

Sonido digital

Podemos definir al sonido como la interpretación que hace nuestro cerebro de las variaciones de presión que genera un objeto vibrante en determinado medio, habitualmente, el aire, sobre el tímpano del oído. El sonido viaja en forma de ondas y cuando mayor es la onda más fuerte es el sonido.

La onda de las partículas de aire forma una onda de sonido.

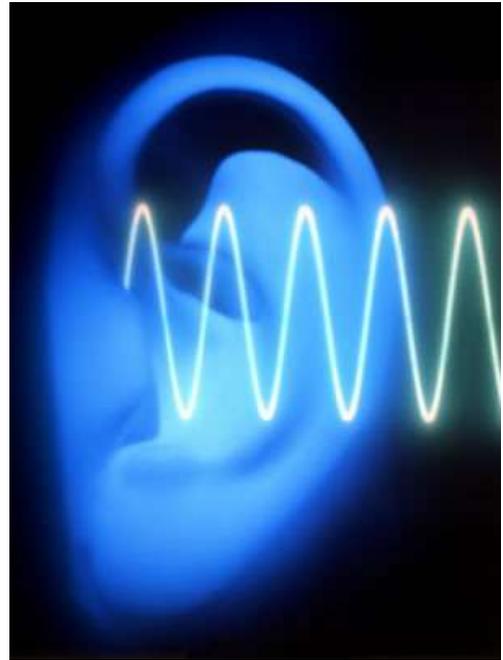
Visualización de las ondas de sonido

Si por ejemplo lanzamos un objeto al agua, una piedra, o simplemente el salto de una rana al agua, vemos que se forman unas ondas en el agua. Las ondas se mueven hacia fuera desde el punto que se lanzó el objeto y lo hace a una velocidad constante, a esta velocidad la denominamos "frecuencia" y a una altura determinada por el tamaño del objeto lanzado. El agua se mueve o



desplaza por la presión que crea el impacto del objeto lanzado, ese movimiento en el agua son lo que denominaríamos las ondas en sonido.

Para que esa vibración sea audible para un ser humano, la oscilación debe producirse entre 20 y 20 mil veces por segundo. Es decir que nuestro rango de audición va de 20 Hertz (ciclos) a 20 Kilohertz.



Características: altura, timbre e intensidad:

ALTURA: determinan si los sonidos son más graves o agudos. Está directamente relacionada con la frecuencia de oscilación. Más ciclos por segundo provocarán un sonido más agudo, menos ciclos, más grave.

TIMBRE: es lo que permite identificar y reconocer una voz o un instrumento musical. Es consecuencia de que los sonidos naturales son complejos y están compuestos por varias frecuencias que se combinan. La frecuencia más grave de ese sonido complejo se la denomina frecuencia base y es la que determina habitualmente la altura del sonido. El resto de las señales se denominan armónicos y son los que le otorgan identidad al sonido.

INTENSIDAD: depende de la diferencia entre las presiones máximas y mínimas que puede alcanzar el sonido. Se mide en decibeles y podemos asociarlas a la idea de volumen. Cuanto mayor amplitud de onda, más elevado será el volumen.

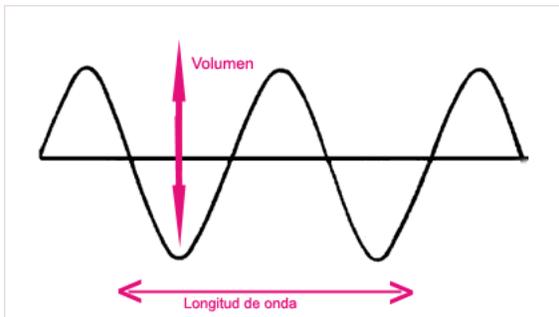
Longitud de onda y la frecuencia

La altura o amplitud de la onda representa su *volumen*, que queda determinado por la potencia de la fuente de sonido, o la fortaleza de la vibración que creó la onda de presión. Se mide en Db o *Decibeles*.

