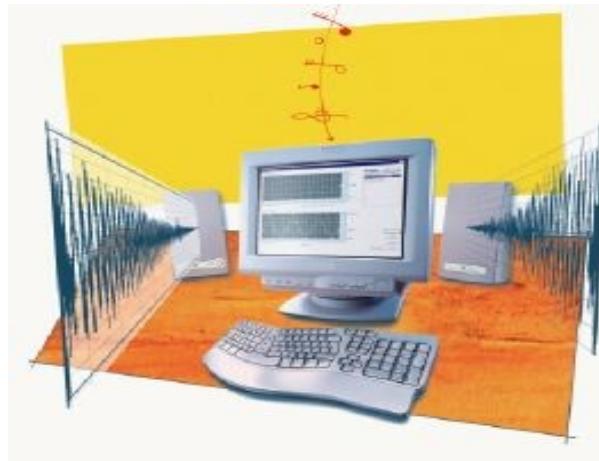
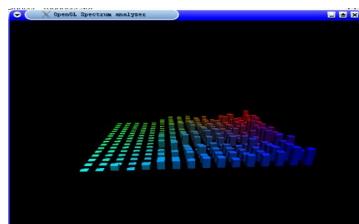
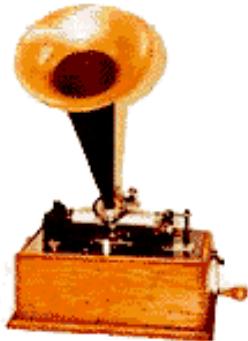


Tema 3

Audio

101010
101010



Sistemas Multimedia Interactivos e Inmersivos

Grado de Ingeniero en Informática

Escola Tècnica Superior de Enginyeria Informàtica

Curso 2021/2022

Manuel Agustí

Objetivos

- Relacionar los fundamentos físicos del sonido y de la percepción humana del audio
- Aplicar los conocimientos sobre digitalización de señales al caso del sonido
- Conocer y comprender los fundamentos de la representación de audio en digital
- Caso de estudio
 - ★ Audio 3D: Seminario de OpenAL
- Casos de estudio relacionados con la voz
 - ★ Síntesis de voz: Seminario de síntesis de voz con Festival
 - ★ Reconocimiento de voz
- Introducir el uso del SDK de cada uno de los casos de estudio en el desarrollo de aplicaciones

Índice

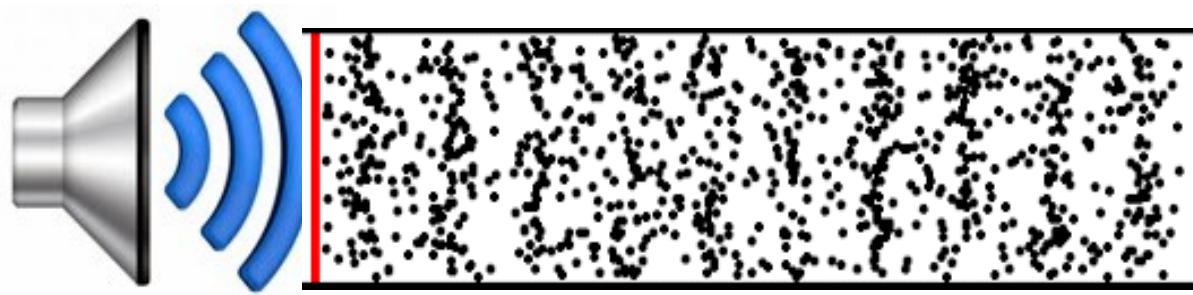
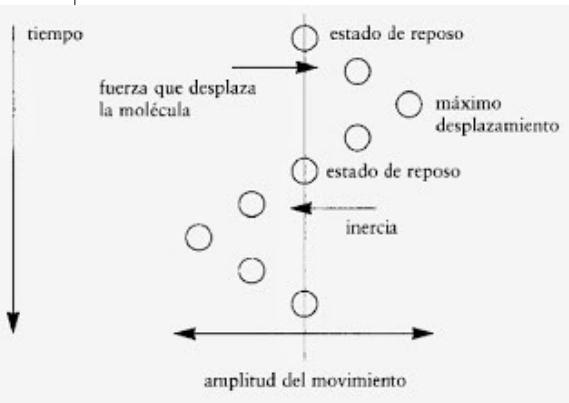
- Introducción
 - Definición y ejemplos
 - Adquisición y representación digital de la información
- Sonido y acústica
 - Dominios de representación
 - Modelado y percepción
 - Introducción al procesado de audio
- Formatos de audio
 - MIDI
 - Forma de ondas
 - Compresión: MP3
- Herramientas software en aplicaciones de escritorio
 - Audio 3D
 - Síntesis y reconocimiento de voz

Índice

- Introducción
 - Definición y ejemplos
 - Adquisición y representación digital de la información
- Sonido y acústica
- Formatos de audio
- Herramientas SW en aplicaciones de escritorio

Naturaleza del sonido

- Def.: Onda de presión.
 - Vibración de las moléculas en un medio (aire).



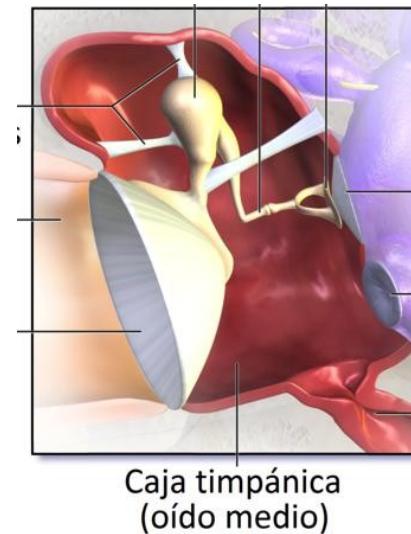
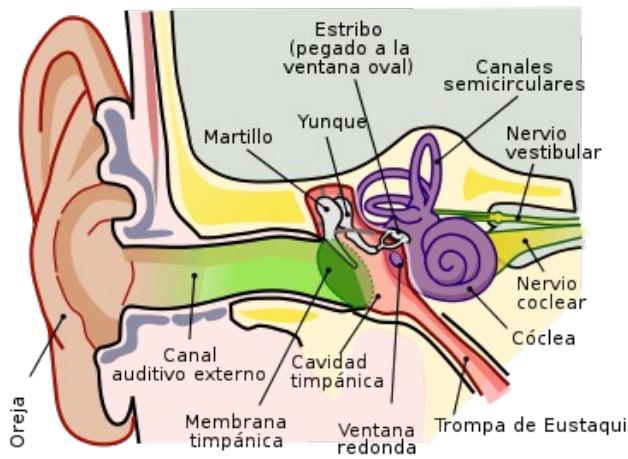
ESTADO	MEDIO	VELOCIDAD DEL SONIDO (m/s)
Gaseoso	Aire (20°C)	340
	Hidrógeno (0°C)	1.286
	Oxígeno (0°C)	317
	Helio (0°C)	972
Líquido	Agua (25°C)	1.493
	Agua de mar (25°C)	1.533
Sólido	Aluminio	5.100
	Cobre	3.560
	Hierro	5.130
	Plomo	1.322
	Caucho	54
Vacio	Vacio	0

Fuente: ¿En qué medio se propaga mejor el sonido?

<<http://sonenacustica.blogspot.com.es/2013/11/en-que-medio-se-propaga-mejor-el-sonido.html>>.

Naturaleza del sonido (II)

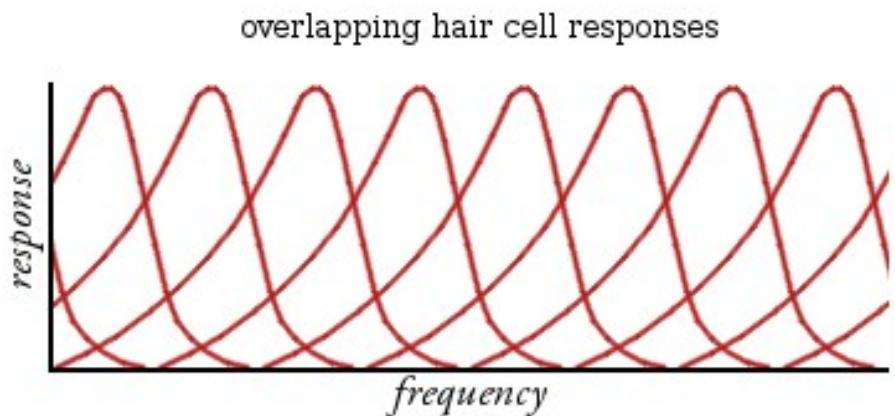
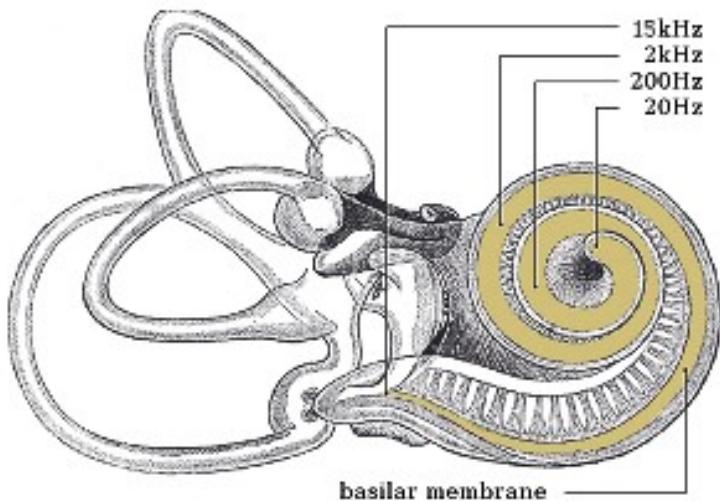
- Onda de presión → vibración de las moléculas en un medio
 - Produce una sensación “auditiva” al perturbar el estado de reposo del oído.



- Intensidad: potencia recibida por unidad de área (W/m^2)
 - En intensidades de sonido → mediciones relativas al mínimo nivel de intensidad audible (10^{-12} W/m^2)
 - P.ej. una fábrica: $10^{-4} / 10^{-12} = 10^8$

Naturaleza del sonido (III)

- Sistema auditivo: respuesta en frecuencia
 - *A sound is inaudible if there are no hair cells tuned to hear it.*

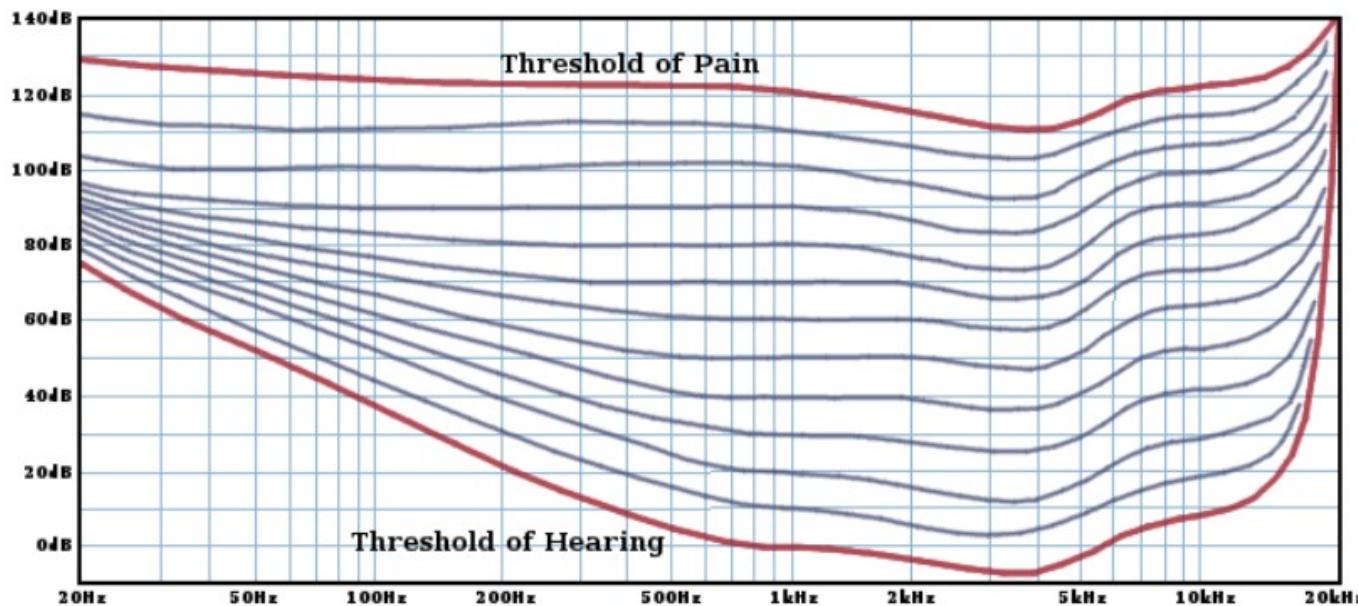


Above left: anatomical cutaway drawing of a human cochlea with the basilar membrane colored in beige. The membrane is tuned to resonate at different frequencies along its length, with higher frequencies near the base and lower frequencies at the apex. Approximate locations of several frequencies are marked.

Above right: schematic diagram representing hair cell response along the basilar membrane as a bank of overlapping filters.

Naturaleza del sonido (IV)

- Sistema auditivo: respuesta en frecuencia
 - *The human hearing range spans 20Hz to 20kHz ... is a generous range. It thoroughly covers the audible spectrum, an assertion backed by nearly a century of experimental data.*



Above: Approximate equal loudness curves derived from Fletcher and Munson (1933) plus modern sources for frequencies > 16kHz. The absolute threshold of hearing and threshold of pain curves are marked in red. Subsequent researchers refined these readings, culminating in the Phon scale and the ISO 226 standard equal loudness curves. Modern data indicates that the ear is significantly less sensitive to low frequencies than Fletcher and Munson's results.

Naturaleza del sonido (V)

- Velocidad de propagación del sonido

- Es la dinámica de propagación de las ondas sonoras.
- En la atmósfera terrestre es de 343,2 m/s o 1234,8 km/h (a 20 °C de temperatura, con 50 % de humedad y a nivel del mar).
- La velocidad del sonido varía en función del medio en el que se trasmite.
- Dado que la velocidad del sonido varía según el medio, se utiliza el número Mach 1 para indicarla.

Velocidad de propagación del sonido en distintos medios

Medio	Velocidad (m/s)
Caucho	60
Aire (14 ° C)	340
Vapor de agua	500
Agua de mar	1.450
Cemento	4.000
Cobre	5.000
Vidrio	5.700
Acero	6.000



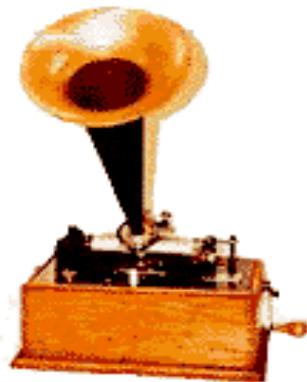
Definiciones (y 2)

- **Bel (B):** se define como el logaritmo decimal de la intensidad de sonido relativa
- **Decibel (dB):** es la unidad más comúnmente empleada en audio
 - ★ $1B = 10dB$
- **Rango dinámico** de un sistema de audio: es la diferencia en dB entre el nivel máximo de sonido que se procesa sin distorsión y el nivel mínimo de sonido que puede ser detectado al final del proceso
 - ★ Es una de las características básicas en un sistema de grabación y reproducción

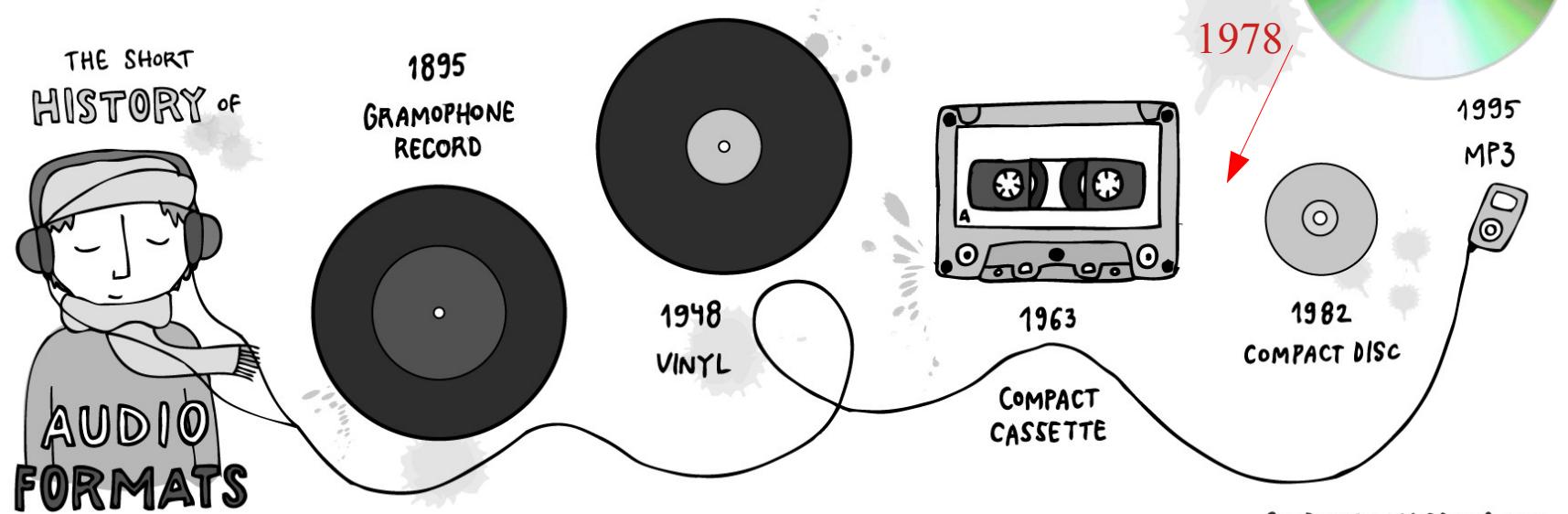
Ejemplos de sonidos y su correspondiente cuantificación en dB

<i>Fuente de sonido</i>	<i>Intensidad (W/m²)</i>	<i>Intensidad relativa</i>	<i>Reacción del que escucha</i>	<i>Intensidad en decibeles (dB)</i>
	10^4	10^{16}		160
Turbina de jet a 10m	10^3	10^{15}	Daño inmediato	150
	10^2	10^{14}	Dolor de oído	140
	10	10^{13}		130
	1	10^{12}	Desagrado	120
Trueno	10^{-1}	10^{11}		110
Cataratas del Niagara	10^{-2}	10^{10}		100
	10^{-3}	10^9		90
Fábrica	10^{-4}	10^8		80
Tráfico de ciudad a 15m.	10^{-5}	10^7		70
Conversación normal a 1m.	10^{-6}	10^6		60
Residencia suburbana	10^{-7}	10^5		50
Biblioteca	10^{-8}	10^4		40
	10^{-9}	10^3		30
Estudio de grabación	10^{-10}	10^2		20
Respiración	10^{-11}	10		10
	10^{-12}	1	Límite audible	0

Adq. y repr. digital de la información



¿Analógico vs Digital?



A. y R. digital de la información (II)

- **Muestreo**
 - * Discretización temporal

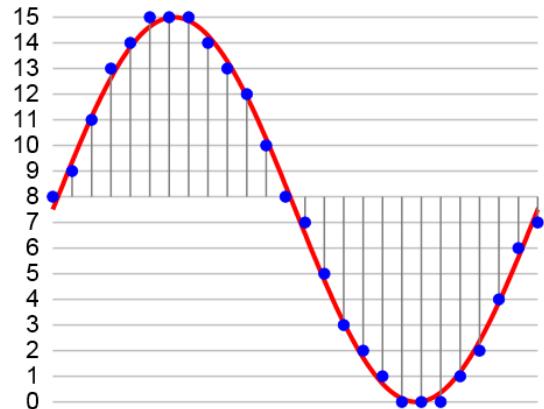


Frecuencia de muestreo

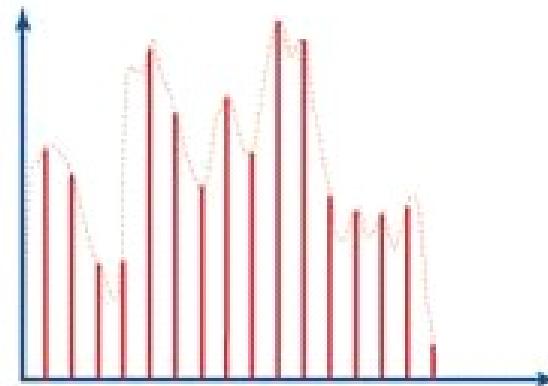
- **Cuantificación**
 - * Discretización valor (amplitud) señal



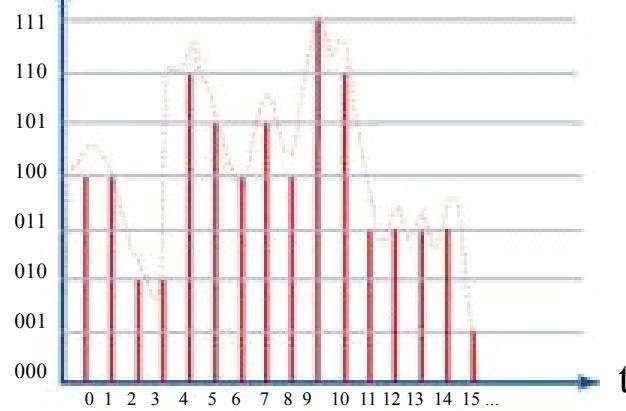
Número de bits



A. y R. digital de la información (III)



A digital



t 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

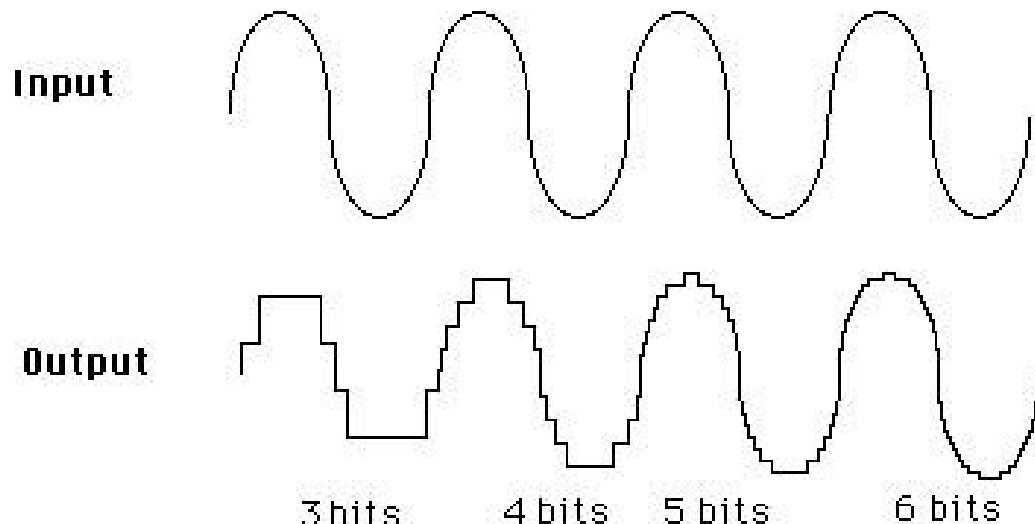
muestras

100	100	010	010	110	101	100	101	100	100	111	111	110	011	011	011	011	001		
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--	--

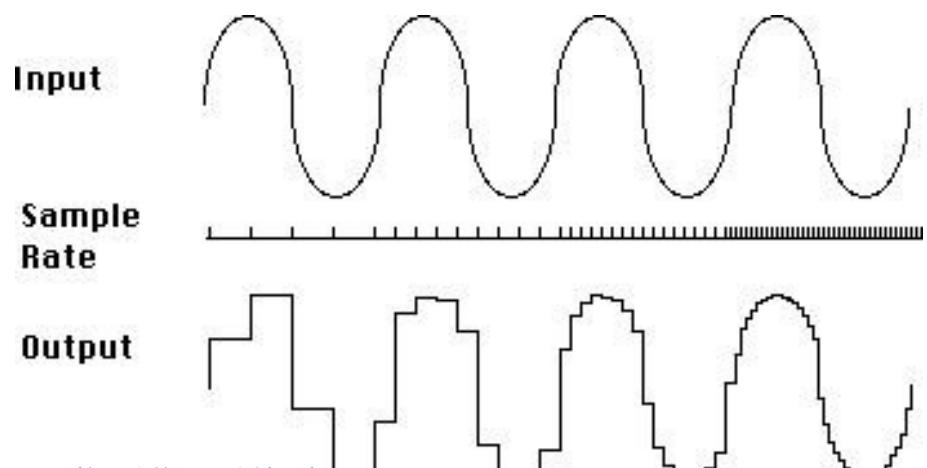
A. y R. digital de la información (IV)

- **Efecto de los parámetros de la digitalización**

- ★ Cuantización



- ★ Frec. de muestreo



A. y R. digital de la información (y V)

La frecuencia de muestreo debe ser superior al doble de la componente frecuencial más alta de la señal a muestrear (Nyquist Theorem)

<i>Formato</i>	<i>Frec. muestreo</i>	<i>Ancho de banda</i>	<i>Rango de frecuencias</i>
Telefonía	8 kHz	3.2 kHz	200-3400 Hz
Teleconferencia	16 kHz	7 kHz	50-7000 Hz
Compact Disc	44.1 kHz	20 kHz	20-20,000 Hz
Digital Audio Tape	48 kHz	20 kHz	20-20,000 Hz

Digitalización de 60 seg. de audio

<i>Velocidad de muestreo</i>	<i>Resolución en bits</i>	<i>Canales Mono ó Estéreo</i>	<i>Tamaño del archivo resultante (Bytes)</i>	<i>Características y posibilidades de la grabación</i>
11 KHz	8	1	660 KB	Esta es la capacidad más baja que se puede emplear en la práctica para obtener resultados útiles. Sonido poco claro y apagado.
11 KHz	8	2	1.3 MB	No es frecuente utilizar 2 canales a esta velocidad de muestreo.
11 KHz	16	1	1.3 MB	Sonido monoaural con calidad radiofónica.
11 KHz	16	2	2.6 MB	El estéreo se nota y mejora ligeramente.
22.05 KHz	8	1	1.3 MB	Sonido monoaural con calidad equiparable a la de los televisores clásicos con sonido analógico. Buena elección para grabaciones de diálogos.
22.05 KHz	8	2	2.6 MB	Elección popular para grabaciones estéreo con calidad aceptable.
22.05 KHz	16	1	2.6 MB	No es una mala elección para voz, pero si reducimos a 8 bits, ahorraremos tamaño y sacrificaremos un poquito de fidelidad.
22.05 KHz	16	2	5.25 MB	Sonido de calidad comparable a la de una buena cinta de sonido.
44.1 KHz	8	1	2.6 MB	Es una elección adecuada para grabaciones monoaurales como diálogos y narraciones con buena calidad.
44.1 KHz	8	2	5.25 MB	La mejor calidad que se podía conseguir en las primeras tarjetas de sonido.
44.1 KHz	16	1	5.25 MB	La elección ideal para narraciones.
44.1 KHz	16	2	10.5 MB	Calidad de grabación de un CD-Audio.

