

Representación de la información en un computador.

## REPRESENTACIÓN DIGITAL DE SONIDOS

1



### Representación de la información en computadores

- L2.0 Sistemas de numeración usuales en informática.
- L2.1 Nociones básicas sobre representación de la información
- L2.2 Representación de textos.
- L2.3 Representación de sonidos.
- L2.4 Representación de imágenes y de video.
- L2.5 Representación de números enteros.
- L2.6 Representación de números reales.
- L2.7 Algoritmos de compresión de datos.

2



## Objetivos

- Las señales de audio hacen referencia a la información contenida en el habla, música y sonidos en general.
- Antes de establecer los objetivos de esta lección les ruego oigan y vean lo que sigue a continuación que corresponde a la primera parte del tercer movimiento:
  - Brahms: Symphony No. 3 / Rattle · Berliner Philharmoniker

3



## Después de escuchar esta bella melodía, nos planteamos analizar y reflexionar sobre:

- La naturaleza de los sonidos.
- Características del sistema de audición humano.
- La forma en qué la información inherente a los sonidos.
  - la transformamos para representarla en forma digital.
  - la almacenamos en forma de fichero.
  - la transmitimos a través de una red de datos.

4



- **Para comprender la naturaleza de los sonidos, le ruego que observe las imágenes que vienen a continuación, donde se incluye:**
  - Imágenes del video vistas anteriormente.
  - Señales que varían de forma continua a los largo del tiempo que es donde se encuentra la **información de acústica**, a la que vamos a dedicar esta lección.
  - Se muestran dos señales porque se ha realizado una grabación con sonido estereofónico (2 canales)
- **Las presentaciones de señales las voy a realizar con la aplicación:**
  - Sony Sound Forge ©

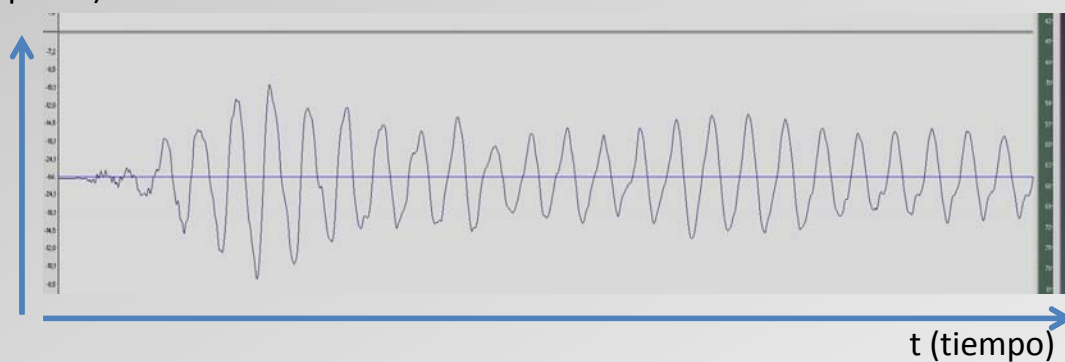
5



**En resumen, el sonido se encuentra en una señal continua,  $s(t)$ , que varia en función del tiempo.**

- **Fragmento de lo escuchado a través de uno de los canales.**

$s(t)$  (amplitud)



6



## Contenidos

- Introducción.
- **Un poco de Teoría de Señales.**
- **Un poco de Física y Fisiología.**
- **Procesos para digitalizar señales de audio.**
- **Archivos de audio.**
- **Transmisión “streaming”.**

7

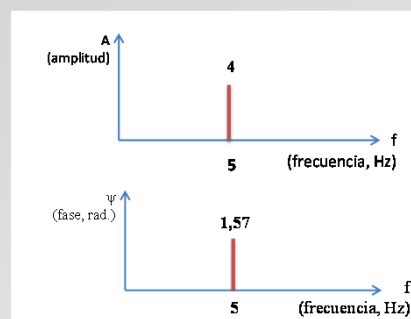
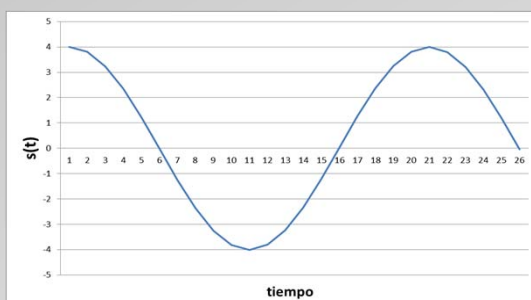


