

Audio digital

Javier Melenchón Maldonado
Francesc Tarrés Ruiz

PID_00186389



Los textos e imágenes publicados en esta obra están sujetos –excepto que se indique lo contrario– a una licencia de Reconocimiento-NoComercial-SinObraDerivada (BY-NC-ND) v.3.0 España de Creative Commons. Podéis copiarlos, distribuirlos y transmitirlos públicamente siempre que citéis el autor y la fuente (FUOC. Fundació para la Universitat Oberta de Catalunya), no hagáis de ellos un uso comercial y ni obra derivada. La licencia completa se puede consultar en <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.es>

Índice

1. Actividad 1. Representación en tiempo y frecuencia de señales de audio.....	7
1.1. Introducción	7
1.1.1. Contenido	7
1.1.2. Metodología	7
1.1.3. Recursos	7
1.2. Objetivos	8
1.3. Guía	9
1.3.1. Introducción	9
1.3.2. Forma de onda temporal y representación en frecuencia	9
1.3.3. Audacity: programa de captura, edición y procesado de audio	20
1.3.4. Audacity: captura de señales y manejo básico de pistas	23
1.3.5. Audacity: formas de onda básicas y representación tiempo frecuencia	28
1.3.6. Superposición de tonos	33
1.3.7. Secuenciación de tonos e interpretación del espectro ...	34
1.3.8. Espectrograma	36
1.3.9. Análisis de espectro de la señal de voz	37
1.3.10. Espectrograma de un silbido	40
1.3.11. Timbre de los instrumentos musicales	42
1.4. Problemas/preguntas	44
2. Actividad 2. Digitalización de señales de audio: muestreo.....	47
2.1. Introducción	47
2.1.1. Contenido	47
2.1.2. Metodología	48
2.1.3. Recursos	48
2.2. Objetivos	49
2.3. Guía	49
2.3.1. La retícula de muestreo	50
2.3.2. La frecuencia de muestreo	54
2.3.3. Análisis en frecuencia de panisAngelicus	56
2.3.4. Análisis en frecuencia de sintoniaOpera	57
2.3.5. Frecuencias de muestreo para señales de voz	58
2.3.6. Señales de banda limitada	59
2.3.7. Teorema del muestreo	60
2.3.8. Aliasing y filtros antialiasing	61
2.3.9. Algunos ejemplos de aliasing	63
2.3.10. El aliasing en vídeo y cine	67

2.3.11. Selección de la frecuencia de muestreo en sistemas de audio	69
2.3.12. Frecuencia de reproducción de la señal de audio	71
2.4. Problemas/preguntas	72
2.5. Evaluación	73
3. Actividad 3. Digitalización de señales de audio:	
cuantificación	76
3.1. Introducción	76
3.1.1. Contenido	76
3.1.2. Metodología	77
3.1.3. Recursos	77
3.2. Objetivos	78
3.3. Guía	80
3.3.1. La retícula de muestreo y el paso de cuantificación	81
3.3.2. Ruido de cuantificación y relación señal ruido	87
3.3.3. La relación señal ruido	89
3.3.4. Codificación PCM (<i>pulse code modulation</i>) de la señal de audio	92
3.3.5. La selección de las palabras código en PCM (opcional)	95
3.3.6. PCM en complemento a 2 (CA2) (opcional)	97
3.3.7. La tasa de bits y el almacenamiento de la codificación PCM	98
3.3.8. Niveles de cuantificación y sistemas de audio	100
3.3.9. Ventajas e inconvenientes de la digitalización de la señal de audio	103
3.4. Problemas/preguntas	105
3.5. Evaluación	107
4. Actividad 4. Filtrado de señales de audio	110
4.1. Introducción	110
4.1.1. Contenido	110
4.1.2. Metodología	111
4.1.3. Recursos	111
4.2. Objetivos	112
4.3. Guía	113
4.3.1. Ejemplos con filtrado paso bajo	116
4.3.2. Ejemplos con filtrado paso alto	121
4.3.3. Combinaciones de filtros en serie y paralelo	123
4.3.4. Filtros paso banda	124
4.3.5. Filtros de <i>shelving</i>	127
4.4. Problemas/preguntas	128
4.5. Evaluación	130
5. Actividad 5. Ecualización de señales de audio	132
5.1. Introducción	132

5.1.1.	Contenido	132
5.1.2.	Metodología	132
5.1.3.	Recursos	133
5.2.	Objetivos	134
5.3.	Guía	135
5.3.1.	Ecualizadores paramétricos	136
5.3.2.	El factor de calidad en los filtros paramétricos	142
5.3.3.	Aplicaciones de los ecualizadores paramétricos	143
5.3.4.	Ecualizadores gráficos	147
5.3.5.	Estructura de un ecualizador gráfico	149
5.4.	Problemas/preguntas	151
5.5.	Evaluación	153
Bibliografía	155

1. Actividad 1. Representación en tiempo y frecuencia de señales de audio

1.1. Introducción

1.1.1. Contenido

Esta actividad debe considerarse como una primera aproximación a las señales de audio digital y a su manejo mediante programas de captura y edición. El objetivo principal es que el estudiante se familiarice con los conceptos de forma de onda y espectro de la señal de audio, aprendiendo las relaciones que existen entre ellos. En la primera parte de la actividad se presentan señales musicales complejas y se examinan las formas de onda, aprendiendo a capturar las señales desde distintos recursos como el CD o el micrófono. Posteriormente, se trabaja con señales tonales, conceptualmente muy intuitivas, para introducir el concepto de análisis de espectro y representación en frecuencia de la señal. Una vez introducidos los conceptos básicos, se extiende para señales de voz e instrumentos musicales.

1.1.2. Metodología

Los conceptos se introducen con ejercicios que el estudiante realiza mediante una aplicación de captura y edición de audio. Los ejercicios son muy guiados y tienen por objetivo que se vayan comprendiendo los conceptos básicos de representación de señales en el dominio del tiempo y en el dominio de la frecuencia. Al final de la actividad se proponen una serie de problemas o ejercicios adicionales que consolidarán los aprendizajes de esta actividad.

1.1.3. Recursos

Los principales recursos utilizados en esta actividad son:

- Software de captura y procesado de audio Audacity (licencia libre).
- Simulador web audiovisual de tonos puros.
- Señales previamente capturadas y presentadas como materiales adicionales a la actividad.
- Propuesta de actividades de búsqueda de información en webs y recursos de Internet para profundizar en los conocimientos adquiridos en esta ac-

tividad (aplicaciones software de ecualización, fabricantes de ecualizadores, etc.).

- Actividades propuestas para verificar la adquisición de los conocimientos de esta actividad.

1.2. Objetivos

Los principales objetivos que se persiguen en esta actividad son:

- 1) Introducir el concepto de forma de onda y representación de la señal en el dominio del tiempo.
- 2) Aprender y familiarizarse con el manejo básico de programas de captura, edición, procesado y reproducción de señales de audio.
- 3) Identificar señales sinusoidales de diferentes frecuencias con el sonido que producen.
- 4) Mezclar diferentes señales sinusoidales interpretando su forma de onda y el sonido que producen.
- 5) Ver e interpretar la representación en frecuencia de señales sinusoidales puras.
- 6) Interpretar la representación en frecuencia de señales complejas.
- 7) Aprender a hacer la captura de señales de voz a partir de micrófonos integrados en el ordenador, analizando la forma de onda y su espectro.
- 8) Aprender a capturar señales musicales a partir de una fuente de CD.
- 9) Introducción al timbre de los instrumentos musicales.
- 10) Representación tiempo-frecuencia de las señales mediante el espectrograma.

Estos objetivos están relacionados con las siguientes **competencias de la asignatura**:

- 1) Capacidad de capturar audio e imágenes de forma eficiente y eficaz.
- 2) Capacidad de digitalizar eficiente y eficazmente contenidos de audio e imagen.
- 3) Capacidad de operar de forma digital con señales digitales de audio e imagen.

Y con las siguientes **competencias generales del grado**:

- 1) Ser capaz de analizar un problema en el nivel de abstracción adecuado a cada situación y aplicar habilidades y conocimientos adquiridos para abordarlo y resolverlo.
- 2) Capturar, almacenar y modificar información de audio, imagen y vídeo digitales aplicando principios y métodos de realización y composición del lenguaje audiovisual.

3) Atender adecuadamente consultas sobre proyectos, tecnologías y mercado de productos multimedia evaluando de forma precisa el entorno de aplicación, los recursos y las alternativas tecnológicas disponibles.

1.3. Guía

1.3.1. Introducción

La señal de audio es una parte muy importante de cualquier producción multimedia, por lo que es fundamental conocer los sistemas que permiten su captura, edición, procesado y producción. No obstante, para poder realizar todas estas operaciones es necesario comprender las características de las señales de audio y los métodos que existen para representar la información que contienen.

En esta actividad nos centraremos principalmente en estudiar dos métodos alternativos para representar las señales de audio: su **forma de onda temporal** y su **espectro**.

Veremos que estas dos formas de representación son complementarias y que podemos extraer información importante de cada una de ellas. También aprenderemos algunas opciones básicas de un programa de captura, edición y procesado de señales de audio. Esta aplicación software se utilizará en todas las actividades de audio de esta asignatura. Explicaremos los fundamentos y herramientas básicas de los programas de captura y edición de audio, centrándonos en concreto en la aplicación que vamos a utilizar durante la asignatura.

Iniciaremos la actividad con una serie de ejercicios exploratorios y, a continuación, continuaremos con la explicación y práctica con una aplicación de edición y captura de audio concreto.

1.3.2. Forma de onda temporal y representación en frecuencia

Los sonidos y las imágenes, tanto estáticas como dinámicas (imágenes en movimiento o vídeo), llegan a nuestros sentidos a través de medios físicos como el aire o la luz. El ser humano ha ideado diferentes alternativas a lo largo de su historia para almacenar la impresión que le producen estos estímulos, desde las primitivas pinturas rupestres hasta los últimos reproductores mp3, pasando por las fotografías o los discos de vinilo. Todas estas tecnologías tienen el objetivo de capturar diferentes series efímeras de estímulos para poder volverlos a percibir a voluntad.

Durante las últimas dos décadas, los estímulos sonoros y visuales están siendo capturados y almacenados en masa a través de las tecnologías de digitalización.

A lo largo de este apartado y de los siguientes empezaremos a estudiar con detalle este proceso de digitalización.

En este apartado empezaremos trabajando un concepto clave que interviene en el proceso de digitalización de contenidos sonoros y visuales y que a la vez forma parte de la naturaleza de los sonidos (y también de las imágenes, a pesar de que no lo trataremos en esta asignatura): el concepto de **frecuencia**. Dado que relacionaremos el sonido con su frecuencia representada sobre una pantalla o papel, también necesitaremos representar los diferentes sonidos de una forma visual; una de estas será la **forma de onda**. Para ayudarnos a desarrollar y trabajar estos conceptos realizaremos una serie de ejercicios exploratorios mediante una simulación web audiovisual. Si bien la explicación podrá seguirse a través de este texto escrito, recomendamos al estudiante que la trabaje a través de la web, puesto que le permitirá incorporar la información sonora a las explicaciones.

