

Capítulo 4

Fuentes de información y codificación de fuente

- Modelado de fuentes de información
- Codificación de fuente sin pérdidas
- Codificación de fuente con pérdidas
- Cuantización
- Técnicas de audio digital

Los sistemas de comunicaciones se diseñan para transmitir información. En cualquier sistema de comunicaciones existe una fuente de información y el propósito del sistema es transmitir la salida de la fuente al destino. En radiodifusión, por ejemplo, la fuente de información es o una fuente de voz, o una fuente de música. En el caso de televisión, la fuente de información es una fuente de video cuya salida es un imán en movimiento. En la transmisión de fax, la fuente de información produce una imagen fija. En la comunicación entre computadoras se transmiten datos binarios o caracteres ASCII y en consecuencia, la fuente puede modelarse como binaria o ASCII. En el almacenamiento de datos binarios en un disco de computadora, la fuente es nuevamente binaria.

Cualquiera tiene una noción intuitiva del significado de información. Sin embargo, el análisis del desempeño de los sistemas de comunicaciones difícilmente podría concebirse sin una medida cuantitativa de información y el modelado matemático de las fuentes de información. Hartley, Nyquist y Shannon fueron los pioneros en la definición cuantitativa de medidas de información. En este capítulo, investigaremos el modelado matemático de las fuentes de información y se proveerá una medida de información. Luego veremos como la salida de una fuente de información puede hacerse más compacta y, en consecuencia, más fácil de transmitir o almacenar.

4.1 Modelado de fuentes de información

La noción intuitiva y común de información se refiere a cualquier conocimiento nuevo sobre algo. Es posible obtener información escuchando, viendo, o por otros medios de percepción. La fuente de información, en consecuencia, produce salidas que son del interés del receptor de la información, quién no conoce esas salidas por adelantado. El rol del diseñador de los sistemas de comunicaciones es asegurar que esta información se transmita al receptor correctamente. Dado que la salida de la fuente de información es una función variante en el tiempo impredecible (si fuera predecible no sería necesario transmitirla), puede modelarse mediante un proceso estocástico. Ya vimos en el capítulo anterior que en los canales de comunicaciones la existencia de ruido introduce una dependencia estadística entre la entrada y la salida del canal. En consecuencia, el diseñador de sistemas de comunicaciones proyecta un sistema que transmite la salida de un proceso estocástico (fuente de información) a un destino a través de un medio estocástico (el canal) y asegura baja distorsión.

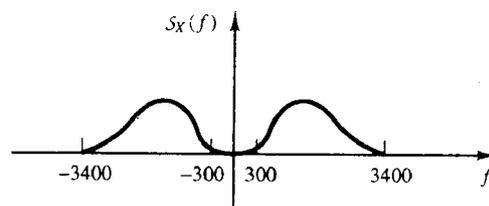


Figura 4.1: Espectro de potencia típico de una señal de voz.

Las fuentes de información pueden modelarse mediante procesos estocásticos, y las propiedades del proceso estocástico dependerán de la naturaleza de la fuente de información. Por ejemplo, cuando modelamos señales de voz, el proceso estocástico resultante tiene toda su potencia en una banda de frecuencias de 300 a 3400 Hz. En consecuencia, la densidad espectral de potencia puede ilustrarse como en la figura 4.1.

Las señales de video se obtienen del barrido de una imagen fija o en movimiento y entonces el ancho de banda dependerá de la resolución requerida. Para la transmisión de TV, dependiendo del sistema empleado (PAL, NTSC o SECAM), esta banda está típicamente entre 0-4.5 MHz y 0-6.5 MHz. Para el caso de telemetría, el ancho de banda dependerá obviamente de la velocidad de cambio de los datos.

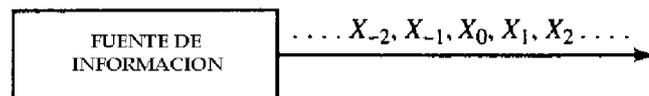


Figura 4.2: Modelo matemático de una fuente de información discreta en tiempo.

El factor común en todos esos procesos es que son de banda limitada, o sea, pueden muestrearse a la velocidad de Nyquist o por encima y reconstruirse de sus valores muestreados. En consecuencia, tiene sentido concentrarse en procesos estocásticos discretos en este capítulo porque todas las fuentes de información de interés pueden modelarse por esos procesos. El modelo matemático de una fuente de información se ilustra en la figura 4.2. Aquí la fuente se modela mediante variables aleatorias X_i definidas como discretas (en la transmisión de datos binarios, por ejemplo) o continuas (voz muestreada por ejemplo). Las propiedades estadísticas del proceso estocástico discreto dependerán de la naturaleza de la fuente de información.

En este capítulo se estudiarán solo modelos simples de fuentes de información. Sin embargo, aún estos modelos simples permiten definir precisamente una medida de información y cotas sobre la representación eficientes de la información.

El modelo más simple de fuente de información que se estudiará es la *fente discreta sin memoria* (fds). una fds es un proceso estocástico de amplitud y tiempo discretos en el cual las X_i son generadas independientemente y con la misma distribución. En consecuencia, una fds genera una secuencia de variables aleatorias distribuidas idénticamente e independientemente (i.i.d.), de forma que tomen valores en un conjunto discreto.

$\mathcal{A} = \{a_1, a_2, \dots, a_N\}$ se definirá como el conjunto de valores que puede tomar la variable aleatoria X , y las probabilidades de tomar cada valor como $p_i = p(X = a_i)$, para todo $i = 1, 2, \dots, N$. Una descripción

completa de la fds está dada por el conjunto \mathcal{A} , denominado el *alfabeto*, y las probabilidades $\{p_i\}_{i=1}^N$.

Ejemplo: Una fuente de información se describe mediante el alfabeto $\mathcal{A} = \{0, 1\}$ y también $p(X_i = 1) = 1 - p(X_i = 0) = p$. Este es un ejemplo de una fds. En el caso especial en que $p = 0.5$, la fuente se denomina *fuentes binaria simétrica*.

4.1.1 Medida de información

Para dar una idea cuantitativa de información, comenzaremos con un modelo básico de una fuente de información y trataremos de definir el contenido de información de la fuente para satisfacer ciertas propiedades intuitivas. Supondremos que la fuente que estamos analizando es una fuente discreta. Las salidas de la fuente deben ser transmitidas a un destino específico donde, en particular, a_1 es la salida más probable y a_N la menos probable. Por ejemplo, es posible imaginar que la fuente representa por un lado la condición del tiempo y por otro la contaminación del aire en cierta ciudad en Enero. En ese caso, \mathcal{A} representa varias combinaciones de condiciones del tiempo y contaminación tal como caluroso y contaminado, muy frío y levemente contaminado, frío y altamente contaminado, frío y medianamente contaminado, muy frío y levemente contaminado, etc. La pregunta es: cual salida provee mayor información, a_1 ó a_N ? (la más probable o la menos probable?). Intuitivamente, saber a_N (o equivalentemente, muy frío y levemente contaminado) revela mayor información. De esto es posible inferir que una medida racional de información para la salida de una fuente debería ser una función decreciente de la probabilidad de ocurrencia. Una segunda propiedad intuitiva de medida de información es que un cambio pequeño en la probabilidad de cierta salida no debería cambiar ostensiblemente la información asociada a esa salida. En otras palabras, la medida de información debería ser una función continua y decreciente de la probabilidad de la salida de la fuente.

Supongamos ahora que la información asociada a la salida a_j puede separarse en dos partes independientes, a_{j1} y a_{j2} , o sea, $X_j = \{X_{(j1)}, X_{(j2)}\}$, $a_j = \{a_{j1}, a_{j2}\}$ y $p(X = a_j) = p(X_{(j1)} = a_{j1})p(X_{(j2)} = a_{j2})$. Esto puede suceder, por ejemplo, si suponemos que la temperatura y la contaminación son independientes y, en consecuencia, cada salida de la fuente puede separarse en dos componentes independientes. Dado que las componentes son independientes, saber información sobre una de las componentes (la temperatura) no provee información sobre la otra componente (contaminación), de donde es posible concluir intuitivamente que la cantidad de información obtenida sabiendo a_j es la suma de la información obtenida de a_{j1} y a_{j2} . De esta discusión es posible concluir que la cantidad de información obtenida de una salida a_j con probabilidad de ocurrencia p_j debe satisfacer las cuatro condiciones siguientes:

1. El contenido de información de la salida a_j depende solamente de la probabilidad de a_j y no del valor de a_j . Esta función se notará con $I(p_j)$ y la llamaremos *información*.
2. La información es una función continua de p_j , o sea, $I(\cdot)$ es una función continua.
3. La información es una función decreciente de su argumento.
4. Si $p_j = p_{(j1)}p_{(j2)}$, entonces $I(p_j) = I(p_{(j1)}) + I(p_{(j2)})$.

Puede probarse que la única función que satisface todas esas propiedades es la función logaritmo, o sea $I(x) = \log(1/x) = -\log(x)$. La base del logaritmo no es importante, aunque define la unidad mediante la cual se mide información. Si la base es 2, la información se expresa en *bits*, y si es el logaritmo natural, la unidad se expresa en *nats*.

Dado que la información obtenida de cada salida de la fuente a_i se define como la información de esa salida, dada por $-\log(p_i)$, es posible definir el contenido de información de la fuente como el promedio ponderado de información de todas sus salidas. La ponderación se justifica porque las diferentes salidas de la fuente pueden tener probabilidades diferentes. En consecuencia, la información asociada a una fuente es el promedio ponderado de la información, o sea $\sum_{i=1}^N p_i I(p_i) = -\sum_{i=1}^N p_i \log p_i$. El contenido de información de la fuente se denomina *entropía*, $H(X)$.

Definición: La entropía de una variable aleatoria discreta X es una función de las probabilidades de ocurrencia de sus muestras, definida como

$$H(X) = \sum_{i=1}^N p_i \log \frac{1}{p_i} = - \sum_{i=1}^N p_i \log p_i$$

Notar que existe algún abuso de notación aquí ya que podríamos esperar que siendo $H(X)$ una función de una variable aleatoria X , sea en consecuencia una variable aleatoria. Sin embargo, $H(X)$ es una función de distribución de masa de probabilidades (pmf) de X , de forma que resulta en un número.

Ejemplo: Para la fuente discreta sin memoria (fds) con probabilidades p y $1 - p$, respectivamente, tenemos que

$$H(X) = -p \log p - (1 - p) \log(1 - p)$$

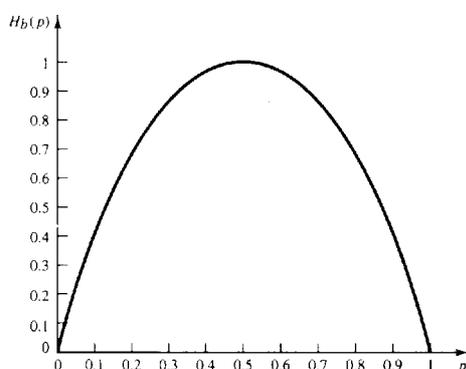


Figura 4.3: La función entropía binaria.

Esta función, notada por $H_b(p)$, se conoce como *función entropía binaria*. La misma se ilustra en la figura 4.3.

Ejemplo: Una fuente con un ancho de banda de 4000 Hz se muestra a la velocidad de Nyquist. Suponiendo que la secuencia resultante puede modelarse aproximadamente por una fds con alfabeto $\mathcal{A} = \{-2, -1, 0, 1, 2\}$ y con probabilidades $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\}$, determinar la velocidad de generación de esa fuente en bits/seg.

Como se tiene que

$$H(X) = \frac{1}{2} \log 2 + \frac{1}{4} \log 4 + \frac{1}{8} \log 8 + 2 \times \frac{1}{16} \log 16 = \frac{15}{8} \text{ bits/muestra}$$

y dado que se tienen 8000 muestras/seg., la fuente produce información a una velocidad de 15 kbits/seg.

En el Apéndice se introduce una discusión adicional de entropía, sus propiedades e interpretación.

4.1.2 Entropía conjunta y condicional

Cuando trabajamos con dos o más variables aleatorias, y en forma similar a lo analizado para probabilidades conjuntas y condicionales, es posible introducir entropía conjunta y entropía condicional. Esos conceptos serán especialmente importantes cuando trabajemos con fuentes con memoria.

Definición: La *entropía conjunta* de dos variables aleatorias discretas $\{X, Y\}$ se define por

$$H(X, Y) = - \sum_{x, y} p(x, y) \log p(x, y)$$

Para el caso de n variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$, se tiene que

$$H(\mathbf{X}) = - \sum_{x_1, x_2, \dots, x_n} p(x_1, x_2, \dots, x_n) \log p(x_1, x_2, \dots, x_n)$$

de donde es posible concluir que la entropía conjunta es simplemente la entropía de una variable aleatoria vectorial.

La entropía condicional de una variable aleatoria X , dada la variable aleatoria Y , puede definirse teniendo en cuenta que si $Y = y$, entonces la pmf de la variable aleatoria X será $p(x|y)$, y la entropía correspondiente será $H(X|Y = y) = - \sum_x p(x|y) \log p(x|y)$, que intuitivamente es la cantidad de incertidumbre en X cuando se conoce $Y = y$. El promedio ponderado de esa cantidad sobre todo y es la incertidumbre en X conocida Y . Esta cantidad se conoce como entropía condicional y se define como sigue:

Definición: La *entropía condicional* de una variable aleatoria X dada la variable aleatoria Y está definida por

$$H(X|Y) = - \sum_{x, y} p(x, y) \log p(x|y)$$

En general se tiene que

$$H(X_n | X_1, \dots, X_{n-1}) = - \sum_{x_1, x_2, \dots, x_n} p(x_1, \dots, x_n) \log p(x_n | x_1, x_2, \dots, x_{n-1})$$

Ejemplo: Usando la regla de la cadena para pmf, $p(x, y) = p(y)p(x|y)$, mostrar que $H(X, Y) = H(Y) + H(X|Y)$. Generalizar este resultado para el caso de n variables aleatorias para obtener

$$H(\mathbf{X}) = H(X_1) + H(X_2|X_1) + \dots + H(X_n|X_1, X_2, \dots, X_{n-1})$$

De la definición de entropía conjunta de dos variables aleatorias se tiene que

$$\begin{aligned} H(X, Y) &= - \sum_{x, y} p(x, y) \log p(x, y) = - \sum_{x, y} p(x, y) \log [p(y)p(x|y)] \\ &= - \sum_{x, y} p(x, y) \log p(y) - \sum_{x, y} p(x, y) \log p(x|y) \\ &= - \sum_y p(y) \log p(y) - \sum_{x, y} p(x, y) \log p(x|y) = H(Y) + H(X|Y) \end{aligned}$$

donde en el último paso se usó que $\sum_x p(x, y) = p(y)$. Esta relación dice que el contenido de información del par (X, Y) es igual al contenido de información de Y más el contenido de información de X después que Y es conocida. Equivalentemente, esto nos dice que se transfiere la misma información conociendo el par (X, Y) , ó primero conociendo Y y luego conociendo la información remanente en X . La prueba para el caso general puede obtenerse por inducción. En este caso, si las n variables son independientes, la relación se reduce a

$$H(\mathbf{X}) = \sum_{i=1}^n H(X_i)$$

Si la variable aleatoria X_n es la salida de una fuente discreta (no necesariamente sin memoria) en el instante n , entonces $H(X_2|X_1)$ es la información nueva provista por la salida de la fuente X_2 cuando ya se conoce X_1 . De la misma forma $H(X_n|X_1, X_2, \dots, X_{n-1})$ es la información nueva en X_n para un observador de la secuencia $(X_1, X_2, \dots, X_{n-1})$. El límite de esta entropía condicional cuando n tiende a infinito se conoce como *taza de entropía* del proceso estocástico.

Definición: La *taza de entropía* de un proceso estocástico discreto se define por

$$H = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n)$$

La tasa de entropía juega el mismo papel que la entropía pero en este caso para fuentes de información con memoria. Es básicamente una medida de la incertidumbre por símbolo de salida de la fuente.

4.2 Teorema de Codificación de Fuente

El teorema de codificación de fuente es uno de los tres teoremas fundamentales de teoría de información introducidos por Shannon. Este teorema establece un límite fundamental sobre la forma en que puede representarse la salida de una fuente de información sin originar una probabilidad de error grande. Se vio que la entropía de una fuente de información es una medida de la incertidumbre o, equivalentemente, el contenido de información de la fuente. En consecuencia, es natural que en la proposición del teorema de codificación de fuente la entropía juegue un papel principal.

La entropía tiene un significado muy intuitivo. Consideremos la observación de las salidas de longitud n de una fds, para n muy grande. En base a la ley de los grandes números (del Capítulo 2) en esta secuencia y con alta probabilidad (o sea, tendiendo a 1 para $n \rightarrow \infty$) el símbolo a_1 se repite aproximadamente np_1 veces, el símbolo a_2 se repite aproximadamente np_2 veces, \dots , y el símbolo a_N se repite aproximadamente np_N veces. Esto significa que para n suficientemente grande, con probabilidad próxima a 1, cada secuencia de la fuente tiene la misma composición y en consecuencia casi la misma probabilidad. Dicho de otra forma, asintóticamente *casi todo es casi igualmente probable*. Las secuencias de símbolos a_i , que notaremos \mathbf{x} , con esa estructura se denominan *secuencias típicas*. La probabilidad de una secuencia típica está dada por

$$\begin{aligned} p(\mathbf{X} = \mathbf{x}) &\cong \prod_{i=1}^N p_i^{np_i} \quad (\text{cada símbolo se supone independiente con probab. } np_i) \\ &= \prod_{i=1}^N 2^{np_i \log p_i} = 2^{n \sum_{i=1}^N p_i \log p_i} = 2^{-nH(X)} \end{aligned}$$

Esto significa que para n suficientemente grande casi todas las secuencias típicas de salida de longitud n son igualmente probables con probabilidad $\cong 2^{-nH(X)}$. Por otro lado, la probabilidad del conjunto de secuencias no típicas es despreciable.

Pero como la probabilidad de ocurrencia de secuencias típicas es aproximadamente 1 y cada secuencia típica tiene probabilidad $2^{-nH(X)}$, el número de secuencias típicas será

$$\text{Nro. sec. típicas} = \frac{1}{2^{-nH(X)}} = 2^{nH(X)}$$

En consecuencia, a pesar que una secuencia de tamaño de alfabeto N puede producir N^n secuencias de longitud n , el número *efectivo* (o sea sin perder una cantidad significativa) de secuencias es $2^{nH(X)}$. La figura 4.4 muestra un diagrama de la propiedad mencionada. Este es un resultado muy importante ya que nos dice que para todo propósito práctico, es suficiente considerar el conjunto de secuencias típicas antes que todo el conjunto de secuencias posibles de la fuente. El error introducido ignorando secuencias no típicas puede hacerse menor que cualquier $\epsilon > 0$ eligiendo n suficientemente grande. Esta es la esencia de la *compresión de datos*, o sea, la práctica de representar la salida de una fuente con el menor número de secuencias comparado con el número total de salidas que la fuente produce.

Del resultado anterior y dado que ϵ es un número positivo arbitrario, puede verse que es posible representar la secuencia típica sin introducir error considerable. Dado que el número de ellas es muy aproximadamente

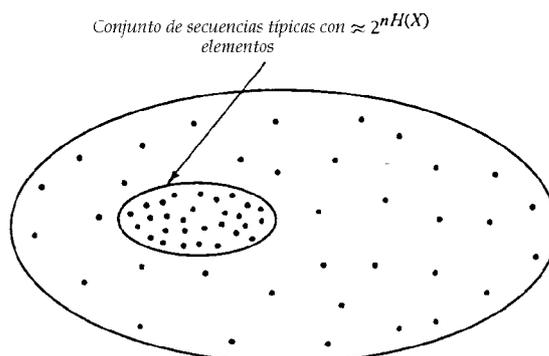


Figura 4.4: Conjunto de secuencias típicas y no típicas.

$2^{nH(X)}$, son necesarios $nH(X)$ bits para representarlas. Sin embargo, esos bits se utilizan para representar secuencias de longitud n . En consecuencia, en promedio, cualquier fuente requiere $H(X)$ bits para una representación casi sin errores. Esto nuevamente justifica la noción de entropía como la cantidad de información por cada secuencia de la fuente.

Hasta aquí se supuso que la fuente es discreta y sin memoria y puede representarse por una variable aleatoria i.i.d. Esa fuente puede comprimirse solo si su pmf no es uniforme, ya que si X tiene distribución uniforme $H(X) = \log N$. Y, en consecuencia, $2^{nH(X)} = 2^{n \log N} = N^n$. Esto significa que el número efectivo de secuencias de salida de la fuente de longitud n es igual al número total de secuencias de la fuente y no existe compresión posible.

No hemos considerado el caso en que la fuente tenga memoria. Para fuentes con memoria, las secuencias no son independientes y las secuencias pasadas introducen información en las futuras. Esto significa que la velocidad a la cual se introduce información nueva disminuye cuando aumenta el número de secuencias. Un ejemplo clásico de esto es el alfabeto Castellano, el cual muestra considerable dependencia entre letras y palabras (una "q" está casi siempre seguida de una "u", etc). La entropía por letra para un texto grande en Castellano es rigurosamente el límite de $H(X_n|X_1, X_2, \dots, X_{n-1})$ para n suficientemente grande (la tasa de entropía discutida anteriormente). Para fuentes estacionarias en general, la tasa de entropía tiene el mismo significado que la entropía para el caso de fds y define el número de secuencias de salida de la fuente para n suficientemente grande, o sea, 2^{nH} , donde H es la tasa de entropía.

Varios estudios con modelos estadísticos del Castellano muestran que la tasa de entropía converge rápidamente, y para $n = 10$ se está muy próximo al límite. Cuando la memoria aumenta, el tamaño del espacio sobre el cual se calculan las probabilidades condicionales aumenta rápidamente, y no es fácil hallar las probabilidades condicionales para obtener la tasa de entropía.

Hasta aquí se discutió una descripción informal del teorema de codificación de fuente. La proposición formal, sin prueba, se presenta a continuación.

Teorema de Codificación de Fuente: Una fuente con tasa de entropía H puede codificarse con un error arbitrariamente pequeño a cualquier velocidad R (en bits/salida) siempre que $R > H$. Por otro lado, si $R < H$, la probabilidad de error no podrá ser arbitrariamente pequeña, independientemente de la complejidad del codificador y decodificador empleados.

Este teorema, solo determina condiciones necesarias y suficientes para la existencia de tal codificación de fuente. No provee ningún algoritmo para diseñar los códigos que logren el desempeño predicho. En la subsección próxima se discutirán dos algoritmos para la compresión de fuentes de información.

4.3 Algoritmos de Codificación de Fuente

En la sección anterior se concluyó que la entropía de la fuente H determina una cota sobre la velocidad a la cual puede comprimirse una fuente para una reconstrucción confiable. Esto significa que a velocidades por encima de la entropía es posible diseñar un código con una probabilidad de error tan pequeña como se desee, mientras que a velocidades por debajo de la entropía tal código no existe. Este resultado importante, sin embargo, no provee algoritmos específicos para diseñar códigos que aproximen esta cota. En esta sección

introduciremos dos algoritmos para diseñar códigos que se desempeñan muy proxicamente a la cota de entropía.

4.3.1 Algoritmo de codificación de fuente de Huffman

En el codificador de Huffman, bloques de longitud fija de salida de la fuente se mapean en bloques binarios de longitud variable. Esto se denomina un codificador *de longitud fija a variable*. La idea es mapear las secuencias de longitud fija que ocurren más frecuentemente a secuencias binarias más cortas y las menos frecuentes a secuencias binarias más largas. En la codificación binaria de longitud variable el problema es la sincronización. Esto significa que debería existir una única forma de dividir la secuencia binaria recibida en palabras de código. El siguiente ejemplo clarifica este punto.

Ejemplo: Supongamos que las salidas posibles de una fuente de información son $\{a_1, a_2, a_3, a_4, a_5\}$, y consideremos los siguientes códigos de esta fuente.

