

Introducción

Las primeras técnicas utilizadas para grabación y procesamiento del sonido eran analógicas: inicialmente cilindros y discos mecánicos y posteriormente hilo magnético, cinta y banda sonora de película. La grabación magnética analógica tal como es conocida en la actualidad, fue desarrollada fundamentalmente bajo el empuje de la segunda guerra mundial, en conjunto con el turborreactor, el radar y la bomba atómica. Desde entonces el procesamiento y grabación analógicos del sonido se ha perfeccionado, estudiando cada defecto del proceso y aplicando las medidas necesarias para reducirlo.

Hoy en día existen además, tecnologías digitales, las que son aplicadas en diversos campos: las imágenes, ya sea fijas o en movimiento (fotografía, cine y vídeo) , las telecomunicaciones (telefonía, correo), la información escrita y por supuesto el campo de la grabación, procesamiento y almacenamiento del sonido; todos ellos implican procesos en donde actualmente se prefiere, por alguna u otra razón, técnicas digitales ya sea para su producción, procesamiento, distribución, transmisión o almacenamiento.

En el presente trabajo, se abordarán, específicamente, aquellos temas relacionados con la grabación digital del sonido, así como algunas formas de almacenamiento y procesamiento de la información en sistemas de audio digital.

1–Digitalización de señales de audio

En un sistema analógico la información sonora está contenida en las infinitas variaciones de algún parámetro continuo, tal como la tensión o la intensidad de flujo magnético. Un determinado parámetro puede ser una exacta representación del original solamente si el proceso de conversión es lineal, por lo que la señal de audio analógica inevitablemente sufre degradaciones, dependiendo del número de etapas o procesos por los que dicha señal atraviesa.

En un sistema de audio digital, la señal es discreta en función del tiempo (corresponde a muestras de la señal original en un intervalo de tiempo) y en función de la amplitud (los valores numéricos de la señal digitalizada se encuentran en pasos discretos). En un sistema de audio digital , la información se encuentra en forma binaria. Las señales enviadas tienen solamente dos estados y cambian en determinados momentos de acuerdo con una señal de reloj estable. Si la señal binaria resulta afectada por el ruido, éste será rechazado en el receptor, ya que solamente se considera si la señal está por encima o por debajo de un determinado umbral. El ruido superpuesto puede desplazar el punto en el que el receptor detecta que ha habido un cambio de estado; la inestabilidad en el tiempo tiene el mismo efecto. Esta inestabilidad es rechazada también, ya que en el receptor la señal es redisparada por un reloj estable, con lo que todos los cambios en el sistema tienen lugar en coincidencia con los flancos de ese reloj. De este modo, la señal binaria resulta al final siempre la misma, aunque atravesase varias etapas.

Por otra parte, en sistemas numéricos es factible incorporar un sistema de corrección de errores, por lo que si llegasen a producirse en la grabación digital distorsiones tales como *drop outs* o interferencias (pueden desaparecer algunos cambios de flujo o aparecer otros que no existan; el resultado es que algunos de los números serán incorrectos), estos podrían ser corregidos.

1.1–Conversión

1.1.1–Tipos de digitalización

Una señal analógica es continua en el tiempo e infinitamente variable en tensión , mientras que una señal digital es discreta en el tiempo y su tensión varía por pasos.

Existen varios métodos para convertir una forma de onda analógica en una secuencia de bits, métodos que están relacionados entre si y en algunos casos, en sistemas de conversión avanzados, pueden trabajar en forma complementaria. La conversión en el tiempo, de continuo a discreto , se conoce como muestreo, el proceso inverso se conoce como reconstrucción. La representación mediante un numero de valor analógico de una muestra se conoce como cuantificación. El proceso de muestreo consiste en que un tren de impulsos de amplitud constante es modulado por la señal de entrada, de aquí el término modulación de impulsos en amplitud (*pulse amplitude modulation*), abreviado como PAM (La elección de la frecuencia de muestreo , es decir, la velocidad a la que debe examinarse la tensión de entrada para transmitir la información de una señal variable; es importante en cualquier sistema: si es demasiado baja, la señal se verá degradada, y si es demasiado alta, el número de muestras a registrar crecerá innecesariamente) . Cuando la altura de estos impulsos es cuantificada y expresada mediante un código numérico, el resultado es conocido como modulación por impulsos codificados (*pulse code modulation*) o PCM. (En principio, esto es equivalente a registrar las variaciones de una tensión, anotando las lecturas de un voltímetro digital cada pocos segundos).

La amplitud de la señal que puede ser transmitida de éste modo depende únicamente de la capacidad del cuantificador y es independiente de la frecuencia de la señal de entrada. Análogamente, la amplitud de las señales no deseadas introducidas en el proceso de cuantificación es también en gran parte independiente de la señal de entrada .

Existe además, un proceso conocido como modulación diferencial por impulsos codificados (*differential pulse code modulation*) , DPCM, donde el parámetro que se cuantifica es la diferencia entre los valores absolutos de la muestra anterior y la actual.

1.1.2–Cuantificación

La señal digitalizada es discreta en el eje vertical (amplitud) ya que los valores numéricos de la señal digitalizada se encuentran expresados en pasos discretos correspondientes a números enteros. (El convertidor análogo digital tiene una salida digital que corresponde a un número entero, y el convertidor digital a análogo tiene una entrada digital que corresponde a un numero entero).

El hecho de que valores enteros para la función digitalizada pueden no representar exactamente todos los valores posibles de la señal análoga (continua), conduce a la idea de error de cuantización. Cualquiera que sea el valor exacto de la señal de entrada, el cuantificador lo expresará como el valor numérico del intervalo en que haya resultado. Cuando éste número llegue al conversor D/A , éste producirá una tensión correspondiente al centro del intervalo. La cuantificación produce , por tanto, un error que no puede exceder de $\pm Q/2$, siendo Q la magnitud del intervalo de cuantificación (en la mayoría de los equipos digitales de audio todos los intervalos de cuantificación son lineales, aplicándose el término de cuantificación uniforme).

El numero de valores enteros distintos que una determinada muestra puede tomar está especificado por el numero de bits, y está dado por la expresión 2^b , donde b es el numero de bits. Por ejemplo, un disco compacto de audio almacena una muestra como una palabra de 16 bit; la palabra toma uno de los $2^{16} = 65536$ valores posibles.

Una forma simple de describir el error de cuantización es en términos de la resolución de un sistema de conversión. Como el sistema es capaz de resolver una unidad entera, y el valor máximo que puede tomar una señal es 2^b (donde b es el numero de bits), se dice que la relación señal – error es entonces 2^b , lo que usualmente se expresa como relación señal – ruido en decibelios, de acuerdo con la expresión:

$$L_s - L_n \text{ (dB)} = 20 \log 2b = 20 b \log 2 = 6b$$

La relación señal – ruido de cuantización para un sistema de 16 bit es entonces $6 \times 16 = 96 \text{ dB}$. Cada bit de resolución adicional agrega 6 dB a la relación. Al igual que otras relaciones establecidas para definir el rango dinámico, esta relación se asume para el mejor de los casos, es decir, donde la señal se encuentra a su máximo valor posible. Al tratar el error de cuantización como un ruido, se asume que la señal original y el tren de impulsos no están en sincronía. En caso de estar sincronizados, el error aparece como una distorsión (fenómeno que será analizado mas adelante).