Para poder utilizar la simulación web con todas sus prestaciones sonoras es necesario disponer de la última versión del navegador *Firefox*. Se pueden utilizar otros navegadores, pero solo podrá experimentarse la parte visual de la simulación.

Ejercicio 1. Aspecto de un sonido sencillo

En este ejercicio aprenderemos a distinguir las diferentes representaciones de un sonido. Aseguraos de tener los altavoces con el volumen activado. Dada la naturaleza de los sonidos implicados, recomendamos no tener un volumen alto, especialmente si utilizamos auriculares.

- **Situad el cursor encima de la caja de más abajo (la titulada como Frecuencias).**

Simulación

- **Intentad mover el cursor dentro de la caja de Frecuencias y responded a las siguientes preguntas (sin leer más allá de las mismas):**

1) ¿Cómo oís el sonido al desplazar el cursor a la derecha? ¿Y si lo desplazáis a la izquierda?

2) ¿Os parece un sonido agradable o molesto? ¿Lo encontráis natural o artificial?

3) ¿Qué creéis que son las animaciones que salen en las cajas Forma de onda y Ondas de presión?

A continuación explicamos los elementos de la simulación, dando respuesta a las preguntas formuladas.

Respuesta

La caja de abajo del todo, donde movéis el ratón, indica la frecuencia seleccionada en el instante actual mediante una línea roja vertical y una caja de texto con un número seguido de las letras *Hz*. *Hz* es la abreviación de hercio, que es una unidad física y se utiliza para indicar de forma numérica una frecuencia; representa un número de vueltas, ciclos o repeticiones por segundo. Un número alto de Hz indica una frecuencia elevada y un número pequeño, una baja. En la simulación podéis ver que los valores están limitados entre 100Hz y 1000Hz, más que suficiente para los propósitos de la misma. Cuando seleccionáis un número elevado, moviendo el ratón a la derecha, percibiréis que el sonido es más agudo que cuando seleccionáis un número pequeño, moviendo el ratón a la izquierda.

La animación que sale bajo **Ondas de presión** es una representación de la densidad del aire por donde se desplaza el sonido. Por desgracia (o por suerte) la presión en el aire no tiene ninguna incidencia en los rayos de luz que la atraviesan y, por lo tanto, no la podemos percibir visualmente en la realidad. La representación dada es una propuesta de lo que veríamos si pudiéramos percibir visualmente los cambios de presión en el aire; hemos asignado el color cian a los puntos con presión máxima (más grande que la atmosférica) y el color blanco a los puntos de presión mínima (menor que la atmosférica). Observad que cuando un sonido es más agudo, es decir, su frecuencia es más elevada, el aire presenta variaciones de presión más rápidas y menos espaciadas. El caso contrario pasa con un sonido más grave (o menos agudo, según como se quiera decir). Estas ondas de presión, al llegar a nuestros oídos, son las responsables de que escuchemos los sonidos tal como los escuchamos. Hemos asociado la descripción de *agudo* a un sonido de elevada frecuencia, como, por ejemplo, hemos asociado la descripción de *negro* al color de las líneas que conforman los bordes de las dos cajas que hay en la simulación.

Sólo queda comentar el contenido de la caja titulada **Forma de onda**. Aquí podéis ver una línea con subidas y bajadas¹ siempre con la misma cadencia (que sólo cambia cuando movéis el ratón en la caja de **Frecuencias**) que se desplaza con el tiempo. Si os fijáis, podréis notar que siempre que en las **Ondas**

Heinrich Rudolf Hertz

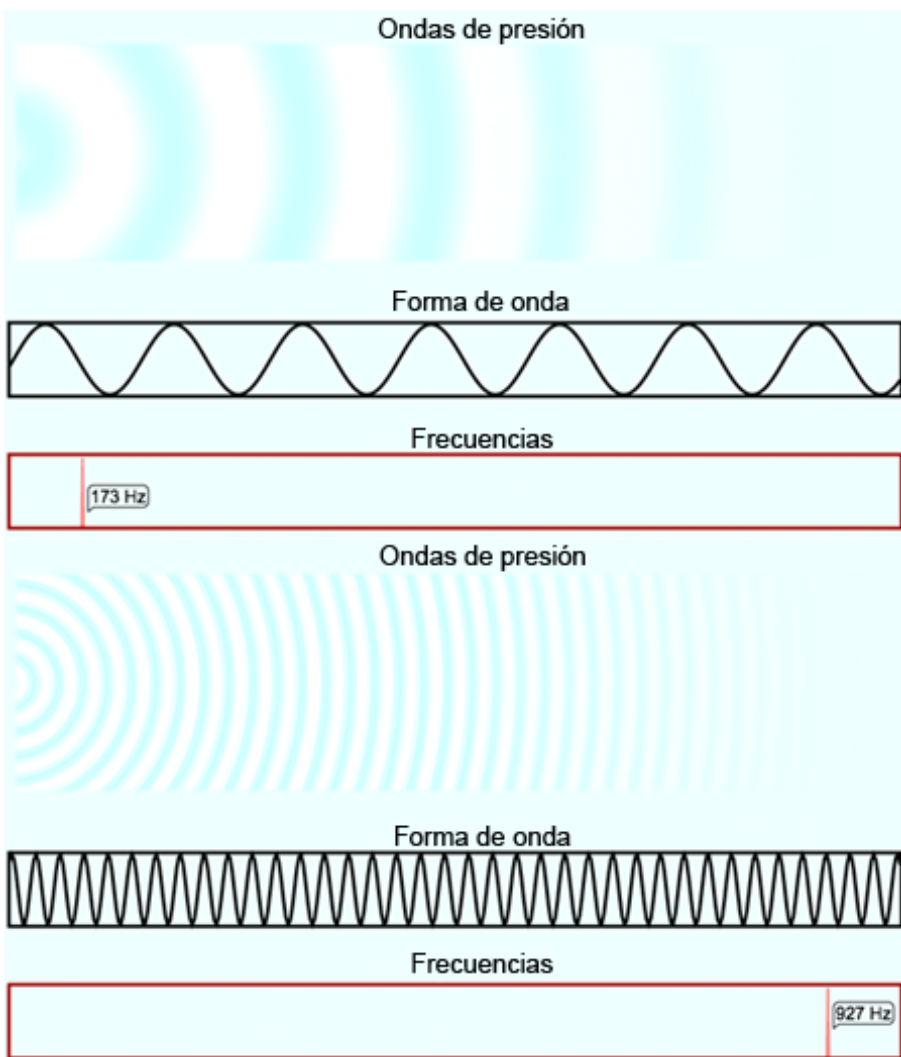
La unidad física de frecuencia hercio se puso en honor a Heinrich Rudolf Hertz, físico alemán descubridor de la propagación de las ondas electromagnéticas.

⁽¹⁾Una subida y bajada completas y seguidas conforma el que se denomina un *ciclo*.

de presión está el color azul, que indica una presión elevada, la línea de la forma de onda se sitúa en lo alto de la caja; al contrario, cuando las ondas de presión muestran un color blanco, que indica una presión baja, la línea de la forma de onda se sitúa en la parte baja de la caja (estas dos correspondencias son más fáciles de observar para frecuencias pequeñas). El resto de tonalidades de azul tienen también su correspondencia con la forma de onda, la cual se sitúa más arriba en tonalidades más cercanas al azul y más abajo en tonalidades menos saturadas y, por lo tanto, más cercanas al blanco. La representación en forma de onda de un sonido es una manera abstracta de representar con una línea² las variaciones de presión en el aire que conforman un sonido y, por lo tanto, el sonido en sí mismo.

⁽²⁾La línea representada por la forma de onda se puede dibujar como una serie de valores numéricos muy seguidos y unidos mediante segmentos muy cortos, dando el efecto de curva que sube y baja. Esta representación tiene unas limitaciones, como se verá más adelante cuando se hable del muestreo.

Figura 1. Aspecto visual mostrado por la simulación con una frecuencia pequeña (arriba) y una elevada (abajo).



Notar el aspecto que toma el sonido de las frecuencias seleccionadas, tanto en forma de ondas de presión en el aire como en la representación abstracta de forma de onda.

Los sonidos que reproduce la simulación y con los que habéis experimentado en este ejercicio reciben el nombre de **tonos puros**. Este tipo de tonos están formados por una única frecuencia (de aquí el calificativo de puros) y son inexistentes en la naturaleza por sí solos y de forma aislada. Habréis notado que son sonidos molestos, puesto que no estamos acostumbrados a percibirlos en la vida diaria, pero tienen gran resonancia dentro de nuestro aparato auditivo. Aun así, todos los sonidos se pueden obtener a partir de tonos puros (o se pueden descomponer en una agregación de tonos puros, depende de cómo se mire).

Ejercicio 2. Aspecto de dos tonos puros simultáneos

En este ejercicio exploraremos el aspecto visual y sonoro que tiene una agregación de diferentes tonos puros. La simulación está preparada para compensar el volumen de reproducción de diferentes tonos puros simultáneamente; aun así, al igual que en el ejercicio anterior, se recomienda partir de un volumen bajo de los altavoces o auriculares y subirlo si no fuera suficiente.

Situad el cursor en la caja de frecuencias y seleccionad un tono puro de una frecuencia 120Hz. Cuando la tengáis seleccionada haced clic con el ratón y a continuación moved el ratón fuera de la caja de frecuencias. Podréis observar que el tono escogido ha quedado fijado

Simulación

Moviendo ahora el ratón dentro de la zona de frecuencias elevadas (por encima de los 800Hz), responded a las siguientes preguntas:

- 1) ¿Qué observáis sobre la forma de onda y ondas de presión? Describid el sonido que oís.
- 2) Si movéis el ratón por debajo de los 800Hz y os vais acercando a la frecuencia seleccionada, ¿continuáis viendo y oyendo lo mismo? Si es que no, describid cuándo notáis un comportamiento diferente y cuál es.

Respuesta

Podréis apreciar que oís los dos tonos, el que habéis fijado a 120Hz y el que tenéis seleccionado al tener el ratón dentro de la caja y por encima de los 800Hz. Los dos tonos viajan juntos por el aire, a través de ondas de presión (y forma de onda) compuestas por los dos tonos. En la forma de onda se puede ver claramente una vibración rápida que va subiendo y bajando. Este aspecto se debe a que las dos formas de onda correspondientes a cada tono puro se han agregado en forma de suma. Los diferentes tonos puros y, en general, sonidos

Ved también

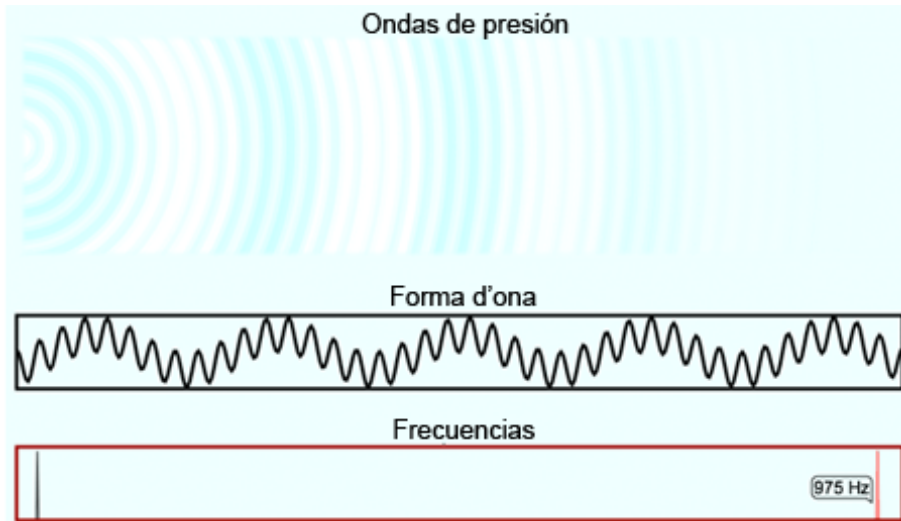
Una prueba de la creación de sonidos a partir de tonos puros la podemos observar en el ejercicio 4; de hecho, en la signatura de *Tratamiento y publicación de audio* se verá cómo crear sonidos a partir de estos elementos.

