestra un ejemplo de cálculo de la entropía de una variable aleatoria con un alfabeto de más símbolos, en concreto de cinco símbolos.

Ejemplo

Una fuente puede ser modelada con el modelo DMS con un alfabeto

$$\mathcal{A}_X = \{-2, -1, 0, 1, 2\},$$

y unas probabilidades

$$p_X(-2) = \frac{1}{2}, p_X(-1) = \frac{1}{4}, p_X(0) = \frac{1}{8}, p_X(1) = \frac{1}{16}, p_X(2) = \frac{1}{16}.$$

En este caso la entropía vale

$$H(X) = \frac{1}{2} \log(2) + \frac{1}{4} \log(4) + \frac{1}{8} \log(8) + 2 \times \frac{1}{16} \log(16) = \frac{15}{8} \text{ bits/símbolo.}$$

Se puede ver que el valor de entropía depende únicamente de las probabilidades de los posibles elementos del alfabeto, y no de los valores concretos del alfabeto. Por ejemplo, una fuente con un alfabeto diferente

$$\mathcal{A}_X = \{0, 1, 2, 3, 4\},$$

pero con el mismo conjunto de valores para las probabilidades, aunque con distinta asignación a cada elemento del alfabeto

$$p_X(0) = \frac{1}{2}, p_X(1) = \frac{1}{4}, p_X(2) = \frac{1}{8}, p_X(3) = \frac{1}{16}, p_X(4) = \frac{1}{16},$$

tiene la misma entropía que la anterior.

5.3.2. Entropía conjunta

La definición de entropía se puede extender a más de una variable aleatoria, lo que tendría aplicación por ejemplo para medir la entropía conjunta de la entrada y la salida de un sistema de comunicaciones. La *entropía conjunta* de dos variables aleatorias X e Y , en general con alfabetos y probabilidades distintos, $\mathcal{A}_X = \{x_i\}_{i=0}^{M_X-1}$, $p_X(x_i)$, y $\mathcal{A}_Y = \{y_j\}_{j=0}^{M_Y-1}$, $p_Y(y_j)$, se define como una extensión trivial de la entropía de una variable aleatoria, considerando en este caso que hay tantos sucesos como casos conjuntos y que la probabilidad de cada uno de ellos está dada por la probabilidad conjunta, lo que lleva a la expresión

$$H(X, Y) = \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \left(\frac{1}{p_{X,Y}(x_i, y_j)} \right).$$

Al igual que para la entropía de una variable aleatoria, dadas las propiedades de la función logaritmo, se puede cambiar el signo e invertir el argumento del logaritmo,

$$H(X, Y) = - \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log (p_{X,Y}(x_i, y_j)).$$

Al igual que la entropía, también se mide en bits o nats por símbolo. El concepto se puede extender a N variables aleatorias. En este caso

$$\mathbf{X} = (X_1, X_1, \dots, X_N),$$

y

$$H(\mathbf{X}) = - \sum_{x_1, x_2, \dots, x_N} p_{\mathbf{X}}(x_1, x_2, \dots, x_N) \cdot \log(p_{\mathbf{X}}(x_1, x_2, \dots, x_N)).$$

En esta notación, el sumatorio indica el N sumatorios contemplando todas las posibles combinaciones de los alfabetos de cada variable aleatoria.

La interpretación de la entropía conjunta no difiere de la de entropía para una variable aleatoria. Al fin y al cabo, un par de variables X e Y pueden considerarse como una única variable aleatoria vectorial con un alfabeto de $M_X \cdot M_Y$ símbolos.

Si las variables aleatorias X e Y son independientes se cumple que

$$p_{X,Y}(x_i, y_j) = p_X(x_i) \times p_Y(y_j).$$

En este caso su entropía conjunta es la suma de las entropías individuales. Esto se había apuntado en la definición de las condiciones que debía cumplir la medida de información, y se demuestra de forma muy sencilla, como puede verse en el siguiente desarrollo

$$\begin{aligned} H(X, Y) &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_X(x_i) \cdot p_Y(y_j) \cdot \log \frac{1}{p_X(x_i) \cdot p_Y(y_j)} \\ &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_X(x_i) \cdot p_Y(y_j) \cdot \log \frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_X(x_i) \cdot p_Y(y_j) \cdot \log \frac{1}{p_Y(y_j)} \\ &= \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log \frac{1}{p_X(x_i)} + \sum_{j=0}^{M_Y-1} p_Y(y_j) \cdot \log \frac{1}{p_Y(y_j)} \\ &= H(X) + H(Y). \end{aligned}$$

Como veremos con más detalle en la siguiente sección, conviene remarcar que esta relación se cumple únicamente bajo la hipótesis de independencia entre las variables aleatorias.

5.3.3. Entropía condicional

El caso en que las variables aleatorias son independientes es en el que la combinación de variables aleatorias produce mayor entropía, ya que si ambas variables no fuesen independientes, el conocimiento del valor de una de ellas eliminaría incertidumbre sobre el valor de la otra. Para medir esa incertidumbre se define la *entropía condicional* de dos variables aleatorias X e Y , $H(X|Y)$, que promedia el valor de la entropía condicional de X dado Y sobre todos los valores del alfabeto de Y , de forma que se define como

$$H(X|Y) = \sum_{j=0}^{M_Y-1} p_Y(y_j) \cdot H(X|Y = y_j).$$

Desarrollando esta expresión para la entropía $H(X|Y = y_j)$ a partir de la distribución condicional de X dada Y , $p_{X|Y}(x_i|y_j)$, se llega a la expresión equivalente

$$\begin{aligned} H(X|Y) &= \sum_{j=0}^{M_Y-1} p_Y(y_j) \sum_{i=0}^{M_X-1} p_{X|Y}(x_i|y_j) \cdot \log \frac{1}{p_{X|Y}(x_i|y_j)} \\ &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_{X|Y}(x_i|y_j)}. \end{aligned}$$

De acuerdo con la regla de Bayes, esta probabilidad cumple la relación

$$p_{X|Y}(x_i|y_j) \cdot p_Y(y_j) = p_{X,Y}(x_i, y_j).$$

En general, se puede extender de forma natural esta definición cuando el condicionamiento es con respecto a varias variable aleatorias

$$\begin{aligned} H(X_N|X_1, X_2, \dots, X_{N-1}) &= \\ &= - \sum_{x_1, x_2, \dots, x_N} p_{\mathbf{X}}(x_1, x_2, \dots, x_N) \log p_{X_N|X_1, X_2, \dots, X_{N-1}}(x_N|x_1, x_2, \dots, x_{N-1}). \end{aligned}$$

La entropía condicional puede interpretarse como una medida de la incertidumbre que se tiene sobre el valor de una variable aleatoria, X , cuando se conoce el valor de otra variable aleatoria, Y . O de otra forma, la comparación entre $H(X)$ y $H(X|Y)$ cuantifica la información que el conocimiento de Y aporta sobre el conocimiento de X . Cuando las variables aleatorias X e Y son independientes, el conocimiento del valor de una de ellas no aporta conocimiento sobre la otra y por tanto no elimina incertidumbre sobre su valor. Por tanto, en este caso

$$H(X|Y) = H(X).$$

Por el contrario, si el conocimiento de Y determina por completo el valor de X , conocido el valor de Y no hay ninguna incertidumbre sobre el valor de X , y la entropía condicional sería $H(X|Y) = 0$.

La entropía conjunta se relaciona con la entropía y con la entropía condicional a través de la siguiente expresión

$$\begin{aligned} H(X, Y) &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_{X,Y}(x_i, y_j)} \\ &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_X(x_i) \cdot p_{Y|X}(y_j|x_i)} \\ &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_{Y|X}(y_j|x_i)} \\ &= \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log \frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_{Y|X}(y_j|x_i)} \\ &= H(X) + H(Y|X). \end{aligned}$$

A través de un desarrollo equivalente, es fácil demostrar que también se cumple la relación

$$H(X, Y) = H(Y) + H(X|Y).$$

Esto no quiere decir que $H(X|Y)$ sea igual a $H(Y|X)$ en general. Esta relación sólo se cumple cuando $H(X) = H(Y)$.

Es decir, que la entropía conjunta se obtiene como la suma de la entropía de una variable aleatoria más la de la otra condicionada a la primera. Se suma por tanto la incertidumbre de una de las variables aleatorias con la que le queda a la otra cuando se conoce la primera. Esto significa, como ya se ha visto anteriormente, que la entropía conjunta sólo será igual a la suma de la entropía de cada una de las variables aleatorias cuando las variables aleatorias sean independientes.

En general, se puede extender esta relación al caso de un mayor número de variables aleatorias, en cuyo caso aplicando la regla de la cadena se tiene la siguiente relación general

$$H(\mathbf{X}) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \dots + H(X_N|X_1, X_2, \dots, X_{N-1}).$$

Cuando (X_1, X_2, \dots, X_N) son variables aleatorias independientes, la entropía conjunta de todas ellas será la suma de la entropía de cada una de las variables aleatorias

$$H(\mathbf{X}) = \sum_{i=1}^N H(X_i).$$

5.3.4. Información mutua

La entropía representa una medida de incertidumbre sobre el valor de una o varias variables aleatorias. Otro concepto que podríamos definir como “complementario” es la denominada *información mutua* entre dos variables aleatorias X e Y , que se denota $I(X, Y)$. La información mutua representa la información que aporta Y sobre el conocimiento de X . La definición formal a partir de las distribuciones marginales de cada variable aleatoria y de su distribución conjunta es

$$I(X, Y) = \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{p_{X,Y}(x_i, y_j)}{p_X(x_i) \cdot p_Y(y_j)},$$

y se mide en bits.

La información mutua es una medida no negativa, $I(X, Y) \geq 0$, que se puede expresar en términos de la entropía y de la entropía condicional, ya que se cumple la siguiente relación:

$$\begin{aligned} I(X, Y) &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{p_{X,Y}(x_i, y_j)}{p_X(x_i) \cdot p_Y(y_j)} \\ &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{p_{X|Y}(x_i, y_j)}{p_X(x_i)} \\ &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_X(x_i)} + \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log(p_{X|Y}(x_i, y_j)) \\ &= \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log \frac{1}{p_X(x_i)} - \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \cdot \log \frac{1}{p_{X|Y}(x_i, y_j)} \\ &= H(X) - H(X|Y). \end{aligned}$$

Un desarrollo equivalente demuestra también que

$$I(X, Y) = H(Y) - H(Y|X).$$

Por otro lado, de la propia definición de la información mutua se tiene la propiedad

$$I(X, Y) = I(Y, X),$$

y teniendo en cuenta la relación entre entropías marginal, condicional y conjunta

$$H(X, Y) = H(Y) + H(X|Y),$$

se comprueba de forma trivial que la información mutua se puede también obtener como

$$I(X, Y) = H(X) + H(Y) - H(X, Y).$$

Estas relaciones entre información mutua y las distintas entropías se representan habitualmente de forma gráfica mediante un *diagrama de Venn*, como el que se muestra en la Figura 5.20.

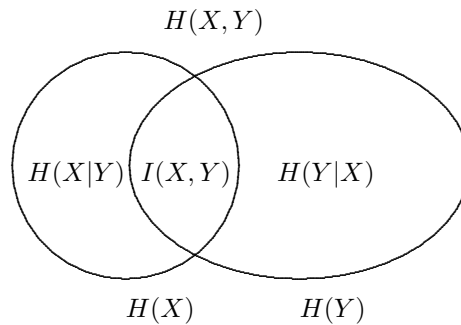


Figura 5.20: Diagrama de Venn que ilustra las relaciones entre entropías e información mutua.

Cada uno de los conjuntos representa la entropía de cada variable, $H(X)$ y $H(Y)$, y la intersección entre ellos es la información mutua $I(X, Y)$, mientras que la unión es la entropía conjunta, $H(X, Y)$. La diferencia entre el conjunto completo y la intersección es la entropía condicional. Si las variables aleatorias son independientes, los conjuntos tendrían intersección nula. Si ambas variables son iguales, la intersección es completa y la entropía condicional es nula. La Figura 5.21 muestra con más claridad como se identifica cada una de las medidas sobre el diagrama de Venn con un sencillo código de colores.

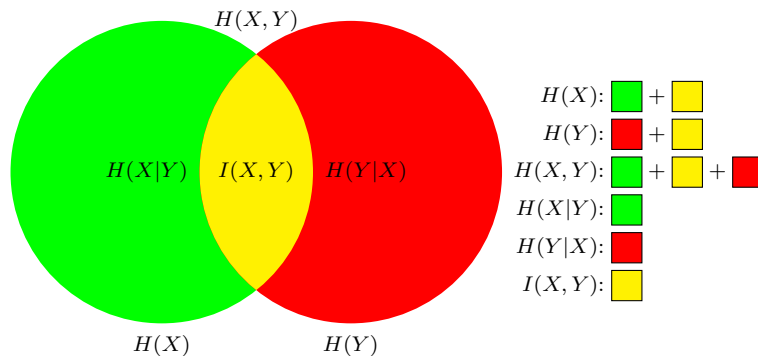


Figura 5.21: Identificación con un sencillo código de colores de las distintas entropías y la información mutua sobre el diagrama de Venn.

A continuación se muestra un ejemplo en el que se calculan las distintas entropías y la información mutua para dos variables aleatorias sencillas.

Ejemplo

Se tienen dos variables aleatorias binarias, X e Y , con idéntico alfabeto $x_0 = y_0 = 0$, $x_1 = y_1 = 1$, y con la siguiente distribución conjunta

$$p_{X,Y}(0,0) = \frac{1}{3}, p_{X,Y}(0,1) = \frac{1}{3}, p_{X,Y}(1,0) = \frac{1}{3}, p_{X,Y}(1,1) = 0.$$

Para calcular la entropía de cada variable aleatoria hay que conocer las distribuciones marginales, que se obtienen de forma sencilla a partir de la distribución conjunta

$$p_X(x_i) = \sum_{j=0}^{M_Y-1} p_{X,Y}(x_i, y_j) \text{ y } p_Y(y_j) = \sum_{i=0}^{M_X-1} p_{X,Y}(x_i, y_j).$$

En este caso

$$p_X(0) = p_Y(0) = \frac{2}{3}, p_X(1) = p_Y(1) = \frac{1}{3}.$$

Por tanto, como en este caso las dos variables tienen la misma distribución, la entropía de ambas variables, parametrizada a través de la función de entropía binaria (al ser las variables aleatorias binarias)

$$H(X) = H(Y) = H_b\left(\frac{2}{3}\right) = H_b\left(\frac{1}{3}\right) = 0,919.$$

La entropía conjunta, aplicando su definición, vendrá dada por

$$H(X, Y) = 3 \times \left(\frac{1}{3} \log_2(3)\right) + 0 \log_2(0) = \log_2(3) = 1,585.$$

Se puede interpretar también este resultado como que (X, Y) es un vector de variables aleatorias con un alfabeto de tres eventos, $(0, 0)$, $(0, 1)$ y $(1, 0)$, todos ellos igualmente probables.

A partir de los resultados anteriores, la entropía condicional se puede obtener a través de la relación

$$H(X|Y) = H(X, Y) - H(Y) = 1,585 - 0,919 = 0,666.$$

Del mismo modo, la información mutua podría obtenerse, por ejemplo, a través de la relación

$$I(X, Y) = H(X) - H(X|Y) = 0,919 - 0,666 = 0,253.$$

La información mutua entre variables aleatorias discretas tiene una serie de propiedades que conviene tener en cuenta. Entre ellas cabe destacar las siguientes:

1. Es siempre no negativa

$$I(X, Y) = I(Y, X) \geq 0.$$

El valor mínimo $I(X, Y) = 0$ se obtiene en el caso en que X e Y son independientes.

2. Su valor máximo está acotado por el valor de la entropía de cada una de las variables aleatorias, con lo que en la práctica está acotado por el valor mínimo de la entropía de las variables aleatorias

$$I(X, Y) \leq \min(H(X), H(Y)).$$

La información mutua nunca puede ser mayor que la medida de información que tiene cada una de las variables.

3. Se puede definir información mutua condicional como el promedio de la información mutua condicionada a cada uno de los posibles valores de la variable aleatoria con respecto a la que se condiciona

$$I(X, Y|Z) = \sum_{i=0}^{M_Z-1} p_Z(z_i) \cdot I(X, Y|Z = z_i).$$

4. La información mutua condicional $I(X, Y|Z)$ también se puede obtener a través de las entropías condicionales como

$$I(X, Y|Z) = H(X|Z) - H(X|Y, Z).$$

5. La regla de la cadena para la información mutua se define a partir de la relación

$$I((X, Y), Z) = I(X, Z) + I(Y, Z|X).$$

6. En general, la regla de la cadena es

$$I((X_1, X_2, \dots, X_N), Y) = I(X_1, Y) + I(X_2, Y|X_1) + \dots + I(X_N, Y|X_1, \dots, X_{N-1}).$$

7. A partir de la definición de información mutua se obtiene la definición de entropía como información mutua de una variable aleatoria consigo misma. Esta relación se demuestra de forma sencilla teniendo en cuenta que la distribución de una variable aleatoria consigo misma toma la forma

$$p_{X,X}(x_i, x_j) = \delta[i - j] \cdot p_X(x_i),$$

por lo que la información mutua de la variable aleatoria X consigo misma es

$$\begin{aligned} I(X, X) &= \sum_{i=0}^{M_X-1} \sum_{j=0}^{M_X-1} \delta[i - j] \cdot p_X(x_i) \cdot \log \frac{\delta[i - j] \cdot p_X(x_i)}{p_X(x_i) \cdot p_X(x_j)} \\ &= \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log \frac{p_X(x_i)}{p_X(x_i) \cdot p_X(x_i)} \\ &= \sum_{i=0}^{M_X-1} p_X(x_i) \cdot \log \frac{1}{p_X(x_i)} \\ &= H(X). \end{aligned}$$

De aquí es de donde proviene el nombre de *auto-información* (la entropía es el promedio de la auto-información), y a la entropía en ocasiones se la conoce con este nombre.

5.3.5. Entropía diferencial e información mutua

Hasta ahora, las medidas que se han presentado hacen referencia a variables aleatorias discretas, lo que sirve para modelar fuentes de información discretas en el tiempo y con un alfabeto discreto. Para estas variables se han presentado la entropía, $H(X)$ y la información mutua $I(X, Y)$, así como las entropías condicional y conjunta, $H(X|Y)$ y $H(X, Y)$.

Para modelar una fuente discreta en el tiempo pero con alfabeto continuo, por ejemplo una fuente de audio muestreada, es preciso utilizar una variable aleatoria continua. En este caso, el análogo de la entropía para variables aleatorias discretas se llama *entropía diferencial*. Sin embargo, esta medida no tiene el significado intuitivo que tenía la entropía, lo que se debe a varios aspectos, como por ejemplo el hecho de que en una variable aleatoria continua, por definición la probabilidad de un número continuo concreto es nula.