En algunos estudios sobre la cuantificación se establece otra relación entre el numero de bits de una palabra y la relación señal ruido: $(6,02b+1,76)\text{dB}$. En dicha aproximación existen dos defectos. Primero, la potencia de ruido calculada tiene un espectro finito y no se ha tenido en cuenta el efecto de sobre éste del filtro de reconstrucción. Segundo y más importante, los cálculos solo son válidos si la función de densidad de probabilidad del error de cuantificación es uniforme. A bajos niveles y particularmente con señales puras o sencillas, el error de cuantificación deja de ser aleatorio y se convierte en una función de la señal de entrada. (en caso que la señal no deseada sea una función de la señal deseada, se debe hablar de distorsión en lugar de ruido) .Al reducirse el nivel de la señal analógica , el error de cuantificación se hace cada vez menos aleatorio, apareciendo ruido de modulación . Cuando en la entrada existe mas de una frecuencia, aparecen productos de intermodulación. El efecto resultante de dicho fenómeno ha sido denominado como granulación.

De lo anterior, se deduce que , mientras el ruido de cuantización es menos aleatorio, dicho ruido tiene relación o es función de la señal de entrada, por lo tanto corresponde a un tipo de distorsión; por otra parte, mientras mas aleatorio sea el ruido de cuantización, se reduce por consiguiente la distorsión.

1.1.2.1–Dither

En la práctica, es difícil tener un error de cuantificación perfectamente determinístico debido al ruido presente en la señal de entrada.. Este ruido tiene un efecto aleatorizador sobre el error de cuantificación, reduciendo la distorsión. En la figura se muestra que el efecto de las fluctuaciones en emborronar en sentido horizontal la función de transferencia del cuantificador.

Cuando la tensión RMS del ruido es un tercio del intervalo de cuantificación, el error de cuantificación se hace tan aleatorio como el ruido, y el proceso de cuantificación resulta lineal.

Esta técnica de aleatorización se conoce como dither; además de linealizar el sistema define la relación señal ruido. Si la señal de entrada no lleva suficiente ruido puede incorporarse una fuente de ruido al conversor. Como fuente de ruido *gaussiano* (aleatorio) puede utilizarse un diodo. Sin embargo, aparece además como resultado un leve ruido de modulación, pero la linealidad de la conversión se mantiene intacta. Mediante procesos conocidos como modelado de ruido o *noise shaping*, el ruido de modulación que aparece al aplicar la técnica de dither es desplazado hacia zonas donde el oído humano es menos sensible, es decir, sobre los 15 KHz. Si se utiliza suficiente dither gaussiano para obtener un sistema lineal, por ejemplo, $Q/3 \text{ RMS}$ (siendo Q la magnitud del intervalo de cuantización), la relación señal ruido puede ahora expresarse como $6,02b \text{ dB}$. En la mayoría de los casos prácticos 6 dB por bit es una valoración adecuada para la relación señal ruido, pero solo si se utiliza dither apropiado.

1.1.2.2–Redither

Es necesario mencionar que, además del proceso de conversión A/D, donde es necesario un sistema de conversión lo más lineal posible; cada vez que la señal ya en el dominio digital es procesada (mezcla, cambio de ganancia, ecualización, procesos dinámicos, etc.), generalmente, debido a los cálculos matemáticos que dichos procesos realizan, aumenta el numero de bits requeridos para representar la

señal. Actualmente, la mayoría de los procesadores profesionales trabajan con resoluciones internas mayores que 16 bit, llegando incluso hasta 48 bit, los que finalmente, para aplicaciones como discos compactos, deben expresarse como palabras de 16 bit. En éstos casos, simplemente truncan la señal a 16 bit (es decir, simplemente eliminar los últimos bits, donde se encuentra la información de menor peso y correspondientemente los niveles mas bajos) da como resultado una pérdida de resolución de la señal, pérdida que ocurriría cada vez que la señal es procesada. En relación con las investigaciones de realizadas por Michael Gerzon (condecorado con medalla de oro por la AES, por sus investigaciones en el campo de la Psicoacústica), el oído humano utiliza precisamente esa información de bajo nivel para reconstruir mentalmente la imagen estéreo, de modo que cualquier compromiso en ésta área se manifiesta como una pérdida en la sensación de espaciosidad y transparencia.

Actualmente, el dither es aplicado en orden de cambiar el carácter del ruido de cuantización, de modo de que sea percibido de manera semejante al hiss análogo. La forma en que trabaja el dither, por ejemplo, para una aplicación muy común como es a la salida de un procesador que internamente trabaja con una resolución de 24 bit (como cuando se procesa audio en un computador), pero que a la salida de dicho procesador entrega un archivo de 16 bit (formato de un disco compacto de audio), es la siguiente: se calcula valores aleatorios y se suman valores aleatorios para cada muestra. Los valores deben ser diferentes para cada el canal izquierdo y derecho, de otro modo se vería comprometida la separación estéreo. Ej:

Se tiene una palabra de 24 bit (cada bit es un 0 o un 1 en notación binaria)

-----16 bits superiores----- -8 bit inferiores-

Palabra de 24 bit MXXX XXXX XXXX XXXW YYYY YYYY

Agregar valores aleatorios ZZZZ ZZZZ

El resultado de la adición de los valores aleatorios (Z) con los 8 últimos bits (Y) se agrega en lugar del bit de menor peso (LSB, representado por la letra W) de la palabra de 16 bit. De este modo, la secuencia de números aleatorios (ruido) se combinan con la información original almacenada en los últimos 8 bits, modulando al LSB (W). De éste modo, el LSB toma valores 1 o 0, dependiendo ahora de la información de nivel bajo almacenada en los 8 últimos bits. El proceso anteriormente mencionado es conocido como *redither*, para diferenciarlo del proceso de *dither* utilizado durante la conversión o grabación digital. (Todos los conversores A/D de 16 bits incorporan dither para linealizar el proceso de conversión).

Los números aleatorios agregados en el proceso anteriormente descrito se traducen como ruido al momento de la conversión D/A, pero, como ya fue mencionado, dicho ruido es modelado, su amplitud tiene que ver solamente con 1 LSB (-96 dB) y corresponde a frecuencias donde el oído es menos sensible, sobre los 15 KHz. Por otra parte, las señales de bajo nivel están representadas por el LSB, por lo que el efecto del ambiente y el decaimiento puede ser oído incluso hasta niveles de alrededor de -115 dB, aunque la relación señal ruido teórica es de 96 dB, el rango dinámico de una señal con *dither* alcanza un poco mas (hasta -115dB), porque el sistema auditivo humano puede percibir señales por debajo del ruido (ruido que, además, está en una banda donde el oído es menos sensible).