Observación

En la simulación web, si queréis volver a empezar podéis pulsar el botón de Reset, o el botón de Pausa para congelar la reproducción. Tened en cuenta que si volvéis a mover el ratón dentro de la caja de frecuencias se deshace la pausa (al igual que si volvéis a pulsar el botón de Pausa).

que viajan por el mismo medio (normalmente el aire, a través de ondas de presión) siempre se suman. Por lo tanto, lo que percibimos normalmente por el oído siempre es una suma de diferentes sonidos.

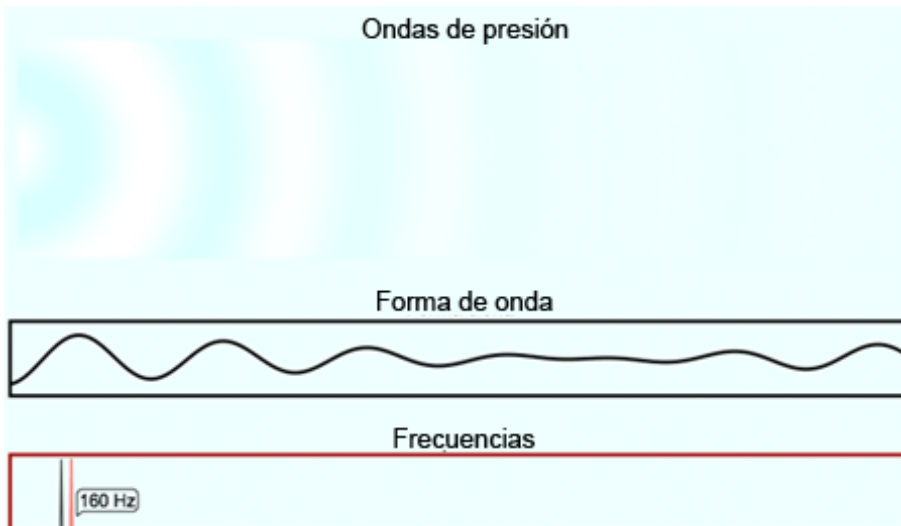
Figura 2. Visualización de dos tonos puros simultáneos de frecuencias muy diferentes



Si vamos reduciendo el valor de la frecuencia del tono seleccionado por el ratón (disminuyéndola progresivamente por debajo de 800Hz), podremos ver cómo la vibración rápida es cada vez más lenta, pero continúa subiendo y bajando al mismo ritmo, ya que la vibración más lenta (la de 120Hz) está fijada. Hay un punto alrededor de los 250Hz en el que visualmente no distinguimos, a partir de la forma de onda, una vibración rápida y otra más lenta; es más, hay un punto alrededor de los 140Hz en el que ya no podemos distinguir los dos tonos. Los dos están mezclados, pero nuestro oído presenta limitaciones físicas, y esta es una de ellas, no poder distinguir tonos de frecuencias cercanas cuando suenan simultáneamente.

Cuando dos tonos se sitúan muy cercanos, lo que se produce es la percepción de un solo tono puro con la presencia de incrementos y decrementos de volumen más o menos rápidos en función de lo cercano que se sitúan los dos tonos. Este tipo de efecto se denomina efecto psicoacústico y se verá con más detalle en próximas asignaturas. Visualmente se puede distinguir también al ver una vibración que aumenta y disminuye de volumen lentamente (más lentamente cuanto más cercanos son los dos tonos). Dado que la forma de onda se representa a una velocidad sensiblemente inferior a la del sonido (de lo contrario no se vería nada), hay que esperar un poco para poder observar este fenómeno sobre la forma de onda y ondas de presión.

Figura 3. Visualización de dos tonos puros muy similares



Se puede apreciar (de forma muy local) cómo la forma de onda aumenta y disminuye de volumen manteniendo una sola frecuencia de oscilación aparente.

La simulación permite añadir hasta 40 tonos puros. Podéis hacer pruebas y experimentar los efectos de la adición de múltiples tonos puros. En próximos ejercicios explotaremos esta característica (concretamente en el ejercicio 4).

Ejercicio 3. Amplitud y sonoridad

En este ejercicio exploraremos una característica más de una onda sonora, la amplitud y cómo afecta esta a la sonoridad.

Cread un tono puro de una frecuencia de 240Hz haciendo clic sobre la caja de Frecuencias y seleccionando la de 240Hz. A continuación, situaos sobre la frecuencia de 320Hz. Oiréis dos tonos diferenciados. Haced clic sobre la frecuencia de 320 y no mováis el ratón; veréis que el texto de la cajita donde se indica la frecuencia cambia a rojo³ y aparece "Amplitud: 1", esto significa que tenéis seleccionado este tono puro. Haciendo clic encima, aumentáis su amplitud. Veréis que el tono de 240Hz se hace más pequeño a medida que aumentáis la amplitud del de 320Hz; esto se debe a cuestiones de representación de la simulación, donde todo se escala según la amplitud mayor. También podéis mover la ruedecita del ratón hacia delante y hacia atrás para aumentar o disminuir la amplitud, respectivamente. Si se os ha movido el ratón deseleccionando el tono de 320Hz, solo tenís que volver a situarlo encima para poder continuar modificando su amplitud.

⁽³⁾Al situar el ratón sobre un tono existente, se selecciona y podemos variar su amplitud.

Simulación

Responded a las siguientes preguntas:

1) Sientes los dos tonos simultáneamente con cualquier cambio de amplitud en el tono de 320Hz? ¿Cuándo dejáis de oír uno u otro? Comentad también la forma de onda que veis al realizar cambios en la amplitud.

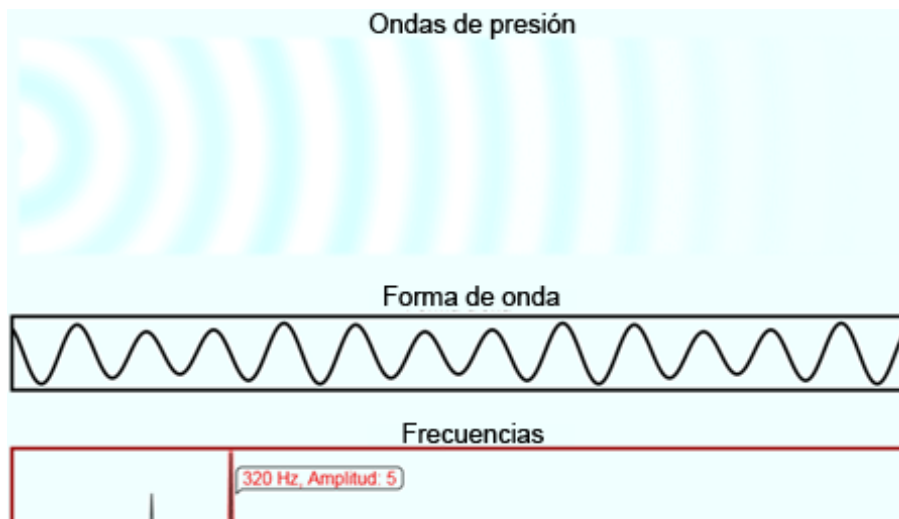
2) Volved a responder las preguntas anteriores (la 1) seleccionando un tono de 940Hz en lugar del de 320Hz (tendréis que pulsar Reset para poder borrar los tonos existentes).

Respuesta

Cuando aumentáis la amplitud del tono de 320Hz (o del de 940Hz) veréis que su forma de onda cobra mucha importancia hasta el punto de casi anular (comparativamente) la del tono de 240Hz; el resultado es una forma de onda que casi parece un solo tono puro a 320Hz (o 940Hz), a pesar de que no lo es nunca, puesto que el tono de 240Hz siempre está presente, aunque sea con poca importancia relativa.

Este hecho tiene impacto sobre nuestro oído. Podréis comprobar que con una amplitud de 4 (o 5, dependiendo de la persona y del volumen del altavoz) el tono de 320Hz anula la sonoridad del tono de 240Hz. No obstante, el tono de 940Hz necesita una amplitud mayor para hacer que dejemos de oír el de 240Hz. Cuanta más diferencia haya entre los dos tonos, más grande tiene que ser uno de los dos para anular acústicamente al otro. Este efecto se denomina enmascaramiento de frecuencias (sonoras) y es la base en la que se sustenta el mp3 para realizar sus compresiones de audio: si una frecuencia no se oye porque tiene una cerca mucho más fuerte, no se guarda la que no se oye (o se elimina), ahorrando espacio de almacenamiento.

Figura 4. Un tono de 240Hz a amplitud 1 no se oye cuando hay un tono de 320Hz a amplitud 5



La forma de onda del tono de 320Hz predomina sobre la de 240Hz, pero visualmente se detecta que no hay un solo tono puro, aunque solo oímos uno.

La amplitud tiene que ver con la sonoridad, a pesar de que no con una razón proporcional. Aumentar la amplitud al doble no hace que oigamos el sonido el doble de fuerte. Para que un sonido suene un punto más alto hay que multiplicar la amplitud por la raíz cuadrada de 10, o lo que es el mismo, por 3,1623 (aproximadamente). Así, si queremos que el tono puro de 320Hz suene el doble de fuerte que el de 240, hay que darle una amplitud de 3; si queremos que suene el triple de fuerte, hay que darle una amplitud de 10; para tenerlo el cuádruplo de fuerte, hace falta una amplitud de 31, y así sucesivamente. Esto se debe a que nuestro oído mide la sonoridad de forma logarítmica y nos permite aumentar mucho el rango de intensidades a oír, respecto a hacerlo con una regla proporcional. La representación visual elegida en la frecuencia, es decir, la altura de los picos, refleja la sonoridad en lugar de la amplitud.

Transformada de Fourier

Como curiosidad, la gráfica que vamos dibujando en la caja de frecuencias recibe el nombre de Transformada de Fourier de la forma de onda dibujada encima, a pesar de que en la aplicación se presenta de manera simplificada.

Ejercicio 4. Aspecto de un sonido complejo

En este ejercicio utilizaremos diferentes tonos puros para crear un sonido más complejo aprovechando las capacidades de la simulación web. Hay que aclarar que al tratarse de una simulación sencilla se obtendrá una aproximación de un sonido conocido, pero no una reproducción exacta.