## Un poco de teoría: señal sinusoidal

$$s(t) = A \cdot \text{sen}(2\pi ft + \varphi) = A \cdot \text{sen}\left(\frac{2\pi}{T}t + \varphi\right)$$

- **Parámetros de una función sinusoidal: A, f, (T=1/f),  $\psi$** 
  - Ejemplo: A=4; f=5 Hz, (T=0,2 s),  $\psi = 1,57$  radianes

$$s(t) = 4 \cdot \text{sen}(0,017 \cdot t + 1,57)$$



8



## Un poco de teoría: análisis de Fourier

- Cualquier señal variable en el tiempo puede considerarse como **suma de señales sinusoidales (“armónicos”)** de distintas frecuencias, fases y amplitudes (Análisis de Fourier).

- Joseph Fourier (matemático y físico francés, 1768 a 1830).

- Si la señal es **periódica**, las frecuencias de los armónicos son discretas :  $nf_0$

$$s(t) = a_0 + \sum_{n=1}^{\infty} [a_n \text{sen}(2\pi n f_0 t + \varphi_n)] \quad a_n = F[s(t), n f_0] ; \varphi_n = G[s(t), n f_0]$$

- Si la señal es **aperiódica** dichas frecuencias son continuas

$$s(t) = \int_{-\infty}^{\infty} A(f) \text{sen}[2\pi f t + \varphi(f)] df \quad A(f) = F[s(t), f] ; \varphi(f) = G[s(t), f]$$

9



## Dominio del tiempo y dominio de la frecuencia

- Se denomina **espectro de frecuencias a la representación de:**

- La variación de amplitudes en función de la frecuencia:  $A(f) = F[s(t), f]$

- La variación de la fase en función de la frecuencia:  $\varphi(f) = G[s(t), f]$

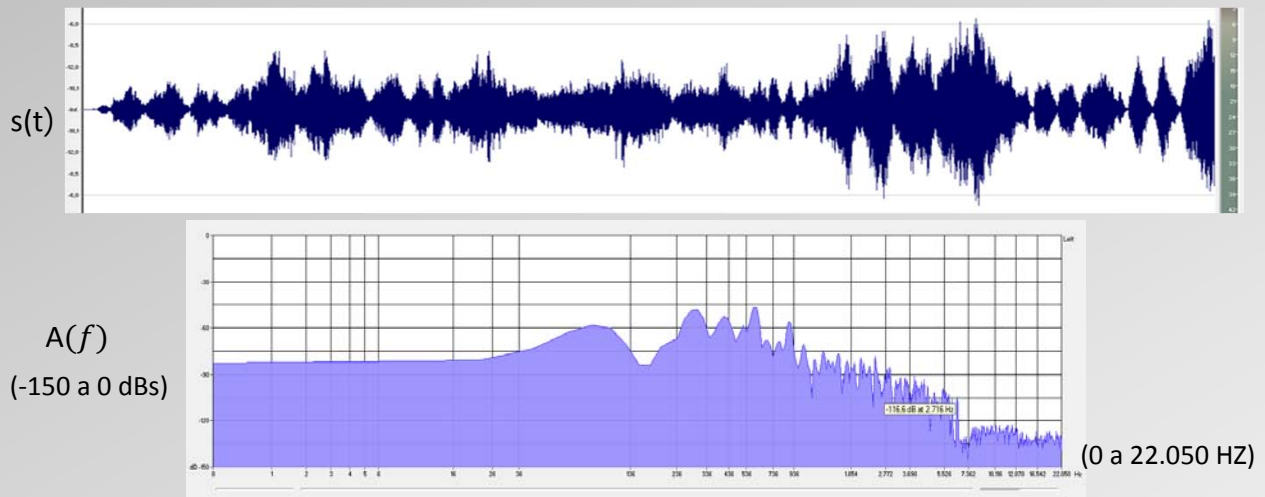
- La misma información hay en la señal en función del tiempo (**dominio del tiempo**), que en las representaciones en función de la frecuencia (**dominio de la frecuencia**).