Siguiendo el anterior ejemplo de lanzar un objeto al agua, cuando mayor sea el objeto lanzado, mayor serán las ondas que se crearan en el agua.

La extensión de una onda se denomina *longitud de onda*, y en número de ondas que se recogen en un mismo punto en un espacio de tiempo, se denomina *frecuencia de la onda*, y esta se mide en Hercios (HZ o KHZ). El oído humano capta un intervalo de frecuencias de 20 HZ y 20.000 HZ.

Las frecuencias bajas de 20 Hz, se denominan infrasonidos y las frecuencias más altas 20.000 Hz, se definen como ultrasonidos.



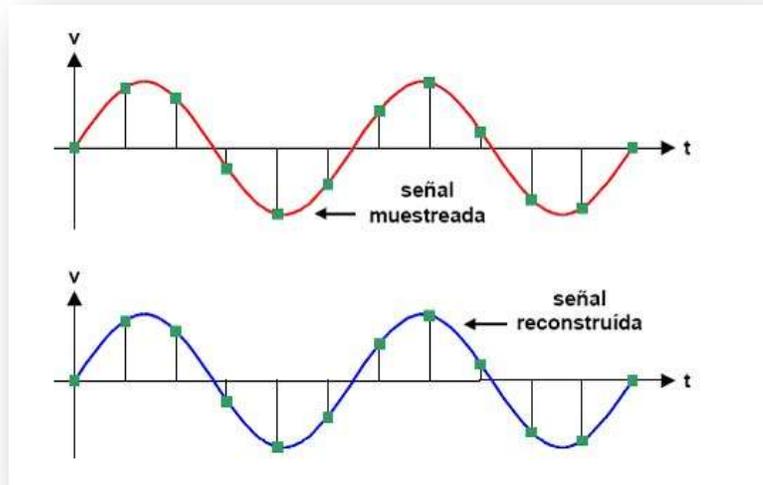
Como se convierte el sonido en digital

El sonido digital es un poco más complejo que el analógico, ya que no se almacena mediante oscilaciones de onda, sino que se ha de convertir en ceros y unos, lenguaje digital (código binario o máquina).

El proceso de digitalización se compone de dos fases: *muestreo y cuantización*. En el muestreo se divide el eje del tiempo en segmentos discretos: la frecuencia de muestreo será la inversa del tiempo que medie entre una medida y la siguiente.

Se trata entonces de la grabación de la altura actual de la onda de sonido a intervalos regulares. La longitud de estos intervalos se denomina, **tasa de muestreo**. Cada nivel o pico es una muestra y cuando las coloca todas juntas ordenadas forman una representación digital de la onda de sonido.

Entonces, la **tasa o frecuencia de muestreo** es el número de muestras por unidad de tiempo que se toman de una señal continua para producir una señal discreta, durante el proceso necesario para convertirla de analógica en digital. Como todas las frecuencias, generalmente se expresa en hercios (Hz, ciclos por segundo) o múltiplos suyos, como el kilohercio (kHz), aunque pueden utilizarse otras magnitudes.



El sonido importado a un programa se representa como onda de sonido. Los picos representan la amplitud o intensidad del sonido durante un espacio de tiempo.

A la conversión del sonido de analógico a digital, se denomina Analog Digital Conversión o Conversión Analógico Digital.

En el momento de la toma de la muestra se realiza la cuantización, que, en su forma más sencilla, consiste simplemente en medir el valor de la señal en amplitud y guardarlo.

Será mayor la fidelidad de esa captura cuanto mayor sea sensibilidad de la escala, que está determinada por la cantidad de bits

Cuantificación entonces es el número de niveles habrá en cada espacio de tiempo.

Las muestras de cada intervalo se convierten en un nivel. Dependiendo de la cantidad de información que haya en cada muestreo, existirán más o menos niveles.

Es evidente que cuantos más bits se utilicen para la división del eje de la amplitud, más "fina" será la partición y por tanto menor el error al atribuir una amplitud concreta al sonido en cada instante.