Utilizando la simulación web, cread siete tonos puros con las siguientes componentes de frecuencia y amplitud:

	Tono puro 1	Tono puro 2	Tono puro 3	Tono puro 4	Tono puro 5	Tono puro 6	Tono puro 7
Frecuencia	130Hz	260Hz	390Hz	520Hz	650Hz	780Hz	910Hz
Amplitud	3	10	16	8	16	10	3

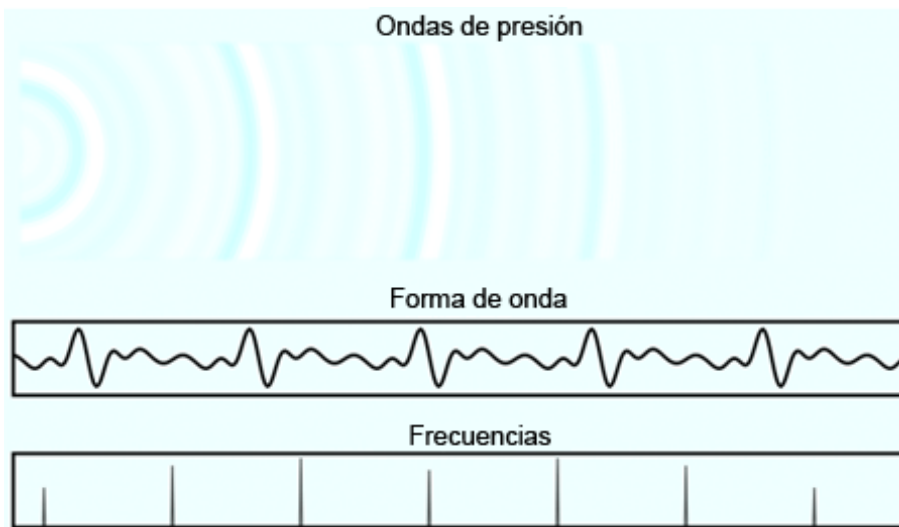
Simulación

Escuchad el sonido producido con atención. ¿Os recuerda algún tipo de sonido conocido? Aquellos con experiencia musical tendréis ventaja al contestar esta pregunta.

Respuesta

El sonido producido por los siete tonos puros propuestos es una aproximación de un Do2 producido por un fagot. No es un sonido exacto, puesto que se necesitarían muchos más tonos puros para recrearlo con precisión, pero podéis ver que con un conjunto limitado de los mismos se puede obtener una buena aproximación. El hecho de crear sonidos complejos a partir de tonos puros forma parte de la disciplina conocida como síntesis de sonido aditiva.

Figura 5. Aproximación del aspecto visual del sonido producido por un fagot tocando un Do₂ en forma de onda, ondas de presión y componentes frecuenciales



Ejercicio 5. Medidas de escala de la simulación

En este ejercicio obtendremos las medidas de escala de las ondas de presión de la simulación web. Al tratarse de una simulación, y como se ha podido experimentar en ejercicios anteriores, la forma de onda se mueve a una velocidad muy inferior a la que tendría en la realidad para poder apreciar visualmente su evolución.

Para hacer este ejercicio tenéis que saber que la ventana que muestra tanto la forma de onda como las ondas de presión va 50 veces más lenta que la velocidad real. Recordemos también que la velocidad del sonido en el aire es de 343 m/s (metros por segundo).

Antes que nada, determinad cuál sería la duración real de una onda que se desplaza por toda la caja de Forma de onda, es decir, ¿cuánto rato se visualizaría dentro de la caja de Forma de onda si la simulación la mostrara a tiempo real y no de forma ralentizada?

Simulación

Una vez conocéis el dato anterior, averigüad cuáles serían las medidas (anchura y altura) en metros, de la zona de las Ondas de presión, en caso de que tuviera medidas reales.

Respuesta

Para responder a la primera pregunta sólo necesitamos el primer dato y medir el tiempo en el que la simulación muestra una forma de onda internamente. Para hacerlo, necesitaremos un cronómetro. Obtendremos la medida experimentalmente y de forma aproximada. Si lo pudiéramos medir con plena exactitud, veríamos que las ondas tardan 2 segundos exactos en atravesar la caja

de forma de onda. Dado que esta velocidad mostrada es 50 veces inferior a la real, significa que en la realidad tardaría 50 veces menos, es decir, 2 segundos entre 50, que son 0,04 segundos. Las ondas atravesarían la caja de **Forma de onda** en 4 centésimas de segundo. Está claro que tenemos que ralentizarlo si queremos apreciar algo visualmente.

Para responder a la segunda pregunta, necesitamos el resultado de la primera y saber que el sonido se desplaza a 343m/s. Si de punta a punta de la caja de **Forma de onda** hay 0,04 segundos en tiempo real, en este tiempo, un punto de presión de aire "viaja" 13,72 metros ($0,04s \times 343m/s$). Así, la anchura de la región simulada de Ondas de presión ocuparía, si se dibujara a escala 1:1, una distancia de 13,72 metros.

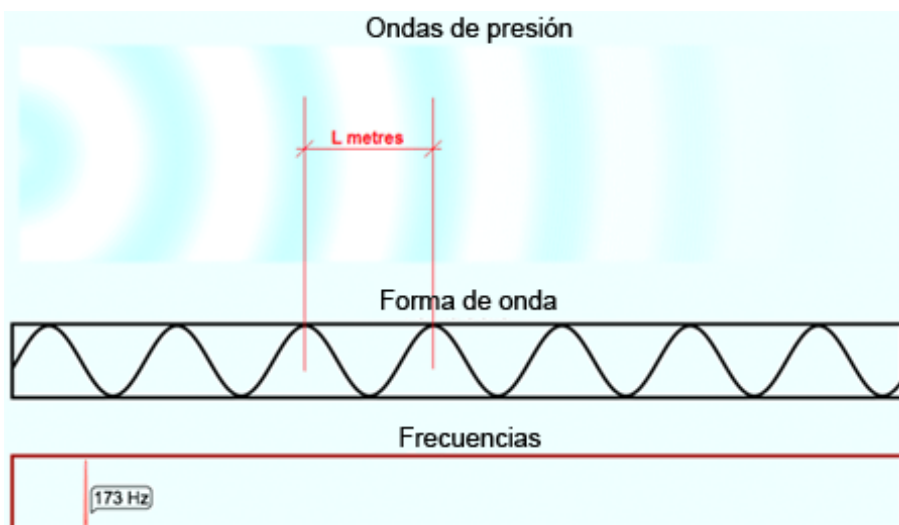
Ejercicio 6. Periodo y longitud de onda

En este ejercicio se identifican dos conceptos asociados directamente con la frecuencia: la longitud de onda y el periodo.

La **longitud de onda** se mide en metros y es la distancia que hay entre dos puntos de una onda que se repiten después de completar un ciclo; por ejemplo, el tiempo que pasa entre dos picos de un tono puro, como se puede ver en la figura 6.

El **periodo** se mide en segundos y es el tiempo que tarda en completarse una longitud de onda, o lo que es lo mismo, el tiempo que tarda en completarse un ciclo; en el caso de un tono puro, el tiempo que tarda en completar una oscilación, o el tiempo que tarda en acabar en el punto más alto desde el punto más alto inmediatamente anterior.

Figura 6. Longitud de onda de una onda sonora medida en metros



Conociendo la definición anterior y recordando que el sonido se desplaza a 343m/s, ¿cuál es la longitud de onda de un tono puro de 1000Hz? ¿Y la de uno de 100Hz? ¿Cuál es su periodo en segundos?

De hecho, si el oído humano capta entre 20Hz y 20000Hz, ¿cuál es la longitud de onda asociada a estos dos límites? ¿Y su periodo?

Respuesta

La longitud de onda de un tono puro de 1.000Hz es de 0,343 metros y la de un tono de 100Hz es de 3,43 metros. El periodo de un tono puro de 1.000Hz es de 0,001 segundos y el de un tono puro de 100Hz es de 0,01 segundos. La longitud de onda de un tono puro de frecuencia de 20Hz es de 17,15 metros y su periodo es de 0,05 segundos. La longitud de onda de un tono puro de frecuencia de 20.000Hz es de 0,01715 metros (o 17,15 milímetros) o tiene un periodo de 0,00005 segundos (o 0,05 milésimas de segundo).

1.3.3. Audacity: programa de captura, edición y procesado de audio

Hay un gran número de aplicaciones para la **captura, edición y procesado** de la señal de audio, entre las que podemos destacar:

- Acoustica,
- Amadeus Pro
- Creative WaveStudio
- GoldWave
- Ardour
- Qtractor
- SoundForge
- SoundBooth
- Wave Editor, etc.

Todas estas aplicaciones permiten la captura de varias pistas de audio procedentes de micrófonos, entradas de línea, CD, ficheros digitales, etc. Es posible cortar y pegar fragmentos, aplicar efectos independientes a cada una de las pistas y mezclar los resultados en una banda sonora final con múltiples pistas. Algunos de estos programas son de uso libre, aunque los más populares (WaveStudio, SoundBooth, SoundForge) sólo suelen proporcionarse con la compra de hardware de digitalización específico o mediante la compra con el distribuidor.

Sistemas operativos

Casi todos los programas anteriores tienen una versión para Windows y algunos disponen de versiones para Mac y/o LINUX.

El programa que vamos a utilizar para la edición y el registro de señales de audio en esta asignatura es **Audacity**, un software libre de código abierto, que lleva ya muchos años en el mercado y que actualmente puede competir en cuanto a funcionalidades y calidad del tratamiento con los productos comerciales más reconocidos.

Descarga

Audacity puede descargarse directamente desde <http://audacity.sourceforge.net>.

Durante la instalación podrá elegir el idioma con el que desea realizarla. Posiblemente se sentirá más cómodo en castellano, aunque si lo instala en inglés, le ayudará a familiarizarse con el tipo de vocabulario utilizado en sistemas de audio.

El **proceso de instalación** es muy simple:

- Para **Windows** debe descargarse el fichero `audacity-win-unicode-x.x.x.exe` (donde `x.x.x` representan la versión del programa) en un directorio desde el que desee realizar la instalación. Al ejecutar el programa le pide permiso para aceptar las condiciones de la licencia GNU y se instala por defecto en el directorio de Archivos de Programas, creando por defecto los iconos del programa en el menú inicio y en el escritorio.
- El procedimiento de instalación para **Mac OS X** y **LINUX** es todavía algo más fácil, ya que al descargar y descomprimir el programa de la aplicación se genera directamente un fichero ejecutable en el directorio de aplicaciones.

Cuando se **inicia el programa** Audacity, aparece una pantalla de ayuda que le permite realizar varios tutoriales para reproducir archivos, grabar voz, editar sonido, exportar e importar formatos de audio, etc. Si lo deseáis, podéis realizar alguno de estos tutoriales, aunque durante esta actividad explicaremos alguna de las funciones anteriores con detalle. Esta **pantalla de bienvenida** se muestra en la figura 1.

Sistemas operativos

Audacity dispone de versiones para los sistemas operativos más utilizados: Windows XP, Windows Vista, Mac OS X, GNU/LINUX.

