- Área: TIC
- Asignatura:
Tecnología de la información

- Título

Teoría de la Información y Transmisión de Datos

Prof.: Ing. Daniel Esteban Vena

- Curso: 6to Año
- Año: 2007

Transmisión de Datos

Técnicas de Transmisión de la Información

Introducción:

El objetivo de un sistema de comunicaciones es hacer lo más eficiente posible la transmisión de la información de un punto a otro dentro de una red de comunicaciones.

El estudio de la teoría de la información puede ser dividida en 3 bloques a saber:

1) Como se mide la información.

2) Cual es la capacidad de un canal de comunicaciones para transferir información.

3) Los aspectos que hacen a la codificación, para que puedan ser utilizados a su máxima capacidad con una tasa de mínimo error.

Medidas de la Información:

La medida de la información esta relacionada con la posibilidad de que una fuente pueda elegir entre varios mensajes posibles.

Cada mensaje podrá tener asociada una probabilidad, de forma tal que cuanto mayor sea la posibilidad de que ese mensaje sea cierto, menor será la información para el usuario.

Por lo tanto, se define como la cantidad de información que el usuario recibe al serle entregado un mensaje, al logaritmo de la inversa de la probabilidad de ocurrencia del suceso contenido en él.

$$I(a) = \text{Log } 1/P(a)$$

Dependiendo de la base del logaritmo, le corresponde una unidad

$$I(a) = \text{Log}_2 \; 1/P(a) \quad \text{======>} \quad \text{Shannon}$$

$I(a) = \text{Log}_e\ 1/P(a)$ ======> NAT.

$I(a) = \text{Log}_{10}\ 1/P(a)$ ======> Hartley

Para sacar el Log. en base 2 N:

$$\text{Log}_2 N = \text{Log}_2 10 \times \text{Log}_{10} N = 3{,}32 \times \text{Log}_{10} N$$

Pero teniendo en cuenta que los medios informáticos trabajan en forma binaria, la probabilidad de un BIT es 0,5

$I_{(0)} = \text{Log}_2\ 1/0{,}5 = 1$ Shannon

$I_{(1)} = \text{Log}_2\ 1/0{,}5 = 1$ Shannon

La norma ISO- 2382 adopta al Shannon como unidad de información y al BIT como identificación del dígito binario.

<u>Entropía:</u>

Se denomina fuente de memoria nula a aquella en que los símbolos que emite son estadísticamente independientes. Suponemos que una fuente de memoria nula puede emitir distintos símbolos, cada uno de ellos con una probabilidad asociada.

Variable aleatoria es una función $I_{(X\ K)}$ donde para cada valor le corresponde un valor de probabilidad $P_{(XK)}$

$$I_{(XK)} = \text{Log}_2\ 1/P_{(XK)} = -\text{Log}_2\ P_{(XK)}$$

Por lo tanto la entropía de una fuente nula será:

$$H = -\sum_{k=1}^{n} P(xk).\log_2 P(xk) \quad [Shannon/símbolo]$$

Por ejemplo: Una fuente de memoria nula puede emitir tres símbolos X1, X2, X3 con probabilidades 1/2, 1/4, 1/8. La entropía de la fuente será:

$$H = -\left[1/2.\ \text{Log}_2\ 1/2 + 1/4.\text{Log}_2\ 1/4 + 1/8.\text{Log}_2\ 1/8\right]$$

Propiedades de la entropía

1) Continuidad:

Pequeños cambios en la probabilidad de un evento le corresponden pequeños valores de la entropía.

2) Simetría:

El orden en que se presentan los eventos no altera la entropía.

3) Aditiva:

Las entropías para un mismo sistema son aditivas.

4) Maximilidad:

Una fuente de memoria nula que puede emitir N símbolos con igual probabilidad tendrá un valor máximo de entropía.

$$H = \log_2 N \qquad [\text{ Shannon/símbolo}]$$

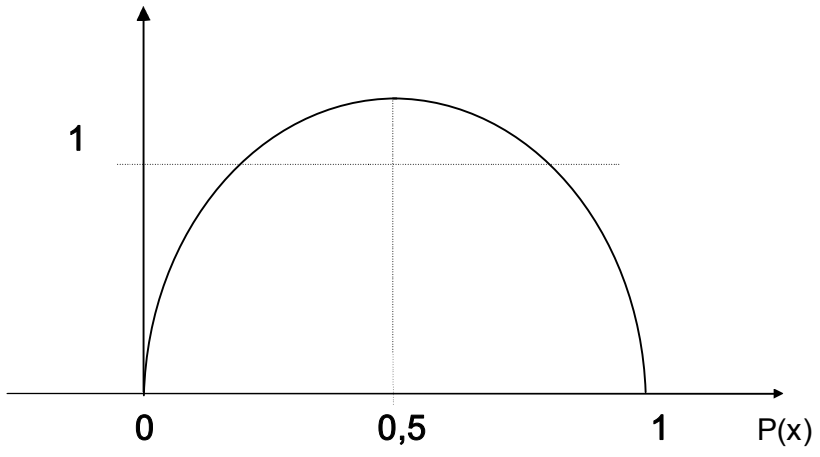
Si por ejemplo calculamos la entropía de una fuente binaria de memoria nula para la cual el 1 se representa con probabilidad X y el 0 con probabilidad (1-X)

$$H(x) = - [P(0) \,.\, \text{Log}_2 P(0) + P(1) \text{ Log}_2 P(1)]$$

Entonces

$$H(x) = - [(1-x) \,.\, \text{Log}_2 (1-x) + x \,.\, \text{Log}_2 x]$$

Graficando:



En consecuencia podemos afirmar que para una fuente discreta, cuyos símbolos son estadísticamente independientes la entropía permite describir la fuente en términos de la información promedio producida por ella. Por ello la información promedio se denomina entropía de la fuente.

Esto es un concepto muy útil dado que los sistemas de comunicaciones no se diseñan para transmitir un mensaje en particular, sino que deben contemplar todos aquellos posibles.

Tasa de Información:

El concepto de entropía por si solo no es suficiente para relacionarlo directamente con la capacidad del canal de comunicación. Por lo tanto definiremos como tasa de información "T" al cociente entre la entropía de la fuente respecto a la duración promedio de los símbolos que esta envía

$$T = H(x) / \tau$$

Donde τ es la duración promedio de los símbolos. se mide en seg./símbolo.

Por lo tanto

$$T = \frac{H(x)}{\tau} = \frac{[Shannon/\ símbolo]}{[seg/\ símbolo]} = [Shannon/\ seg]$$

Si consideramos como caso particular una fuente digital de elementos binarios equiprobables la cantidad de bits coincide con la cantidad de shannon, por lo cual podemos expresar la tasa de información en bits por segundo (bps).

Por ejemplo para el sistema telegráfico tenemos:

P(punto) = 2/3

P(Raya) = 1/3

τ(punto) = 0,2 S

τ(raya) = 0,4 S

$H = 2/3 \log_2 3/2 + 1/3 \log_2 3 = 0.92$ bits/símbolo.