1.1.3-La señal digitalizada en función del tiempo

1.1.3.1- Muestreo y Aliasing

La señal muestreada, llamada para éste caso $y(t)$, corresponde al producto de la señal de entrada, llamada $x(t)$ y un tren de pulsos llamado $s(t)$, donde dicho tren es un tren de funciones delta

regularmente espaciadas por un intervalo T_s . (Las funciones delta se llaman también delta dirac y representan pulsos que sólo existen en un tiempo dado (sólo en un instante dado) por ello, para muestrear señales de audio convenientemente, el pulso de muestreo debe ser lo más parecido a este tipo de función, de modo que al multiplicar la señal por la delta, sólo queda el valor de la señal en ese instante de tiempo). La función delta se caracteriza por:

$\delta(t) = 0$ en todo t excepto en $t = t_0$ donde

$$\int_{-\infty}^{\infty} \delta(t) dt = 1$$

El producto de ésta operación representa la acción de un conversor A/D. Por lo tanto

$$y(t) = x(t) s(t)$$

donde $s(t)$ es:

$$s(t) = \sum_{m=-\infty}^{\infty} \delta(t - t_m)$$

donde el punto t_m es el m -ésimo tiempo de muestreo mT_s . El factor T_s le da a la función $s(t)$ dimensiones de unidad y sirve para hacer de la transformada de Fourier de $s(t)$ una función más simple.

La información entregada por y esta en forma de números discretos, un conjunto $\{y_m\}$. La función $y(t)$ no tiene existencia en los reales. Esto es importante ya que ilustra las limitaciones que afectan al procesamiento de la señal. La función $y(t)$ sería también la señal de salida si simplemente fuera pasara por un convertidor D/A sin ningún otro tipo de procesamiento. Entre los puntos de muestreo la señal es cero.

Como $y(t)$ es el producto de la señal de entrada y el tren de muestreo, la transformada de Fourier de $y(t)$, $Y(t)$ estará relacionada con la transformada de Fourier de la señal de entrada, $X(t)$, en convolución con la transformada de Fourier del tren de impulsos $S(t)$,

$$Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega') S(\omega - \omega') d\omega' \quad (\text{integral entre } -\infty \text{ y } +\infty)$$

La función $S(\omega)$ esta dada por la transformada de Fourier de $s(t)$

$$S(\omega) = \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t - mT_s) e^{-j\omega t} dt \quad (t - mT_s)$$

(integral entre $-\infty$ y $+\infty$; sumatoria de $m = -\infty$ hasta $m = +\infty$)

$$S(\omega) = \sum_{m=-\infty}^{\infty} e^{-j\omega m T_s} \quad (\text{sumatoria de } m = -\infty \text{ hasta } m = +\infty)$$

haciendo los reemplazos correspondientes

$$S(\omega) = \sum_{n=-\infty}^{\infty} e^{-j\omega n T_s} \quad (\text{sumatoria de } n = -\infty \text{ hasta } n = +\infty), \text{ donde } T_s = 2\pi / \omega_s$$

Entonces

$$Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega') \sum_{n=-\infty}^{\infty} e^{-j(\omega - \omega') n T_s} d\omega' = \sum_{n=-\infty}^{\infty} X(\omega - \omega_n) \quad (\text{integral entre } -\infty \text{ y } +\infty; \text{ sumatorias de } n = -\infty \text{ hasta } n = +\infty)$$

(integral entre $-\infty$ y $+\infty$; sumatorias de $n = -\infty$ hasta $n = +\infty$)

La ecuación $Y(n)$ dice que la transformada de Fourier de la señal digitalizada está dada por repetidas versiones de la transformada de Fourier de la señal de entrada. El término en la sumatoria cuando $n=0$ es precisamente $X(f)$, la transformada de Fourier de la señal entrante. Para los otros valores de n , aparecen frecuencias laterales ($n f_s$), relacionadas por n con la frecuencia de la señal de entrada:

La figura muestra la señal de entrada $X(t)$ y la señal muestreada $Y(t)$, donde se asume que: la señal de entrada tiene valores reales y puede ser representada entonces, por un gráfico en 2 dimensiones. (En principio, no hay problema si X es complejo, solamente habría una figura más complicada); y se asume que los valores que toma $X(nT)$ para cada valor de n , no se superponen o interfieren entre sí. Dicha superposición no ocurre si la frecuencia más alta del espectro $X(f)$ es menor que $f_s/2$ (donde f_s es la frecuencia del tren de impulsos de muestreo).

El criterio de Nyquist dice que, cuando una señal es muestreada, la frecuencia más alta en la señal debe ser menor que la mitad de la frecuencia de muestreo. (en la práctica, si lo que se desea es digitalizar señales dentro de todo el espectro audible, es decir, de 20Hz a 20KHz, bastaría con utilizar una frecuencia de muestreo de 40.001 Hz. Sin embargo, esto no es posible debido, entre otras cosas, a que los filtros utilizados durante el proceso de reconstrucción D/A no son ideales y poseen una pendiente, de modo que si solamente se aplicase el criterio de Nyquist, existiría superposición entre los valores que toma $Y(nT)$ para cada valor de n . Es por esto que, en aplicaciones de audio profesional se utiliza como por lo menos, la frecuencia de 44.1KHz, que es un poco mayor que el mínimo exigido para cumplir con el criterio de Nyquist).

Si el criterio de Nyquist se cumple en el proceso de muestreo (con las consideraciones prácticas pertinentes), no existe pérdida de información. Un análisis de frecuencia de $y(t)$ será correspondiente con el análisis de la función de entrada, mientras que dicho análisis sea acotado a frecuencias menores que $f_s/2$ (es decir, frecuencias menores que 22,05KHz). Sin embargo y tal como se observa en la figura anterior, en frecuencias sobre $f_s/2$ la función $y(t)$ tiene componentes espectrales que la función de entrada $x(t)$ no tiene. Dicha frecuencias superiores ($n f_s$) son eliminadas mediante filtros durante el proceso de reconstrucción D/A, principalmente pensando en la posibilidad de seguir procesando la señal en etapas posteriores, etapas que, de no tener una respuesta lineal, provocarían una enorme distorsión de intermodulación.