Observación

Las figuras que se muestran pueden no coincidir exactamente con las de la versión de Audacity que estéis utilizando, el idioma en el que hayáis hecho la instalación o el sistema operativo. En general, las diferencias deberían ser mínimas y se debería poder seguir la actividad sin la menor dificultad. Las figuras que presentamos durante esta actividad proceden de la versión en castellano del sistema operativo Vista.

Figura 1. Pantalla de bienvenida del programa Audacity



Se recomienda realizar alguno de los tutoriales si se tienen dificultades con el manejo del programa y el seguimiento de esta actividad.

Si pulsamos el botón de Aceptar, aparece la **pantalla principal** del programa Audacity que se muestra en la figura 2.

Figura 2. Pantalla principal del programa Audacity



El programa dispone de varios **menús en la parte superior** de la ventana en los que encontramos las funciones habituales:

- **Menú Archivo:** dispone de opciones para crear nuevos archivos, abrir y guardar archivos, guardar proyectos y exportar e importar señales de audio en diferentes formatos.
- **Menú Editar:** permite realizar funciones de recorte y pega con las pistas de audio, dividir pistas, gestionar silencios, etc.
- **Menú Ver:** se utiliza principalmente para determinar la fracción de la señal de audio que se muestra en la ventana, controlando por tanto la escala horizontal (zoom) con la que se presentan las gráficas.

- **Menú Control:** gestiona las grabaciones, reproducciones, pausas, mezclas, ir al principio o al final de pistas, etc.
- **Menú Pistas:** permite añadir nuevas pistas al proyecto, alinearlas, silenciar algunas pistas y añadir etiquetas a las pistas.
- **Menú Generar:** se pueden generar distintos tipos de señales artificiales (sintéticas). Este menú será muy utilizado en esta actividad para generar señales sencillas que nos ayuden a interpretar las relaciones existentes entre la representación en tiempo y la representación en frecuencia.
- **Menú Efecto:** en este menú se dispone de diferentes posibilidades para generar efectos sobre la señal de audio. Las funciones que incluye dependen mucho de la versión del programa, ya que actualmente las actualizaciones del programa se están centrande principalmente en la incorporación de efectos adicionales. En esta asignatura sólo veremos algunos de los efectos conceptualmente más sencillos de interpretar.
- **Menú Analizar:** se utiliza para determinar la representación en frecuencia de la señal (espectro). Dispone además de otras herramientas de análisis que permiten encontrar silencios de forma automática, determinar el ritmo y el compás musical, etc.
- **Menú Ayuda:** incluye una interfaz para la búsqueda de ayuda por temas y palabras clave, el acceso a la pantalla de bienvenida y los tutoriales, la información de versión y un test de rendimiento de la máquina.

Además de estas opciones de menú, el programa dispone de varios **botones** que permiten un acceso rápido a las funciones más habituales. En la figura 2 se muestran algunas de las opciones que se utilizan con mayor frecuencia, como son los botones de reproducción y los controles de volumen de la salida y del micrófono.

En los ejercicios siguientes iremos presentando algunos ejemplos sobre las características de las señales de audio, a la vez que se expondrá cómo se realizan algunas de las funciones de Audacity.

1.3.4. Audacity: captura de señales y manejo básico de pistas

En este ejercicio vamos a aprender cómo se captura una señal de audio procedente de un CD y una señal de voz procedente de un micrófono gestionando la manipulación básica de las pistas de audio e interpretando las formas de onda.

1) Para empezar seleccionad un CD de audio cualquiera.

2) Para capturar la señal de audio de un CD, introducid el CD y seleccionad en el menú **Archivo** la opción **Importar** y después **Audio**.

Aparecerá el explorador de archivos.

3) Seleccionad el CD y la pista de audio que deseáis capturar.

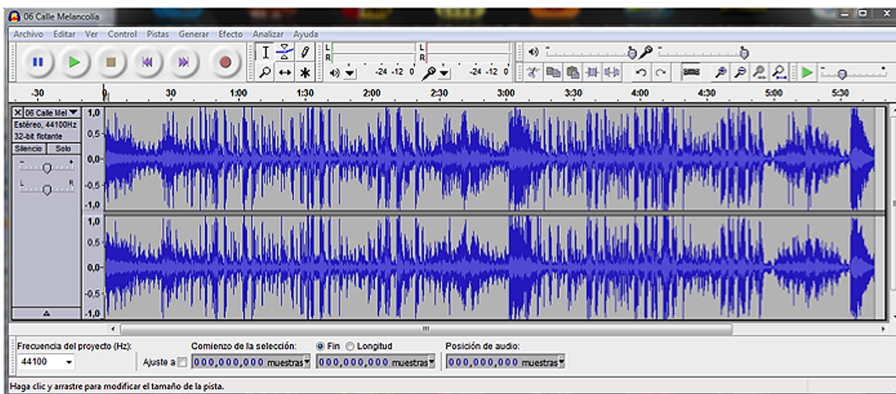
Esta operación resulta más sencilla en unos sistemas operativos que en otros.

- Si estáis utilizando **Mac OS X** o **LINUX** simplemente debéis seleccionar el tema de interés y de forma automática el sistema operativo seleccionará los conversores adecuados y cargará el archivo en Audacity.
- El sistema operativo **Windows** os pedirá que primero paséis del formato **CDA**, propio del CompactDisc Audio, al formato **WAV**. Hay diferentes formas de hacer este proceso. Posiblemente, la más sencilla es abrir el reproductor de **Windows Media**, seleccionar **Herramientas/Opciones** y en la pestaña de **Copiar Música desde CD** seleccionar el formato **WAV** (sin pérdida). Después, se trata de seleccionar el tema del disco que se pretende convertir a WAV y hacer la copia (con el botón **Iniciar Copia desde CD**). Por defecto, el tema seleccionado se convertirá al formato WAV y se archivará en el directorio predefinido de Windows **Mi Música**.

4) Ahora, para abrirlo con Audacity debemos seleccionar en el menú **Archivo** la opción **Importar/Audio**. Se abrirá la ventana del explorador de archivos y deberemos seleccionar el tema que acabamos de convertir a WAV.

En la figura 3 se muestra una captura de la pantalla de Audacity una vez se ha realizado la carga del tema musical. Se trata de un archivo estéreo, por lo que aparecen las formas de onda de la señal del canal derecho y la del canal izquierdo. En la parte superior de la pista estéreo está representado un eje de tiempos que en nuestro ejemplo llega prácticamente a los 6 minutos. Los ejes verticales de las gráficas están normalizados entre -1.0 y 1.0, lo que significa que la señal siempre estará comprendida entre estos dos valores.

Figura 3. Pantalla de Audacity una vez cargado un tema musical estéreo completo (aprox. 6 minutos)



5) Puede reproducirse el tema musical pulsando el botón de **Play** (en verde).

A medida que se realiza la reproducción observará el desplazamiento del cursor, que indica en qué momento de la reproducción nos encontramos. Observad que la amplitud que toman las señales en el eje vertical depende de la intensidad sonora. Así, en los instantes de poca intensidad sonora vemos que los niveles de las señales son bajos y que aumentan cuando aumenta la intensidad. En efecto, la forma de onda que se representa en la pantalla se corresponde con la tensión que, debidamente amplificadas, enviaríamos a un altavoz para la reproducción de la señal de audio.

En la línea inferior se nos indica que la frecuencia de muestreo del proyecto es de 44.100 Hz. Esto significa que estamos trabajando en un formato de audio digital que toma un total de 44.100 muestras por segundo de la señal (anlizaremos con mayor detalle este parámetro en la siguiente actividad de audio). También en la línea inferior se nos indica la posición del cursor en formato numérico. Por defecto se nos muestra el número de muestra en el que está situada la selección de la señal de audio.

Para manejar más fácilmente el archivo y realizar algunas pruebas sobre edición, vamos a seleccionar un fragmento de unos 30 segundos.

6) Con el ratón marcad un fragmento de la señal de aproximadamente 30 segundos (hacia la mitad del tema).

7) Una vez seleccionados estos 30 segundos, en el menú **Editar**, seleccionad **Recortar** (Ctrl+T).

Observaréis que desaparece toda la información que no estaba seleccionada.

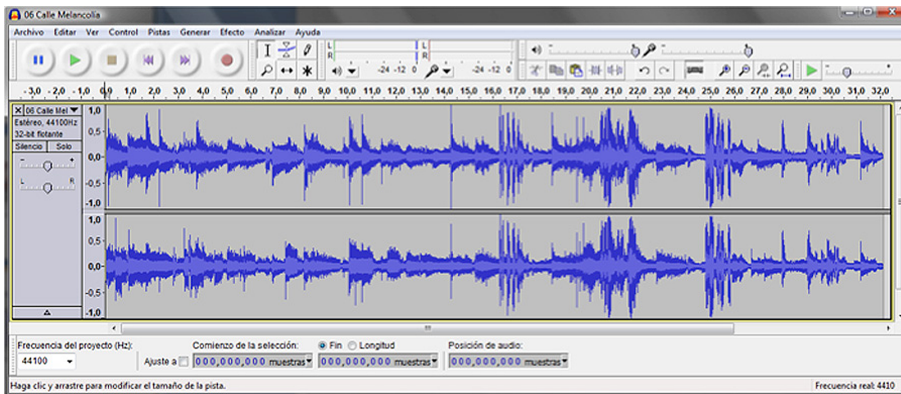
8) Para situar este segmento al principio de la pista, podemos hacer **Cortar** (Ctrl+X), después situar el cursor al principio de la pista y realizar un **Pegar** (Ctrl+V).

9) Ahora, en el menú **Ver** ejecutad el comando **Ajustar a la ventana** (Ctrl+F), con lo que se mostrará el detalle de las formas de onda en toda la ventana del Audacity.

10) **Reproducir** el archivo y observad cómo estamos trabajando sólo con un fragmento del archivo original.

En la figura 4 mostramos el resultado que hemos obtenido en nuestro ejemplo. Ahora, el eje horizontal tiene una duración aproximada de 32 segundos.

Figura 4. Recorte y selección de un fragmento de unos 30 segundos



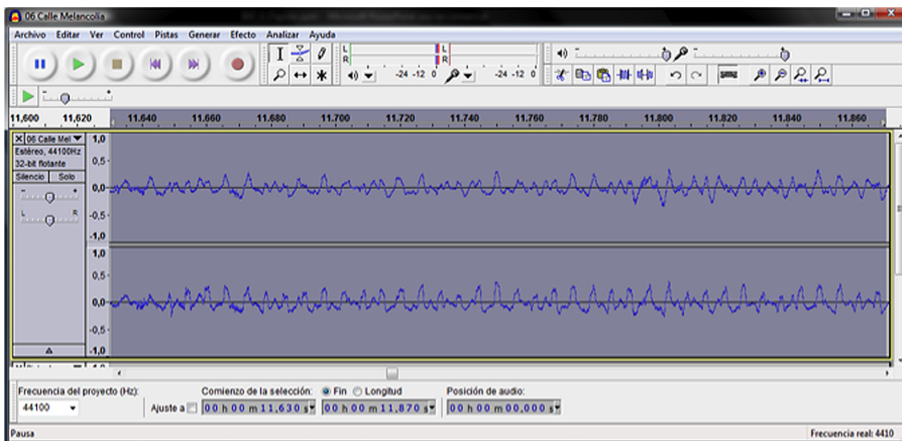
11) Podemos ver con mayor detalle la forma de onda de la señal seleccionando un fragmento con el ratón (seleccionad aproximadamente la mitad de un segundo o menos) y mostrando en la pantalla el detalle seleccionado mediante el comando **Ver/Ampliar la Selección** (Ctrl+E). Podéis volver a visualizar la pista completa con Ctrl+F.