Formalmente, la entropía diferencial de una variable aleatoria continua X , con una función densidad de probabilidad $f_X(x)$, se define como

$$h(X) = \int_{-\infty}^{\infty} f_X(x) \log \frac{1}{f_X(x)} dx,$$

donde de nuevo se considera $0 \log(1/0) = 0$.

Ejemplo

Se desea determinar la entropía diferencial de una variable aleatoria uniformemente distribuida en un intervalo $[0, a]$.

Utilizando directamente la definición de entropía diferencial, y teniendo en cuenta que la función densidad de probabilidad de la variable aleatoria vale $1/a$ entre 0 y a , la entropía vale

$$h(X) = \int_0^a \frac{1}{a} \log(a) dx = \log(a).$$

A partir de este ejemplo se pueden observar algunas propiedades interesantes:

1. Para $a < 1$ se tiene $h(X) < 0$, lo que va en contra de la propiedad de no negatividad de la entropía de una variable aleatoria discreta.
2. Para $a = 1$, se tiene $h(X) = 0$, y en este caso X no es determinista, con lo que tiene un cierto grado de incertidumbre. Esto va también en contra de las propiedades de la entropía para variables aleatorias discretas.

Ejemplo

Se tiene una variable aleatoria X con una función densidad de probabilidad gaussiana con media nula y varianza σ^2

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}.$$

Utilizando logaritmos neperianos, la entropía diferencial para esta variable aleatoria en nats se calcula como

$$\begin{aligned} h(X) &= - \int_{-\infty}^{\infty} f_X(x) \ln \frac{1}{\sqrt{2\pi\sigma^2}} dx - \int_{-\infty}^{\infty} f_X(x) \ln \left(e^{-\frac{x^2}{2\sigma^2}} \right) dx \\ &= \ln(\sqrt{2\pi\sigma^2}) + \frac{\sigma^2}{2\sigma^2} = \ln(\sqrt{2\pi\sigma^2}) + \frac{1}{2} \\ &= \frac{1}{2} \ln(2\pi e\sigma^2) \text{ nats.} \end{aligned}$$

Para llegar a esta expresión se han utilizado las siguientes propiedades sobre la distribución gaussiana

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \text{ y } \int_{-\infty}^{\infty} x^2 f_X(x) dx = \sigma^2.$$

Cambiando la base del logaritmo a 2, se tiene la entropía diferencial en bits

$$h(X) = \frac{1}{2} \log_2(2\pi e\sigma^2) \text{ bits.}$$

Dependiendo del valor de la varianza, en particular comparándolo con un valor $\sigma^2 = \frac{1}{2\pi e}$, esta entropía puede tomar valores positivos, negativos o nulo.

Al igual que para variables aleatorias discretas, también se definen para variables aleatorias continuas entropías conjuntas y condicionales. Se define la *entropía diferencial conjunta* entre dos variables aleatorias X e Y como

$$h(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log \frac{1}{f_{X,Y}(x, y)} dx dy.$$

En cuanto a la *entropía diferencial condicional*, su definición es

$$h(X|Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \log \frac{1}{f_{X|Y}(x|y)} dx dy.$$

A menudo se utiliza la definición alternativa pero equivalente

$$h(X|Y) = \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} f_{X|Y}(x|y) \log \frac{1}{f_{X|Y}(x|y)} dx dy.$$

Se puede observar que son las extensiones naturales de las definiciones para variables aleatorias discretas. Por tanto se cumplen las mismas relaciones. En concreto,

$$h(X, Y) = h(Y) + h(X|Y).$$

Del mismo modo se define también la *información mutua* para variables aleatorias continuas como

$$I(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \log \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)} dx dy,$$

que también se mide en bits, y donde para evitar ambigüedades se define $0 \log \frac{0}{0} = 0$.

Como en el caso de variables aleatorias discretas, la información mutua se puede expresar en función de las entropías

$$I(X, Y) = h(Y) - h(Y|X) = h(X) - h(X|Y) = h(X) + h(Y) - h(X, Y).$$

Al contrario que en el caso de la entropía diferencial para variables aleatorias continuas, donde no se mantiene la interpretación intuitiva de entropía para variables aleatorias discretas como medida de incertidumbre o de información, para la información mutua si se mantiene y tiene el mismo significado; es decir, la información mutua indica el conocimiento que aporta una variable sobre la otra. Además se mantienen la mayoría de propiedades básicas de la información mutua para variables aleatorias discretas. En particular, dada su definición se cumplen las siguientes propiedades:

1. $I(X, Y) \geq 0$, es decir, es una función no negativa.
2. $I(X, Y) = 0$ sólo si las variables X e Y son independientes.
3. $I(Y, X) = I(X, Y)$.

5.4. Capacidad de canal

Una vez definidos los conceptos de entropía e información mutua, en esta sección se va a tratar de determinar la cantidad máxima de información que se puede transmitir por un canal haciendo uso de esas medidas de información. En primer lugar se introducirá el concepto de codificación de canal como mecanismo para lograr una comunicación fiable a través de un cierto canal no fiable, para a continuación definir la capacidad de un canal y estudiar cómo se obtiene dicho valor para dos tipos de canales: el canal digital modelado mediante un DMC y el canal gaussiano.

5.4.1. Codificación de canal para una transmisión fiable

El principal objetivo cuando se transmite información sobre cualquier canal de comunicaciones es la *fiabilidad*. Esta fiabilidad en sistemas de comunicaciones digitales se mide mediante la probabilidad de error en el receptor. Como en todo canal de comunicaciones, aparte de las posibles distorsiones, se introduce ruido. A primera vista puede parecer que la probabilidad de error siempre estará acotada por un valor no nulo que dependerá del nivel de ruido, es decir

$$P_e \geq K = f(\sigma^2),$$

donde σ^2 es la potencia del ruido a la entrada del receptor. Sin embargo, como ya se esbozó en la introducción, un resultado fundamental de la teoría de la información dice que la transmisión fiable, entendiéndose por transmisión fiable aquella en la que se tiene una probabilidad de error por debajo de cualquier límite fijado, es posible incluso en canales ruidosos siempre que la velocidad de transmisión o tasa de transmisión esté por debajo de un determinado valor denominado *capacidad de canal*. Este resultado fue presentado por Shannon en 1948 y es lo que se denomina *teorema de codificación de canal*, o en inglés *noisy channel-coding theorem*. A modo de resumen, lo que dice el teorema de codificación de canal es que *la limitación básica que introduce el ruido en un canal de comunicaciones no está en la fiabilidad de la comunicación, sino en la velocidad de la misma*. Así, será posible obtener una comunicación con una probabilidad de error arbitrariamente baja siempre y cuando se transmita información a una tasa por debajo de un valor límite que dependerá de las características del canal.

En primer lugar se va a ver cómo es posible obtener este límite de una forma intuitiva, siguiendo la misma línea argumental que se sigue en [Proakis y Salehi, 2002], para luego formularlo en términos de la teoría de la información.

La Figura 5.22 muestra un canal discreto sin memoria (DMC) con un alfabeto de entrada formado por cuatro elementos, $\mathcal{A}_X = \{a, b, c, d\}$ y el mismo alfabeto de salida $\mathcal{A}_Y = \mathcal{A}_X$.

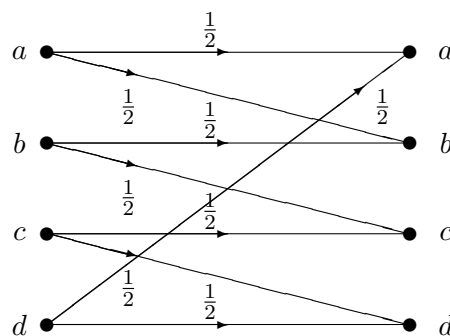


Figura 5.22: Ejemplo de un canal discreto sin memoria.

Dado que no se trata de un canal ideal, si se utiliza este canal para la transmisión en la forma habitual, el receptor no podrá identificar con probabilidad de error nula el símbolo transmitido a la vista del símbolo recibido. Cuando en la salida se tiene el valor a , el receptor no puede discernir si se ha transmitido a o b , ya que la transmisión de ambos puede producir a la salida del canal el símbolo a . Lo mismo ocurre con el resto de posibles valores de salida. Si se observa b , no es posible discernir sin error si se debe a la transmisión de a o de b , y así sucesivamente. Por tanto, hay una probabilidad de error que viene fijada por las características del canal, en este caso por las probabilidades de transición que lo definen.

Sin embargo, dadas las especiales características de este canal, va a ser posible transmitir información sin errores a través del mismo. Es evidente que la imposibilidad de discernir sin error el símbolo transmitido a la vista del valor de salida se debe a que los conjuntos que forman los posibles valores de salida asociados a la transmisión de cada símbolo se “solapan”. Así, existe un solapamiento entre las salidas que se pueden dar cuando se transmite a (que pueden ser a y b) y cuando se transmite b (que pueden ser b y c). Por eso cuando se observa b en la salida no es posible saber con absoluta certeza qué símbolo se ha transmitido. Pero para este canal es posible elegir un subconjunto de elementos del alfabeto de entrada cuyas salidas no se solapan. Por ejemplo, si sólo se transmiten los símbolos a y c , a la vista de la salida no habrá ambigüedad posible sobre el símbolo transmitido; si en la salida se tienen a o b , se sabe con certeza que el símbolo transmitido es a ; lo mismo sucede si en la salida se tienen c o d , en cuyo caso se tiene la certeza de que se ha transmitido c . Por tanto, transmitiendo ese subconjunto de posibles valores de X la probabilidad de error es nula. El precio que se paga por esta fiabilidad es la tasa de transmisión. Transmitiendo los 4 símbolos que forman parte del alfabeto de X en cada uso del canal se transmiten dos bits de información. Sin embargo, si sólo se transmiten 2 posibles valores, en cada uso del canal se transmitirá únicamente un bit de información (recuerde que el número de bits de información por símbolo transmitido es $\log_2 M$, siendo M el número de elementos del alfabeto de símbolos que se transmite).

El mecanismo utilizado en este ejemplo para poder transmitir con probabilidad de error nula ilustra la idea fundamental subyacente en el teorema de codificación de canal para poder transmitir de forma fiable: utilizar en la transmisión sólo aquellos símbolos cuyas salidas correspondientes sean disjuntas. Aquí es preciso matizar que el objetivo de la codificación de canal no es en realidad conseguir una transmisión con probabilidad de error nula, sino con una probabilidad de error por debajo de un cierto valor, que puede ser arbitrariamente bajo. Así, para conseguir una transmisión, si no con probabilidad de error nula, con una probabilidad de error arbitrariamente baja, se podrán utilizar en la transmisión sólo aquellos símbolos que, si no tienen salidas asociadas disjuntas, tengan salidas cuyo solapamiento se produzca con una probabilidad suficientemente baja.

Un problema que surge con esta idea para la transmisión con probabilidad de error arbitrariamente baja, es que en la práctica los canales reales no tienen un comportamiento como el de la Figura 5.22, donde hay símbolos de entrada cuyas salidas no se solapan; y en la gran mayoría de los casos ni siquiera hay un subconjunto de símbolos cuyas salidas se solapan con una probabilidad suficientemente baja. Sin embargo, la teoría de la codificación de canal propone un sencillo mecanismo para generar artificialmente una situación similar a esta. Aunque este mecanismo se puede utilizar para canales digitales con alfabetos M -arios, por simplicidad se ilustrará este mecanismo poniendo como ejemplo un caso binario; en concreto el modelo de canal binario simétrico o BSC, en el que se denotará mediante ε la probabilidad de error de bit (BER), como se muestra en el diagrama de rejilla de la izquierda en la Figura 5.23.

Si se observa el canal BSC y se intenta aplicar el procedimiento aplicado sobre el canal de la Figura 5.22, se ve que directamente no es posible, ya que las salidas de los dos símbolos se solapan por completo con una probabilidad arbitraria (dada por ε) y porque además sólo hay dos símbolos (si sólo se transmite uno de ellos, no habrá en realidad información en la transmisión). En general, no va a ser posible aplicar ese procedimiento de forma directa para casi ningún canal DMC real. Para utilizar esta idea, lo que se hace es aplicarla no directamente sobre el canal sino sobre el denominado *canal extendido* de orden n . El canal extendido de orden n se define como un canal en el que se agrupan bloques de n símbolos para formar símbolos extendidos (de orden n), dando lugar a unos alfabetos de entrada \mathcal{A}_X^n y \mathcal{A}_Y^n . La idea es transmitir la información no en cada uso individual del canal, sino de forma conjunta en n usos del canal. El diagrama de la derecha de la Figura 5.23 ilustra la idea de la extensión de orden $n = 3$ para el BSC. La información se

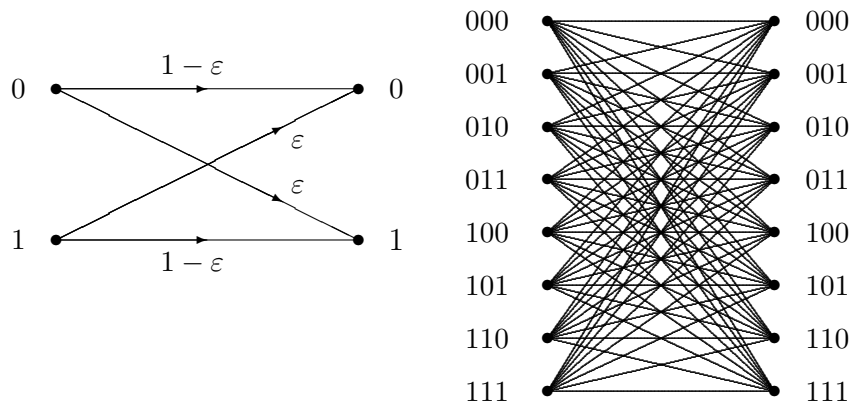


Figura 5.23: Canal binario simétrico (BSC) y su correspondiente canal extendido de orden $n = 3$.

transmitirá por bloques haciendo 3 usos del canal, de forma que el alfabeto de este canal extendido está ahora formado por 8 posibles símbolos extendidos, correspondientes a los 8 posibles valores que pueden tomar los símbolos del canal original en los 3 usos que definen el canal extendido. Se ha pasa por tanto de trabajar con un sistema con un alfabeto de dos símbolos

$$\mathcal{A}_X = \mathcal{A}_Y = \{0, 1\},$$

a un sistema extendido con alfabeto de 2^n símbolos, en este caso $2^3 = 8$ símbolos

$$\mathcal{A}_X^3 = \mathcal{A}_Y^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}.$$

Las probabilidades de transición sobre este canal extendido se obtienen como productos de las n probabilidades de transición a través del canal original asociadas a cada caso sobre el canal extendido. Si se definen los símbolos extendidos como vectores de n elementos, representados estadísticamente por las variables aleatorias vectoriales \mathbf{X} y \mathbf{Y} , con alfabetos

$$\mathbf{X} \in \{\mathbf{x}_i\}_{i=0}^{M_X^n - 1}, \mathbf{Y} \in \{\mathbf{y}_j\}_{j=0}^{M_Y^n - 1}$$

con

$$\mathbf{x}_i = [x_i[0], x_i[1], \dots, x_i[n - 1]], \mathbf{y}_j = [y_j[0], y_j[1], \dots, y_j[n - 1]]$$

las probabilidades de transición sobre los símbolos extendidos son

$$p_{\mathbf{Y}, \mathbf{X}}(\mathbf{y}_j | \mathbf{x}_i) = \prod_{\ell=0}^{n-1} p_{Y|X}(y_j[\ell] | x_i[\ell]).$$

En la figura no se han etiquetado las ramas con las probabilidades de transición por falta de espacio en la misma, pero dichas probabilidades se obtienen de forma muy sencilla, habiendo 4 posibles valores:

- Si el símbolo extendido de entrada y el de salida coinciden, la probabilidad de transición asociada es

$$p_{\mathbf{Y}, \mathbf{X}}(\mathbf{y}_i | \mathbf{x}_i) = (1 - \varepsilon)^3.$$

- Si entre el símbolo extendido de entrada y el de salida la distancia de Hamming es 1 (cambia sólo un bit), la probabilidad de transición asociada es

$$p_{\mathbf{Y}, \mathbf{X}}(\mathbf{y}_j | \mathbf{x}_i) = \varepsilon \cdot (1 - \varepsilon)^2.$$

- Si entre el símbolo extendido de entrada y el de salida la distancia de Hamming es 2 (cambian dos bits), la probabilidad de transición asociada es

$$p_{\mathbf{Y},\mathbf{X}}(\mathbf{y}_j|\mathbf{x}_i) = \varepsilon^2 \cdot (1 - \varepsilon).$$

- Por último, si entre el símbolo extendido de entrada y el de salida la distancia de Hamming es 3 (cambian los tres bits), la probabilidad de transición asociada es

$$p_{\mathbf{Y},\mathbf{X}}(\mathbf{y}_j|\mathbf{x}_i) = \varepsilon^3.$$

Básicamente, la probabilidad de que un bit coincida entre entrada y salida es $1 - \varepsilon$, y la de que sea distinta es ε , de donde de forma directa se obtienen estas cuatro probabilidades.

Evidentemente, para probabilidades de error de bit ε bajas, las dos primeras probabilidades son mucho mayores que las dos últimas (es mucho más probable tener cero o un error sobre los tres bits, que tener dos o tres errores). Se podrían por tanto dividir las transiciones en transiciones de alta probabilidad y de baja probabilidad. Esta división se muestra en la Figura 5.24 mediante un código de colores: en negro se representan las transiciones de alta probabilidad y en verde las de baja probabilidad.

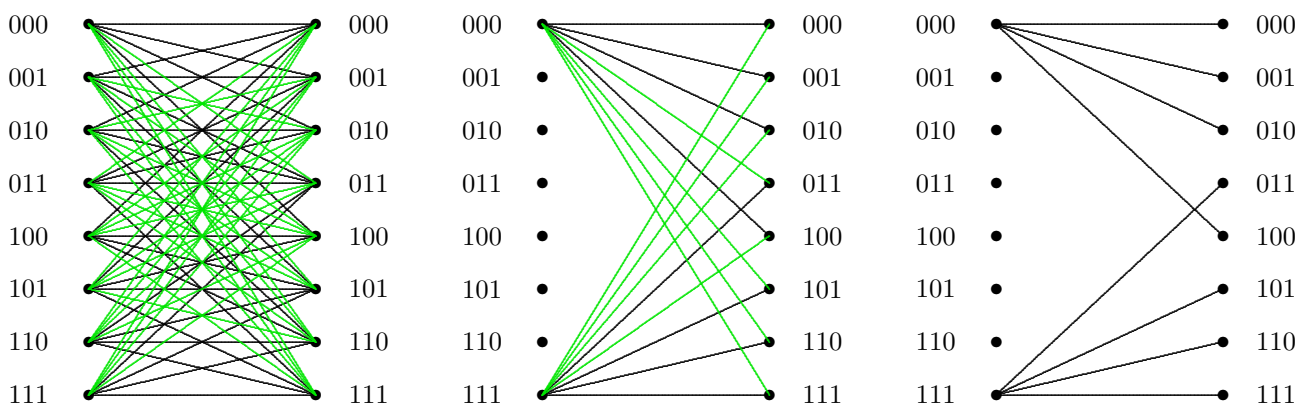


Figura 5.24: Distinción entre transiciones de alta (negro) y de baja (verde) probabilidad y selección de un subconjunto de símbolos extendidos sin solapamiento en los enlaces e alta probabilidad.