- Es fácil pasar del dominio del tiempo al dominio de la frecuencia.

10



**Dominio del tiempo y de la frecuencia de:  
Brahms: Symphony No. 3 / Rattle · Berliner Philharmoniker**



11



## Concepto de ancho de banda

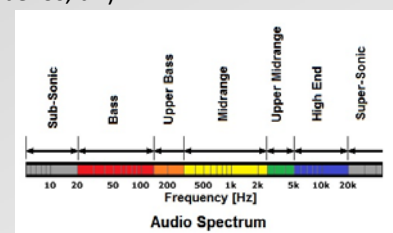
- **Ancho de banda de una señal:**
  - Es el rango de frecuencias que ocupa:  $B = f_h - f_l$ 
    - $f_h$  y  $f_l$  son las frecuencias de corte (frecuencias en las que cae la amplitud,  $A$ , a un 70,7% de sus valores centrales,  $A_0$ )
  - Algunos anchos de banda
    - Voz humana y música: 0,1 a 22 KHz
    - Voz a través de teléfono: 0,2 a 3,4 KHz
    - Señal de video (NTCS TV): 0 a 4,6 MHz
    - Señal de video (HDTV/USA): 24,9 MHz
    - Señal digital con  $r$  símbolos por segundo,  $B=r/2$
- **Ancho de banda de transmisión**
  - Rango de frecuencias que un medio puede transmitir o procesar sin atenuar. Es una medida de la velocidad que admite el medio, ya que cuanto mayor es el ancho de banda más rápidamente con el tiempo pueden cambiar las señal que transitan por él.
  - Para evitar distorsiones, siempre el ancho de banda de transmisión debe de ser igual o mayor que el ancho de banda de la señal o conjunto de señales que transitan por él.

12



## Un poco de física y de fisiología

- Una señal de audio (voz, música o cualquier sonido) es una **onda de presión** que se transmite a través de un **medio**.
  - La **intensidad sonora** ( $I$ ) es la potencia acústica transmitida por unidad de área ( $\text{W}/\text{m}^2$ ).
- Uno de los sentidos más sensibles de los humanos es el oído:
  - Umbral de audición:  $I_0 = 10^{-12} \text{ W}/\text{m}^2$
  - Sensación de dolor:  $1 \text{ W}/\text{m}^2$
  - Sensibilidad logarítmica: se utiliza una escala logarítmica (decibelios, dB):
    - $I_{dB} = 10 \cdot \log_{10} \frac{I}{I_0}$
  - Rango de frecuencias audibles: 20 Hz a 20 KHz



13



## Procesos para digitalización de una señal de audio

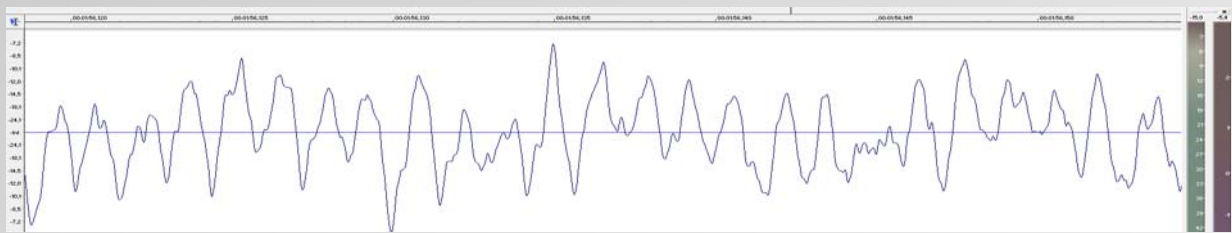
- **Captación**
- **Adaptación:**
  - La señal captada se amplifica o atenúa para encajarla dentro dos valores límites, p. e. entre  $-5$  voltios y  $+5$  voltios.
- **Muestreo.**
- **Digitalización.**
- **Codificación/compresión.**