$\tau = 2/3 \cdot 0,2 + 1/3 \cdot 0,4 = 0,27$ seg.

$T = H/\tau = 0,92 / 0,25 = 3,4$ bps

Otro ejemplo:

Se tiene una fuente de 32 símbolos equiprobables cada uno compuesto por 5 bits. La duración de cada símbolo es de 1 seg. y además sabemos que la información se transmite en bloques de 20 símbolos separados por dos pulsos cada uno.

Calcular la tasa de información del sistema.

$H_{(block)}$ = 5 bits/simb x 20 = bits/símbolo.

$\tau_{(block)}$ = 20 x 1 mseg/símbolo + 2 x 2,5 ms/simb= 25 ms/simb

$T = H/\tau$ = 100 bits/símbolo = 4000 bps

Definición de Byte:

Se denomina byte al número de bits necesarios para poder representar un carácter.

Definición de palabra:

Se define como palabra a un número de caracteres fijos (bytes), que un computador trata como unidad, cuando se los transfiere entre sus distintas unidades:

6 bits por byte ======> long. de palabra de 12, 24, 36, 72.

8 bits por byte ======> long. de palabra de 8, 16, 32, 64.

Concepto de bloque:

El bloque es un conjunto formado por algunos bits, que reciben un tratamiento único a los efectos de la transferencia que una computadora realiza entre su memoria y sus periféricos.

Medidas de la velocidad en la transmisión de la información

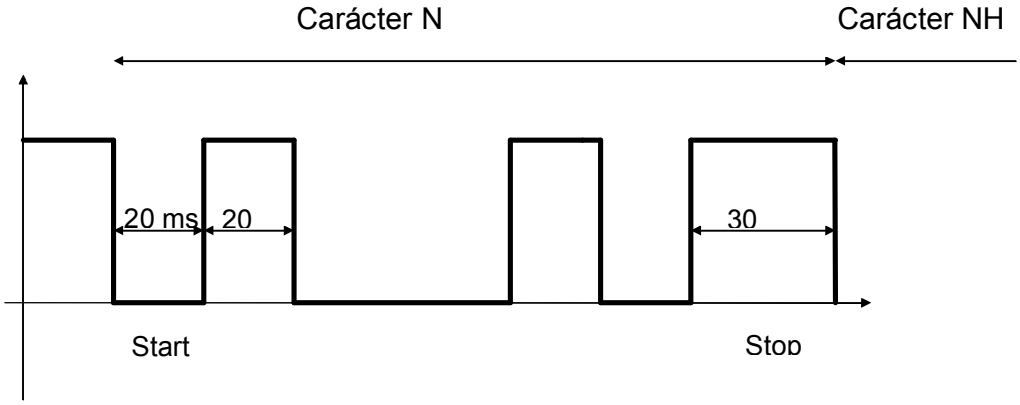
Velocidad de modulación:

Se denomina velocidad de modulación a la inversa de la medida del intervalo de tiempo nominal mas corto, entre dos instantes significativos sucesivos de la señal modulada (recomendación X-15 CCITT). También se la suele definir como la velocidad de modulación a la inversa del tiempo que dura el elemento mas corto de la señal que se utiliza para crear un pulso.

$$Vm = 1/\tau(seg.) = [baudio]$$

También se la suele llamar Vm => Velocidad de señalización.

Por ejemplo vamos a estudiar la transmisión de un servicio telex que trabaja en modo asincrónico.



Utilizo el código BAUDOT , en donde la señal de arranque y datos es de 20 ms y la parada 30 ms.

$$Vm = 1/\tau = 1/20 \text{ ms} = 50 \text{ baudios}$$

Velocidad binaria o velocidad de transmisión

Si bien CCITT define a cada una de ellas o bien UIT solamente adopta la velocidad binaria, llevan conceptos equivalentes.

Velocidad binaria: Velocidad global de transmisión expresada en **bps.**

Se denomina velocidad de transmisión en un canal de datos al numero de dígitos binarios transmitidos en la unidad de tiempo, independientemente de que halla o no información.

En las transmisiones asincrónicas este concepto carece de sentido por cuanto la separación entre caracteres puede ser variable. En este caso es recomendable utilizar el concepto de velocidad de modulación.

Por lo tanto, la velocidad binaria o de transmisión solo se usa en los sistemas sincrónicos.

Para un enlace de *m* canales y de *n* niveles, la velocidad de transmisión será:

$$Vt = \sum_{i=1}^{m} \frac{1}{Ti} \bullet \log_2 ni$$

Donde Ti es la menor duración teórica de un elemento y ni es el número de estados significativos de la modulación.

Para un canal que transmita en modo serie lo expresado queda

$$Vt = \frac{1}{T} \bullet \log_2 n$$

Simplificamos más la ecuación si consideramos que la información es binaria

$Vt = 1 / T$

Para el caso de modulaciones de cuatro, ocho, dieciséis:

$Vt = 2 .1/ T$; $4.1/ T$; $8.1/ T$

Si lo relacionamos con *Vm*

$Vt = 2\ Vm$; $3\ Vm$; $4\ Vm$

entonces

$\log_2 4 = 2$; $\log_2 8 = 3$; $\log_2 16 = 4$

<u>Velocidad de transferencia de datos</u>

Se denomina velocidad de transferencia de datos al número medio de bits por unidad de tiempo que se transmiten entre equipos correspondientes a un sistema de transmisión.

$$Vtd = \frac{\text{Número total de bits transmitidos}}{\text{tiempo empleado}}$$

Análogamente podemos extender este concepto a bytes, caracteres, palabras o bloques y medir el tiempo en seg., minutos u horas.

Vdt = bytes/segundo

Vdt = bloques/hora

También podemos hablar de la velocidad real de transferencia de datos como el número medio de bits por unidad de tiempo (recomendación V-7 CCITT). La definición es exactamente la misma a excepción de que cuando consideramos Vrtd solo se miden los bits, bytes, palabras sin errores de transmisión.

De esta manera podemos definir la eficiencia del sistema de comunicaciones como:

$$E = \frac{Vrtd}{Vtd}$$

Transmisión en modo serie

Se denomina transmisión en modo serie a aquella en que los bits que componen cada carácter se transmiten en *n* ciclos de 1 BIT cada uno. La secuencia de los bits transmitidos se efectúa siempre, por orden de pasos crecientes, es decir, al revés de como se escriben las cifras en el sistema de numeración binario.

Se entiende por sincronismo, que tanto la fuente como el colector de datos, tengan una base de tiempo común, de forma que reconozcan inequívocamente la transmisión de un uno o un cero.

La transmisión puede hacerse de dos maneras: Sincrónica y Sincrónica.

En el procedimiento asincrónico, cada carácter a ser transmitido es delimitado por un BIT denominado de cabecera (start) y uno o dos bits denominados de terminación o parada (stop). El BIT de arranque tiene funciones de sincronización de los relojes del transmisor y del receptor. El o los bits de parada se usan para separar un carácter del siguiente.