Letra	Probabilidad	Código 1	Código 2	Código 3	Código 4
a_1	$p_1 = \frac{1}{2}$	1	1	0	00
a_2	$p_2 = \frac{1}{4}$	01	10	10	01
a_3	$p_3 = \frac{1}{8}$	001	100	110	10
a_4	$p_4 = \frac{1}{16}$	0001	1000	1110	11
a_5	$p_5 = \frac{1}{16}$	00001	10000	1111	110

En el primer código cada palabra termina con un 1. En consecuencia, tan pronto como el decodificador observa un 1, reconoce que la palabra de código ha terminado y una nueva comenzará. Esto significa que el código es *autosincronizante*. En el segundo código cada palabra comienza con un 1. Por eso, observando un 1, el decodificador reconoce que una palabra de código nueva y en consecuencia, el bit previo era el último de la palabra anterior. Este código es también autosincronizante pero no es tan conveniente como el primero porque con este es necesario esperar hasta recibir el primer bit de la palabra siguiente para reconocer que una nueva palabra ha comenzado. En el código 1 se reconoce el último bit sin necesidad de recibir el primer bit de la próxima palabra de código. Los códigos 1 y 2 son de esa manera *univocamente decodificables*. Sin embargo, solo el código 1 es *instantáneo*. Los códigos 1 y 3 tienen la propiedad interesante que ninguna palabra es prefijo de otra, se dice que satisfacen la *condición de prefijo*.

Dada una secuencia (alfabeto) $\{a_0, a_1, \dots, a_{M-1}\}$ generada por una fuente discreta sin memoria, con $\{p_0, p_1, \dots, p_{N-1}\}$, definimos como **condición de prefijo** a las siguientes reglas:

- Sea $\{m_{k_1}, m_{k_2}, \dots, m_{k_N}\}$ la k -ésima palabra asignada al símbolo a_k , con $m_k = 0$ ó 1 , N es la longitud de la palabra.
- La parte inicial de la palabra se representa por $m_{k_1} \dots m_{k_i}$, $i \leq N$.
- Cualquier secuencia construida de la parte inicial se denomina **prefijo**.
- El **código de prefijo** se define como aquel en que la palabra no es prefijo de ninguna otra.

Puede demostrarse que una condición necesaria y suficiente para que un código sea univocamente decodificable e instantáneo es que satisfaga la condición de prefijo. Esto significa que esos códigos son univocamente decodificables e instantáneos. Sin embargo, el código 3 tiene la ventaja de tener menor longitud promedio. En realidad, para el código 1 la longitud promedio es

$$E[L] = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + 5 \times \frac{1}{16} = \frac{31}{16}$$

y para el código 3 se tiene que

$$E[L] = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + 4 \times \frac{1}{16} = \frac{30}{16}$$

El código 4 tiene como mayor desventaja no ser unívocamente decodificable. Por ejemplo, la secuencia 110110 puede decodificarse de dos formas distintas, como a_5a_5 o como $a_4a_2a_3$. Los códigos que no son unívocamente decodificables no son deseables y deberían evitarse en la práctica. De la discusión anterior es posible concluir que el más deseable de los códigos anteriores es el 3. Este es un ejemplo de un *código de Huffman*, ha ser discutido brevemente a continuación.

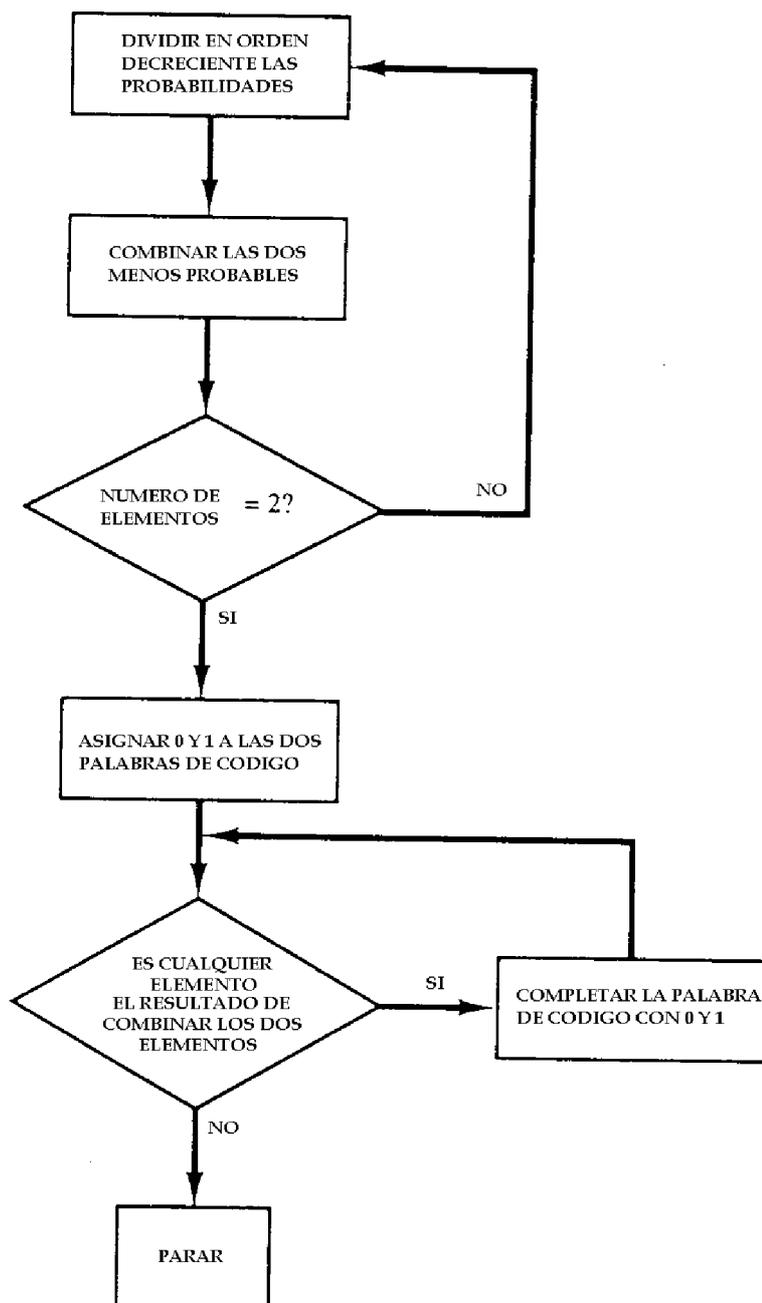


Figura 4.5: Algoritmo de codificación de Huffman.

Como mencionado anteriormente, la idea en la codificación de Huffman es elegir una longitud de palabra tal que las secuencias más probables tengan longitudes más cortas. Si podemos mapear cada salida de la fuente de probabilidad p_i en una palabra de código de longitud aproximadamente $\log \frac{1}{p_i}$ y al mismo tiempo asegurar decodificabilidad, es posible lograr una longitud promedio de palabra de aproximadamente $H(X)$. Los códigos de Huffman son unívocamente decodificables e instantáneos con longitud promedio de palabra de código mínima. En ese sentido son óptimos. La optimalidad se entiende sobre todos los códigos que satisfacen la condición de prefijo (y en consecuencia son unívocamente decodificables e instantáneos).

Algoritmo de codificación de Huffman

1. Clasificar las salidas de la fuente en orden decreciente de acuerdo a sus probabilidades.
2. Mezclar las dos salidas menos probables en una única salida cuya probabilidad es la suma de las probabilidades correspondientes.
3. Si el número de salidas remanentes es 2, entonces ir al próximo paso; de lo contrario ir al paso 1.
4. Asignar arbitrariamente 0 y 1 como palabras de código a las dos salidas remanentes.
5. Si una salida es el resultado de mezclar las dos salidas en un paso precedente, completar la palabra actual con un 0 o un 1 para obtener la palabra de código para las salidas precedentes y repetir el paso 5. Si no existe salida precedida por otra salida en el paso precedente, parar.

La figura 4.5 muestra un diagrama de flujo de este algoritmo.

Ejemplo: Diseñar un código de Huffman para la fuente dada en el ejemplo anterior. El diagrama en árbol de la figura 4.6 resume los pasos de diseño para la construcción del código y las palabras del código resultante.

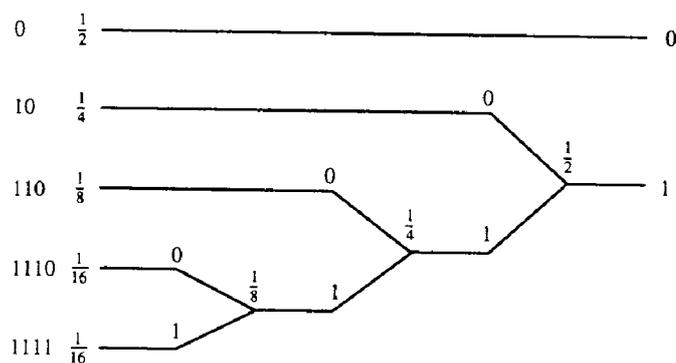


Figura 4.6: Diagrama en árbol de la codificación Huffman.

Puede mostrarse que la longitud promedio de un código de Huffman, definida por

$$\bar{R} = \sum_{x \in X} p(x)l(x)$$

donde $l(x)$ es la longitud de la palabra de código correspondiente a la salida de la fuente x , satisface la desigualdad

$$H(X) \leq \bar{R} < H(X) + 1$$

Entonces, si en lugar de letras aisladas, el código de Huffman fué diseñado para secuencias de letras fuente de longitud n (n -ésima extensión de la fuente, ver Apéndice) se verifica que

$$H(\mathbf{X}^n) \leq \bar{R}_n < H(\mathbf{X}^n) + 1$$

donde \bar{R}_n es la longitud promedio de palabra para una secuencia de fuente extendida y, en consecuencia, $\bar{R} = \frac{1}{n}\bar{R}_n$. En el caso en que la fuente con la que se está trabajando es sin memoria, se tiene que $H(\mathbf{X}^n) = nH(X)$, de forma que sustituyendo se tiene que

$$H(X) \leq \bar{R} < H(X) + \frac{1}{n}$$

En consecuencia, para n suficientemente grande \bar{R} puede hacerse tan próxima a $H(X)$ como deseado. Es obvio también que un comportamiento similar es esperable en el caso de fuentes discretas con memoria aproximando ahora a la tasa de entropía.

Ejemplo: Una fds con salidas equiprobables y alfabeto $\mathcal{A} = \{a_1, a_2, a_3\}$ tiene el código de Huffman diseñado en la figura 4.7.

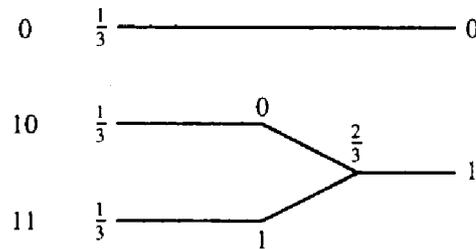


Figura 4.7: Codificación de Huffman.

La entropía de la fuente es $H(X) = \log 3 = 1.585$ y $\bar{R} = \frac{5}{3} = 1.667$. Si se usa secuencias de dos letras, se tendrá una fuente

$$\mathcal{A}^2 = \{(a_1, a_1), (a_1, a_2), (a_1, a_3), \dots, (a_3, a_2), (a_3, a_3)\}$$

con probabilidades $\{\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}\}$. Un código Huffman para esta fuente se muestra en el diagrama de la figura 4.8.

En nuestro tratamiento, solo se consideraron códigos de Huffman binarios. El algoritmo para diseñar códigos de Huffman binarios puede generalizarse fácilmente a un algoritmo para el diseño de códigos de Huffman multinivel.

4.3.2 Algoritmo de codificación de fuente de Lempel-Ziv

Como se discutió, los códigos de Huffman son óptimos en el sentido que para una fuente determinada proveen un código de prefijo con una longitud de bloque promedio mínima. Independientemente de esto, existen problemas importantes en la implementación de un codificador de Huffman. Uno de esos problemas es que el código de Huffman depende fuertemente de la distribución de probabilidades de la fuente. La estadística de la fuente debe ser conocida a priori para diseñar el codificador. Si podemos observar solo las salidas de la fuente, entonces será necesario un proceso de codificación en dos pasos. En el primer paso,

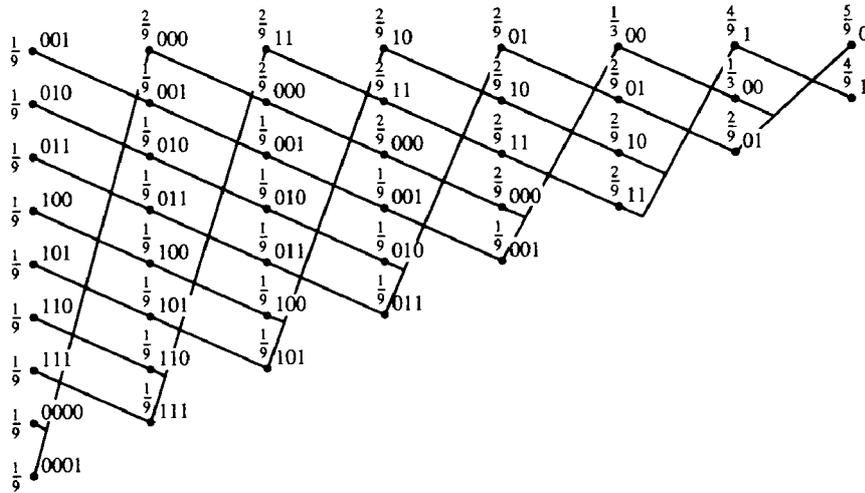


Figura 4.8: Diagrama en árbol de la codificación de Huffman.

estimamos la estadística de la fuente (la cual puede ser lenta en casos donde la fuente tiene memoria o se trata de aplicar el codificador a una extensión de la fuente). En el segundo paso se realiza la codificación. Otro problema con la codificación de Huffman es que si el código se diseña para bloques fuente de longitud 1, el código utilizará variaciones en la frecuencia de las salidas de la fuente y no la memoria de la fuente. Si se desea utilizar también la memoria de la fuente, es necesario diseñar la longitud de los bloques mayor o igual a 2 y esto hace crecer exponencialmente la complejidad del algoritmo. Por ejemplo, codificando caracteres ASCII con longitud de bloque 1 requiere un árbol con 256 nodos terminales, pero si la longitud del bloque es 2, el tamaño del árbol y en consecuencia la complejidad de la codificación será mucho mayor. En ciertas aplicaciones, como en el almacenamiento magnético o los medios ópticos donde se desea altas velocidades de transferencia de información, la complejidad y velocidad de la codificación de Huffman es un cuello de botella.

El algoritmo de Lempel-Ziv pertenece a la clase de *algoritmos universales de codificación de fuente*, o sea algoritmos que son independientes de la estadística de la fuente. Este algoritmo es un esquema de codificación de una *longitud variable* a una *longitud fija*. Esto significa que cualquier secuencia de salida de la fuente se divide unívocamente en partes de longitud variable, y esas partes se codifican usando palabras de código de igual longitud. La división se realiza identificando partes de la menor longitud posible que no existan anteriormente en la secuencia. Para este propósito, el divisor observa la salida de la fuente y siempre que la nueva secuencia de salida de la fuente coincide con alguna de las partes existentes no se agrega una parte nueva y se incluye una nueva letra a la secuencia. En cuanto la nueva salida de la fuente es diferente de las partes disponibles, esta se reconoce como una nueva parte y se codifica. El esquema de codificación es simple. La nueva parte es la concatenación de una parte previa y la nueva salida de la fuente. Para codificarla, *se concatenan la expansión binaria del orden lexicográfico de la parte previa y el bit nuevo*. Por ejemplo, supongamos que se desea codificar la siguiente secuencia:

0100001100001010000010100000110000010100001001001

Dividiendo la secuencia mediante las reglas anteriores se obtienen las siguientes partes

0, 1, 00, 001, 10, 000, 101, 0000, 01, 010, 00001, 100, 0001, 0100, 0010, 01001

Es posible verificar que todas las partes son diferentes y que cada parte es una parte previa concatenada con la nueva salida de la fuente. El número de partes es 16. Esto significa que para cada parte serán necesarios 4 bits, más un bit extra para representar la nueva salida de la fuente. La secuencia anterior se puede codificar como

0000 0, 0000 1, 0001 0, 0011 1, 0010 0, 0011 0, 0101 1, 0110 0,

0001 1, 1001 1, 1000 1, 0101 0, 0110 1, 1010 0, 0100 0, 1110 1

La tabla siguiente resume los resultados del procedimiento completo.

Ubicación en el diccionario	Contenido del diccionario	Código
1	0001	0 0000 0
2	0010	1 0000 1
3	0011	00 0001 0
4	0100	001 0011 1
5	0101	10 0010 0
6	0110	000 0011 0
7	0111	101 0101 1
8	1000	0000 0110 0
9	1001	01 0001 1
10	1010	010 1001 0
11	1011	00001 1000 1
12	1100	100 0101 0
13	1101	0001 0110 1
14	1110	0010 1010 0
15	1111	0010 0100 0
16		1110 1

Esta representación difícilmente parece interpretarse como un esquema de compresión de datos debido a que una secuencia de longitud 49 se ha mapeado en una secuencia de longitud 80. Pero a medida que la longitud de la secuencia original se incrementa el propósito de compresión de este algoritmo se vuelve evidente. Puede demostrarse que para una fuente estacionaria y ergódica y cuando la longitud de la secuencia aumenta, el número de bits en la secuencia comprimida se aproxima a $nH(X)$, donde $H(X)$ es la tasa de entropía de la fuente. La descompresión de la secuencia codificada es directa y puede realizarse muy fácilmente.

Un problema con el algoritmo de Lempel-Ziv es elegir el número de partes. En el ejemplo se tienen 16 partes y en consecuencia se requieren 4 bits para representar cada parte. En general, cualquier número fijo de partes resultará tarde o temprano pequeño y causará el subdimensionamiento del diccionario. Por ejemplo, si fuese necesario continuar la codificación anterior para letras adicionales, no sería posible incluir nuevas partes en el diccionario porque se asignaron 4 bits para la representación de los elementos de ese diccionario y ya se tiene el diccionario completo. Para resolver este problema, tanto el codificador como el decodificador deben purgar de sus diccionarios aquellos elementos que no son ya útiles y se los debe substituir por nuevos. El algoritmo Lempel-Ziv se utiliza ampliamente en la práctica para comprimir archivos de computadora. Algunas implementaciones de estos algoritmos son los utilitarios "compress" y "uncompress" del sistema operativo UNIX y numerosos algoritmos bajo el sistema operativo MS-DOS (ZIP, ZOO, LZH, ARJ, etc.).

4.4 Teoría de Velocidad - Distorsión

En secciones anteriores se vió que es posible codificar la salida de una fuente de información sin memoria a una velocidad arbitrariamente próxima a su entropía y aún así mantener la confiabilidad en la salida. La reconstrucción confiable de la fuente significa que la probabilidad de error puede hacerse próxima a cero cuando la longitud del bloque tiene a infinito. En muchos casos, sin embargo, transmitir a velocidades próximas a la entropía no es posible. Por ejemplo, si se tiene una capacidad de almacenamiento limitado y la entropía de la fuente de información es mayor que la capacidad de almacenamiento, es imposible recuperar la salida de la fuente de los datos almacenados libre de errores. En casos como este, alguna técnica de compresión *con pérdidas* deberá utilizarse y se introducirá alguna *distorsión*.

Un ejemplo de este último caso es la codificación de la salida de una fuente analógica. Las muestras de una fuente analógica son números reales, y para representarlo se requiere un número infinito de bits. En consecuencia, en un sistema digital práctico no existe forma de transmitir o almacenar datos analógicos sin pérdida de información. Para codificar muestras analógicas, las muestras deben cuantizarse, y los valores cuantizados se codifican luego utilizando alguno de los algoritmos de codificación de fuente.

Se discutirán a continuación metodologías de compresión de datos sujetas a un criterio de fidelidad. Se introducirán primero algunas cantidades necesarias para el análisis de este caso y se definirá luego la función velocidad-distorsión.

4.4.1 Información Mutua

La entropía de una variable aleatoria discreta X conocida la variable aleatoria Y se notará por $H(X|Y)$. Entonces, si la entropía inicial de la variable aleatoria X es $H(X)$, luego $H(X) - H(X|Y)$ es la incertidumbre remanente sobre X después que se ha removido la variable Y . En otras palabras, $H(X) - H(X|Y)$ es la cantidad de información provista por la variable Y en relación a X . Esta cantidad juega un papel importante tanto en la codificación de fuente como en la codificación de canal, que se discutirá en el capítulo 8, y se denomina *información mutua* entre dos variables aleatorias.

Definición: La *información mutua* entre dos variables aleatorias discretas X e Y , $I(X; Y)$ se define por

$$I(X; Y) = H(X) - H(X|Y)$$

Ejemplo: Sean X e Y dos variables aleatorias binarias con $p(X = 0, Y = 0) = \frac{1}{3}$, $p(X = 1, Y = 0) = \frac{1}{3}$ y $p(X = 0, Y = 1) = \frac{1}{3}$. Hallar $I(X; Y)$ en este caso.

De las probabilidades condicionales $p(X = 0) = p(Y = 0) = \frac{2}{3}$ y en consecuencia $H(X) = H(Y) = H_b(\frac{2}{3}) = 0.919$. Por otro lado, el par (X, Y) puede considerarse un vector aleatorio uniformemente distribuido con tres valores $(0, 0)$, $(1, 0)$ y $(0, 1)$. Entonces, $H(X, Y) = \log 3 = 1.585$. Así se tiene que $H(X|Y) = H(X, Y) - H(Y) = 1.585 - 0.919 = 0.666$ y finalmente $I(X; Y) = H(X) - H(X|Y) = 0.919 - 0.666 = 0.253$.

La información mutua tiene ciertas propiedades importantes a tener en cuenta que se resumirán a continuación:

1. $I(X; Y) \geq 0$ con la igualdad si y solo si X e Y son independientes.
2. $I(X; Y) \leq \min(H(X), H(Y))$.
3. $I(X; Y) = \sum_{x,y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}$.
4. $I(X; Y) = H(X) + H(Y) - H(X, Y)$.
5. $I(X; Y|Z)$ es la información mutua condicional, definida por $I(X; Y|Z) = H(X|Z) - H(X|Y, Z)$.
6. $I(X; Y|Z) = \sum_z p(z)I(X; Y|Z = z)$.
7. $I(XY; Z) = I(X; Z) + I(Y; Z|X)$. Esta es la regla de la cadena para la información mutua.
8. En general, $I(X_1, \dots, X_n; Y) = I(X_1; Y) + I(X_2; Y|X_1) + \dots + I(X_n; Y|X_1, \dots, X_{n-1})$.

La figura 4.9 representa la relación entre entropía, entropía condicional e información mutua.

4.4.2 Entropía diferencial

Hasta aquí se definió entropía e información mutua para fuentes discretas. Si se trabaja con una fuente discreta en tiempo de alfabeto continuo cuyas salidas son números reales, no existe una magnitud que tenga el significado intuitivo de la entropía. En el caso continuo se utiliza en general otra cantidad que se asemeja a la entropía y que se denomina *entropía diferencial*. En realidad, para reconstruir la salida de una fuente continua en forma confiable, se requerirá un número infinito de bits por salida de la fuente dado que cualquier salida de la fuente es un número real y la expansión binaria de un número real requiere infinita cantidad de bits.