Hecha esa distinción, ahora es posible buscar un subconjunto de símbolos extendidos a la entrada cuyas salidas sean disjuntas en cuanto a las transiciones de alta probabilidad, lo que sucede por ejemplo con los símbolos extendidos 000 y 111, como se ilustra en la figura. Ahora, cuando se transmite uno de los dos símbolos extendidos, si se producen sólo las transiciones más probables, a partir de la salida se identificará correctamente el símbolo extendido transmitido. Obviamente, también se puede producir alguna transición de baja probabilidad, en cuyo caso se producirá un error al identificar el símbolo extendido transmitido. En ese caso, la probabilidad de error vendrá dada por la probabilidad de que se produzca alguna de las transiciones poco probables, que para este ejemplo será

$$P_e = 1 \times (1 - \varepsilon)^3 + 3 \times \varepsilon \cdot (1 - \varepsilon)^2.$$

Este valor se obtiene teniendo en cuenta que las transiciones de baja probabilidad son las que correspondientes a dos o tres errores de bit en la transmisión, y que sólo hay un patrón de tres errores y tres de dos errores para cada bloque de tres bits transmitidos.

Como ejemplo numérico para ilustrar el beneficio obtenido, si un sistema de comunicaciones binario tiene una BER $\varepsilon = 0,1$, uno de cada 10 bits será erróneo (porcentaje del 10%). Usando

como base ese sistema y una extensión como la del ejemplo de orden $n = 3$, la probabilidad de error pasará a ser $P_e = 0,028$; es decir, se puede conseguir un porcentaje del 2,8 % de bits erróneos usando como base para la transmisión un sistema con un porcentaje de error del 10 %. Para el caso $\varepsilon = 0,01$ el porcentaje de bits erróneos del sistema será del 1 %. En ese caso $P_e = 2,98 \times 10^{-4}$, lo que significa un porcentaje de error de aproximadamente el 0,03 %.

Esta técnica permite reducir la probabilidad de error de un sistema, pero naturalmente lo hace a costa de algo. Y ese algo es la tasa de transmisión de información. Si se usa el sistema sin la extensión, en cada uso del canal se transmite un bit de información. Si se usa el sistema extendido, como sólo se transmiten 2 de los 8 posibles símbolos extendidos, cada uno de ellos transportará un bit de información real, por ejemplo 000 transportará el “0” y 111 transportará el “1”. De esta forma, en cada 3 usos del canal se enviará un sólo bit de información, con lo que la tasa efectiva, que sería 1/3 (1 bit de información por cada 3 usos del canal), ha disminuido respecto a la transmisión directa sin la extensión.

La elección del subconjunto de elementos sin solapamientos en transiciones de baja probabilidad no tiene por qué restringirse a dos símbolos extendidos. Sería por ejemplo posible hacer una extensión de orden $n = 5$, y elegir 4 símbolos extendidos cuyas transiciones de alta probabilidad (definidas como aquellas en las que se producen cero errores o un error de bit sobre los cinco bits transmitidos) no se solapen; podrían ser por ejemplo los símbolos 00000, 10101, 01110 y 11011. En este caso, al haber 4 símbolos extendidos, cada 5 usos del canal se enviarán 2 bits de información real, por lo que la tasa de transmisión efectiva de información será 2/5.

Codificación de canal

En general a esta técnica basada en la definición de símbolos extendidos de orden n que se utiliza para poder realizar una transmisión con una probabilidad de error suficientemente baja sobre un sistema que inherentemente tiene una probabilidad de error mayor se le denomina *codificación de canal*. La técnica no está limitada a su utilización sobre sistemas binarios, como en los ejemplos que se han utilizado para su presentación, sino que se puede emplear también sobre sistemas M -arios. La Figura 5.25 ilustra el funcionamiento de un sistema que utiliza la codificación de canal como mecanismo para controlar la probabilidad de error del sistema.

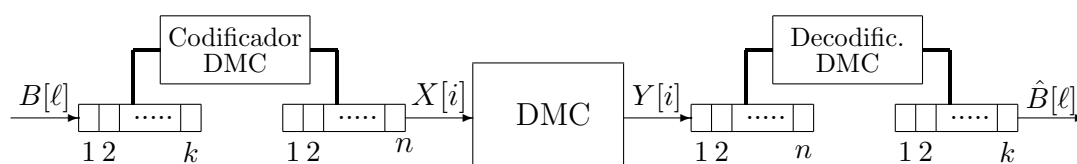


Figura 5.25: Codificador y decodificador de canal para la transmisión sobre un canal digital modelado mediante un canal discreto sin memoria.

Un sistema de codificación de canal está compuesto por un codificador en el lado del transmisor, que lleva a cabo la *codificación de canal* y un decodificador en el lado del receptor, que lleva a cabo la *decodificación de canal*. El codificador tiene como entrada un vector de k símbolos M -arios, que define un alfabeto de entrada, formalmente denominado *conjunto de índices*, de M^k elementos. Este conjunto de índices estaría compuesto en los ejemplos anteriores por los bloques de información efectiva que se va a transmitir en cada n usos del canal. Ante cada entrada, el codificador genera como salida un vector de n símbolos de entrada al DMC, haciendo corresponder a cada una de las combinaciones de k símbolos M -arios que forman el conjunto de índices, una combinación de

n símbolos del DMC ó *palabra código*. En los ejemplos anteriores, estas palabras código serían el subconjunto de símbolos extendidos cuyas salidas se solapan con baja probabilidad. En estas condiciones, se dice que este es un código $\mathcal{C}(k, n)$, donde al valor n se le suele llamar longitud del código. El conjunto de todas las palabra código (M^k), se denomina *diccionario del código*. El decodificador funciona de forma recíproca. Formalmente un código queda definido por el conjunto de índices y las funciones de codificación y decodificación. Los diccionarios del código para los dos ejemplos anteriores, que en ambos casos eran códigos binarios, serían los que se muestran en la Tabla 5.1.

Conjunto de índices	Palabras código	Conjunto de índices	Palabras código
0	000	00	00000
1	111	01	10101
		10	01110
		11	11011

Código de ejemplo $\mathcal{C}(1, 3)$
 Código de ejemplo $\mathcal{C}(2, 5)$

Tabla 5.1: Diccionarios del código para los dos ejemplos binarios que se han utilizado para ilustrar la técnica de codificación de canal.

Un parámetro muy importante en todo código de canal es el que define el cociente entre el número de símbolos a la entrada del codificador, k , y el número de veces que se utiliza el DMC para transmitir la palabra código, n , ya que define la cantidad de información real que se transmite en cada uso de canal. Este parámetro se denomina *tasa de transmisión* o simplemente *tasa del código*. Habitualmente se denota por R y se mide en símbolos por uso del DMC

$$R = \frac{k}{n} \text{ símbolos por uso.}$$

Para sistemas binarios, las unidades de esta tasa del código son bits por uso del canal.

Teorema de codificación de canal

Al presentar el principio de la codificación de canal se ha visto que hay un compromiso entre la tasa de transmisión o tasa del código y las prestaciones obtenidas. De forma natural surgen algunas preguntas respecto a las prestaciones que se pueden obtener con esta técnica. ¿Es posible reducir la probabilidad de error tanto como se desee utilizando esta técnica sobre cualquier tipo de canal? ¿Qué limitaciones imponen las características del canal sobre las prestaciones que puede ofrecer la capacidad de canal? La respuesta a estas preguntas está en el denominado *teorema de codificación de canal*, presentado por Claude Shannon en 1948.

El teorema de codificación de canal demuestra que existe un límite para la máxima tasa a la que se puede transmitir por un canal discreto sin memoria, que es el que se denomina *capacidad de canal*, y que se obtiene formalmente como el valor máximo, sobre todas las posibles distribuciones para el alfabeto de entrada del canal, de la información mutua entre la entrada y la salida del canal, es decir

$$C = \max_{p_X(x_i)} I(X, Y),$$

donde $I(X, Y)$ es la información mutua entre la entrada X y la salida Y del canal. El teorema además demuestra los siguientes aspectos:

1. Si la tasa de transmisión R es menor que la capacidad del canal C , entonces para cualquier valor arbitrariamente bajo $\delta > 0$ existe un código con una longitud de bloque n suficientemente larga cuya probabilidad de error es menor que δ .
2. Si $R > C$, la probabilidad de error de cualquier código con cualquier longitud de bloque está limitada por un valor no nulo que depende de las características del canal.
3. Existen códigos que permiten alcanzar la capacidad del canal $R = C$.

Es importante hacer aquí una aclaración sobre el tercer punto. El teorema, aunque demuestra que es posible alcanzar dicha capacidad, no responde a la pregunta de cómo se pueden obtener en la práctica dichos códigos. En un problema práctico no se va a alcanzar dicha capacidad, sino que en general se utilizarán códigos que estarán por debajo de la misma.

En la siguiente sección se estudiará cómo calcular la capacidad de canal en primer lugar para un canal digital, y a continuación para el canal gaussiano. El diseño y análisis de códigos de canal prácticos no entra dentro del ámbito de esta asignatura, sino que se estudiarán en la asignatura “*Comunicaciones Digitales*”.

5.4.2. Capacidad de canal para el canal digital

En primer lugar se va a estudiar el caso del canal digital binario, en el que se agrupan n bits para formar los nuevos símbolos extendidos que se agruparán a la entrada y salida del sistema para implementar la codificación de canal. Para este caso se va a obtener la máxima cantidad de información que se podrá transmitir de forma fiable a través del canal de dos formas: mediante una explicación intuitiva basada en la definición de transiciones de alta probabilidad en la línea de la explicación utilizada anteriormente para explicar el principio de la codificación de canal, tal y como se presenta en [Proakis y Salehi, 2002]; y mediante la definición presentada por Shannon en el teorema de codificación de canal.

Aplicando la ley de los grandes números, para valores de n suficientemente grandes, cuando por una canal binario con probabilidad de error de bit ε se transmite una secuencia de n bits, la salida tendrá con gran probabilidad $n \times \varepsilon$ bits erróneos; es decir, la secuencia recibida tendrá con alta probabilidad $n \times \varepsilon$ bits distintos con respecto a la secuencia transmitida. El número de posibles secuencias de n bits que se diferencian en $n \cdot \varepsilon$ bits viene dado por el número combinatorio

$$\binom{n}{n \cdot \varepsilon}.$$

Teniendo en cuenta que un número combinatorio se puede obtener como

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

y utilizando la aproximación de Stirling para los números factoriales, que viene dada por

$$n! \approx n^n e^{-n} \sqrt{2\pi n}$$

ese número combinatorio puede aproximarse como

$$\binom{n}{n \cdot \varepsilon} \approx 2^{n H_b(\varepsilon)},$$

donde $H_b(\varepsilon)$ es la función de entropía binaria con argumento ε . Esto significa que para cada posible secuencia de n bits transmitidos hay aproximadamente $2^{n H_b(\varepsilon)}$ secuencias altamente probables en la salida del sistema.

Por otro lado, el número total de secuencias de longitud n bits altamente probables a la salida, depende de la incertidumbre de cada uno de los bits que forman la secuencia, medida a través de su entropía, que a la vez dependerá de la probabilidad de tener un uno o un cero en la salida. Si ambos símbolos son equiprobables, todas las secuencias posibles tendrán la misma probabilidad, y el número de secuencias altamente probables será 2^n . Pero si los bits en la salida no son igualmente probables, la probabilidad de las secuencias de salida será diferente para cada secuencia. Por ejemplo, si el bit “0” es menos probable que el bit “1”, las secuencias con muchos ceros serán menos probables. En ese caso, el número de secuencias de bits altamente probables se puede aproximar a partir de la medida de entropía de la salida, $H(Y)$, que será $H(Y) = H_b(p_Y(0))$, utilizando la expresión

$$2^{n H(Y)}.$$

Por tanto, el número máximo de secuencias de n bits en la entrada sin solapamiento entre las salidas altamente probables que generan en la salida, y que se denotará como M_{ss} , será el cociente entre el número de secuencias altamente probables en la salida y el número de secuencias que con alta probabilidad se generan en la salida cuando se transmite una cierta secuencia en la entrada, es decir

$$M_{ss} = \frac{2^{nH(Y)}}{2^{nH_b(\varepsilon)}} = 2^{n(H(Y)-H_b(\varepsilon))}.$$

El número de bits de información que se puede asociar a esas M_{ss} secuencias sin solapamiento es

$$\log_2 M_{ss} = n(H(Y) - H_b(\varepsilon)) \text{ bits de información.}$$

Por tanto la tasa del código resultante será el cociente entre los bits de información y el número de usos del canal

$$R = \frac{\log_2 M_{ss}}{n} = H(Y) - H_b(\varepsilon).$$

El máximo valor posible de esta tasa R es el que define la denominada *capacidad de canal*, que se denota como C . La entropía de los bits a la salida del canal binario será máxima cuando los bits “0” y “1” son igualmente probables, en cuyo caso $H(Y) = 1$. Por tanto, de forma intuitiva se ha llegado a que la capacidad del canal para un canal digital binario simétrico con probabilidad de error de bit ε es

$$C = 1 - H_b(\varepsilon) \text{ bits por uso del canal.}$$

A este resultado se ha llegado de forma intuitiva para el caso del canal BSC. A continuación se obtendrá el mismo resultado a partir de la teoría de la información. Para ello se calcula la información mutua entre la entrada y la salida del canal y a través de ella se intentará averiguar qué parte de la información se transmite y qué parte se pierde a su paso por el canal.

Para calcular la información entre la entrada, X , y la salida, Y , del canal DMC es necesario conocer las distribuciones de ambas variables. Conociendo la de la entrada, $p_X(x_i)$, como se conocen las probabilidades de transición, se conoce la distribución conjunta de la entrada y la salida

$$p_{X,Y}(x_i, y_j) = p_{Y|X}(y_j|x_i) \cdot p_X(x_i).$$

A partir de esta, la distribución de la salida se obtiene como

$$p_Y(y_j) = \sum_{i=0}^{M_X-1} p_{X,Y}(x_i, y_j) = \sum_{i=0}^{M_X-1} p_{Y|X}(y_j|x_i) \cdot p_X(x_i).$$

De este modo se tiene la caracterización completa entrada/salida. Tenga en cuenta que para un BSC

$$p_{Y|X}(y_j|x_i) = \begin{cases} 1 - \varepsilon & \text{si } j = i \\ \varepsilon & \text{si } j \neq i \end{cases}$$

A partir de estas distribuciones se puede calcular la información mutua entre entrada y salida $I(X, Y)$, por ejemplo a través de las relaciones con las distintas entropías, mediante

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X).$$

La información mutua entre la entrada y la salida representa la información que aporta la salida del canal sobre la entrada, o la incertidumbre que se elimina sobre el valor de la entrada cuando se conoce la salida: en definitiva, la información que es transmitida por el canal.

Para aclarar esta idea se utilizará $I(X, Y) = H(X) - H(X|Y)$ y se analizarán dos casos extremos del DMC más sencillo, el BSC. En este caso, el mejor canal posible es aquel libre de errores, es decir, $\varepsilon = 0$, o alternativamente $\varepsilon = 1$. Tenga en cuenta que en un sistema binario equivocarse siempre es una forma de acertar siempre, sólo hay que cambiar la decisión para ello. El peor caso posible es el caso en el que $\varepsilon = 1/2$ (cualquier ε mayor que $1/2$ se puede asimilar, cambiando la decisión, a un caso $1 - \varepsilon$).

Para el caso $\varepsilon = 0$, la distribución conjunta entrada salida es

$p_{X,Y}(x_i, y_j)$	x_0	x_1	$p_Y(y_j)$
y_0	$p_X(x_0)$	0	$P_X(x_0)$
y_1	0	$p_X(x_1)$	$P_X(x_1)$

o teniendo en cuenta que no se producen errores, $p_Y(y_i) = p_X(x_i)$ o simplemente se puede establecer la igualdad $Y = X$. Por tanto,

$$I(X, Y) = I(X, X) = H(X),$$

lo que significa que $H(X|Y) = 0$.

Por otro lado, la distribución conjunta cuando $\varepsilon = 1/2$ es

$p_{X,Y}(x_i, y_j)$	x_0	x_1	$p_Y(y_j)$
y_0	$p_X(x_0)/2$	$p_X(x_1)/2$	$1/2$
y_1	$p_X(x_0)/2$	$p_X(x_1)/2$	$1/2$

En este caso, Y presenta una distribución equiprobable independientemente de la distribución de X , como era de esperar. Esto quiere decir que X e Y son estadísticamente independientes, y la probabilidad conjunta se puede poner como

$$p_{X,Y}(x_i, y_j) = p_X(x_i) \cdot p_Y(y_j).$$

Como ya se ha deducido con anterioridad, si X e Y son independientes, su información mutua es nula,

$$I(X, Y) = 0,$$

lo que significa que $H(X|Y) = H(X)$.

De estos dos casos se pueden extraer las siguientes conclusiones:

1. La información mutua entre entrada y salida del canal es la cantidad de información que pasa de la entrada a la salida cuando el canal es utilizado. En el caso en que la probabilidad de error es nula, pasa toda la información ($I(X, Y) = H(X)$), y en el caso en que la entrada y la salida son estadísticamente independientes se “pierde” toda la información ($I(X, Y) = 0$).
2. $H(X|Y)$ puede interpretarse como la información que se “pierde” en el canal, y así la información que “atravesará” el canal, $I(X, Y)$, es igual a la información que hay a la entrada, $H(X)$, menos la que se pierde, $H(X|Y)$. Cuando la probabilidad de error es nula la pérdida es nula, y cuando la entrada y la salida son estadísticamente independientes, la pérdida es total, es decir, igual a la información a la entrada del canal.

Estas conclusiones se pueden extender a cualquier DMC con alfabetos de entrada y salida de M_X y M_Y símbolos, respectivamente.

Ahora bien, la información mutua entre la entrada y la salida del canal depende de la distribución de probabilidades a la entrada. Si deseamos saber cuál es la máxima cantidad de información capaz de atravesar un determinado canal, es necesario considerar todas las posibles distribuciones y la que produzca la mayor información mutua será la óptima, y la información mutua para esa distribución será la máxima información capaz de atravesar el canal, es decir, la *capacidad de canal*.

Formalmente, se define la capacidad de canal, C , de un canal discreto sin memoria, como

$$C = \max_{p_X(x_i)} I(X, Y).$$

Sus unidades son bits (o bits por uso de canal).