La importancia del criterio de Nyquist se hace evidente cuando dicho criterio no se cumple, es decir, cuando existen componentes espectrales en la señal de entrada cuya frecuencia es mayor que la mitad de la frecuencia de muestreo. Como ejemplo, si se considera un tono puro de frecuencia 1000Hz que es muestreado con una frecuencia de 1333 Hz. De acuerdo con la ecuación:

$$Y(nT) = \sum_{k=-\infty}^{+\infty} X(f - k f_s) e^{-j 2 \pi k n T f_s} = X(f - n f_s)$$

(integral entre $- \infty$ y $+$; sumatorias de $n = - \infty$ hasta $n = + \infty$), el espectro consiste en señales laterales alrededor de armónicos de 1333Hz. Las señales laterales se encuentran desplazadas por la frecuencia de la señal original, ± 1000 Hz. El espectro contiene ahora componentes de 333Hz, mientras que la entrada no tenía dicha frecuencia. Este fenómeno es conocido como aliasing. Al reproducir la señal $y(t)$ de éste ejemplo en un convertidor D/A, la componente de 333Hz aparecerá como una notoria distorsión que, además, no tiene relación armónica con la señal original.

En orden de eliminar los efectos del aliasing, los sistemas de digitalización incluyen filtros pasa bajos, llamados filtros anti aliasing, en el circuito electrónico que precede al conversor A/D. Este filtro tiene como propósito eliminar todas las componentes con frecuencias sobre la mitad de la frecuencia de muestreo en la señal de entrada. En la práctica, los filtros anti aliasing normalmente poseen una frecuencia de corte muy próxima a la mitad de la frecuencia de muestreo, lo que significa que deben

tener una pendiente muy abrupta.

En muchos equipos el filtro anti aliasing y el filtro de reconstrucción tienen las mismas características.: presentan un rizado tanto en la banda de paso como en la banda de rechazo. De acuerdo con estudios como los de Lagadec y Stockham, los filtros con rizado causan un efecto de dispersión: la señal de salida resulta desestabilizada en el tiempo y aparecen pre ecos en algunas señales. Para reducir éste efecto, se utilizan filtros ligeramente distintos para el filtro anti aliasing y de reconstrucción.

Para lograr que la pendiente del filtro sea abrupta, generalmente se conectan en cascada varias secciones de filtrado con la misma frecuencia de corte, por lo que la respuesta de fase de éstas secciones se acumula(los filtros pasivos generalmente poseen inductancias). Si bien es posible construir un filtro de respuesta de fase lineal, libre de rizado y con una atenuación requerida en la banda de rechazo, por la complejidad de diseño y el costo, resulta mas eficiente utilizar técnicas como el sobremuestreo (oversampling).

En la figura siguiente, puede comprobarse que el empleo de una frecuencia de muestreo elevada da lugar a la aparición de un amplio intervalo entre la banda base y la primera banda lateral.

En éste caso los filtros anti aliasing y de reconstrucción pueden tener una caída suave, que causa menor perjuicio en la linealidad de la respuesta de fase en la banda base, Y además puede utilizarse una configuración Butterworth, que no presenta rizado ni dispersión.

1.1.3.2– Sobremuestreo

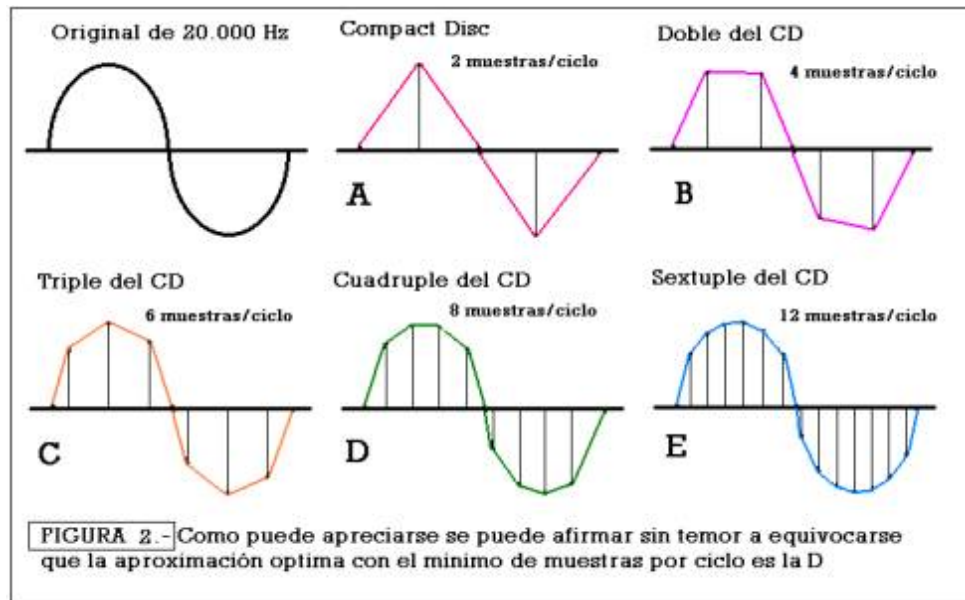
En principio, sobremuestreo significa que la interface entre la señal análoga y digital se realiza con una frecuencia de muestreo más alta que la utilizada para almacenar los datos digitales. Como ejemplo, en un reproductor de discos compactos con sobremuestreo de cuatro veces, tres muestras sucesivas son interpoladas en la salida, entre lecturas sucesivas de las muestras almacenados como datos en el disco. Mientras en el disco entrega 44.100 muestras (por canal) cada segundo, el mecanismo de interpolación entrega 176.400 muestras por segundo al convertor D/A. La interpolación se realiza agregando tres muestras con valor cero por cada valor de entrada, y procesando el resultado a través de un filtro digital. Los coeficientes en el filtro digital ponderan cuatro muestras de entrada como un promedio para generar con esto una muestra de salida interpolada. Los filtros digitales pueden ser muy precisos, no tienen ruido y no presentan distorsión de fase. Se aplica un filtro digital pasa bajo al tren de muestreo de 44.1KHz en la frecuencia de corte cercana a 20KHz, tal como lo requiere el teorema de Nyquist. En la etapa siguiente, la frecuencia de muestreo del convertidor D/A es cuatro veces mayor y los requerimientos en el filtro de reconstrucción análogo solo son cortar en $176.400/2$, proceso que es más fácil de realizar para la mayoría de los filtros.

Un principio similar se aplica en la grabación. El convertidor A/D trabaja con una frecuencia de muestreo mayor que la frecuencia con que eventualmente serán almacenados lo datos. De éste modo, un filtro pasa bajos análogo común puede ser usado como filtro anti aliasing. Un dispositivo conocido como *decimator* , que descarta muestras antes del almacenamiento, incluye un filtro digital que se encarga de remover las componentes de entrada con frecuencias mayores que la mitad de la frecuencia con que son almacenadas las muestras y que podrían provocar aliasing en la señal almacenada.

1.1.3.3 Elección de la Frecuencia de Muestreo

Una frecuencia de muestreo elevada puede ser adecuada para transmitir una forma de onda como la que se observa en la figura (a), pero en la figura (b) se detecta que, la frecuencia de salida tiene una forma de onda y frecuencia diferente a la señal original.