En la figura 5 se muestra el resultado de ampliar la forma de onda en unos 0,25 segundos. Ahora es posible ver el detalle de la representación de la señal, es decir, de la tensión que enviaríamos a un altavoz para su reproducción.

Observación

Observad que el manejo de los comandos Cortar, Pegar, Copiar es análogo al de un procesador de textos. Simplemente, en vez de movernos a través de un documento nos estamos moviendo en las pistas de audio.

Figura 5. Ampliación de un fragmento de unos 240 ms para visualizar el detalle de la forma de onda



Ahora vamos a registrar una pista de voz utilizando el micrófono incorporado en el ordenador.

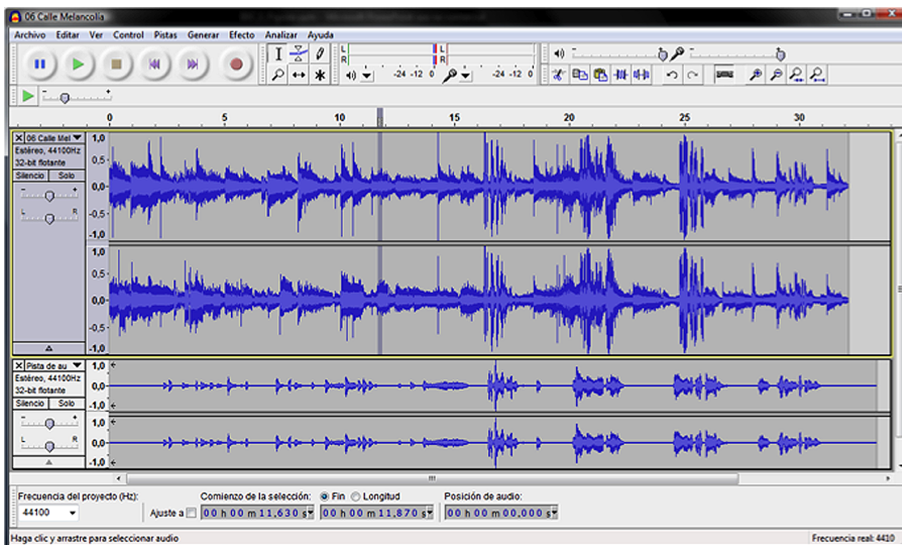
1) Situada el cursor al inicio de la pista de audio, pones unos auriculares y dispones a grabar su voz en una nueva pista de audio. Para ello, simplemente tenéis que pulsar el botón de **REC** (botón rojo).

Por los auriculares escucharéis la pista original mientras que por el micrófono se registrará vuestra voz. Acompañad al cantante o hablad mientras se efectúa la grabación.

2) Ahora podéis **reproducir** y observaréis cómo el programa realiza la mezcla de todas las pistas.

El resultado que hemos obtenido en nuestro ejemplo se muestra en la figura 6.

Figura 6. Detalle de la pista de audio procedente del CD más la pista de audio procedente de una grabación de voz



Podéis actuar sobre el resultado de la mezcla mediante el panel de control de pistas que aparece junto a las pistas. Este panel proporciona información sobre el tipo de señal (estéreo, 44100 Hz, 32-bit flotante) y dispone de varios controles para la mezcla. El botón Silencio deja la pista en silencio durante la reproducción, el botón Solo desactiva el resto de pistas durante la reproducción, escuchando sólo la pista de interés. Las dos barras de desplazamiento horizontales son para el control de volumen individual de la pista y para el control de balance entre el canal derecho o el izquierdo. Si queréis aplicar efectos a una pista, deberéis seleccionar con el cursor la pista deseada y aplicar alguno de los efectos disponibles en el menú efecto (probad de aplicar por ejemplo un efecto de eco a la pista de voz que acabáis de registrar).

Si lo deseáis podéis incorporar nuevas pistas realizando más grabaciones. Puede incluir señales de línea externas si dispone de conectores en su ordenador. Se sugiere que se realicen varias pruebas con diversas pistas, manipulando la información, aplicando funciones de recortar y pegar fragmentos de audio, modificar los niveles de las señales, aplicar efectos a las diferentes pistas, etc.; hasta que se encuentre cómodo con la aplicación y el manejo de las pistas.

3) Finalmente, una vez ajustados todos los parámetros de las pistas podéis realizar la grabación del resultado final mediante el comando **Exportar** que encontraréis en el menú **Archivo**.

1.3.5. Audacity: formas de onda básicas y representación tiempo frecuencia

En este ejercicio vamos a generar algunas formas de onda básicas, escuchando sus resultados y analizando su composición en frecuencia.

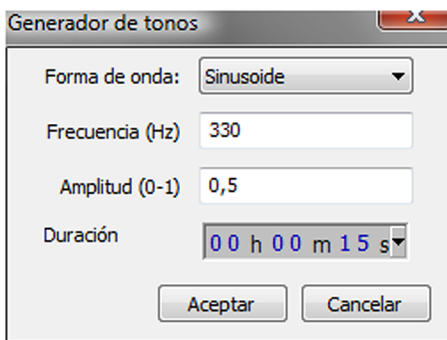
El objetivo es introducir la representación espectral de una señal que, como veremos a lo largo de la asignatura, puede resultar muy útil para interpretar la composición de los sonidos y generar distintos tipos de filtros para acondicionar las señales.

Vamos a empezar un nuevo proyecto de Audacity.

1) Podemos arrancar de nuevo la aplicación o ejecutar la opción **Nuevo** en el menú **Archivo** (Ctrl+N).

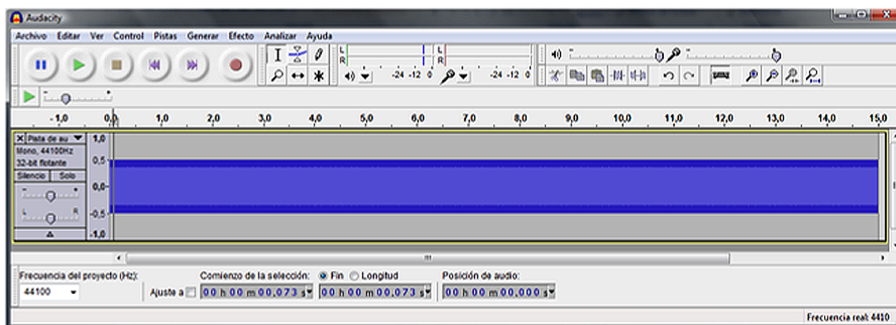
2) Ahora iremos al menú **Generar** y seleccionaremos la opción **Tono**, con lo que aparecerá el menú de generación de tonos representado en la figura 7. Seleccionaremos una forma de onda sinusoidal, con una frecuencia de 330 Hz, una amplitud de 0,5 y una duración de 15 segundos (los mismos parámetros que se muestran en la figura 7).

Figura 7. Cuadro de diálogo para la generación de un tono sinusoidal



Escuchad la señal. Se trata de un tono puro que se corresponde aproximadamente con la nota musical Mi3. En principio, al generar la señal se visualizará en la ventana la forma de onda en toda su longitud, es decir, 15 segundos. Como la señal tiene una frecuencia de 330 Hz, esto significa que en un segundo tendremos un total de 330 periodos de la sinusoide. Así pues, estamos viendo un total de $330 \times 15 = 4950$ periodos en la ventana que se representa en la figura 8. Evidentemente, la resolución de la pantalla gráfica del ordenador no nos permite distinguir la forma sinusoidal de la señal.

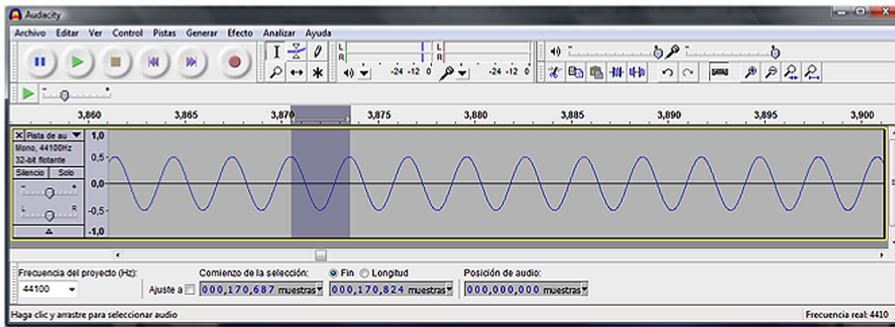
Figura 8. Representación gráfica de la forma de onda de la señal tonal durante los 15 s



Para visualizar la forma de onda sinusoidal, podemos seleccionar un fragmento corto de la señal y visualizarlo en la ventana de Audacity (Ctrl+E).

3) Seleccionad un fragmento y ampliadlo hasta obtener una gráfica con unas características parecidas a las que se muestran en la figura 9, donde puede apreciarse el carácter sinusoidal de la señal. En esta figura también hemos seleccionado un periodo completo de la señal. Tal y como nos indica la barra inferior, el comienzo de la selección va desde la muestra 170.687 hasta la muestra 170.824. Así pues, podemos concluir que el número de muestras de un periodo es de 137.

Figura 9. Representación de un fragmento de la señal sinusoidal y selección de un periodo básico



De acuerdo con este resultado, podríamos estimar la frecuencia de la señal a partir de su número de muestras. En efecto, como estamos tomando un total de 44.100 muestras por segundo. Observad que 44.100 muestras por segundo es un número muy elevado de muestras. Significa que el tiempo que transcurre entre dos muestras es tan solo de $1/44100$ segundos, es decir, 0,00002268 segundos, o lo que es lo mismo, 0,02268 milisegundos o 22,68 microsegundos. Como el número de muestras que tiene el periodo de repetición básico de la señal es de 137 muestras, el tiempo equivalente será de $137/44100 = 3,106$ milisegundos. La frecuencia es el inverso del periodo, por lo que obtendremos una estimación de la frecuencia $44.100/137 = 321,89$ Hz.

El resultado obtenido no coincide con la frecuencia teórica, pero esto es debido al error que cometemos al seleccionar un único periodo basándonos en la representación gráfica de la forma de onda.