Algunas propiedades de la capacidad de canal, que derivan de su definición a través de la información mutua, son las que definen las cotas para su valor mínimo y máximo:

1. $C \geq 0$, ya que $I(X, Y) \geq 0$.
2. $C \leq \log M_X$, ya que $C = \max I(X, Y) \leq \max H(X) = \log M_X$.
3. $C \leq \log M_Y$, por la misma razón.

Para algunos casos sencillos, como el canal binario simétrico, es posible calcular de forma directa la capacidad del canal. Para el BSC, la información mutua es

$$\begin{aligned} I(X, Y) &= H(Y) - H(Y|X) \\ &= H(Y) - \sum_{i=0}^1 p_X(x_i) H(Y|X = x_i) \\ &= H(Y) - \sum_{i=0}^1 p_X(x_i) \left(- \sum_{j=0}^1 p_{Y|X}(y_j|x_i) \log p_{Y|X}(y_j|x_i) \right) \\ &= H(Y) - \sum_{i=0}^1 p_X(x_i) (-\varepsilon \log(\varepsilon) - (1 - \varepsilon) \log(1 - \varepsilon)) \\ &= H(Y) - \sum_{i=0}^1 p_X(x_i) H_b(\varepsilon) \\ &= H(Y) - H_b(\varepsilon). \end{aligned}$$

Se comprueba que se llega al mismo resultado obtenido con anterioridad. El máximo de esta información mutua se obtiene, dado que el parámetro ε es fijo y no se puede actuar sobre el, cuando la entropía $H(Y)$ es máxima. Esto se produce para una distribución de salida equiprobable, que para el BSC equivale a tener una distribución de entrada equiprobable. En este caso, $H(Y) = 1$ y la capacidad de canal vale por tanto

$$C = 1 - H_b(\varepsilon).$$

Es fácil representar esta dependencia con la probabilidad de error teniendo en cuenta la forma de variación de $H_b(\varepsilon)$. En un canal con probabilidad de error nula podemos transmitir un bit por uso del canal, mientras que en un canal con probabilidad de error $1/2$ no se puede enviar ninguna información.

Hay que tener en cuenta que el problema del cálculo de la capacidad de canal en general se puede plantear como un problema de maximización con restricciones. Maximización de la información mutua, que depende de las probabilidades de entrada, con las restricciones impuestas por dichas probabilidades, ya que deben siempre cumplirse las siguientes $2M_X + 1$ restricciones:

- $p_X(x_i) \geq 0$, para $i = 0, 1, \dots, M_X - 1$.
- $p_X(x_i) \leq 1$, para $i = 0, 1, \dots, M_X - 1$.
- $\sum_{i=0}^{M_X-1} p_X(x_i) = 1$.

En algunos canales simples será posible calcular la capacidad de canal de forma analítica. Para otro tipo de canales más complicados, a veces no es posible encontrar de forma sencilla la solución de forma analítica. En este caso se recurre a técnicas numéricas, que barren todas las posibles distribuciones de entrada hasta encontrar la que hace máxima la información mutua.

Ejemplo

Se pretende encontrar la capacidad del canal mostrado en la Figura 5.26.

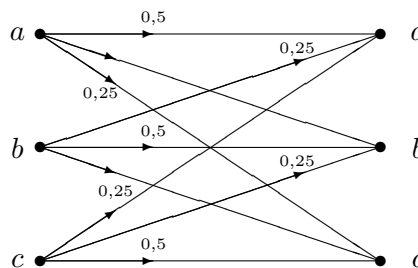


Figura 5.26: Un ejemplo de canal DMC

Para obtener la capacidad, lo primero es obtener una expresión de la información mutua. Una opción es a través de $H(Y)$ y $H(Y|X)$. De la relación entrada salida se observa que para los tres casos, dado $X = a$, $X = b$ o $X = c$, la variable aleatoria de salida Y , tiene tres símbolos con probabilidades $1/2$, $1/4$ y $1/4$ respectivamente. Por tanto,

$$H(Y|X = a) = H(Y|X = b) = H(Y|X = c) = \frac{1}{2} \log_2(2) + 2 \times \frac{1}{4} \log_2(4) = 1,5.$$

Como para todos los símbolos se obtiene lo mismo, independientemente de $P_X(x_i)$ se tiene que

$$H(Y|X) = 1,5.$$

Por tanto

$$I(Y, X) = H(Y) - 1,5$$

Para maximizar $I(X, Y)$, hay que maximizar $H(Y)$, y esto supone encontrar la distribución de probabilidad de la variable aleatoria de entrada que produce salidas igualmente probables. En este caso esto sucede para símbolos de entrada igualmente probables. En este caso

$$H(Y) = \log_2(3) = 1,585.$$

Por tanto la capacidad del canal es

$$C = 1,585 - 1,5 = 0,085 \text{ bits/uso del canal.}$$

Se observa que es relativamente baja, pero es normal si se tiene en cuenta que transmitiendo símbolos a través de este canal se falla un 50 % de las veces.

5.4.3. Capacidad de canal para el canal gaussiano

Una vez estudiado el canal digital y cómo se obtiene su capacidad, en esta sección se extiende el estudio para el caso del canal gaussiano, donde la salida del mismo en un instante dado es una variable aleatoria continua en lugar de una variable aleatoria discreta. Se tiene en cuenta en este análisis que en un sistema de comunicaciones real existen habitualmente dos importantes limitaciones en cuanto al uso de los recursos:

- Una limitación en la potencia máxima que puede transmitirse, que se aplica sobre la señal de comunicaciones que genera el transmisor.
- Una limitación en el ancho de banda de la señal transmitida.

Por eso en esta sección se estudiará cuál es la máxima cantidad de información que puede transmitirse de forma fiable en un sistema de comunicaciones digitales a través de un canal gaussiano cuando la potencia máxima de la señal transmitida está limitada a P_X wattios y el ancho de banda disponible son B Hz. Al igual que en la sección anterior, el análisis se realizará de dos formas. Por un lado se hará una demostración más intuitiva en términos del número de símbolos que pueden transmitirse con baja probabilidad de solapamiento de sus salidas, en la línea del tratamiento seguido en [Proakis y Salehi, 2002], y por otro se obtendrá también la capacidad de canal mediante un desarrollo basado en la teoría de la información, siguiendo la línea del análisis realizado en [Artés-Rodríguez et al., 2007].

Capacidad del canal gaussiano: demostración intuitiva

En un canal gaussiano la salida del canal se modela mediante una variable aleatoria Y que se relaciona con la entrada, modelada mediante una variable aleatoria X , a través de la relación aditiva

$$Y = X + Z,$$

donde Z es una variable aleatoria que modela el efecto del ruido, con una distribución gaussiana, de media nula y con varianza (potencia de ruido) P_Z .

La idea a seguir para obtener la capacidad del canal es similar a la seguida en el canal digital. Se va a muestrear las señales transmitida y recibida en n instantes de tiempo con el objetivo de ver cuál es el máximo número de posibles valores de la señal transmitida que da lugar en la salida a valores con baja probabilidad de solapamiento en el límite cuando n tiende a infinito. La restricción de potencia en la señal transmitida implica que si se tienen n realizaciones de la variable aleatoria X , que se agrupan en un vector \mathbf{x}

$$\mathbf{x} = \{x_1, x_2, \dots, x_n\},$$

lo que modelaría el valor de la señal transmitida en n instantes, para un valor de n suficientemente grande, se cumple

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \leq P_X.$$

En este caso x_i no denota el alfabeto de una variable aleatoria discreta, sino las n realizaciones de la variable aleatoria continua X que modela la amplitud de la señal transmitida, $s(t)$, en n instantes de tiempo.

Para bloques de longitud n de valores a la entrada, salida y de ruido, agrupados en los vectores \mathbf{x} , \mathbf{y} y \mathbf{z} , respectivamente, se puede escribir la relación vectorial

$$\mathbf{y} = \mathbf{x} + \mathbf{z}.$$

Si n es suficientemente grande, por la ley de los grandes números, la limitación en potencia del término de ruido implica la restricción

$$\frac{1}{n} \sum_{i=1}^n z_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2 \leq P_Z.$$

Finalmente, dada la independencia entre X y Z , la potencia de Y será la suma de las potencias de X y Z , es decir $P_Y = P_X + P_Z$, de tal modo que sobre las n muestras de la salida se tendrá la restricción

$$\frac{1}{n} \sum_{i=1}^n y_i^2 \leq P_Y = P_X + P_Z.$$

Teniendo en cuenta que la suma de coordenadas al cuadrado de un vector proporciona su norma al cuadrado, las restricciones que imponen la potencia de la señal transmitida y la potencia de ruido se pueden escribir de la siguiente forma

$$\|\mathbf{x}\|^2 \leq nP_X, \quad \|\mathbf{y} - \mathbf{x}\|^2 \leq nP_Z, \quad \|\mathbf{y}\|^2 \leq n(P_X + P_Z).$$

La interpretación geométrica de las restricciones que imponen la potencia de la señal transmitida y la potencia de la señal de ruido, conduce a las siguientes conclusiones cuando n se incrementa de forma asintótica

- La representación vectorial de las n muestras de la señal transmitida, \mathbf{x} , se encuentra localizada en una hiper-esfera n -dimensional de radio $\sqrt{nP_X}$ centrada en el origen.
- La representación vectorial de las n muestras de la señal de salida, \mathbf{y} , se encuentra en una hiper-esfera n -dimensional de radio $\sqrt{nP_Z}$ y centrada en torno a la representación vectorial de las n muestras de la señal transmitida, \mathbf{x} .

- La representación vectorial de las n muestras de la señal recibida, \mathbf{y} , se encuentra localizada en una hiper-esfera n -dimensional de radio $\sqrt{n(P_X + P_Z)}$ centrada en el origen.

Haciendo uso de esta interpretación geométrica, el cálculo de la capacidad del canal equivale a encontrar cuantas secuencias distintas de n muestras de la señal transmitida \mathbf{x} se pueden obtener de tal forma que las salidas a que dan lugar no se solapen sobre el espacio de salida. Obviamente, si se cumple esta condición, entonces las secuencias de salida se pueden decodificar de forma fiable. Por tanto, la pregunta a la que hay que responder para establecer el valor de la capacidad del canal es: *¿Cuántas esferas de radio $\sqrt{nP_Z}$ se pueden empaquetar en una esfera de radio $\sqrt{n(P_X + P_Z)}$?* La Figura 5.27 ilustra este planteamiento del problema para el cálculo de la capacidad de canal.

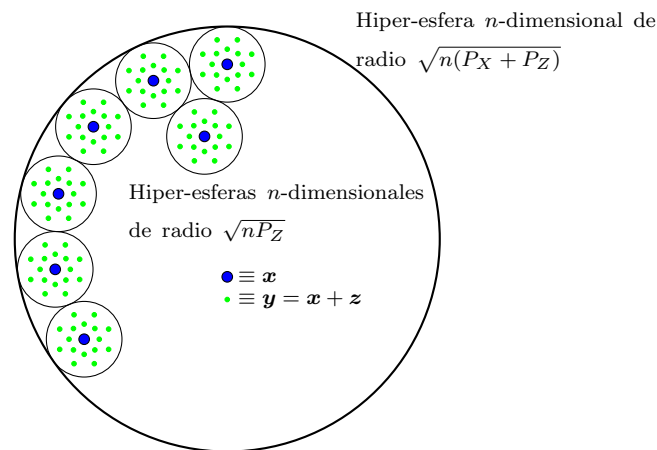


Figura 5.27: Interpretación geométrica del rango de las n muestras en la salida cuando se transmite un cierto valor para las n muestras de la entrada para el cálculo de la capacidad de un canal gaussiano.

La respuesta, de forma aproximada, se puede obtener mediante la relación entre los volúmenes de las dos hiper-esferas. El volumen de una hiper-esfera de dimensión n y radio r es proporcional a su radio elevado a la potencia n ,

$$V_n = K_n r^n,$$

donde K_n es una constante independiente del radio que depende de la dimensión del espacio. En este caso, el número máximo de símbolos, entendidos como bloques de n muestras de la señal transmitida, que pueden ser transmitidos con fidelidad a través de un canal gaussiano se aproxima mediante el cociente entre el volumen de la esfera en que está contenida la representación de las n muestras de la salida del canal, que tiene radio $\sqrt{n(P_X + P_Z)}$, y el volumen de la esfera en que está contenida la representación de la salida cuando se ha transmitido un cierto símbolo \mathbf{x} , que tiene radio $\sqrt{nP_Z}$. Este cociente es por tanto

$$M_{ss} = \frac{K_n (n(P_X + P_Z))^{n/2}}{K_n (nP_Z)^{n/2}} = \left(\frac{P_X + P_Z}{P_Z} \right)^{n/2} = \left(1 + \frac{P_X}{P_Z} \right)^{n/2}.$$

Como con M_{ss} secuencias sin solapamiento se pueden codificar $\log_2 M_{ss}$ bits de información, la capacidad del canal gaussiano con restricción de potencia P_X y potencia de ruido P_Z viene dada

por el cociente entre este número de bits y el número de usos del canal, que será n , es decir

$$\begin{aligned} C &= \frac{\log_2 M_{ss}}{n} = \frac{1}{n} \frac{n}{2} \log_2 \left(1 + \frac{P_X}{P_Z} \right) \\ &= \frac{1}{2} \log_2 \left(1 + \frac{P_X}{P_Z} \right). \end{aligned}$$

Ahora hay que recordar que al trabajar con una limitación en el ancho de banda, en el receptor se filtra la señal recibida con un filtro ideal de ancho de banda B Hz para minimizar el efecto del ruido, con lo que la potencia del ruido filtrado es

$$P_Z = N_0 B.$$

Por tanto la capacidad del canal gaussiano es

$$C = \frac{1}{2} \log_2 \left(1 + \frac{P_X}{N_0 B} \right) \text{ bits/uso.}$$

Si se multiplica este resultado por el número de usos (transmisiones) por segundo, que cumpliendo el teorema de Nyquist son $2B$, se obtiene la capacidad del canal en bits/s,

$$C = B \log_2 \left(1 + \frac{P}{N_0 B} \right) \text{ bits/s.}$$

Esta es la conocida fórmula de Shannon para la capacidad de un canal con ruido aditivo blanco y gaussiano.

Capacidad del canal gaussiano a través de la información mutua

A este mismo resultado se puede llegar también a través de una derivación más formal basada en la teoría de la información [Artés-Rodríguez et al., 2007]. En el modelo de canal gaussiano la relación entre la entrada y la salida viene dada por

$$Y = X + Z,$$

donde Z es una variable aleatoria gaussiana, de media nula y varianza $P_Z = N_0 B$. De este modo la distribución condicional de Y dado X , $f_{Y|X}(y|x)$, es una distribución gaussiana de media x y varianza $\sigma^2 = P_Z$. La capacidad del canal se obtendrá a través de la información mutua de la misma forma que para el canal digital, maximizando la información mutua entre la entrada y la salida del canal, pero incluyendo la restricción en la potencia máxima para la señal transmitida (sin esta restricción, la capacidad sería teóricamente infinita). Es decir, se define la capacidad de canal como

$$C = \max_{f_X(x) | E[X^2] \leq P_X} I(X, Y),$$

donde la restricción en potencia viene dada por la limitación $E[X^2] \leq P_X$.

La información mutua entre entrada y salida se puede calcular a través de su relación con las entropías diferenciales

$$I(X, Y) = h(Y) - h(Y|X) = h(X + Z) - h(Z).$$

La entropía diferencial del ruido se calcula de forma muy sencilla, ya que su distribución es gaussiana de media nula y varianza $\sigma^2 = P_Z$, caso que ya se ha considerado con anterioridad, por lo que

$$h(Z) = \int_{-\infty}^{\infty} f_Z(z) \log_2 \frac{1}{f_Z(z)} dz = \frac{1}{2} \log_2 2\pi e P_Z.$$

Esta entropía diferencial depende únicamente de la varianza del término de ruido, $P_Z = N_0 B$. Por tanto la información mutua se puede escribir como

$$I(X, Y) = h(X + Z) - \frac{1}{2} \log_2 2\pi e P_Z.$$

Para obtener su máximo se hará uso de la siguiente propiedad sobre la entropía diferencial:

- Si una variable aleatoria tiene un valor fijo de varianza, la función densidad de probabilidad que hace máxima su entropía diferencial es la distribución gaussiana.

La demostración de esta propiedad se puede ver, por ejemplo, en [Artés-Rodríguez et al., 2007], Capítulo 9, página 565.

Dado que X y Z son estadísticamente independientes y Z tiene media nula, la varianza de Y es la suma de las varianzas de X e Y , por lo que

$$E[Y^2] = E[(X + Z)^2] = E[X^2] + E[Z^2] \leq P_X + P_Z.$$

Como la varianza de Y está acotada, la capacidad se alcanza cuando Y tiene una distribución gaussiana, en cuyo caso el valor máximo de la información mutua, y por tanto la capacidad, es

$$C = \frac{1}{2} \log_2 2\pi e (P_X + P_Z) - \frac{1}{2} \log_2 2\pi e \sigma^2 = \frac{1}{2} \log_2 \left(1 + \frac{P_X}{P_Z} \right).$$

Este resultado es el mismo que el obtenido anteriormente.

5.5. Límites en un sistema digital de comunicaciones

En la sección anterior se ha obtenido la capacidad de canal para el canal gaussiano

$$C = B \cdot \log_2 \left(1 + \frac{P_X}{N_0 B} \right) \text{ bits/s.}$$

Además de por supuesto del valor de N_0 , el valor de la capacidad de canal depende de dos parámetros relevantes de todo sistema de comunicaciones: la potencia de la señal transmitida, P_X , y el ancho de banda disponible, B Hz. En esta sección en primer lugar se analizará la dependencia de la capacidad de canal con estos dos parámetros del sistema de comunicaciones.

En cuanto a la potencia de la señal transmitida, el análisis es sencillo. Si se incrementa la potencia transmitida, se incrementa la capacidad de canal: obviamente, cuanto mayor es la potencia, más niveles (o símbolos de longitud n muestras) se pueden poner, y más bits/uso son posibles. Sin embargo hay que tener en cuenta que el incremento sigue una ley logarítmica, de modo que para poder obtener un incremento lineal en la capacidad, es preciso un incremento exponencial en la potencia transmitida. En cualquier caso, teóricamente es posible incrementar la capacidad hasta el infinito incrementando la potencia de la señal transmitida.