Esto lleva a la idea de que, el criterio de Nyquist es solo el comienzo del proceso que debe seguirse para alcanzar un valor aceptable de la frecuencia de muestreo. Por ejemplo, si se desea muestrear una señal senoidal de 15KHz (frecuencia, por cierto presente en los armónicos y sobre armónicos de algunos instrumentos y que dan cierto carácter o identidad al sonido), la reconstrucción de dicha forma de onda se ve seriamente deteriorada al utilizar una frecuencia de muestreo de 44,1KHz, tal como en los discos compactos de audio que se comercializan actualmente. El caso mas extremo, pese a que éste fenómeno afecta a todas las frecuencias altas (es decir, de 12 KHz hacia arriba, aproximadamente), es el caso de la forma de onda resultante para una señal senoidal de 20KHz, donde apenas se alcanzan a tomar 2 muestras para cada ciclo. Si se utilizan frecuencias de muestreo mayores, la reconstrucción de la forma de onda en la salida resulta mas natural o mas fiel a la forma de onda original.



Como se mencionó anteriormente, la pendiente de los filtros de reconstrucción disponibles obliga a los diseñadores a elevar la frecuencia de muestreo por encima del valor teórico de Nyquist. En equipos de consumo, cuanto más baja sea la frecuencia de muestreo, mejor, ya que el costo del medio se reduce. Con productos profesionales existe la necesidad de operar a velocidad variable para realizar correcciones de tono. Cuando se reduce la velocidad de un magnetófono digital, la frecuencia de muestreo de la señal reproducida por la cinta disminuye, por lo que, con una frecuencia de muestreo mínima la primera frecuencia imagen puede ser suficientemente baja como para atravesar el filtro de reconstrucción. Este problema se soluciona elevando la frecuencia de muestreo sin cambiar la respuesta de los filtros.

De esto se deduce que los magnetófonos con ajuste de velocidad, generalmente los de cabeza estacionaria, deben utilizar una frecuencia de muestreo superior.

2.-Almacenamiento de la información

2.1.Principios básicos de la grabación digital

Los principios básicos de la grabación digital son simples. Dado que el medio presenta solamente dos estados (0 o 1, H o L), la forma de onda de grabación en una cinta será típicamente una corriente cuya dirección se invierte, pero cuya magnitud permanece constante. Para proporcionar la mejor S/N en la reproducción , la corriente requerida es un poco menor que la necesaria para saturar la cinta, ya que la saturación produce campos de interferencia alrededor de la cabeza y diafonía sobre las pistas adyacentes.

La tarea inicial de los circuitos de reproducción es reconstruir la forma de onda grabada. La amplitud de la señal no es determinante, lo que importa es el momento en que la corriente de escritura, por tanto, el flujo en la cinta, se invierte. Este momento puede determinarse localizando los picos de los impulsos de reproducción. Con altas velocidades de datos esto puede realizarse satisfactoriamente, diferenciando la señal y buscando los pasos por cero.

En este caso, el sistema produce ruido entre los picos. Este problema se soluciona mediante un detector de picos con puerta, en el que solo se considera los cruces por cero de un impulso que excede del umbral

2.2.-El Casete digital doméstico

El formato SDAT utiliza cinta de 3,81 mm de ancho y 10 m de espesor, montada en un casete de aproximadamente el tamaño de un casete compacto, con un contenido de unos 130 m. Al igual que en el RDAT(el SDAT y el RDAT son producto de la Conferencia DAT), hay varios modos de operar que permiten al usuario grabar con frecuencias de muestreo de 48 y 32 KHz, y tratar cintas pregrabadas a 44,1KHz. El formato SDAT utiliza códigos de producto Reed– Solomon, y en él las palabras de código a lo largo de las pistas actúan como detectores de error.

2.3 El Disco Compacto

La capa de información del disco compacto es un espejo ópticamente plano sobre el cuál se levantan algunos abultamientos microscópicos. Una fina cubierta de aluminio hace a ésta capa reflectante.

Cuando un fino haz de luz se enfoca sobre la capa de información, la presencia de los abultamientos afecta a la forma en que la luz es reflejada, detectándose éstas variaciones en la lectura del disco.

La lectura óptica del CD es un proceso sin contacto físico y , por lo tanto no hay mecanismo de desgaste. El sistema óptico se enfoca sobre la capa de información, la cuál está lo suficientemente por debajo de la superficie como para que el polvo y los arañazos caigan fuera del punto de enfoque y sus efectos sean minimizados. Combinando una potente técnica de corrección de errores, el resultado es un dispositivo altamente resistente al uso. El acceso a la banda deseada es rápido, y en el CD se mejora mediante el uso de subcódigos introducidos entre las secuencias de datos, lo que permite que el comienzo de cada programa sea localizado de una manera precisa. El pequeño tamaño del disco permite la fabricación de reproductores portátiles y de uso en automóviles.

El proceso del disco compacto desde el cortado, pasando por el prensado y la lectura no produce degradación en la señal musical, ya que simplemente transporta una serie de números que son exactamente aquellos que fueron registrados en el master. La única etapa en la reproducción donde pueden aparecer diferencias subjetivas en la calidad del sonido en el convertidor D/A.

Los pasos principales en la fabricación de un disco compacto se muestran a continuación:

El paso inicial consiste en producir un disco de cristal virgen de alrededor de 220mm de diámetro y 6mm de espesor. El disco se lava primero con una solución alcalina, después con un disolvente de fluorocarbono y se seca antes de pulir su superficie. Después se lleva a cabo un proceso de limpieza usando una mezcla de agua desionizada y alcohol isopropilo, con un lavado final de fluocarbono. A continuación se inspeccionan las irregularidades de la superficie del disco, mediante un rayo láser y monitoreando la reflexión según rota el disco. Los discos rechazados vuelven a ser pulidos, los discos aceptados pasan y se les aplica una capa adhesiva seguida de una cubierta fotoresistente. El espesor de la capa se controla con precisión, ya que afectará a los abultamientos del disco final y, por tanto, se lleva a cabo un rastreo óptico para comprobar que no hay defectos en esta capa, lo cual causaría errores de datos en el producto final.

El proceso de cortado se muestra a continuación:

Un láser de ion argón o de cadmio helio se enfoca sobre la capa resistente según revoluciona el disco. El enfoque se obtiene mediante otro láser de neón helio separado que comparte el mismo sistema óptico que el láser anterior. La capa resistente es insensible a la longitud de onda del láser. La intensidad del láser se controla con un dispositivo conocido como modelador acústico óptico, que es excitado por un codificador. Cuando el dispositivo está en estado de reposo, la luz pasa a través de él, pero cuando la superficie se excita con vibraciones de alta frecuencia, la luz se dispersa. A medida que aumenta el radio de la pista, la velocidad de rotación disminuye proporcionalmente, tal que la velocidad lineal del haz sobre el disco permanece constante. (esto genera un tiempo de reproducción bastante mayor del que se obtendría con una velocidad de rotación constante). Debido a las diminutas dimensiones de la estructura de la pista, el cortado debe realizarse con extrema precisión. Se usa refrigeración por aire en el eje y en la cabeza del láser y la máquina completa lleva amortiguadores para evitar vibraciones de la estructura que afecten al formato de la pista.