Definición: La *entropía diferencial* de una variable aleatoria continua X con pdf $f_X(x)$ se notará por $h(X)$ y está definida por

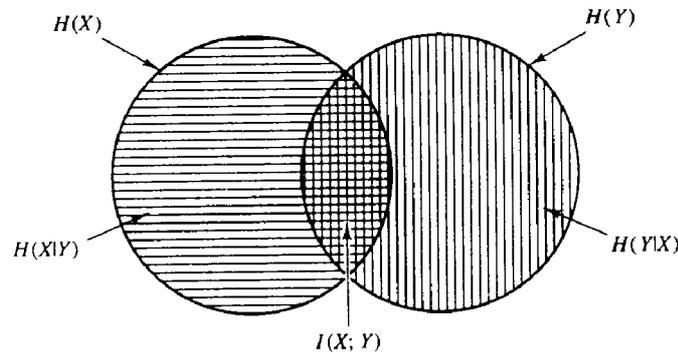


Figura 4.9: Entropía, entropía condicional e información mutua.

$$h(X) = - \int_{-\infty}^{\infty} f_X(x) \log f_X(x) dx$$

Ejemplo Determinar la entropía diferencial de una variable aleatoria X uniformemente distribuida en $[0, a]$. Utilizando la definición,

$$h(X) = - \int_0^a f_X(x) \log \frac{1}{a} dx = \log a$$

Puede concluirse que para $a < 1$ se tiene que $h(X) < 0$, en contraste con la no negatividad de la entropía discreta. También, para $a = 1$, $h(X) = 0$ sin ser X determinística. Esto también está en contraste con las propiedades de la entropía discreta.

Ejemplo: Determinar la entropía diferencial de una variable aleatoria Gaussiana con media cero y varianza σ^2 .

La pdf es $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$. En consecuencia, usando logaritmo natural, la entropía diferencial será

$$\begin{aligned} h(X) &= - \int_{-\infty}^{\infty} \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) f(x) dx - \int_{-\infty}^{\infty} \ln \left(e^{-\frac{x^2}{2\sigma^2}} \right) f(x) dx \\ &= \ln(\sqrt{2\pi\sigma^2}) + \frac{\sigma^2}{2\sigma^2} = \frac{1}{2} \ln(2\pi\sigma^2) \end{aligned}$$

donde se utilizó que $\int_{-\infty}^{\infty} f(x) dx = 1$ y $\int_{-\infty}^{\infty} x^2 f(x) dx = \sigma^2$. Cambiando el logaritmo a base 2 se tiene que

$$h(X) = \frac{1}{2} \log(2\pi e \sigma^2)$$

La extensión de la definición de entropía diferencial a variables aleatorias conjuntas y entropía diferencial condicional es directa. Para dos variables aleatorias se tiene que

$$h(X, Y) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \log f(x, y) dx dy$$

y

$$h(X|Y) = h(X, Y) - h(Y)$$

La información mutua entre dos variables aleatorias continuas conjuntas X e Y se define en forma similar al caso discreto como

$$I(X; Y) = h(Y) - h(Y|X) = h(X) - h(X|Y)$$

A pesar que la entropía diferencial no tiene la interpretación intuitiva de la entropía discreta, puede demostrarse que la información mutua de variables aleatorias continuas tiene básicamente la misma interpretación que la información mutua de variables aleatorias discretas, o sea, es la información provista por una variable aleatoria sobre otra variable aleatoria.

4.4.3 Función velocidad - distorsión

Volviendo al problema original de representar una fuente continua con un número finito de bits/símbolo, la pregunta es: si la salida de la fuente va a ser comprimida y representada por cierto número de bits/símbolo, cuán próximas pueden ser las versiones comprimida y original? Esta pregunta no es aplicable solo a fuentes continuas. Un problema similar puede formularse para fuentes discretas. Suponiendo una fuente discreta con entropía $H(X)$. Esto significa que es posible transmitir la salida de esta fuente usando $H(X)$ bits/símbolo de salida de la fuente con probabilidad de error arbitrariamente baja. Supongamos además que usar $H(X)$ bits/símbolo es muy alto y no es tolerable (por ejemplo, en almacenamiento en discos magnéticos, el espacio de disco requerido sería muy grande). Entonces el número de bits/símbolo disponibles es menor que $H(X)$, por lo que la recuperación libre de errores no es posible y algunos errores serán inevitables. La pregunta es: para un número dado de bits/símbolo, cuál es la mínima tasa de error alcanzable y como puede lograrse este mínimo? La pregunta también puede establecerse en la dirección opuesta, o sea, cuál es el número mínimo de bits/símbolo requerido para reproducir la fuente con cierto nivel de distorsión?

Para presentar formalmente la discusión anterior es necesario definir la noción de distorsión. Distorsión en la reproducción de una fuente es una medida de la fidelidad o proximidad de la reproducción en relación a la salida original de la fuente. En la reproducción de alta fidelidad, la señal reproducida está muy próxima a la señal original y la distorsión es baja, mientras que en la reproducción de baja fidelidad existe mayor "distancia" entre la señal original y la reproducción, y en consecuencia una mayor distorsión. Una *medida de distorsión* mide cuán lejos están la señal original y su reproducción. Es posible imaginar diferentes medidas de distorsión entre una señal $x(t)$ y su reproducción $\hat{x}(t)$, por ejemplo, $\max_t |x(t) - \hat{x}(t)|$, $\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t) - \hat{x}(t)| dt$ y $\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} (x(t) - \hat{x}(t))^2 dt$ son tres medidas de distorsión.

Una buena medida de distorsión debe satisfacer dos propiedades. Primero, debe ser una buena aproximación al proceso de percepción, y segundo, debe ser lo suficientemente simple como para ser matemáticamente tratable. La primera propiedad simplemente dice que, por ejemplo, si se está trabajando con transmisión de voz y en la percepción de la voz la fase de la señal no es un factor crucial, la medida de distorsión no debe ser dependiente de la reproducción exacta de la fase. Por otro lado, si se trabaja con percepción de imagen, la fase juega un rol importante y en consecuencia esto debe reflejarse en la elección de la medida de distorsión. Hallar una medida de distorsión que satisfaga ambas propiedades no es usualmente una tarea simple.

En general, una medida de distorsión es la distancia entre x y su reproducción \hat{x} , notada por $d(x, \hat{x})$. En el caso discreto una medida de distorsión utilizada comúnmente es la *distorsión de Hamming*, definida por

$$d_H(x, \hat{x}) = \begin{cases} 1, & x \neq \hat{x} \\ 0, & \forall x \end{cases}$$

En el caso continuo se utiliza frecuentemente la *distorsión de error cuadrático* definida por

$$d(x, \hat{x}) = (x - \hat{x})^2$$

Se supone también que trabajamos con una *medida de distorsión por símbolo*, lo que significa que la distorsión entre secuencias es el promedio de la distorsión entre sus componentes, o sea

$$d(\mathbf{x}, \hat{\mathbf{x}}^n) = \frac{1}{n} \sum_{i=1}^n d(x_i, \hat{x}_i)$$

Esta suposición significa que la posición del "error" en la reproducción no es importante y que la distorsión es independiente del contexto.

Como la salida de la fuente es un proceso estocástico, $d(\mathbf{x}, \hat{\mathbf{x}}^n)$ es una variable aleatoria. Definimos la distorsión de la fuente como el valor esperado de esta variable aleatoria,

$$D = E[d(\mathbf{X}, \hat{\mathbf{X}}^n)] = \frac{1}{n} \sum_{i=1}^n E[d(X_i, \hat{X}_i)] = E[d(X, \hat{X})]$$

donde en el último paso se utilizó la suposición de estacionaridad de la fuente (independencia de las distribuciones del índice i).

Con todas esas definiciones la pregunta original puede reescribirse como sigue: dada una fuente de información sin memoria con alfabeto \mathcal{X} y distribución de probabilidades $p(x)$, un alfabeto de reproducción $\hat{\mathcal{X}}$ y una medida de distorsión $d(x, \hat{x})$, cual es el número mínimo de bits/salida de la fuente R requerido para garantizar que la distorsión promedio entre la secuencia de salida y la correspondiente reproducción no exceda cierto D especificado? Es obvio que R es una función decreciente de D , o sea, si es necesaria alta fidelidad de reproducción (bajo D) es necesario que R sea alto. La relación entre R y D se expresa por la *función velocidad - distorsión*. El teorema siguiente introduce la forma general de la función velocidad-distorsión (sin demostración).

Teorema El número mínimo de bits/símbolo de la fuente requerido para reproducir una fuente sin memoria con distorsión menor o igual a D se denomina la *función velocidad-distorsión*, $R(D)$, que está definida por

$$R(D) = \min_{p(\hat{x}|x): E\{d(X, \hat{X})\} \leq D} I(X; \hat{X})$$

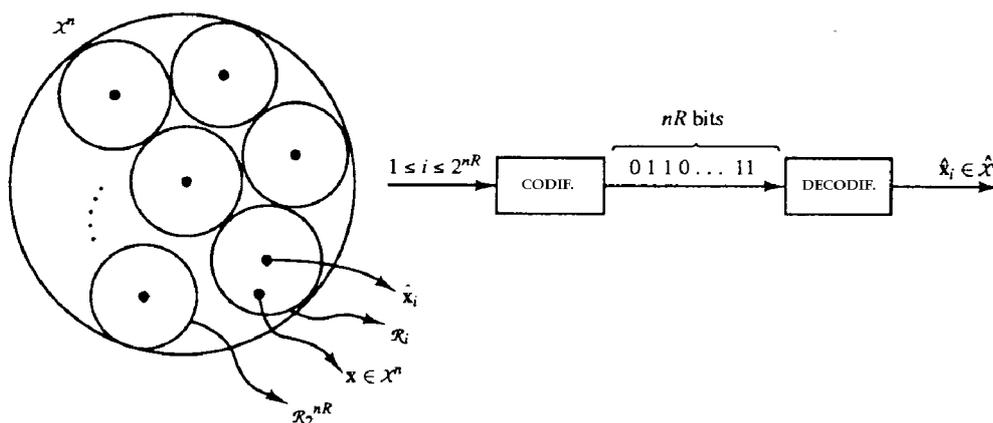


Figura 4.10: Representación esquemática del teorema de velocidad - distorsión.

La figura 4.10 es una representación esquemática de este teorema. El espacio de salidas de la fuente de longitud n , \mathcal{X}^n , se divide en 2^{nR} regiones. Si la salida de la fuente \mathbf{x}^n cae en la región i , se transmite por el decodificador la representación binaria de i . Dado que $1 < i \leq 2^{nR}$, la representación binaria es de longitud nR , y la codificación se realiza a una velocidad de R bits/salida de la fuente. El decodificador, después de recibir la representación binaria de i , genera una secuencia predeterminada $\hat{\mathbf{x}}^n$ tal que su distancia promedio (distorsión) de la secuencia \mathbf{x} en la región i es mínima. Esta es la mejor representación de las secuencias asociadas a la región i . Para grandes valores de R se tendrá un gran número de regiones y en consecuencia la representación será muy precisa (cuantización fina), mientras que si R es pequeña, el número de regiones

es pequeño y la distorsión es grande (cuantización gruesa). Existirán dos casos extremos. El primero sucede cuando hay una sola región ($R = 0$). En este caso el punto de representación es en cierto sentido (a ser clarificado más adelante) el centro de masa del espacio completo de la entrada. El segundo caso extremo es cuando cada región consiste en una única salida de la fuente. En este caso, R toma su valor máximo, y la distorsión es cero. Del teorema de velocidad-distorsión es posible concluir que si la función velocidad - distorsión está especificada para una fuente y una medida de distorsión, entonces se conoce el número mínimo de bits/símbolo de la fuente requerido para reconstruir la fuente con esa medida de distorsión. Especificada la velocidad, es posible determinar la distorsión mínima alcanzable si se utiliza una codificación a esa velocidad. Debe enfatizarse que, como en el caso del teorema de codificación de fuente, los resultados indicados por la función velocidad-distorsión son "límites fundamentales" en el sentido que pueden lograrse solo asintóticamente y con una complejidad creciente en su implementación.

Ejemplo: Para una fuente binaria sin memoria con $P(X_i = 1) = 1 - P(X_i = 0) = p$, y con distorsión de Hamming puede mostrarse que la función velocidad-distorsión está dada por

$$R(D) = \begin{cases} H_b(p) - H_b(D) & 0 \leq D \leq \min\{p, 1-p\} \\ 0, & \forall D \end{cases}$$

1. Suponiendo $p = 0.5$, cuántos bits/salida de la fuente serán necesarios para transmitir esta información con probabilidad de error a lo sumo igual a 0.25?

De la definición de distorsión de Hamming, la probabilidad de error y la distorsión promedio coinciden. En consecuencia $P_e = D = 0.25$ y como $p = 0.5$ se tiene un caso en que $0 \leq D \leq \min\{p, 1-p\}$. Esto significa que $R(0.25) = H_b(0.5) - H_b(0.25)$, lo que resulta en $R \cong 0.189$.

2. Con $p = 0.5$ y un canal que puede transmitir 0.75 bits/por cada salida de la fuente, cual es la probabilidad de error mínima alcanzable?

Para $R = 0.75$ es necesario resolver la ecuación $H_b(p) - H_b(D) = 0.75$, donde $H_b(p) = H_b(0.5) = 1$ y en consecuencia $H_b(D) = 0.25$, lo cual determina $P_e = D = 0.042$.

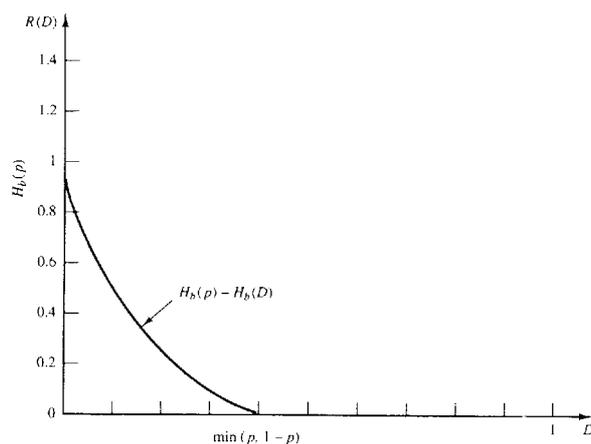


Figura 4.11: Función velocidad - distorsión para una fuente binaria con distorsión de Hamming.

Un gráfico de la función velocidad-distorsión para una fuente binaria se muestra en la figura 4.11. Para distorsión cero (probabilidad de error nula), se tiene que $R(D) = H_b(p)$, lo cual concuerda con el teorema de codificación de fuente. Suponiendo $p < 0.5$, para $D = p$ se tiene que $R(D) = 0$, o sea, podemos reproducir la fuente con una distorsión p sin transmisión, haciendo que el vector de reproducción sea cero. Esto significa que $D = p_e = P[X \neq \hat{X}] = P[X \neq 0] = P[X = 1] = p$.

Puede mostrarse también que para una fuente Gaussiana de media cero y varianza σ^2 y con una medida de distorsión cuadrática, la función velocidad-distorsión estará dada por

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} & 0 \leq D \leq \sigma^2 \\ 0, & \forall D \end{cases}$$

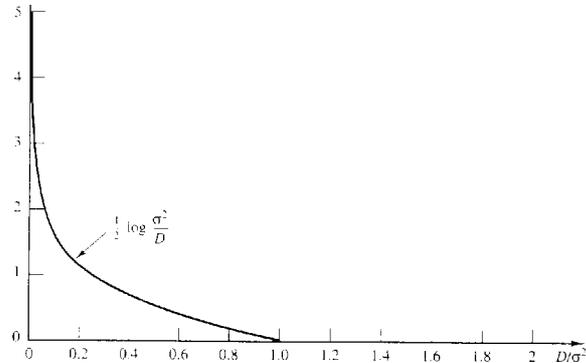


Figura 4.12: Función velocidad - distorsión para una fuente Gaussiana con distorsión media cuadrática.

Un gráfico de la función velocidad-distorsión se muestra en la figura 4.12. Una pregunta interesante ahora es la siguiente: en que factor decrece la distorsión si la velocidad de codificación de la fuente se incrementa en 1 bit? La respuesta, por supuesto, dependerá de la función velocidad-distorsión de la fuente (la cual a su vez dependerá de la estadística de la fuente y de la medida de distorsión). Considerando el caso de una fds Gaussiana de media cero con medida de distorsión cuadrática, y como $0 < D \leq \sigma^2$ y $R(D) = \frac{1}{2} \log \frac{\sigma^2}{D}$ es posible obtener la *función distorsión-velocidad* como $D(R) = \sigma^2 2^{-2R}$. Obviamente, incrementar R en 1, disminuirá D por un factor de 4, o equivalentemente en 6 dB. Esto significa que cada bit de transmisión por salida de la fuente reduce la distorsión en 6 dB.

Ejemplo: En la representación de una fuente Gaussiana de media cero y varianza unitaria, cual es la mínima distorsión alcanzable si se emplean 8 bits/salida de la fuente? En que factor se disminuye la distorsión si se utilizan 16 bits/salida de la fuente?

Utilizando la relación $D(R) = \sigma^2 2^{-2R}$, con $R = 8$ y $\sigma = 1$, se tiene que $D = \frac{1}{2^{16}} \cong 1.52 \times 10^{-5}$. Si se utilizan, en lugar de 8, 16 bits, la distorsión se reduce en 48 dB o por un factor de 4^8 .

4.5 Cuantización

En la sección previa se presentó el problema de representación de fuentes analógicas, que requiere teóricamente un número infinito de bits/salida de la fuente. En consecuencia, siempre existirá alguna distorsión en la transmisión de fuentes analógicas, y el objetivo es minimizar dicha distorsión. Se introdujo también la función velocidad-distorsión, que define un límite fundamental en el compromiso entre velocidad de codificación y distorsión. En esta sección se investigarán esquemas prácticos para representar la salida de una fuente analógica a bajas velocidades y, al mismo tiempo, sin introducir distorsión excesiva. Como ya se vió, el límite fundamental prometido por la función velocidad - distorsión puede aproximarse solo asintóticamente, o sea, mediante la utilización de codificadores y decodificadores muy complejos. El codificador observa salidas de la fuente de longitud n , $\mathbf{x} \in \mathcal{X}^n$ y las mapea a secuencias de representación de longitud n , $\hat{\mathbf{x}} \in \hat{\mathcal{X}}^n$, siendo 2^{nR} el número de estas secuencias y en consecuencia R bits/salida de la fuente serán necesarios para su transmisión. Cuanto más grande sea el valor de n , más próximo al límite de velocidad-distorsión operará el sistema. Esto significa que un esquema efectivo de cuantización debería trabajar con bloques de salida de la fuente antes que con salidas individuales. Los cuantizadores que operan sobre bloques de salida de la fuente se denominan *cuantizadores vectoriales*, en oposición a los *cuantizadores escalares*, que cuantizan cada salida separadamente.

Además de clasificar a los cuantizadores como escalares o vectoriales, es posible clasificarlos (y en general a los codificadores de fuente) sobre la base del método general para comprimir los datos como *codificadores*

de forma de onda o codificadores de análisis - síntesis. Los codificadores de forma de onda comprimen la señal analógica en base a varios métodos. En estos casos el mecanismo por el cual se genera la señal no es importante, sino que sus características son más relevantes (tales como ancho de banda, densidad espectral de potencia, propiedades estadísticas, etc.). Como el mecanismo mediante el cual se genera la forma de onda no es importante, los resultados son muy robustos y pueden aplicarse a todas las fuentes independientemente de su naturaleza. En los codificadores de análisis-síntesis no se comprime y transmite la señal directamente. En su lugar, se adopta un modelo para la producción de la señal y se comprimen y transmiten los parámetros del modelo. Por ejemplo, en la codificación de voz el mecanismo mediante el cual se produce la voz puede modelarse como un filtro variante en el tiempo excitado por ruido blanco o una secuencia de impulsos. En este tipo de codificadores, se cuantizan los parámetros del filtro y sus entradas (en forma escalar o vectorial) y luego se transmiten. En el receptor, se genera un filtro que simula el comportamiento del tracto vocal y se lo excita con la entrada apropiada de forma que resulte una replica de la señal requerida. Esta aproximación está basada en un modelo y no tiene la generalidad de la codificación de forma de onda. Por otro lado, los esquemas de cuantización basados en modelos alcanzan relaciones de compresión mejores que en el caso de los de forma de onda.

4.5.1 Cuantización escalar

En la cuantización escalar cada salida individual de la fuente se cuantiza en cierto número de niveles que después son codificados a una secuencia binaria. Cada salida de la fuente es un número real en general, pero aún así la transmisión de números reales requiere un número infinito de bits. En consecuencia, es necesario mapear el conjunto de los números reales en un conjunto finito y al mismo tiempo minimizar la distorsión introducida. En la cuantización escalar el conjunto de los números reales \mathcal{R} se particiona en N subconjuntos disjuntos \mathcal{R}_k , $1 \leq k \leq N$. A cada subconjunto \mathcal{R}_k corresponde un punto de representación \hat{x}_k . Si la salida de la fuente en el instante i , x_i pertenece a \mathcal{R}_k , entonces se representa por \hat{x}_k , que es la *versión cuantizada* de x_i . \hat{x}_k luego se representa mediante una secuencia binaria y se transmite. Dado que existen N posibilidades para los niveles cuantizados, $\log N$ bits son suficientes para codificar esos niveles en secuencias binarias (N se elige generalmente como una potencia de 2). En consecuencia, el número de bits requeridos para transmitir cada salida de la fuente es $R = \log N$ bits. El precio pagado por disminuir de infinitos niveles a $\log N$ es, por supuesto, la distorsión introducida.

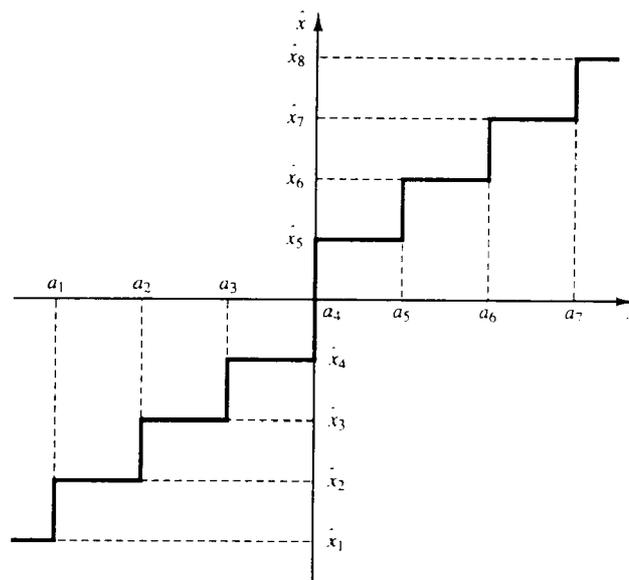


Figura 4.13: Ejemplo de un esquema de cuantización de 8 niveles.

La figura 4.13 muestra un ejemplo de un esquema de cuantización de 8 niveles. En este esquema las 8 regiones están definidas como $\mathcal{R}_1 = (-\infty, a_1]$, $\mathcal{R}_2 = (-a_1, a_2]$, \dots , $\mathcal{R}_8 = (-a_7, +\infty]$. El punto de representación (o valor cuantizado) en cada región es \hat{x}_i , como se muestra en la figura. La función de cuantización Q está definida por

$$Q(x) = \hat{x}_i \quad \forall x \in \mathcal{R}_i$$

Esta función también se muestra en la figura 4.13, de donde puede concluirse que la función cuantización es no lineal y no invertible. Esto se debe a que todos los puntos en \mathcal{R}_i se mapean en un solo punto \hat{x}_i . Dado que la función cuantización no es inversible, se pierde alguna información en el proceso de cuantización y esta pérdida de información no es recuperable.

Si se utiliza una medida de distorsión cuadrática, se tiene que

$$d(x, \hat{x}) = (x - Q(x))^2 = \tilde{x}^2$$

donde $\tilde{x} = x - \hat{x} = x - Q(x)$. Dado que X es una variable aleatoria, lo mismo serán \hat{X} y \tilde{X} , en consecuencia

$$D = E[d(X, \hat{X})] = E[X - Q(X)]^2$$

Ejemplo: Se tiene una fuente Gaussiana estacionaria de media cero y densidad espectral de potencia

$$S_x = \begin{cases} 2, & |f| < 100 \text{ Hz} \\ 0, & \forall f \end{cases}$$

La fuente se muestrea a la velocidad de Nyquist y cada muestra se cuantiza utilizando el cuantizador de 8 niveles de la figura 4.13 con $a_1 = -60$, $a_2 = -40$, $a_3 = -20$, $a_4 = 0$, $a_5 = 20$, $a_6 = 40$, $a_7 = 60$ y $\hat{x}_1 = -70$, $\hat{x}_2 = -50$, $\hat{x}_3 = -30$, $\hat{x}_4 = -10$, $\hat{x}_5 = 10$, $\hat{x}_6 = 30$, $\hat{x}_7 = 50$, $\hat{x}_8 = 70$. Cuales son la distorsión y velocidad resultantes?