Por ejemplo, una señal muestreada a **8 BITS OFRECEN 256 NIVELES DE CUANTIZACIÓN Y A 16, ENTREGA 65536 NIVELES**. El margen dinámico de la audición humana es de unos 100 dB. La división del eje se puede realizar a intervalos iguales o según una determinada función de densidad, buscando más



resolución en ciertos tramos si la señal que se trata tiene más componentes en cierta zona de intensidad, como veremos en las técnicas de codificación. El proceso completo se denomina habitualmente **PCM** (Pulse Code Modulation)

Un ejemplo es el de los **CD DE MÚSICA**, presentan un muestreo de **16 BITS, ESTO REPRESENTA 65.536 NIVELES DISTINTOS Y 44.1 KHZ.**, por segundo.

Tasa de Muestreo, pulsos por segundo

La tasa de muestreo se expresa en Hercios, que son las muestras o pulsos por segundo.

Por ejemplo si tenemos un archivo de sonido de 40 KHz, hacemos referencia al tiempo que existe entre muestra y muestra. 40 KHz se traduce en 40.000 hercios, muestras por segundo (Un KHz. equivale a 1000 hercios).

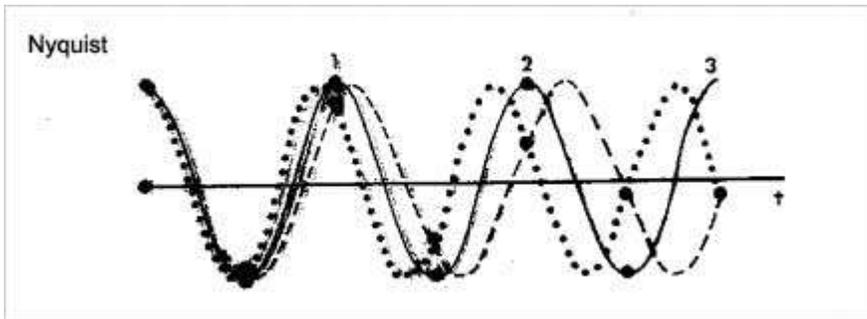
Tasa de muestreo recomendable

En audio, la máxima audiodiferencia perceptible para el oído humano joven y sano está en torno a los 20 kHz, por lo que teóricamente una frecuencia de muestreo de 40000 sería suficiente para su muestreo; no obstante, el estándar introducido por el CD, se estableció en 44100 muestras por segundo. La frecuencia de muestreo ligeramente superior permite compensar los filtros utilizados durante la conversión analógica-digital.

Hay que tener en cuenta que no todas las fuentes sonoras se aproximan a los 20 kHz que corresponden a esta frecuencia máxima; la mayoría de los sonidos está muy por debajo de ésta. Por ejemplo, si se va a grabar la voz de una soprano, la máxima frecuencia que la cantante será capaz de producir no tendrá armónicos de nivel significativo en la última octava (de 10 a 20 kHz), con lo que utilizar una frecuencia de muestreo de 44100 muestras por segundo sería innecesario (se estaría empleando una capacidad de almacenamiento extra que se podría economizar).

El estándar del CD-Audio está fijado en 44100 muestras por segundo, pero esto no significa que esa sea la frecuencia que utilizan todos los equipos. Los sistemas domésticos de baja calidad pueden utilizar tasas de 22050 muestras por segundo o de 11025 muestras por segundo (limitando así la frecuencia de los componentes que pueden formar la señal). Además, las tarjetas de sonido de los equipos informáticos utilizan frecuencias por encima o por debajo de este estándar, muchas veces seleccionándolas en función de las necesidades concretas (sobre todo, en aplicaciones de audio profesional).

Ley Nyquist

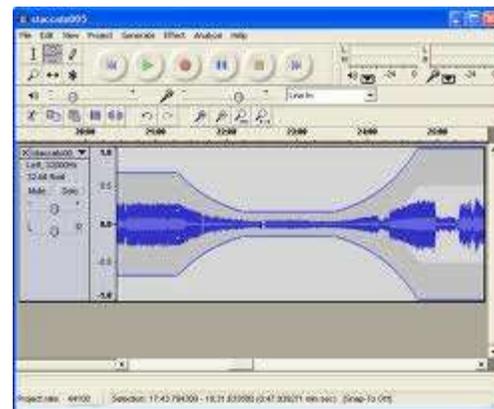


Existe una ley descubierta por Nyquist, que dice que si la tasa de muestreo es dos veces la frecuencia más alta que contiene el sonido, este se puede muestrear o samplar sin pérdida de calidad.