Concepto de *bitrate*

- Tasa de transferencia de bits
- Cantidad de información digital / tiempo
 - * Generalmente se expresa en kbits/segundo
kbits/s, kb/s o kbps.
- Se suelen utilizar múltiplos decimales (SI)
 - * $1000 \text{ bit/s} = 1 \text{ kbit/s}$ (un kilobit o mil bits por segundo)
 - * $1000 \text{ kbit/s} = 1 \text{ Mbit/s}$ (un megabit o un millón de bits por segundo)
 - * $1000 \text{ Mbit/s} = 1 \text{ Gbit/s}$ (un gigabit o mil millones de bits por segundo)
- Suele ser un indicador de la calidad de la información MM

Tarea

¿Cómo calcular el *bitrate*?

¿Cómo calcular el tamaño (teórico) de un fichero de audio?

Índice

- Introducción
- Sonido y acústica
 - Dominios de representación
 - Modelado y percepción
 - Introducción al procesado de audio
- Formatos de audio
- Herramientas software en aplicaciones de escritorio

Sonido y Acústica

Sonido

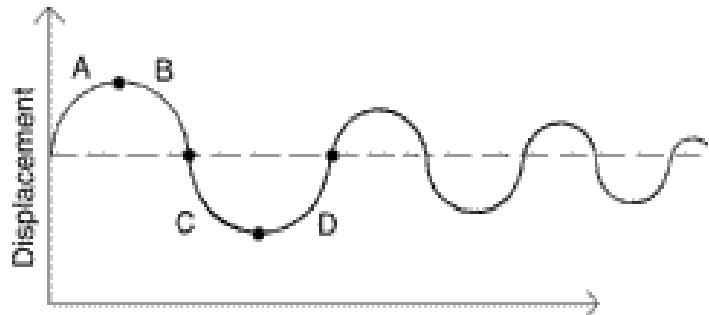
- Impresión fisiológica producida en el oído por las vibraciones elásticas de los cuerpos

Audio

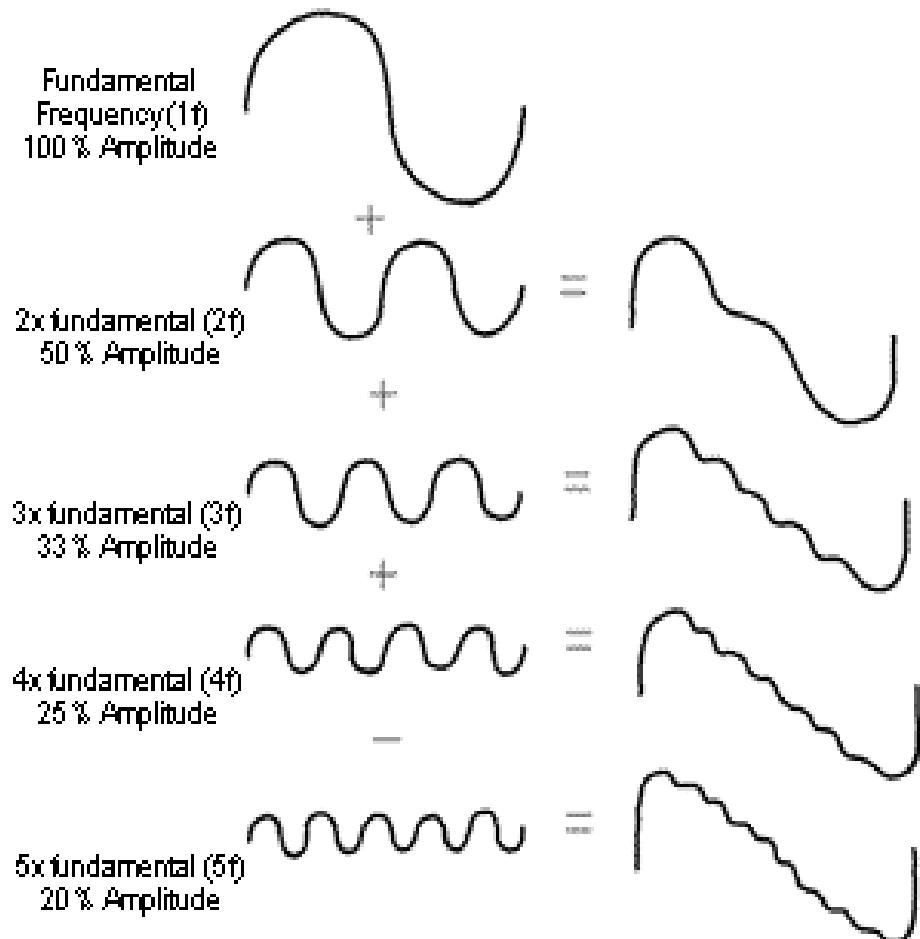
- Adjetivo: perceptible por el oído humano
- Sustantivo: parte de un equipo relativa al sonido

Sonido

- Vibración
 - ★ de un punto

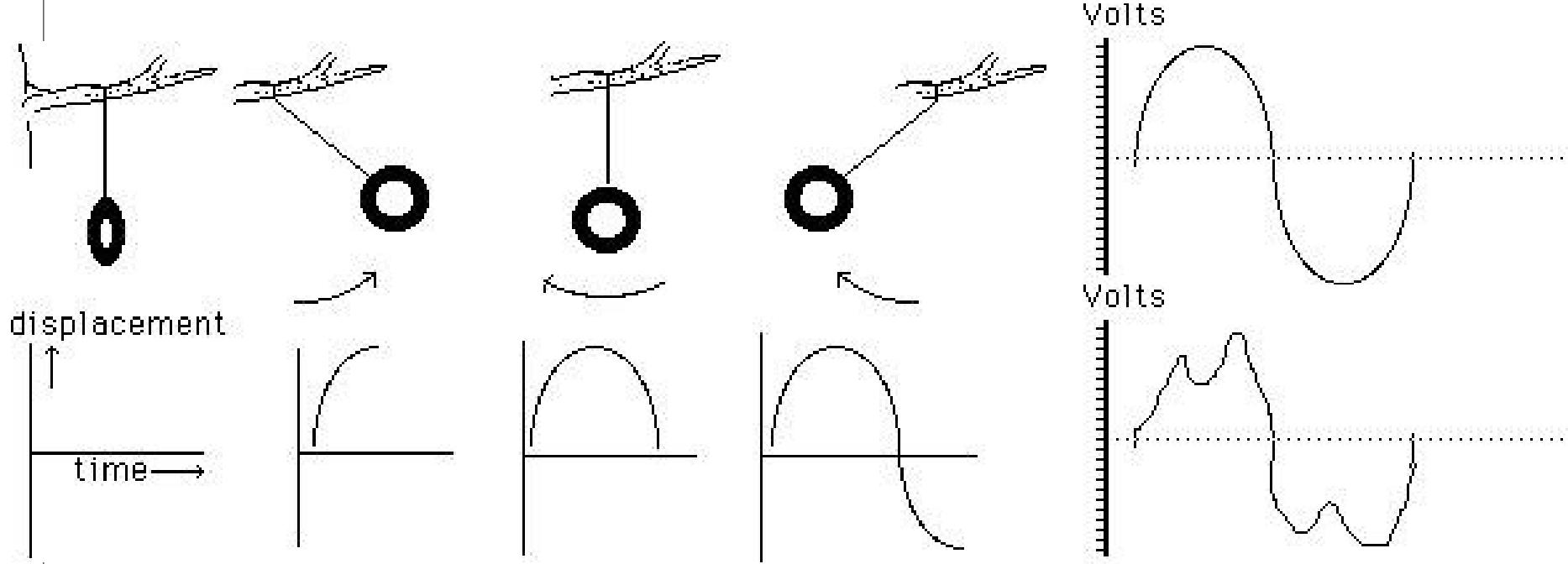


- ★ de un cuerpo:
 - Frecuencia Fundamental
 - Armónicos



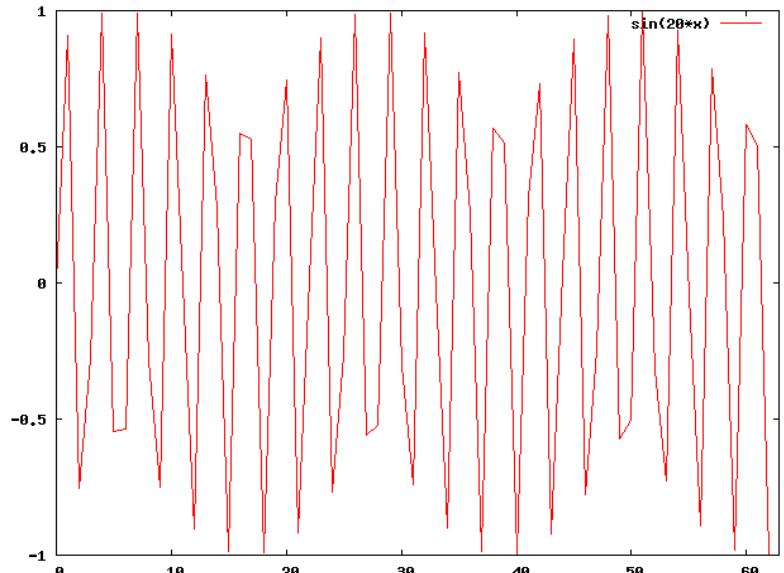
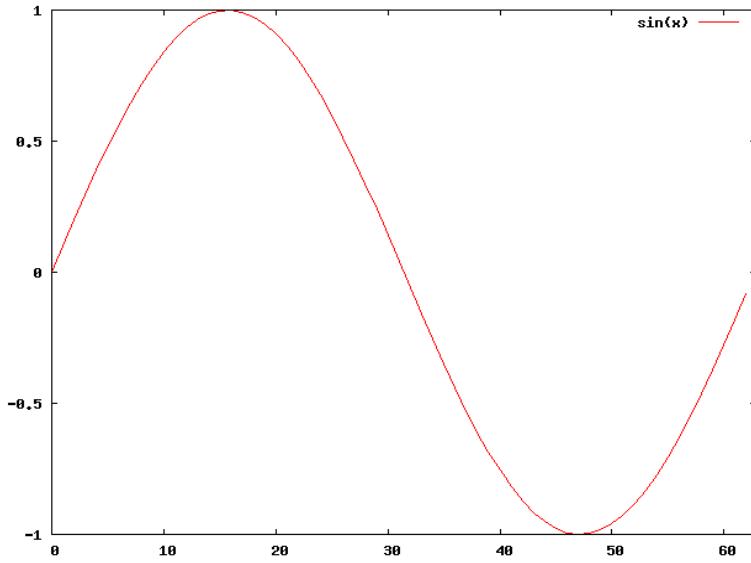
Características del sonido

- Características de las señales sonoras:
 - ★ **Sonoridad (loudness)**: es la intensidad subjetiva, depende principalmente de las **amplitudes** de las comp. del sonido.
 - ★ **Altura (pitch)**: **frecuencia percibida**
 - ★ **Timbre** : freq. fundamental y armónicos (calidad del sonido). Pueden tener la misma amplitud y freq. fdtal. pero suenan ◊



Sonido

- Sonido
 - ★ Intensidad + timbre + frecuencia
 - Graves vs agudos.



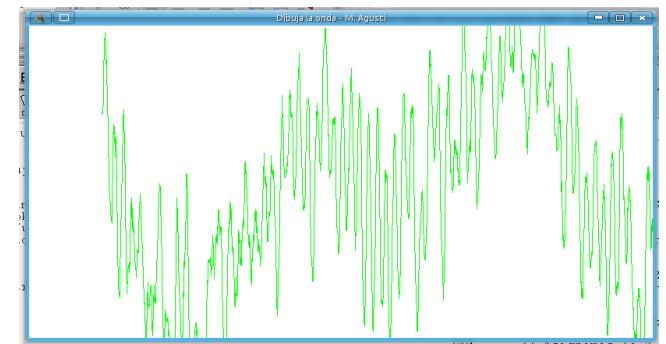
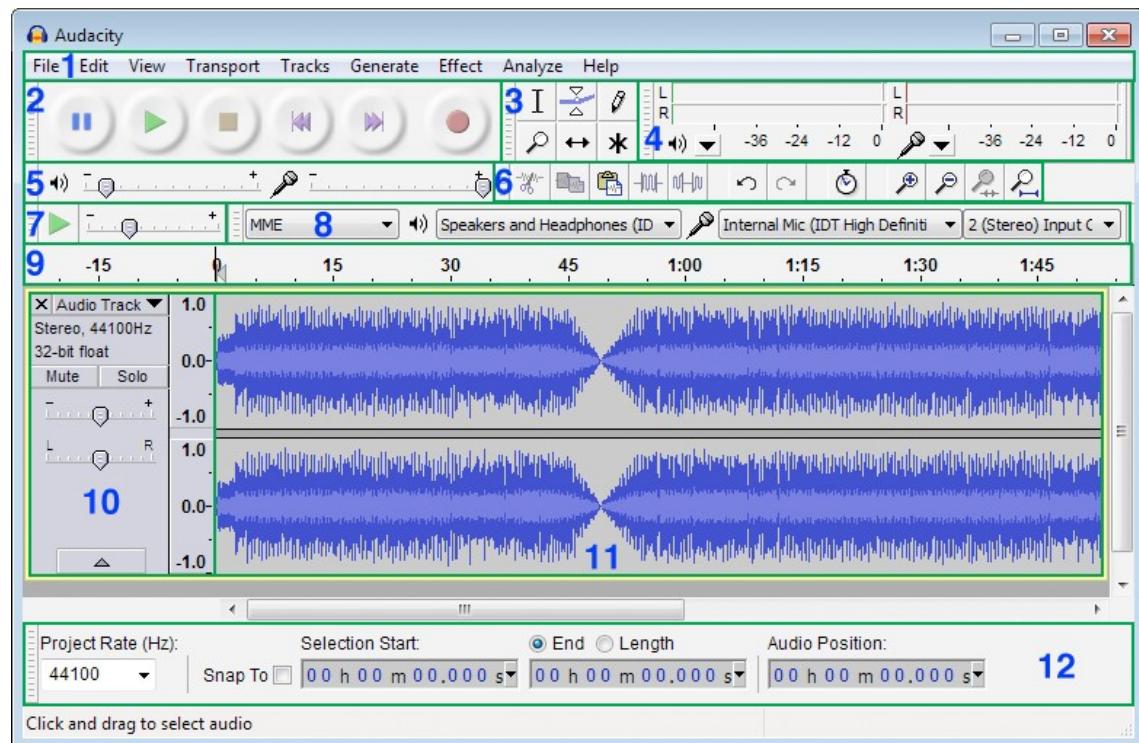
- 20 Hz .. 20k Khz:: [infrasonidos .. ultrasonidos].
- Dominio del tiempo vs dominio de la frecuencia

Dominio de representación

- Del tiempo

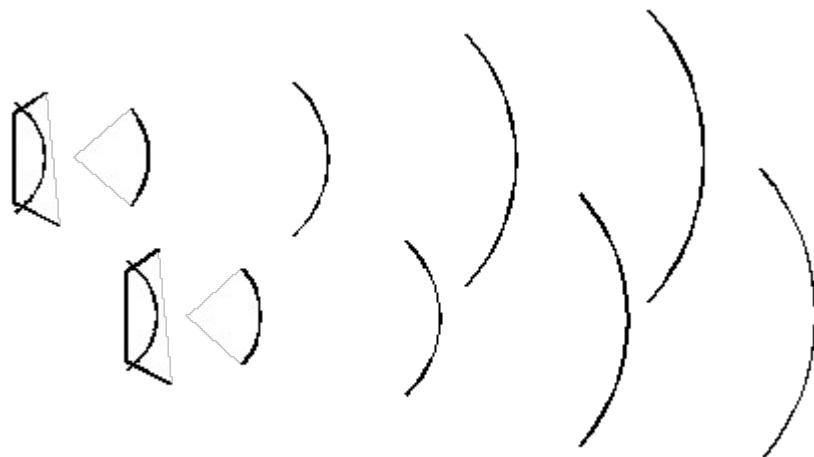
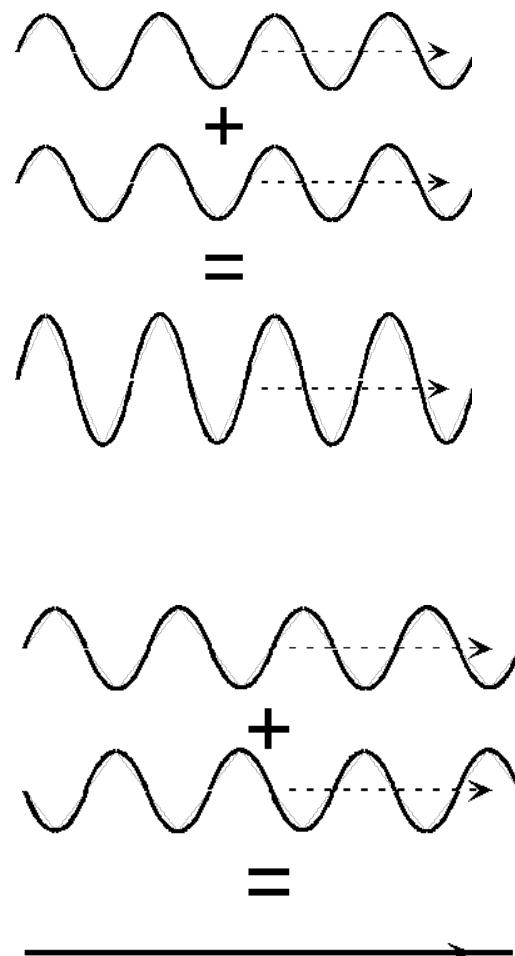
- Modelado del sonido en el tiempo

- ★ Efecto de eco
- ★ Cancelación de eco
- ★ ...
- ★ Y visualizarlo



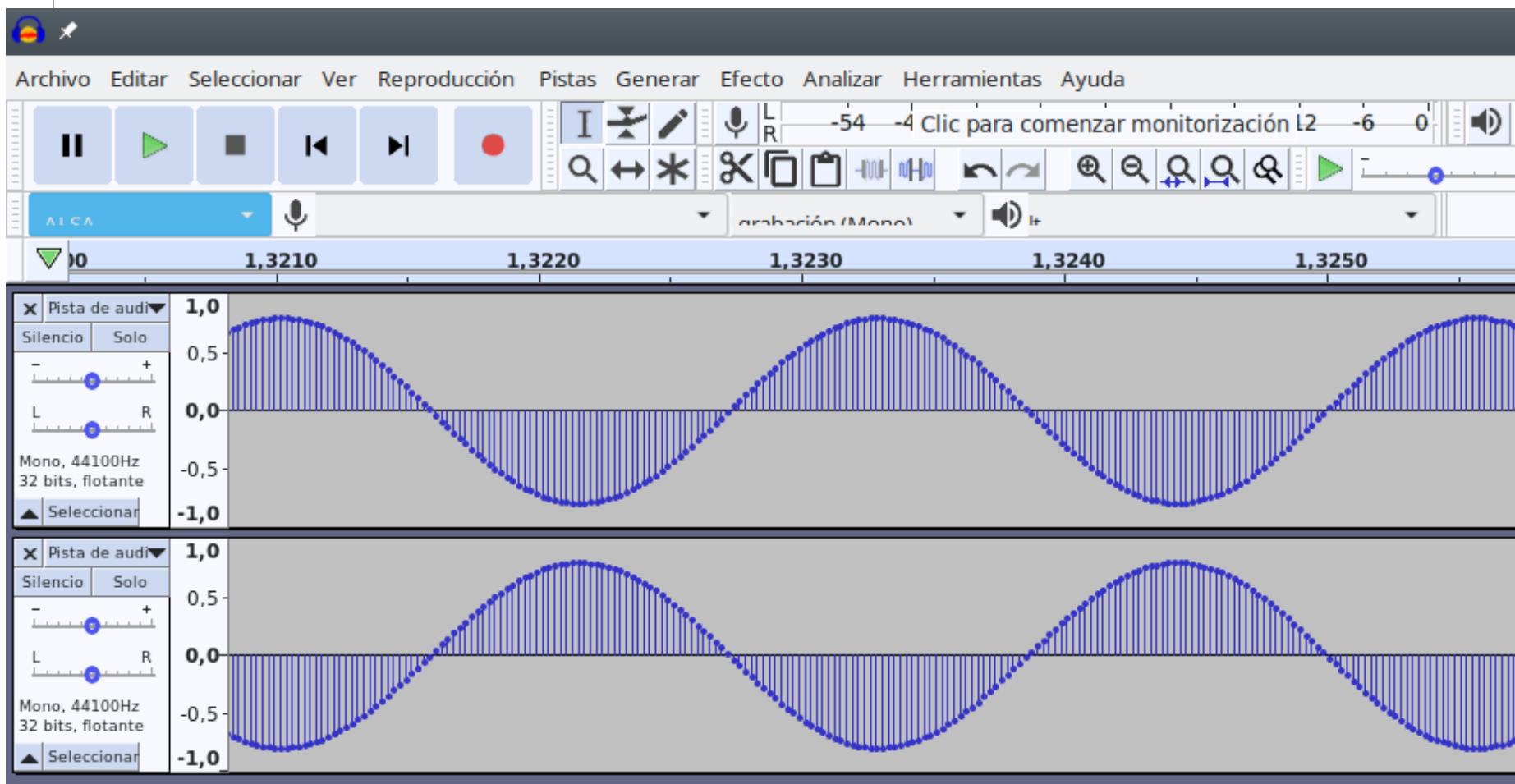
Dominio de representación (II)

- Sonidos en desfase



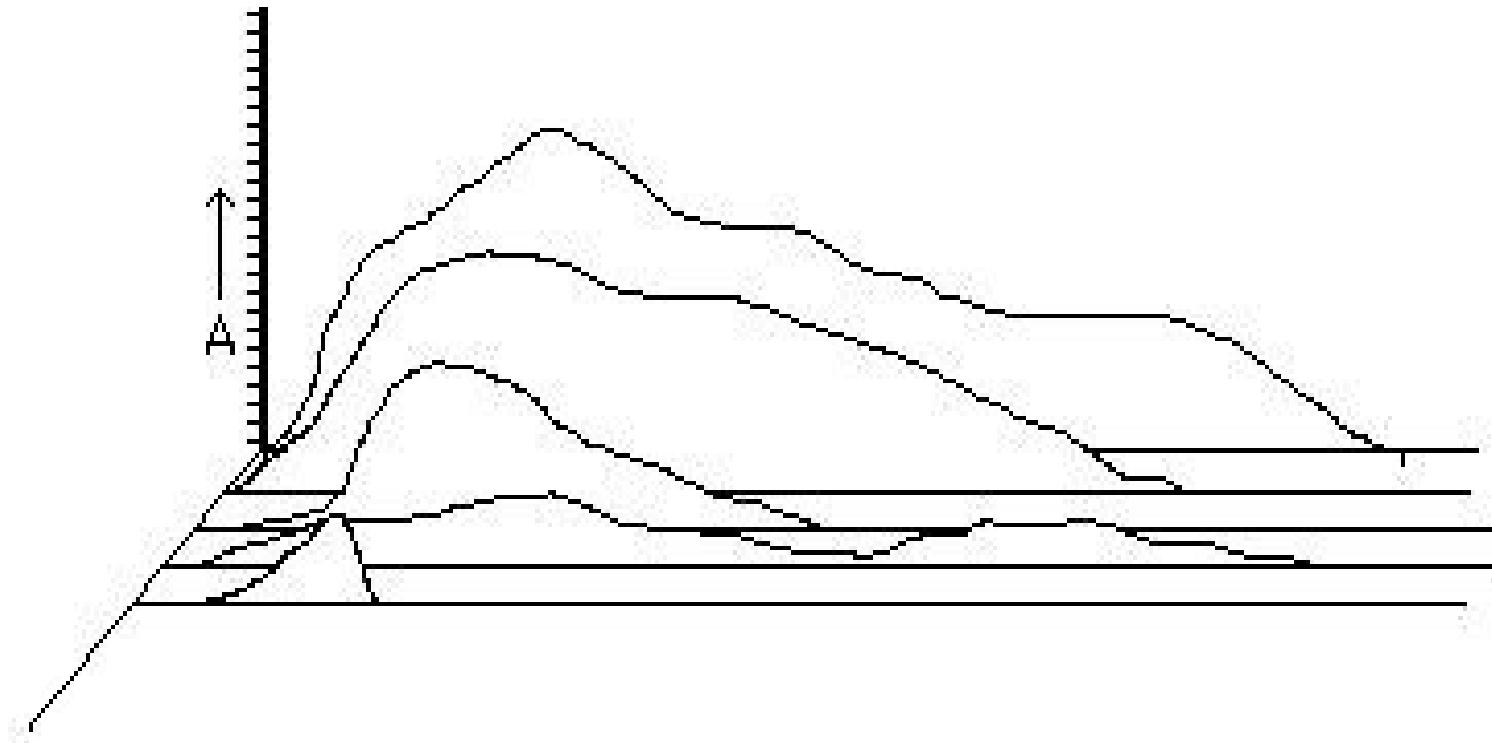
Dominio de representación (III)