La frecuencia de muestreo es $f_s = 200$ Hz. Cada muestra es una variable aleatoria Gaussiana de media cero y varianza

$$\sigma^2 = E(X_i^2) = R_X(\tau)|_{\tau=0} = \int_{-\infty}^{\infty} S_X(f) df = \int_{-100}^{100} 2 df = 400$$

Dado que cada muestra se cuantiza a 8 niveles, entonces se requerirán $\log 8 = 3$ bits por muestra, de forma que la velocidad es

$$R = 3f_s = 600 \text{ bits/seg}$$

Para hallar la distorsión es necesario evaluar $E(X - \hat{X})^2$ para cada muestra. Es posible mostrar que siempre que el proceso sea estacionario $E(X - \hat{X})^2 = P_{\tilde{X}(t)}$, o sea, la potencia del proceso $\tilde{X}(t)$, o equivalentemente, el promedio temporal de distorsión introducido por cuantización. Pero

$$D = E[X - \hat{X}]^2 = \int_{-\infty}^{\infty} (x - Q(x))^2 f_X(x) dx$$

donde $f_X(x)$ es la función densidad de probabilidad de la variable aleatoria X . De aquí se tiene que

$$D = \sum_{i=1}^8 \int_{\mathcal{R}_i} (x - Q(x))^2 f_X(x) dx$$

o equivalentemente

$$D = \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 f_X(x) dx + \sum_{i=2}^7 \int_{a_{i-1}}^{a_i} (x - \hat{x}_i)^2 f_X(x) dx + \int_{a_7}^{\infty} (x - \hat{x}_8)^2 f_X(x) dx$$

donde $f_X(x) = \frac{1}{\sqrt{2\pi 400}} e^{-x^2/800}$. Sustituyendo $\{a_i\}_{i=1}^7$ y $\{\hat{x}_i\}_{i=1}^8$ en esta integral y evaluando el resultado con una tabla de la función Q, se obtiene que $D \cong 33.38$.

Notar que si hubieramos utilizado 0 bits/salida de la fuente, entonces la mejor estrategia debería ser hacer la señal reconstruida igual a cero. En este caso la distorsión sería $D = E(X - 0)^2 = \sigma^2 = 400$. Este esquema de cuantización y la transmisión de 3 bits/salida de la fuente conduce a una reducción de la distorsión a 33.38. Es también interesante comparar el resultado anterior con el predicho por la cota de velocidad-distorsión. Sustituyendo $R = 3$ y $\sigma = 20$ en $R = \frac{1}{2} \log \frac{\sigma^2}{D}$ se obtiene $D = 6.25$. Obviamente, el esquema de cuantización simple mostrado en la figura 4.13 está lejos del óptimo. Las razones para este pobre desempeño son tres. Primero y principal, la cota de velocidad-distorsión es asintótica y se cumple para el mapeo óptimo en *bloques* de salidas de la fuente cuando la longitud del bloque tiende a infinito, mientras que en el ejemplo, se emplea un cuantizador escalar operando sobre las salidas individuales de la fuente. La segunda razón es que, aún con el cuantizador escalar con 8 niveles no se ha realizado ningún intento para diseñar en forma óptima el cuantizador, eligiendo los $\{a_i\}$ y \hat{x}_i apropiadamente. La tercera razón es que después de la cuantización las 8 salidas $\{\hat{x}_i\}_{i=1}^8$ no son equiprobables y pueden aún ser comprimidas. En este ejemplo,

$$p(\hat{x}_i) = \int_{a_{i-1}}^{a_i} \frac{1}{\sqrt{2\pi 400}} e^{-x^2/800} \quad 2 \leq i \leq 7$$

$$p(\hat{x}_1) = p(\hat{x}_8) = \int_{a_7}^{\infty} \frac{1}{\sqrt{2\pi 400}} e^{-x^2/800}$$

lo que resulta en $p(\hat{x}_1) = p(\hat{x}_8) = 0.0014$, $p(\hat{x}_2) = p(\hat{x}_7) = 0.0214$, $p(\hat{x}_3) = p(\hat{x}_6) = 0.1359$ y $p(\hat{x}_4) = p(\hat{x}_5) = 0.3414$. Utilizando el teorema de codificación de fuente es posible concluir que la salida del cuantizador puede comprimirse hasta $H(\hat{X}) = 2.105$ bits/salida de la fuente. Esto significa que tiene más sentido comparar 33.38 con el valor obtenido de la función distorsión con $R = 2.105$, lo que resulta en $D = 21.61$ (diferente de 6.25).

En el ejemplo anterior, se eligió como medida de desempeño $E[X - Q(X)]^2$, lo que se denomina *distorsión media cuadrática* o *ruido de cuantización*. Una medida de desempeño más significativa es una versión normalizada del ruido de cuantización con respecto a la potencia de la señal original.

Definición: Si la variable aleatoria X se cuantiza a $Q(X)$, la *relación señal a ruido de cuantización* (SQNR) se define por

$$SQNR = \frac{E[X^2]}{E[X - Q(X)]^2}$$

Cuando se trabaja con señales, la potencia del ruido de cuantización es

$$P_{\hat{X}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} E[X(t) - Q(X(t))]^2 dt$$

y la potencia de señal es

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

de donde la SQNR es

$$SQNR = \frac{P_X}{P_{\hat{X}}}$$

Puede mostrarse que si $X(t)$ es estacionario entonces esta ecuación es igual a la definición, donde X es una variable aleatoria representando $X(t)$ en cualquier punto.

Cuantización Uniforme

Los cuantizadores uniformes son los ejemplos más simples de cuantizadores escalares. En un cuantizador uniforme la línea real entera se particiona en N regiones. Todas las regiones excepto \mathcal{R}_1 y \mathcal{R}_N tienen igual longitud, Δ . Esto significa que para todo i , $1 \leq i \leq N-1$ se tiene que $a_{i+1} - a_i = \Delta$. La figura 4.13 es un ejemplo de un cuantizador uniforme de 8 niveles. En un cuantizador uniforme la distorsión está dada por

$$D = \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 f_X(x) dx + \sum_{i=1}^{N-2} \int_{a_1+(i-1)\Delta}^{a_1+i\Delta} (x - \hat{x}_{i+1})^2 f_X(x) dx + \int_{a_1+(N-2)\Delta}^{\infty} (x - \hat{x}_N)^2 f_X(x) dx$$

De aquí puede observarse que D es una función con $N = 2$ parámetros de diseño, o sea, a_1 , Δ y $\{\hat{x}_i\}_{i=1}^N$. Para diseñar el cuantizador uniforme óptimo, es necesario diferenciar D con respecto a las variables anteriores y hallar los valores que minimizan D .

Algunas suposiciones adicionales permiten simplificar la expresión anterior. Si se supone que $f_X(x)$ es una función par de x (función densidad de probabilidad simétrica), entonces el mejor cuantizador también tendrá propiedades de simetría. Esto significa que para N par, se tendrá $a_i = -a_{N-i} = -(\frac{N}{2} - i)\Delta$ para todo $1 \leq i \leq \frac{N}{2}$ (de donde $a_{\frac{N}{2}} = 0$), y $\hat{x}_i = -\hat{x}_{N+1-i}$ para $1 \leq i \leq \frac{N}{2}$. En ese caso se tendrá

$$D = 2 \int_{-\infty}^{(-\frac{N}{2}+1)\Delta} (x - \hat{x}_1)^2 f_X(x) dx + 2 \sum_{i=1}^{\frac{N}{2}-1} \int_{(-\frac{N}{2}+i)\Delta}^{(-\frac{N}{2}+i+1)\Delta} (x - \hat{x}_{i+1})^2 f_X(x) dx$$

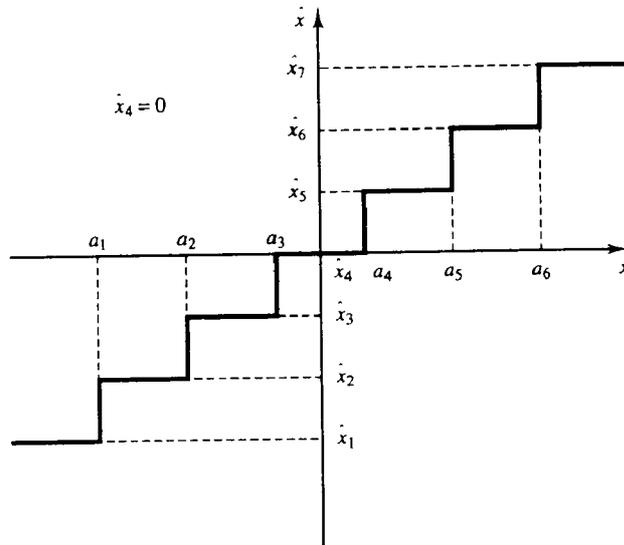


Figura 4.14: Cuantizador uniforme de siete niveles.

Cuando N es impar, se tendrá la situación mostrada en la figura 4.14. En este caso $a_i = -a_{N-i} = (-\frac{N}{2} + i)\Delta$ para todo $1 \leq i \leq \frac{N-1}{2}$ y $\hat{x}_i = -\hat{x}_{N+1-i}$ para $1 \leq i \leq \frac{N+1}{2}$, de donde $\hat{x}_{\frac{N+1}{2}} = 0$. La distorsión estará dada por

$$D = 2 \int_{-\infty}^{(-\frac{N}{2}+1)\Delta} (x - \hat{x}_1)^2 f_X(x) dx + 2 \sum_{i=1}^{\frac{N-3}{2}} \int_{(-\frac{N}{2}+i)\Delta}^{(-\frac{N}{2}+i+1)\Delta} (x - \hat{x}_{i+1})^2 f_X(x) dx + \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} x^2 f_X(x) dx$$

La minimización de la distorsión en esos casos, a pesar que mucho más simple si comparado con el caso general, es aún una tarea tediosa y se realiza principalmente mediante técnicas numéricas. La tabla 4.1 muestra los niveles de espaciamiento para cuantización óptima para variables aleatorias Gaussianas de media cero y varianza unitaria cuando los \hat{x} se eligen en los puntos medios de las regiones de cuantización.

Cuantización No Uniforme

Si se relaja la condición de que las regiones de cuantización (excepto la primera y la última) tengan igual longitud, es posible minimizar la distorsión con una menor cantidad de restricciones, de donde resulta que el cuantizador no uniforme se desempeña mejor comparado con el cuantizador uniforme con el mismo número de niveles. Suponiendo que se desea diseñar el cuantizador óptimo con N niveles de cuantización sin otras restricciones sobre las regiones, la distorsión será:

$$D = \int_{-\infty}^{a_1} (x - \hat{x}_1)^2 f_X(x) dx + \sum_{i=1}^{N-2} \int_{a_i}^{a_{i+1}} (x - \hat{x}_{i+1})^2 f_X(x) dx + \int_{a_{N-1}}^{\infty} (x - \hat{x}_N)^2 f_X(x) dx$$

En esta expresión existen un total de $2N - 1$ variables sobre las cuales es necesario realizar la minimización de D . Diferenciando respecto a los a_i se obtiene

$$\frac{\partial}{\partial a_i} D = f_X(a_i)[(a_i - \hat{x}_i)^2 - (a_i - \hat{x}_{i+1})^2] = 0$$

de donde resulta

$$a_i = \frac{1}{2}(\hat{x}_i + \hat{x}_{i+1}) \quad (4.1)$$

Este resultado significa que, en un cuantizador óptimo, *los límites de las regiones de cuantización son los puntos medios de los valores cuantizados*. Dado que la cuantización se realiza en base a la distancia mínima, cada valor x se cuantiza al $\{\hat{x}\}_{i=1}^N$ más próximo. Para determinar los valores cuantizados \hat{x}_i es posible diferenciar D con respecto a \hat{x}_i y definir $a_0 = -\infty$ y $a_N = \infty$. De esta forma se obtiene

$$\frac{\partial}{\partial \hat{x}_i} D = \int_{a_{i-1}}^{a_i} 2(x - \hat{x}_i) f_X(x) dx = 0$$

de donde resulta

$$\begin{aligned} \hat{x}_i &= \frac{\int_{a_{i-1}}^{a_i} x f_X(x) dx}{\int_{a_{i-1}}^{a_i} f_X(x) dx} = \frac{\int_{a_{i-1}}^{a_i} x f_X(x) dx}{P(a_{i-1} < X \leq a_i)} = \int_{a_{i-1}}^{a_i} x \frac{f_X(x)}{P(a_{i-1} < X \leq a_i)} dx \\ &= \int_{a_{i-1}}^{a_i} x f_X(x|a_{i-1} < X \leq a_i) dx \\ &= E[X|a_{i-1} < X \leq a_i] \end{aligned} \quad (4.2)$$

donde se utilizó la definición de densidad de probabilidad condicional

$$f_X(x|a_{i-1} < X \leq a_i) = \begin{cases} \frac{f_X(x)}{P(a_{i-1} < X \leq a_i)} & a_{i-1} < x \leq a_i \\ 0, & \forall x \end{cases}$$

La ecuación anterior muestra que en un cuantizador óptimo *el valor cuantizado (o punto de representación) para una región debe elegirse en el centroide (valor esperado condicional) de esa región*. Las ecuaciones (4.1) y (4.2) determinan las condiciones necesarias para que un cuantizador escalar sea óptimo y se conocen como condiciones de Lloyd-Max. El criterio para cuantización óptima puede resumirse como sigue:

Nro. niveles N	Paso Δ	Error medio Cuadrático D	Entropía $H(\hat{x})$
1	–	1.00000	0.000
2	1.5960	0.36340	1.000
3	1.2240	0.19020	1.536
4	0.9957	0.11880	1.904
5	0.8430	0.08218	2.183
6	0.7334	0.06065	2.409
7	0.6508	0.04686	2.598
8	0.5860	0.03744	2.761
9	0.5338	0.03069	2.904
10	0.4908	0.02568	3.032
11	0.4546	0.02185	3.148
12	0.4238	0.01885	3.253
13	0.3972	0.01645	3.350
14	0.3739	0.01450	3.440
15	0.3534	0.01289	3.524
16	0.3352	0.01154	3.602
17	0.3189	0.01040	3.676
18	0.3042	0.009430	3.746
19	0.2909	0.008594	3.811
20	0.2788	0.007869	3.874
21	0.2678	0.007235	3.933
22	0.2576	0.006678	3.990
23	0.2482	0.006185	4.045
24	0.2396	0.005747	4.097
25	0.2315	0.005355	4.146
26	0.2240	0.005004	4.194
27	0.2171	0.004687	4.241
28	0.2105	0.004401	4.285
29	0.2044	0.004141	4.328
30	0.1987	0.003905	4.370
31	0.1932	0.003688	4.410
32	0.1881	0.003490	4.449
33	0.1833	0.003308	4.487
34	0.1787	0.003141	4.524
35	0.1744	0.002986	4.560
36	0.1703	0.002843	4.594

Tabla 4.1: Cuantizador uniforme óptimo para una fuente Gaussiana

1. Los límites de las regiones de cuantización son los puntos medios de los valores cuantizados correspondientes (regla del vecino más próximo).
2. Los valores cuantizados son los centroides de las regiones de cuantización.

A pesar que esas reglas son muy simples, no resultan en soluciones analíticas para el diseño del cuantizador óptimo. El método usual para diseñar el cuantizador óptimo es comenzar con un conjunto de regiones de cuantización y luego hallar los valores cuantizados siguiendo el segundo criterio. Luego se diseñan nuevas regiones de cuantización para los nuevos valores cuantizados y se alterna entre los dos pasos hasta que la distorsión no cambia más hallá de un valor predefinido de un paso a otro. En base a este método es posible diseñar el cuantizador no uniforme óptimo para varias estadísticas de fuente.

4.5.2 Cuantización vectorial

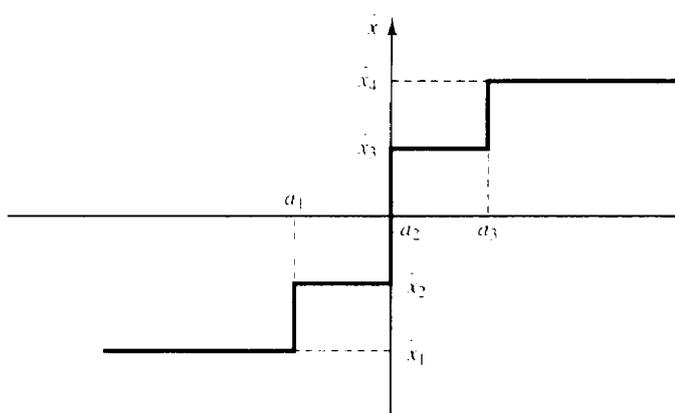


Figura 4.15: Cuantizador escalar de cuatro niveles.

En la cuantización escalar cada salida de la fuente discreta en tiempo (la cual es usualmente el resultado del muestreo de una fuente continua en tiempo) se cuantiza separadamente y luego se codifica. Por ejemplo, si se utiliza un cuantizador escalar de 4 niveles y se codifica cada nivel en 2 bits se estará utilizando 2 bits por cada salida de la fuente. Este esquema de cuantización se muestra en la figura 4.15.

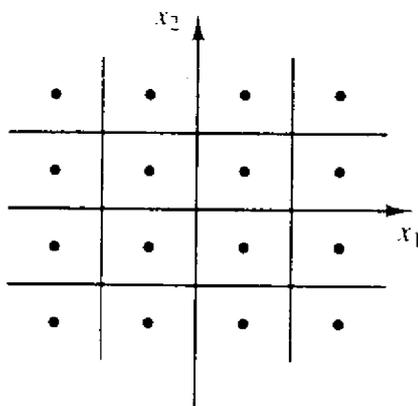


Figura 4.16: Cuantizador escalar de 4 niveles aplicado a dos muestras.

Si se considera ahora dos muestras de salida de la fuente a la vez, y se interpreta esas dos muestras como un punto en el plano, el cuantizador escalar particiona el plano en 16 regiones de cuantización como mostrado en la figura 4.16. Puede observarse que las regiones en el espacio bidimensional tienen todas forma rectangular. Si se permiten regiones de una forma arbitraria es posible obtener mejores resultados. Esto

significa que se está cuantizando 2 salidas de la fuente a la vez utilizando 16 regiones, lo cual es equivalente a 4 bits/ 2 salidas de la fuente o 2 bits/cada salida de la fuente. En consecuencia, el número de bits/salida de la fuente para cuantizar dos muestras a la vez es igual al número de bits/salida de la fuente obtenidos para el caso escalar. Si se relajan las requerimientos de tener regiones rectangulares, el desempeño mejorará. Si se toman ahora tres muestras a la vez y se cuantiza el espacio tridimensional en 64 regiones, se tendrá aún menor distorsión con el mismo número de bits/salida de la fuente. La idea de la cuantización vectorial es tomar bloques de salidas de la fuente de longitud n y diseñar el cuantizador en el espacio n -dimensional antes que realizar la cuantización basandose en muestras individuales en el espacio unidimensional.

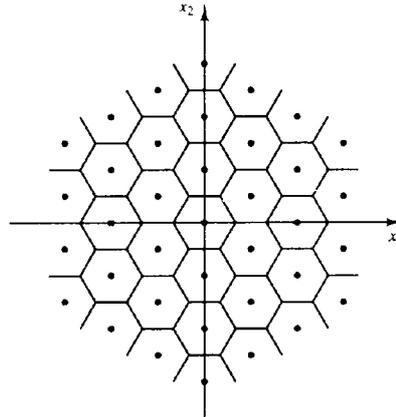


Figura 4.17: Cuantización vectorial en dos dimensiones.

Suponiendo que las regiones de cuantización en el espacio n -dimensional se notan por \mathcal{R}_i , $1 \leq i \leq K$. Esas K regiones particionan el espacio n -dimensional. Cada bloque de salidas de la fuente de longitud n se nota por $\mathbf{x} \in \mathbb{R}^n$ y si $\mathbf{x} \in \mathcal{R}_i$ entonces está cuantizado a $Q(\mathbf{x}) = \hat{\mathbf{x}}_i$. La figura 4.17 muestra este esquema de cuantización para $n = 2$ y $K = 37$. Dado que existe un total de K valores cuantizados, $\log K$ bits serán suficientes para representarlos. Esto significa que serán necesarias $\log K$ bits/ n salidas de la fuente o, que la velocidad de codificación de la fuente será

$$R = \frac{\log K}{n} \quad \text{bits/salida de la fuente}$$

El cuantizador vectorial óptimo de dimensión n y el número de niveles K son aquellos que resultan de elegir la región \mathcal{R}_i y los valores de cuantización $\hat{\mathbf{x}}_i$ de forma que la distorsión sea mínima. Aplicando el mismo procedimiento que en el caso de cuantización escalar, es posible obtener el siguiente criterio para diseñar un cuantizador vectorial óptimo:

1. La región \mathcal{R}_i es el conjunto de todos los puntos del espacio n -dimensional próximos a $\hat{\mathbf{x}}_i$, o sea para $i \neq j$,

$$\mathcal{R}_i = \{\mathbf{x} \in \mathcal{R}^n : \|\mathbf{x} - \hat{\mathbf{x}}_i\| < \|\mathbf{x} - \hat{\mathbf{x}}_j\|, \forall j \neq i\}$$

2. $\hat{\mathbf{x}}_i$ es el centroide de la region \mathcal{R}_i , donde

$$\hat{\mathbf{x}}_i = \frac{1}{p(\mathbf{X} \in \mathcal{R}_i)} \int \int \cdots \int_{\mathcal{R}_i} \mathbf{x} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

Una aproximación práctica al diseño de cuantizadores vectoriales se basa en la misma idea utilizada en el diseño de cuantizadores escalares óptimos. Comenzando a partir de un conjunto de regiones de cuantización es posible obtener los vectores cuantizados óptimos para esas regiones utilizando el segundo criterio, luego reparticionando el espacio utilizando el primer criterio y se repiten estos pasos hasta alcanzar una distorsión despreciable.

La cuantización vectorial tiene gran aplicación en la codificación de voz e imagen y se han propuesto numerosos algoritmos para reducir su complejidad computacional. Puede probarse que para fuentes ergódicas y estacionarias, el desempeño del cuantizador vectorial óptimo se aproxima al desempeño óptimo dado por la función velocidad-distorsión cuando n es suficientemente grande.

4.6 Codificación de forma de onda

Los esquemas de codificación de forma de onda se diseñan para reproducir la señal de salida de la fuente en el destino con la menor distorsión posible. En estas técnicas no se presta atención al mecanismo que produce la forma de onda, y toda la atención se dirige a la reproducción de la salida de la fuente en el destino con alta fidelidad. Dado que la estructura de la fuente no juega ningún rol en el diseño de los codificadores de forma de onda y solo las propiedades de la señal afectan el diseño, los codificadores de forma de onda son robustos y pueden utilizarse con una variedad de fuentes, siempre que las formas de onda producidas por esas fuentes tengan ciertas similitudes.

4.6.1 Modulación de pulsos codificados (PCM)



Figura 4.18: Diagrama en bloques de un sistema PCM.

La Modulación de pulsos Codificados (PCM) es la forma más simple y antigua de codificación de forma de onda. Un codificador PCM está formado por tres partes básicas: un muestreador, un cuantizador y un codificador. Un bloque funcional de un sistema PCM se muestra en la figura 4.18.