La necesidad de obtener mayor rendimiento entre los bits útiles, es decir los que contienen información, hizo pensar en la posibilidad de transmitir menor cantidad de bits para el proceso de resincronización entre emisor y receptor.

Paralelamente el desarrollo de tecnologías más sofisticadas permitió la transmisión conjunta de señales de sincronismo con los datos. De esta manera nacieron los procedimientos de transmisión sincrónicos. En este caso, existen dos relojes uno en el TX y otro en el RX. En el modo sincrónico, la información útil es transmitida entre dos grupos denominados delimitadores. Estos grupos uno es llamado encabezado (header) que se encarga de resincronizar los relojes y uno de terminación.

Los bloques a ser transmitidos tienen un tamaño que oscilan entre 128 y 1024 bytes. La señal de sincronismo en el extremo fuente, puede ser generada por el equipo terminal de datos o por el módem. Sin embargo, cualquiera sea el que genere dicha señal, la misma será común para ambos equipos y en ambos extremos de la línea.

Compresión de datos

Los medios de comunicaciones necesarios para realizar transferencia de datos, han requerido de hardware, módems, más veloces y sistemas de multicanalización de canales de comunicaciones más eficientes, que usando modernas técnicas estadísticas han ido mejorando sus performances con el objeto de abaratar los costos de comunicación.

En todo este esquema un método que ha ido ganando mercado rápidamente son los sistemas de compresión de datos que permiten reducir el volumen de estos, disminuyendo las necesidades de almacenamiento en discos y abaratando las transmisiones, usando códigos mas sofisticados, sin necesidad de aumentar el BW de los canales de comunicación , se logra transferir mayor cantidad de información en tiempos menores.

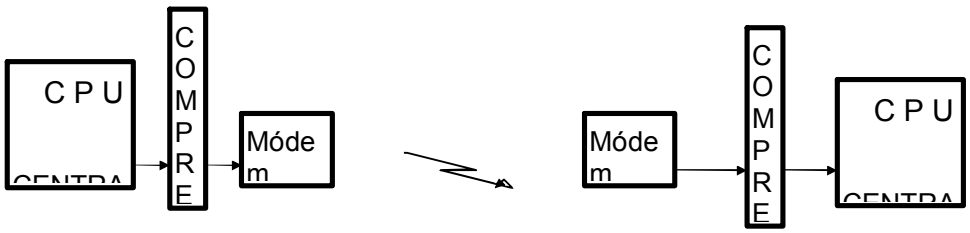
Se denomina compresión de datos a las técnicas lógicas o físicas que permiten que un conjunto de datos de una determinada longitud, pueda ser reducido en su tamaño sin alterar el significado de la información que la misma contiene.

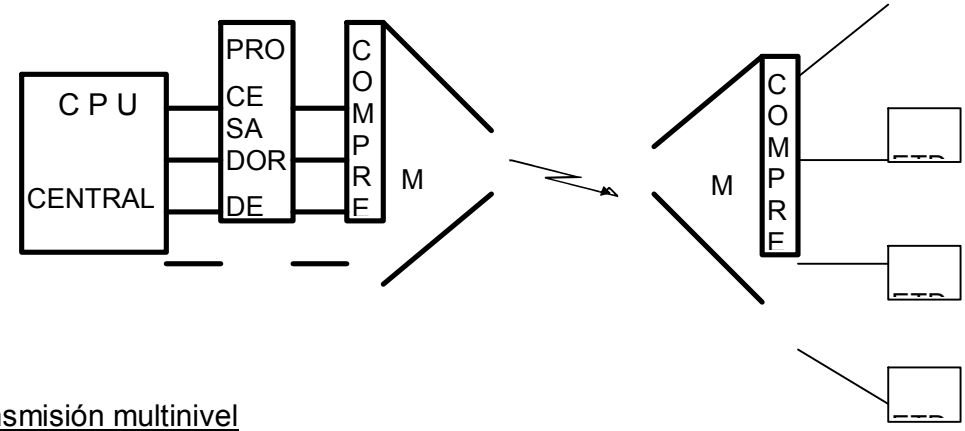
Índice de compresión X

$$X = \frac{\text{Longitud original de un conjunto de datos}}{\text{Longitud comprimida del conjunto de datos}}$$

Se denomina factor de mérito de compresión

P = 1 / X





Transmisión multinivel

Una de las maneras de aumentar la velocidad de transmisión sin aumentar la velocidad de modulación de manera que no aumente la tasa de errores y el BW es mediante un aumento en los niveles.

$$Vt = 1 / T \cdot \log_2 n$$

Donde *n* es el número de niveles. Cuando *n* es mayor a 2, entonces la transmisión es denominada multinivel.

Para n=2	4 niveles (dibits)
para n=3	8 niveles (tribits)
para n=4	16 niveles (cuadribits)

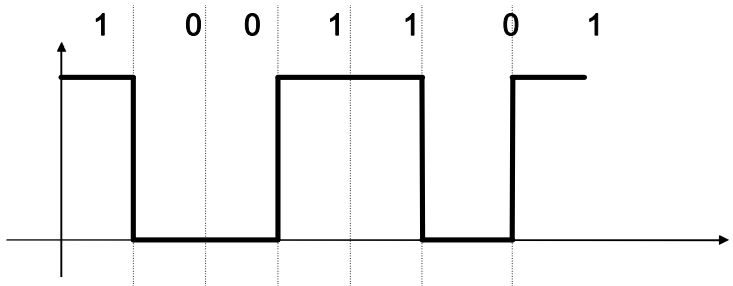
La mayoría de los equipos módems, que transmiten a velocidades de más de 2400 bps, emplean estas formas para aumentar la velocidad de transmisión.

Si consideramos como ejemplo una transmisión de 9600 bps necesitaríamos cambiar la cadencia cada 104 μs. Si usamos cuadribits, podríamos trabajar a velocidades de modulación de 2400 baudios lo que representa una cadencia de 416 μs

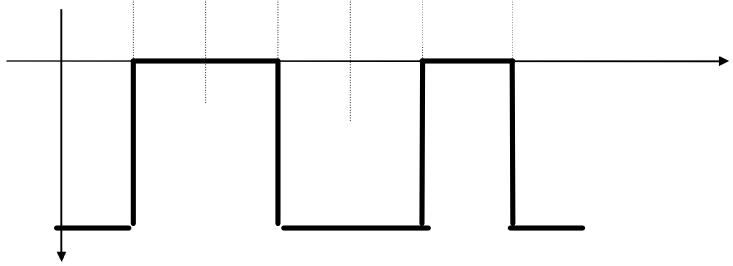
1- Señales en Banda Base

Señal unipolar:

Se dice que una señal es unipolar cuando un dígito binario 0 o 1 está representado por una misma polaridad ya sea positiva o negativa.