- Modelado del sonido en el tiempo
 - ★ Cancelación de sonidos



Dominio de representación (y II)

- De la frecuencia



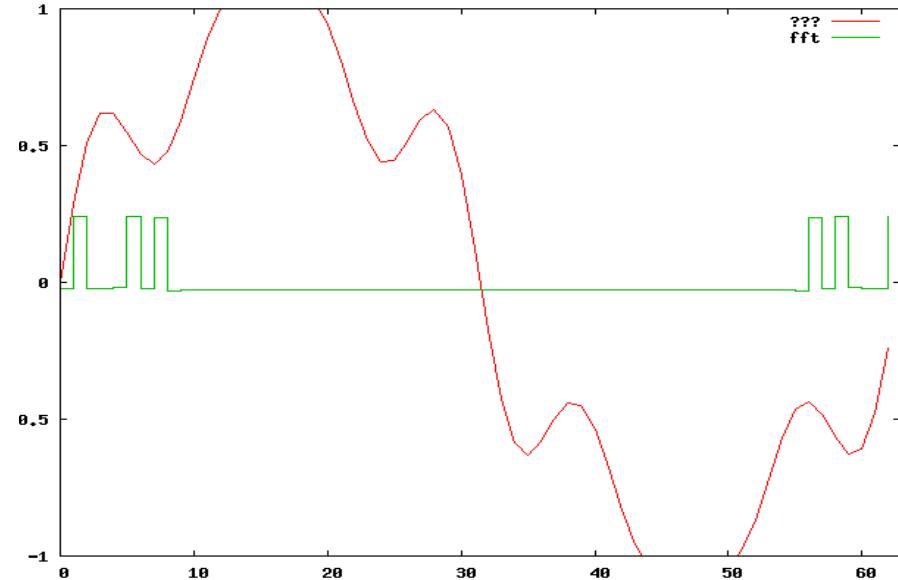
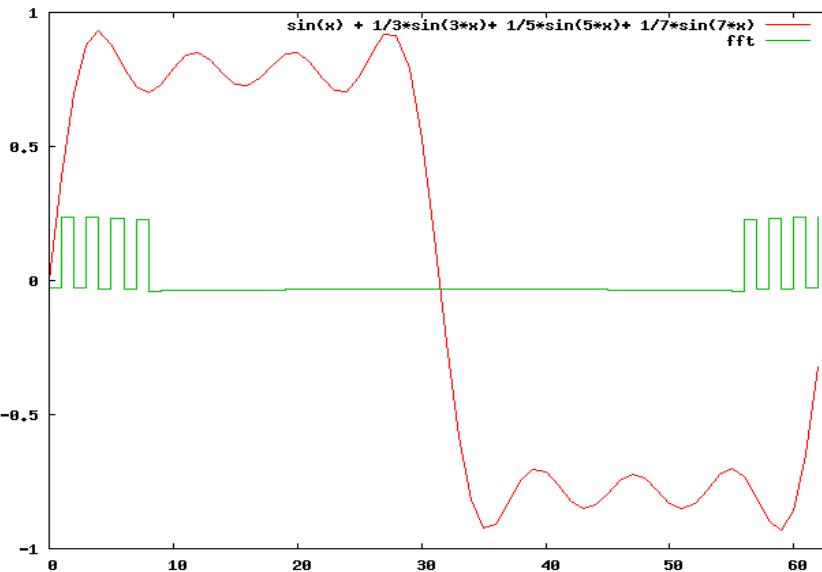
Dominio de representación (y II)

- Transformada de Fourier
 - **Tiempo ↔ Frecuencia**



Modelado y percepción

- Modelado del sonido en frecuencia:
 - ★ Transformada de Fourier
 - Frecuencia fundamental y armónicos

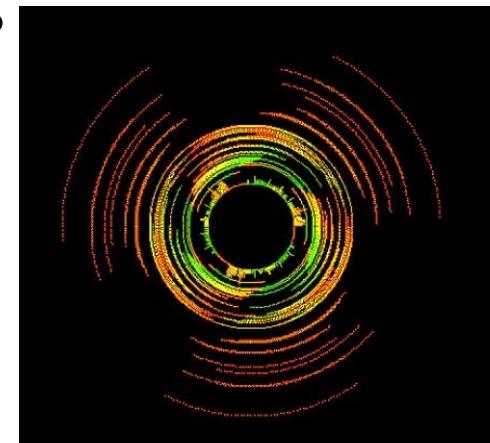


Modelado y percepción (II)

- Modelado del sonido en frecuencia
 - ★ *EQU Graphical Equalizer XMMS Effect Plugin*
 - Sólo MP3 <<http://equ.sourceforge.net>>

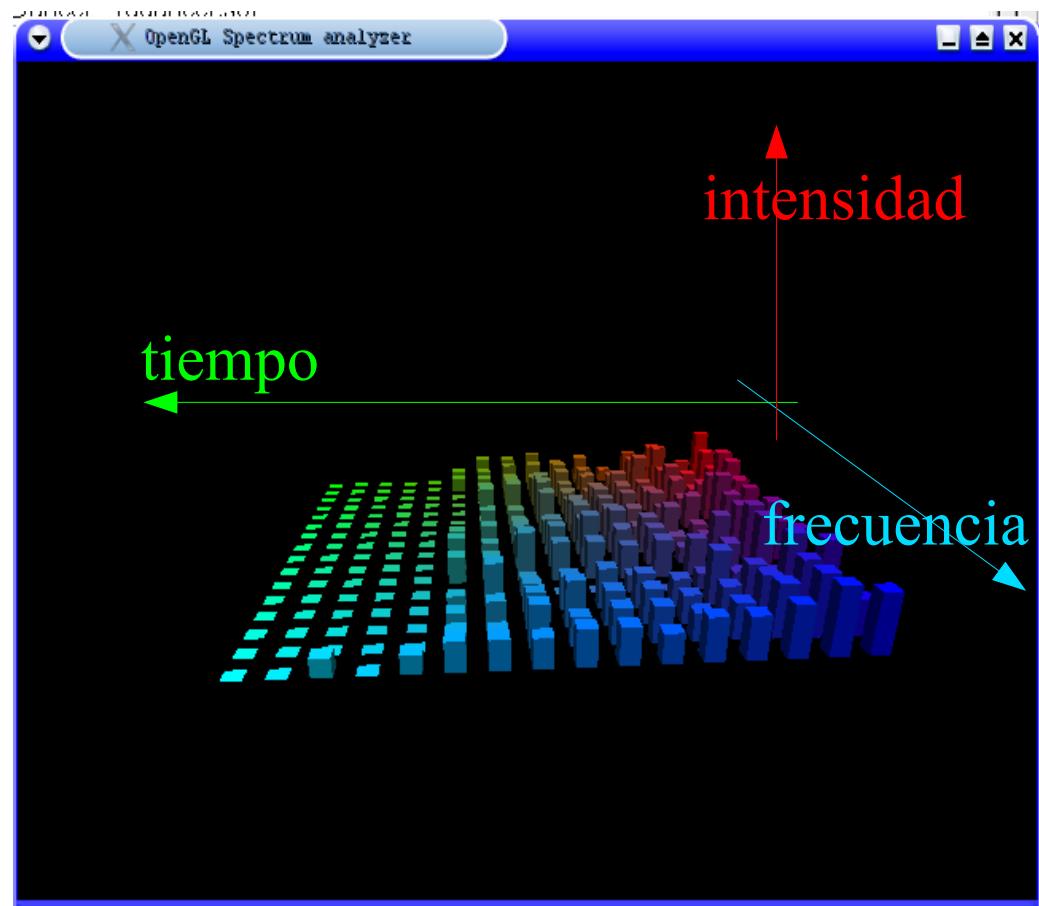
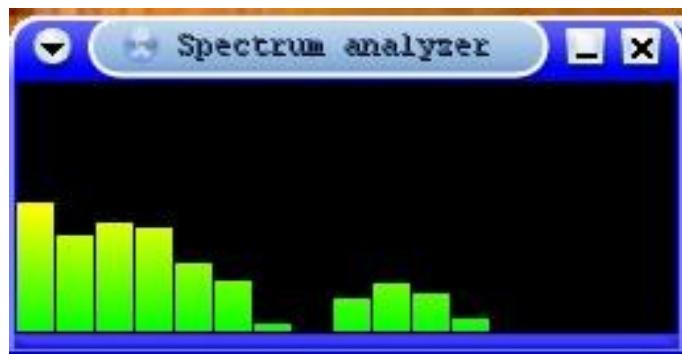
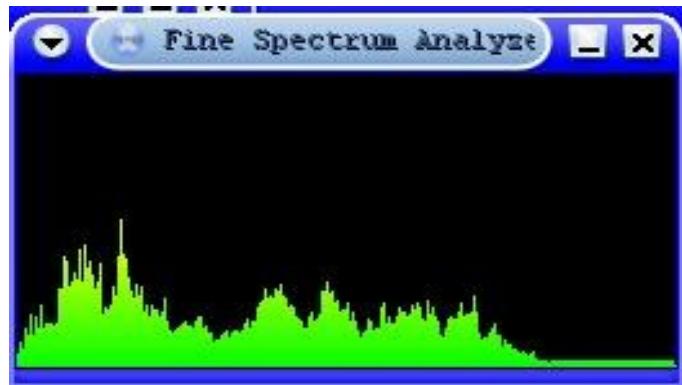


- ★ Visualizar el espectro de frecuencias



Modelado y percepción (y III)

- Modelado del sonido en frecuencia
 - ★ XMMS: visualizar el espectro de frecuencias

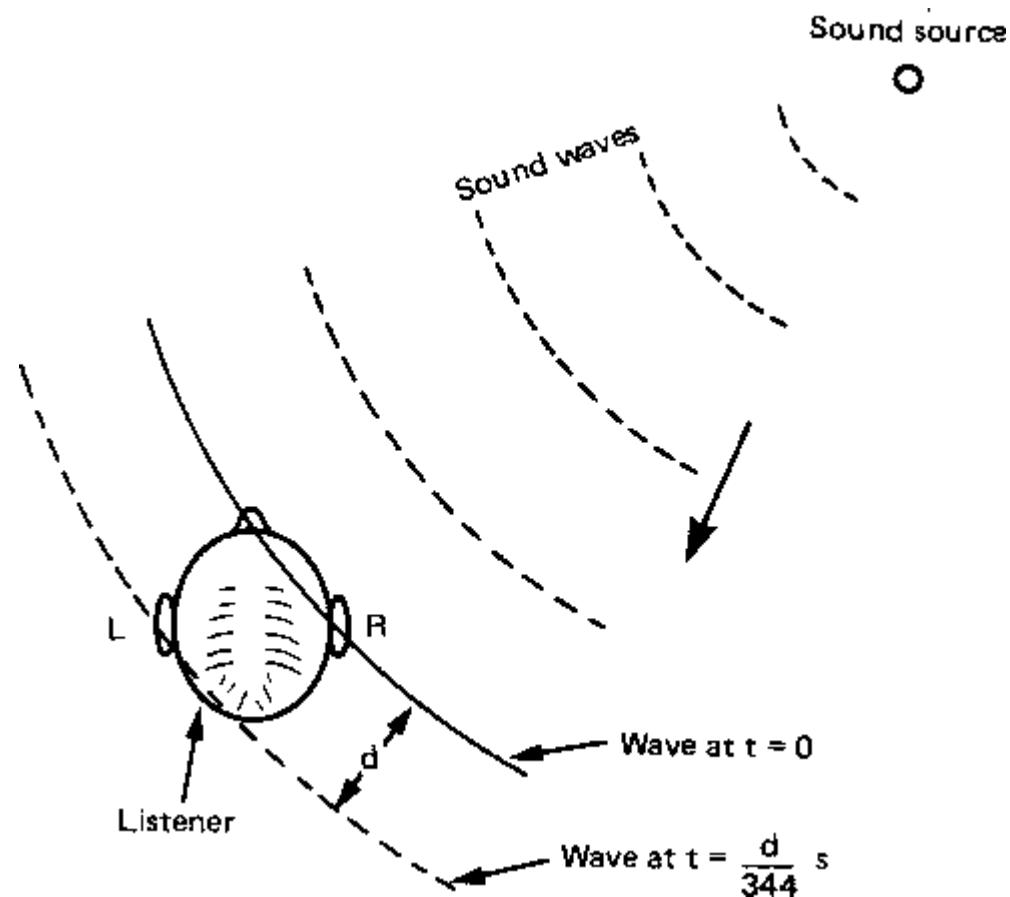


Acústica

- Cómo suena ...
 - * ¿Un árbol que cae en medio del bosque y no hay nadie allí para oírlo?
 - * ¿Una sirena al pasar?
- Estudio del sonido y la percepción: audio
 - * El fenómeno físico
 - * ... y el sistema de percepción humano interpreta
 - Psicoacústica
 - Oído humano

Acústica (2)

- Y el sistema de percepción interpreta:
 - ★ 20 .. 20k Hz,
 - ★ edad,
 - ★ volumen.
 - ★ bel → decibel \Rightarrow 0dB
 - Escala logarítmica:
20 dB .. 135 dB
 - ★ Dirección del sonido

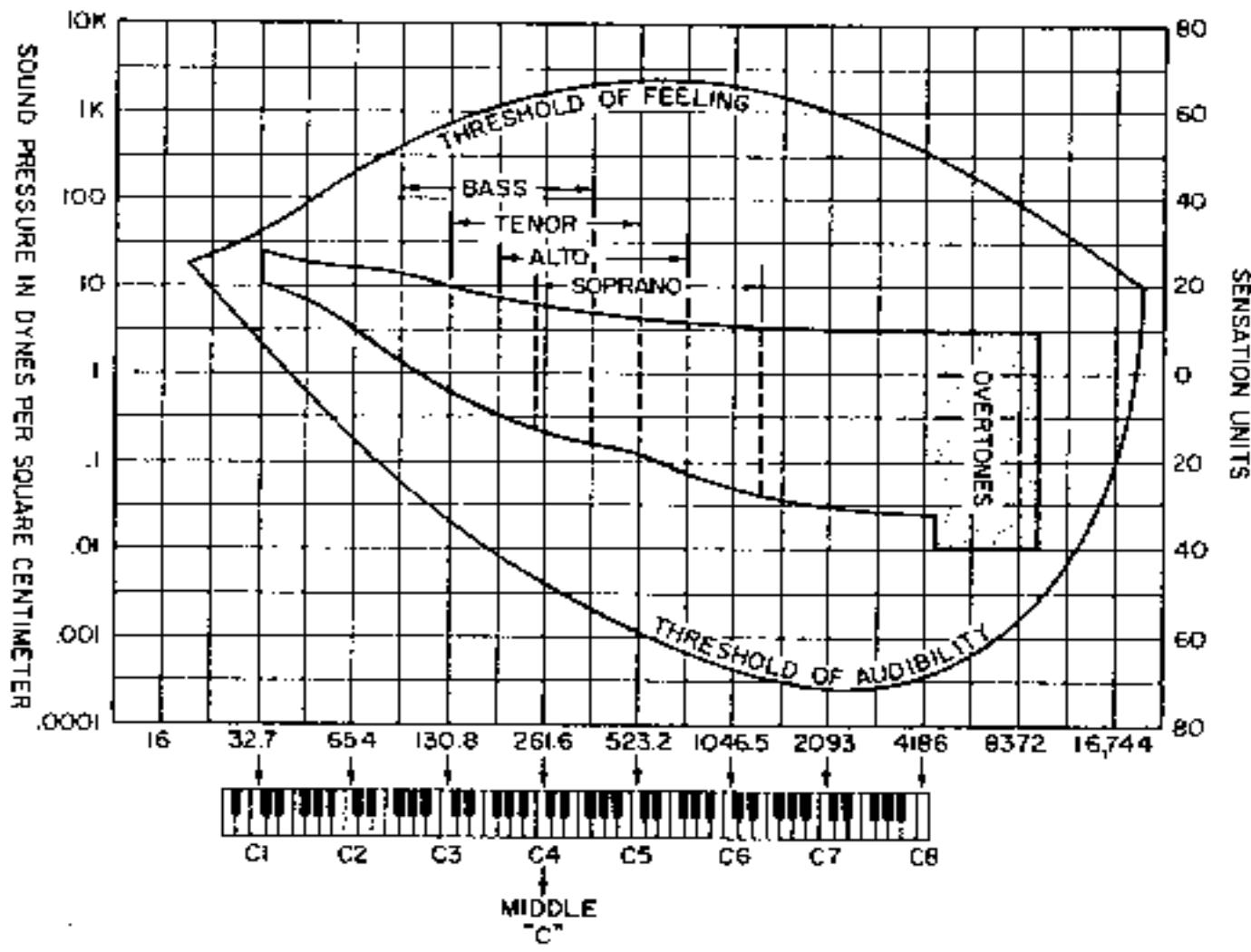


Percepción: Psicoacústica

- El habla es una forma de onda compleja
 - ★ Las vocales y los sonidos graves son bajas frec.
 - ★ Las consonantes son altas frecuencias
- El ser humano es más sensible a las bajas frecuencias
 - ★ La región más importante es 2 kHz a 4 kHz
- La percepción depende también de las condiciones ambientales
- Algunos sonidos son enmascarados por otros cercanos en frecuencias (MPEG).

Respuesta del oído humano

- 20 .. 20k Hz :: 2 .. 4 kHz



Pérdida auditiva con la edad

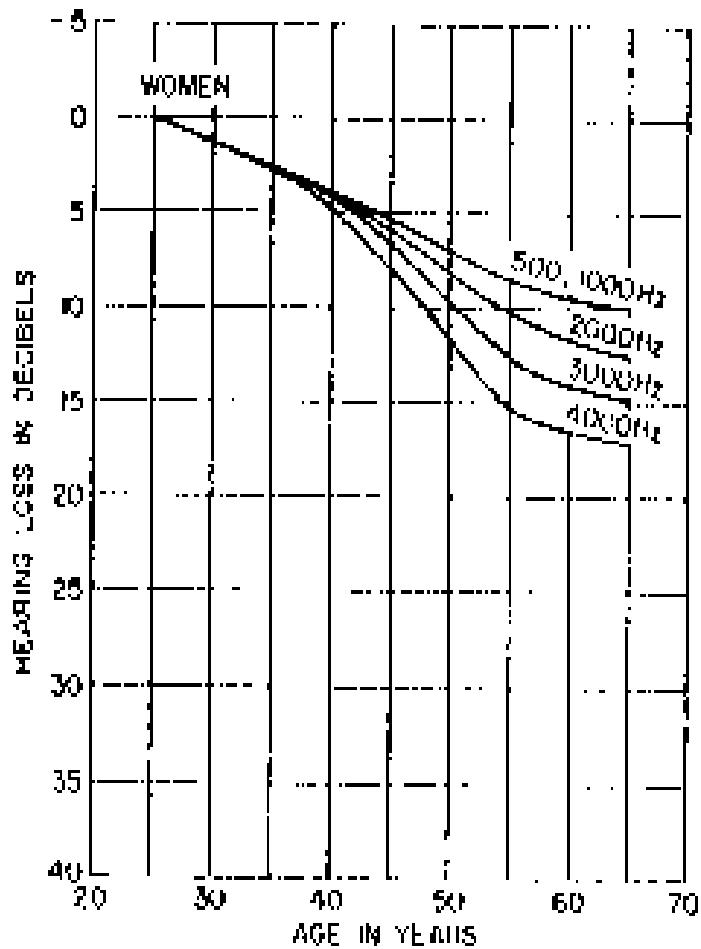


Fig. 1-99B. Typical hearing losses for women. (Courtesy, General Radio Co.)

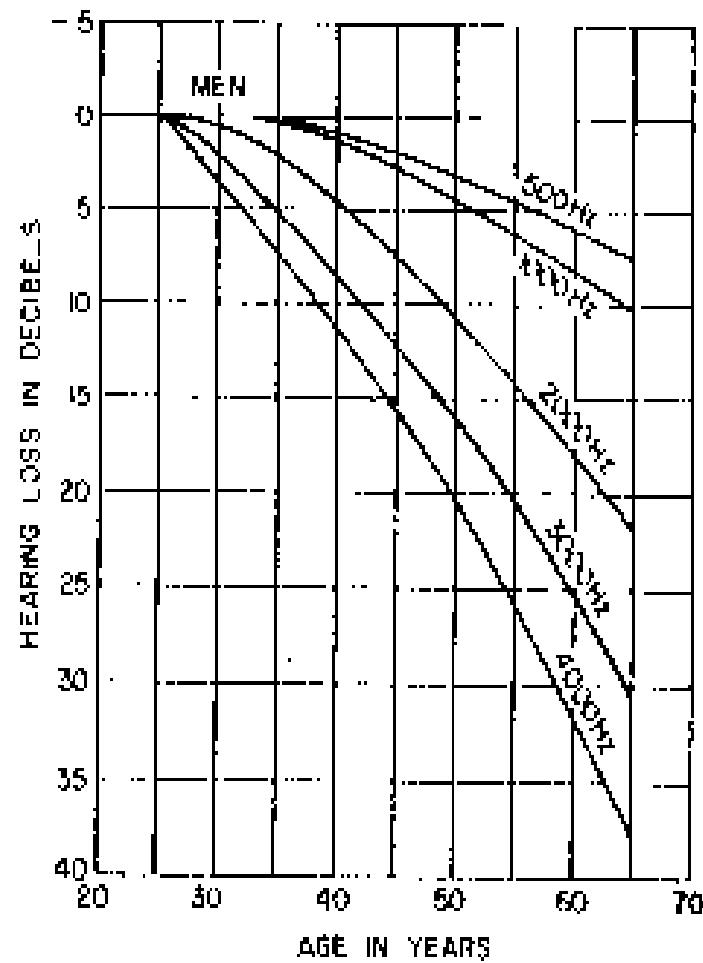


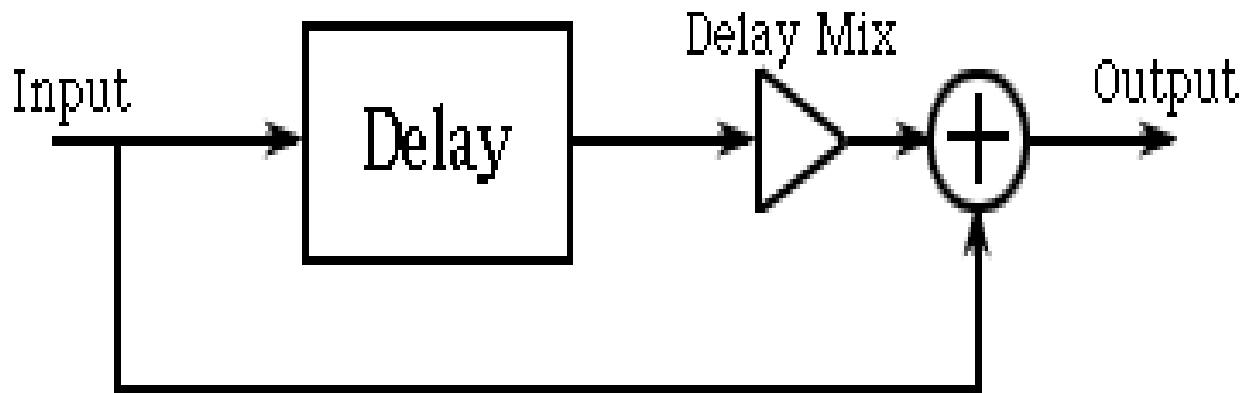
Fig. 1-99A. Typical hearing losses for men. (Courtesy, General Radio Co.)

Introducción al procesado de audio

- Efectos básicos
 - ★ Retraso (*delay*), coros (*chorus*) y reverberación (*reverb*) y ecualizar.
 - ★ Modulación de señal (envolvente) y filtros.
- Recreación de efectos reales
 - ★ Doppler
 - ★ Audio 3D
- Síntesis
 - ★ FM, tabla de ondas, modelado, aditiva, sustractiva, modular, ...