La señal que entra al muestreador es de ancho de banda limitado, W . Usualmente existe un filtro con ancho de banda W previo al muestreador para evitar que frecuencias superiores a W entren al muestreador. Este filtro se denomina el *filtro de premuestreo*. El muestreo se realiza a una velocidad superior a la de Nyquist para permitir una banda de salvaguarda. Los valores muestreados entran luego a un cuantizador escalar. El cuantizador es o un cuantizador uniforme, que resulta en un sistema PCM uniforme, o un cuantizador no uniforme. La elección del cuantizador se basa en las características de la salida de la fuente. La salida del cuantizador luego se codifica en una secuencia binaria de longitud ν , donde $N = 2^\nu$ es el número de niveles de cuantización.

PCM uniforme

En la aplicación de PCM uniforme se supone que el rango de la entrada es $[-x_{max}, x_{max}]$ y el número de niveles de cuantización N es una potencia de 2, $N = 2^\nu$. A partir de esto, la longitud de cada región de cuantización es

$$\Delta = \frac{2x_{max}}{N} = \frac{x_{max}}{2^{\nu-1}}$$

Los valores cuantizados en PCM uniforme se eligen como los puntos medios de las regiones de cuantización, y en consecuencia el error $\tilde{x} = x - Q(x)$ es una variable aleatoria en el intervalo $(-\frac{\Delta}{2}, \frac{\Delta}{2}]$. En las aplicaciones usuales de PCM, el número de niveles N es alto y el rango de variaciones de la señal de entrada (x_{max}) es pequeño. Esto significa que la longitud de cada región de cuantización (Δ) es pequeña y, con esas suposiciones, en cada región de cuantización el error $\tilde{X} = X - Q(X)$ puede aproximarse adecuadamente mediante una variable aleatoria uniformemente distribuida en $(-\frac{\Delta}{2}, \frac{\Delta}{2}]$. La distorsión introducida por la cuantización (ruido de cuantización) será

$$E[\tilde{X}^2] = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \frac{1}{\Delta} \tilde{x}^2 d\tilde{x} = \frac{\Delta^2}{12} = \frac{x_{max}^2}{3N^2} = \frac{x_{max}^2}{3 \times 4^\nu}$$

donde ν es el número de bits por muestra de la fuente. La relación señal a ruido de cuantización (SQNR) será entonces

$$SQNR = \frac{\overline{X^2}}{\overline{\tilde{X}^2}} = \frac{3 \times N^2 \overline{X^2}}{x_{max}^2} = \frac{3 \times 4^\nu \overline{X^2}}{x_{max}^2}$$

Si normalizamos X con \tilde{X} , o sea, $\tilde{X} = \frac{X}{x_{max}}$, entonces

$$SQNR = 3 \times N^2 \overline{\tilde{X}^2} = 3 \times 4^\nu \overline{\tilde{X}^2}$$

Notar que por definición $|\tilde{X}| \leq 1$ y en consecuencia $\overline{\tilde{X}^2} \leq 1$. Esto significa que $3N^2 = 3 \times 4^\nu$ es un límite superior a la SQNR en PCM uniforme. Esto también significa que la SQNR de PCM uniforme se deteriora cuando se incrementa el rango dinámico de la entrada debido a que un incremento en ese rango dinámico resulta en un incremento de $\overline{\tilde{X}^2}$. Para reducir la sensibilidad en relación al rango dinámico de la fuente se utiliza en general PCM no uniforme.

Expresando la SQNR en dB se obtiene

$$SQNR|_{dB} \cong P_{\tilde{X}}|_{dB} = 6\nu + 4.8$$

De aquí puede notarse que cada bit extra (aumentando ν en uno) incrementa la SQNR en 6 dB.

Ejemplo: Cual es la SQNR resultante de una señal uniformemente distribuida sobre $[-1, 1]$ cuando se emplea PCM uniforme con 256 niveles?

Dado que $x_{max} = 1$, entonces $\tilde{X} = X$ y $P_{\tilde{X}} = \int_{-1}^1 \frac{1}{2} x^2 dx = \frac{1}{2}$. En consecuencia, usando $\nu = \log_2 256 = 8$, se tiene que

$$SQNR = 3 \times 4^\nu \overline{\tilde{X}^2} = 65536 \quad (\cong 48.16 \text{ dB})$$

Los requerimientos de ancho de banda de los sistemas de transmisión de pulsos digitales, de los cuales PCM es un ejemplo, serán estudiados en el Capítulo 6. Aquí discutiremos brevemente algunos resultados concernientes con los requerimientos de ancho de banda de un sistema PCM. Si una señal tiene un ancho de banda W , entonces el número mínimo de muestras para reconstrucción perfecta de la señal está dado por el teorema del muestreo, $2W$ muestras/seg. Si se requiere alguna banda de salvaguarda, entonces el número de muestras por segundo es f_s , generalmente mayor que $2W$. Se utilizan ν bits por cada muestra, y en consecuencia se requerirán νf_s bits/seg para la transmisión de la señal PCM. En el caso de muestreo a la velocidad de Nyquist, esto es igual a $2\nu W$ bits/seg. El requerimiento de ancho de banda mínimo para la transmisión de R bits/seg es (según se estudiará en el Capítulo 6) $\frac{R}{2}$. En consecuencia el requerimiento de ancho de banda mínimo de un sistema PCM es

$$BW = \frac{\nu f_s}{2}$$

lo cual, usando la velocidad de Nyquist, determina el requerimiento mínimo absoluto de ancho de banda dado por

$$BW = \nu W$$

Esto significa que un sistema PCM expande el ancho de banda de la señal original por un factor igual a ν .

PCM no uniforme

Siempre que la estadística de la señal de entrada sea próxima a una distribución uniforme, el sistema PCM uniforme trabajará adecuadamente. Sin embargo, en la codificación de ciertas señales como por

ejemplo voz, la distribución de la entrada está lejos de ser uniformemente distribuida. Para una señal de voz en particular existe mayor probabilidad de amplitudes pequeñas y menor probabilidad de amplitudes grandes. En consecuencia, tiene sentido para el diseñador un cuantizador con más regiones de cuantización en amplitudes bajas y menos regiones en amplitudes altas. El cuantizador resultante será un cuantizador no uniforme.

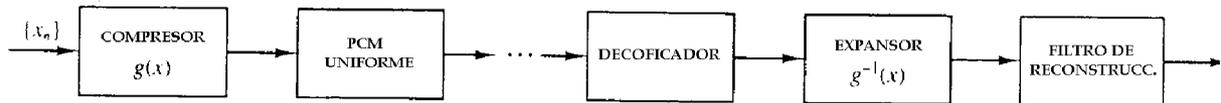


Figura 4.19: Diagrama en bloques de un sistema PCM no uniforme.

El método usual para efectuar cuantización no uniforme es pasar primero las muestras a través de un elemento no lineal que comprima las amplitudes grandes (reduzca el rango dinámico de la señal) y entonces efectuar una cuantización uniforme sobre la salida. En el receptor, se aplica la inversa (expansión) de esta operación no lineal para obtener el valor de la muestra. Esta técnica se denomina *compresión-expansión*. Un diagrama de bloques de este sistema se muestra en la figura 4.19.

Existen dos tipos de compresores - expansores ampliamente utilizados para codificación de voz. El compresor-expansor de ley μ , utilizado en EEUU y Canadá emplea la función logarítmica en la transmisión, con $|x| \leq 1$,

$$g(x) = \frac{\log(1 + \mu|x|)}{\log(1 + \mu)} \operatorname{sgn}(x)$$

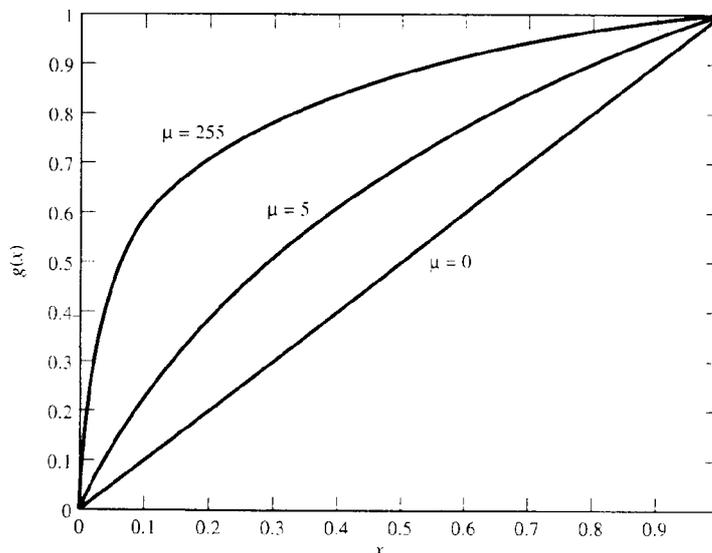
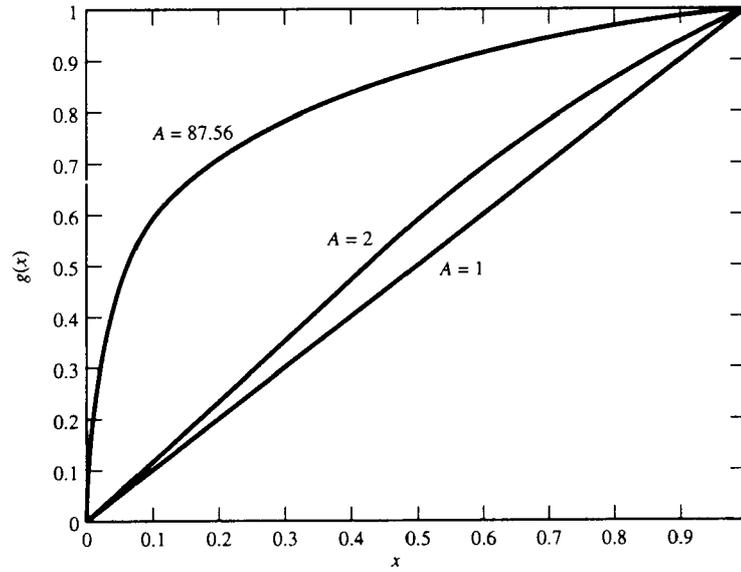


Figura 4.20: Característica del compander de ley μ .

El parámetro μ controla la cantidad de compresión y expansión. El sistema estándar en EEUU y Canadá emplea un compresor con $\mu = 255$, seguido de un cuantizador uniforme con 128 niveles (7 bits/muestra). La utilización de un compresor-expansor en este sistema mejora el desempeño del sistema en alrededor de 24 dB. Un gráfico de la característica de un compresor-expansor de ley μ se muestra en la figura 4.20.

El segundo compresor logarítmico más ampliamente utilizado es el compresor-expansor de ley A . La característica de este compresor-expansor se muestra en la figura 4.21 y está dada por

Figura 4.21: Característica del compander de ley A .

$$g(x) = \begin{cases} \frac{1+\ln A|x|}{1+\ln A} \operatorname{sgn}(x), & \frac{1}{A} \leq |x| \leq 1 \\ \frac{A|x|}{1+\ln A} \operatorname{sgn}(x), & 0 \leq |x| \leq \frac{1}{A} \end{cases}$$

donde A se elige igual a 87.56.

Diseño óptimo del compander

Anteriormente se estudiaron los criterios para el diseño óptimo de cuantizadores. La aproximación a la solución del problema de diseño óptimo de un compresor-expansor es similar. La solución da una idea aproximada de las características del compresor-expansor ideal, que cuando es seguido de un cuantizador uniforme, determina una solución próxima a la óptima. Definiendo $a_0 = -x_{max}$ y $a_N = +x_{max}$ se tiene que

$$D = \sum_{i=1}^N \int_{a_{i-1}}^{a_i} (x - \hat{x}_i)^2 f_X(x) dx$$

Si el número de regiones de cuantización es grande y la función densidad de probabilidad es suficientemente suave, es posible asumir en cada región una distribución uniforme y, en consecuencia, $\hat{x}_i = \frac{a_{i-1} + a_i}{2}$. Sustituyendo \hat{x}_i en la relación anterior se tiene que

$$D \cong \sum_{i=1}^N f_X(a_{i-1}) \frac{\Delta_i^2}{12} \quad (4.3)$$

donde $\Delta_i = a_i - a_{i-1}$. Notar que la entrada al compresor es una muestra cuantizada no uniformemente con regiones de cuantización de tamaño Δ_i y la salida es una muestra uniformemente cuantizada con regiones de cuantización del mismo tamaño Δ . De la figura 4.22 es posible ver que

$$g'(a_{i-1}) \cong \frac{\Delta}{\Delta_i} \quad (4.4)$$

Suponiendo que la función $g(x)$ mapea el intervalo $[-x_{max}, x_{max}]$ en $[-y_{max}, y_{max}]$, entonces las condiciones de borde serán

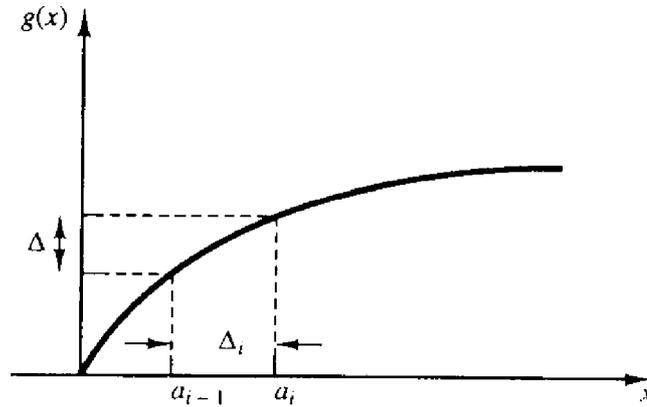


Figura 4.22: Relación entrada - salida del compander.

$$\begin{cases} g(-x_{max}) = -y_{max} \\ g(x_{max}) = y_{max} \end{cases}$$

Sustituyendo (4.4) en (4.3) se obtiene

$$D \cong \sum_{i=1}^N f_X(a_{i-1}) \left[\frac{\Delta}{g'(a_{i-1})} \right]^2 \Delta_i$$

usando $\Delta = \frac{y_{max}}{2^{\nu-1}}$ y suponiendo que N es muy grande se tiene que

$$D \cong \frac{y_{max}^2}{3 \times 4^\nu} \int_{-\infty}^{\infty} \frac{f_X(x)}{|g'(x)|^2} dx$$

donde $g(x)$ es la característica del compresor. La relación anterior determina la distorsión cuando se emplea un compresor con característica $g(x)$. Es posible minimizar D para obtener el compresor óptimo. El compresor óptimo resultante tiene la siguiente forma

$$g(x) = y_{max} \left[\frac{2 \int_{-\infty}^x [f_X(\eta)]^{\frac{1}{3}} d\eta}{\int_{-\infty}^{\infty} [f_X(\eta)]^{\frac{1}{3}} d\eta} - 1 \right]$$

y la distorsión resultante será

$$D = \frac{1}{12 \times 4^\nu} \left[\int_{-\infty}^{\infty} [f_X(\eta)]^{\frac{1}{3}} d\eta \right]^3$$

4.6.2 Modulación de pulsos codificados diferencial (DPCM)

En un sistema PCM, después de muestrear la señal, cada muestra se cuantiza independientemente utilizando un cuantizador escalar. Esto significa que los valores de las muestras previas no tienen efecto sobre la cuantización de las muestras nuevas. Sin embargo, como se mencionó anteriormente, cuando un proceso estocástico limitado en ancho de banda se muestrea a la frecuencia de Nyquist o más rápido, las muestras obtenidas son variables aleatorias usualmente correlacionadas. La excepción a esto es cuando el espectro del proceso es plano dentro del ancho de banda de interés. Esto significa que las muestras previas transportan alguna información sobre la próxima muestra, y esta información puede emplearse para mejorar el desempeño del sistema PCM. Por ejemplo, si los valores de las muestras previas fueran pequeños, con alta probabilidad

el próximo valor de muestra será también pequeño y, en consecuencia, no es necesario cuantizar un amplio rango de valores para lograr un buen desempeño.

En su forma más simple, en la modulación de pulsos codificados diferencial (DPCM), se cuantiza la diferencia entre dos muestras adyacentes. Dado que dos muestras adyacentes están altamente correlacionadas, su diferencia tiene variaciones pequeñas y, en consecuencia, para lograr un cierto nivel de desempeño se requerirán menor cantidad de bits para cuantizarla. Esto significa que DPCM puede lograr niveles de desempeño similares a PCM pero con menores velocidades de transmisión.

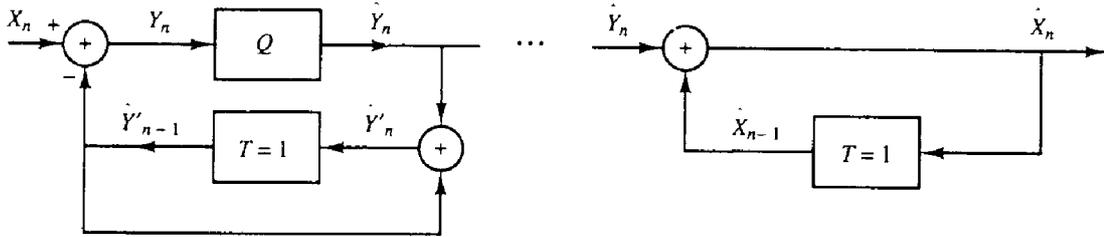


Figura 4.23: Un esquema DPCM simple.

La figura 4.23 muestra un diagrama en bloques de este esquema DPCM simple. Como puede observarse en la figura, la entrada al cuantizador no es simplemente $X_n - X_{n-1}$ sino $X_n - \hat{Y}'_{n-1}$. Veremos que \hat{Y}'_{n-1} está próximamente relacionada con X_{n-1} y la elección anterior tiene la ventaja de evitar la acumulación del ruido de cuantización. La entrada al cuantizador, Y_n , se cuantiza mediante un cuantizador escalar (uniforme o no uniforme) para producir \hat{Y}_n . Utilizando las relaciones

$$Y_n = X_n - \hat{Y}'_{n-1} \quad \hat{Y}'_n = \hat{Y}_n + \hat{Y}'_{n-1} \quad (4.5)$$

obtenemos el error de cuantización entre la entrada y la salida del cuantizador como

$$\hat{Y}_n - Y_n = \hat{Y}_n - (X_n - \hat{Y}'_{n-1}) = \hat{Y}_n - X_n + \hat{Y}'_{n-1} = \hat{Y}'_n - X_n \quad (4.6)$$

En el receptor se tiene que

$$\hat{X}_n = \hat{Y}_n + \hat{X}_{n-1} \quad (4.7)$$

Comparando (4.5) y (4.7) puede observarse que \hat{Y}'_n y \hat{X}_n satisfacen la misma ecuación con la misma excitación \hat{Y}_n . En consecuencia, si las condiciones iniciales de \hat{Y}'_n y \hat{X}_n se eligen iguales, ellas serán iguales. Por ejemplo, si hacemos $\hat{Y}'_{-1} = \hat{X}_{-1} = 0$, entonces para todo n tendremos que $\hat{Y}'_n = \hat{X}_n$. Sustituyendo esto en (4.6) se obtiene

$$\hat{Y}_n - Y_n = \hat{X}_n - X_n$$

Esto muestra que el error de cuantización entre X_n y su reproducción \hat{X}_n es el mismo que el error de cuantización entre la entrada y la salida del cuantizador. Sin embargo, el rango de variaciones de Y_n es usualmente mucho menor comparado con el de X_n y en consecuencia, Y_n puede cuantizarse con menos bits.

En una versión más compleja de DPCM, en lugar de utilizar solo la muestra previa, se utilizan las últimas p muestras para predecir el valor de la próxima. Luego se cuantiza la diferencia entre la muestra X_n y su valor predicho. Se emplea usualmente un predictor lineal de la forma $\sum_{i=1}^p a_i X_{n-i}$, y los coeficientes del predictor a_i se eligen para minimizar el error medio cuadrático entre la muestra X_n y el valor predicho

$$D = E\left[X_n - \sum_{i=1}^p a_i X_{n-i}\right]^2$$

Expandiendo y suponiendo que X_n es un proceso estacionario se obtiene

$$D = R_X(0) - 2 \sum_{i=1}^p a_i R_X(i) + \sum_{i=1}^p \sum_{j=1}^p a_i a_j R_X(i-j)$$

Para minimizar D , se diferencia en relación a los a_i y se iguala a cero. Después de diferenciar se obtiene

$$\sum_{i=1}^p a_i R_X(i-j) = R_X(j) \quad 1 \leq j \leq p$$

donde R_X es la función autocorrelación del proceso X_n . Resolviendo el sistema de ecuaciones anterior (denominado usualmente *ecuaciones de Youle-Walker*), es posible hallar el conjunto de coeficientes del predictor óptimo $\{a_i\}_{i=1}^p$.

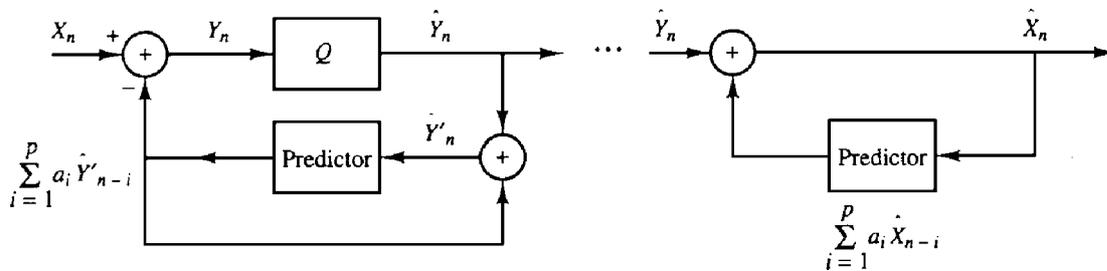


Figura 4.24: Un esquema DPCM general.

La figura 4.24 muestra un diagrama en bloques de un sistema DPCM general. Este diagrama en bloques es muy similar al de la figura 4.23. La única diferencia es que el retardo $T + 1$ se ha sustituido con el filtro de predicción $\sum_{i=1}^p a_i X_{n-i}$. Un análisis similar al anterior muestra que $\hat{Y}_n - Y_n = \hat{X}_n - X_n$. Debido al uso de un predictor de p pasos, se está utilizando más información en la predicción de X_n y en consecuencia, el rango de variaciones de Y_n será menor. Esto a su vez significa que pueden lograrse velocidades de codificación aún menores en este caso. Los sistemas DPCM encuentran amplia aplicación en la compresión de imagen y voz.

4.6.3 Modulación Delta (ΔM)

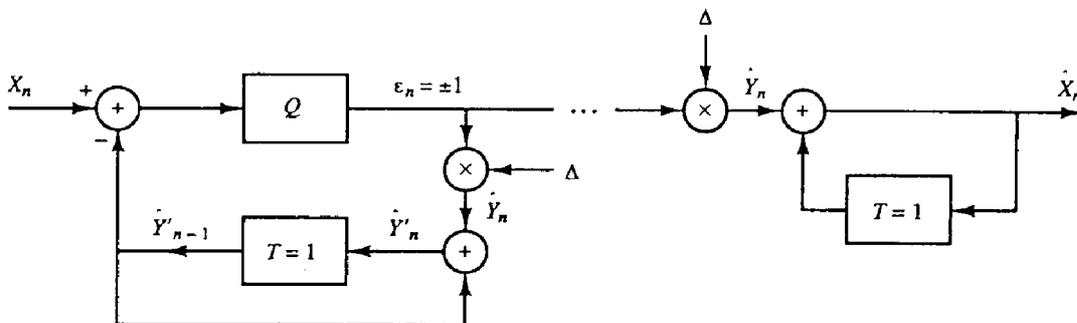


Figura 4.25: Modulación Delta.

La modulación Delta (ΔM) es una versión simplificada del esquema DPCM mostrado en la figura 4.23. En ΔM se utiliza un cuantizador de 1 bit (2 niveles) con magnitudes $\pm\Delta$. Un diagrama en bloques de un

sistema ΔM se muestra en la figura 4.25. El mismo análisis que se aplicó en el sistema DPCM simple es válido aquí.