Poniendo por ejemplo la frecuencia más alta que podemos escuchar es de 20 KHz. (20.000 hz). Esto se traduce a que si muestreamos un sonido a 40 KHz, nadie notará la diferencia entre un sonido original a un sonido muestreado. Los discos compactos contienen sonido muestreado a 40 KHz., un nivel máximo de muestreo, dando como resultado una copia digital perfecta.

Edición de audio digital

El archivo de audio digitalizado y almacenado en un disco regrabable puede ser modificado y se le pueden aplicar diferentes efectos, generalmente llamados filtros. Es posible realizar desde sencillas operaciones como sacar una parte de audio, repetir tramos, cambiar el volumen, hasta otras más compleja como destacar sonidos o disimularlos.



Hay una gran cantidad de programas disponibles como el SoundForge o el CoolEdit. Sin embargo el Audacity, un editor de descarga y uso gratuito, muy completo y apto para casi todo tipo de procesamiento del audio.

El Audacity permite editar y mezclar varios canales simultáneamente, en un entorno gráfico, intuitivo y amigable.

Tamaño del sonido digitalizado sin compresión

Si muestreamos 3 minutos de un sonido a 44100 muestras por segundo con una profundidad de 16 bits por canal ocuparemos el siguiente espacio en archivo de computadora:

- 3 minutos x 60 segundos= 180 segundos
- 180 seg x 44100 muestras por seg.= 7.938.000 muestras
- 7.930.000 muestras x 16 bits= 127.008.000 bits
- 127.008.000 x 2 canales= 254.016.000 bits
- 254.016.000 bits / 8 bytes = 31.752.000 bytes
- 31.752.000 bytes = 31 Mb aprox.

Tasa de transferencia o Bitrate

Se trata de la cantidad de bits que se pueden transmitir o procesar durante un tiempo determinado. Habitualmente se mide por segundo. Así, podemos mencionar que existen tasas de transferencia para enlaces de internet, por ejemplo, de 512 kbps (kilobits por segundo).

En el caso de los archivos de audio o video, la tasa de transferencia determina la calidad con la que podrá ser reproducido el archivo.

En el ejemplo anterior, será necesaria una tasa de transferencia de 1411kbps, que surge de:

- 44.100 muestras por segundo x 16 bits= 705.600 bits por segundo
- 705.600 bits por segundo x 2 canales= 1.411.200 bits por segundo
- 1.411.200 bps = 1.411 kbps = 1,41 Mbps

De modo que, por ejemplo, para ser transmitido ese archivo por internet sin interrupciones se requeriría de una conexión de al menos 2 Megas (2Mbps)

Formatos de almacenamiento de audio digital

WAV: almacena todo los bits que se han digitalizado. Es un formato de compresión sin pérdida.

Para este tipo de archivos, el sonido es uno solo. Es decir que aunque provenga de fuentes diversas, y luego el oído humano pueda reconocerlas, para la codificación digital es un único sonido

MIDI: En lugar de almacenar un muestreo del sonido digital, almacena las instrucciones para que una computadora interprete y reproduzca sonido. Se puede equiparar con la noción de partitura. La calidad final ependerá de la capacidad de la máquina para reproducir la complejidad del sonido.

Es pequeño y mucho más editable porque se puede descomponer la interpretación de cada instrumento.

MP3: funciona con la lógica del WAV pero comprime aprovechando los límites de la capacidad de audición humana. Hay pérdida de calidad. Hay diferentes calidades de compresión.