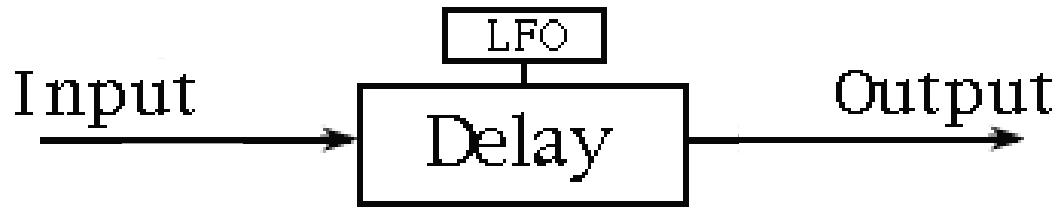
Efectos de audio: básicos

- Retraso (*delay*):



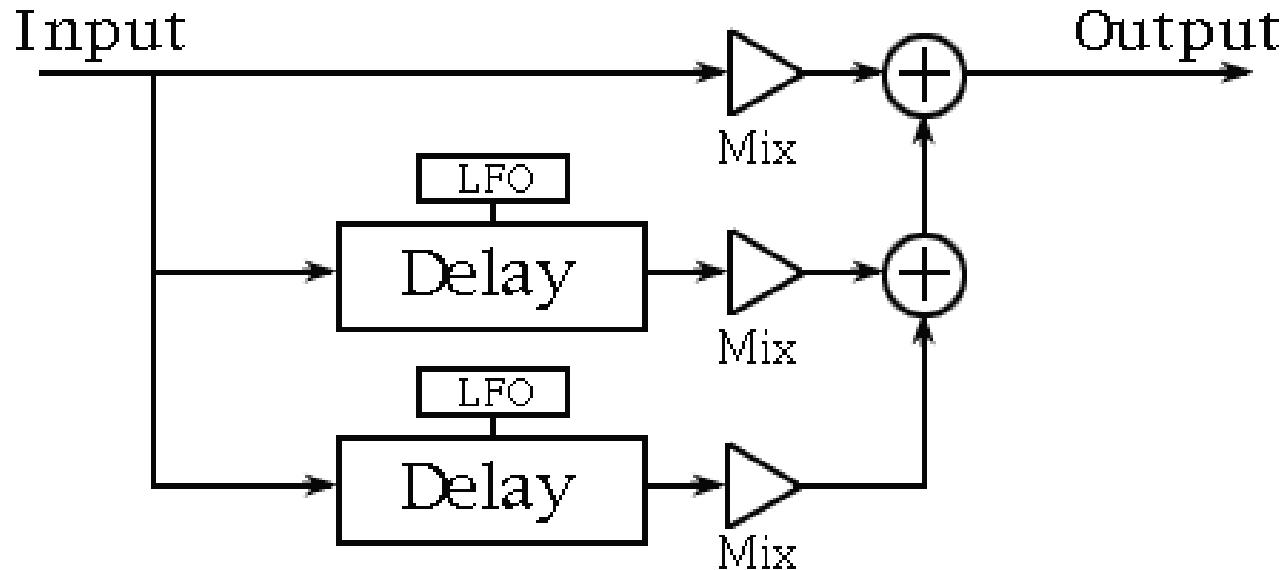
Efectos de audio: básicos (II)

- Eco (Echo)



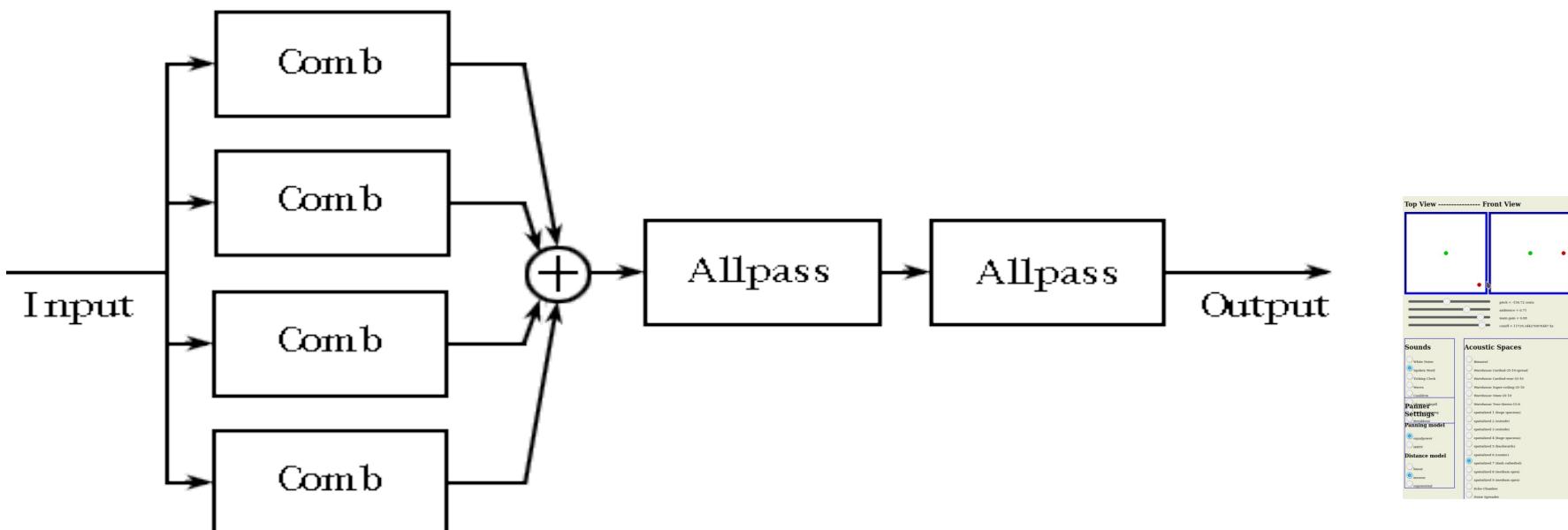
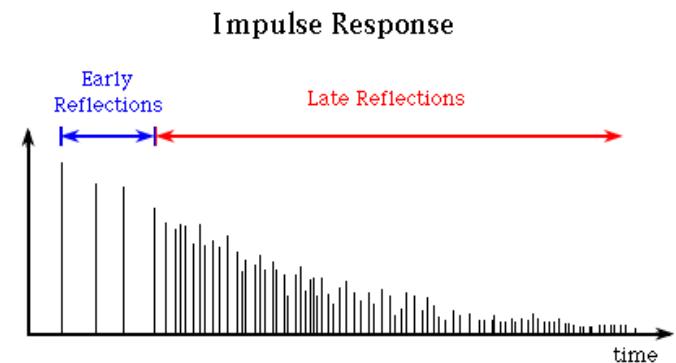
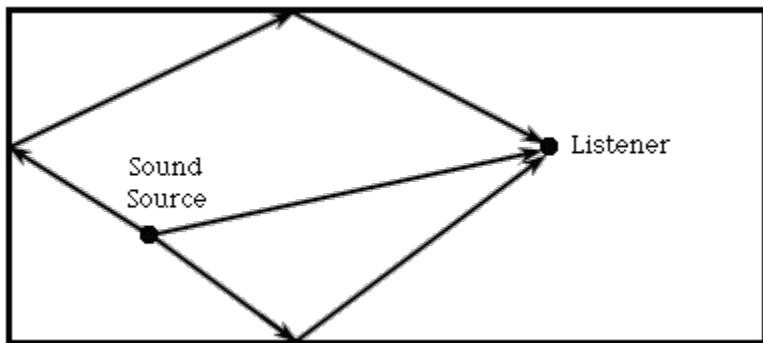
Efectos de audio: básicos (III)

- Coros (*chorus*):



Efectos de audio: básicos (IV)

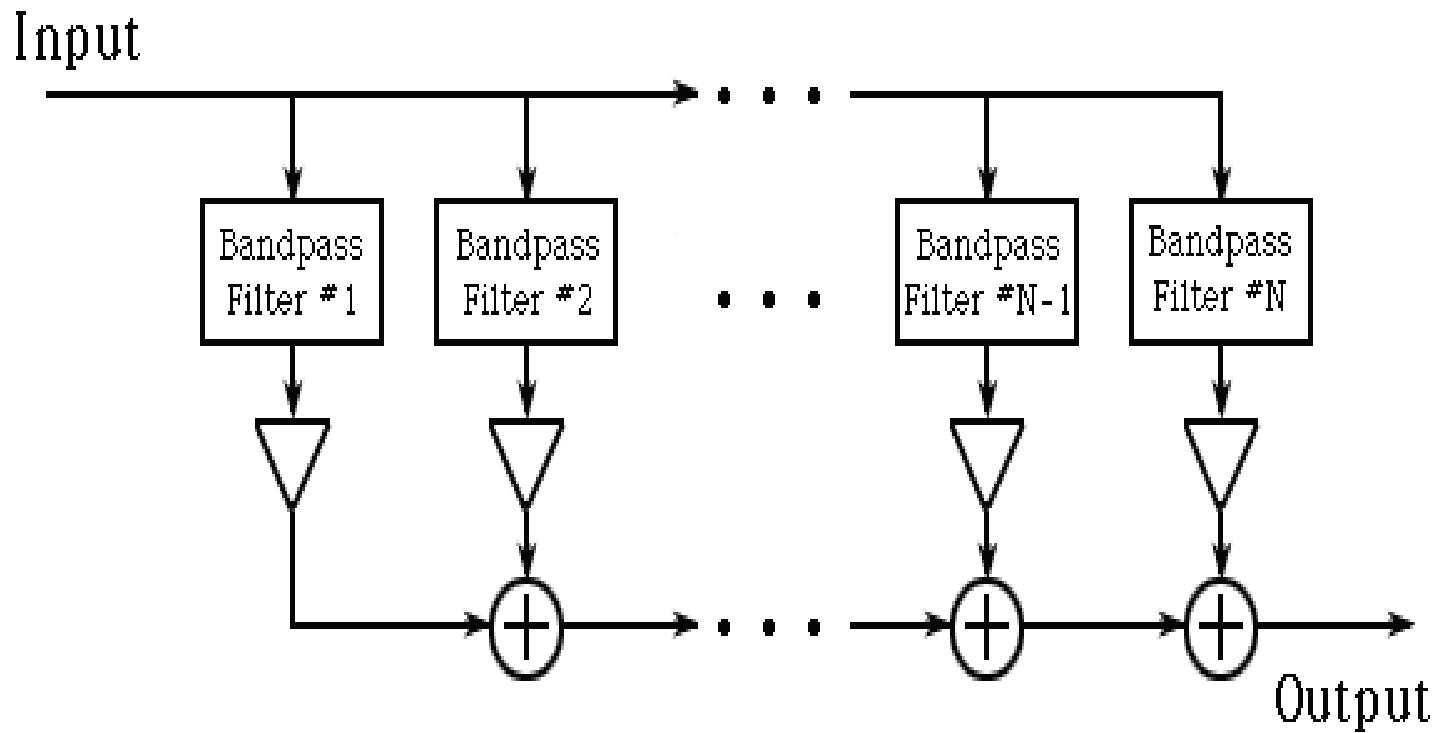
- Reverberación:



Fuente: Harmony Central®: Effects Explained: Reverberation <<http://www.harmony-central.com/Effects>>,
Reverberation <http://valy.ika.eu/DIY/soundsystem/_upload_by_VeeHell/know2how/Audio_Tutorials_Library_By_LooPus/fx_basics/ambiance_fx/Reverberation.htm>
Panning and Reverberation <<https://googlechromelabs.github.io/web-audio-samples/archive/demos/simple.html>>

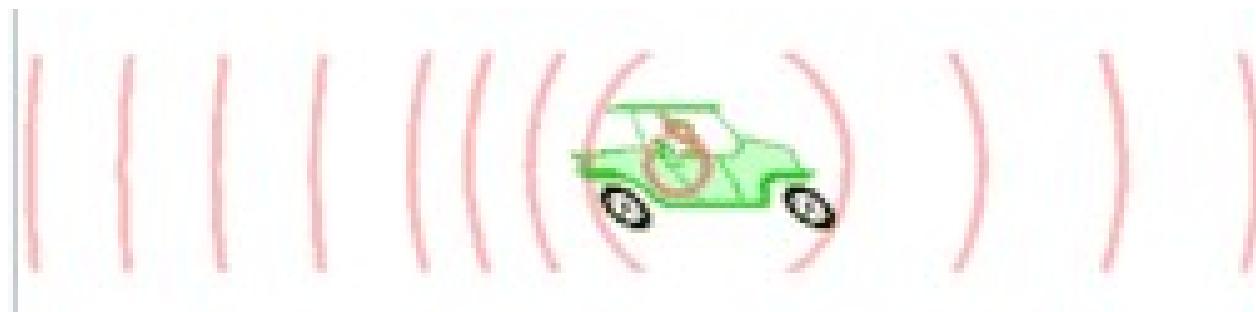
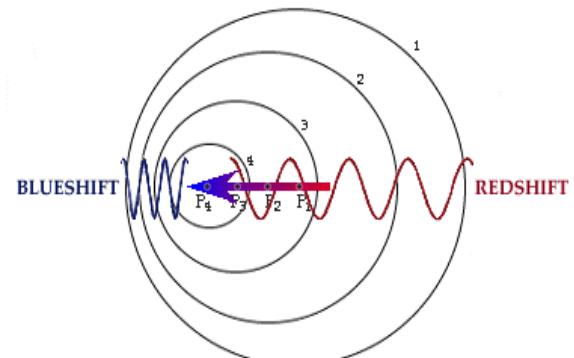
Efectos de audio: básicos (VI)

- Ecualizador:



Efecto reales: Doppler

- Fenómeno físico y recreación
 - Relación entre frecuencia emitida y recibida.



$$\bullet \text{Doppler}_{\text{OpenAL}} = f(V_{\text{Emisor}}, V_{\text{Oyente}}, V_{\text{Sonido}}) \rightarrow f' = \frac{f \cdot (V_{\text{sonido}} - (k * V_{\text{Oyente}}))}{(V_{\text{Sonido}} - (k * V_{\text{Emisor}}))}$$

Efectos reales: audio 3D

- Fenómeno físico y recreación

- ★ El sonido envolvente (*surround*) se refiere al uso de múltiples canales de audio para provocar efectos de localización espacial del sonido a la audiencia, ya sea proveniente de una película o de una banda sonora.
- ★ Recreación: implementación



Efectos reales: audio 3D (II)

- Fenómeno físico y recreación

- ★ El sonido envolvente (*surround*) se refiere al uso de múltiples canales de audio para provocar efectos de localización espacial del sonido a la audiencia, ya sea proveniente de una película o de una banda sonora.

- ★ Recreación: implementación

- ★ Plugins

- Xmms: sound enhancing plugins

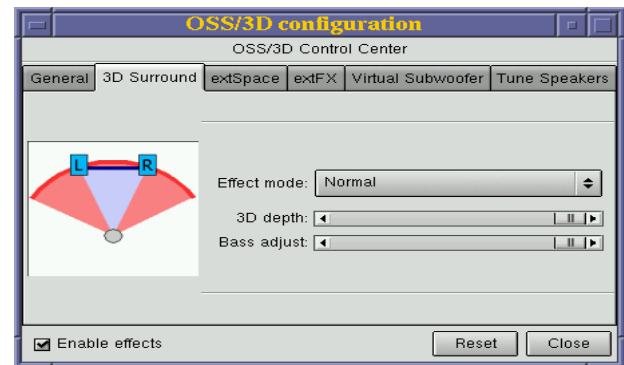
- 4Front Technologies OSS/3D

- ★ Motores de audio

- OpenAL

- FMOD

- ★ Sistema hardware



Efectos reales: audio 3D (III)

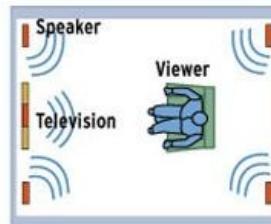
- Fenómeno físico y recreación

- ★ Recreación: implementación
 - ★ Plugins
 - ★ Motores de audio
 - ★ Sistema *hardware*

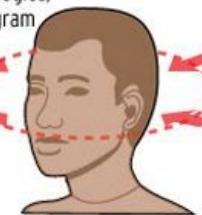
- *Canales*

- *Monofónico.*
 - *Estereofónico*
 - *Cuadrifonía o multicanal*

Most home theater systems today consist of 2, 5 or 7 speakers (plus a subwoofer) to create a **surround sound experience**.

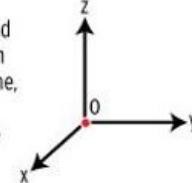


Using existing technologies, audio mixers can program sound to come out of **channels** to the front, back or sides to give that feeling of sound **coming from around you**.

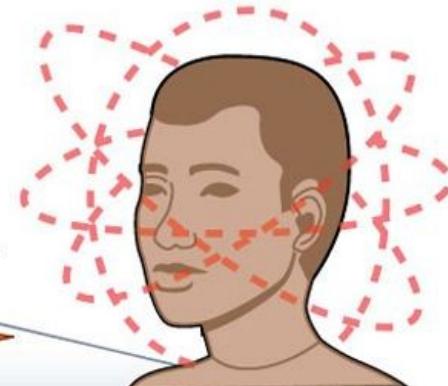


SRS Labs is pushing a fundamental paradigm shift that eliminates the concept of **channels** in favor of **three dimensional audio**.

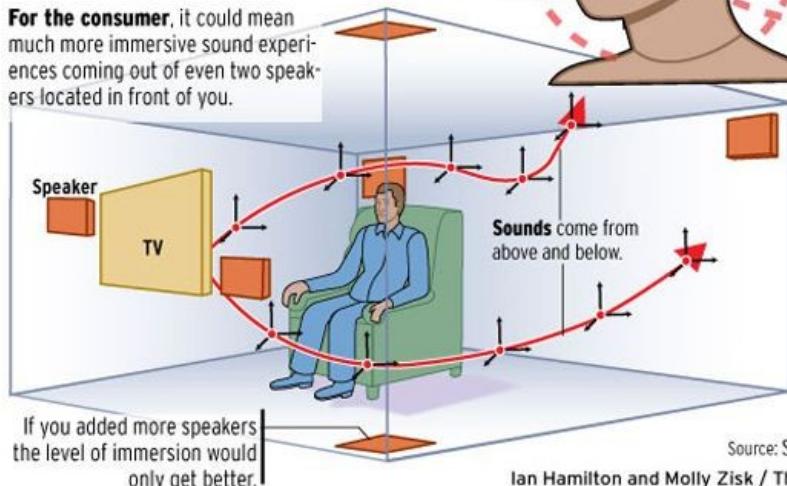
Pieces of recorded sound are defined as objects in **3D space**. And, over time, the audio can move through a room in three dimensions.



With audio defined in 3D space an **audio mixer** could create one mix that is "future-proofed" for any and all speaker configurations.



For the consumer, it could mean much more immersive sound experiences coming out of even two speakers located in front of you.

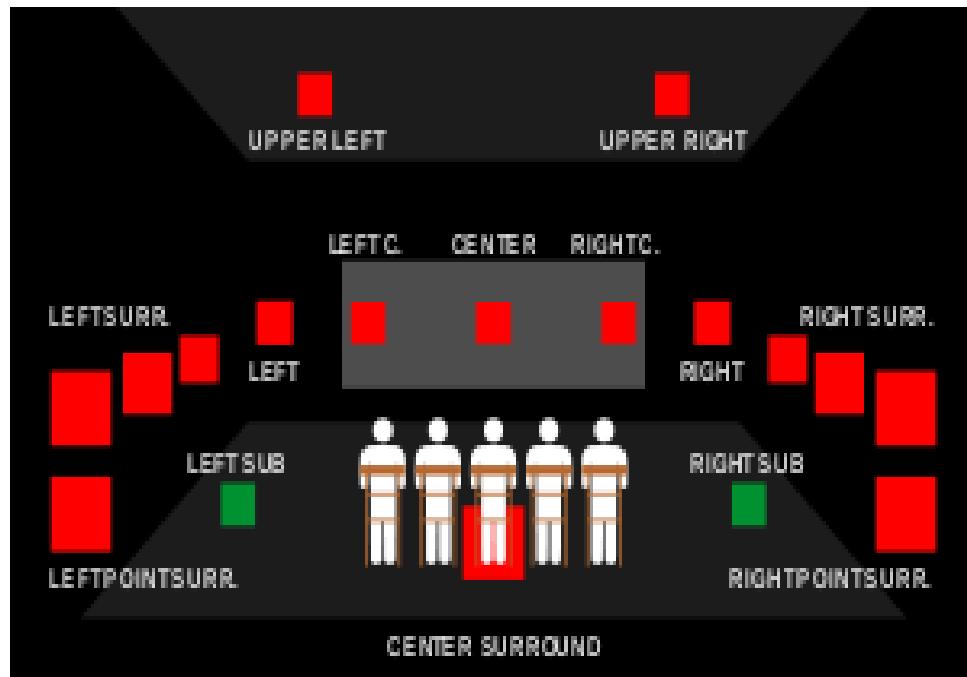
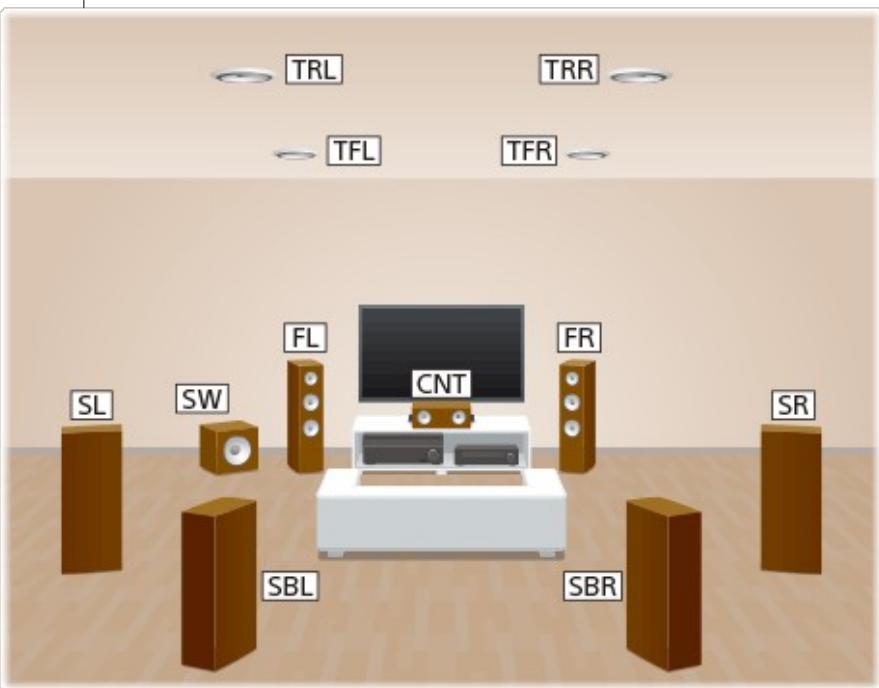


Source: SRS Labs, Inc.