Dado que en ΔM se emplea solo 1 bit/muestra, el ruido de cuantización será alto a menos que el rango dinámico de Y_n sea muy bajo. Esto a su vez significa que X_n y X_{n-1} deben tener un coeficiente de correlación muy alto. Para tener alta correlación entre X_n y X_{n-1} es necesario muestrear a una velocidad mucho mayor que la de Nyquist. En consecuencia, la frecuencia de muestreo en ΔM es usualmente mucho mayor que la de Nyquist pero, dado que el número de bits/muestra es solo 1, el número total de bits/seg requerido para transmitir una señal es menor que el de un sistema PCM.

Una ventaja principal de ΔM es la estructura muy simple del sistema. En el receptor se tiene la siguiente relación para la reconstrucción de \hat{X}_n

$$\hat{X}_n - \hat{X}_{n-1} = \hat{Y}_n$$

Resolviendo esta ecuación para \hat{X}_n y suponiendo condiciones iniciales nulas, se obtiene

$$\hat{X}_n = \sum_{i=0}^n \hat{Y}_i$$

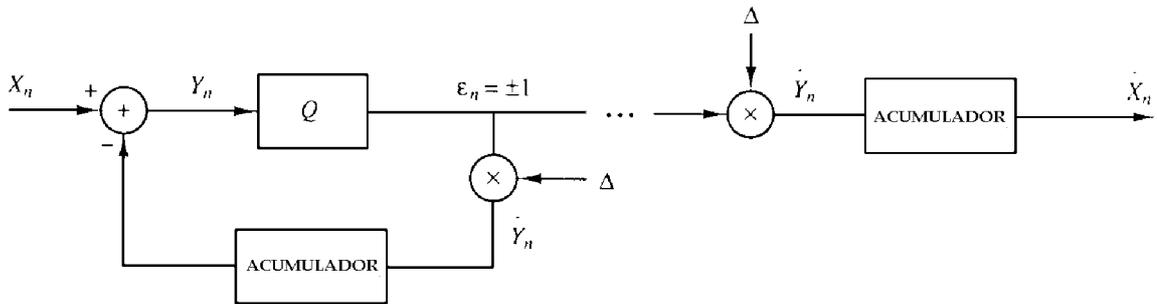


Figura 4.26: Modulación delta con integradores.

Esto significa que para obtener \hat{X}_n son necesarios solo los valores de \hat{Y}_n . Si los valores de las muestras se representan mediante impulsos, el acumulador puede representarse por un integrador simple. Esto simplifica el diagrama en bloques del sistema ΔM al de la figura 4.26.

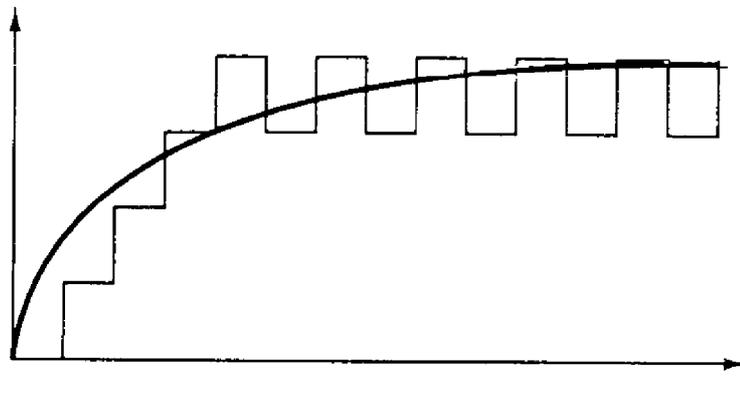


Figura 4.27: Δ grande y el ruido granular.

El tamaño del paso Δ es un parámetro muy importante en el diseño de un sistema ΔM . Grandes valores de Δ hacen que el modulador siga cambios rápidos en la señal de entrada pero al mismo tiempo causa un

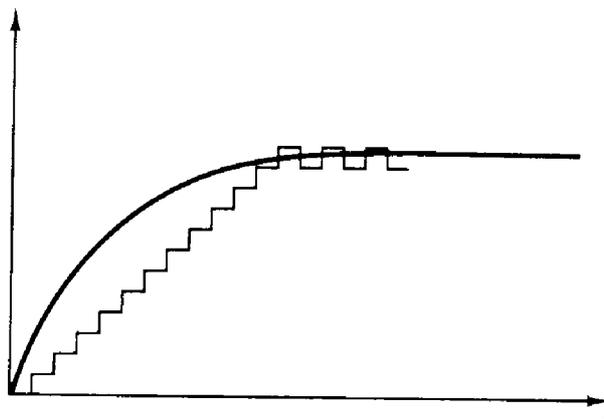


Figura 4.28: Δ pequeño y distorsión de sobrecarga de pendiente.

excesivo ruido de cuantización cuando la entrada cambia lentamente. Esto se muestra en la figura 4.27. Como puede verse en esta figura, para grandes valores de Δ y cuando la entrada varía lentamente se produce un gran ruido de cuantización, conocido como *ruido granular*. El caso de un Δ muy pequeño se muestra en la figura 4.28. En este caso se encuentran problemas cuando la entrada tiene cambios muy rápidos. En ese caso, le toma mayor tiempo a la salida seguir a la entrada y se produce un ruido de cuantización excesivo en este período. Este tipo de distorsión se denomina *de sobrecarga*.

Modulación Delta adaptiva

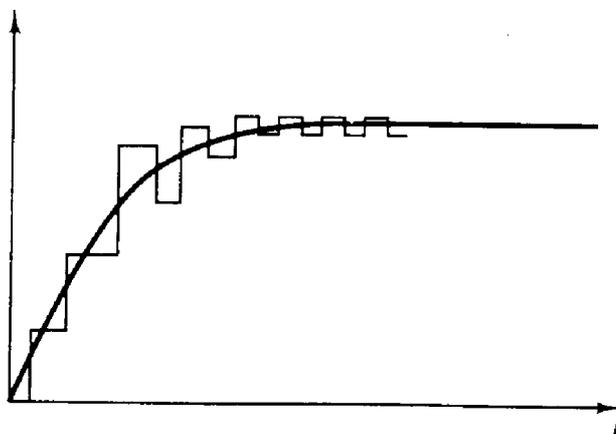


Figura 4.29: Desempeño de ΔM adaptiva.

Vimos que un tamaño de paso muy grande provoca ruido granular y un tamaño de paso muy pequeño produce distorsión de sobrecarga. Esto significa que una buena elección de Δ es un valor de compromiso, aunque en algunos casos el desempeño del mejor valor de compromiso (por ejemplo, aquel que minimice la distorsión media cuadrática) no es satisfactorio. Una solución que trabaja bien en esos casos es cambiar el tamaño del paso de acuerdo a los cambios en la entrada. Si la entrada tiende a cambiar rápidamente, el tamaño del paso se elige grande tal que la salida pueda seguir a la entrada también rápidamente y no resulte distorsión de sobrecarga. Cuando la entrada es de variación lenta, el tamaño del paso se cambia a un valor pequeño para prevenir el ruido granular. Tales cambios en el tamaño del paso se muestran en la figura 4.29.

Para cambiar adaptativamente el tamaño del paso es necesario diseñar un mecanismo para reconocer pendientes grandes y pequeñas. Si la pendiente de la entrada es pequeña, la salida del cuantizador \hat{Y}_n alterna entre valores de $+\Delta$ y $-\Delta$, como mostrado en la figura 4.29. Este es el caso cuando la principal fuente de ruido es el granular y es necesario disminuir el tamaño del paso. Sin embargo, en el caso de

sobrecarga de pendiente, la salida no puede seguir a la entrada rápidamente y la salida del cuantizador será una secuencia de pasos decrecientes o incrementales. Por lo anterior puede concluirse que el signo de dos \hat{Y}_n sucesivas es un buen criterio para cambiar el tamaño del paso. Si las dos salidas sucesivas tienen el mismo signo, el tamaño del paso debería incrementarse, y si son de signos opuestos, debería disminuirse.

Una regla particularmente simple para cambiar el tamaño del paso está dada por

$$\Delta_n = \Delta_{n-1} K^{\epsilon_n \times \epsilon_{n-1}}$$

donde ϵ_n es la salida del cuantizador antes del escalamiento por el tamaño del paso y K es una constante mayor que uno. Se ha verificado que en el rango de 20-60 kbits/seg, con una elección de $K = 1.5$, el desempeño de un sistema ΔM adaptivo es de 5 a 10 dB mejor que el de un sistema ΔM simple cuando se aplica a fuentes de voz.

4.7 Técnicas de análisis-síntesis

En contraste con la codificación de forma de onda, las técnicas de análisis - síntesis son métodos basados en un modelo para el mecanismo que produce la señal. Los parámetros del modelo que son extraídos de la señal fuente se cuantizan, se codifican y se transmiten al receptor. En el receptor y en base a la información recibida se sintetiza el mismo modelo y se utiliza para generar una salida similar a la señal original. Estos sistemas se utilizan principalmente para codificación de voz y en esta sección se discutirá brevemente uno de ellos conocido como *codificación lineal predictiva* (LPC).

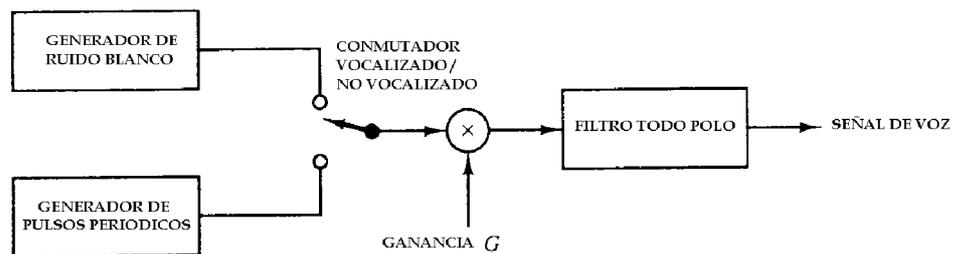


Figura 4.30: Modelo del mecanismo de generación de la voz.

La voz se produce como resultado de la excitación del tracto vocal por las cuerdas vocales. Este mecanismo puede modelarse como un filtro variante en el tiempo (el tracto vocal) excitado por un generador de señales. El tracto vocal es una combinación de la garganta, la boca, la lengua, los labios, la nariz y el cambio de forma durante la generación de la voz, y es modelado en consecuencia como un sistema variante en el tiempo. Las propiedades de la señal de excitación dependen fuertemente del tipo de sonido de la voz, ya sea *vocalizado* o *no vocalizado*. Para sonidos vocalizados la excitación puede modelarse como una secuencia periódica de impulsos de frecuencia f_0 , cuyo valor dependerá del locutor. El recíproco de esta frecuencia, $\frac{1}{f_0}$ se denomina *período de pitch*. Para sonidos no vocalizados la excitación se modela adecuadamente como un ruido blanco. Este modelo se muestra en la figura 4.30. El filtro del tracto vocal se modela usualmente con un filtro todo polo descrito por la siguiente ecuación a diferencias:

$$X_n = \sum_{i=1}^p a_i X_{n-i} + Gw_n$$

donde w_n es la secuencia de entrada (ruido blanco o impulsos), G es un parámetro de ganancia, $\{a_i\}$ son los coeficientes del filtro, y p es el número de polos del filtro. El proceso w_n que representa aquella parte de X_n no contenida en las p muestras previas se denomina *proceso de innovación*. Las señales de voz son muy aproximadamente estacionarias por cortos períodos de tiempo, del orden de 20-30 mseg. Este comportamiento característico se obtiene de la observación que el tracto vocal no puede cambiar instantáneamente. En consecuencia, sobre intervalos de 20-30 mseg. los coeficientes del filtro todo polo pueden suponerse fijos. En el codificador entonces se dispone del almacenamiento de 20-30 mseg. de voz con lo cual se estiman los

parámetros del modelo $\{a_i\}$, el tipo de señal de excitación (ruido blanco o impulsos), el período de pitch (si es un sonido vocalizado) y la ganancia G .

Para discutir este procedimiento es posible suponer que la señal de voz se filtra a 3 kHz y se muestrea a 8000 muestras/seg. Las muestras de señal se subdividen en bloques de 160 muestras, correspondientes a intervalos de 20 msec. $\{x_n, 1 \leq n \leq 160\}$ es la secuencia de muestras de cada bloque. El codificador debe medir los parámetros del modelo a ser transmitidos al receptor.

Para determinar los coeficientes del filtro en el codificador se utiliza predicción lineal. Un predictor lineal de orden p es un filtro digital todo ceros con entrada $\{x_n\}$ y salida

$$\hat{x}_n = \sum_{k=1}^p a_k x_{n-k} \quad \text{para } 1 \leq n \leq N$$

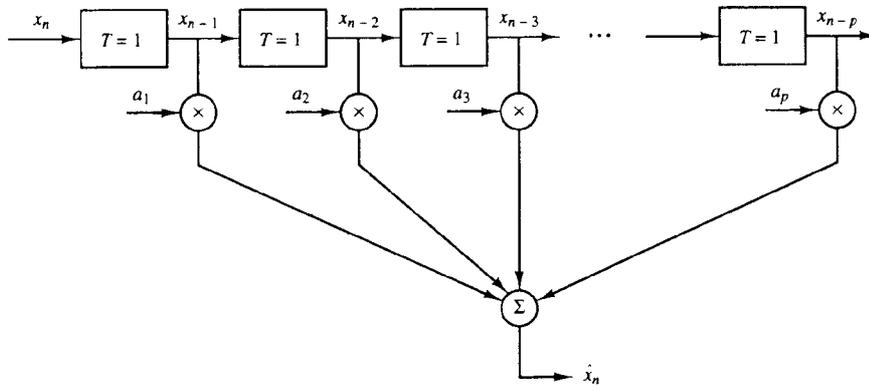


Figura 4.31: Diagrama en bloques funcional de predicción lineal.

donde se supuso que fuera del intervalo de interés $x_n = 0$. La figura 4.31 ilustra con un diagrama en bloques el proceso de predicción. La diferencia entre la muestra de voz real x_n y el valor predicho \hat{x}_n constituye el error de predicción e_n , o sea

$$e_n = x_n - \hat{x}_n = x_n - \sum_{k=1}^p a_k x_{n-k}$$

Para extraer tanta información cuanto sea posible de los valores previos de x_n se eligen los coeficientes $\{a_n\}$ de forma que el error cuadrático promedio, o sea

$$\xi_p = \frac{1}{N} \sum_{n=1}^N e_n^2 = \frac{1}{N} \sum_{n=1}^N \left(x_n - \sum_{k=1}^p a_k x_{n-k} \right)^2$$

sea mínimo. Diferenciando ξ_p con respecto a cada uno de los coeficientes del filtro $\{a_i\}$ e igualando la derivada a cero se obtiene un conjunto de ecuaciones lineales en los coeficientes del filtro, o sea

$$\frac{1}{N} \sum_{n=1}^N x_n x_{n-i} = \frac{1}{N} \sum_{n=1}^N \sum_{k=1}^p a_k x_{n-i} x_{n-k} \quad \text{para } 1 \leq i \leq p$$

Como se supuso que fuera del intervalo de estacionaridad $1 \leq n \leq N$, $x_n = 0$, es posible escribir la relación anterior como

$$\begin{aligned} \frac{1}{N} \sum_{n=-\infty}^{\infty} x_n x_{n-i} &= \frac{1}{N} \sum_{n=-\infty}^{\infty} \sum_{k=1}^p a_k x_{n-i} x_{n-k} \\ &= \sum_{n=-\infty}^{\infty} a_k \left[\frac{1}{N} \sum_{k=1}^p x_{n-i} x_{n-k} \right] \end{aligned}$$

Definiendo ahora $\hat{R}_i = \frac{1}{N} \sum_{n=-\infty}^{\infty} x_n x_{n-i}$ es posible escribir la ecuación anterior como

$$\hat{R}_i = \sum_{k=1}^p a_k \hat{R}_{i-k} \quad \text{para } 1 \leq i \leq p$$

la cual define las ecuaciones de *Youle-Walker*. Utilizando una notación matricial se puede reescribir

$$\mathbf{r} = \hat{\mathbf{R}}\mathbf{a}$$

donde \mathbf{a} es el vector de los coeficientes del predictor lineal, $\hat{\mathbf{R}}$ es una matriz de $p \times p$ cuyo elemento (i, j) -ésimo es \hat{R}_{i-j} , y \mathbf{r} es un vector cuyos componentes son los \hat{R}_i . Puede verificarse fácilmente de la definición que $\hat{R}_i = \hat{R}_{-i}$, de forma que $\hat{\mathbf{R}}$ es una matriz simétrica. También es obvio que todos los elementos de $\hat{\mathbf{R}}$ que están sobre la diagonal o subdiagonales son iguales. Tal matriz se denomina *Toeplitz* y asociada a ella existen algoritmos recursivos eficientes para calcular su inversa. Uno de esos métodos es el bien conocido *algoritmo de Levinson-Durbin*.

Con la elección óptima de los coeficientes del predictor el error cuadrático promedio mínimo toma la forma

$$\xi_p^{min} = \hat{R}_0 - \sum_{k=1}^p \hat{R}_k$$

De acuerdo al modelo de producción de la voz

$$\xi_p^{min} = \frac{1}{N} \sum_{n=1}^N \left[x_n - \sum_{k=1}^p a_k x_{n-k} \right]^2 = G^2 \frac{1}{N} \sum_{n=1}^N w_n^2$$

Normalizando la secuencia de excitación $\{w_n\}$ tal que $\frac{1}{N} \sum_{n=1}^N w_n^2 = 1$, obtenemos el valor de la ganancia del predictor: $G = \sqrt{\xi_p^{min}}$.

La estimación del tipo de excitación (impulsos o ruido) y la estimación del período de pitch en el caso de la excitación de tren de impulsos puede lograrse mediante varios algoritmos. Una aproximación simple es transformar los datos de voz al dominio frecuencia y procurar picos pronunciados en el espectro de la señal. Si el espectro exhibe picos en alguna frecuencia fundamental f_0 , la excitación se define como un tren de pulsos de período $\frac{1}{f_0}$. Si el espectro de las muestras de voz no exhibe picos, la excitación se define como ruido blanco.

Los coeficientes del filtro de predicción, la ganancia, la información de vocalización y el pitch se cuantizan y transmiten al receptor por cada bloque de voz muestreada. La señal de voz se sintetiza a partir de esos parámetros utilizando el modelo del sistema como ilustrado en la figura 4.30. Típicamente la información de vocalización requiere 1 bit, la frecuencia de pitch se representa mediante 6 bits, el parámetro de ganancia puede representarse mediante 5 bits utilizando expansión logarítmica y los coeficientes de predicción requieren de 8-10 bits por coeficiente. En base a la codificación por predicción lineal, la voz puede comprimirse a tasas tan bajas como 2400 bits/seg. Es posible también utilizar cuantización vectorial cuando se cuantizan los parámetros LPC. Esto reduciría aún más la tasa de transmisión. En contraste, PCM aplicada a la voz tiene una tasa de transmisión de 56000 bits/seg.

LPC se utiliza ampliamente en la codificación de voz para reducir el ancho de banda. Mediante cuantización vectorial de los parámetros LPC puede obtenerse buena calidad de voz a tasas de alrededor de

4800 bits/seg. Una versión de LPC con cuantización vectorial ha sido adoptada como un estándar para la compresión de voz en sistemas telefónicos celulares móviles. La codificación eficiente de voz es un área muy activa de investigación y se espera aún mayores reducciones en la tasa de transmisión en los codificadores comerciales en los próximos años.

4.8 Transmisión y grabación de audio digital (*)

Las señales de audio constituyen una gran parte de las comunicaciones diarias. Hoy en día miles de estaciones de radio emiten señales de audio en forma analógica. La calidad de radiodifusión de señales de voz es generalmente aceptable siempre que la voz sea inteligible. Por otro lado, la calidad de las señales de música que se difunden sobre radio AM es de baja calidad relativa debido a que el ancho de banda de la señal transmitida está restringido por las regulaciones. La radiodifusión FM de señales analógicas provee mayor fidelidad utilizando un ancho de banda de canal significativamente mayor. Es esperable que, en el futuro próximo, la radiodifusión comercial de señales de audio sea convertida también a la forma digital.

En la transmisión de señales de audio sobre canales telefónicos, la conversión de transmisión analógica a digital, la cual ha sido llevada a cabo en las últimas tres décadas, está casi completa. A continuación, se describirán algunos de los desarrollos actuales en la codificación de señales de audio para la transmisión telefónica.

La industria del entretenimiento ha experimentado los cambios más dramáticos y se ha beneficiado de la conversión de las señales de audio a la forma digital. El desarrollo del disco compacto (CD) y el grabador de audio digital han vuelto los sistemas de grabación analógica previous basicamente obsoletos. Se utilizará el sistema de reproducción de CD como un caso de estudio de métodos sofisticados de codificación/decodificación de fuente y canal que se han desarrollado en los últimos años.

4.8.1 Audio digital en sistemas de transmisión telefónica (*)

Casi toda la transmisión de señales de voz sobre canales telefónicos es actualmente digital. La codificación de señales de voz para su transmisión sobre canales telefónicos ha sido un tópico de intensas investigaciones sobre los últimos 50 años y continúa siendolo todavía. En ese período se ha desarrollado una amplia variedad de métodos de codificación de fuente/voz que se utilizan aún en nuestros días.

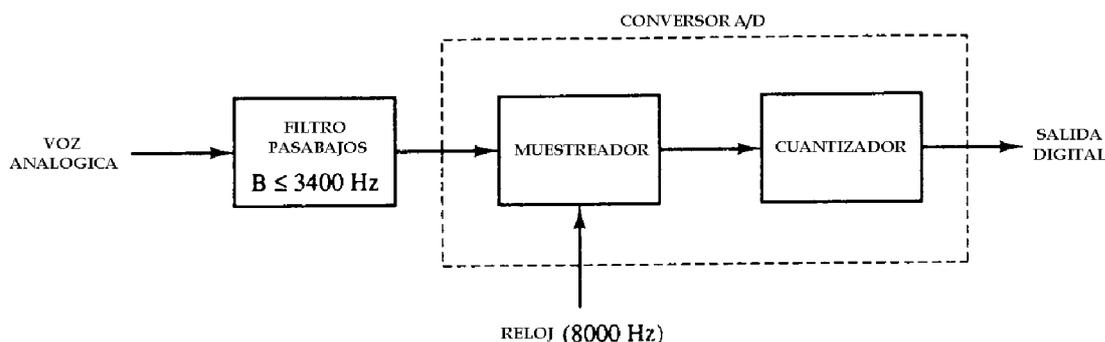


Figura 4.32: Conversión analógica a digital de señales de voz.

La configuración general de un codificador de voz se muestra en la figura 4.32. Dado que el contenido de frecuencias de la señal de voz está limitado a 3400 Hz, la señal se pasa primero por un filtro pasabajos antialiasing y luego se muestrea. Para asegurar que el aliasing es despreciable se elige una frecuencia de muestreo de 8000 Hz o mayor. Las muestras analógicas se cuantizan y se representan en forma digital para su transmisión, sobre el canal telefónico.

PCM y DPCM se utilizan ampliamente como métodos de codificación para transmisión digital de la voz. La compresión logarítmica con $\mu = 255$ se utiliza frecuentemente para lograr la cuantización no uniforme. La velocidad de transmisión típica para PCM es de 64 kbits/seg mientras que para DPCM es de 32 kbits/seg.

La codificación y decodificación de PCM y DPCM se efectúan generalmente en la central de conmutación telefónica (central office), donde se conectan las líneas de cada abonado telefónico en un área geográfica común al sistema de transmisión. Las señales de voz codificadas en PCM o DPCM son transmitidas de una

central de conmutación a otra en forma digital sobre lo que se denomina líneas troncales que son capaces de transportar señales de voz digitalizadas de varios abonados. El método para la transmisión simultánea de varias señales sobre un mismo canal de comunicación se denomina *multiplexado*. En el caso de la transmisión PCM o DPCM, las señales de diferentes abonados se multiplexan en tiempo, de ahí el nombre de *multiplexado por división en tiempo* (TDM). En TDM, se selecciona un intervalo de tiempo T_f como un cuadro (frame). Cada cuadro se subdivide en N subintervalos de duración T_f/N , donde N corresponde al número de usuarios que harán uso común del canal de comunicación. Luego, a cada abonado que desee utilizar el canal para transmisión se le asigna un subintervalo dentro de cada cuadro. En PCM cada usuario transmite una muestra de 8 bits en cada subintervalo.