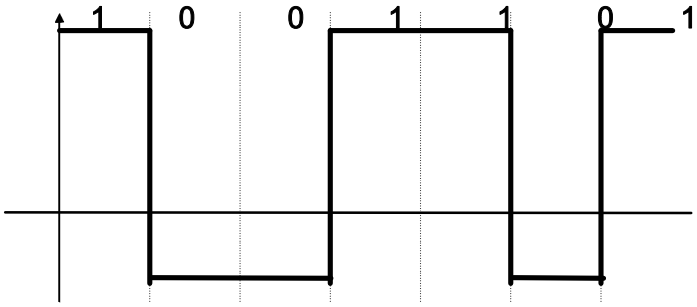
Señal unipolar positiva



Señal unipolar negativa

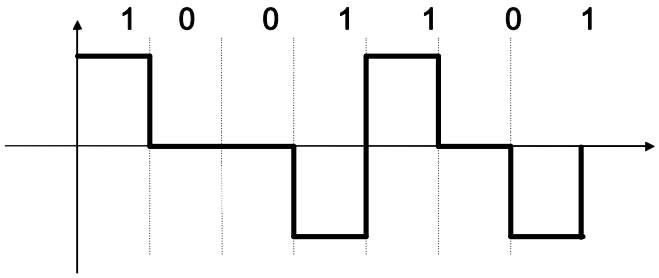
2- Señal polar:

Se dice que una señal es polar cuando los valores que representan dígitos binarios 0 y 1 se originaron como consecuencia de un valor máximo positivo +Vi y un valor negativo -Vi.

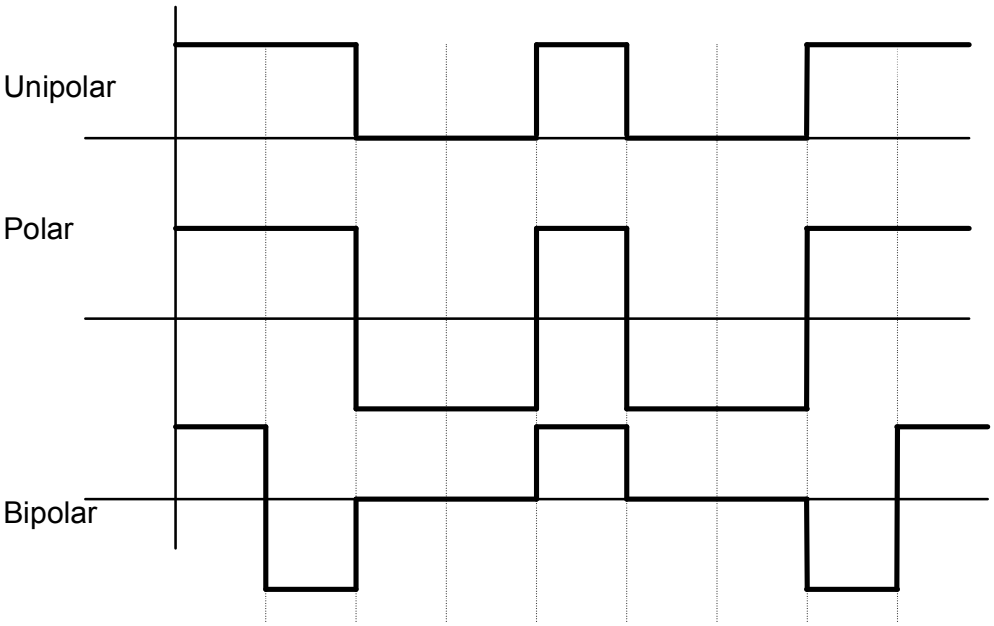


3- Señal bipolar:

Se dice que una señal es bipolar cuando los dígitos binarios en el caso de 1 toman polaridad alternada y el cero siempre es cero.



Vamos a hacer otro ejemplo teniendo en cuenta las definiciones vistas. Sea un número 110010011 en los casos vistos.



Transmisión en Banda Base

Por lo general se utiliza este tipo de transmisión por su bajo costo. Los equipos que realizan este tipo de transmisión se los denomina *módem de banda base* .

Cuando deseamos enviar señales en esta modalidad la señal es previamente codificada de forma tal de reducir los componentes de continua que todo pulso rectangular asimétrico pueda presentar. La existencia de transformadores y de amplificadores, si fuera el cero, no dejarían pasar la componente de continua provocando una deformación importante de la señal hecho que aumentaría la tasa de error.

Un ejemplo de utilización de esto es la red local o LAN, que transmite fundamentalmente en banda base debido a que los medios de comunicación no pertenecen a redes públicas. por lo tanto toda señal en banda base debe ser codificada para adaptarla a la línea de transmisión. De allí surge la existencia de diversos códigos de línea o códigos de banda base.

Características particulares de las transmisiones en Banda Base

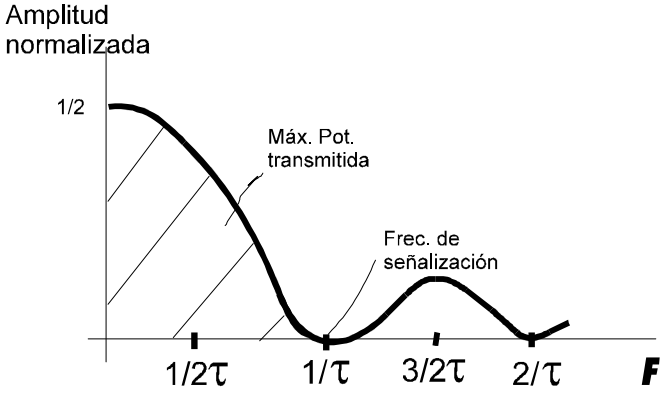
La señal más simple para la transmisión es la señal NRZ (non return to cero).

La transmisión de un 1 ===> emisión de un pulso.

La transmisión de un 0 ===> no emisión de un pulso.

Por lo tanto decimos que es una señal unipolar o también conocida como ON/OFF.

Si analizamos un pulso en el dominio de la frecuencia tenemos

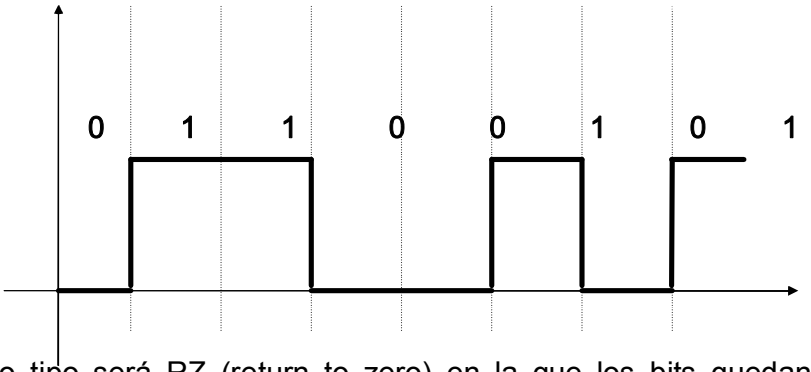


Pero esta señal al ser de baja frecuencia la hace muchas veces una señal incompatible para la transmisión dado que existen transformadores de acoplamiento y por otra parte estos transformadores por lo general se utilizan como acoplamiento eléctrico.