14



## Proceso de captación

- La **captación** se realiza mediante de un micrófono u otro transductor generándose, a partir de la señal de presión una **señal analógica** (eléctrica).
- Una señal analógica puede tomar cualquier valor dentro de un determinado intervalo continuo.
  - El video de demostración:
    - Tiene una duración total de 00:02:53,949.
    - A continuación se muestra un fragmento de unos 20 ms (00:01:56,120 a 00:01:56,140):

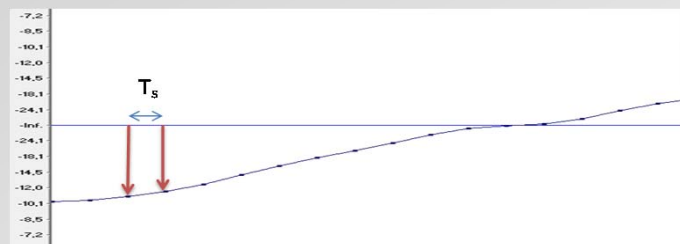


15



## Proceso de muestreo

- En principio tendríamos que almacenar la amplitud de la señal de todos los instantes de tiempo comprendidos entre 0 y 2' 53,949".
- Como hay infinitos instantes (puntos), resulta imposible.
- Solución → muestreo
  - Se toman muestras de la señal regularmente con un periodo de muestreo  $T_s$  (p.e. 22,7 $\mu$ s).
    - La frecuencia de muestreo será:  $F_s = \frac{1}{T_s}$  (p.e. 44.100 muestras por segundo)



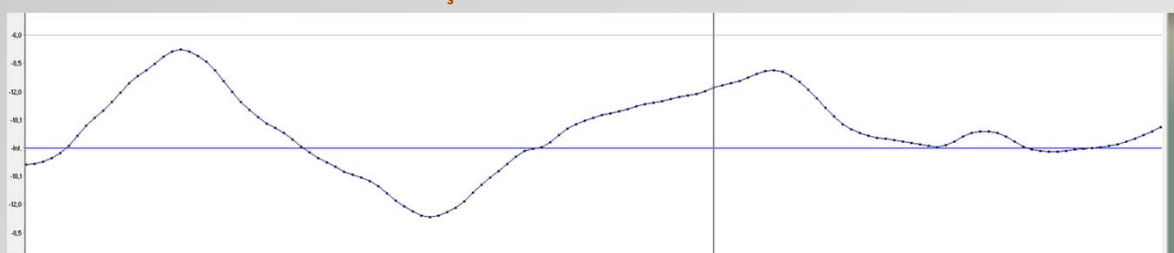
16





- **La frecuencia de muestreo debe ser lo suficientemente alta como para que se pueda reconstruir (“interpolar”) la señal original a partir de las muestras.**
  - Pero, cuanto mayor sea la frecuencia de muestreo, más muestras habrá que almacenar y más ocupará o tardará en transmitirse nuestro fichero.
    - Hay un compromiso entre calidad y capacidad ocupada por el fichero.
    - Se demuestra (**Teorema de muestreo de Nyquist-Shannon**) que para recuperar con máxima calidad la señal original es suficiente con que la frecuencia de muestreo sea el doble del ancho de banda de la señal original:

- $F_s \geq 2 \cdot B$



17



## Proceso de digitalización o conversión analógica a digital

- **El valor de la amplitud de cada muestra se convierte a binario.**
  - El nº de bits por muestra,  $n_s$ , nos indica la precisión de la conversión de analógico a digital:
    - 8 bits: 256 niveles de amplitud
    - 16 bits: 65.536 niveles.
  - Cuantos más bits por muestra, mayor calidad pero mayor ocupación del archivo de datos