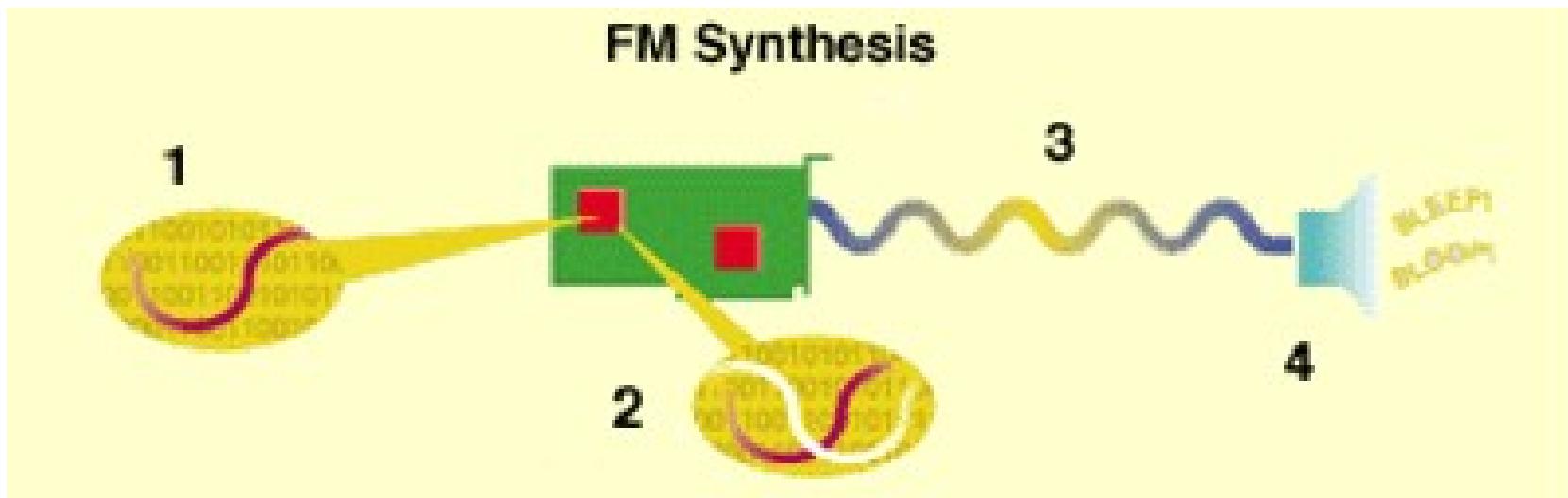
Ian Hamilton and Molly Zisk / The Register

Efectos reales: audio 3D (IV)

- Ej. de sistema envolvente
 - ★ Dolby Atmos



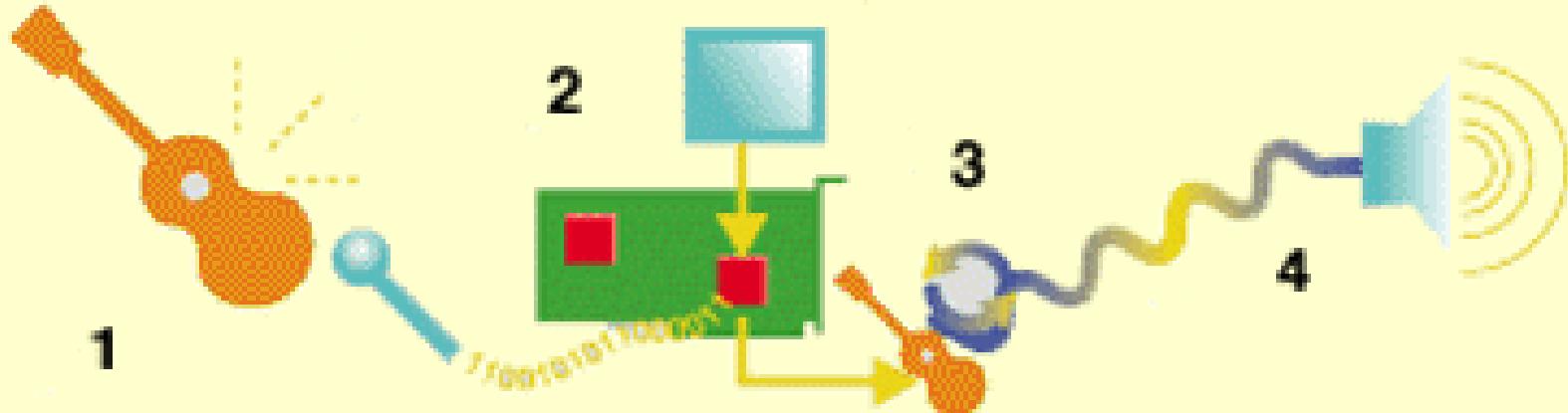
Síntesis de sonido: métodos



- 1. El chip de síntesis genera ondas senoidales
- 2. Dos o más ondas se combinan para formar señales más complejas... (modulación)
- 3. ...que son convertidas a señales analógicas
- Los sonidos generados por modulación de frecuencia tienden a parecer artificiales

Síntesis de sonido: métodos (II)

Wavetable Synthesis

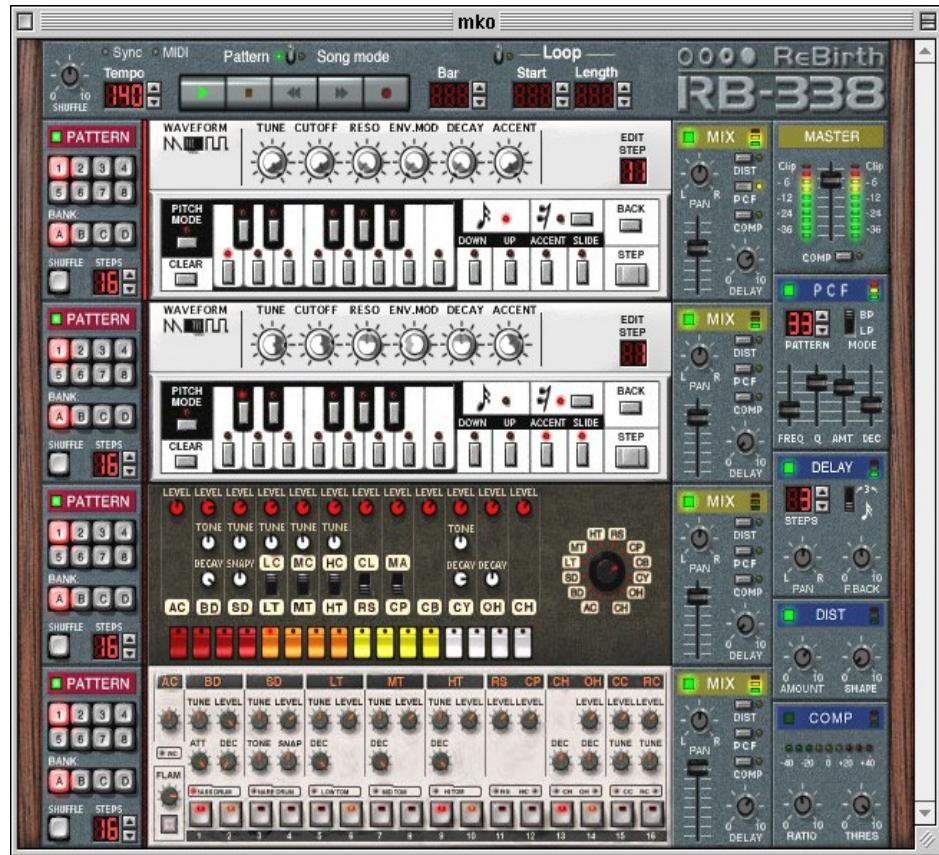


1. Los sonidos de instrumentos se graban en estudio
 2. Pequeñas muestras se guarda en ROM (tabla de ondas)
 3. Las aplicaciones los utilizan.
- La calidad es buena si lo son las muestras. Sin embargo, para reproducir con fidelidad un sonido no basta con guardar una grabación, ya que cuando se toca una nota diferente, no sólo se cambia la frecuencia del sonido sino otros parámetros importantes.

Síntesis de sonido: métodos (IV)

- Sintetizadores Software

- ★ Síntesis aditiva, sustractiva, modular, granular, ...
- ★ Generador de ruido
 - Ejemplos de sonidos naturales:
 - *SimplyNoise*
 - *NoiseMachine*



Síntesis de sonido: métodos (y VI)

- Ejemplo: generador de ruido

- *Noise machine*

- *150+ sound generators available*

- *Ejemplos*

- *Cafe Restaurant, Japanese Garden, Gregorian Chant, Medieval Forest, Examination Time, ...*

- *Rain Noise*

- *Custom-shaped Online Noise Machines*

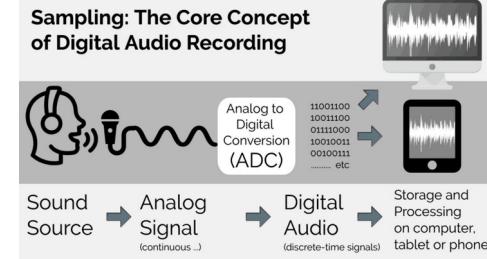
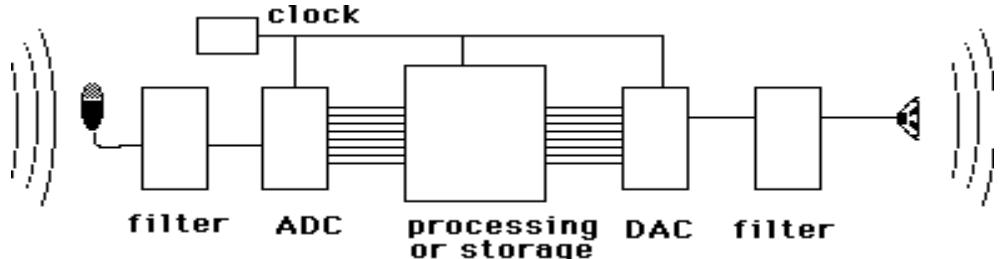
- *Frequency-Shaped Natural Rain Noise Generator*



Adquisición y reproducción

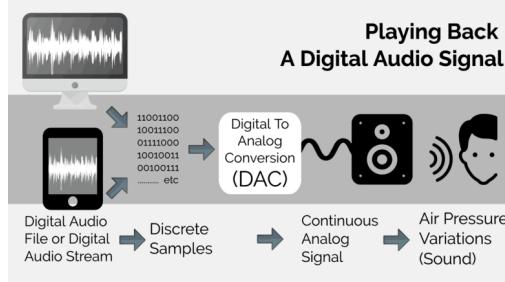
- Adquisición

- Un micrófono vibra de acuerdo con la presión ejercida por las ondas sonoras sobre su membrana.
 - Se transforma esta energía mecánica en energía eléctrica.



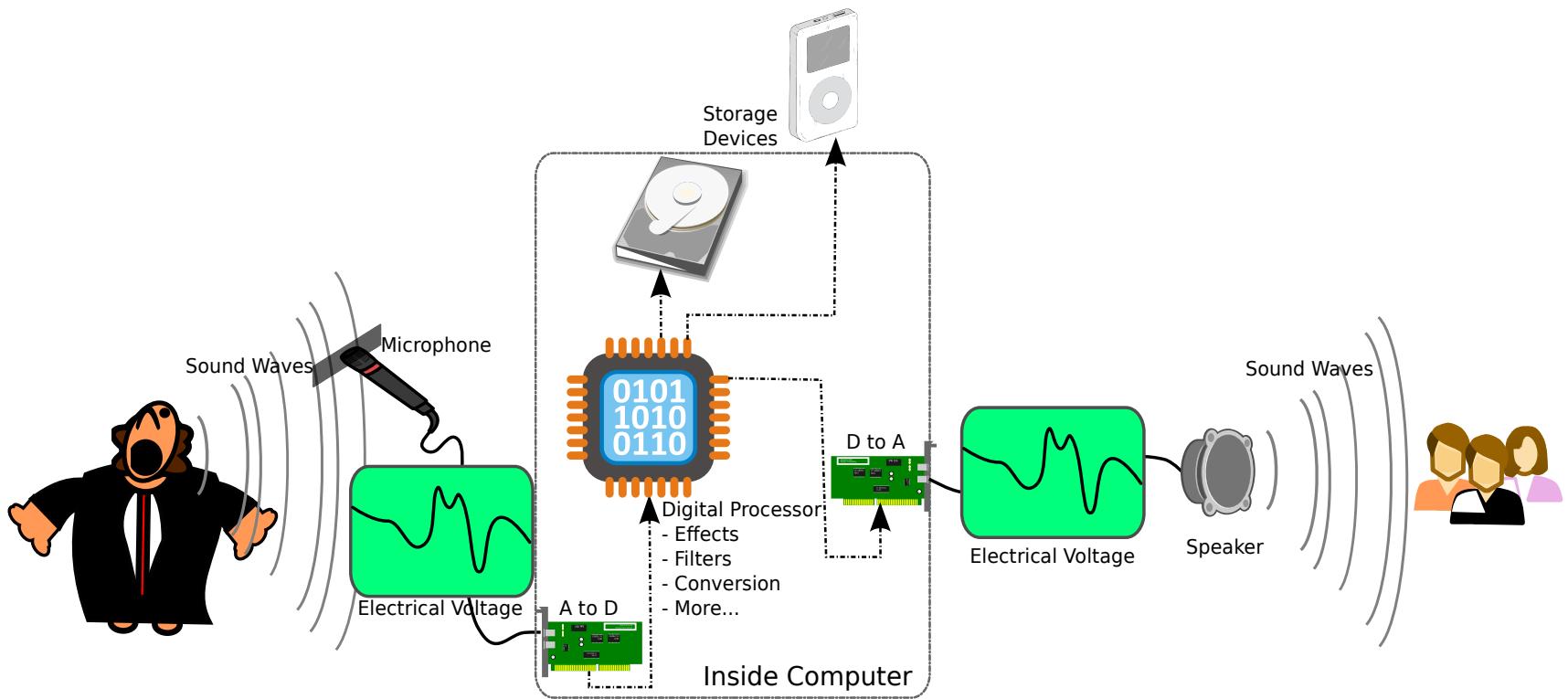
- Reproducción

- Un altavoz transforma la energía eléctrica en ondas sonoras



Adquisición y reproducción (II)

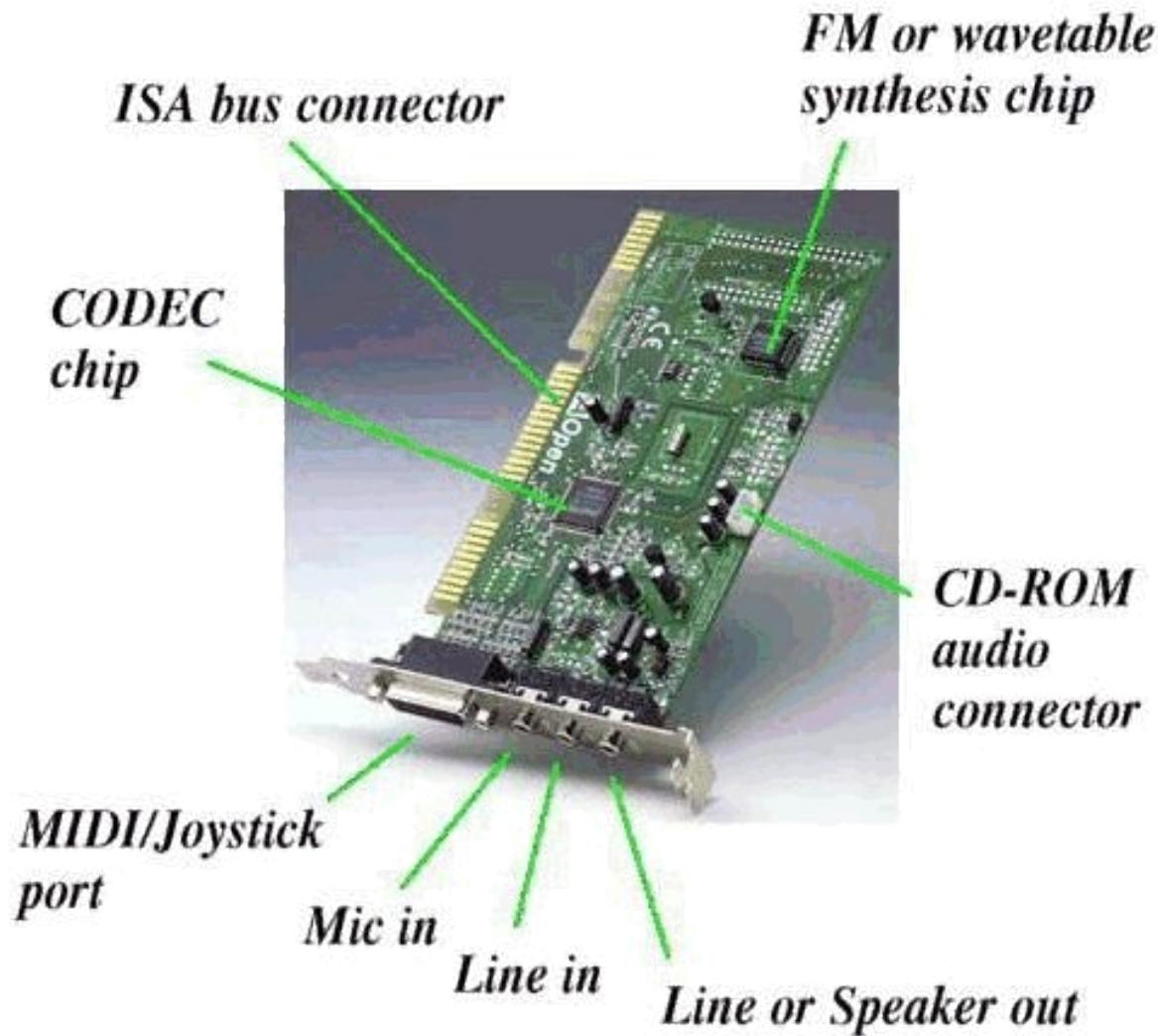
- El sonido en un computador



Adquisición y reproducción (III)

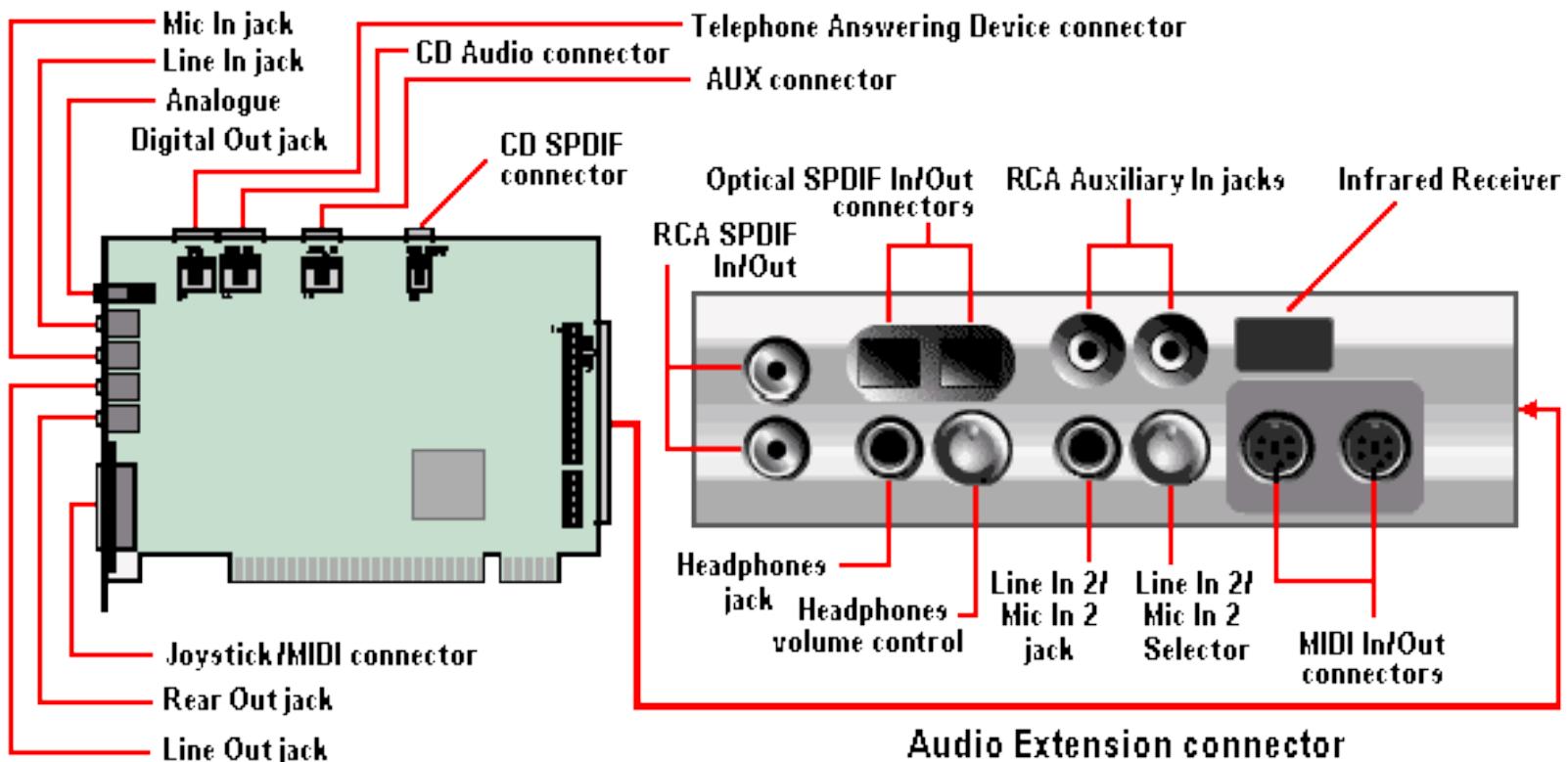
- Tarjetas de sonido: características

- ★ método de síntesis
- ★ n° bits
- ★ freq. muestreo
- ★ 32, 64, ... voces
- ★ estéreo/multicanal
- ★ *full duplex* y *enhanced*
- ★ *DSP: sonido 3-D alterando*
 - Amplitud
 - Frec. (Doppler)
 - Fase



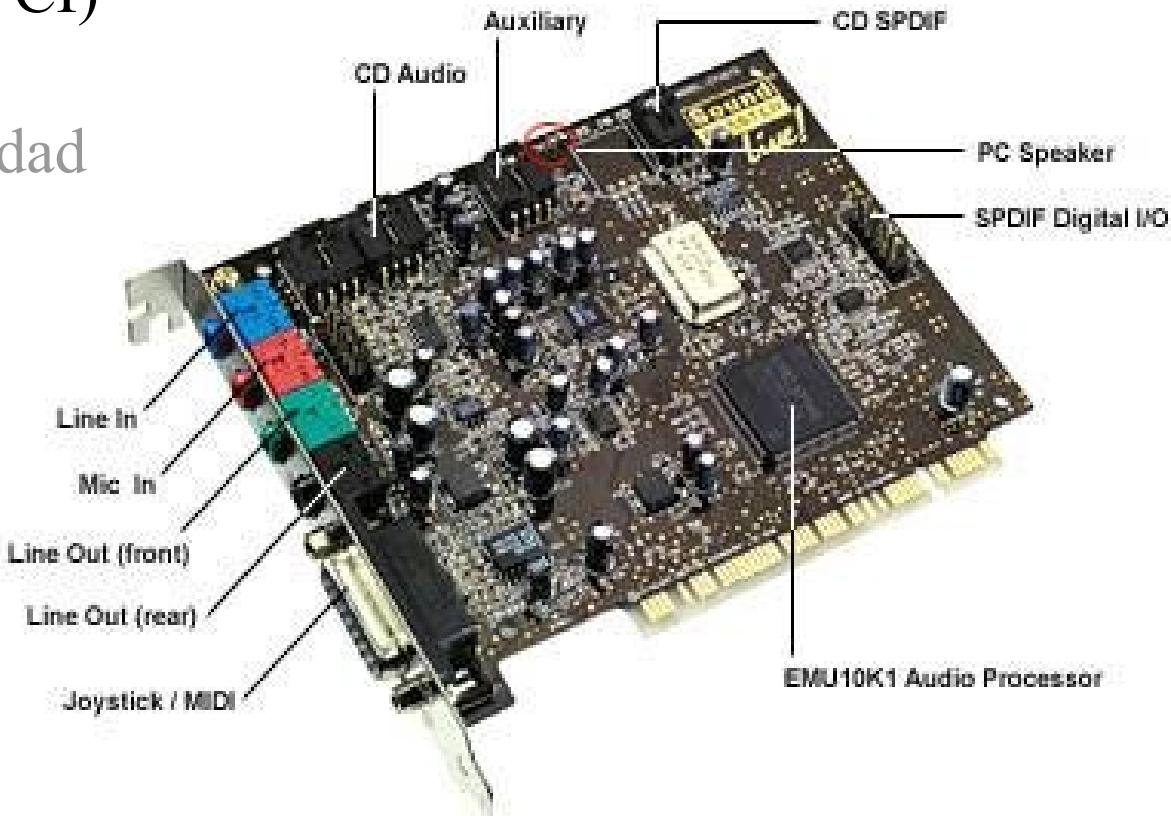
Tarjetas de sonido (IV)

- Conectividad



Adquisición y reproducción (V)

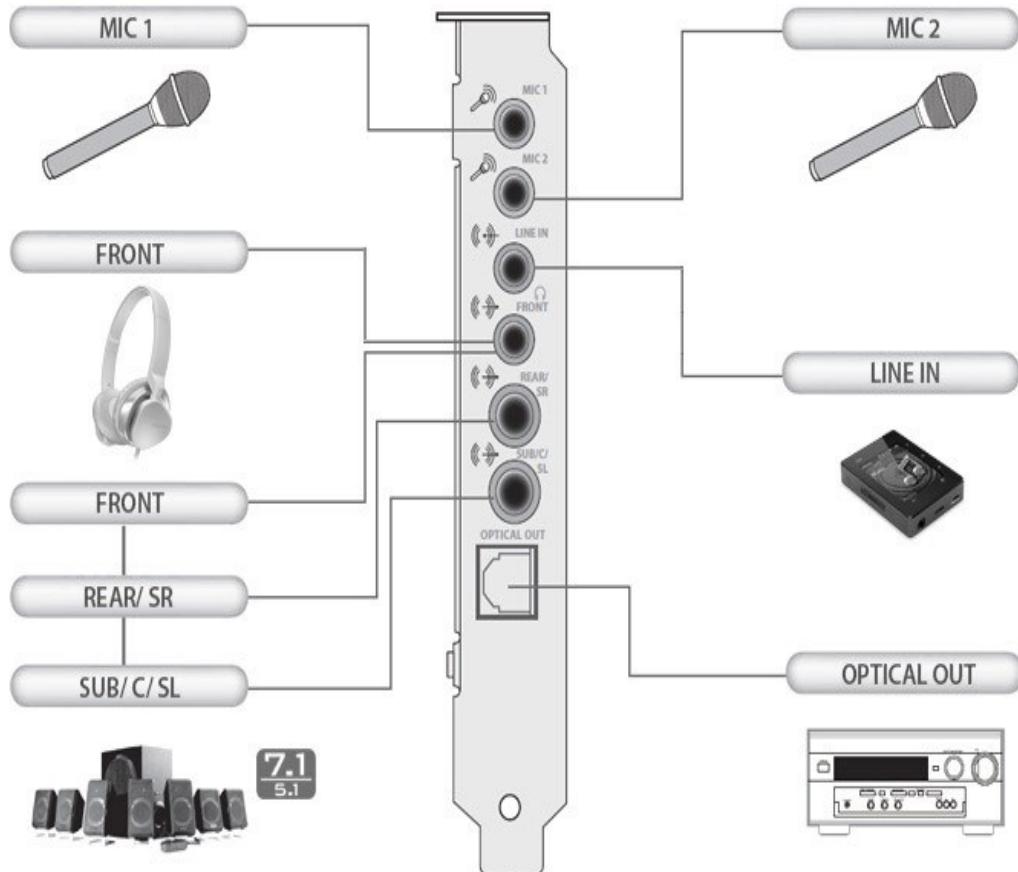
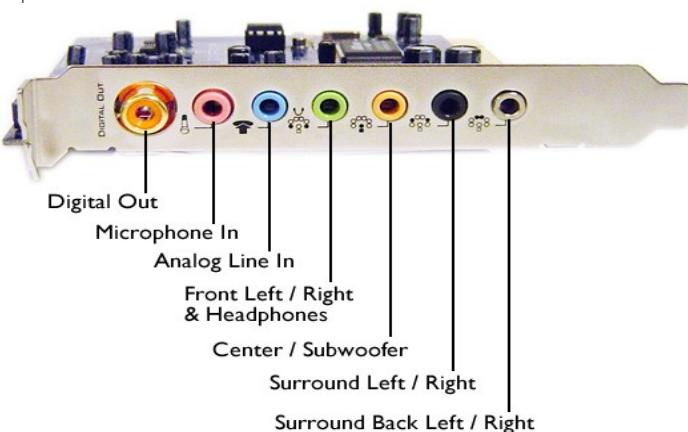
- Tarjetas de sonido:
 - ★ Creative SB Live! (PCI)
 - Bloques
 - Conectividad