El efecto del ancho de banda en la capacidad del canal es diferente. Incrementar B tiene dos efectos contrapuestos. Por un lado, un mayor ancho de banda incrementa la velocidad de transmisión, pero por otro lado incrementa el nivel del ruido y por tanto reduce las prestaciones. Haciendo llegar B al límite infinito, y aplicando la regla de L'Hopital, se obtiene el límite de la capacidad de canal cuando el ancho de banda tiende a infinito

$$\lim_{B \rightarrow \infty} C = \frac{P_X}{N_0} \log_2(e) = 1,44 \frac{P_X}{N_0}.$$

Este resultado significa que, al contrario de lo que sucede con la potencia de la señal transmitida, el incremento del ancho de banda del canal por sí sólo no puede aumentar la capacidad hasta cualquier valor deseado, sino que hay un límite máximo alcanzable que depende de la relación señal a ruido (P_X/N_0).

El teorema de codificación de canal de Shannon plantea un límite máximo en la tasa de transmisión de información con un sistema de comunicaciones digitales. Esto quiere decir que en un sistema de comunicaciones práctico, siempre se debe cumplir que la tasa de transmisión efectiva (definida sobre los bits de información) esté por debajo de la capacidad de canal, $R < C$. Para el caso del canal gaussiano esto implica que la velocidad de transmisión binaria, R_b bits/s, tiene que estar por debajo de la capacidad del canal. Definiendo la relación señal a ruido como la relación entre la potencia de la señal transmitida y N_0

$$SNR = \frac{P_X}{P_Z} = \frac{P_X}{N_0 B},$$

esto significa que se debe cumplir la relación

$$R_b < B \log \left(1 + \frac{P}{N_0 B} \right) \text{ bits/s.}$$

Dividiendo los dos lados de esta desigualdad por B , definiendo la tasa binaria espectral, o eficiencia espectral, como

$$\eta = \frac{R_b}{B} \text{ bits/s/Hz,}$$

se obtiene el siguiente límite para la eficiencia espectral en un sistema de comunicaciones práctico

$$\eta < \log_2 \left(1 + \frac{P_X}{N_0 B} \right).$$

Si ahora se define la energía media por bit como el cociente entre la potencia de la señal transmitida y la tasa de transmisión binaria

$$E_b = \frac{P_X}{R_b},$$

y la correspondiente relación entre E_b y N_0 ,

$$\frac{E_b}{N_0} = \frac{SNR}{\eta},$$

se puede reescribir la limitación de la tasa binaria espectral o eficiencia espectral en términos de esta relación como

$$\eta < \log_2 \left(1 + \eta \frac{E_b}{N_0} \right).$$

Esta restricción que relaciona la relación señal a ruido y la eficiencia espectral del sistema se escribe en la literatura con varias expresiones alternativas equivalentes, como por ejemplo

$$\frac{E_b}{N_0} > \frac{2^\eta - 1}{\eta}.$$

Esta expresión indica que hay un valor mínimo de relación E_b/N_0 que es necesario para poder tener una comunicación fiable. En el caso de esta expresión, si se calcula su límite cuando η tiende a infinito se tiene

$$\lim_{\eta \rightarrow \infty} \frac{E_b}{N_0} = \ln 2 = 0,693 \approx -1,6 \text{ dB.}$$

La curva que define el valor límite de la tasa binaria espectral como una función de la relación E_b/N_0 , que se representa en la Figura 5.28, divide el plano η vs E_b/N_0 en dos regiones. En una región, debajo de la curva, es posible una comunicación fiable (entendida como una comunicación en la que el uso de técnicas de codificación de canal permite reducir la probabilidad de error hasta niveles arbitrariamente bajos). En la otra, por encima de la curva, no es posible tener una comunicación fiable. Las prestaciones de cualquier sistema se pueden representar mediante un punto en el plano de esta curva. Cuanto más cerca está el punto de la curva, más cerca está el rendimiento de ser óptimo.

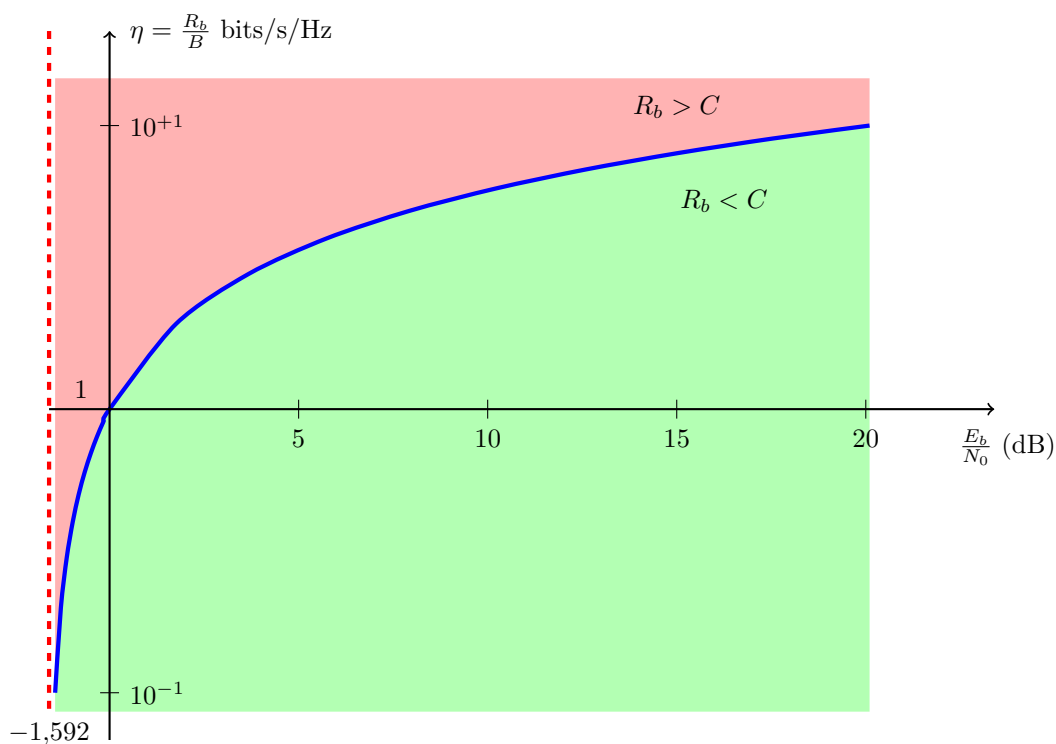


Figura 5.28: Valor límite de la tasa binaria espectral η (bits/s/Hz) frente a relación E_b/N_0 para un canal gaussiano. Se distingue la región donde es posible una comunicación fiable (en verde) de la región donde no es posible (en rojo).

Se puede ver que esta curva, cuando η tiende a cero, la relación E_b/N_0 tiende al valor anteriormente calculado

$$\frac{E_b}{N_0} = \ln 2 = 0,693 \approx -1,6 \text{ dB.}$$

Este es un mínimo absoluto para una comunicación fiable, es decir, que para que se pueda tener una comunicación fiable, la relación E_b/N_0 ha de estar por encima de ese límite.

En esta figura se pueden también apreciar dos cosas en cuanto al valor de η :

1. Cuando $\eta \ll 1$, el ancho de banda es grande y la única limitación aparece por la potencia. En este caso se habla de *sistemas limitados en potencia*. En este caso hay que utilizar constelaciones sencillas.

2. Cuando $\eta \gg 1$ el ancho de banda del canal es pequeño y se habla de *sistemas limitados en ancho de banda*. En este caso se utilizan constelaciones muy densas (p.e. 256-QAM).

La expresión anterior indica un límite mínimo que debe alcanzar la relación entre la energía de la señal y la del ruido para poder transmitir con una cierta eficiencia espectral de forma fiable. Aunque en la literatura ese límite se suele expresar a través de la relación E_b/N_0 , en otras ocasiones se expresa a partir de la relación señal a ruido, en cuyo caso la expresión resultante es

$$SNR > 2^n - 1.$$

Dado este límite mínimo de SNR, en ocasiones se define la denominada “*relación señal a ruido normalizada*”, que básicamente compara la relación señal a ruido de trabajo con ese mínimo nivel requerido

$$SNR_{norm} = \frac{SNR}{2^n - 1}.$$

Por la propia definición de esta relación señal a ruido normalizada, en un sistema práctico debe tomar valores mayores que la unidad, o lo que es lo mismo, mayores que 0 decibelios. Se trata por tanto de una medida relativa que indica lo cerca o lejos que se encuentra el sistema del valor límite de funcionamiento en la zona fiable.

5.6. Ejercicios

Ejercicio 5.1 Una variable aleatoria tiene un alfabeto $\mathcal{A}_X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ con probabilidades

$$p_X(x_1) = 0,1, p_X(x_2) = 0,2, p_X(x_3) = 0,3, p_X(x_4) = 0,05, p_X(x_5) = 0,15, p_X(x_6) = 0,2.$$

Calcule la entropía de la variable aleatoria y compare dicha entropía con la de una variable aleatoria con el mismo alfabeto pero uniformemente distribuida.

Ejercicio 5.2 X es una variable aleatoria binaria con alfabeto $\mathcal{A}_X = \{0, 1\}$ y con probabilidades $p_X(0) = p$ y $p_X(1) = 1 - p$. Y es una variable aleatoria binaria con el mismo alfabeto que depende de X como

$$p_{Y|X}(1|0) = p_{Y|X}(0|1) = \varepsilon.$$

- Calcule $H(X)$, $H(Y)$, $H(Y|X)$, $H(X, Y)$, $H(X|Y)$ e $I(X, Y)$.
- Para un valor ε fijo, ¿qué valor de p maximiza $I(X, Y)$?
- Para un valor p fijo, ¿qué valor de ε minimiza $I(X, Y)$?

Ejercicio 5.3 La entrada de un canal discreto sin memoria tiene un alfabeto de 7 símbolos $\mathcal{A}_X = \{x_0, x_1, x_2, x_3, x_4, x_5, x_6\}$ con probabilidades

$$p_X(x_0) = 0,05, p_X(x_1) = 0,1, p_X(x_2) = 0,1, p_X(x_3) = 0,15, p_X(x_4) = 0,05, p_X(x_5) = 0,25, p_X(x_6) = 0,3.$$

respectivamente. El canal discreto sin memoria tiene la siguiente matriz de canal

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

- Calcule la entropía de la entrada del canal, X , y compárela con la de una variable aleatoria con el mismo alfabeto pero símbolos equiprobables.
- Calcule la entropía de la variable aleatoria a la salida del canal, Y , y compárela con la de una variable aleatoria con el mismo alfabeto pero con símbolos equiprobables.
- Calcule la entropía conjunta $H(X, Y)$, las entropías condicionales, $H(X|Y)$ y $H(Y|X)$, y la información mutua $I(X, Y)$.

Ejercicio 5.4 El canal mostrado en la Figura 5.29 modela el denominado canal binario con borrado (BEC, del inglés *binary erasure channel*). Representa un modelo en el que se decide un símbolo cuando se tiene cierta certidumbre sobre el mismo, pero se marca como dudoso (borrado) cuando no hay suficiente certidumbre. Calcule la capacidad del canal y representéla como una función del parámetro ε que determina la probabilidad de borrado.

Ejercicio 5.5 Se tiene el canal discreto sin memoria representado en la Figura 5.30.

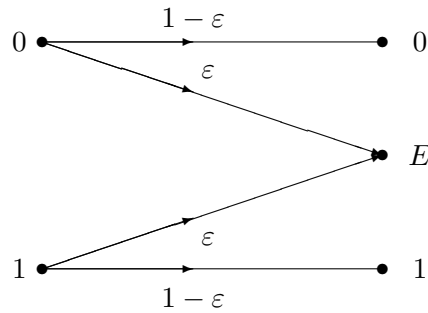


Figura 5.29: Representación del canal BEC del Ejercicio 5.4.

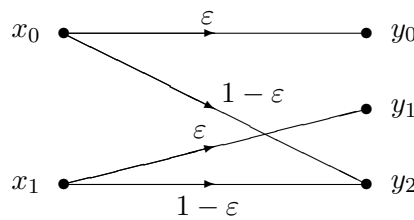


Figura 5.30: Canal para el Ejercicio 5.5.

- Calcule la capacidad del canal (valor y probabilidades de entrada para las que se alcanza).
- Suponiendo que el símbolo x_0 se transmite con un probabilidad p , calcule la entropía conjunta entrada salida, $H(X, Y)$, y la entropía condicional de la entrada conocida la salida, $H(X|Y)$ en función de los parámetros ε y p . Dibújelo, en función de p , para un valor $\varepsilon = 0,25$.

Ejercicio 5.6 Se tienen los canales que se muestran en la Figura 5.31.

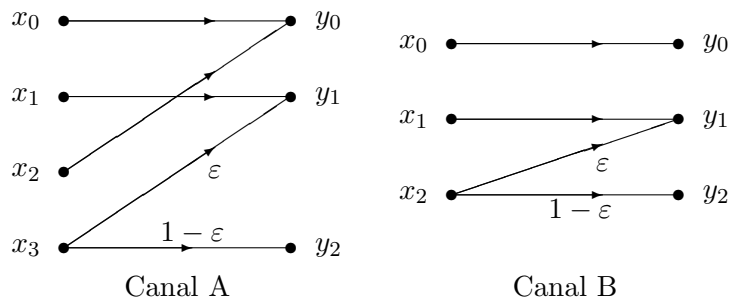


Figura 5.31: Canales para el Ejercicio 5.6.

- Calcule la capacidad del Canal A para $\varepsilon = 0$ y para $\varepsilon = 1$.
- Para el Canal A, si los símbolos de entrada son equiprobables, $p_X(x_i) = 1/4, i = 0, \dots, 3$, se conoce que la información mutua entre la entrada y la salida del canal es $I(X, Y) = 1,069$ y que la entropía de la salida del canal es $H(Y) = 1,272$. Calcule el valor del parámetro ε y la entropía condicional $H(X|Y)$.
- Para el Canal B, se sabe que para las probabilidades *a priori*, $p_X(x_i)$, que maximizan la información mutua entre la entrada y salida del canal, $I(X, Y)$, se tiene una entropía condicional $H(Y|X) = 0,207$ bits para $\varepsilon = 1/4$. Calcule la capacidad del canal y las probabilidades de los símbolos de entrada para las que se obtiene.

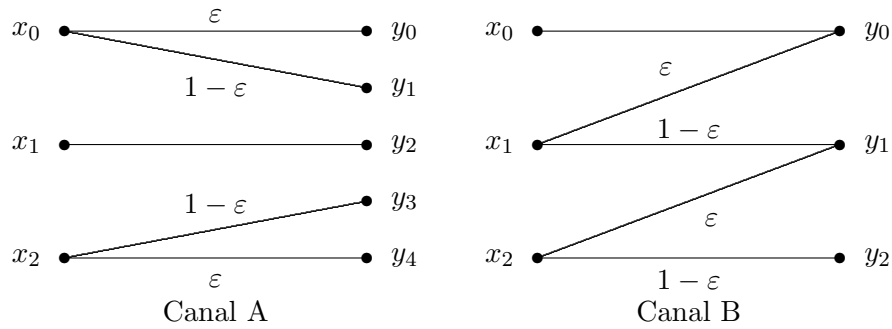


Figura 5.32: Canales para el Ejercicio 5.7.

Ejercicio 5.7 Se tienen los canales que se muestran en la Figura 5.32.

- Calcule la capacidad del Canal A (valor y probabilidades de los símbolos de entrada para la que se alcanza).
- Para el Canal B, si los símbolos de entrada son equiprobables, se sabe que la entropía conjunta vale $H(X, Y) = 2,1258$ bits/símbolo. Calcule el valor de ϵ , $H(Y)$, $I(X, Y)$, y $H(X|Y)$.

Ejercicio 5.8 Se tiene el canal que se muestra en la Figura 5.33.

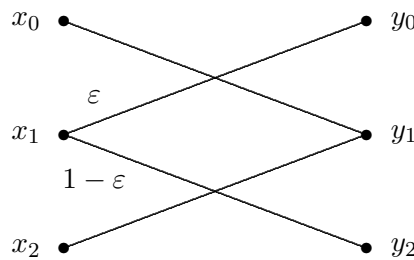


Figura 5.33: Canales para el Ejercicio 5.8.

- Calcule la capacidad del canal de la figura y representéla como una función de ϵ para $\epsilon \in [0, 1]$.
- Si se define $p = p_X(x_1)$ y se hace $p_X(x_2) = 0$, calcule las expresiones de la entropía condicional $H(X|Y)$ y de la conjunta, $H(X, Y)$, en función de p y ϵ , y representélas en función de p para $\epsilon = 1/2$.

Ejercicio 5.9 Se tiene el canal que se muestra en la Figura 5.34.

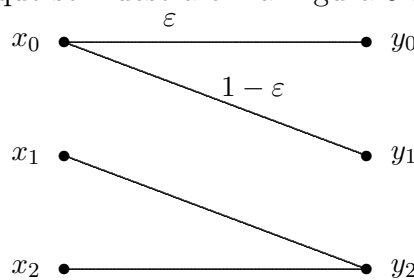


Figura 5.34: Canales para el Ejercicio 5.9.

- Calcule la capacidad del canal de la figura y representéla como una función de ϵ para $\epsilon \in [0, 1]$.

- b) Si se hace $p_X(x_1) = p_X(x_2) = p$, calcule las expresiones de la entropía condicional $H(X|Y)$ y de la conjunta, $H(X, Y)$, en función de p y ε , y represéntelas en función de p para $\varepsilon = 1/2$.

Ejercicio 5.10 El canal discreto sin memoria (DMC) que se muestra en la Figura 5.35 corresponde al modelo estadístico de canal de un sistema de comunicaciones que utiliza la constelación de cuatro símbolos de la misma figura.

- Escriba la matriz de canal asociada a la representación esquemática del DMC, compárela con la matriz de canal para dicha constelación, explique qué aproximación se ha considerado sobre un sistema de comunicaciones con dicha constelación, y obtenga el valor de ε para ese caso (suponga transmisión sobre un canal gaussiano de densidad espectral de potencia de ruido $N_0/2$).
- Calcule $H(Y|X)$, $H(X|Y)$, $H(X, Y)$ e $I(X, Y)$ para el canal de la figura si los símbolos de entrada son equiprobables, y a partir de las expresiones analíticas obtenga el valor de ε que hace mínimas $H(Y|X)$ y $H(X|Y)$, y máxima $I(X, Y)$.
- Calcule la capacidad de canal para los valores mínimo y máximo que puede tomar ε teniendo en cuenta la estructura del DMC en este caso.

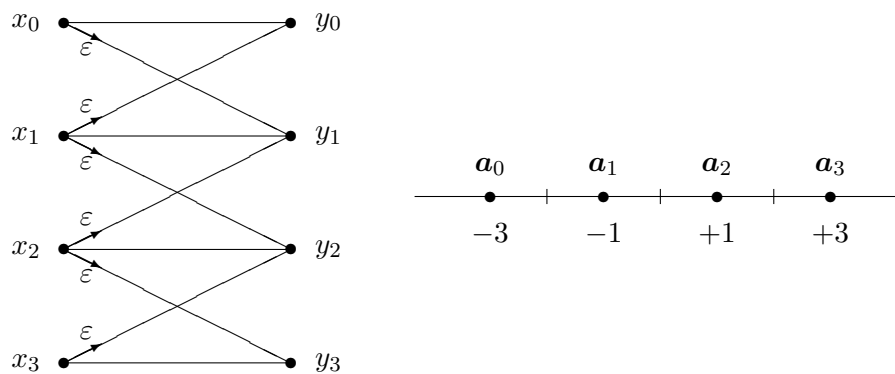


Figura 5.35: Canales para el Ejercicio 5.10.