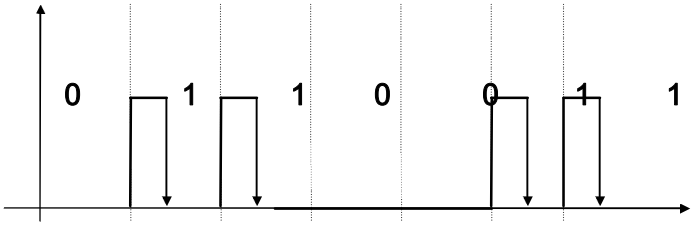
Otro inconveniente que presentan las señales ON/OFF ocurre cuando se tiene que enviar una señal de reloj o sincronismo conjuntamente con datos, esta señal no posee potencia suficiente en la frecuencia de señalización y en consecuencia impide recuperar la señal de reloj.

A continuación, clasificaremos las señales en banda base.

Además de clasificarla por el tipo de polaridad la podemos clasificar mediante el ancho del pulso. Como por ejemplo la señal NRZ que el BIT ocupa la totalidad del intervalo.



Otro tipo será RZ (return to zero) en la que los bits quedan representados por pulsos que ocupan una parte del intervalo significativo.



<u>Códigos usuales para la transmisión en banda base</u>

Los códigos más usados son:

1) Unipolar sin retorno a cero (NRZ)

2) Unipolar con retorno a cero (RZ)

3) Polar sin retorno a cero (NRZ)

4) Polar con retorno a cero (RZ)

5) Bipolar sin retorno a cero (NRZ)

6) Bipolar con retorno a cero (RZ)

7) Codificación diferencial

8) Manchester.

9) Manchester diferencial.

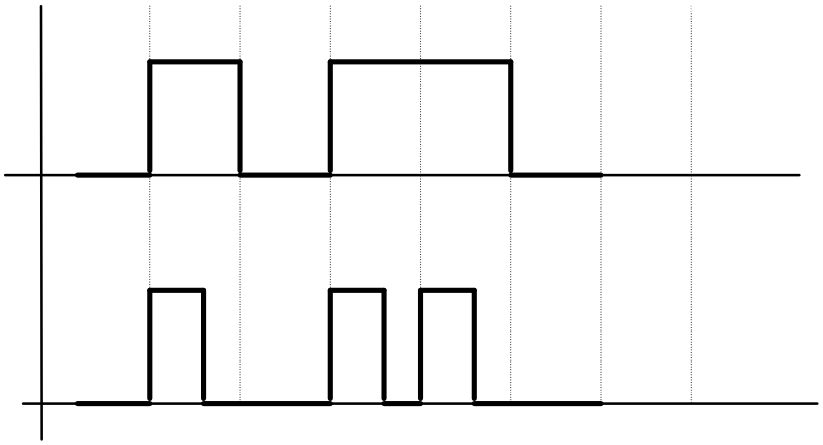
10) Miller.

11) HDB3

12) Código 4B 3T (4 binario 3 ternario)

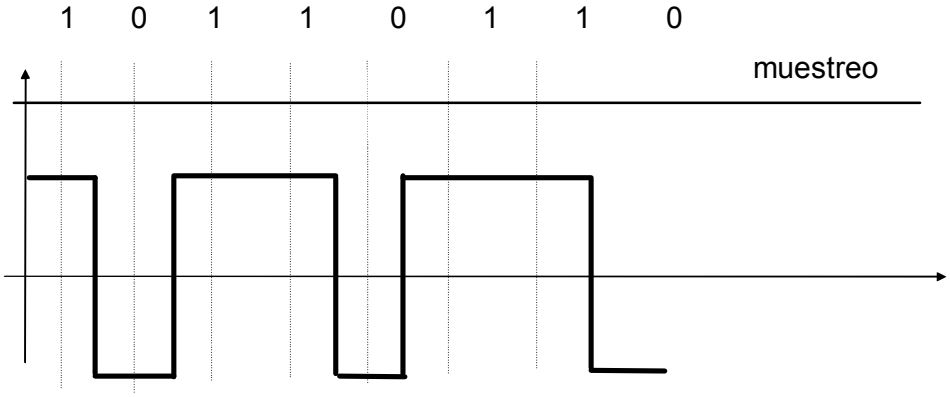
1) y 2)

El estado de cada BIT esta representado por un pulso de corriente continua y un estado de corriente nulo. En este tipo de señales debemos realizar un muestreo para determinar el valor de cada BIT. La transmisión de un 1 se la suele llamar nivel de marca y la transmisión de un cero nivel de espacio.



3)

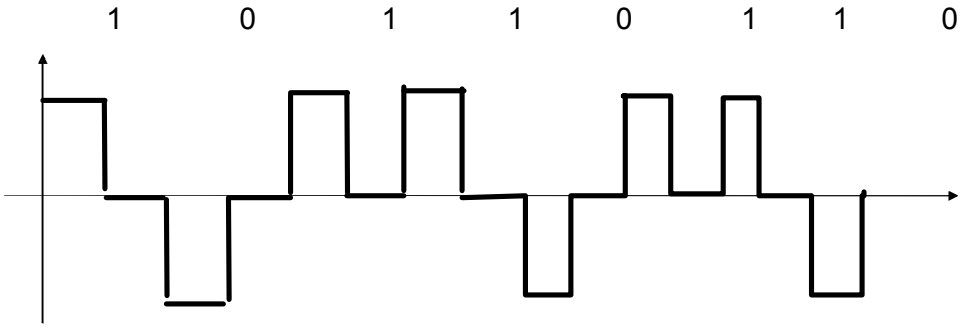
El estado de cada BIT queda determinado por una corriente continua positiva y otra negativa. En este tipo de señales, si bien se pierde el sincronismo se tiene la ventaja que tiene menor BW dado que los pulsos son más anchos. Aquí también tenemos que realizar muestreo para determinar la presencia de cada BIT.



4)

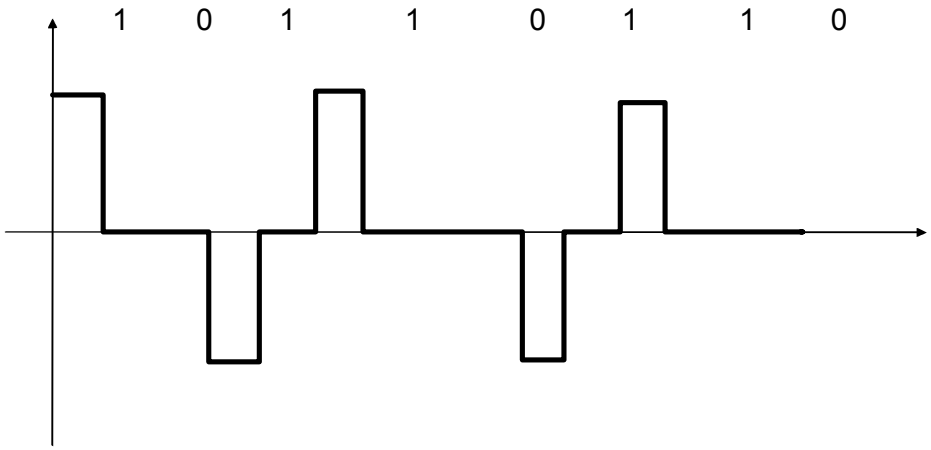
Si consideramos un BIT que represente un 1 tendremos que durante un instante habrá corriente negativa y luego retornará a cero.