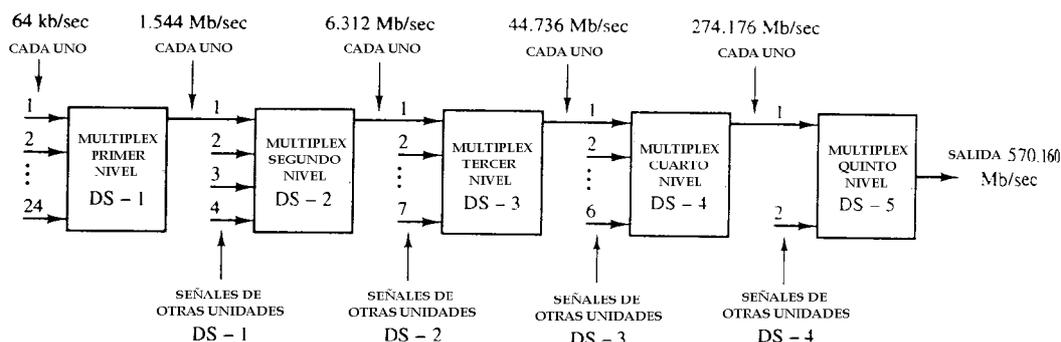


Figura 4.33: Jerarquía digital para sistemas de comunicaciones telefónicas.

En la transmisión digital de voz sobre líneas telefónicas via PCM existe una jerarquía TDM estándar que se ha establecido para acomodar abonados múltiples. En el primer nivel de la jerarquía TDM se multiplexan 24 señales de abonado en una única secuencia de 1.544 Mb/s (24 × 64 kbits más algunos bits adicionales con propósitos de control). La señal combinada TDM resultante se denomina usualmente un canal DS-1. En el segundo nivel de TDM, se multiplexan cuatro canales DS-1, para formar un canal DS-2, con una velocidad de 6.312 Mb/s. En el tercer nivel de la jerarquía, se combinan siete canales DS-2 mediante TDM para producir un canal DS-3, el cual tiene una velocidad de 44.736 Mb/s. Más allá de DS-3, existen dos niveles más de la jerarquía TDM, como se ilustra en la figura 4.33.

En los sistemas celulares de radio móvil para la transmisión de señales de voz, el ancho de banda disponible por usuario es pequeño y no puede soportar las altas velocidades requeridas por los métodos de codificación de forma de onda tal como PCM y DPCM. Para esta aplicación, se utiliza el método de análisis - síntesis con base en LPC para estimar el conjunto de parámetros del modelo a partir de segmentos cortos de la señal de voz. Los parámetros del modelo de la voz se transmiten sobre el canal utilizando cuantización vectorial. De esta forma, puede lograrse con LPC una velocidad de transmisión en el rango de 4800-9600 bits/seg.

En los sistemas celulares de comunicaciones móviles, la estación base en cada celda sirve como interfaz para el sistema telefónico terrestre. La compresión LPC de voz se requiere solamente para la transmisión de radio entre el usuario móvil y la estación base en cualquier celda. En la interfaz de la estación base, la señal de voz codificada por LPC se convierte en analógica y se remuestrea y digitaliza utilizando PCM o DPCM para la transmisión sobre el sistema telefónico terrestre. En consecuencia, es posible notar que en la transmisión de un abonado móvil a un abonado fijo existirán dos tipos diferentes de codificación analógica digital, mientras que en la comunicación de voz entre dos móviles servidos por estaciones base diferentes existirán cuatro conversiones entre los dominios analógicos y digital.

4.8.2 Grabación de audio digital (*)

Historicamente, la grabación de audio se volvió una realidad con la invención del fonógrafo durante la segunda mitad del siglo 19. El fonógrafo ha tenido una vida de aproximadamente 100 años antes de ser suplantado por el disco compacto, el cual fué introducido en 1982. Durante ese período de 100 años se ha presenciado la introducción de una amplia variedad de métodos de almacenamiento, el más popular de los cuales ha probado ser el larga duración (LP), introducido en 1948. Los LP proveían un método de almacenamiento de audio analógico de calidad relativamente alta.

A pesar de su amplia aceptación y popularidad, los métodos de almacenamiento de audio analógicos tienen cierto número de limitaciones, incluyendo un rango dinámico limitado (típicamente de alrededor de 70

dB) y una relación señal a ruido relativamente baja (típicamente de alrededor de 60 dB). En comparación, el rango dinámico de una orquesta musical está en el rango de 100 a 120 dB. Esto significa que, para almacenar la música en forma analógica, el rango deberá ser reducido para ser compatible con el del sistema de almacenamiento. De otra forma en niveles de música bajos el ruido sería audible y si se desea evitar eso, la saturación ocurrirá en niveles de música altos. La compresión del rango dinámico durante la grabación se realiza usualmente controlando el volumen de la señal de entrada.

La grabación de audio digital y su posterior reproducción permiten mejorar la fidelidad de la música grabada incrementando el rango dinámico y la relación señal a ruido. Además, la grabación digital es generalmente más durable y no se deteriora con las reproducciones, como en las grabaciones analógicas. Se describirá más abajo un sistema CD como un ejemplo de un sistema de audio digital comercialmente exitoso. La tabla siguiente provee una comparación de algunas especificaciones importantes de una grabación de LP y de un sistema CD. Las ventajas de este último son claramente evidentes.

Especificación	Grabación LP	Sistema CD
Resp. en Frecuencia	30 Hz - 20 kHz	20 Hz - 20 kHz
Rango dinámico	70 dB @ 1 kHz	> 90 dB
SNR	60 dB	> 90 dB
Distorsión armónica	1-2 %	0.005 %
Durabilidad	Resp. en frec. se degrada con reproduc.	Permanente
Vida útil	500-600 horas	5000 horas

Desde el punto de vista de sistemas, el sistema CD involucra la mayoría de los elementos de un sistema de comunicaciones digitales moderno. Esto incluye conversión analógico-digital (A/D) y digital-analógica (D/A), interpolación, modulación/demodulación y codificación/decodificación de canal. Un diagrama en bloques general de los elementos de un sistema CD de audio digital se ilustra en la figura 4.34. Se describirán las principales características del codificador y decodificador de fuente.

Las dos señales de audio de los micrófonos L y R en el estudio de grabación o sala de conciertos se muestrean y digitalizan pasándolas por un conversor A/D. La banda de frecuencias audibles del sonido está limitada a aproximadamente 20 kHz. En consecuencia, la frecuencia de Nyquist correspondiente es de 40 kHz. Para permitir una banda de salvaguarda y prevenir el aliasing, la frecuencia de muestreo en un sistema CD ha sido seleccionada en 44.1 kHz. Esta frecuencia es compatible con el equipamiento de grabación de video que se utiliza comúnmente para la grabación digital de señales de audio sobre cinta magnética.

Las muestras de las señales L y R se cuantizan utilizando PCM uniforme con 16 bits/muestra. De acuerdo a la fórmula de SQNR, la cuantización uniforme de 16 bits resulta en una SQNR de más de 90 dB. Además, la distorsión armónica alcanzada es de 0.005 %. Los bytes PCM de la grabación digital se codifican para protegerlos contra errores en el proceso de lectura y luego se pasan al modulador.

En el modulador se adiciona información de control digital y display, incluyendo una tabla de los contenidos del disco. Esta información adicional permite la programación del reproductor de CD. El tipo de modulación y demodulación y codificación de canal se describirán en los Capítulos 5, 6 y 8.

Se almacena ópticamente la señal digital del modulador utilizando un láser en la superficie de un disco plástico que está revestido con un material fotoresistente. De esto resulta un disco patrón que se utilizará para producir CD por medio de una serie de procesos que en definitiva convierten la información en marcas delgadas (pits) en un disco plástico. El disco se reviste con aluminio reflectivo y luego con una capa protectora.

En el reproductor de CD se utiliza un láser para barrer ópticamente un surco sobre el disco a una velocidad constante de 1.25 m/seg y de esa forma se lee la señal grabada digitalmente. Luego, las señales L y R se demodulan y pasan a través de un decodificador de canal. La señal de audio digital se convierte a una señal de audio analógica mediante un conversor D/A.

Las señales de audio digital L y R en el conversor D/A tienen una precisión de 16 bits. En principio, la conversión digital a analógica de las dos señales de 16 bits con una frecuencia de muestreo de 44.1 kHz es relativamente simple. Sin embargo, la implementación práctica de un conversor D/A de 16 bits es muy cara. Por otro lado, es posible conseguir fácilmente conversores D/A de 12 bits (o menos) muy fácilmente. El problema es la elaboración de un método para conversión D/A que emplee baja resolución y, en consecuencia, un conversor D/A de bajo costo, mientras que mantenga la precisión de 16 bits en la señal de audio digital.

La solución práctica a este problema es expandir el ancho de banda de la señal de audio digital mediante *sobremuestreo* utilizando *interpolación* y *filtrado digital* antes de la conversión analógica. Una aproximación

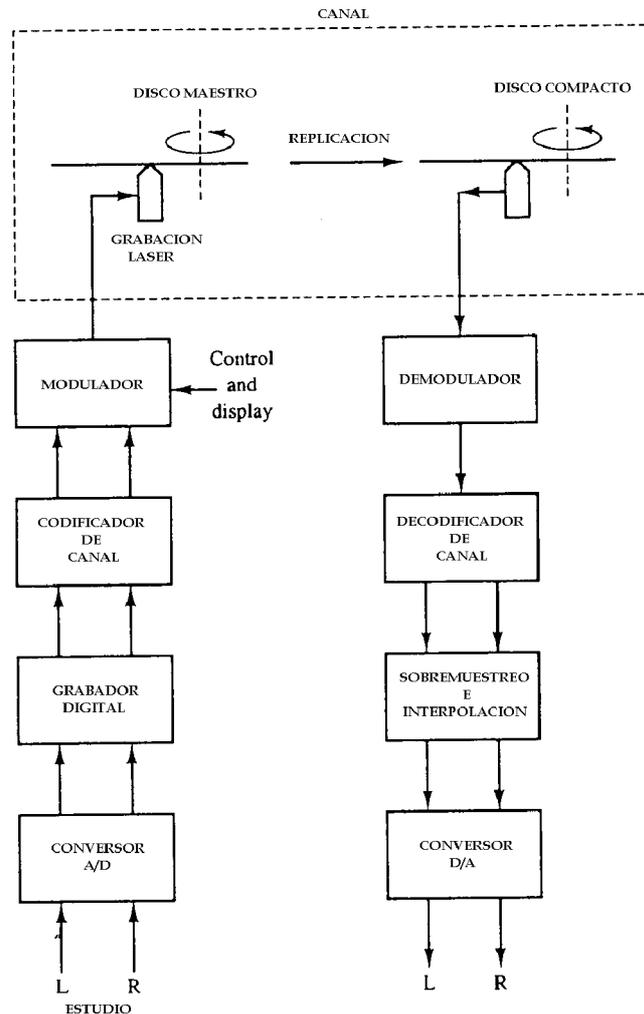


Figura 4.34: Sistema de audio digital CD.

a la solución básica se muestra en el diagrama de bloques de la figura 4.35. Las señales de audio digital L y R de 16 bits se muestrean a una frecuencia superior por un factor múltiplo U , insertando $U - 1$ ceros entre muestras sucesivas de la señal de 16 bits. Este proceso incrementa efectivamente la frecuencia de muestreo a $U \times 44.1$ kHz. Las señales resultantes L y R se filtran luego mediante un filtro FIR digital que produce una salida de alta velocidad y alta resolución. La combinación del rellenado de ceros y el filtrado es un método práctico de realizar un *interpolador digital*. El filtro FIR se diseña para tener fase lineal y un ancho de banda de aproximadamente 20 kHz. Su propósito es eliminar las imágenes del espectro creadas por el proceso de rellenado de ceros y se suele denominar *filtro anti-imágen*.

Observando las señales de audio digital de alta velocidad y precisión L y R de la salida del filtro FIR es posible concluir que las muestras sucesivas son aproximadamente iguales, difiriendo solamente en los bits menos significativos. En consecuencia, es posible representar muestras sucesivas de las señales de audio digital mediante diferencias entre muestras y, de esta forma, reducir su rango dinámico. Si el aumento de la frecuencia de muestreo por un factor U es suficientemente grande puede utilizarse modulación Delta para reducir la salida cuantizada a una precisión de 1 bit/muestra. De esta forma, el conversor D/A se simplifica considerablemente. Un factor de incremento $U = 256$ es el más frecuentemente elegido en la práctica. Esto eleva la frecuencia de muestreo a 11.2896 MHz.

Es necesario tener en cuenta la configuración general para el sistema de modulación Delta convencional como se muestra en la figura 4.36. Se eliminará el integrador del decodificador a la entrada del modulador. Esto tiene dos efectos. Primero, esto pre-enfatiza las bajas frecuencias de la señal de entrada y, de esta forma, se incrementa la correlación de la señal. Segundo, esto simplifica el decodificador del modulador

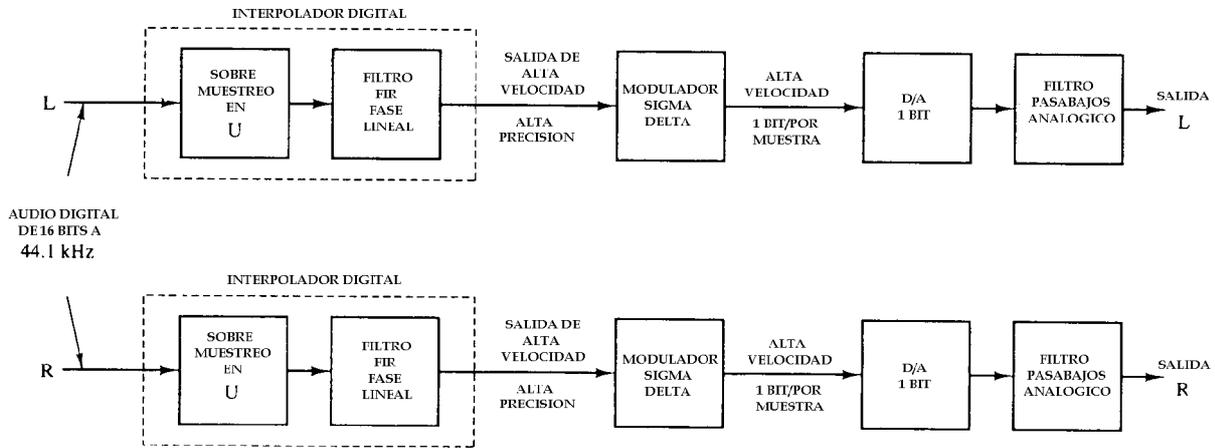


Figura 4.35: Conversión D/A mediante sobremuestreo e interpolación.

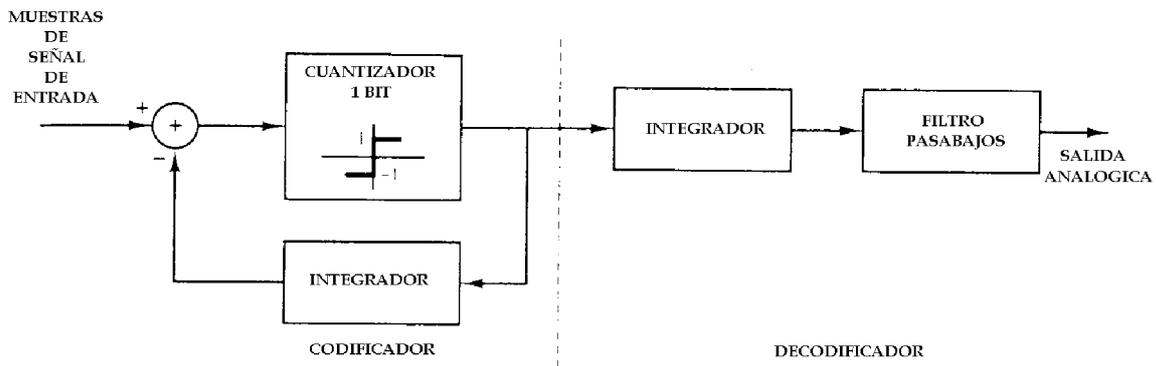


Figura 4.36: Sistema Δ convencional.

dato que el diferenciador (el sistema inverso) requerido en el decodificador es cancelado por el integrador. En consecuencia, el decodificador se reduce a un simple filtro pasabajos. Además, los dos integradores en el codificador pueden reemplazarse por un único integrador colocado antes que el cuantizador. El sistema resultante, que se muestra en la figura 4.37, se denomina *modulador sigma-delta* (SDM). La figura 4.38 ilustra un SDM que emplea un integrador digital simple (SDM de primer orden) con una función transferencia

$$H(z) = \frac{z^{-1}}{1 - z^{-1}}$$

De esta forma el SDM simplifica el proceso de conversión D/A ya que requiere un D/A de un bit seguido por un filtro analógico convencional (por ejemplo, un filtro Butterworth) para proveer protección antialiasing y suavización de señal. Los filtros analógicos de salida tienen una banda pasante de aproximadamente 20 kHz y eliminan de esta forma cualquier ruido por encima de la banda de señal deseada. En reproductores de CD modernos tanto el interpolador, como el SDM, el conversor D/A de 1 bit y el filtro pasabajos de suavización se implementan generalmente en un único circuito integrado.

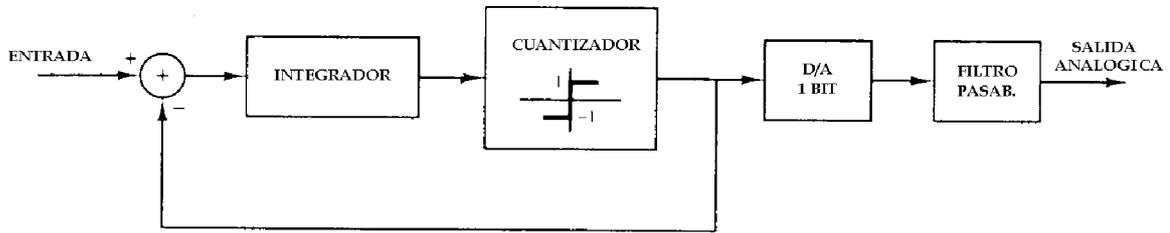


Figura 4.37: Configuración básica de SDM seguida por un convertor D/A de 1 bit y un filtro pasabajos de suavización.

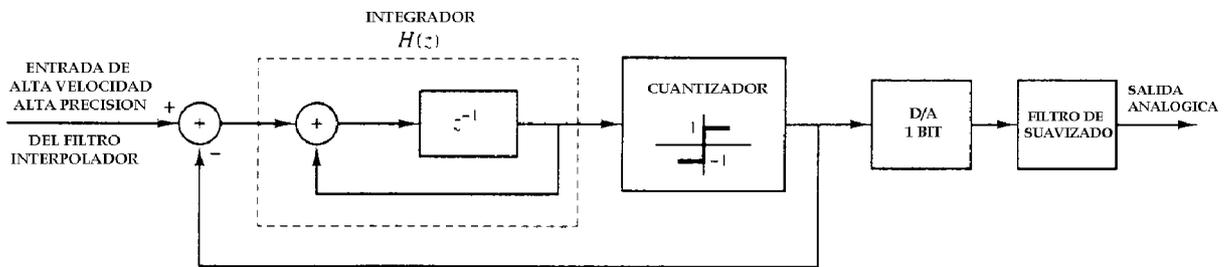


Figura 4.38: Convertor D/A SDM con un integrador de primer orden.

4.9 Apéndice: Entropía

Entropía es una medida de la incertidumbre de una variable aleatoria. En este apéndice se discutirá la noción de entropía y se desarrollarán algunas de sus propiedades más importantes. Se mostrará entonces que la entropía cuantifica la cantidad de información requerida para especificar la salida de una realización de una variable aleatoria. Finalmente se discutirá el método de máxima entropía, de gran utilidad en canales de comunicaciones, que sirve para caracterizar variables aleatorias cuando solo se conoce algunos parámetros estadísticos de ellas, como media y varianza.

4.9.1 Entropía de una variable aleatoria

Sea X una variable aleatoria discreta con alfabeto $\mathcal{S}_X = \{1, 2, \dots, K\}$ y pmf $P(X = k)$. Interesa cuantificar la incertidumbre de un evento $A_k = \{X = k\}$. Claramente, la incertidumbre de A_k es baja si la probabilidad de A_k es próxima a uno, y será alta si la probabilidad de A_k es pequeña. La medida de incertidumbre que satisface esos dos requerimientos es la siguiente

$$I(X = k) = \ln \frac{1}{P(X = k)} = -\ln P(X = k) \quad (4.8)$$

A partir de la figura 4.39 es posible notar que $I(X = k) = 0$ si $P(X = k) = 1$, y $I(X = k)$ aumenta cuando disminuye $P(X = k)$. La entropía de una variable aleatoria X se define como el valor esperado de la incertidumbre de sus realizaciones, o sea

$$H_X = E[I(X = k)] = \sum_{k=1}^K P(X = k) \ln \frac{1}{P(X = k)} = -\sum_{k=1}^K P(X = k) \ln P(X = k) \quad (4.9)$$

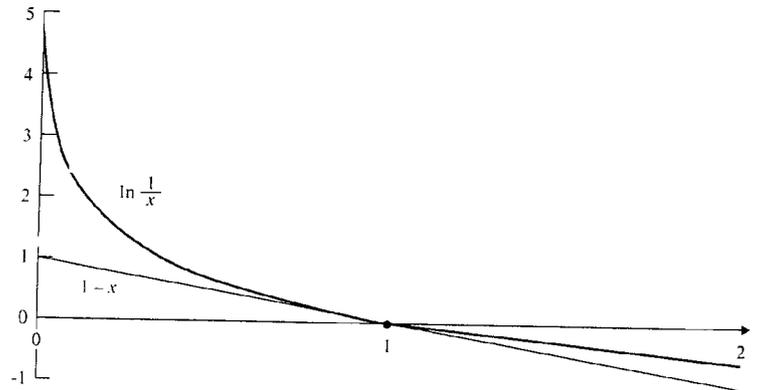


Figura 4.39: $\ln(1/x) \geq 1 - x$.

Notar también que en esta definición, $I(X)$ es una función de una variable aleatoria. La entropía tendrá unidades de *bits* cuando el logaritmo sea de base 2. Cambiar la base del logaritmo es equivalente a multiplicar la entropía por una constante, dado que $\ln(x) = \ln 2 \log_2 x$.

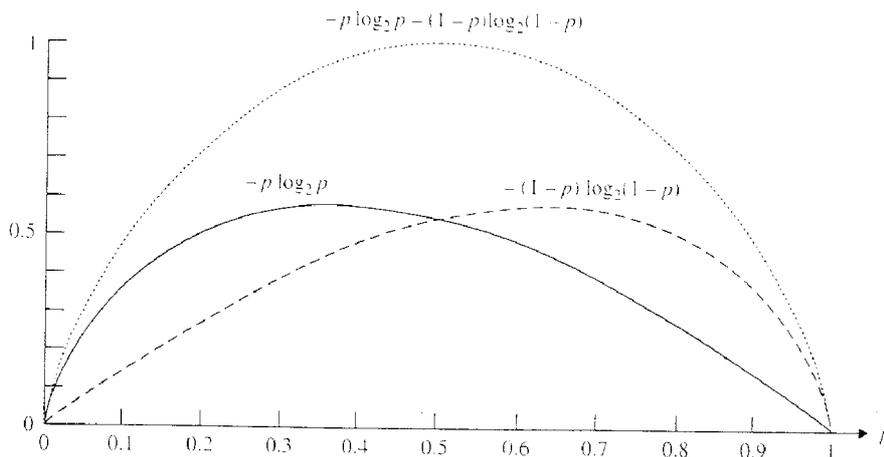


Figura 4.40: Entropía de una variable aleatoria binaria.