Ejercicio 5.11 El canal discreto sin memoria (DMC) que se muestra en la Figura 5.36 corresponde al modelo estadístico de canal de un sistema de comunicaciones que utiliza la constelación de cuatro símbolos de la misma figura.

- Escriba la matriz de canal asociada a la representación esquemática del DMC, compárela con la matriz de canal para dicha constelación, explique qué aproximación se ha considerado sobre un sistema de comunicaciones con dicha constelación, y obtenga el valor de ε para ese caso (suponga transmisión sobre un canal gaussiano de densidad espectral de potencia de ruido $N_0/2$).
- Calcule $H(Y|X)$, $H(X|Y)$, $H(X, Y)$ e $I(X, Y)$ para el canal de la figura si los símbolos de entrada son equiprobables, represéntelas en función de ε , y obtenga el valor de ε que hace máximas $H(X, Y)$, $H(Y|X)$ y $H(X|Y)$, y mínima $I(X, Y)$.
- Calcule la capacidad del canal de la figura en función de ε .

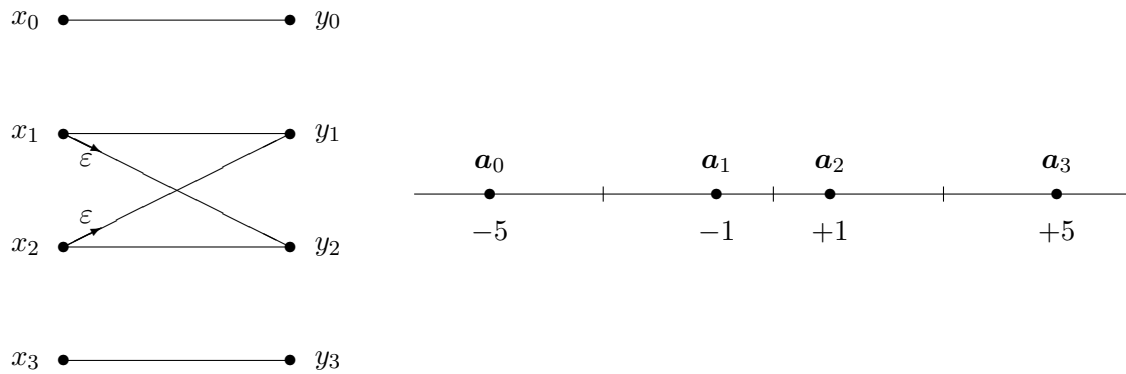


Figura 5.36: Canales para el Ejercicio 5.11.

Ejercicio 5.12 Un canal discreto sin memoria (DMC) con alfabeto de entrada $\mathcal{A}_X = \{x_0, x_1, x_2, x_3\}$ y alfabeto de salida $\mathcal{A}_Y = \{y_0, y_1, y_2, y_3\}$ viene dado por la siguiente matriz de canal

$$P = \begin{bmatrix} 1 - \varepsilon_0 & \varepsilon_0 & 0 & a \\ 0 & 1 - \varepsilon_1 & b & 0 \\ 0 & c & 1 - \varepsilon_1 & 0 \\ d & 0 & \varepsilon_0 & 1 - \varepsilon_0 \end{bmatrix}.$$

- a) Determine el valor de las constantes a, b, c y d , y calcule $H(Y), H(X, Y), H(Y|X), H(X|Y)$, e $I(X, Y)$ para el canal si los símbolos de entrada son equiprobables.
- b) Calcule la capacidad del canal de la figura en función de ε_0 y de ε_1 .

Ejercicio 5.13 Un canal discreto sin memoria (DMC) con alfabeto de entrada $\mathcal{A}_X = \{x_0, x_1, x_2\}$ y alfabeto de salida $\mathcal{A}_Y = \{y_0, y_1, y_2\}$ viene dado por la siguiente matriz de canal

$$P_{Y|X} = \begin{bmatrix} 1 - \varepsilon_0 & \varepsilon_0 & 0 \\ \varepsilon_1 & 1 - 2\varepsilon_1 & \varepsilon_1 \\ 0 & \varepsilon_0 & 1 - \varepsilon_0 \end{bmatrix}$$

- a) Calcule $H(Y), H(Y|X), H(X|Y)$, e $I(X, Y)$ para el canal si los símbolos de entrada son equiprobables.
- b) Calcule la capacidad del canal de la figura en función de ε_0 si se ha fijado $p_X(x_1)=0$.
- c) Calcule la capacidad del canal de la figura en función de ε_0 y de ε_1 e indique la distribución de entrada para la que se obtiene.
- d) Si $\varepsilon_0 = \varepsilon_1 = \varepsilon$ para qué valor de ε se alcanza la máxima capacidad. Explique por qué es esta capacidad la máxima alcanzable en un canal con 3 símbolos de entrada.

Ejercicio 5.14 Un canal discreto sin memoria (DMC) con alfabeto de entrada $\mathcal{A}_X = \{x_0, x_1, x_2, x_3\}$ y alfabeto de salida $\mathcal{A}_Y = \{y_0, y_1, y_2, y_3\}$ viene dado por la siguiente matriz de canal

$$P = \begin{bmatrix} 1 - \varepsilon_1 & \varepsilon_1 & a & b \\ \varepsilon_1 & 1 - \varepsilon_1 & c & d \\ a & c & 1 - \varepsilon_0 & \varepsilon_0 \\ b & d & \varepsilon_0 & 1 - \varepsilon_0 \end{bmatrix}.$$

- a) Determine el valor de las constantes a, b, c y d , y ponga un ejemplo de sistema digital de comunicaciones (teniendo en cuenta la constelación transmitida) para el que esta matriz de canal sea una buena aproximación.
- b) Para el caso en el que los símbolos x_0 y x_1 se transmiten con igual probabilidad, $p_X(x_0) = p_X(x_1) = p/2$, y los símbolos x_2 y x_3 también tienen la misma probabilidad, distinta en general de las de los otros dos, calcule $H(X), H(Y), H(X, Y), H(Y|X)$, y $H(X|Y)$, en función de ϵ_0, ϵ_1 y p .
- c) Calcule la capacidad del canal en función de ϵ_0 y de ϵ_1 e indique la distribución de entrada para la que se obtiene.

Ejercicio 5.15 Se definen dos sistemas de comunicaciones basados en dos canales discretos sin memoria (DMC). El sistema A tiene como alfabeto de entrada $\mathcal{A}_X = \{x_0, x_1\}$ y como alfabeto de salida $\mathcal{A}_Y = \{y_0, y_1\}$. El comportamiento de todo el sistema viene dado por la probabilidad conjunta de los símbolos de entrada y de salida $P_{X,Y}(x_i, y_j)$ donde $i, j \in \{0, 1\}$. Del mismo modo, el sistema B tiene como alfabeto de entrada $\mathcal{A}_Y = \{y_0, y_1\}$ y como alfabeto de salida $\mathcal{A}_Z = \{z_0, z_1\}$. El comportamiento del sistema viene dado por la probabilidad conjunta de los símbolos de entrada y de salida $P_{Y,Z}(y_i, z_j)$ donde $i, j \in \{0, 1\}$

$P_{X,Y}(x_i, y_j)$	x_0	x_1
y_0	α	$(1 - \alpha)\epsilon$
y_1	0	$(1 - \alpha)(1 - \epsilon)$

$P_{Y,Z}(y_i, z_j)$	y_0	y_1
z_0	$\beta(1 - \epsilon')$	0
z_1	$\beta\epsilon'$	$1 - \beta$

Si queremos estudiar las prestaciones de cada uno de los sistemas de forma independiente con el fin de compararlos:

- a) Obtenga $H(X), H(Y), H(X, Y), H(Z)$ y $H(Y|Z)$.
- b) Para cada uno de los canales obtenga respectivamente $P_{Y|X}(y_j|x_i)$ y $P_{Z|Y}(z_j|y_i)$ y dibuje el esquema del DMC asociado en cada uno de los casos. Identifique símbolos de entrada, símbolos de salida y probabilidades de transición.
- c) Calcule la información mutua para el sistema A.

Si se decide concatenar los dos canales obtenidos (A-B):

- d) Obtenga el DMC equivalente identificando las probabilidades de transición en cada caso.
- e) Determine los valores de ϵ y ϵ' que llevan al canal equivalente a comportarse como un canal binario simétrico (BSC).

Ejercicio 5.16 Se tiene el canal discreto sin memoria mostrado en la Figura 5.37. La distribución de probabilidades de los símbolos de entrada se denota como $p_X(x_i) \equiv p_i$ con $i \in \{0, 1, 2\}$.

- a) Escriba la matriz de canal correspondiente a este canal discreto equivalente.
- b) Calcule las entropías $H(X), H(Y), H(X, Y), H(Y|X)$ y $H(X|Y)$ y la información mutua $I(X, Y)$ para la siguiente distribución de entrada:

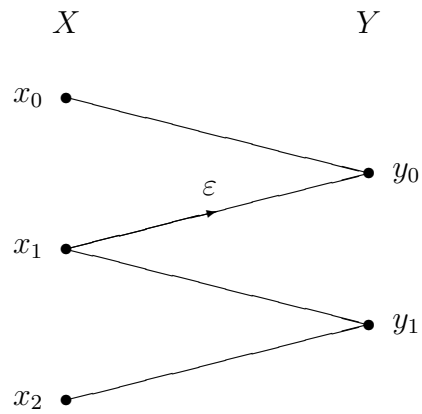


Figura 5.37: Representación en diagrama de rejilla para el DMC del Ejercicio 5.16.

- I) Símbolos de entrada equiprobables: $p_0 = p_1 = p_2 = \frac{1}{3}$.
 - II) Sólo se transmite el símbolo x_1 : $p_0 = p_2 = 0$, $p_1 = 1$.
- c) Obtenga la capacidad del canal, indicando la distribución de entrada para la que se obtiene, en los siguientes casos
- I) Para el caso $\varepsilon = \frac{1}{2}$.
 - II) Si se elimina el símbolo x_2 del alfabeto de X (o lo que es lo mismo, para $p_X(x_2) = 0$).
NOTA: En este apartado, ε no toma el valor del apartado anterior, es un valor constante arbitrario.

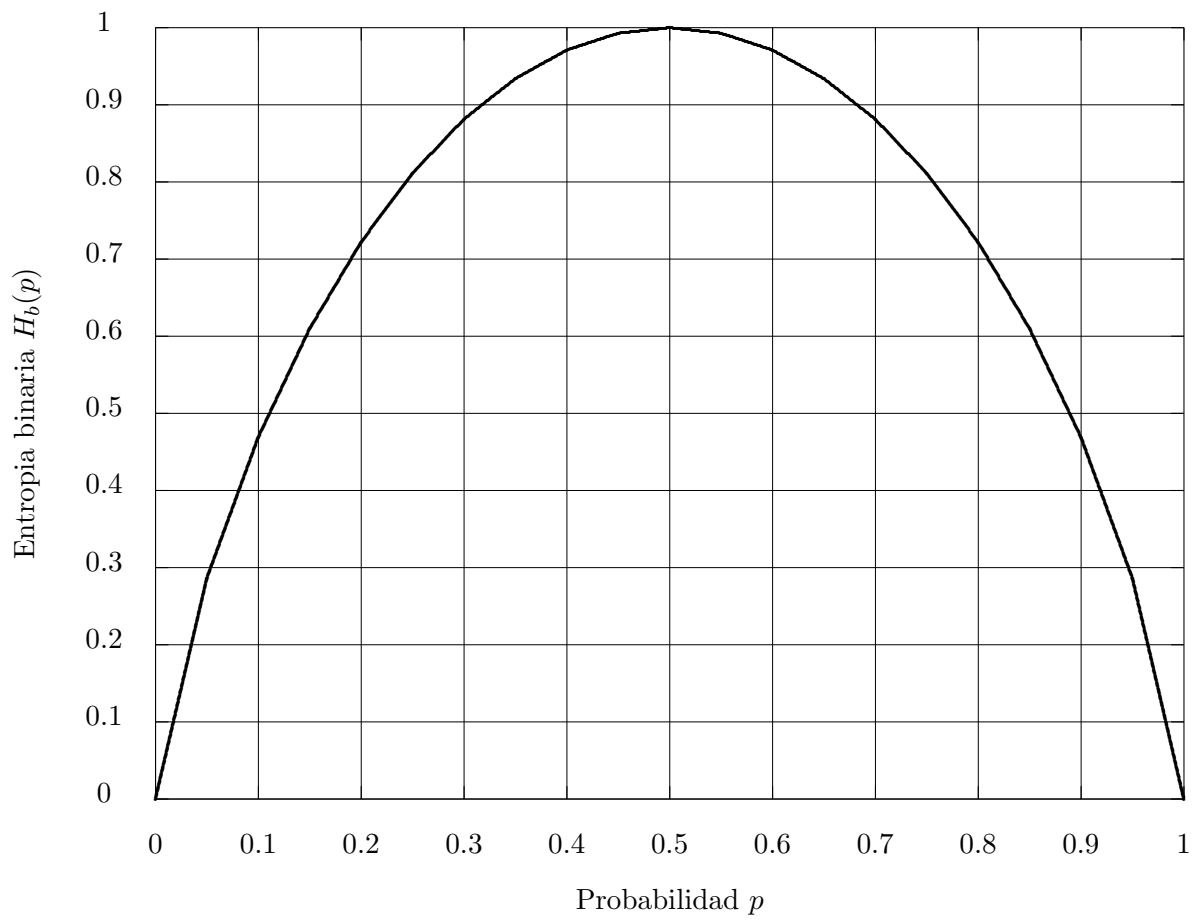


Figura 5.38: Función de entropía binaria, $H_b(p)$.

5.7. Soluciones de los ejercicios

Ejercicio 5.1 *Solución*

$$H(X) = 2,4087 \text{ bits/símbolo.}$$

En el caso en que los símbolos son igualmente probables

$$H(X) = \log_2(6) = 2,585 \text{ bits/símbolo.}$$

Ejercicio 5.2 *Solución*

a) Definiendo $\alpha = p(1 - \varepsilon) + (1 - p)\varepsilon$

$$H(X) = H_b(p) = -p \log_2(p) - (1 - p) \log_2(1 - p).$$

$$H(Y) = H_b(\alpha) = -\alpha \log_2(\alpha) - (1 - \alpha) \log_2(1 - \alpha).$$

$$H(Y|X) = H_b(\varepsilon).$$

$$H(X, Y) = H_b(p) + H_b(\varepsilon).$$

$$H(X|Y) = H_b(p) + H_b(\varepsilon) - H_b(\alpha)$$

$$I(X, Y) = H_b(\alpha) - H_b(\varepsilon).$$

b) $p = \frac{1}{2}$

c) $\varepsilon = \frac{1}{2}$

Ejercicio 5.3 *Solución*

a) $H(X) = 2,5282 \text{ bits/símbolo.}$ Para símbolos equiprobables

$$H(X) = \log_2(7) = 2,8074 \text{ bits/símbolo.}$$

b) $H(Y) = 1,4060 \text{ bits/símbolo.}$ Para símbolos equiprobables

$$H(Y) = \log_2(3) = 1,585 \text{ bits/símbolo}$$

c)

$$H(X, Y) = 2,5282 \text{ bits/símbolo.}$$

$$H(X|Y) = 1,1222 \text{ bits/símbolo.}$$

$$H(Y|X) = 0 \text{ bits/símbolo.}$$

$$I(X, Y) = 1,4060 \text{ bits/símbolo.}$$

Ejercicio 5.4 *Solución*

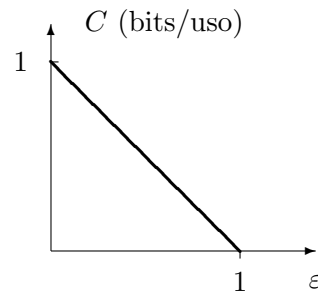


Figura 5.39: Capacidad del canal en función de ε , Ejercicio 5.4.

La capacidad del canal es

$$C = (1 - \varepsilon) \text{ bits/uso,}$$

y que esta capacidad se alcanza para unas probabilidades a priori

$$p_X(0) = p_X(1) = \frac{1}{2}.$$

La capacidad se representa en función de ε en la Figura 5.39.

Ejercicio 5.5 Solución

a) $C = \varepsilon$, para $p_X(x_0) = p_X(x_1) = \frac{1}{2}$.

b)

$$H(X, Y) = H_b(p) + H_b(\varepsilon)$$

$$H(X|Y) = (1 - \varepsilon)H_b(p)$$

Ejercicio 5.6 Solución

a) Para $\varepsilon = 0$

$$C = \log_2(3) = 1,585 \text{ bits/uso.}$$

Las probabilidades de entrada para las que se alcanza la capacidad del canal son las que cumplen que

$$p_X(x_1) = p_X(x_3) = \frac{1}{3}, \quad [p_X(x_0) + p_X(x_2)] = \frac{1}{3}.$$

Para $\varepsilon = 1$

$$C = \log_2(2) = 1 \text{ bit/uso.}$$

Las probabilidades de entrada para las que se alcanza la capacidad del canal son las que cumplen

$$[p_X(x_0) + p_X(x_2)] = \frac{1}{2}, \quad [p_X(x_1) + p_X(x_3)] = \frac{1}{2}.$$

b) $\varepsilon = 0,75$

$$H(X|Y) = 0,931 \text{ bits/símbolo.}$$

c) Las probabilidades para las que se obtiene la capacidad del canal son

$$p_X(x_0) = 0,403, \quad p_X(x_1) = 0,339, \quad p_X(x_2) = 0,258.$$

La capacidad del canal es el máximo de la información mutua

$$C = H(Y) - H(Y|X) = 1,308 \text{ bits/uso}$$

-o-O-o-

Ejercicio 5.7 Solución

a) La capacidad es

$$C = \log_2(3) = 1,585 \text{ bits/uso}$$

para

$$p_X(x_0) = p_X(x_1) = p_X(x_2) = \frac{1}{3}.$$

b) Cualquiera de estos dos valores es posible

$$\varepsilon = \begin{cases} 0,25 \\ 0,75 \end{cases}.$$

$$H(Y) = \begin{cases} 1,5546, & \varepsilon = 0,25 \\ 1,2807, & \varepsilon = 0,75 \end{cases}.$$

$$I(X, Y) = \begin{cases} 1,0137, & \varepsilon = 0,25 \\ 0,7398, & \varepsilon = 0,75 \end{cases}.$$

$$H(X|Y) = H(X, Y) - H(Y) = \begin{cases} 0,5712, & \varepsilon = 0,25 \\ 0,8451, & \varepsilon = 0,75 \end{cases}.$$

Ejercicio 5.8 Solución

a) La capacidad de canal es $C = 1$ bit/uso y se alcanza para

$$p_X(x_1) = 1/2, \quad p_X(x_0) + p_X(x_2) = 1/2.$$

b) $H(X|Y) = 0$ y $H(X, Y) = H_b(p) + p \cdot H_b(\varepsilon)$. Para $\varepsilon = \frac{1}{2}$, $H(X, Y) = H_b(p) + p$.