Este tipo de señales se denominan autosincronizantes debido a que en el receptor, la señal de reloj queda unívocamente sincronizada por la cadencia de los pulsos, positivos y negativos que llegan desde el transmisor.



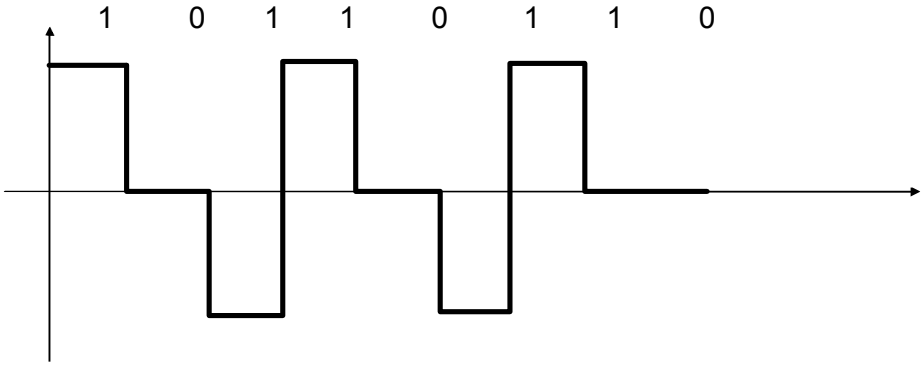
5)

En este tipo de señales se utiliza bipolaridad solamente cuando se transmiten unos, y la transmisión de ceros corresponde a señales de no corriente.



6) Bipolar sin retorno a cero (AMI)

Este tipo de código, también denominado AMI (alternative mark inversion), presenta la ventaja de utilizar pulsos de mayor duración que los bipolares RZ, siendo en consecuencia menor el BW y desde el punto de vista del sincronismo presenta iguales características que los anteriores.

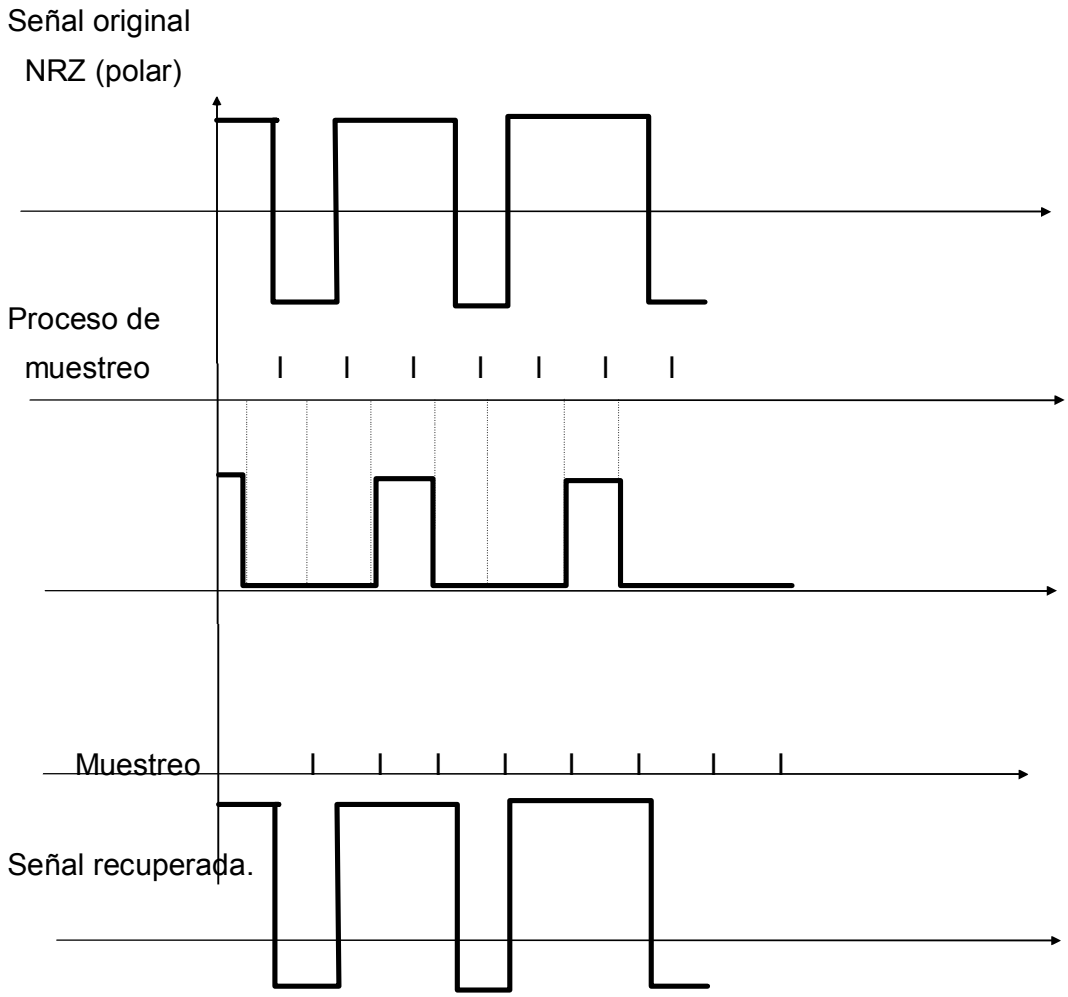


7) Codificador diferencial

En este tipo de codificación tienen lugar dos etapas. Una para formar en el TX la señal diferencial a transmitir por el medio físico y otra para armar la señal recuperada en el RX.

Primero una señal del tipo NRZ (polar) debe ser muestreada. En el instante del muestreo se detecta un 1 se produce un cambio de estado o transición. Cuando lo que se detecta es un cero, significa una no transición.

Para recuperar la señal original se debe efectuar un nuevo muestreo de la onda recibida, comparándose la polaridad de muestreo adyacente. Si ha habido una transición corresponde un "1", caso contrario "cero".