★ Creative SB X-Fi PCI

Adquisición y reproducción (VI)

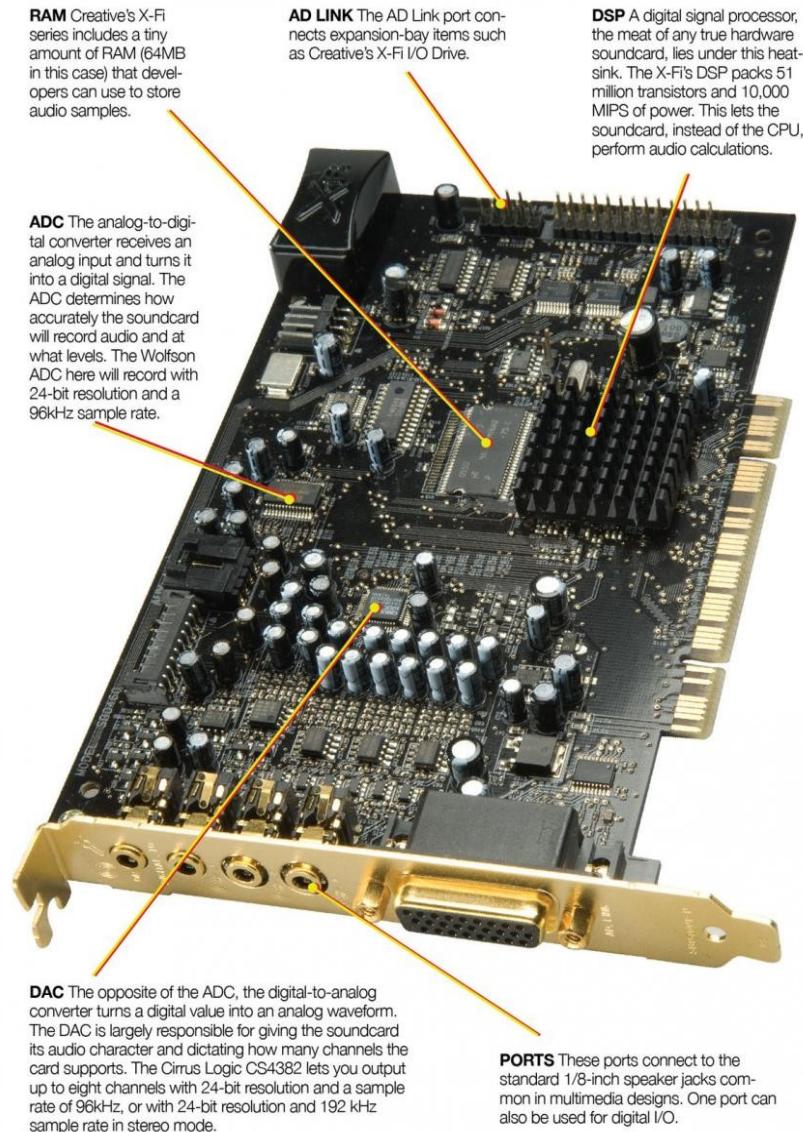
- Tarjetas de sonido:
 - ★ Creative SB Live! (PCI)
 - Bloques
 - Conectividad



★ Creative SB X-Fi PCI

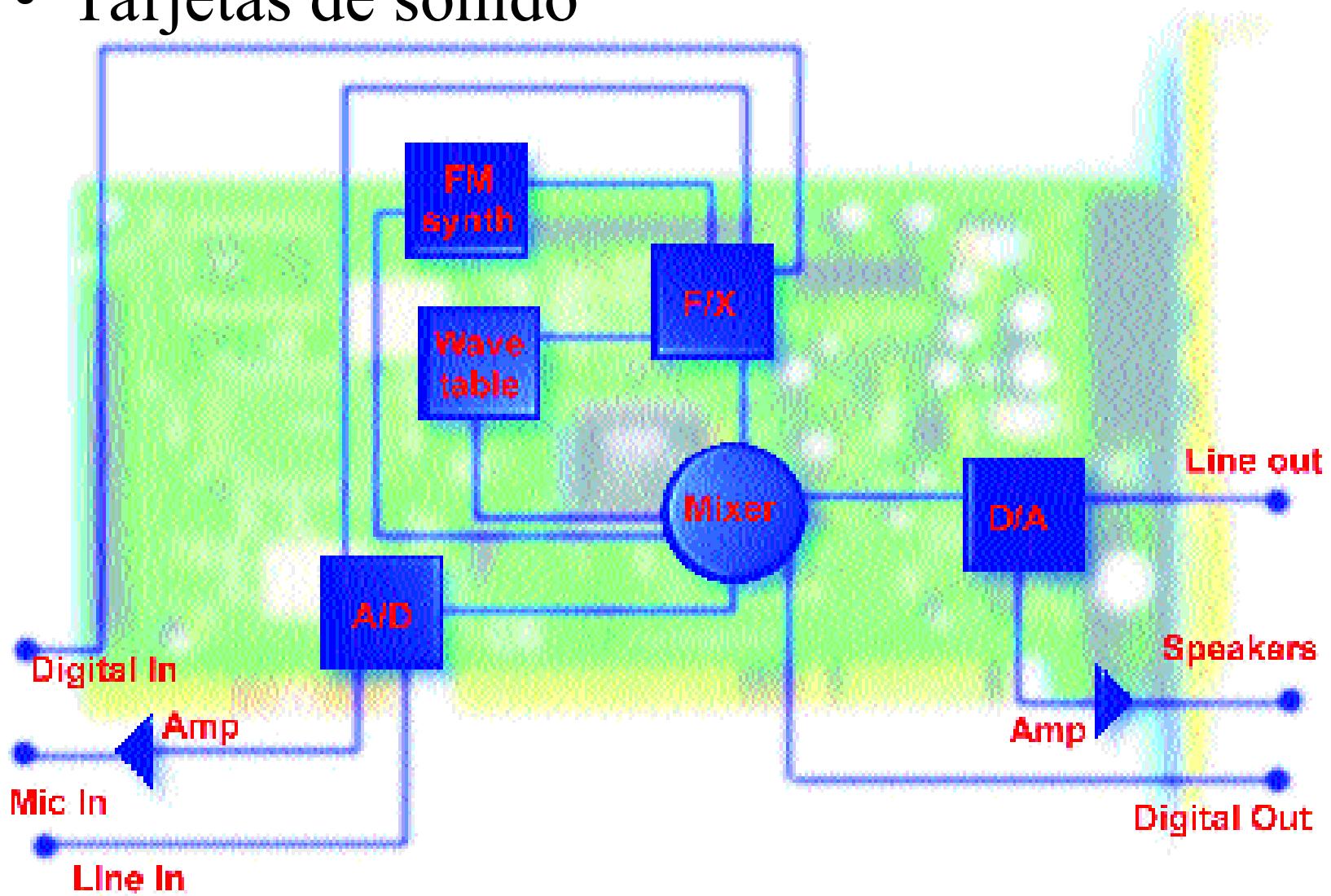
Adquisición y reproducción (VII)

- Tarjetas de sonido:
 - ★ Creative SB Live! (PCI)
 - ★ Creative SB X-Fi PCI



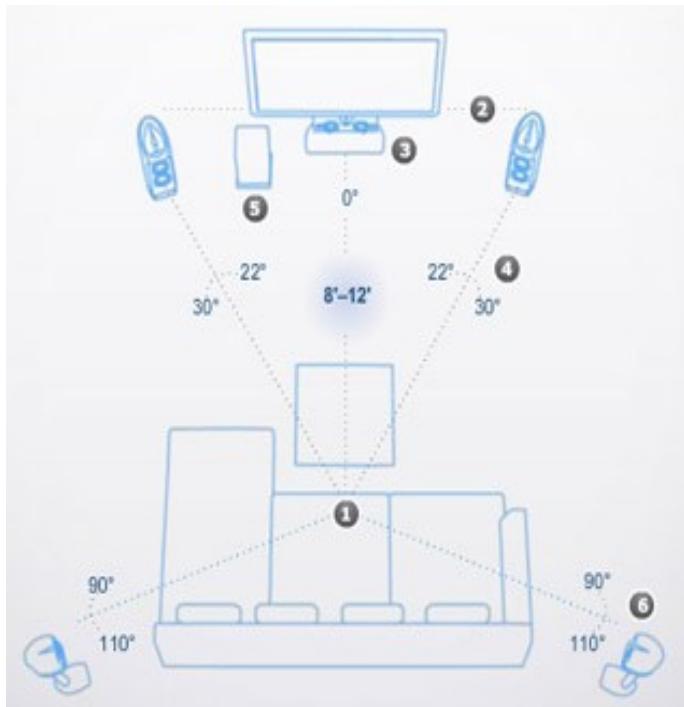
Adquisición y reproducción (VIII)

- Tarjetas de sonido



Adquisición y reproducción (IX)

- Sistemas de salida de audio
 - * Altavoces y auriculares multicanal

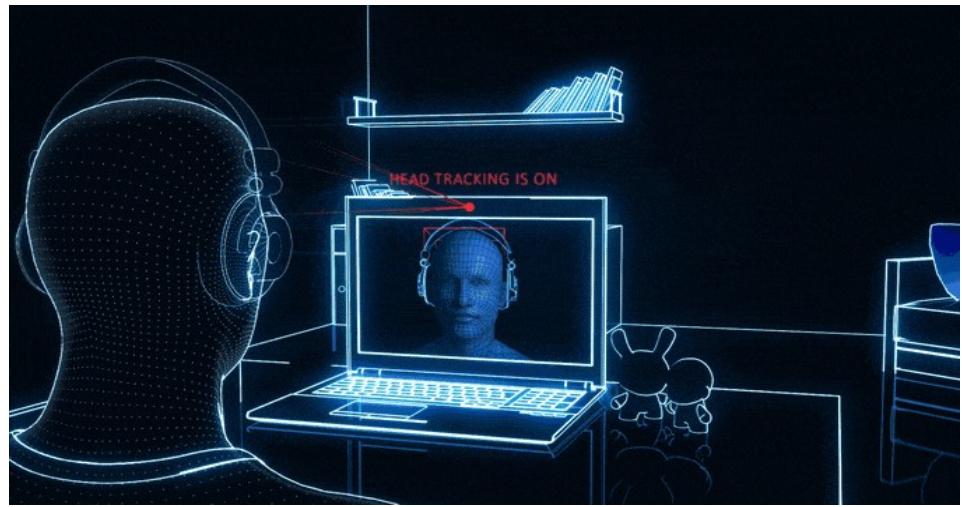


* Sistemas inmersivos



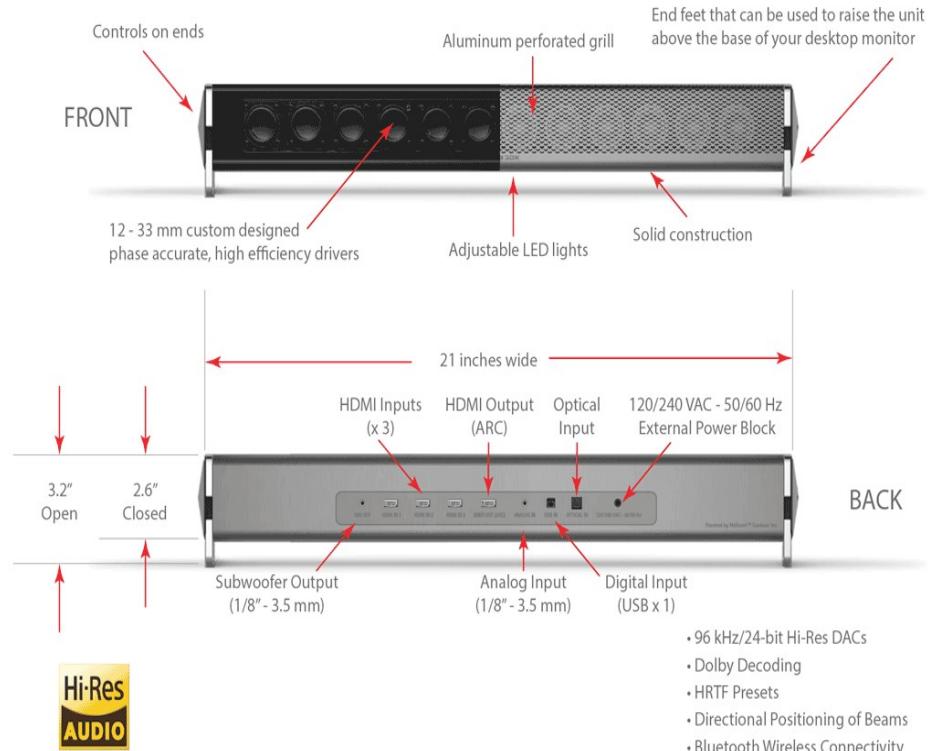
Adquisición y reproducción (X)

- Sistemas de salida de audio
 - ★ Sistemas inmersivos: auriculares
 - Seguimiento de la posición del usuario
 - App



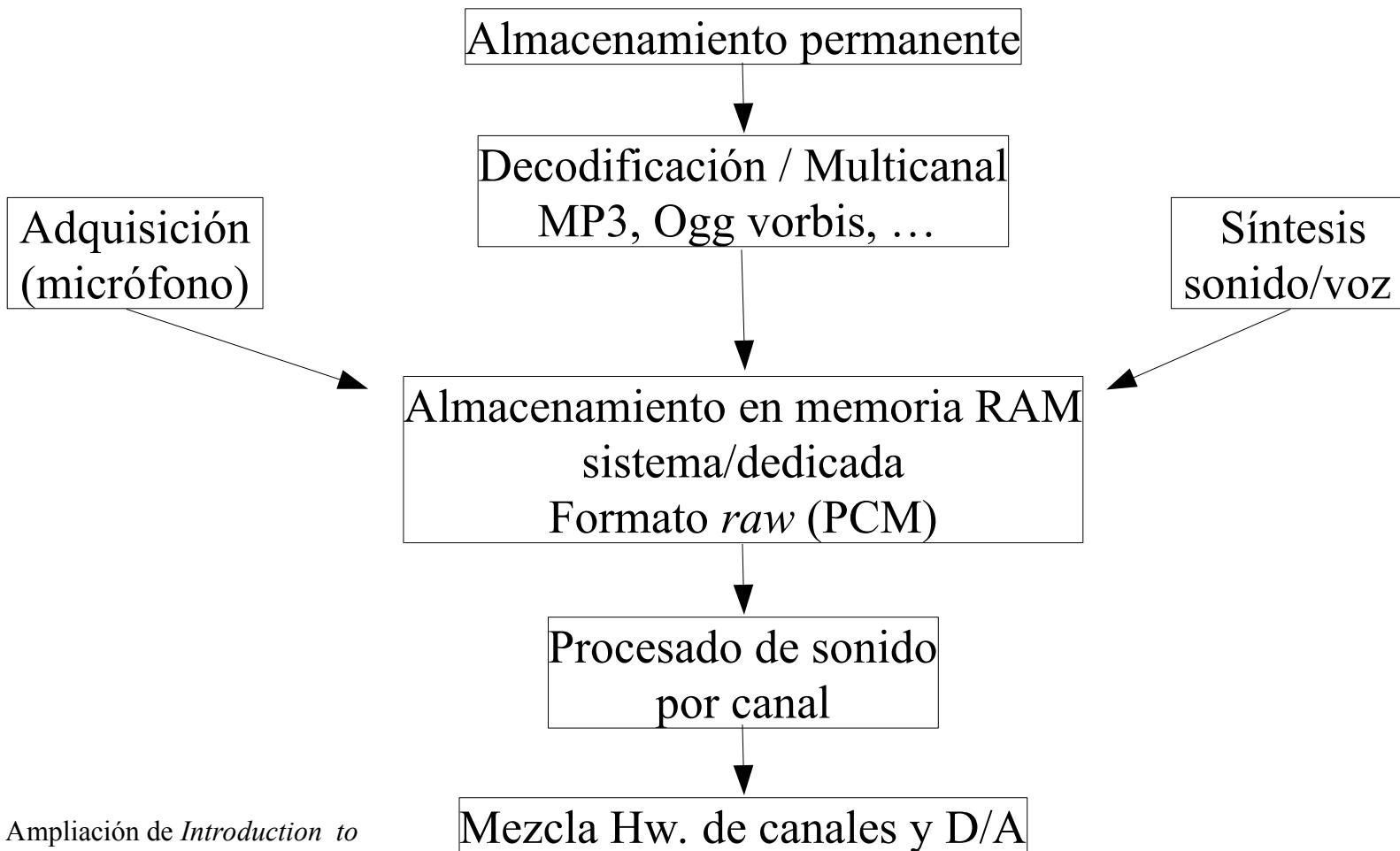
Adquisición y reproducción (XI)

- Sistemas de salida de audio
 - Sistemas inmersivos: altavoces
 - Ej. Yarra3D
 - Barra
 - *App*



Adquisición y reproducción (y XII)

- El sonido en un computador
 - ★ Perspectiva del desarrollador

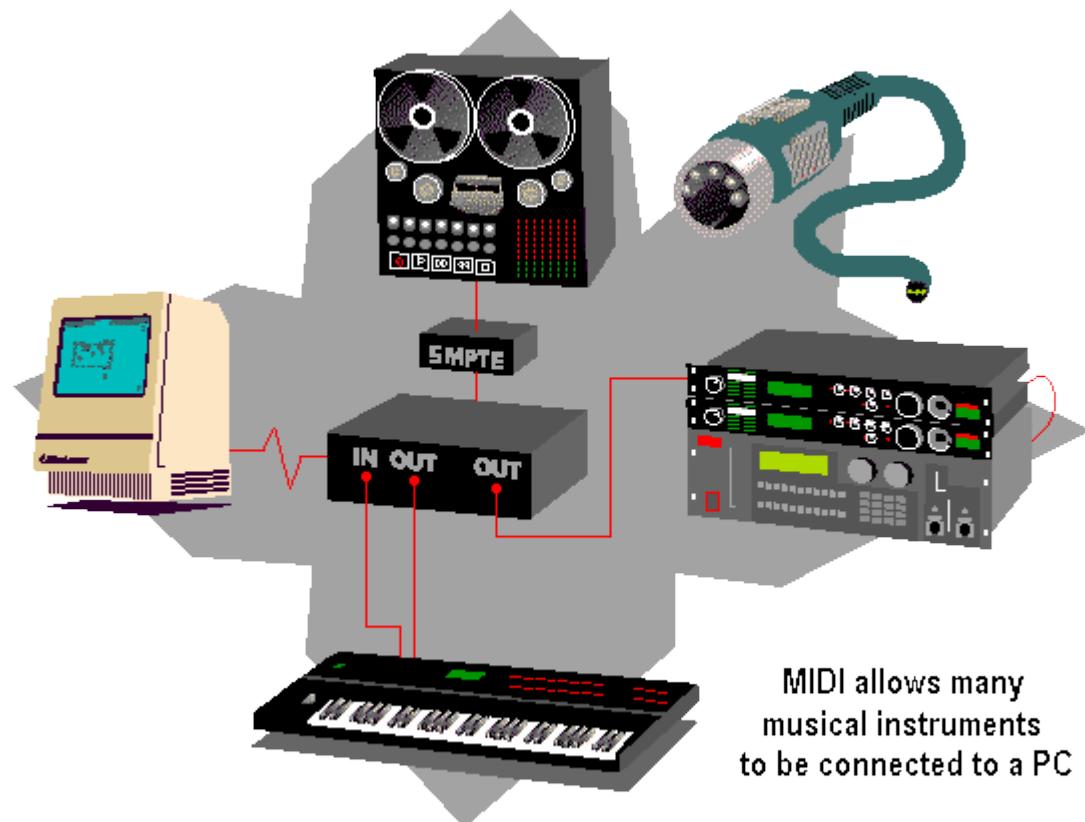


Índice

- Introducción
- Sonido y acústica
- Formatos de audio
 - MIDI
 - Forma de ondas
 - Compresión: MP3
- Herramientas software en aplicaciones de escritorio

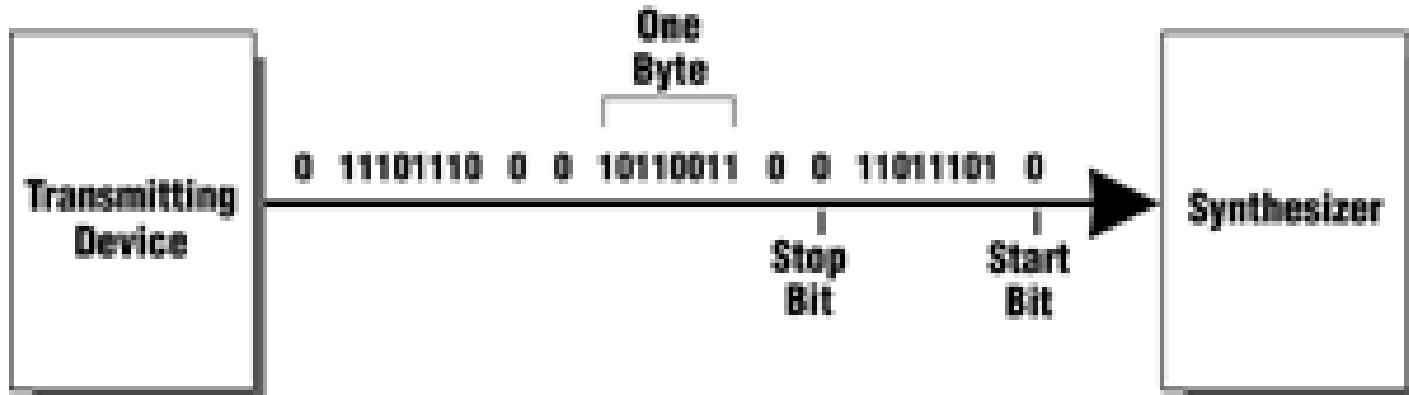
El estándar MIDI

- *Musical Instrument Digital Interface (1983)*
 - ★ Protocolo diseñado para la grabación y reproducción de música en sintetizadores digitales.
- Multiplataforma:
 - independiente de dispositivo
 - y
 - resolución



El estándar MIDI (2)

- Componentes Hard. y Soft.:
 - ★ Transmisión serie asíncrona; tramas de 10 bits; 31.25 Kbs; unidireccional. DIN-5



- ★ Instrucciones que indican a un dispositivo qué sonidos debe generar (\cong partitura y músicos).
- ★ ¿Unicamente sonido?
 - Nota, volumen, instrumento, fuerza (velocidad), ...