Ejercicio 5.9 Solución

a) $C = 1$ bit/uso, para $p_X(x_0) = \frac{1}{2}$ y $p_X(x_1) + p_X(x_2) = \frac{1}{2}$.

b)

$$H(X, Y) = 2p + H_b(2p) + (1 - 2p)H_b(\varepsilon)$$

$$H(X|Y) = 2p$$

Ejercicio 5.10 Solución

a)

$$\mathbf{P}^{DMC} = \begin{bmatrix} 1 - \varepsilon & \varepsilon & 0 & 0 \\ \varepsilon & 1 - 2\varepsilon & \varepsilon & 0 \\ 0 & \varepsilon & 1 - 2\varepsilon & \varepsilon \\ 0 & 0 & \varepsilon & 1 - \varepsilon \end{bmatrix}$$

$$\mathbf{P}^{Const} = \begin{bmatrix} 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{1}{\sqrt{N_0/2}}\right) & 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{3}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & 1 - 2Q\left(\frac{1}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{3}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{3}{\sqrt{N_0/2}}\right) & 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right) \end{bmatrix}$$

La aproximación que se ha hecho es que sólo se considera la probabilidad de error entre símbolos que estén a distancia mínima, despreciándose la probabilidad de cometer errores con símbolos a $2d_{min}$ o $3d_{min}$. Eso significa que

$$Q\left(\frac{k}{\sqrt{N_0/2}}\right) \approx 0, \text{ para } k > 1,$$

y que

$$\varepsilon = Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

b)

$$H(Y|X) = \frac{1}{2}H_b(\varepsilon) + \frac{1}{2}H_b(2\varepsilon) + \varepsilon$$

$$H(X|Y) = \frac{1}{2}H_b(\varepsilon) + \frac{1}{2}H_b(2\varepsilon) + \varepsilon$$

$$H(X, Y) = 2 + \frac{1}{2}H_b(\varepsilon) + \frac{1}{2}H_b(2\varepsilon) + \varepsilon$$

$$I(X, Y) = 2 - \frac{1}{2}H_b(\varepsilon) - \frac{1}{2}H_b(2\varepsilon) - \varepsilon$$

El valor de ε que minimiza $H(Y|X)$ y $H(X|Y)$ y maximiza $I(X, Y)$ es $\varepsilon = 0$ (en cuyo caso $H(X|Y) = H(Y|X) = 0$, e $I(X, Y) = 2$).

c) Los valores mínimo y máximo que puede tomar ε son $\varepsilon = 0$ y $\varepsilon = \frac{1}{2}$, respectivamente.

i) Para $\varepsilon = 0$, $C = 2$ bits/uso para $p_X(x_i) = \frac{1}{4}$, $i = 0, 1, 2, 3$.

ii) Para $\varepsilon = \frac{1}{2}$, $C = 1$ bits/uso para $p_X(x_0) = p_X(x_3)$ y $p_X(x_1) = p_X(x_2)$.

Ejercicio 5.11 Solución

a) La matriz de canal para el DMC es

$$\mathbf{P}^{DMC} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \varepsilon & \varepsilon & 0 \\ 0 & \varepsilon & 1 - \varepsilon & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Para la constelación, la matriz de canal correspondiente es

$$\mathbf{P}^{Const.} = \begin{bmatrix} 1 - Q\left(\frac{2}{\sqrt{N_0/2}}\right) & Q\left(\frac{2}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right) - Q\left(\frac{8}{\sqrt{N_0/2}}\right) & Q\left(\frac{8}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{2}{\sqrt{N_0/2}}\right) & 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{2}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{4}{\sqrt{N_0/2}}\right) & Q\left(\frac{4}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{4}{\sqrt{N_0/2}}\right) & Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{4}{\sqrt{N_0/2}}\right) & 1 - Q\left(\frac{1}{\sqrt{N_0/2}}\right) - Q\left(\frac{2}{\sqrt{N_0/2}}\right) & Q\left(\frac{2}{\sqrt{N_0/2}}\right) \\ Q\left(\frac{8}{\sqrt{N_0/2}}\right) & Q\left(\frac{5}{\sqrt{N_0/2}}\right) - Q\left(\frac{8}{\sqrt{N_0/2}}\right) & Q\left(\frac{2}{\sqrt{N_0/2}}\right) - Q\left(\frac{5}{\sqrt{N_0/2}}\right) & 1 - Q\left(\frac{2}{\sqrt{N_0/2}}\right) \end{bmatrix}.$$

La aproximación que se ha hecho es que la probabilidad de transición con símbolos más alejados de distancia 2 es nula, o lo que es lo mismo, que sólo se producen errores entre los símbolos \mathbf{a}_1 y \mathbf{a}_2 , lo que sobre los términos de la matriz anterior significa

$$Q\left(\frac{k}{\sqrt{N_0/2}}\right) \approx 0, \text{ para } k > 1.$$

Por tanto

$$\varepsilon = Q\left(\frac{1}{\sqrt{N_0/2}}\right).$$

b)

$$H(X, Y) = 2 + \frac{1}{2} \cdot H_b(\varepsilon)$$

$$H(Y|X) = \frac{1}{2} \cdot H_b(\varepsilon).$$

$$H(X|Y) = \frac{1}{2} \cdot H_b(\varepsilon).$$

$$I(X, Y) = 2 - \frac{1}{2} \cdot H_b(\varepsilon).$$

La representación de estas funciones se muestra en la Figura 5.40.

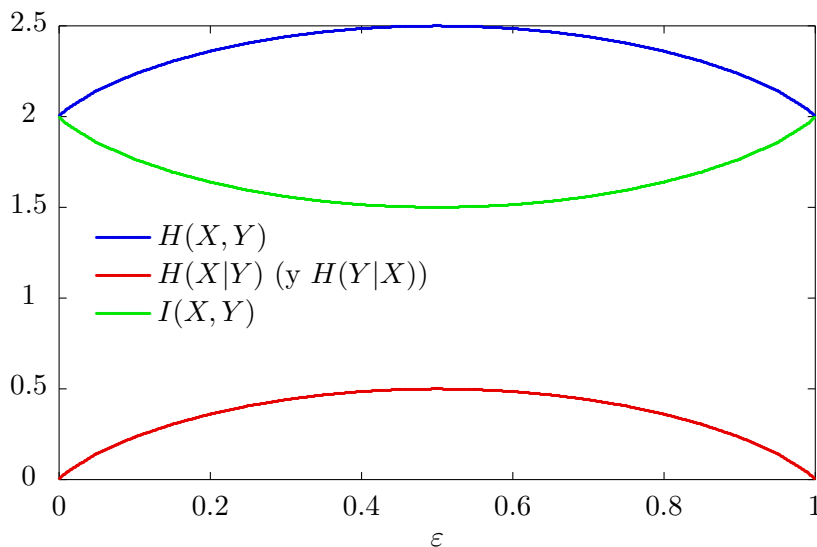


Figura 5.40: Representación de $H(X, Y)$, $H(X|Y)$ (que es igual que $H(Y|X)$) e $I(X, Y)$ en función de ε .

Los valores máximos de $H(X, Y)$, $H(X|Y)$, y $H(Y|X)$ se obtienen para $\varepsilon = 1/2$, lo mismo que ocurre con el valor mínimo de $I(X, Y)$.

c) La capacidad de canal es

$$C = 1 + H_b\left(\frac{2^{H_b(\varepsilon)}}{1 + 2^{H_b(\varepsilon)}}\right) - \left(\frac{1}{1 + 2^{H_b(\varepsilon)}}\right) \cdot H_b(\varepsilon),$$

valor que se obtiene para

$$p_X(x_0) = p_X(x_3) = \frac{2^{H_b(\varepsilon)-1}}{1 + 2^{H_b(\varepsilon)}}, \quad p_X(x_1) = p_X(x_2) = \frac{1}{2 \cdot (1 + 2^{H_b(\varepsilon)})}.$$

Ejercicio 5.12 *Solución*

1. $a = d = 0, b = c = \varepsilon_1$.

$$H(Y) = 2 + \frac{1}{2} \left[(1 - \varepsilon_0) \log_2 \frac{1}{1 - \varepsilon_0} + (1 + \varepsilon_0) \log_2 \frac{1}{1 + \varepsilon_0} \right].$$

Si se tiene en cuenta que $\frac{1}{2}(1 + \varepsilon_0) = 1 - \frac{1}{2}(1 - \varepsilon_1)$, se podrían reagrupar términos y escribir

$$H(Y) = 1 + H_b \left(\frac{1}{2}(1 - \varepsilon_0) \right).$$

$$H(X, Y) = 2 + \frac{1}{2}H_b(\varepsilon_0) + \frac{1}{2}H_b(\varepsilon_1)$$

$$H(Y|X) = \frac{1}{2}H_b(\varepsilon_0) + \frac{1}{2}H_b(\varepsilon_1)$$

$$H(X|Y) = 1 + \frac{1}{2}H_b(\varepsilon_0) + \frac{1}{2}H_b(\varepsilon_1) - H_b \left(\frac{1}{2}(1 - \varepsilon_0) \right)$$

$$I(X, Y) = 1 + H_b \left(\frac{1}{2}(1 - \varepsilon_0) \right) - \frac{1}{2}H_b(\varepsilon_0) - \frac{1}{2}H_b(\varepsilon_1)$$

2. La capacidad del canal es

$$C = 1 + H_b(p(1 - \varepsilon_0)) - pH_b(\varepsilon_0) - (1 - p)\frac{1}{2}H_b(\varepsilon_1) \text{ bits/uso}$$

para

$$p = \frac{1}{(1 - \varepsilon_0) \cdot \left[1 + 2^{\frac{H_b(\varepsilon_0) - H_b(\varepsilon_1)}{1 - \varepsilon_0}} \right]}.$$

La distribución para la que se consigue la capacidad es

$$p_X(x_0) = p_X(x_3) = \frac{p}{2}, \quad p_X(x_1) = p_X(x_2) = \frac{1 - p}{2}$$

Ejercicio 5.13 *Solución*

- a)

$$H(Y) = \frac{2}{3}(1 - \varepsilon_0 + \varepsilon_1) + H_b \left(\frac{2}{3}(1 - \varepsilon_0 + \varepsilon_1) \right)$$

$$H(Y|X) = \frac{2}{3}H_b(\varepsilon_0) + \frac{1}{3}H_b(2\varepsilon_1) + \frac{2}{3}\varepsilon_1.$$

$$H(X, Y) = \log_2(3) + \frac{2}{3}H_b(\varepsilon_0) + \frac{1}{3}H_b(2\varepsilon_1) + \frac{2}{3}\varepsilon_1.$$

$$H(X|Y) = H(X, Y) - H(Y)$$

$$I(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

- b) En este caso el canal que nos queda es el *binary erasure channel* (BEC)

$$C = (1 - \varepsilon_0) \text{ bits/símbolo, para } p_X(x_0) = p_X(x_2) = \frac{1}{2}.$$

c)

$$C = p(1 - \varepsilon_0) + (1 - p)2\varepsilon_1 + H_b(p(1 - \varepsilon_0) + (1 - p)2\varepsilon_1) - pH_b(\varepsilon_0) - (1 - p)(H_b(2\varepsilon_1) + 2\varepsilon_1)$$

para el valor

$$p = \frac{(1 - 2\varepsilon_1) \cdot 2^{\frac{H_b(\varepsilon_0) - H_b(2\varepsilon_1) - 2\varepsilon_1}{2\varepsilon_1 + \varepsilon_0 - 1}} - \varepsilon_1}{(1 - 2\varepsilon_1 - \varepsilon_0) \cdot 2^{\frac{H_b(\varepsilon_0) - H_b(2\varepsilon_1) - 2\varepsilon_1}{2\varepsilon_1 + \varepsilon_0 - 1}} + \frac{1}{2} \cdot (1 - \varepsilon_0 - 2\varepsilon_1)}$$

siendo la distribución de entrada

$$p_X(x_0) = p_X(x_2) = \frac{p}{2}, \quad p_X(x_1) = 1 - p$$

d) Si se tiene que $\varepsilon_0 = \varepsilon_1$, es muy fácil ver que para $\varepsilon_0 = \varepsilon_1 = 0$ se tiene un canal ideal, en el que $H(Y|X) = 0$, y la capacidad se alcanza con símbolos de entrada equiprobables, que generan símbolos de salida equiprobables y por tanto la capacidad vale $\log_2(3)$, que es el máximo valor posible para un sistema con tres entradas y tres salidas.

Ejercicio 5.14 Solución

a) $a = b = c = d = 0$. Son dos canales BSC en paralelo con distintas probabilidades de error. Por lo tanto sería un buen modelo para un sistema que enviara bits a través de dos subsistemas independientes.

b)

$$H(X) = 1 + H_b(p).$$

$$H(Y) = 1 + H_b(p)$$

$$H(X, Y) = 1 + H_b(p) + pH_b(\varepsilon_1) + (1 - p)H_b(\varepsilon_0)$$

$$H(Y|X) = pH_b(\varepsilon_1) + (1 - p)H_b(\varepsilon_0)$$

$$H(X|Y) = pH_b(\varepsilon_1) + (1 - p)H_b(\varepsilon_0)$$

c)

$$C = 1 + H_b(p) - pH_b(\varepsilon_1) - (1 - p)H_b(\varepsilon_0)$$

con

$$p = \frac{1}{1 + 2H_b(\varepsilon_1) - H_b(\varepsilon_0)}$$

siendo la distribución de entrada para la que se alcanza

$$p_X(x_0) = p_X(x_1) = \frac{p}{2}, \quad p_X(x_2) = p_X(x_3) = \frac{1 - p}{2}$$

Ejercicio 5.15 Solución

a)

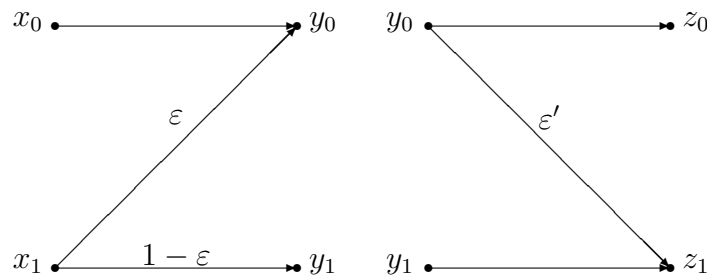
$$H(X) = H_b(\alpha), \quad H(Y) = H_b((1 - \alpha)(1 - \varepsilon)), \quad H(Z) = H_b(\beta(1 - \varepsilon'))$$

$$H(X, Y) = H_b(\alpha) + (1 - \alpha)H_b(\varepsilon)$$

$$H(Y|Z) = H_b(\beta) + \beta H_b(\varepsilon') - H_b(\beta(1 - \varepsilon'))$$

b) Las probabilidades de transición, dadas en la matriz de canal, son

$$P_{Y|X} = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1 - \varepsilon \end{bmatrix}, \quad P_{Z|Y} = \begin{bmatrix} 1 - \varepsilon' & \varepsilon' \\ 0 & 1 \\ 1 - \varepsilon' & \varepsilon' \end{bmatrix}$$



c) $I(X, Y) = H_b((1 - \alpha)(1 - \varepsilon)) - (1 - \alpha)H_b(\varepsilon)$

d)

$$P_{Z|X} = P_{Y|X} \times P_{Z|Y} = \begin{bmatrix} 1 - \varepsilon' & \varepsilon' \\ \varepsilon(1 - \varepsilon') & \varepsilon\varepsilon' + 1 - \varepsilon \end{bmatrix}$$

e) La concatenación será un BSC si $\varepsilon = \frac{\varepsilon'}{1 - \varepsilon'}$

Ejercicio 5.16 Solución

a) La matriz de canal es

$$P = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1 - \varepsilon \\ 0 & 1 \end{bmatrix}.$$

b) Valores de las entropías en los dos escenarios

I)

$$H(X) = \log_2 3 = 1,585 \text{ bits/símbolo.}$$

$$H(Y) = H_b\left(\frac{1}{3}(1 + \varepsilon)\right) = H_b\left(\frac{1}{3}(2 - \varepsilon)\right) \text{ bits/símbolo.}$$

$$H(Y|X) = p_1 \cdot H_b(\varepsilon) = \frac{1}{3} \cdot H_b(\varepsilon) \text{ bits/símbolo.}$$

$$H(X, Y) = \log_2 3 + \frac{1}{3} \cdot H_b(\varepsilon) \text{ bits/símbolo.}$$

$$H(X|Y) = \log_2 3 + \frac{1}{3} \cdot H_b(\varepsilon) - H_b\left(\frac{1}{3}(1 + \varepsilon)\right) \text{ bits/símbolo.}$$

$$I(X, Y) = H_b\left(\frac{1}{3}(1 + \varepsilon)\right) - \frac{1}{3} \cdot H_b(\varepsilon) \text{ bits/símbolo.}$$

II)

$$H(X) = 2 \times 0 \log_2 0 + 1 \log_2 1 = 0 \text{ bits/símbolo,}$$

$$H(Y) = H_b(\varepsilon) \text{ bits/símbolo.}$$

$$H(Y|X) = p_1 \cdot H_b(\varepsilon) = H_b(\varepsilon) \text{ bits/símbolo.}$$

$$H(X, Y) = 0 + H_b(\varepsilon) = H_b(\varepsilon) \text{ bits/símbolo.}$$

$$H(X|Y) = 0 \text{ bits/símbolo,}$$

$$I(X, Y) = 0 \text{ bits/símbolo.}$$

c) La capacidad de canal, para cada caso, vale

I) $C = \max_{p_X(x_i)} I(X, Y) = 1$ bit/uso, valor que se alcanza para $p = 0$, lo que significa que

$$p_0 = \frac{1}{2}, p_1 = 0, p_2 = \frac{1}{2}.$$

II)

$$C = H_b \left(\frac{1}{1 + 2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}} \right) - \frac{1}{(1-\varepsilon) \cdot \left(1 + 2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}\right)} \cdot H_b(\varepsilon),$$

que se obtiene para

$$p_X(x_0) = 1 - p = 1 - \frac{1}{(1-\varepsilon) \cdot \left(1 + 2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}\right)}, p_X(x_1) = p = \frac{1}{(1-\varepsilon) \cdot \left(1 + 2^{\frac{H_b(\varepsilon)}{1-\varepsilon}}\right)}.$$

Apéndice A

Tablas de interés

A continuación se presentan algunas tablas de interés en la asignatura. En concreto, se presentan la tabla con la función $Q(x)$, de especial utilidad en el cálculo de probabilidades de error, y unas tablas con algunas de las transformadas de Fourier más comunes y las propiedades básicas de dicha transformada.