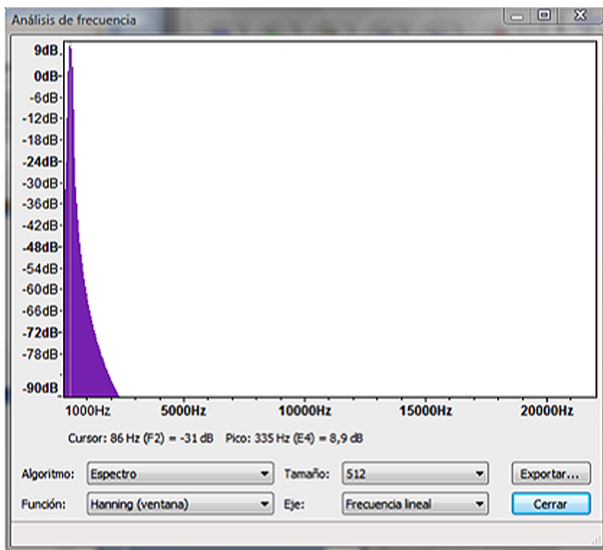
Actividad

Como ejercicio se propone que repitáis la selección de un periodo varias veces y observad cómo el número de muestras estimado varia, ya que es difícil realizar una estimación con mucha precisión.

Veamos una forma alternativa de estimar la frecuencia de la señal basándonos en el **análisis espectral**. Para realizar el análisis de la señal, volved a visualizar toda la señal completa (Ctrl+F).

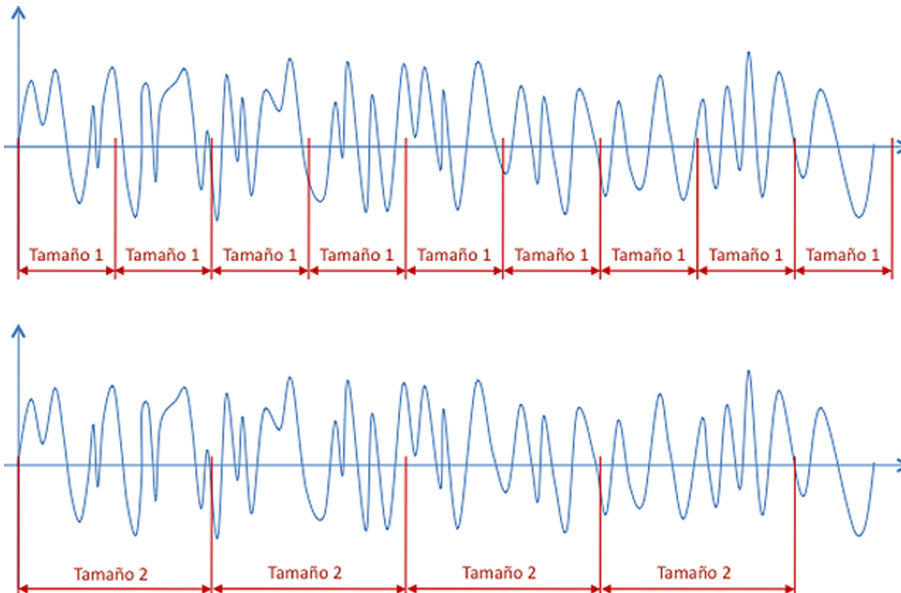
1) Seleccionad un fragmento de varios segundos de la señal y en el menú **Ana-lizar** seleccionad la opción **Análisis de espectro**. Obtendréis como resultado la gráfica de la figura 10 en la que se representa una estimación de las frecuencias que componen nuestra señal. Esta estimación se realiza con un método matemático denominado transformada de Fourier. Vemos que la estimación que hemos obtenido por defecto tampoco es excesivamente potente, ya que, tal y como se indica al pie de la figura, se detecta un pico en la frecuencia de 335 Hz (que tampoco se corresponde con la real de 330 Hz). En los siguientes párrafos vamos a ver cómo podemos mejorar esta estimación de la frecuencia modificando algunos parámetros.

Figura 10. Estimación de frecuencia obtenida a partir del análisis de espectro de Audacity



2) En la parte inferior de la ventana de **Análisis de frecuencia** vemos que se dispone de cuatro cajas con opciones que podemos seleccionar. La primera permite seleccionar varios algoritmos de análisis. En nuestro caso, siempre utilizaremos el **Análisis de espectro** para estudiar la composición en frecuencia de las señales. La segunda caja de selección permite ajustar el tamaño de la ventana de análisis. Este es un factor clave para mejorar la resolución del análisis de espectro de la señal. Para comprender el significado de este parámetro consideraremos el ejemplo que se presenta en la figura 11.

Figura 11. Análisis de Fourier realizado sobre fragmentos de la señal de audio

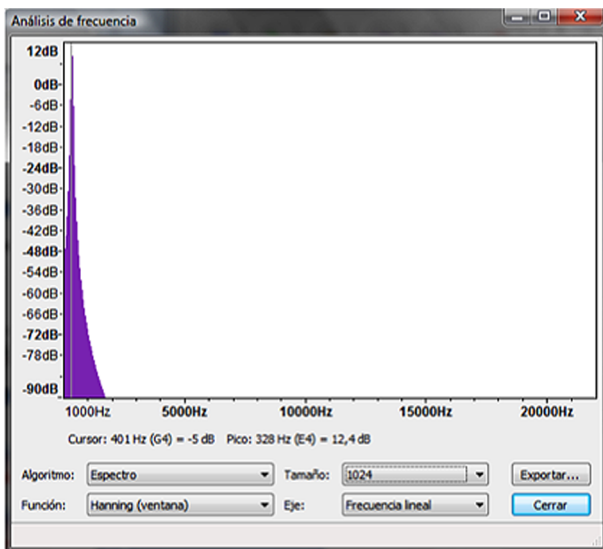


En este caso tenemos representada una señal con una selección de un elevado número de muestras, desde el principio al final de la gráfica. La estimación del espectro se hace dividiendo la señal en pequeños fragmentos (también denominados ventanas) que tienen un tamaño prefijado (512 muestras en el ejemplo de la figura 10). Para cada uno de estos fragmentos se realiza una

estimación del espectro (transformada de Fourier) y los resultados obtenidos se promedian, dando lugar a una gráfica como la mostrada en la figura 10. Así pues, lo que visualizamos es el promedio de todos los espectros calculados para las diferentes ventanas. En nuestro ejemplo, la transformada de Fourier se realiza sólo sobre un número prefijado de 512 muestras.

Si el tiempo durante el que analizamos la señal es pequeño no tendremos demasiada precisión en estimar las frecuencias que componen la señal. En nuestro caso, 512 muestras se corresponden con un tiempo de análisis de $512/44100 = 11,61$ milésimas de segundo. Si aumentamos el tamaño de la ventana al doble (1.024 muestras) estaremos analizando la señal durante un tiempo mayor, por lo que, sin entrar en los detalles de los algoritmos matemáticos de la transformada de Fourier, parece natural que podamos estimar las frecuencias que la componen con mayor precisión. En efecto, en la figura 12 mostramos el resultado del análisis en frecuencia que obtenemos aplicando una ventana de tamaño de 1.024 muestras sobre el mismo fragmento de la señal. El pico es mucho más agudo y su posición es más precisa (el máximo está en 328 Hz). Probad con una ventana de 2.048 muestras y veréis cómo la precisión aún mejora (331 Hz).

Figura 12. Análisis en frecuencia con ventanas de 1.024 muestras



Precisión en la frecuencia de la señal

Obsérvese que si ponemos una ventana suficientemente grande podemos determinar con gran precisión la frecuencia de la señal.

El párrafo anterior parece sugerir que cuanto mayor es la longitud de la ventana mejor será la estimación del espectro. Esto es cierto si las características de la señal no varían, como es nuestro caso. En efecto, tenemos un **tono puro** que se mantiene en las mismas condiciones durante 15 segundos completos, por lo tanto, en principio, podríamos utilizar toda esta información para estimar el espectro de la señal. En cambio, con señales musicales o de voz, las características varían de forma muy rápida en el tiempo, los cambios son dinámicos y en ocasiones puede interesar que estas ventanas sean más pequeñas para estudiar cómo evolucionan las características del espectro a lo largo del tiempo.

Resumiendo, dependiendo de las características de las señales puede que estemos más interesados en utilizar fragmentos de **gran longitud** (señal estacionaria) o fragmentos de **corta longitud** (señales variantes). En cada caso deberemos seleccionar la longitud apropiada.

3) Además del tamaño de la ventana también podemos seleccionar la función que por defecto tiene seleccionado el valor Hanning. En principio no modificaremos este valor para las diferentes estimaciones de frecuencia que realizaremos en esta asignatura. La función representa la importancia que se da a las muestras dentro de la ventana. Así, con una ventana rectangular (una de las opciones que podemos elegir), todas las muestras de la ventana tienen la misma importancia. En una ventana de Hanning las muestras centrales tienen más peso en los algoritmos que realizan la estimación de los espectros.

4) Finalmente, también podemos seleccionar el tipo de escala que se aplica al eje horizontal. Por defecto aparece la escala lineal, aunque en audio, muchas veces es habitual utilizar una escala logarítmica. La razón por la que se utiliza una escala logarítmica es que el oído se comporta de forma logarítmica, es decir, tenemos mucha más resolución para discriminar señales de baja frecuencia que para distinguir señales de alta frecuencia. Probad a modificar el escalado en el eje horizontal a logarítmico. En este caso, la posición de las frecuencias en el eje horizontal se corresponde con la percepción real que tenemos de las frecuencias de las señales.

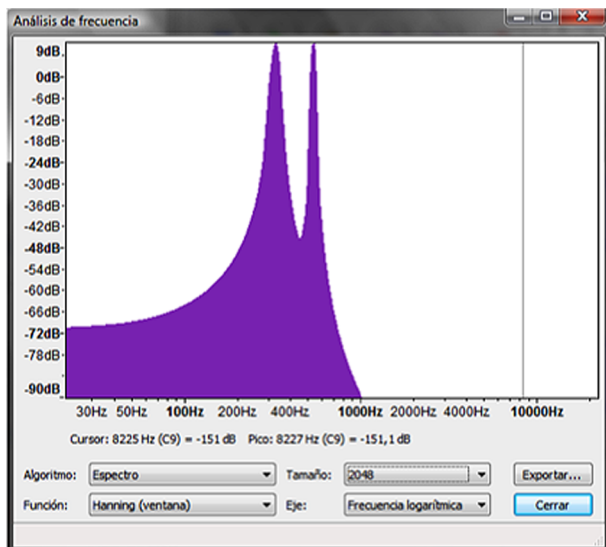
1.3.6. Superposición de tonos

Ahora vamos a generar un nuevo tono a una frecuencia de 540 Hz.

1) Para ello, acceded al menú **Pistas** y seleccionad **Añadir nueva/Pista de audio**. Sitúad el cursor al inicio de esta nueva pista y generad un tono sinusoidal de 15 segundos, con una amplitud de 0.5 y una frecuencia de 540 Hz. Escuchad la señal y comprobad cómo el oído puede distinguir perfectamente entre las dos frecuencias.

2) Seleccionad ahora con el cursor un fragmento de varios segundos de las dos señales y haced el análisis de espectro (para seleccionar las dos pistas, debéis crear un rectángulo con el ratón que incluya las dos pistas (utilizad Ctrl+A si queréis seleccionar todas las señales de todas las pistas). Usad varios valores del **tamaño** de la ventana de análisis y pasad de frecuencia lineal a logarítmica para ver los efectos sobre la señal. Comprobad que las estimaciones de frecuencia son correctas, siempre que el tamaño de la ventana sea suficientemente grande.

Figura 13. Ejemplo de superposición de 2 tonos



Muestra de uno de los resultados de los espectros que hemos obtenido con este ejemplo.