Ejemplo. Entropía de una variable aleatoria binaria. Sea $\mathcal{S}_X = \{0, 1\}$ y $p = P(X = 0) = 1 - P(X = 1)$. La figura 4.40 muestra $-p \ln p$, $-(1 - p) \ln(1 - p)$, y la entropía de la variable aleatoria $H_X = h(p) = -p \ln(p) - (1 - p) \ln(1 - p)$ en función de p . Notar que $h(p)$ es simétrica alrededor de $p = 1/2$ y que alcanza el máximo en $p = 1/2$. Notar también como las incertidumbres para los eventos $\{X = 0\}$ y $\{X = 1\}$ varía en forma complementaria: cuando $P(X = 0)$ es muy pequeña (o sea, alta incertidumbre) entonces $P(X = 1)$ está próxima a uno (o sea, alta certeza) y viceversa. De esta forma la mayor incertidumbre promedio ocurre cuando $P(X = 0) = P(X = 1) = 1/2$.

H_X puede interpretarse como la incertidumbre promedio que puede resolverse mediante la observación de X . Esto sugiere que si se está diseñando un experimento binario (por ejemplo una pregunta si/no) entonces la incertidumbre promedio que se resuelve se maximizará si las dos salidas del experimento se diseñan para tener igual probabilidad.

Ejemplo. Reducción de la entropía a través de información parcial. La representación binaria de una

variable aleatoria X toma valores del conjunto $\{000, 001, 010, \dots, 111\}$ con iguales probabilidades. Hallar la reducción en la entropía de X dado el evento $A = \{X \text{ comienza con } 1\}$.

La entropía de X será

$$H_X = -\frac{1}{8} \log \frac{1}{8} - \frac{1}{8} \log \frac{1}{8} - \dots - \frac{1}{8} \log \frac{1}{8} = 3 \text{ bits.}$$

El evento A implica que X está en el conjunto $\{100, 101, 110, 111\}$, de forma que la entropía de X dado A será

$$H_{X|A} = -\frac{1}{4} \log \frac{1}{4} - \dots - \frac{1}{4} \log \frac{1}{4} = 2 \text{ bits.}$$

De esta forma, la reducción de entropía será $H_X - H_{X|A} = 3 - 2 = 1$ bit.

Sean $\mathbf{p} = (p_1, p_2, \dots, p_K)$ y $\mathbf{q} = (q_1, q_2, \dots, q_K)$ dos pmf. La *entropía relativa* de \mathbf{q} con respecto a \mathbf{p} se define por

$$H(p; q) = \sum_{k=1}^K p_k \ln \frac{1}{q_k} - H_x = \sum_{k=1}^K p_k \ln \frac{p_k}{q_k} \quad (4.10)$$

La entropía relativa es no negativa, e igual a cero si y solo si $p_k = q_k$ para todo k , o sea

$$H(p; q) \geq 0 \quad \text{con igualdad si y solo si } p_k = q_k, \text{ para } k = 1, \dots, K. \quad (4.11)$$

Se utilizará esta propiedad varias veces en esta sección.

Para demostrar que la entropía relativa es no negativa se utilizará la desigualdad $\ln(1/x) \geq 1 - x$ (con igualdad si y solo si $x = 1$), como mostrado en la figura 4.39. La ecuación anterior entonces puede escribirse como

$$H(p; q) = \sum_{k=1}^K p_k \ln \frac{p_k}{q_k} \geq \sum_{k=1}^K p_k \left(1 - \frac{q_k}{p_k}\right) = \sum_{k=1}^K p_k - \sum_{k=1}^K q_k = 0.$$

Para que se verifique la igualdad en esta expresión debe ser $p_k = q_k$ para $k = 1, \dots, K$.

Sea X cualquier variable aleatoria con $\mathcal{S}_X = \{1, 2, \dots, K\}$ y pmf \mathbf{p} . Haciendo $q_k = 1/K$ en (4.11) se tendrá que

$$H(p; q) = \ln K - H_X = \sum_{k=1}^K p_k \ln \frac{p_k}{1/K} \geq 0,$$

lo cual implica que para cualquier variable aleatoria X con $\mathcal{S}_X = \{1, 2, \dots, K\}$,

$$H_X \leq \ln K \quad \text{con igualdad si y solo si } p_k = \frac{1}{K}, \text{ para } k = 1, \dots, K. \quad (4.12)$$

De forma que la máxima entropía alcanzable por la variable aleatoria X es $\ln K$, y este máximo se alcanza cuando todas las realizaciones son equiprobables.

La ecuación (4.12) muestra que la entropía de variables aleatorias con \mathcal{S}_X finito es siempre finita. Por otro lado, también muestra que cuando el tamaño de \mathcal{S}_X aumenta, la entropía puede disminuir sin límite. Puede mostrarse que algunas variables aleatorias de un alfabeto infinito tienen entropía finita.

Para variables aleatorias continuas se tiene que $P(X = x) = 0$ para todo x . En consecuencia, debido a (4.8), la incertidumbre para cada evento $\{X = x\}$ es infinita, y de (4.9) se concluye que la *entropía de*

variables aleatorias continuas es infinita. El ejemplo siguiente da una idea como puede usarse el concepto de entropía para variables aleatorias continuas.

Ejemplo. Entropía de una variable aleatoria continua cuantizada. Sea X una variable aleatoria continua que toma valores en el intervalo $[a, b]$. Es posible suponer que el intervalo $[a, b]$ se divide en un gran número de subintervalos K de longitud Δ . Sea $Q(X)$ el punto medio del subintervalo que contiene a X . Hallar la entropía de Q .

Sea x_k el punto medio del subintervalo k , entonces $P(Q = x_k) = P(X \in \text{subintervalo } k) = P(x_k - \Delta/2 < X < x_k + \Delta/2) \cong f_X(x_k)\Delta$, y de esta forma

$$\begin{aligned} H_Q &= \sum_{k=1}^K P(Q = x_k) \ln P(Q = x_k) \cong - \sum_{k=1}^K f_X(x_k)\Delta \ln(f_X(x_k)\Delta) \\ &= -\ln(\Delta) - \sum_{k=1}^K f_X(x_k) \ln(f_X(x_k))\Delta \end{aligned} \quad (4.13)$$

La ecuación anterior muestra que existe un compromiso entre la entropía de Q y el error de cuantización $X - Q(X)$. Cuando se disminuye Δ el error disminuye, pero la entropía aumenta sin límite, confirmando nuevamente que la entropía de una variable aleatoria continua es infinita.

En la expresión final para H_X de (4.13), cuando Δ se aproxima a cero, la primera parte tiende a infinito, pero la segunda parte se aproxima a una integral que puede ser finita en algunos casos. La *entropía diferencial* se define a partir de esa aproximación como

$$H_X = - \int_{-\infty}^{\infty} f_X(x) \ln f_X(x) dx = -E[\ln f_X(X)]$$

En la expresión anterior se reutilizó la notación de H_X sobreentendiendo que se está trabajando con la entropía diferencial de una variable aleatoria continua.

Ejemplo. Entropía diferencial de una variable aleatoria uniforme. La entropía diferencial para X uniforme en $[a, b]$ será

$$H_X = -E \left[\ln \frac{1}{b-a} \right] = \ln(b-a)$$

Ejemplo. Entropía de una variable aleatoria gaussiana. La entropía diferencial cuando X es una variable aleatoria Gaussiana de media cero y varianza σ^2 será

$$H_X = -E[\ln f_X(X)] = -E \left[\ln \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(X-m)^2}{2\sigma^2} \right] = \frac{1}{2} \ln(2\pi\sigma^2) + \frac{1}{2} = \frac{1}{2} \ln(2\pi e\sigma^2)$$

La función entropía y la entropía diferencial difieren en varios aspectos fundamentales. Más adelante se verá que la entropía de una variable aleatoria discreta tiene una interpretación operacional bien definida, como el número promedio de bits de información requeridos para especificar el valor de esa variable. La entropía diferencial no posee esta interpretación operacional. Además, la función entropía no cambia cuando se mapea la variable aleatoria X en otra Y mediante una transformación inversible. Nuevamente, la entropía diferencial no posee esta propiedad. Independientemente de ello, la entropía diferencial posee algunas propiedades útiles. La entropía diferencial aparece naturalmente en problemas que involucran reducción de entropía. Además, la entropía relativa de variables aleatorias continuas, definida como

$$H(f_X; f_Y) = \int_{-\infty}^{\infty} f_X(x) \ln \frac{f_X(x)}{f_Y(y)} dx, \tag{4.14}$$

no cambia bajo transformaciones inversibles.

4.9.2 Entropía como medida de información

Sea X una variable aleatoria discreta con $\mathcal{S}_X = \{1, 2, \dots, K\}$ y pmd $p_k = P(X = k)$. Es posible suponer que el experimento que produce X es realizado por Juan, que intenta comunicar la realización a María respondiendo a una serie de preguntas con respuestas de tipo si/no. El interés apunta a la caracterización del número promedio mínimo de preguntas requeridas para identificar X .

Ejemplo. Una urna contiene 16 bolitas: 4 están rotuladas con "1", 4 están rotuladas con "2", 2 están rotuladas con "3", 2 están rotuladas con "4" y las remanentes están rotuladas con "5", "6", "7" y "8". Juan toma una bolita de la urna aleatoriamente. Discutir cuales estrategias puede utilizar María para saber el rótulo de la bolita a través de una serie de preguntas si/no. Comparar el número promedio de preguntas realizadas con la entropía de X .

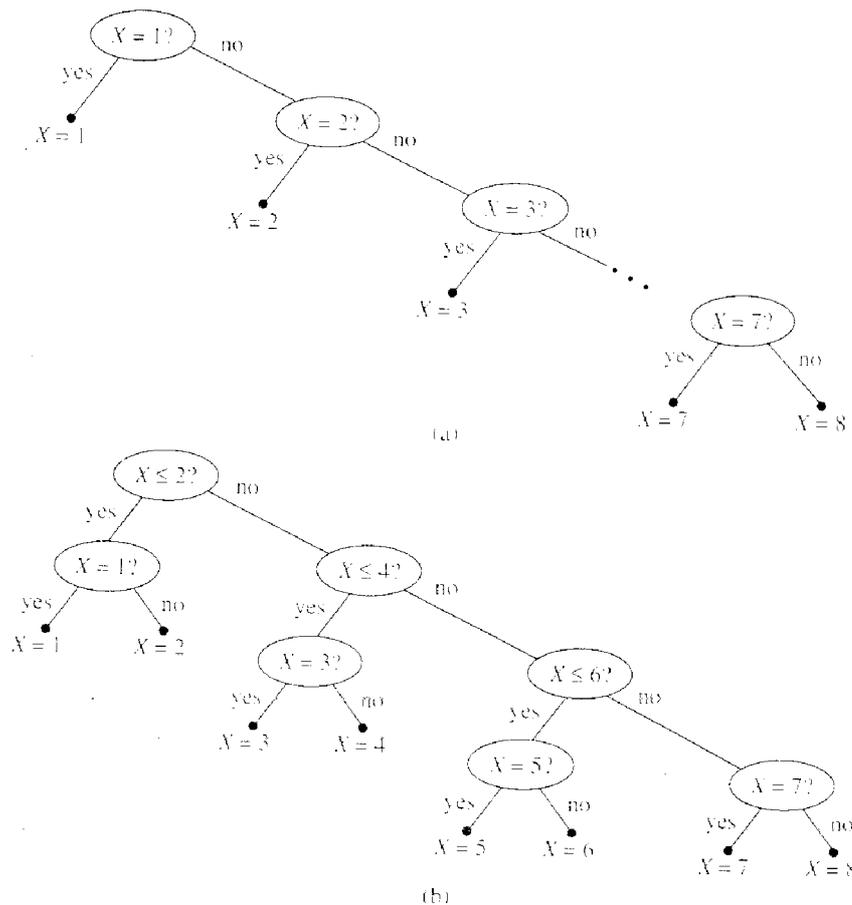


Figura 4.41: Dos estrategias para encontrar el valor de X a través de una serie de preguntas si/no.

Si X es la variable aleatoria asociada al rótulo sobre las bolitas, entonces $\mathcal{S}_X = \{1, 2, \dots, 8\}$ y la pmf será $\mathbf{p} = (1/4, 1/4, 1/8, 1/8, 1/16, 1/16, 1/16, 1/16)$. Se compararán las dos estrategias ilustradas en las figuras 4.41a y 4.41b.

La serie de preguntas en la figura 4.41a tiene en cuenta que la probabilidad de que $\{X = k\}$ disminuye con k . De esta forma es razonable preguntarse X es igual a 1?, X es igual a 2? y así sucesivamente hasta

que la respuesta sea sí. Si L es el número de preguntas hasta que la respuesta es afirmativa, entonces el número promedio de preguntas será

$$\begin{aligned} E[L] &= 1 \left(\frac{1}{4}\right) + 2 \left(\frac{1}{4}\right) + 3 \left(\frac{1}{8}\right) + 4 \left(\frac{1}{8}\right) + 5 \left(\frac{1}{16}\right) + 6 \left(\frac{1}{16}\right) + 7 \left(\frac{1}{16}\right) + 7 \left(\frac{1}{16}\right) \\ &= 51/16 \end{aligned}$$

La serie de preguntas en la figura 4.41b usa una observación realizada en ejemplos anteriores, tal que las preguntas sí/no deberían diseñarse para que las dos respuestas sean equiprobables. Las preguntas en la figura 4.41b satisfacen ese requerimiento. El número promedio de preguntas es

$$\begin{aligned} E[L] &= 2 \left(\frac{1}{4}\right) + 2 \left(\frac{1}{4}\right) + 3 \left(\frac{1}{8}\right) + 3 \left(\frac{1}{8}\right) + 4 \left(\frac{1}{16}\right) + 4 \left(\frac{1}{16}\right) + 4 \left(\frac{1}{16}\right) + 4 \left(\frac{1}{16}\right) \\ &= 44/16 \end{aligned}$$

De esta forma la segunda serie de preguntas logra un mejor desempeño.

Finalmente la entropía de X será

$$H_X = -\frac{1}{4} \log_2 \left(\frac{1}{4}\right) - \frac{1}{4} \log_2 \left(\frac{1}{4}\right) - \frac{1}{8} \log_2 \left(\frac{1}{8}\right) - \dots - \frac{1}{16} \log_2 \left(\frac{1}{16}\right) = 44/16$$

que coincide con el desempeño de la segunda serie de preguntas.

El problema de diseñar la serie de preguntas para identificar la variable aleatoria X es exactamente el mismo que el problema de codificar la salida de una fuente de información. Cada salida de la fuente de información es una variable aleatoria X , y la tarea del codificador es mapear cada salida posible en una cadena única de dígitos binarios. Es posible ver esta correspondencia teniendo en cuenta los árboles de la figura 4.41 e identificando cada respuesta sí/no con 0/1. La secuencia de ceros y unos desde el nodo superior al nodo terminal define entonces la cadena binaria ("palabra de código") para cada salida. Se concluye entonces que el problema de hallar la mejor serie de preguntas sí/no es el mismo que el de hallar el código binario que minimiza la longitud promedio de palabra de código.

En el resto de esta sección se discutirá aspectos asociados con teoría de información. El primero de ellos definirá que la longitud promedio de la palabra de código de cualquier código no puede ser menor que la entropía. El segundo resultado es que si la pmf de X consiste en potencias de 1/2 entonces existirá un árbol de codificación que alcanza la entropía. Finalmente, mediante la codificación de grupos de realizaciones de X es posible lograr una palabra de código de longitud promedio arbitrariamente próxima a la entropía. De esta forma, *la entropía de X representa el número mínimo promedio de bits requeridos para establecer la realización de X .*

Primero se mostrará que la longitud promedio de palabra de código para cualquier árbol de codificación no puede ser menor que la entropía. De la figura 4.41 es posible notar que el conjunto de longitudes $\{\ell_k\}$ asociadas a las palabras de código para cualquier árbol binario completo satisface

$$\sum_{k=1}^K 2^{-\ell_k} = 1. \quad (4.15)$$

Para ver esto, es posible extender el árbol a la misma profundidad que la palabra de código más larga, como ilustrado en la figura 4.42. Si se "poda" el árbol en un nodo de profundidad ℓ_k , se removerá una fracción $2^{-\ell_k}$ de los nodos en el pie del mismo. Notar que el dual también es cierto: si un conjunto de longitudes de palabra de código satisface la ecuación anterior, entonces es posible construir un árbol de codificación con esas longitudes.

Considerar luego la diferencia entre la entropía y $E[L]$ para cualquier árbol de codificación binario, o sea

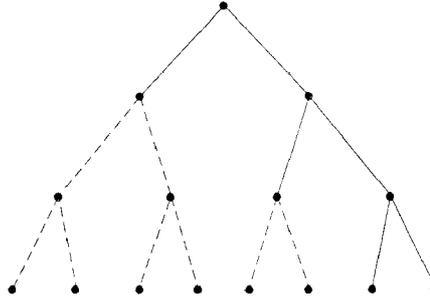


Figura 4.42: Extensión de un árbol de codificación a un árbol completo.

$$\begin{aligned}
 E[L] - H_X &= \sum_{k=1}^K \ell_k P(X = k) + \sum_{k=1}^K P(X = k) \log_2 P(X = k) \\
 &= \sum_{k=1}^K P(X = k) \log_2 \frac{P(X = k)}{2^{-\ell_k}}
 \end{aligned}$$

donde se ha expresado la entropía en bits. Esta ecuación es la entropía relativa de (4.10) con $q_k = 2^{-\ell_k}$. De esta forma con (4.11) se verifica que

$$E[L] \geq H_X \quad \text{con igualdad si y solo si } P(X = k) = 2^{-\ell_k} \quad (4.16)$$

Así, el número promedio de preguntas para cualquier árbol de codificación (y en particular el mejor árbol de codificación) no puede ser menor que la entropía de X . En consecuencia, es posible utilizar a la entropía H_X como una cota para verificar el desempeño de cualquier codificación.

La ecuación (4.16) implica también que si las realizaciones de X tienen todas probabilidades iguales a potencia de $1/2$, entonces es posible hallar un árbol de codificación que logra la entropía. Si $P(X = k) = 2^{-\ell_k}$, entonces se asigna la salida k a la palabra de código de longitud ℓ_k . Es posible mostrar que siempre se puede hallar un árbol de codificación con esas longitudes teniendo en cuenta que la suma de todas las probabilidades es igual a uno, y en consecuencia las longitudes de las palabras de código satisfacen (4.15). La ecuación (4.16) implica entonces que $E[L] = H_X$.

Se alcanzó entonces el objetivo de verificar que la entropía de una variable aleatoria X representa el número mínimo promedio de bits requeridos para identificar su valor. En relación a variables aleatorias continuas, estas suponen valores en un alfabeto infinito, de forma que se requerirán en general un número infinito de bits para especificar el valor de la variable aleatoria. De esta forma, la interpretación de entropía como el número promedio de bits requeridos para especificar una variable aleatoria implica inmediatamente que las variables aleatorias continuas tienen entropía infinita. Esto implica que cualquier representación de una variable aleatoria continua que utiliza un número finito de bits tendrá asociada inherentemente algún tipo de aproximación.

4.9.3 El método de máxima entropía

Sea X una variable aleatoria con $\mathcal{S}_X = \{x_1, x_2, \dots, x_K\}$ y pmf desconocida $p_k = P(X = k)$. Es posible suponer la estimación de la pmf de X dado el valor esperado de alguna función $g(X)$ de X , o sea

$$\sum_{k=1}^K g(x_k) P(X(x_k)) = c \quad (4.17)$$

Por ejemplo, si $g(X) = X$ entonces $c = E[g(X)] = E[X]$, y si $g(X) = (X - E[X])^2$ entonces $c = \text{var}[X]$. Claramente este problema es indeterminado dado que el conocimiento de esos parámetros no es suficiente para especificar la pmf unívocamente. El *método de máxima entropía* permite una solución aproximada a este problema procurando la pmf que maximiza la entropía sujeta a la restricción de (4.17).

Suponiendo llevar a cabo el problema de maximización utilizando multiplicadores de Lagrange, se obtiene

$$H_X + \lambda \left(\sum_{k=1}^K g(x_k) P(X=x_k) - c \right) = \sum_{k=1}^K P(X=x_k) \ln \frac{P(X=x_k)}{C e^{-\lambda g(x_k)}} \quad (4.18)$$

donde $C = e^c$. Notar que si $\{C e^{-\lambda g(x_k)}\}$ forma una pmf, entonces la expresión anterior es el valor negativo de la entropía relativa de esta pmf con respecto a \mathbf{p} . A partir de la ecuación (4.11) se concluye que (4.18) es siempre menor o igual a cero con igualdad si y solo si $P(X=x_k) = C e^{-\lambda g(x_k)}$. Se mostrará ahora que además esto conduce a la solución de máxima entropía.

Es posible suponer que la variable aleatoria X tiene pmf $p_k = C e^{-\lambda g(x_k)}$ donde C y λ se eligen de forma de satisfacer (4.17) y obviamente de forma que p_k sea una pmf. Entonces X tendrá una entropía

$$H_X = E[-\ln P(X)] = [-\ln C e^{-\lambda g(x_k)}] = -\ln C + \lambda E[g(X)] = -\ln C + \lambda c \quad (4.19)$$

Es posible comparar ahora la entropía de (4.19) con la de alguna otra pmf q_k que también satisfaga la restricción de (4.17). Considerar la entropía relativa de \mathbf{p} con respecto a \mathbf{q} , o sea

$$\begin{aligned} 0 \leq H(\mathbf{q}; \mathbf{p}) &= \sum_{k=1}^K q_k \ln \frac{q_k}{p_k} = \sum_{k=1}^K q_k \ln p_k + \sum_{k=1}^K q_k (-\ln C + \lambda g(x_k)) \\ &= -\ln C + \lambda c - H(q) = H_X - H(q) \end{aligned}$$

De esta forma $H_X \geq H(q)$ y \mathbf{p} alcanza la mayor entropía.

Ejemplo. Sea X una variable aleatoria con $\mathcal{S}_X = \{0, 1, \dots\}$ y valor esperado $E[X] = m$. Hallar la pmf que maximiza la entropía.

En este ejemplo, $g(X) = X$, de forma que

$$p_k = C e^{-\lambda k} = C \alpha^k$$

donde $\alpha = e^{-\lambda}$. Claramente X es una variable aleatoria geométrica con media $m = \alpha/(1-\alpha)$ de donde $\alpha = m/(m+1)$. De esto se obtiene que $C = 1 - \alpha = 1/(m+1)$.

Cuando se trabaja con variables aleatorias continuas, el método de máxima entropía maximiza la entropía diferencial dada por

$$- \int_{-\infty}^{\infty} f_X(x) \ln f_X(x) dx$$

El parámetro de información tiene la forma

$$c = E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (4.20)$$

La expresión de entropía relativa de (4.14) y la metodología empleada para variables aleatorias discretas pueden utilizarse en este caso para mostrar que la pdf $f_X(x)$ que maximiza la entropía diferencial entrará la forma

$$f_X(x) = C e^{-\lambda g(x)} \quad (4.21)$$

donde C y λ se eligen de forma de satisfacer la restricción y tal que la integral de esta última ecuación sea igual a uno.

Ejemplo. Es posible suponer que la variable aleatoria continua X tiene varianza conocida $\sigma^2 = E[(X-m)^2]$, donde la media no está especificada. Hallar la pdf que maximiza la entropía de X .

La ecuación (4.21) implica que la pdf tendrá la forma

$$f_X(x) = Ce^{-\lambda/(x-m)^2}$$

Es posible satisfacer el requerimiento de (4.20) haciendo que

$$\lambda = \frac{1}{2\sigma^2} \quad C = \frac{1}{\sqrt{2\pi\sigma^2}}$$

de forma que se obtiene una pdf Gaussiana con varianza σ^2 . Notar que la media m es arbitraria; o sea, cualquier elección de m produce una pdf que maximiza la entropía diferencial.