El disco virgen se trata luego con un revelado que endurece la resistencia de las áreas no expuestas. Un lavado elimina las áreas expuestas y crea los huecos de la superficie de la capa resistente. El proceso se detiene cuando un láser de monitorización detecta la profundidad correcta del hueco. Se da una capa de plata por evaporación para hacerlo eléctricamente conductivo. De manera similar a la producción de los discos de vinilo, se fabrica un disco padre por niquelado y se vuelve a niquelar para producir un disco madre. A partir de un disco madre se fabrica un determinado número de discos matrices que serán usados para estampar los discos. El proceso de estampado transfiere la estructura del hueco al material del disco transparente. Dado que cada paso del proceso crea una imagen especular, el disco será una exacta reproducción del original y tendrá huecos sobre su superficie. Una fina capa de aluminio se aplica a la capa reflectante seguida por una capa protectora de laca. El disco se centra ajustando ópticamente el recorrido de pista y el agujero central puede ser perforado y ajustado periféricamente, aunque algunos estampadores producen el disco ya perforado y ajustado completamente. La etiqueta se coloca sobre la cara de laca y puede imprimirse o pegarse directamente. Como el disco se leerá por el lado opuesto a la etiqueta, los huecos aparecerán ahora como abultamientos desde la cara de lectura.

La siguiente figura muestra las especificaciones mecánicas del CD.

Dentro de un diámetro total de unos 120 mm, el área de programa ocupa una banda de 33mm de ancho, entre 50 y 116mm de diámetro. Las áreas de entrada y de salida aumentan el ancho de ésta banda a unos 35,5mm. Como el salto de pista es una constante de 1,6 m, habrá:

$$35,6 \times 1000 / 1,6 = 22.188$$

pistas que cruzan el radio del disco. Como la pista es una espiral continua, la longitud total de las pistas resultará de multiplicar el número anterior por la circunferencia media

$$\text{longitud} = 2 \times 22.188 \times (58,5 + 23) / 2 = 5,7 \text{ km}$$

El ancho de banda (datos por segundo) requerido por el audio digital es tal, que es obligatorio utilizar grabaciones de alta densidad si se desea obtener un tiempo de reproducción razonable, y esto implica longitudes de onda cortas. El uso de un láser enfocado a distancia sobre el disco permite que las grabaciones de longitudes de onda corta se reproduzcan sin contacto físico, mientras que las grabaciones magnéticas requieren un contacto físico que implica mecanismos sometidos a desgaste.

La capa de información del CD se lee a través del espesor del disco. Esta solución hace que el haz de lectura penetre a través de un área lo más grande posible.

A pesar del diminuto tamaño del punto de enfoque de, aproximadamente 1,2 m de diámetro, la luz penetra a través de un círculo de 0,7mm de diámetro. Como resultado de esto. Las partículas de polvo que eventualmente existen en la superficie del disco deberían ser muy grandes como para alcanzar a oscurecer el haz de luz. Este método tolera muy bien los arañazos en la superficie, y en caso extremo puede aplicarse un pulidor de metales. Por el contrario, el lado de la etiqueta es mucho más vulnerable que el lado de lectura, ya que la capa de laca solo alcanza a 30 m de espesor. Por ésta razón no es recomendable escribir sobre la etiqueta del CD, ya que la presión ejercida ,por ejemplo, por un bolígrafo, puede distorsionar la capa de información; y se ha encontrado que los disolventes utilizados en los lapiceros marcadores penetran la laca y originan daños en los datos.

3.–Codificación y mecanismos de corrección de errores

El proceso de grabación en un CD necesita de un proceso de codificación, en parte para incorporar mecanismos de corrección de errores, información anexa distinta a la información de audio existente y para facilitar el proceso de escritura sobre la superficie del disco. Durante el proceso de lectura se realiza el proceso inverso, la decodificación.

Existen dos tipos de errores que pueden aparecer en el CD:

- *errores aleatorios* , definidos como errores que se presentan aislados o en grupos de no más de 17Bit. Un caso típico es un bulto mal impreso en el disco.
- *errores de ráfaga* que se presentan en grupos mayores de 17Bit. Pueden ser producidos, por ejemplo, por un arañazo.

En el CD se usan bits de paridad (los llamados bytes P y Q) y entrelazado de la información, que siguen el código CIRC, y modulación EFM. Los bytes de paridad P y Q Se calculan a partir del código Reed–Solomon e introducen redundancia para la corrección de errores aleatorios. El entrelazado es un reordenamiento de la información y se realiza con la idea de que, al reponer las muestras en sus posiciones originales, los errores del disco quedarán distribuidos entre datos válidos, siendo muy probable poder corregirlos con los bytes de la paridad. Mediante la modulación EFM se intenta pasar la información existente a combinaciones de bits con pocas transiciones entre cero y uno para facilitar la grabación física en el disco.

La unidad básica de información son 12 muestras de sonido (cada una de 16 bits, como se ha visto), 6 muestras del canal derecho y 6 del canal izquierdo, que son $6 \times 2 \times 16 = 192$ bits=24 bytes. Estos 24 bytes forman un bloque, llamado *inicial* , y se codifican según el código CIRC.

3.1–Codificación CIRC

.

El Cross Interleave Reed–Solomon Code (CIRC) es un método de detección y corrección de errores que consiste en añadir 8 bytes de paridad, según la codificación Reed–Solomon, en dos pasos, llamados C1 y C2 (4 bytes en cada paso), separados por un entrelazado. El cálculo de la paridad es idéntico en ambos pasos. La diferencia está en cuántos bytes se usan para obtenerlos. El C2 usa 24 bytes (los 24 del bloque inicial) y la C1 usa 28 bytes (24 del bloque inicial más los 4 bytes de paridad C2).

La codificación Reed–Solomon permite corregir hasta 2 bytes de una palabra de 28 o 32 bytes. Para explicar como se logra ésta corrección, es necesario introducir el concepto de distancia mínima *dmin* de un código: es el número *mínimo* de símbolos que se han de cambiar en una palabra de código para pasar de una palabra válida a otra también válida (no necesariamente contiguas). Con *dmin* =5, que es precisamente la que tiene la codificación Reed–Solomon, es posible corregir dos símbolos erróneos. En

este caso los símbolos son bytes y la palabra, a la que añadir la paridad, es de 24 (C2) o 28 (C1) bytes. Cada palabra (sin paridad) legal varía de una a otra en por lo menos un símbolo (1 byte). Así que $d_{min} = 1$. Es la paridad la que da la distancia deseada: el código Reed–Solomon calcula los 4 bytes de manera que varíen de una palabra a otra y permitan $d_{min} = 5$. Esto permite finalmente corregir 2 bytes en palabras de 28 y 32 bytes (que son las de 24 y 28 bytes más los 4 bytes de paridad).