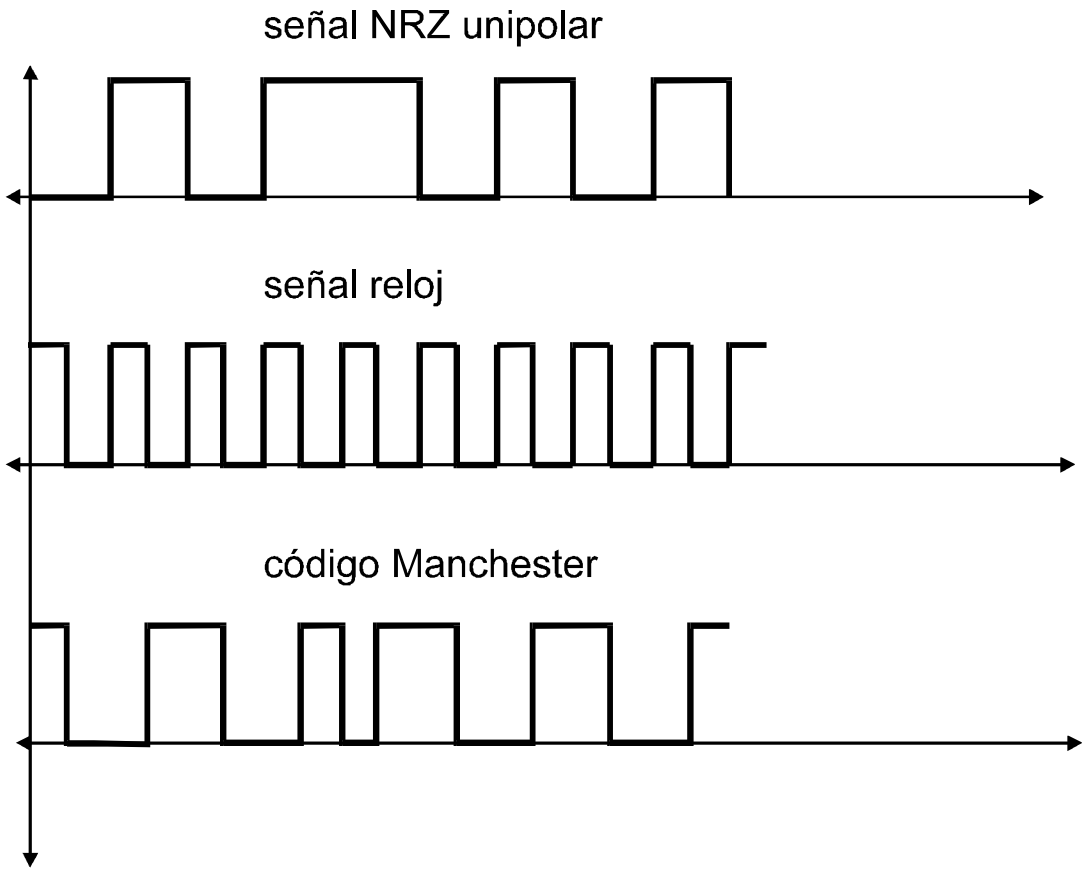


8) Código Manchester.

En este caso el BIT "1" se representa por una transición positiva en la mitad del intervalo significativo y un cero con una transición negativa en la misma ubicación.

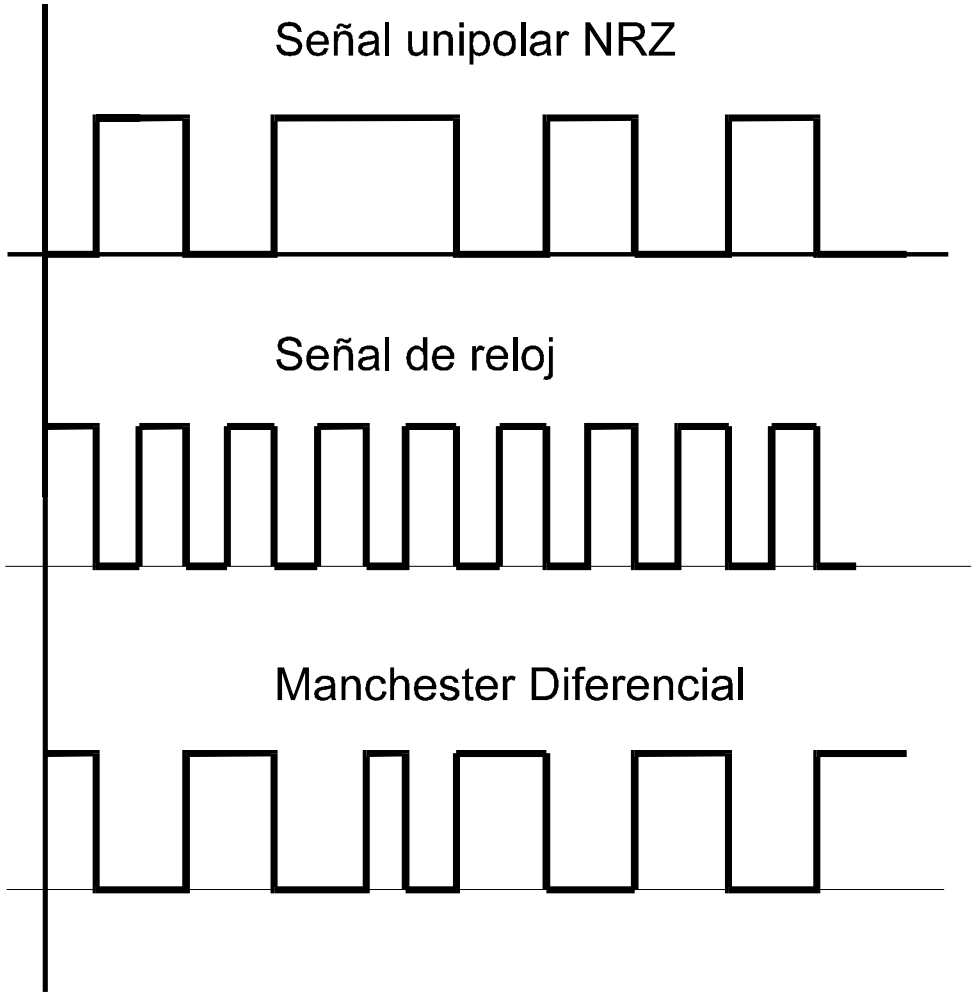
En este tipo de codificación no se utiliza la diferencia de valor de los niveles para representar los bits, sino que se emplean las fases positivas y negativas de los pulsos.

Esta técnica posibilita la solución al problema de recuperación de la señal reloj.



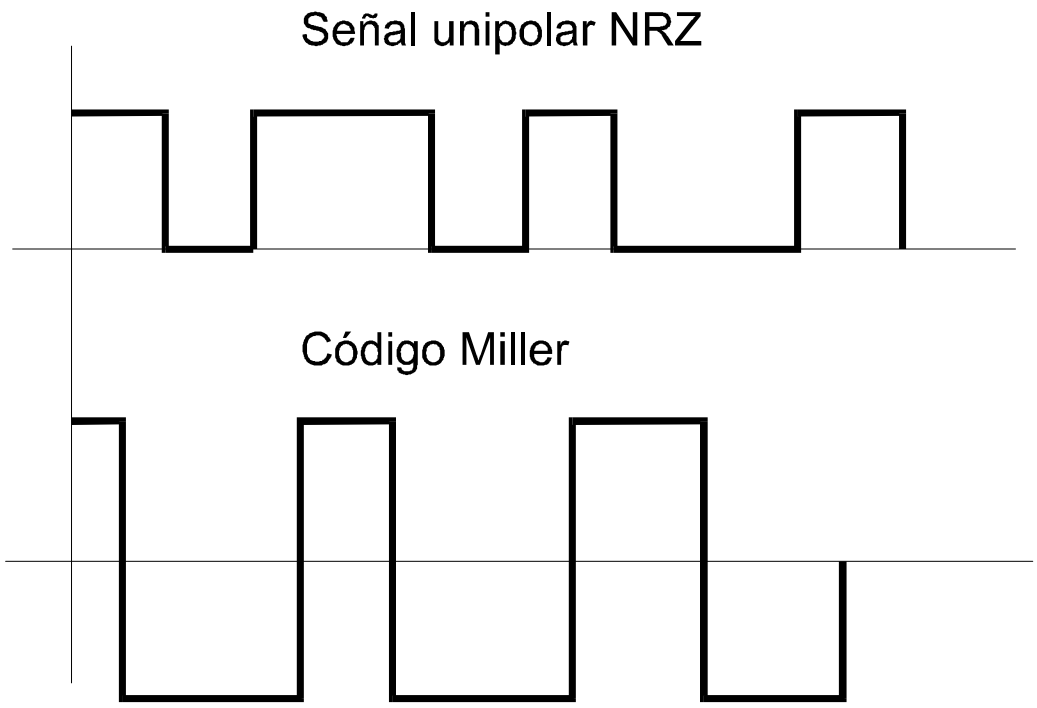
9) Código Manchester Diferencial.