18



## Codificación de las muestras

- Una vez obtenidos los valores binarios de las muestras hay que codificarlos de acuerdo con un determinado formato.
- Con frecuencia, después de la codificación, o simultáneamente a ella, se realiza un proceso de compresión de la información.
- Una vez que se produce la información digital acústica, ésta:
  - O bien se graba en forma de **archivo**
  - O bien se **transmite** por una red de datos
- En el momento de reproducir audio digitalizado, hay que hacer el proceso inverso:
  - Decomprimir y decodificar la señal
  - Reproducir la a través de un altavoz u otro medio, al mismo ritmo en que se generó.
- Hay circuitos o programas que se denominan **CODEC** que realizan ambas funciones:
  - CODifica/dEcodifiCA
  - COMprime/DEComprime

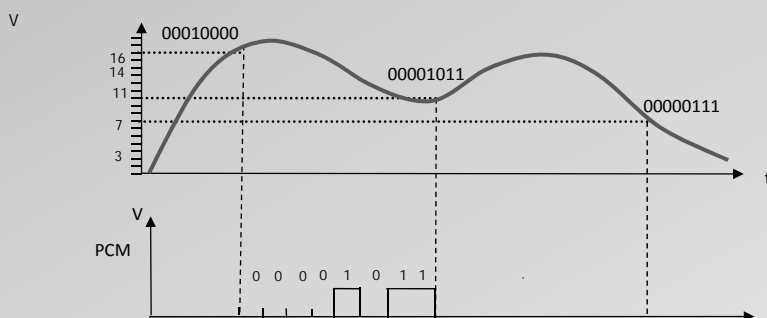


19



La mayoría de los formatos de codificación están basados en PCM (*Pulse Code Modulation*):

- En el intervalo entre la muestra  $n$  y la  $n+1$  (tiempo  $T_s$ ) se genera un tren de impulsos que corresponde al valor binario de la muestra  $n$ .



$f_s$ : frecuencia de muestreo  
 $n_s$ : nº de bits por muestra  
 $c$ : nº de canales

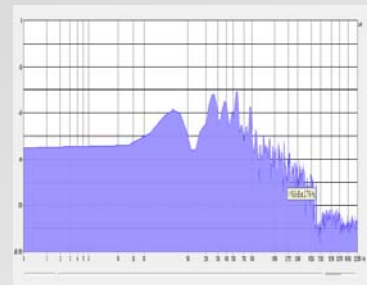
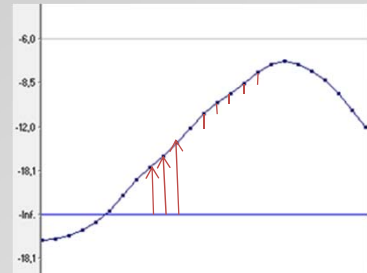
$$\text{Caudal o tasa de bits (bit rate): } R_{bps} = f_s \frac{\text{muestras}}{\text{segundo}} \cdot n_s \frac{\text{bits}}{\text{muestra}} \cdot c$$

20



## Tipos de CODEC usuales:

- **PCM (Pulse Code Modification):**
  - se transmiten o graba un tren de pulsos correspondientes a las amplitudes de cada muestra (o a sus logaritmos).
- **DPCM (Differential Pulse Code Modulation):**
  - Se almacena la diferencia con la muestra anterior (4 bits)
- **ADPCM (Adaptive Differential Pulse Modulation):**
  - Se predice el nuevo valor, y se almacena el error entre el valor predicho y el real (las diferencias necesitan menos bits)
- **$\mu$ -law, similar a ADPCM**
- **MPEG Audio Capa-III: para formatos MP2, MP3 y AAC**
  - varía el número de bits y la frecuencia en función del rango de frecuencias: se tienen en cuenta las características del oído humano
- **Particulares:**
  - Microsoft (Windows Media Audio), de Real Networks, de Apple, etc.