Dominio del Tiempo ($x(t)$) Dominio de la Frecuencia ($X(j\omega)$)

$\delta(t)$	1
1	$2\pi \cdot \delta(\omega)$
$\delta(t - t_0)$	$e^{-j\omega t_0}$
$e^{j\omega_0 t}$	$2\pi \cdot \delta(\omega - \omega_0)$
$\cos(\omega_0 t)$	$\pi \cdot \delta(\omega - \omega_0) + \pi \cdot \delta(\omega + \omega_0)$
$\text{sen}(\omega_0 t)$	$\frac{\pi}{j} \cdot \delta(\omega - \omega_0) - \frac{\pi}{j} \cdot \delta(\omega + \omega_0)$
$\Pi\left(\frac{t}{T}\right) = \begin{cases} 1, & t \leq \frac{T}{2} \\ 0, & t > \frac{T}{2} \end{cases}$	$T \cdot \text{sinc}\left(\frac{\omega T}{2\pi}\right)$
$\text{sinc}\left(\frac{t}{T}\right)$	$T \cdot \Pi\left(\frac{\omega T}{2\pi}\right)$
$\Lambda\left(\frac{t}{T}\right) = \begin{cases} 1 - \frac{ t }{T}, & t \leq T \\ 0 & t > T \end{cases}$	$T \cdot \text{sinc}^2\left(\frac{\omega T}{2\pi}\right)$
$\text{sinc}^2\left(\frac{t}{T}\right)$	$T \cdot \Lambda\left(\frac{\omega T}{2\pi}\right)$
$u(t)$	$\frac{1}{j\omega} + \pi \cdot \delta(\omega)$
$e^{-\alpha t} \cdot u(t), \alpha > 0$	$\frac{1}{\alpha + j\omega}$
$t \cdot e^{-\alpha t} \cdot u(t), \alpha > 0$	$\frac{1}{(\alpha + j\omega)^2}$
$\frac{t^{n-1}}{(n-1)!} \cdot e^{-\alpha t} \cdot u(t), \alpha > 0$	$\frac{1}{(\alpha + j\omega)^n}$
$e^{-\alpha t }$	$\frac{2\alpha}{\alpha^2 + \omega^2}$
$e^{-\pi t^2}$	$e^{-\pi f^2} = e^{-\frac{\omega^2}{4\pi}}$
$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \\ 0, & t = 0 \end{cases}$	$\frac{2}{j\omega}$
$\delta'(t)$	$j\omega$
$\delta^{(n)}(t)$	$(j\omega)^n$
$\frac{1}{t}$	$-j\pi \cdot \text{sgn}(\omega)$
$\sum_{n=-\infty}^{\infty} \delta(t - nT)$	$\frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi k}{T}\right)$

Tabla A.1: Tabla de transformadas de Fourier (en $j\omega$)

Dominio del Tiempo ($x(t)$) Dominio de la Frecuencia ($X(j\omega)$)

$a \cdot x(t) + b \cdot y(t)$	$a \cdot X(j\omega) + b \cdot Y(j\omega)$
$x(t - t_0)$	$e^{-j\omega t_0} \cdot X(j\omega)$
$e^{j\omega_0 t} \cdot x(t)$	$X(\omega - \omega_0)$
$x^*(t)$	$X^*(-\omega)$
$x(-t)$	$X(-\omega)$
$x(a \cdot t)$	$\frac{1}{ a } X\left(\frac{\omega}{a}\right)$
$x(t) * y(t)$	$X(j\omega) \cdot Y(j\omega)$
$x(t) \cdot y(t)$	$\frac{1}{2\pi} \cdot X(j\omega) * Y(j\omega)$
$\frac{d}{dt} x(t)$	$j\omega \cdot X(j\omega)$
$\int_{-\infty}^t x(t) dt$	$\frac{1}{j\omega} \cdot X(\omega) + \pi \cdot X(0) \cdot \delta(\omega)$
$t \cdot x(t)$	$j \frac{d}{d\omega} X(j\omega)$
$x(t)$ real	$X(j\omega) = X^*(-j\omega)$

Relación de Parseval para señales aperiódicas

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(j\omega)|^2 d\omega$$

Propiedad de dualidad

$$f(u) = \int_{-\infty}^{\infty} g(v) \cdot e^{-juv} dv$$

$$g(t) \xleftrightarrow{TF} f(\omega)$$

$$f(t) \xleftrightarrow{TF} 2\pi \cdot g(-\omega)$$

Tabla A.2: Tabla de propiedades de la transformada de Fourier.

x	$Q(x)$	x	$Q(x)$	x	$Q(x)$	x	$Q(x)$
0.00	5.00000×10^{-1}	2.05	2.01822×10^{-2}	4.10	2.06575×10^{-5}	6.15	3.87415×10^{-10}
0.05	4.80061×10^{-1}	2.10	1.78644×10^{-2}	4.15	1.66238×10^{-5}	6.20	2.82316×10^{-10}
0.10	4.60172×10^{-1}	2.15	1.57776×10^{-2}	4.20	1.33457×10^{-5}	6.25	2.05226×10^{-10}
0.15	4.40382×10^{-1}	2.20	1.39034×10^{-2}	4.25	1.06885×10^{-5}	6.30	1.48823×10^{-10}
0.20	4.20740×10^{-1}	2.25	1.22245×10^{-2}	4.30	8.53991×10^{-6}	6.35	1.07657×10^{-10}
0.25	4.01294×10^{-1}	2.30	1.07241×10^{-2}	4.35	6.80688×10^{-6}	6.40	7.76885×10^{-11}
0.30	3.82089×10^{-1}	2.35	9.38671×10^{-3}	4.40	5.41254×10^{-6}	6.45	5.59251×10^{-11}
0.35	3.63169×10^{-1}	2.40	8.19754×10^{-3}	4.45	4.29351×10^{-6}	6.50	4.01600×10^{-11}
0.40	3.44578×10^{-1}	2.45	7.14281×10^{-3}	4.50	3.39767×10^{-6}	6.55	2.87685×10^{-11}
0.45	3.26355×10^{-1}	2.50	6.20967×10^{-3}	4.55	2.68230×10^{-6}	6.60	2.05579×10^{-11}
0.50	3.08538×10^{-1}	2.55	5.38615×10^{-3}	4.60	2.11245×10^{-6}	6.65	1.46547×10^{-11}
0.55	2.91160×10^{-1}	2.60	4.66119×10^{-3}	4.65	1.65968×10^{-6}	6.70	1.04210×10^{-11}
0.60	2.74253×10^{-1}	2.65	4.02459×10^{-3}	4.70	1.30081×10^{-6}	6.75	7.39226×10^{-12}
0.65	2.57846×10^{-1}	2.70	3.46697×10^{-3}	4.75	1.01708×10^{-6}	6.80	5.23096×10^{-12}
0.70	2.41964×10^{-1}	2.75	2.97976×10^{-3}	4.80	7.93328×10^{-7}	6.85	3.69250×10^{-12}
0.75	2.26627×10^{-1}	2.80	2.55513×10^{-3}	4.85	6.17307×10^{-7}	6.90	2.60013×10^{-12}
0.80	2.11855×10^{-1}	2.85	2.18596×10^{-3}	4.90	4.79183×10^{-7}	6.95	1.82643×10^{-12}
0.85	1.97663×10^{-1}	2.90	1.86581×10^{-3}	4.95	3.71067×10^{-7}	7.00	1.27981×10^{-12}
0.90	1.84060×10^{-1}	2.95	1.58887×10^{-3}	5.00	2.86652×10^{-7}	7.05	8.94589×10^{-13}
0.95	1.71056×10^{-1}	3.00	1.34990×10^{-3}	5.05	2.20905×10^{-7}	7.10	6.23784×10^{-13}
1.00	1.58655×10^{-1}	3.05	1.14421×10^{-3}	5.10	1.69827×10^{-7}	7.15	4.33890×10^{-13}
1.05	1.46859×10^{-1}	3.10	9.67603×10^{-4}	5.15	1.30243×10^{-7}	7.20	3.01063×10^{-13}
1.10	1.35666×10^{-1}	3.15	8.16352×10^{-4}	5.20	9.96443×10^{-8}	7.25	2.08386×10^{-13}
1.15	1.25072×10^{-1}	3.20	6.87138×10^{-4}	5.25	7.60496×10^{-8}	7.30	1.43884×10^{-13}
1.20	1.15070×10^{-1}	3.25	5.77025×10^{-4}	5.30	5.79013×10^{-8}	7.35	9.91034×10^{-14}
1.25	1.05650×10^{-1}	3.30	4.83424×10^{-4}	5.35	4.39771×10^{-8}	7.40	6.80922×10^{-14}
1.30	9.68005×10^{-2}	3.35	4.04058×10^{-4}	5.40	3.33204×10^{-8}	7.45	4.66701×10^{-14}
1.35	8.85080×10^{-2}	3.40	3.36929×10^{-4}	5.45	2.51849×10^{-8}	7.50	3.19089×10^{-14}
1.40	8.07567×10^{-2}	3.45	2.80293×10^{-4}	5.50	1.89896×10^{-8}	7.55	2.17629×10^{-14}
1.45	7.35293×10^{-2}	3.50	2.32629×10^{-4}	5.55	1.42835×10^{-8}	7.60	1.48065×10^{-14}
1.50	6.68072×10^{-2}	3.55	1.92616×10^{-4}	5.60	1.07176×10^{-8}	7.65	1.00490×10^{-14}
1.55	6.05708×10^{-2}	3.60	1.59109×10^{-4}	5.65	8.02239×10^{-9}	7.70	6.80331×10^{-15}
1.60	5.47993×10^{-2}	3.65	1.31120×10^{-4}	5.70	5.99037×10^{-9}	7.75	4.59463×10^{-15}
1.65	4.94715×10^{-2}	3.70	1.07800×10^{-4}	5.75	4.46217×10^{-9}	7.80	3.09536×10^{-15}
1.70	4.45655×10^{-2}	3.75	8.84173×10^{-5}	5.80	3.31575×10^{-9}	7.85	2.08019×10^{-15}
1.75	4.00592×10^{-2}	3.80	7.23480×10^{-5}	5.85	2.45787×10^{-9}	7.90	1.39452×10^{-15}
1.80	3.59303×10^{-2}	3.85	5.90589×10^{-5}	5.90	1.81751×10^{-9}	7.95	9.32558×10^{-16}
1.85	3.21568×10^{-2}	3.90	4.80963×10^{-5}	5.95	1.34071×10^{-9}	8.00	6.22096×10^{-16}
1.90	2.87166×10^{-2}	3.95	3.90756×10^{-5}	6.00	9.86588×10^{-10}		
1.95	2.55881×10^{-2}	4.00	3.16712×10^{-5}	6.05	7.24229×10^{-10}		
2.00	2.27501×10^{-2}	4.05	2.56088×10^{-5}	6.10	5.30342×10^{-10}		

 Tabla A.3: Función $Q(x)$

IGUALDADES TRIGONOMÉTRICAS

$$\begin{aligned}
 \operatorname{sen}^2(a) + \operatorname{cos}^2(a) &= 1 \\
 \tan^2(a) + 1 &= \sec^2(a) \\
 \cot^2(a) + 1 &= \operatorname{csc}^2(a) \\
 \operatorname{cos}(a) \cdot \operatorname{cos}(b) &= \frac{1}{2} \cdot \operatorname{cos}(a - b) + \frac{1}{2} \cdot \operatorname{cos}(a + b) \\
 \operatorname{sen}(a) \cdot \operatorname{sen}(b) &= \frac{1}{2} \cdot \operatorname{cos}(a - b) - \frac{1}{2} \cdot \operatorname{cos}(a + b) \\
 \operatorname{cos}(a) \cdot \operatorname{sen}(b) &= \frac{1}{2} \cdot \operatorname{sen}(a + b) - \frac{1}{2} \cdot \operatorname{sen}(a - b) \\
 \operatorname{sen}(a \pm b) &= \operatorname{sen}(a) \cdot \operatorname{cos}(b) \pm \operatorname{cos}(a) \cdot \operatorname{sen}(b) \\
 \operatorname{cos}(a \pm b) &= \operatorname{cos}(a) \cdot \operatorname{cos}(b) \mp \operatorname{sen}(a) \cdot \operatorname{sen}(b) \\
 \operatorname{sen}(a) + \operatorname{sen}(b) &= 2 \cdot \operatorname{sen}\left(\frac{a+b}{2}\right) \cdot \operatorname{cos}\left(\frac{a-b}{2}\right) \\
 \operatorname{cos}(a) + \operatorname{cos}(b) &= 2 \cdot \operatorname{cos}\left(\frac{a+b}{2}\right) \cdot \operatorname{cos}\left(\frac{a-b}{2}\right) \\
 \operatorname{sen}(a) - \operatorname{sen}(b) &= 2 \cdot \operatorname{cos}\left(\frac{a+b}{2}\right) \cdot \operatorname{sen}\left(\frac{a-b}{2}\right) \\
 \operatorname{cos}(a) - \operatorname{cos}(b) &= -2 \cdot \operatorname{sen}\left(\frac{a+b}{2}\right) \cdot \operatorname{sen}\left(\frac{a-b}{2}\right) \\
 \operatorname{sen}^2(a) - \operatorname{sen}^2(b) &= \operatorname{sen}(a+b) \cdot \operatorname{sen}(a-b) \\
 \operatorname{cos}^2(a) - \operatorname{sen}^2(b) &= \operatorname{cos}(a+b) \cdot \operatorname{cos}(a-b) \\
 \operatorname{cos}(a) &= \frac{e^{+ja} + e^{-ja}}{2}, \quad \operatorname{sen}(a) = \frac{e^{+ja} - e^{-ja}}{2j} \\
 e^{ja} &= \operatorname{cos}(a) + j \operatorname{sen}(a)
 \end{aligned}$$

INTEGRALES DE FUNCIONES TRIGONOMÉTRICAS

$$\begin{aligned}
 \int \operatorname{cos}(a \cdot t) dt &= \frac{1}{a} \operatorname{sen}(a \cdot t) \\
 \int \operatorname{sen}(a \cdot t) dt &= -\frac{1}{a} \operatorname{cos}(a \cdot t) \\
 \int \operatorname{cos}^2(a \cdot t) dt &= \frac{t}{2} + \frac{1}{2a} \cdot \operatorname{cos}(a \cdot t) \cdot \operatorname{sen}(a \cdot t) = \frac{t}{2} + \frac{1}{4a} \cdot \operatorname{sen}(2a \cdot t) \\
 \int \operatorname{sen}^2(a \cdot t) dt &= \frac{t}{2} - \frac{1}{2a} \cdot \operatorname{cos}(a \cdot t) \cdot \operatorname{sen}(a \cdot t) = \frac{t}{2} - \frac{1}{4a} \cdot \operatorname{sen}(2a \cdot t) \\
 \int \operatorname{cos}(a \cdot t) \cdot \operatorname{cos}(b \cdot t) dt &= \frac{1}{2} \cdot \frac{\operatorname{sen}((a-b) \cdot t)}{a-b} + \frac{1}{2} \cdot \frac{\operatorname{sen}((a+b) \cdot t)}{a+b} \\
 \int \operatorname{sen}(a \cdot t) \cdot \operatorname{sen}(b \cdot t) dt &= \frac{1}{2} \cdot \frac{\operatorname{sen}((a-b) \cdot t)}{a-b} - \frac{1}{2} \cdot \frac{\operatorname{sen}((a+b) \cdot t)}{a+b} \\
 \int \operatorname{cos}(a \cdot t) \cdot \operatorname{sen}(b \cdot t) dt &= \frac{1}{2} \cdot \frac{\operatorname{cos}((a-b) \cdot t)}{a-b} - \frac{1}{2} \cdot \frac{\operatorname{cos}((a+b) \cdot t)}{a+b}
 \end{aligned}$$

Tabla A.4: Igualdades trigonométricas e integrales trigonométricas.

Dominio del Tiempo ($x(t)$)	Dominio de la Frecuencia ($X(f)$)
$\delta(t)$	1
1	$\delta(f)$
$\delta(t - t_0)$	$e^{-j2\pi f t_0}$
$e^{j2\pi f_0 t}$	$\delta(f - f_0)$
$\cos(2\pi f_0 t)$	$\frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)$
$\text{sen}(2\pi f_0 t)$	$\frac{1}{2j}\delta(f - f_0) - \frac{1}{2j}\delta(f + f_0)$
$\Pi(t) = \begin{cases} 1, & t < \frac{1}{2} \\ \frac{1}{2}, & t = \pm\frac{1}{2} \\ 0, & \text{en otro caso} \end{cases}$	$\text{sinc}(f)$
$\text{sinc}(t)$	$\Pi(f)$
$\Lambda(t) = \begin{cases} t + 1, & -1 \leq t < 0 \\ -t + 1, & 0 \leq t < 1 \\ 0, & \text{en otro caso} \end{cases}$	$\text{sinc}^2(f)$
$\text{sinc}^2(f)$	$\Lambda(f)$
$e^{-\alpha t}u(t), \alpha > 0$	$\frac{1}{\alpha + j2\pi f}$
$te^{-\alpha t}u(t), \alpha > 0$	$\frac{1}{(\alpha + j2\pi f)^2}$
$e^{-\alpha t }$	$\frac{2\alpha}{\alpha^2 + (2\pi f)^2}$
$e^{-\pi t^2}$	$e^{-\pi f^2}$
$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ -1, & t < 0 \\ 0, & t = 0 \end{cases}$	$\frac{1}{j\pi f}$
$u(t)$	$\frac{1}{2}\delta(f) + \frac{1}{j2\pi f}$
$\delta'(t)$	$j2\pi f$
$\delta^{(n)}(t)$	$(j2\pi f)^n$
$\frac{1}{t}$	$-j\pi \text{sgn}(f)$
$\sum_{n=-\infty}^{\infty} \delta(t - nT_0)$	$\frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_0}\right)$

Tabla A.5: Tabla de transformadas de Fourier (en f)

Bibliografía

- [Artés-Rodríguez et al., 2007] A. ARTÉS-RODRÍGUEZ, F. PÉREZ-GONZÁLEZ, J. CID-SUEIRO, R. LÓPEZ-VALCARCE, C. MOSQUERA-NARTALLO, y F. PÉREZ-CRUZ. *Comunicaciones Digitales*. Pearson Educación. 2007.
- [Proakis y Salehi, 2002] J. G. PROAKIS y M. SALEHI. *Communication System Engineering*. Prentice-Hall, Upper Saddle River, New Jersey, segunda edición. 2002.