El estándar MIDI (y 3): descripción

- Descripción:

- ★ *General MIDI (GM)*

- Estándar para el esquema de numeración de los programas MIDI (sonidos que es capaz de generar un sintetizador): 1-128
 - *General Standard (GS)*
 - *Roland: GM + variaciones sonidos + efectos*

- ★ *Standard MIDI File,*

- ★ *Sample Dump Standard,*

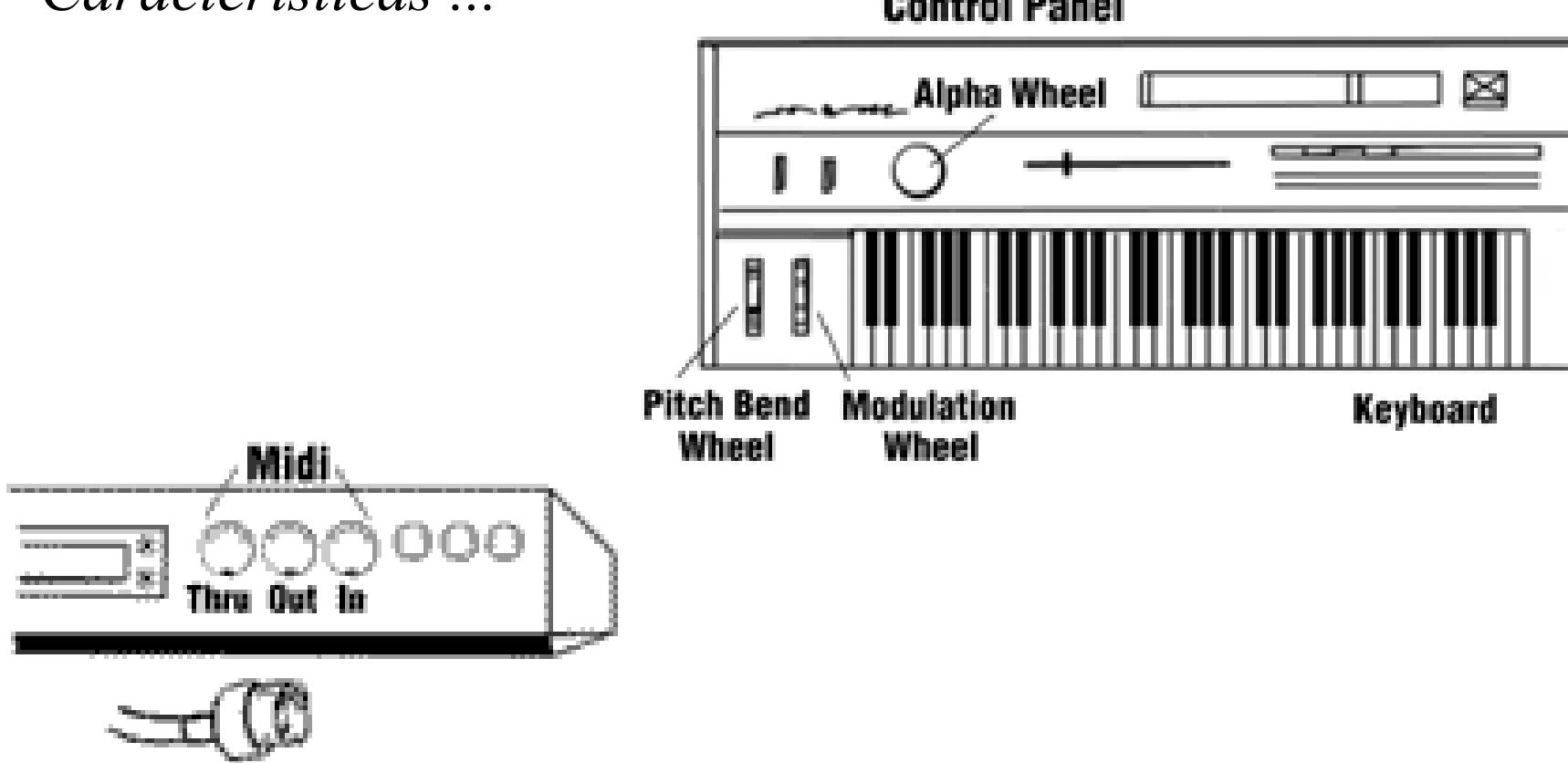
- ★ *MIDI Show Control*

(iluminación, proyectores, fuegos artificiales, telones, plataformas, etc)



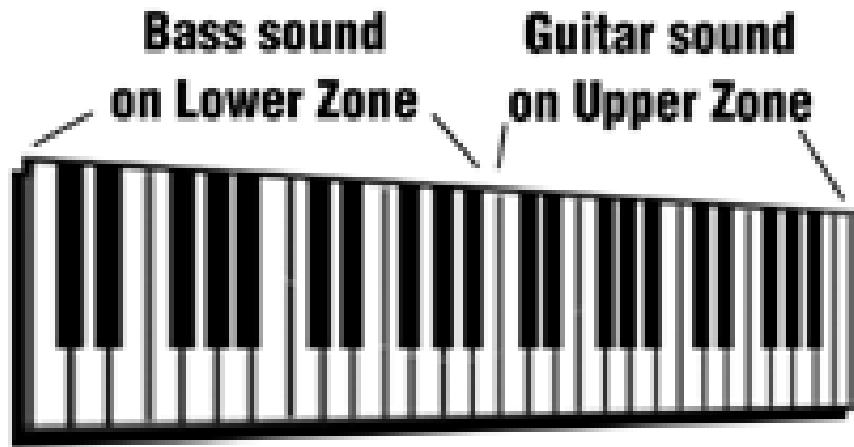
Dispositivos MIDI: Sintetizadores

- *Controllers & sound modules*
- *MIDI In, Out y Thru*
- *Características ...*



Dispositivos MIDI: Sintetizadores

- Características
 - ★ Sensibilidad,
 - ★ Polifonía,
 - ★ Multímbrico
 - ★ Modo de recepción y *keyboard splitting*



- ★ Programables: combinar sonidos (instrumentos), filtros, modificar sonidos (amplitud, frecuencia y timbre)

Otros dispositivos MIDI

- *Samplers*
 - ★ Dispositivos que permiten ... utilizando órdenes MIDI.
- *Guitar & Wind Controllers*
- *MIDI Drum machines*
- *Tambor MIDI
(Pere Pascual)*



Ride cymbal | Toms | Bass drum | Snare | Hi-hat | Crash



Formatos de formas de ondas

- 2 estilos
 - ★ con cabecera (frec. muestreo, nº de bits, estéreo/mono...)
 - ★ sólo datos (*raw*)
- Codificaciones lineales y logarítmicas:
ADPCM, μ -*law*, *A*-*law*
- *MOD*: *muestras cortas en PCM y partitura (interpretación musical)*

Formatos de almacenamiento

- Formatos sin compresión (o comprimidos sin pérdidas)

Extensión	Nombre	Origen	Comentarios
.au ó .snd	AU mu-law	NeXT, Sun	Frecuencia de muestreo variable. Soportado por todas las plataformas
.aif(f)	AIFF AIFC	Apple (Mac) SGI	Frec. de muestr., tamaño de la muestra y núm de canales variables. AIFC: comprimido.
.iff	IFF/8SVX	Amiga	Frec. de muestr. y núm de canales variables. Sólo 8 bits.
.voc	VOC	Soundblaster	Frec. de muestr. variable. Sólo 8 bits, 1 canal.
.wav	RIFF, WAVE	Microsoft	Frec. de muestr., tamaño de la muestra y núm de canales variables.
.mod .nst		Amiga	

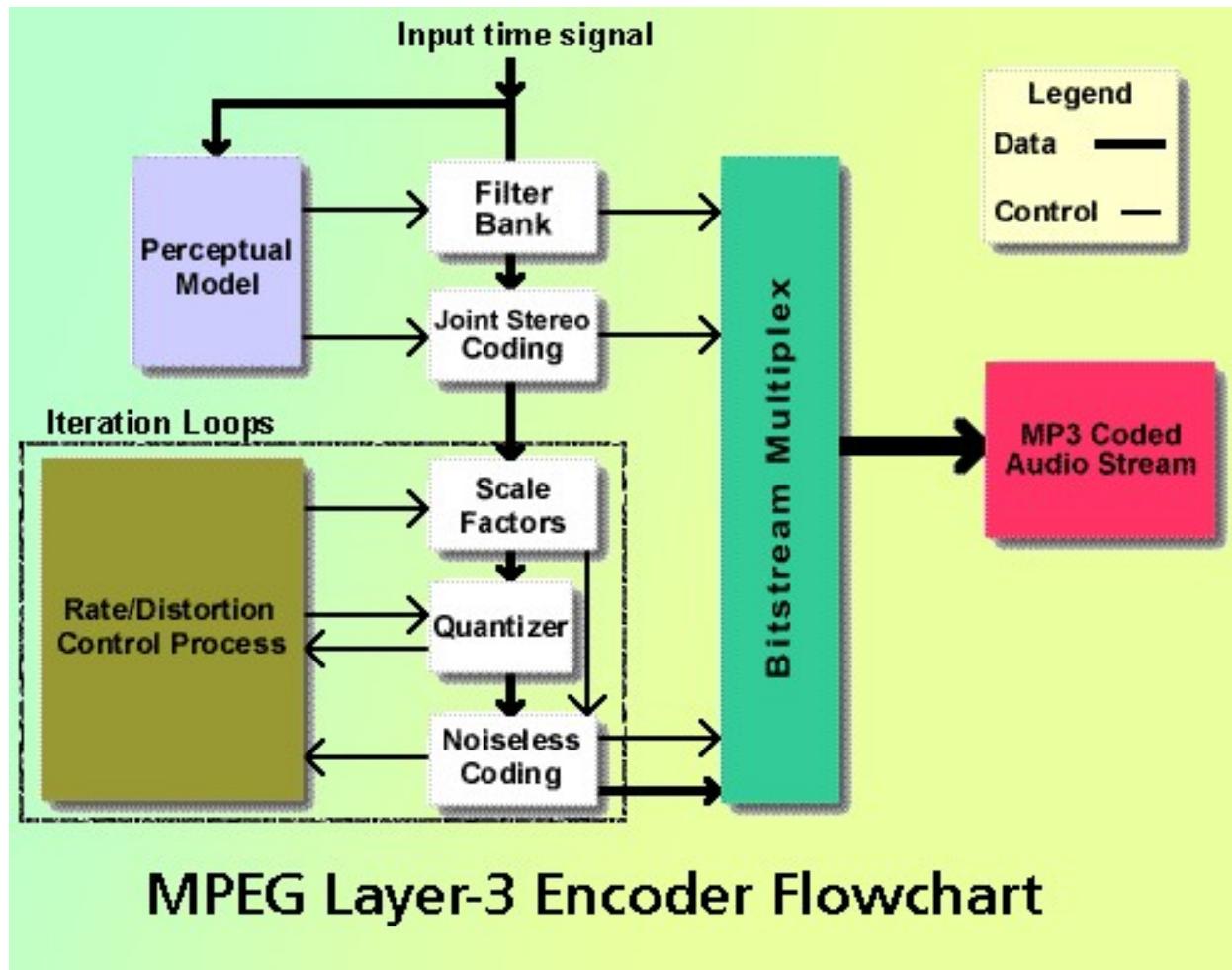
- FLAC (*Free Lossless Audio Compression*)

Formatos de almacenamiento (2)

- Formatos con compresión con pérdidas
 - ✖ **MP3** o *MPEG-1 Layer 3* (y otros que siguen el estándar MPEG): de 2:1 a 120:1, ideal para música con 128 kbps (.mp3)
 - ✖ *AAC or Advanced Audio Coding*: potenciado por Apple, es parte de las especificaciones de MPEG-2 y MPEG-4 (.aac, .m4a, .m4p)
 - ✖ *MP3-Pro*: mejora a altas frecuencias permitiendo bitrates de 64 kbps (.mp3, .mp3pro)
 - ✖ *Real Audio*: audio continuo de Real Networks, hasta 60:1, ideal para voz (.ra, .rm, or .ram)
 - ✖ *Windows Media Audio*: propio de Microsoft, puede codificar a 64kbps con calidad CD (.wma)
 - ✖ **Ogg Vorbis**: sin patentes, de características similares a MP3-Pro (.ogg)
 - ✖ **Opus**: *sin patentes, de características similares a Vorbis y AAC (.opus)*

MP3

- Compresión ← psicoacústica



MPEG Audio Layer-3

- Historia
 - ★ 1987: Fraunhofer Institute for Integrated Circuits, Applied Electronics Center, IIS-A
 - ★ IIS → ISO-MPEG Audio Layer-3
 - IS 11172-3 y IS 13818-3
 - ★ Sin reducir la cantidad de información:
 - 1 seg. estéreo con calidad CD: 16 bits y 44.1 kHz
 - → 1400 Mbit ⇒ MP3 (factor 12)
 - ★ Bases de la teoría de la codificación de la Percepción del sonido



MPEG: *The big picture*

- ISO & IEC → 1988
Moving Picture Experts Group
 - ★ *steps*
 - *MPEG-1 “Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to 1.5 Mbit/s”*
 - *MPEG-2 “Generic Coding of Moving Pictures and Associated Audio Information” (MPEG-3, HDTV)*
 - *MPEG-4 “Coding of Audio-Visual Objects”*

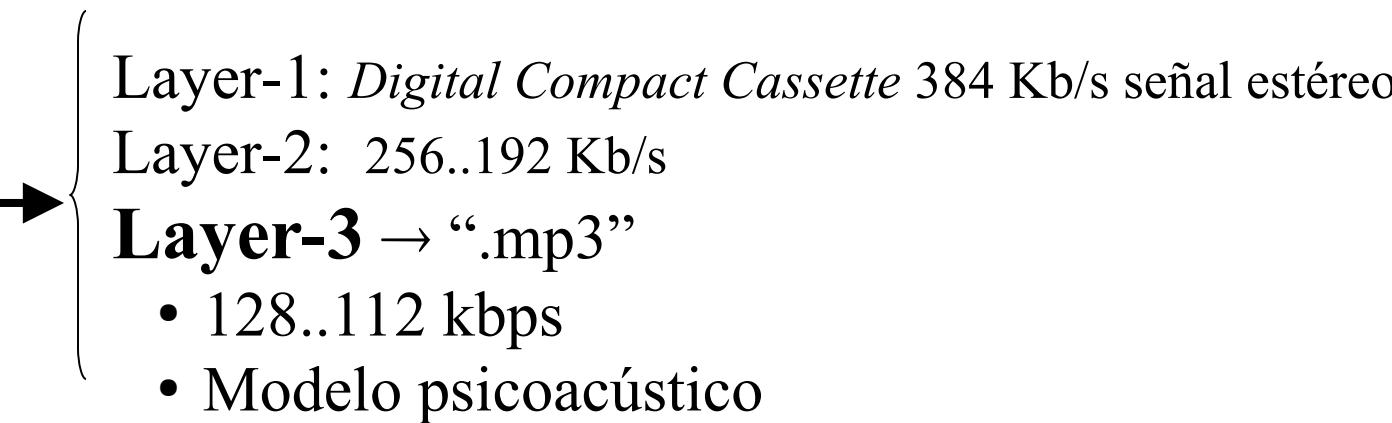
MPEG: *The big picture* (y 2)

- MPEG-1 (1992, ISO 11172):

★ partes

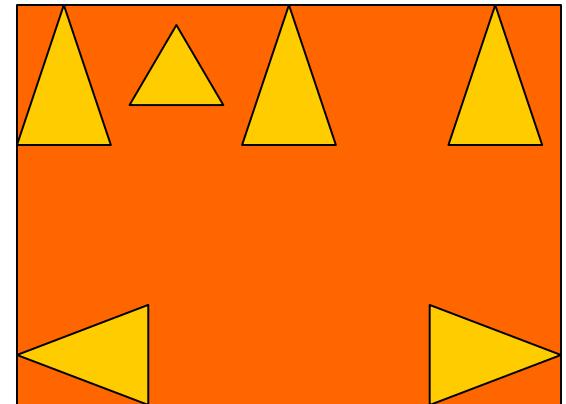
- IS-11172-1 *System*
- IS-11172-1 *Video*
- IS-11172-1 *Audio*
- IS-11172-1 *Compliance Testing*
- DTR-11172-1 *Software Simulation*

codecs

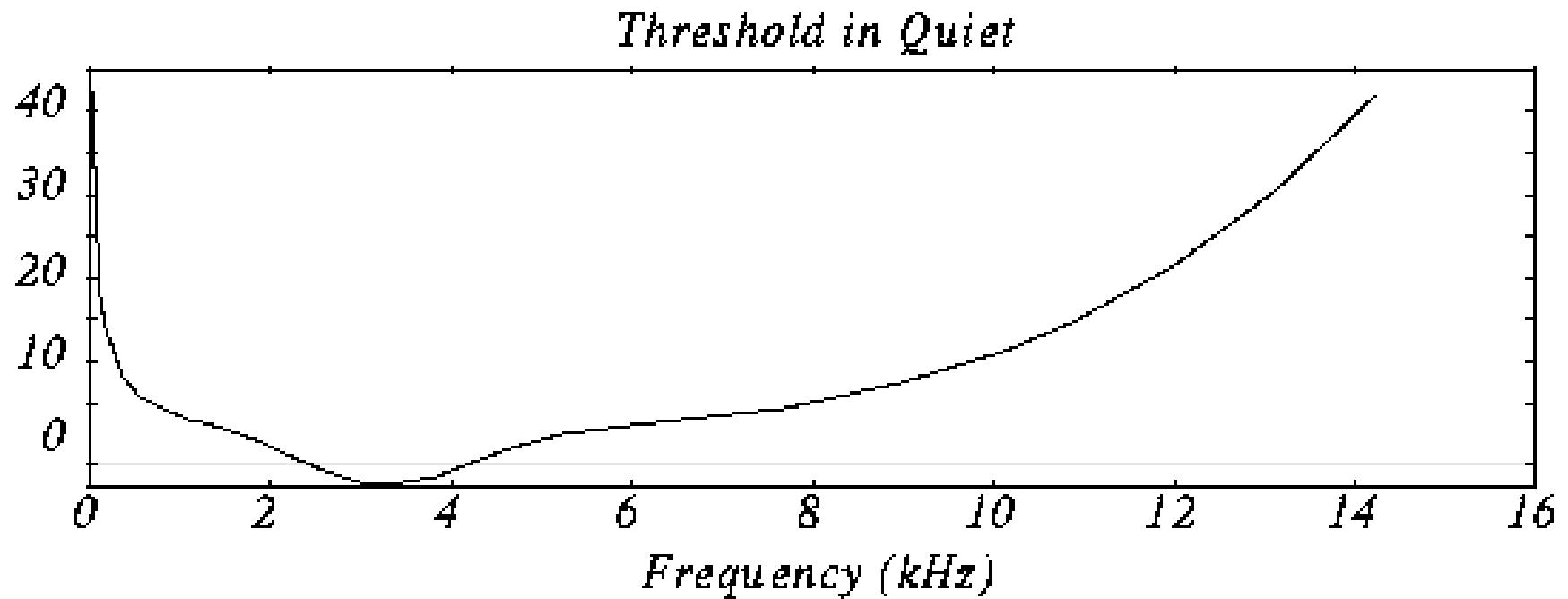


MPEG Audio

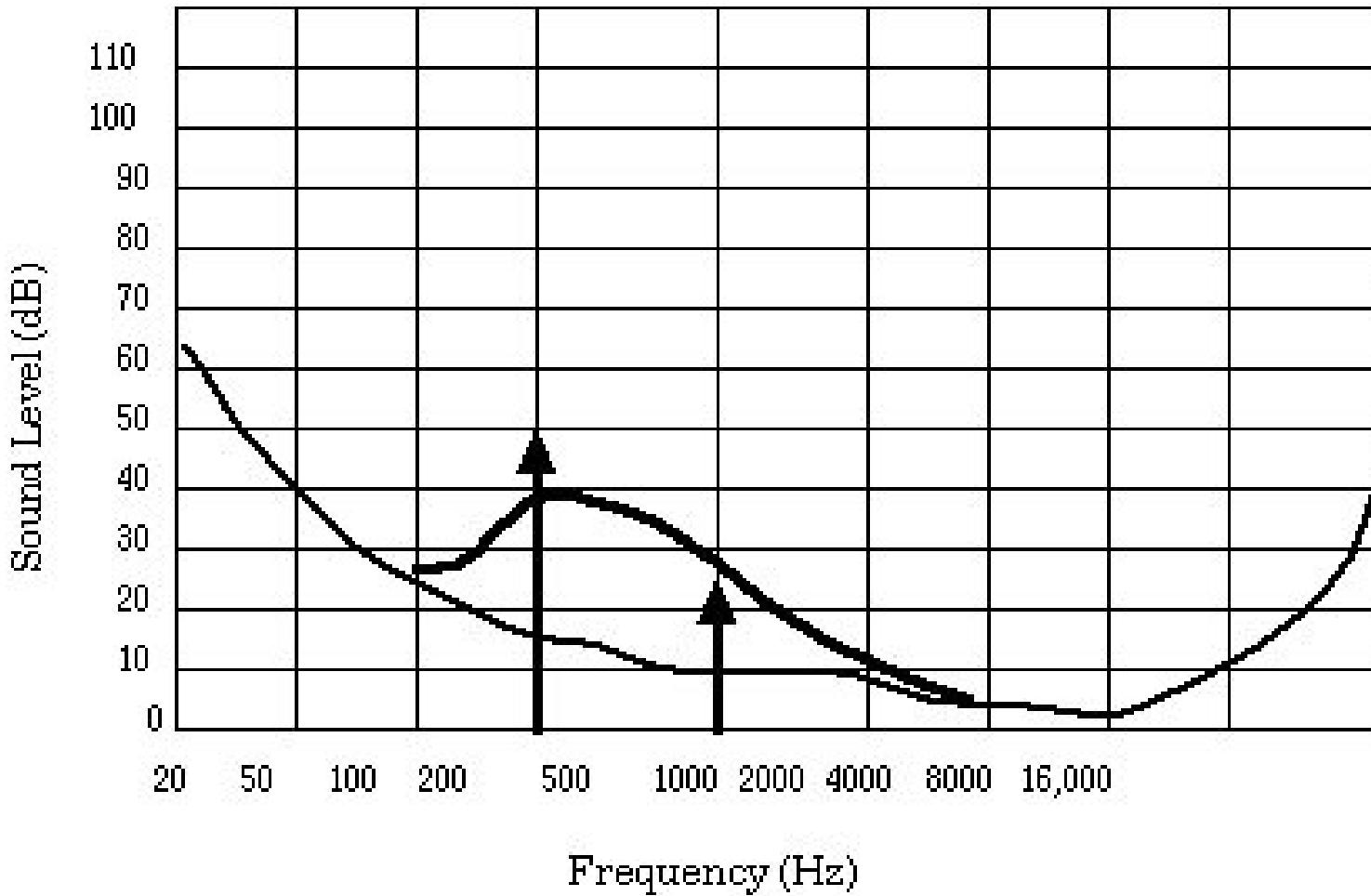
- MPEG - 1 Audio (1992)
 - ★ Vídeo y Audio en CD-ROM 1x
 - ★ Audio: mono, estéreo
 - Frecs. muestreo: 32, 44.1, 48 kHz.
- MPEG-2 Audio
 - ★ Sonido Multicanal: 5.1
 - ★ No es compatible hacia atrás
 - 1-48 canales;
 - 8..96k Hz freq. muestreo; ratios de 8..160 kbits/s
- MPEG-4 Audio
 - ★ Integra codificación natural y sintética
 - ★ Localización espacial (3D)
 - ★ Ratios estándares de 2..64 kbits/s



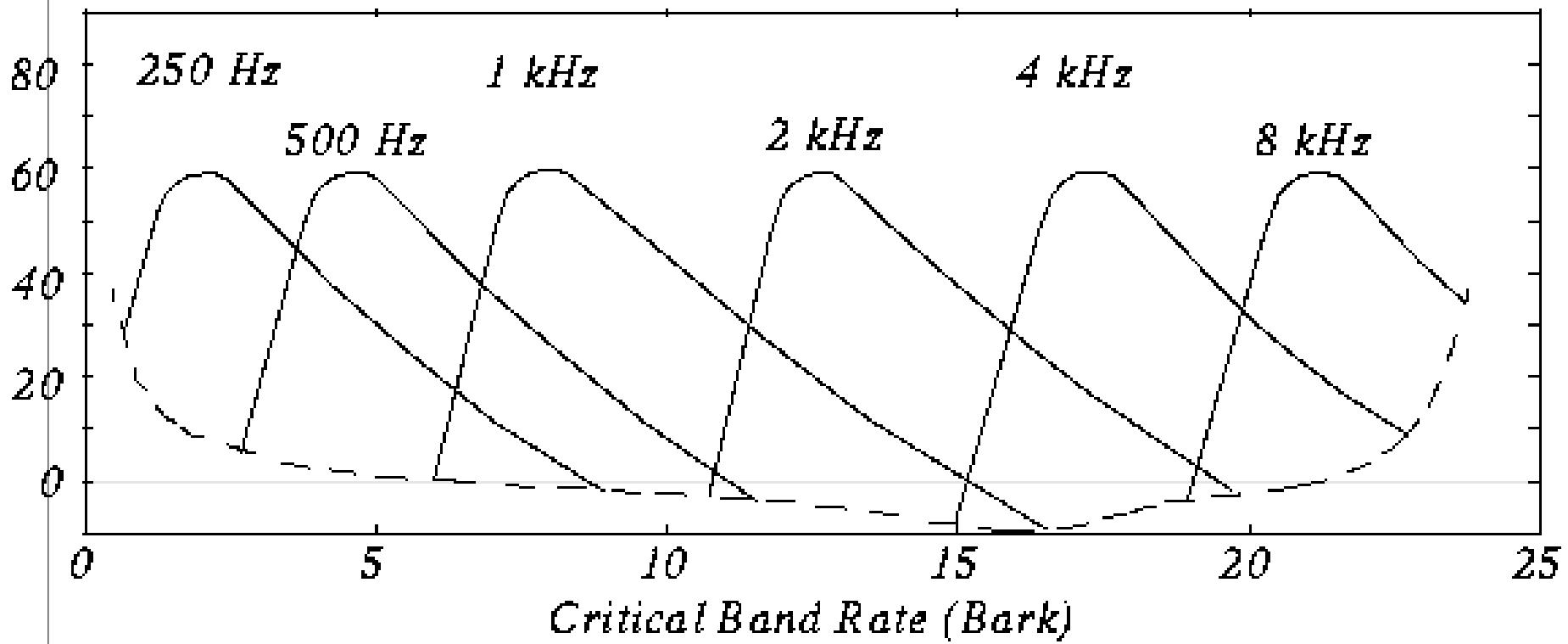
La curva de sensibilidad del oído



El fenómeno de enmascaramiento



Escala perceptual BARK

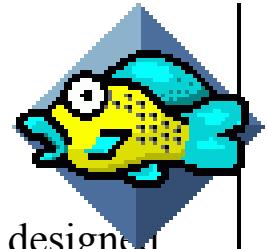


MPEG Audio Layer-3

- MPEG Layer-3
 - ★ Codificación perceptual de audio y compresión psicoacústica para eliminar la información superflua (las partes redundantes e irrelevantes de la señal de sonido).
 - ★ MDCT (*Modified Discrete Cosine Transform*)
 - ★ ¡Fin de las patentes!
 - Abril de 2017: "el programa de licencias de MP3 de Technicolor para ciertas patentes y software de Technicolor y Fraunhofer IIS relacionados con el MP3 ha expirado".