21



## Ejemplos de calidades de audio

Aplicación	Frecuencia muestreo $f_s$ (KHz)	Nº bits/muestra	Nº de canales	Tasa de bits $T_b$ (Kbps)	Capacidad en 1 minuto
Telefonía	8	8	1	60 Kbps	480 KB
Radio AM	11	8	1	88 Kbps	644 KB
Radio FM	22,05	8	2	705.6 Kbps	5,17 MB
CD	44,1	16	2	1,38 Mbps	10,34 MB
Sonido DVD y en TDT	48	16	2	1,42 Mbps	11,25 MB

Características del ejemplo utilizado (3ª Sinf. de Brahms)

$$R_{bps} = f_s \frac{\text{muestras}}{\text{segundo}} \cdot n_s \frac{\text{bits}}{\text{muestra}} \cdot c$$

44.100 Hz	16 bit	Stereo	00:02:53,949	39.912,4 MB
-----------	--------	--------	--------------	-------------

22



## Archivos de audio

- **Para almacenar de una señal de sonido:**

– Los valores obtenidos en la conversión (**muestras codificadas**) se almacenan en posiciones consecutivas

– Antes de las muestras se incluye una **cabecera**, con información sobre el tipo de fichero y sus parámetros (**metadatos**.)

- **Capacidad de almacenamiento necesaria para los datos en función del tiempo:**

$$C(t) = \frac{f_s \cdot n_s \cdot c \cdot t}{8} \quad (\text{Bytes})$$

- **Son muy importantes las técnicas de compresión de datos**

– ¡1 minuto en calidad CD necesita  $C(60) = 10$  MB!

23



## Formatos de sonidos

Formato	Exten.	CODEC	Comp.	Observ.
AIFF	.aif, .aiff	PCM <i>u otros</i>	no	MAC,
AU	.au	u-law <i>u otros</i>	no	SUN/Next
WAV	.wav	PCM <i>u otros</i>	no	
Audio-CD (CDDA)	--	PCM	no	
FLAC	.flac		$C \approx 58\%$	Sin pérdidas
Windows Media Audio (WMAL)	.wma		$C \approx 58\%$	Sin pérdidas
MP3 (MEPEG-1 Audio Layer 3)	.mp3	MPEG Audio-III	$C \approx 9\%$	
Windows Media Audio	.wma	<i>Propietario</i>	$C \approx 9\%$	Microsoft
AAC (Advanced Audio Coding)	.m4a		$C \approx 8\%$	
Vorbis	.ogg		$C \approx 8\%$	

24



## Transmisión de audio

- Los archivos de audio se pueden transmitir como cualquier otro archivo
- Si se desea oír en “tiempo real” se hace una transmisión en forma de ráfagas (“**streaming**”)
  - Es necesario que el ancho de banda del canal sea suficiente para transmitir el caudal de datos necesario ( $\geq 2 \cdot R_s$ )

25



## Reproducción

- Al irse recibiendo la información se va almacenando temporalmente en el sistema de reproducción (**buffer**) .
- Para oír correctamente una información digitalizada de audio:
  1. Debe recibirse por el canal de datos la información “a tiempo” para oírla correctamente (al mismo ritmo en que se produjo)
  2. El equipo o programa de “**reproducción**” debe reconstrucción las señales originales a una velocidad adecuada.
    - Eliminación de cabeceras
    - Descompresión, etc.

26



## Recepción *streaming* de Youtube



Instante que se esta visualizando

Instante hasta el que ya ha llegado información

La barra gris siempre debe ir por delante de la roja. Si no es así es debido a que el canal no tiene ancho de banda suficiente o el servidor que suministra la información es lento

<http://www.youtube.com/watch?v=5-HcTI8zuBo>

27



## Resumen y conclusiones

- **Introducción.**
  - Brahms: Symphony No. 3 / Rattle · Berliner Philharmoniker  
<http://www.youtube.com/watch?v=5-HcTI8zuBo>
- **Un poco de Teoría de Señales.**
- **Un poco de Física y Fisiología.**
- **Procesos para digitalizar señales de audio:**
  - Captación de la señal.
  - Muestreo.
  - Digitalización.
  - Codificación.
- **Archivos de audio.**
- **Transmisión “streaming”.**

28