MP3

MP3 es un formato de datos que debe su nombre a un algoritmo de codificación llamado MPEG 1 Layer 3, el cual, a su vez, es un sistema de compresión de audio que permite almacenar sonido con una calidad similar a la de un CD y con un índice de compresión muy elevado, del orden de **1:11**. En la práctica, esto significa que en un CD-Rom se pueden grabar unos 11 CD de audio, es decir, unas 150 canciones aproximadamente.



El sistema de codificación que utiliza MP3 **ES UN ALGORITMO DE PÉRDIDA**. Es decir, el sonido original y el que obtenemos posteriormente no son idénticos. Esto se debe a que MP3 aprovecha **LAS DEFICIENCIAS DEL OÍDO HUMANO** y elimina toda aquella información que no somos capaces de percibir. Se han realizado multitud de estudios de percepción acústica descubriendo que hay una serie de efectos que pueden ayudar a la codificación del sonido con el objetivo de reducir todo lo posible la cantidad de información inútil o redundante.

Los más importantes son:

LOS LÍMITES DE AUDICIÓN. Nuestro oído solo trabaja con frecuencias que van entre los 20 Hz y los 20 Khz aproximadamente, con lo que las frecuencias restantes son descartables.

EFFECTO DE ENMASCARAMIENTO. Es aquel que se produce cuando dos señales de frecuencia similar se superponen. Entonces solo podemos percibir aquella que posee más volumen y, por lo tanto, la de volumen menor es susceptible de ser eliminada.

REDUNDANCIA DE ESTÉREO. Existen redundancias entre los componentes tonales y no tonales del sonido en los dos canales estéreo, y, además, por debajo de una cierta frecuencia el oído humano no es capaz de percibir la direccionalidad del sonido, por lo cual por debajo de estas frecuencias es posible incluso codificar un solo canal, junto con información complementaria para restaurar la sensación espacial para el otro canal.

La digitalización de la señal mediante PCM es la forma más simple de codificación de la señal, y es la que utilizan tanto los CD como los sistemas DAT. Como toda digitalización, añade ruido a la señal, generalmente indeseable. Como hemos visto, **CUANTOS MENOS BITS SE UTILICEN EN EL**

MUESTREO Y LA CUANTIZACIÓN, MAYOR SERÁ EL ERROR al aceptar valores discretos para la señal continua, esto es, mayor será el ruido.

Para evitar que el ruido alcance un nivel excesivo hay que emplear un gran número de bits, de forma que a 44,1 KHz. y utilizando 16 bits para cuantizar la señal, uno de los dos canales de un CD produce más de 700 kilobits por segundo (kbps). Como veremos, gran parte de esta información es innecesaria y ocupa un ancho de banda que podría liberarse, a costa de aumentar la complejidad del sistema decodificador e incurrir en cierta pérdida de calidad. El compromiso entre ancho de banda, complejidad y calidad es el que produce los diferentes estándares del mercado.

Calidad	muestreo	Bits/muestra	Modo	Tasa de transferencia	Frecuencia
Teléfono	8000hz	8	mono	64kbps	200-3400hz
Radio AM	11025hz	8	mono	88kbps	
Radio FM	22050hz	16	Estéreo	705kbps	
CD	44100hz	16	Estéreo	1411kbps	20-20000

Los números de esta tabla hacen referencia a los formatos de audio sin compresión. Cuando aplicamos una compresión como la que genera MP3 la tasa de transferencia o bitrate funciona como parámetro para indicar la calidad de la compresión.

De ese modo, el programa de compresión establece una relación entre la tasa de transferencia que le vamos a indicar y la calidad final del sonido del archivo una vez comprimido.

Los estándares de uso indican que se puede establecer la siguiente relación entre la tasa de transferencia (o Bitrate) y la calidad final para la compresión en MP3:

- **4 Kbps** es el mínimo del estándar
- **8 Kbps** es la calidad telefónica
- **32 Kbps** corresponde a la calidad de radio AM
- **96 Kbps** corresponde a la calidad de radio FM
- **128 Kbps** es la calidad cercana al CD, muy común en MP3
- **192 Kbps** corresponde a la calidad CD en formato MP3
- **320 Kbps** es la máxima calidad para el formato MP3