La alternativa libre



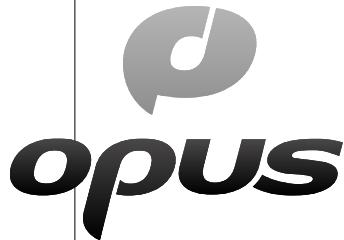
- *Xiph.Org Foundations*
 - ★ OGG Vorbis

- Ogg is a patent-free, fully open multimedia bitstream container format designed for efficient streaming and file compression (storage).
 - *Ogg Vorbis is a fully open, non-proprietary, patent-and-royalty-free, general-purpose compressed audio format for mid to high quality (8kHz-48.0kHz, 16+ bit, polyphonic) audio and music at fixed and variable bitrates from 16 to 128 kbps/channel. This places Vorbis in the same competitive class as audio representations such as MPEG-4 (AAC), and similar to, but higher performance than MPEG-1/2 audio layer 3, MPEG-4 audio (TwinVQ), WMA and PAC.*
 - *The bitstream format for Vorbis I was frozen Monday, May 8th 2000. All bitstreams encoded since will remain compatible with all future releases of Vorbis.*

★ Librerías

- *The libvorbis reference implementation provides both a standard encoder and decoder under a BSD license.*
 - *The Tremor reference decoder provides an integer-only implementation of the decoder for embedded devices*

La alternativa libre (y III)



- Opus Interactive Audio Codec (OPUS) - Xiph.Org Foundations
 - Opus is a totally open, royalty-free, highly versatile audio codec. Opus is unmatched for interactive speech and music transmission over the Internet, but is also intended for storage and streaming applications. It is standardized by the Internet Engineering Task Force (IETF) as RFC 6716 which incorporated technology from Skype's SILK codec and Xiph.Org's CELT codec.
 - Características
 - Opus can handle a wide range of audio applications, including **Voice over IP, videoconferencing, in-game chat, and even remote live music performances**. It can scale from low bitrate narrowband speech to very high quality stereo music. Supported features are:
 - **Bitrates from 6 kb/s to 510 kb/s. Sampling rates** from 8 kHz (narrowband) to 48 kHz (fullband). Frame sizes from 2.5 ms to 60 ms
 - Support for both constant bitrate (CBR) and variable bitrate (VBR). Audio bandwidth from narrowband to fullband. Support for speech and music
 - Support for mono and stereo. Support for **up to 255 channels** (multistream frames). Dynamically adjustable bitrate, audio bandwidth, and frame size. Good loss robustness and packet loss concealment (PLC)
 - Floating point and fixed-point **implementation**.

Índice

- Introducción
- Sonido y acústica
- Formatos de audio
- Herramientas software en aplicaciones de escritorio
 - Seminario de Audio 3D con OpenAL
 - Seminario de síntesis y reconocimiento de voz

Herramientas software

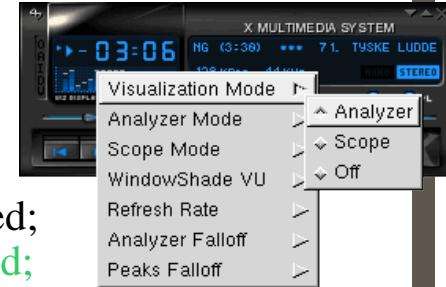
- *Tipos:*
 - ✖ *Componentes (plugins)*
 - ✖ *Librerías y SDK*
 - ◆ *De bajo nivel*
 - ◆ *Específicas de formatos*
 - ✖ *Motores de audio y SDK*
 - ✓ *Renderizar audio 3D*
 - ✓ *Procesar audio*
 - ✓ *De síntesis de voz*
 - ✓ *De reconocimiento de voz*

Herramientas software (II)

- *Plugin* de audio
 - ✗ (XMMS) www.xmms.org
 - ✗ Ejemplo (bibliografía de IMD)

```
typedef struct _VisPlugin {
```

```
    void *handle;
    char *filename;
    int xmms_session;
    char *description;
    int num_pcm_chs_wanted;
    int num_freq_chs_wanted;
```



```
    void (*init)(void);
    void (*cleanup)(void);
    void (*about)(void);
    void (*configure)(void);
    void (*disable_plugin)(struct _VisPlugin *);
    void (*playback_start)(void);
    void (*playback_stop)(void);
    void (*render_pcm)(gint16 pcm_data[2][512]);
    void (*render_freq)(gint16 freq_data[2][256]);
}
```

```
} VisPlugin;
```

Herramientas software (III)

- Librerías de bajo nivel:

- ★ **SDL** <<https://www.libsdl.org/>>



- *Simple DirectMedia Layer is a cross-platform development library designed to provide low level access to audio, keyboard, mouse, joystick, and graphics hardware via OpenGL and Direct3D.*

- ★ **libao** <<http://xiph.org/ao/>>

- Libao is a cross-platform audio library that allows programs to output audio using a simple API on a wide variety of platforms. It currently supports: NULL, WAVE, RAW, OSS, ALSA, PulseAudio, Mac OS X, Windows >= 98, NAS,

- ★ **libsnd** <<http://www.mega-nerd.com/libsndfile/>>



- *Lbsndfile is a C library for reading and writing files containing sampled sound through one standard library interface.*

- ★ **SFML** <<http://www.sfml-dev.org/>>

- ★ **PortAudio** <<http://www.portaudio.com/>>



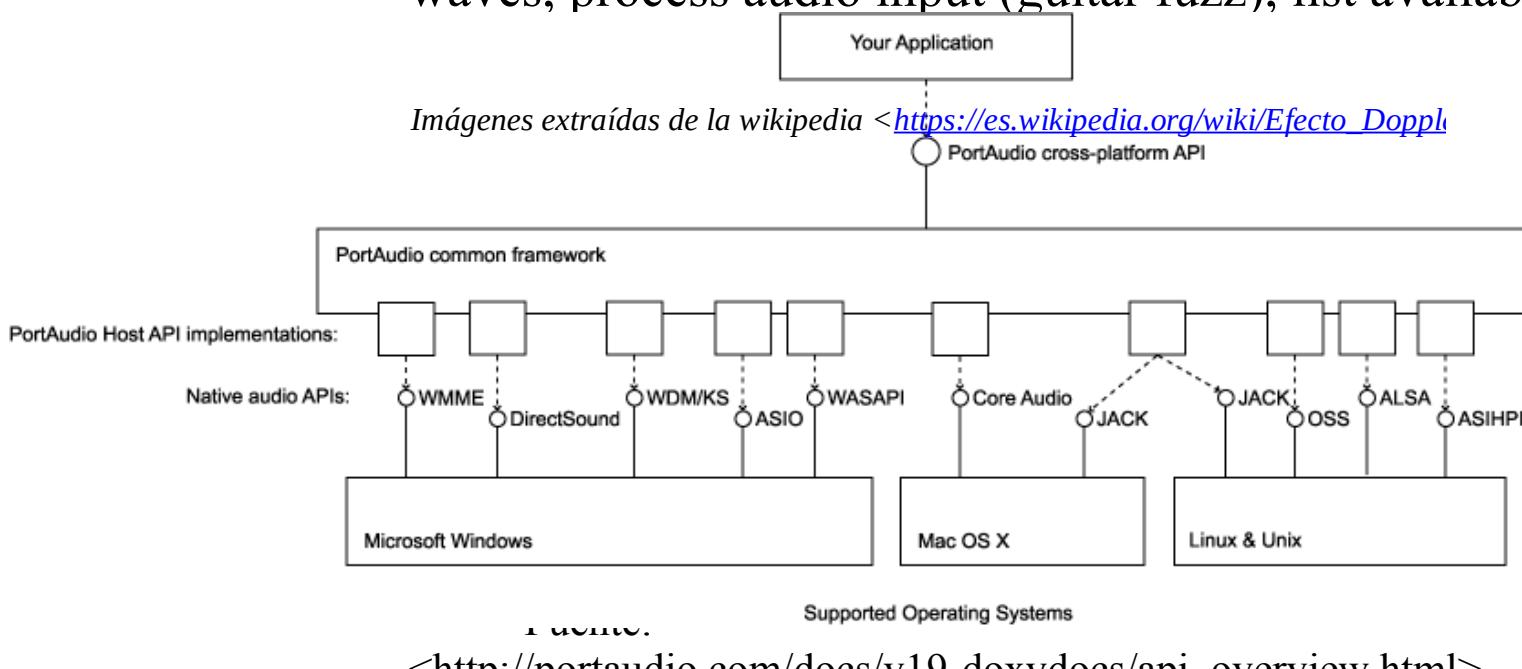
, ...

Herramientas software (IV)

- Librerías:

- ★ PortAudio <<http://www.portaudio.com/>>

- *PortAudio is a free, cross-platform, open-source, audio I/O library.*
 - ... recording and/or playing sound using a simple callback function or a blocking read/write interface, ... play sine waves, process audio input (guitar fuzz), list available



Herramientas software (VI)

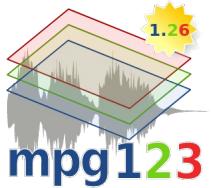
- Librerías específicas de formatos

- ★ Xiph.Org Foundation

- Vorbis /Tremor <<https://www.xiph.org/vorbis/>>
 - Opus
 - FLAC



- ★ Libmpg123 <<https://www.mpg123.de/api/>>



- Librerías de alto nivel:: motores

- ★ OpenAL
 - ★ FMOD
 - ★ Web Audio API (Web Speech API)

Herramientas software

- Estándar OpenAL <<http://www.openal.org>>
 - Historia:
 - *Loki Entertainment Software* --> 1.0 (~ 2000)

 - Creative Labs --> 1.1 (~ 2005)

 - *OpenAL Sample Implementation (the SI)*
 - *OpenAL.org* --> 1.1 (~ 2018)

 - Efficient
OpenAL renders 3D sound quickly.
Your buffer will be perfectly filled.
 - Simple
Ever written OpenGL code? You
already know how to work with
OpenAL.
 - Open API
Remember EAX and A3D? Forget
them. OpenAL is vendor neutral.

Herramientas software

- Librerías de alto nivel:: motores
 - API: *Open Audio Library*
 - Interfaz de programación multiplataforma para audio multicanal 3D
 - Objetos:
 - Los *buffers*: audio en formato PCM
 - 8 ó 16bits, 1 ó 2 canales.
 - Una fuente (*source*): datos de audio en una posición espacial
 - La velocidad, posición y dirección e intensidad del sonido.
 - El oyente (*listener*): donde "proyectar" la "escena"
 - La velocidad, posición y dirección del mismo, así como la ganancia asociada a todos los sonidos.
 - Aunque se pueden definir varios oyentes sólo uno puede estar activo.



Herramientas software (VIII)

- Librerías de alto nivel:: motores

- ★ FMOD <<http://www.fmod.org/>>

- FMOD Core: API

- *FMOD is the fastest, most powerful and easiest to use sound system on Windows, Linux, and Windows CE there is, and now Macintosh, GameCube, PS2 & Xbox!*
 - *FMOD supports 3d sound, midi, mods, mp3, ogg vorbis, wma, aiff, recording, obstruction/occlusion, cd playback (analog or digital), cd ripping, mmx, internet streaming, dsp effects, spectrum analysis, user created samples and streams, synchronization support, ASIO, EAX 2&3, C/C++/VB/Delphi and more.*
 - *General sound interface*

- FMOD Studio

- *Visual interface*



Ejemplo de uso de FMOD Core

```
//Play Sound Example Copyright (c), Firelight Technologies Pty, Ltd 2004-2019. (playsound.cpp)
#include "fmod.hpp"
#include "common.h"

int FMOD_Main() {
    FMOD::System *system; FMOD::Sound *sound1, *sound2, *sound3;
    FMOD::Channel *channel = 0; FMOD_RESULT result;
    unsigned int version; void *extradriverdata = 0;

    Common_Init(&extradriverdata);
    // Create a System object and initialize
    result = FMOD::System_Create(&system);     ERRCHECK(result);
    result = system->getVersion(&version);      ERRCHECK(result);
    if (version < FMOD_VERSION) {
        Common_Fatal("FMOD lib version %08x doesn't match header version %08x", version, FMOD_VERSION); }

    result = system->init(32, FMOD_INIT_NORMAL, extradriverdata); ERRCHECK(result);
    result = system->createSound(Common_MediaPath("drumloop.wav"), FMOD_DEFAULT, 0, &sound1); ERRCHECK(result);
    result = sound1->setMode(FMOD_LOOP_OFF);
    //drumloop.wav has embedded loop points which automatically makes looping turn on,
    // so turn it off here. We could have also just put FMOD_LOOP_OFF in the above CreateSound call.
    ERRCHECK(result);

    result = system->createSound(Common_MediaPath("jaguar.wav"), FMOD_DEFAULT, 0, &sound2); ERRCHECK(result);
    result = system->createSound(Common_MediaPath("swish.wav"), FMOD_DEFAULT, 0, &sound3);   ERRCHECK(result);

    // Main loop
    do {
        Common_Update();
        if (Common.BtnPress(BTN_ACTION1)) {
            result = system->playSound(sound1, 0, false, &channel); ERRCHECK(result);
            if (Common.BtnPress(BTN_ACTION2)){
                result = system->playSound(sound2, 0, false, &channel); ERRCHECK(result);
                if (Common.BtnPress(BTN_ACTION3)) {
                    result = system->playSound(sound3, 0, false, &channel); ERRCHECK(result);}
            }
        }
    } while (true);
}
```

Ejemplo de uso de FMOD Core (y II)

```
...
result = system->update();ERRCHECK(result);
{ unsigned int ms = 0, lenms = 0; bool playing = 0, paused = 0; int channelsplaying = 0;
if (channel) {
    FMOD::Sound *currentsound = 0;
    result = channel->isPlaying(&playing);
    if ((result != FMOD_OK) && (result != FMOD_ERR_INVALID_HANDLE) && (result != FMOD_ERR_CHANNEL_STOLEN))
    { ERRCHECK(result); }

    result = channel->getPaused(&paused);
    if ((result != FMOD_OK) && (result != FMOD_ERR_INVALID_HANDLE) && (result != FMOD_ERR_CHANNEL_STOLEN))
    { ERRCHECK(result); }

    result = channel->getPosition(&ms, FMOD_TIMEUNIT_MS);
    if ((result != FMOD_OK) && (result != FMOD_ERR_INVALID_HANDLE) && (result != FMOD_ERR_CHANNEL_STOLEN))
    { ERRCHECK(result); }
    channel->getCurrentSound(&currentsound);
    if (currentsound) {
        result = currentsound->getLength(&lenms, FMOD_TIMEUNIT_MS);
        if ((result != FMOD_OK) && (result != FMOD_ERR_INVALID_HANDLE) && (result != FMOD_ERR_CHANNEL_STOLEN))
        { ERRCHECK(result); }
    }
}
system->getChannelsPlaying(&channelsplaying, NULL);
[...]
}
Common_Sleep(50);
} while (!Common_BtnPress(BTN_QUIT));

// Shut down
result = sound1->release(); ERRCHECK(result);
result = sound2->release(); ERRCHECK(result);
result = sound3->release(); ERRCHECK(result);
result = system->close(); ERRCHECK(result);
result = system->release(); ERRCHECK(result);
Common_Close();
return 0;
}
```

Herramientas software (X)

- Librerías de alto nivel:: motores web

- ★ Web Audio

- *W3C Candidate Recommendation Snapshot, 14 January 2021 <<https://www.w3.org/TR/webaudio/>>*
 - *This specification describes a high-level Web API for processing and synthesizing audio in web applications.*



- ★ Web Speech API <<https://wicg.github.io/speech-api/>>



- *Web Platform Incubator Community Group. It is not a W3C Standard nor is it on the W3C Standards Track.*
 - *This specification defines a JavaScript API to enable web developers to incorporate speech recognition and synthesis into their web pages. It enables developers to use scripting to generate text-to-speech output and to use speech recognition as an input for forms, continuous dictation and control. The JavaScript API allows web pages to control activation and timing and to handle results and alternatives.*

Herramientas software (XI)

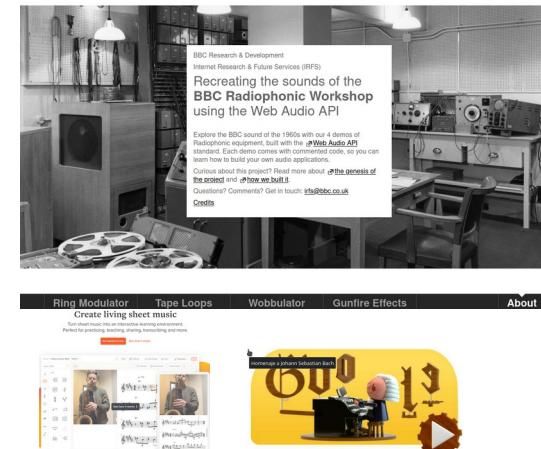
- Librerías de alto nivel:: motores web

- ★ Web Audio: demos

- “Web Audio Demos” <<https://webaudio.github.io/demo-list/>>
 - “Web Audio API: Advanced Sound for Games and Interactive Apps”. <<https://webaudioapi.com/samples/>>
 - “Web Audio Demos” <<https://webaudiodemos.appspot.com/>>
 - “Recreating the sounds of the BBC Radiophonic Workshop using the Web Audio API” <<https://webaudio.prototyping.bbc.co.uk/>>



Web Audio Demos			
Playground (a,i) The Web Audio Playground helps developers visualize how the graph nodes in the Web Audio API work. Shown at I/O 2012.	Vocoder (a,i,m) This complex audio processing app (shown at I/O 2012) implements a 28-band vocoder - a "robotic voice" processor.	"Analog" Synth (a,m) This application implements a polyphonic "analog" synthesizer, with a classic voice architecture. Playable via Web MIDI or onscreen keyboard.	Input Effects (a,i) This demo lets you play with few common effects on the audio inputs.
AudioRecorder (a,i) This is sample code for recording audio from live input and downloading them as WAV files, built on Matt Diamond's excellent Recorder.js .	Oscilloscope (a) Oscilloscope waveform display for Web Audio, and pulse-width duty cycle control on a square wave.	Monosynth (a,m) Example of a monophonic Web MIDI/Web Audio synth, with no UI.	wubwubwub (a,m) This application implements a dual DJ deck, specifically intended to be driven by a Numark MIDI controller.
Pitch Detection (i) This application performs phonic autocorrelating detection in realtime. This is suitable for a guitar tuner or other complex waveform source.	The Game of Life (m) Conway's Game of Life, on an 8x8 grid, interfacing with a Novation Launchpad controller via Web MIDI.	Happy Drums (a,m) Quick addition of MIDI controller support (using a Livid Instruments CNTRLR) to the Shiny Drum Machine .	TemplateSynth (a,m) Simple template for Web Audio synthesis, with a polyphonic voice architecture, on-screen keyboard (including touch), and Web MIDI.
TouchPad (a,t) Pointer/multi-touch input to control Web Audio. Load it on a touch device, touch and slide around - to open/close the filter and modify Q.	Volume Meter (i) This is a simple code example of how to properly implement a sample-playback volume meter. Eventually should support looping and voice control, but simple code example of how to properly implement a clip-detecting volume meter.	Sampler (a,m) This is an in-progress example of how to implement a sample-playback synthesizer.	Slide Deck Slide deck on Web Audio and Web MIDI.



Herramientas software (y XII)

- Librerías de alto nivel:: motores
 - ★ Web Speech API: demos

web-speech-api

A repository for demos illustrating features of the Web Speech API. See [Web_Speech_API](#) for more details.

Speech color changer demo

[Run recognition demo live](#)

Tap the screen then say a colour — the grammar string contains a large number of HTML keywords to choose from, although we've removed most of the multiple word colors to remove ambiguity. We did keep goldenrod, cos, well.

This current works only on Chrome/Chrome Mobile. To get this code running successfully, you'll need to run the code through a web server (localhost will work.)

Phrase matcher demo

Speak the phrase and then see if the recognition engine successfully recognises it — this is another demo that relies on speech recognition, written for a research team at the University of Nebraska at Kearney.

This current works only on Chrome/Chrome Mobile. To get this code running successfully, you'll need to run the code through a web server (localhost will work.)

[Run phrase matcher demo live](#)

Speak easy synthesis demo

[Run synthesis demo live](#)

Type words in the input then submit the form to hear it spoken. You can also select the different voices available on the system, and alter the rate and pitch.

This currently works in Chrome and Firefox.

Web Speech API Demonstration

Click on the microphone icon and begin speaking for as long as you like.



English ▾

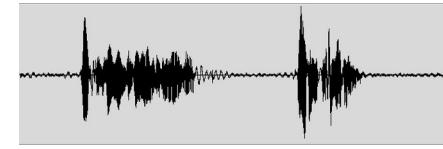
United States ▾

Seminario de OpenAL

- OpenAL
 - Estándard
 - API en 3 niveles: AL, ALC y ALUT
 - Documentación y extensiones
 - Conexión con otras librerías
 - Ejemplos de uso
 - Hello, World!
 - Generación de sonidos simples
 - Generación de sonidos desde fichero
 - Audio posicional
 - Efecto Doppler
 - Streaming

Herramientas software

- Síntesis y reconocimiento de voz
 - Festival
 - The Festival Speech Synthesis System
(Centre for Speech Technology Research)
 - CMU Sphinx
 - Carnegie Mellon University.
Speech at CMU <<http://www.speech.cs.cmu.edu/>>
 - Open Source Speech Recognition Toolkit
 - Speaker-independent large vocabulary
continuous speech recognizer with
industrial strength.



Seminario de síntesis de voz

- Festival
 - Introducción a Festival y a la síntesis de voz
 - Introducción práctica a la síntesis de voz con Festival
 - Modo interactivo
 - Dentro de una aplicación
 - Otras posibilidades de Festival
 - Sable
 - Flite

Bibliografía y enlaces

- *The Tapeless Studio: online magazine of audio recording on the PC.*
- *Short MPEG-2 Description*
(drogo.cselt.stet.it/ufv/leonardo/mpeg/standards/mpeg-2/mpeg-2.htm)
- *Frecuently Askef Questions about MPEG Audio Layer-3*
(www.iis.fhg.de/amm/techinf/layer3faq/index.html)
- *MPEG Audio FAQ's Version 8*
(drogo.cselt.stet.it/ufv/leonardo/mpeg/faq/faq-audio.htm)
- *MPEG Audio Web Page*
(www.tnt.uni-hannover.de/project/mpeg/audio)

Bibliografía y enlaces (2)

- **Introducción a la Codificación de Audio Digital (MPEG y DOLBY AC-3)**
Francisco García López (www.timagazine.net; Diciembre 1997)
- *Grupos de noticias*
 - ★ *rec.music.makers*
 - ★ *rec.audio.pro*
 - ★ *alt.music.midi*
 - ★ *alt.binaries.sounds.midi*
 - ★ *comp.music.midi*
- *Steve Rabin, 2010, Introduction to Game Development*
- *OpenAL* <<http://connect.creativelabs.com/openal/>>
- *Festival* <<http://www.cstr.ed.ac.uk/projects/festival/>>
- *CMU Sphinx* <<http://cmusphinx.sourceforge.net/>>

Bibliografía y enlaces (y 3)

- *M. Agustí. (2011). Introducción al procesado de audio mediante OpenAL.* <http://hdl.handle.net/10251/12694>.
- *M. Agustí. (2011). Introducción al empleo de técnicas de audio posicional mediante OpenAL.* <http://hdl.handle.net/10251/12697>.
- *M. Agustí. (2011). Efectos de audio básicos mediante OpenAL.* <http://hdl.handle.net/10251/12696>.
- *M. Agustí. (2012). Uso del micrófono para captura de audio en OpenAL.* <http://hdl.handle.net/10251/17547>.
- *M. Agustí. (2018). Extensiones para OpenAL: efectos ambientales.* <http://hdl.handle.net/10251/105664>
- *M. Agustí. (2018). Extendiendo OpenAL con OGG Vorbis.* <http://hdl.handle.net/10251/109210>.
- *M. Agustí. (2018). Extendiendo OpenAL con SDL. Caso de estudio MP3.* <http://hdl.handle.net/10251/105383>.
- *M. Agustí. (2018). OpenAL: efecto Doppler. Posicionamiento y velocidad del sonido.* <http://hdl.handle.net/10251/104052>.
- *M. Agustí. (2018). Reproducción de ficheros Opus con OpenAL: precarga vs "streaming".* <http://hdl.handle.net/10251/109211>.
- *M. Agustí. (2018). OpenAL y OpenGL: escuchar y ver el sonido.* <http://hdl.handle.net/10251/105550>.
- *M. Agustí. (2019). Introducción al uso del API de bajo nivel de FMOD.* <http://hdl.handle.net/10251/123301>.
- *Ivars Badía, A.; M. Agustí. (2011). Interacción con OpenCV: detección de movimiento para realizar un instrumento virtual con OpenCV + OpenAL.* <http://hdl.handle.net/10251/12684>.