Este código se caracteriza porque para la transmisión de un cero se efectúa una transición negativa en la mitad del intervalo significativo, mientras que para la transmisión de un uno no se efectúa ninguna transición, pero si se alterna la polaridad con la cual se representa



10) Código Miller.

Este código emplea para la transmisión de un "1" , la transición en la mitad del intervalo significativo. Para el caso de un cero, existe una transición al final del intervalo si el BIT siguiente es cero o de lo contrario no habrá ninguna transición.



El código Miller permite reducir considerablemente la contribución de las bajas frecuencias y garantiza un número mínimo de transiciones como para recuperar la señal de reloj y tiene un BW menor que el código Manchester.

11) Código HDB-3

El código High Density Binary se basa en el denominado AMI (Alternative Mark Invertion), el cual es un código bipolar NRZ.

El cero siempre es cero y el 1 con polaridad alternada. Este tipo de señal tiene el inconveniente que cuando viene una larga secuencia de "0" puede ser que se pierda la posibilidad de recuperar la señal reloj.

Es por ello que para limitar estas secuencias se efectúan violaciones a la polaridad como se realiza HDB-3 que permite la transmisión de 3 bits ceros consecutivos e inserta un BIT uno denominado BIT de violación.

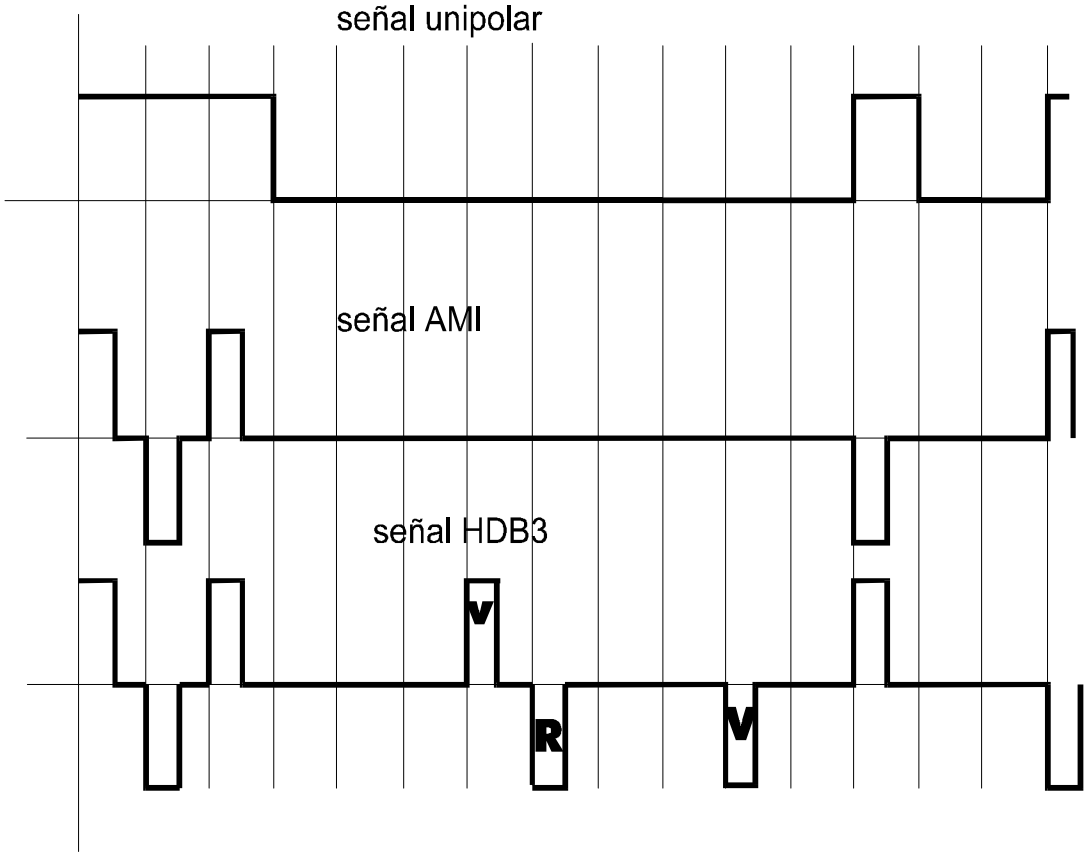
En el código HDB-3 cuando se tiene 4 ceros seguidos, se reemplaza la secuencia por " 000V o R00V "

El pulso V = 1 se denomina violación y R tiene siempre igual polaridad que V y se denomina pulso de relleno. Para decidir que secuencia emplear se debe contar la cantidad de ceros existentes entre la última violación y la actual:

Si la cantidad de ceros es par =====> 000V

Si la cantidad de ceros es impar====> R00V

El primer pulso V de la serie siempre lleva la misma polaridad del BIT transmitido. Esto sirve para que al recibirlo, se puede interpretar que se trata de una secuencia de ceros y no de datos. Los pulsos de violación se transmiten con polaridad alternada. De existir pulso de R lleva la misma polaridad que V.



12) Código 4B-3T (4 binario, 3 terciario).

El código HDB-3 es el que frecuentemente se utiliza hasta 34 Mbps sobre cables de cobre. Para velocidades mayores sobre cables coaxiales se utiliza el 4B -3T (por ej: 140 Mbps) Este código reduce la transmisión de 4 bits de información a 3 niveles señal terciaria.

Regla de formación.

Señal binaria	Señal terciaria		
0000	0	-1	+1
0001	-1	+1	0
0010	-1	0	+1
0011	0	+1	-1
0100	+1	-1	0
0101	+1	0	-1
0110	+1	-1	+1
0111	0	+1	+1
1000	0	+1	0
1001	0	0	+1
1010	-1	+1	+1
1011	+1	0	0
1100	+1	0	+1
1101	+1	+1	0
1110	+1	+1	-1
1111	+1	+1	+1