Hay otro procedimiento que permite corregir hasta 4 bytes (en general $d_{min} - 1$) en una palabra de 28 o 32 bytes. La condición para aplicar este procedimiento es restrictiva: hay que conocer a priori en qué posiciones de la palabra están los bytes erróneos. En el supuesto que existe una palabra Y1 con 4 bytes marcados como erróneos, dado que la distancia entre palabras de código válidas Xi es cinco, sólo habrá una palabra X1 en la que coincidan los bytes válidos, y Y1 se considerará un X1 erróneo. Las posiciones con bytes erróneos se llaman posiciones de borrado (*erasure position*).

La secuencia completa de la creación de una trama CIRC es : a) retardo y ordenación, b) codificación Reed–Solomon C2 (cálculo de la paridad Q), c) entrelazado, d) codificación Reed–Solomon C1 (cálculo de la paridad P) y e) retardo con inversión lógica de paridad. En la decodificación la secuencia es la inversa. En el CIRC los bloques de 24 bytes se codifican en paralelo, es decir, se tratan los bytes de 24 en 24.

a.– Retardo y ordenación.

Primero, se retardan las muestras pares dos bloques. Cuando se habla de retardar muestras en n bloques se ha de entender que es un retardo de la señal digital de $n \times D$, con $D = 1 / [88200 \text{ (muestras / segundo)}] \times 12 \text{ (muestras / bloque)} / [24 \text{ (bytes / bloque)}] = 5.67 \mu\text{s}$. Más adelante los bloques tendrán 28 y 32 bytes, números que han de sustituir a 24 en el anterior retardo. El siguiente paso es un reordenamiento (simétrico entre las 6 primeras y las 6 últimas muestras) para separar las pares de las impares. Esto permite, en la decodificación, distribuir entre bytes correctos los bytes erróneos no corregibles con las paridades P y Q (pero marcados como erróneos por ellas) para su interpolación.

b.– Codificación C2 (adición de la paridad Q).

Ahora, los 24 bytes del bloque ordenado se usan para calcular 4 bytes de paridad, llamada paridad Q. La paridad Q está diseñada para corregir los errores de ráfaga marcados por la paridad P y también los errores aleatorios que no se pudieron corregir con ella. Los 4 bytes Q se sitúan en el centro de los 24 bytes del bloque primario para aumentar la distancia entre muestras pares e impares, y así mejorar la interpolación en caso de una ráfaga de errores. A este bloque de 28 bytes le llamaremos bloque codificado C2.

c.– Entrelazado.

Después, se realiza un entrelazado cruzado: los 28 bytes que forman el bloque codificado C2 se retardan con diferentes períodos múltiplos de 4 bloques: el primer byte del bloque no se retarda, el segundo se retarda 4 bloques, el tercero 8 bloques y así hasta el byte vigésimo octavo, que se retarda en 108 bloques. De esta manera cada uno de los 28 bytes se almacena en otros tantos bloques distribuidos entre 109 bloques y los 28 bytes del bloque resultante proceden de 28 bloques codificados C2 diferentes. Este entrelazado está diseñado para, en la decodificación, aislar los errores de ráfaga, es decir, convertirlos en errores aleatorios. Al decodificar primero se usa la paridad P, se desentrelaza, y luego se usa la Q. Por tanto es la paridad Q la que puede corregir los errores de ráfaga que, en éste paso, son convertidos en errores aleatorios.

d.– Codificación C1 (adición de la paridad P).

De los 28 bytes resultantes se calculan 4 bytes más de paridad según CIRC C1. La paridad P corrige la mayor parte de errores aleatorios y detecta y aísla los errores de ráfaga para que los corrija la paridad Q. En el decodificador, si se detectan más errores, se marca todo el bloque con un *erasure flag* para que la paridad Q actúe sobre ellos.

e.– Retardo con inversión lógica de paridad.

Los bytes impares se retardan en un bloque. Esto evita que los errores aleatorios corrompan más de un byte por bloque. Los bytes de paridad P y Q, como último paso, son invertidos lógicamente para tener siempre bytes diferentes de cero. Esto ayuda en la lectura de datos cuando hay zonas de silencio de audio. El bloque de 32 bytes resultante es la *trama CIRC*.

3.2– Subcódigo.

Después de la codificación CIRC se añade un byte de subcódigo por bloque de 32 bytes. Los 8 bits de subcódigo se llaman P, Q, R, S, T, U, V y W. El CD sólo usa los bits P y Q (que no tienen nada que ver con las paridades P y Q). Estos bits de subcódigo incluyen información sobre el número de canciones del disco, su comienzo, final y duración, el *Copyright*, etc. Los bits de subcódigo se usan agrupando 98 bloques para obtener $98 \times 8 = 784$ bits.

3.3– Modulación EFM.

EFM significa *modulación de ocho a catorce*: los 33 bytes (264 bits) escogidos en grupos de 8 en 8 bits se traducen a grupos de catorce bits, llamados *channel bits*. ¿Qué factores se tienen en cuenta para escoger las palabras de 14 bits? La asignación debe ser biyectiva y debe minimizar el número de transiciones entre ceros y unos.

Para que los ceros y unos (abultamientos y no abultamientos) tengan longitudes controladas, se exige que las palabras de 14 bits tengan más de 2 pero menos de 10 ceros seguidos. De las posibles $2^{14} = 16384$ combinaciones 267 satisfacen este criterio. Sólo se necesitan $2^8 = 256$.

3.4 – Bits de unión.

Los grupos de 14 bits se unen con 3 bits. Dos de los bits de unión son siempre 0, para evitar tener unos sucesivos entre las palabras de 14 bits. El tercer bit, que puede ser cero o uno, se añade para ayudar a la sincronización del reloj y para reducir los componentes de baja frecuencia de la señal digital. En la demodulación los tres bits de unión se desechan.

Cada byte de la trama CIRC de 32 bytes, al modularlo y añadirle los bits de unión, pasa a tener 17 bits. Este grupo de 17 bits es el que se grabará en el disco. Los errores aleatorios fueron definidos como aquellos errores que como máximo tenían 17 bits. Entonces, un error aleatorio como máximo afectará a dos símbolos contiguos de 17 bits. El codificador CIRC causa retardo en los bytes impares en 1 bloque como último paso antes de invertir las paridades P y Q. Así que al leer el disco, si se comete un error de 17 bits, después de la demodulación EFM y de desechar los bits de unión, los dos bytes contiguos se asignan a bloques diferentes al ser retardados los símbolos pares. Por tanto como máximo, ante un error aleatorio, se pierde un byte por bloque. Si ocurre un error de ráfaga se perderán más bytes, y la paridad Q, en la medida de lo posible, los corregirá.

3.5– Adición de la sincronización y formato de la trama.

Los grupos de 564 bits (33×14 *channel bits* + 34×3 *bits de unión*) resultantes deben ser diferenciados o delimitados unos de otros. Se usa por tanto una palabra de 24 bits de sincronización al inicio de cada

grupo, que es única y no puede confundirse con ningún dato. La palabra en cuestión, a la que se habrá de añadir los bits de unión, es 1000 0000 0001 0000 0000 0010: 3 transiciones separadas por 10 ceros. El formato de la trama final que se graba en el disco está formado por 588 bits: 24 bits de sincronismo, 336 (24 x 14) bits de datos, 112 (8 x 14) bits de corrección de errores, 14 bits de subcódigo y 102 (34 x 3) bits de unión.

3.6– Velocidad de transferencia de la información.

El proceso de muestreo se realiza a 44.1 kHz, obteniendo de cada muestra 32 bits, 16 del canal izquierdo y otros 16 del derecho. Por lo tanto, la fuente da datos a una velocidad de $44.1 \times 1000 \times 32 = 1411200$ bits/s o 172.26 kbytes/s (1 kbyte = 1024 bytes). Esta será la velocidad de transferencia que tendrán los datos (la música) una vez leídos y decodificados en el lector de CD.

La velocidad de transferencia de la trama final del disco al lector es mayor, ya que en el mismo tiempo tenemos que enviar mucha más información. La trama se crea en todo el proceso de codificación anteriormente explicado. En este proceso hemos pasado de los 192 (24 x 8) bits de datos iniciales a los 588 bits del formato final, de manera que la velocidad de transferencia de bits al lector hecha por el cabezal son $1411200 \times 588 / 192 = 4321800$ bits/s, o 527.5 kbytes/s.

A continuación se presentan dos ejemplos de errores posibles en la lectura del disco: que se lean dos bytes contiguos erróneos (error aleatorio) y que se lean 23 bytes contiguos erróneos (error de ráfaga).

a.–2 bytes erróneos. Después de leer la información del disco, de desechar los bits de unión y de demodular la EFM, el bloque de 32 bytes se retarda y se invierten las paridades. El retardo hace que los bytes dejen de ser contiguos y haya solo un error por bloque. El bloque pasa al decodificador C1. Al comprobar los bytes de paridad P detecta un byte erróneo. El byte podría ser un byte de paridad, sin merma de efectividad. Como se ha comentado antes, el código Reed–Solomon permite corregir hasta 2 bytes, así que corrige el byte detectado. El bloque se desentrelaza. En el decodificador C2 se quita la paridad Q y en el último paso se reordena y retarda, para obtener 24 bytes de datos correctos.

b.– 23 bytes erróneos . En este caso, después del retardo tenemos un bloque con 11 bytes erróneos y otro con 12 bytes erróneos. El decodificador C1 detecta más de 2 bytes erróneos por bloque y marca cada bloque con un *erasure flag* , sin poder corregirlos. Los bytes erróneos, por tanto, están indicados como tales para el resto de etapas. Al desentrelazar los retardos que se aplican son los inversos que en el entrelazado, es decir, el primer byte se retarda 108 bloques, el segundo 104, y así hasta el último, que no se retarda. El resultado es que los bytes erróneos (marcados con un *erasure flag* y por tanto en posiciones conocidas) quedan distribuidos entre bytes correctos, y los bloques resultantes tienen menos de 3 bytes erróneos, y son corregibles por el decodificador C2.

Con es sistema CIRC se pueden corregir hasta 16 bloques contiguos o 512 bytes. Cuando el error es mayor se opta por la interpolación de los bytes erróneos si éstos, que están marcados y por tanto en posiciones conocidas, se encuentran entre datos correctos. Si incluso la interpolación es imposible, el decodificador opta por pasar los bytes erróneos como silencio. Estos silencios suelen ser muy cortos y normalmente no se notan.

4 – Conclusiones

Las principales ventajas de la grabación digital de audio, pueden resumirse como sigue:

- La calidad de una grabación digital de audio es independiente de la cabeza y del soporte en un sistema correctamente diseñado. La respuesta de frecuencia, linealidad y ruido están determinados exclusivamente por la calidad de los procesos de conversión.

- Una grabación digital es una serie de números y como tal puede ser copiada sin degradación, en un número indefinido de generaciones.
- El uso de técnicas de corrección de errores elimina los efectos causados por los *dropouts* ; el empleo de un corrector de base de tiempos en la reproducción permite eliminar efectos de *wow* y *flutter*, además de permitir la sincronización precisa entre varios equipos. (punto sumamente importante)
- Existen además, en el campo del audio digital ,enormes posibilidades de procesamiento y edición no destructiva y no lineal, posibilidades que no existen en el campo de la grabación analógica.
- No existe comparación posible entre la relación costo –beneficio entre las tecnologías análogas y digitales. Si bien existen equipos de grabación analógica con excelentes prestaciones (buena relación señal ruido, excelente respuesta de frecuencia y baja distorsión), el costo de dichos equipos es sumamente elevado, sobretodo al compararlo con el costo de una sistema de grabación digital profesional.

Si bien todavía pueden existir algunas ventajas subjetivas que benefician a la grabación analógica, dichas ventajas se refieren al estado actual de la grabación digital, estado que ha sido mantenido artificialmente principalmente por la industria del disco compacto (16 bit, 44,1KHz). Los avances técnicos tanto en los procesos de conversión y en otros campos anexos, como es la industria de los computadores, permiten vislumbrar nuevas posibilidades para el audio digital, donde, por ejemplo, elevando la frecuencia de muestreo se logre por fin una reconstrucción fiel de la forma de onda original para todo el espectro musical. (en el entendido que existen frecuencias que, pese a no estar dentro des rango de audición humana, pueden ser percibidas por el ser humano mediante algún otro mecanismo). Entre los nuevos formatos que están a la espera de ser introducidos al mercado, es posible mencionar el SACD (super audio cd), con un sistema de digitalización diferente al actual PCM, llamado DSD (la principal característica de éste lenguaje tiene que ver con el error de cuantización); y el DVD– audio (digital versatile disc– audio), ambos sistemas trabajan con frecuencias de muestreo superiores a 80.000 Hz y con resolución equivalente a 21 bit en el caso del SACD (resolución que, por cierto no es comparable con la resolución actual de los sistemas PCM) y 24 bit el caso del DVD–audio y ofrecen posibilidades de reproducción multicanal.

Bibliografía

- El arte del audio digital / John Watkinson, Instituto oficial de radiotelevisión española, 1989.
- Signals, sound, and sensation / William M.Hartmann, Springer–Verlag, 1998.

Digitalización de Señales de Audio