Observación

Observad cómo ahora el espectro muestra claramente la superposición de los dos tonos.

1.3.7. Secuenciación de tonos e interpretación del espectro

Este ejemplo es parecido al anterior pero en lugar de poner los dos tonos simultáneamente, en pistas paralelas, vamos a ponerlos uno detrás del otro usando la misma pista.

1) Partimos del ejercicio anterior y borramos la segunda pista (basta con marcar la cruz en la caja que tiene los controles de volumen y balance de la pista).

2) Después ponemos el cursor al final de la primera pista, la que tiene un tono con 330 Hz y 15 segundos de duración. Generamos ahora un tono de 540 Hz, con 15 segundos de duración. Como resultado tendremos una única pista con una duración de 30 segundos. En los 15 primeros segundos tenemos un tono de 330 Hz y en los siguientes, uno de 540 Hz. Si escuchamos la señal observaremos el cambio de frecuencia.

3) Seleccionemos ahora un fragmento de audio que incluya los dos tonos y aproximadamente durante el mismo tiempo. Por ejemplo, seleccionad el audio entre el segundo 12 y el 18. Realizad el análisis de espectro.

Veamos qué ocurre cuando el fragmento que seleccionamos incluye el doble de tiempo de una señal que de la otra.

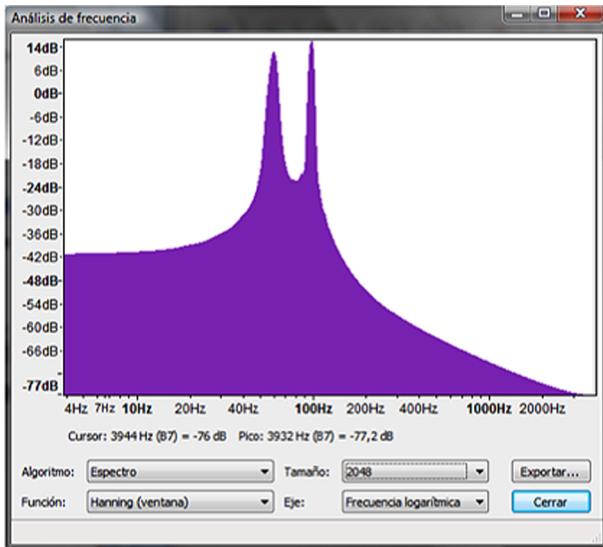
4) Seleccionad por ejemplo entre el segundo 12 y el 21. Ahora tenemos una selección donde el primer tono está presente durante 3 segundos mientras el segundo tiene una duración de 6 segundos.

Reflexión

El resultado que deberíais obtener es el que ya habéis obtenido en el ejercicio 3. ¿Podéis explicar por qué?

5) Haced el análisis de frecuencia. Ahora deberíais tener un resultado como el de la figura 14, donde la amplitud del pico situado en la frecuencia de 540 Hz es mayor que la del pico situado en 330 Hz. Intentad razonar por qué uno de los picos tiene mayor potencia que el otro.

Figura 14. Espectro obtenido cuando hacemos el análisis tomando un mayor periodo de tiempo de la señal de 540 Hz.



6) Finalmente, seleccionad un fragmento de señal en el que sólo esté presente uno de los tonos. Veremos que ahora en la estimación de espectro sólo aparece el tono que hemos seleccionado.

La explicación de estos resultados se corresponde con la idea que hemos explicado en la figura 11 al comentar el procedimiento de cálculo de la estimación de espectro. En efecto, la señal se divide en fragmentos y se calcula el espectro de cada uno de los fragmentos. Todos estos espectros son posteriormente promediados y presentados en la pantalla. Esto significa:

- a) Cuando hemos tomado el mismo tiempo de los dos tonos en el promedio han participado tantos espectros correspondientes a uno de los tonos como del otro, por lo tanto, han aparecido los dos picos con la misma importancia.
- b) Cuando uno de los tonos ha intervenido el doble de tiempo que el otro se han promediado más espectros correspondientes al segundo tono que al primero, por lo que el pico ha tenido una mayor amplitud.
- c) Cuando seleccionamos un fragmento en el que sólo tenemos un tono, sólo participa este tono en el promedio.

Por tanto, es importante tener en cuenta que la gráfica que nos presenta el análisis de espectro es una gráfica promediada y que debemos interpretarla correctamente en cada momento.

Esencialmente, el análisis de espectro nos proporciona una estimación de las características frecuenciales de la señal durante el tiempo de duración de la ventana de análisis. Esto nos permite ver en una sola gráfica los diferentes elementos de frecuencia que componen la señal, de un solo vistazo podemos conocer la composición en frecuencia de una señal compleja. No obstante, si la señal cambia sus características a lo largo del tiempo no sabemos en qué instantes se producen estos cambios. En general, para ver con precisión los instantes en los que se producen los cambios necesitamos recurrir a la **representación temporal**. El espectrograma, que veremos en el siguiente ejercicio, constituye un excelente compromiso entre la resolución en tiempo (que nos proporciona la representación temporal) y la resolución en frecuencia (que nos proporciona la estimación de espectro). El espectrograma es una representación en 3D de cómo los espectros van evolucionando a lo largo del tiempo. Veremos los detalles de esta representación en el ejercicio siguiente.

Observación

Nótese, por ejemplo, que si no examinamos con mayor detalle la señal, podríamos concluir que el resultado que hemos obtenido en el primer apartado es debido a que tenemos dos tonos que están activos al mismo tiempo.

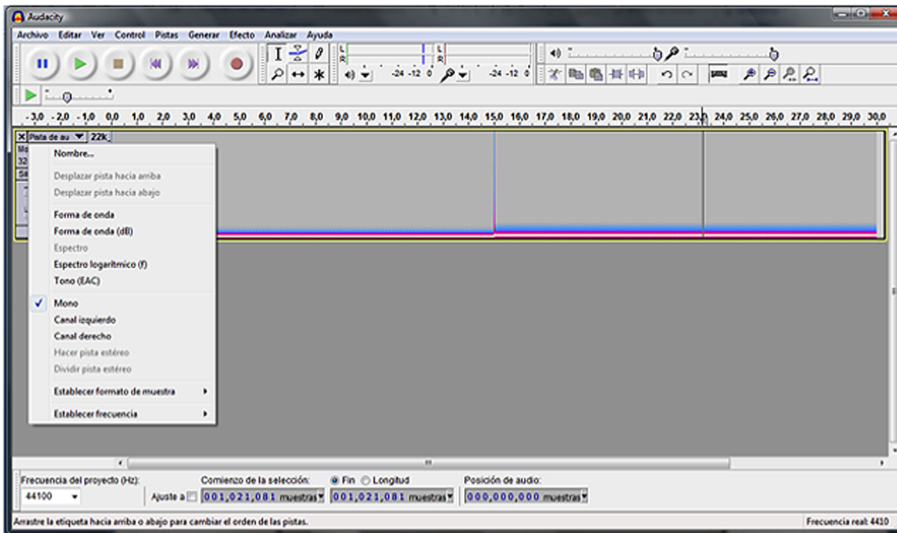
1.3.8. Espectrograma

Una alternativa al análisis de espectro que hemos realizado en el apartado anterior es el espectrograma. Esencialmente, el espectrograma consiste en representar todos los espectros que se van obteniendo al analizar cada uno de los fragmentos de una señal audio. Elaboraremos este primer ejemplo del espectrograma con la misma señal que hemos generado en el apartado anterior.

1) Para visualizar el espectrograma de una señal de audio con el programa Audacity, debemos pulsar en el desplegable de **Pista de audio** (ver figura 15) y seleccionar la opción Espectro.

- La representación que obtenemos debe ser interpretada correctamente y con sumo cuidado. En el eje horizontal seguimos teniendo el tiempo mientras que en el eje vertical tenemos la frecuencia. La frecuencia en nuestro ejemplo va desde 0 a 22 kHz. Los espectros se representan en colores, asignando los colores más claros en aquellas frecuencias en las que se producen amplitudes o picos significativos.
- La representación del espectrograma también puede realizarse en un eje logarítmico de frecuencias. Para ello, debe seleccionarse la opción de Espectro logarítmico en el desplegable de la pista de audio.

Figura 15. Representación del espectrograma de la señal



Observación

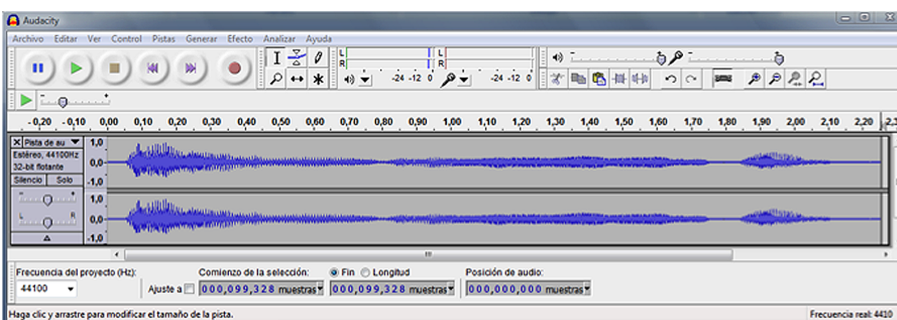
Obsérvese en nuestro ejemplo cómo después de los 15 segundos la raya blanca (el pico de frecuencia) se desplaza ligeramente hacia arriba debido al cambio de la frecuencia.

1.3.9. Análisis de espectro de la señal de voz

En este ejercicio vamos a examinar el espectro de la señal de voz. Para ello proponemos que grabéis vuestra propia voz diciendo la palabra *saludo* lentamente. En la figura 16 se muestra una posible forma de onda obtenida con esta palabra.

Cuando digitalicéis vuestra propia versión, veréis que los resultados no coinciden ni en forma de onda, ni en estimación de frecuencia con los que presentamos en el ejemplo. Esto es debido a que cada persona tiene sus propias características de generación de voz. De hecho, podemos reconocer a las personas por su voz debido a estas características específicas de cada uno.

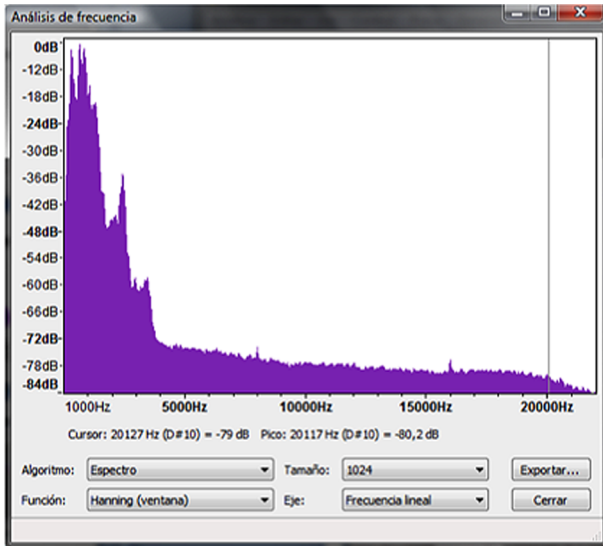
Figura 16. Forma de onda de la señal de voz con la palabra *saludo*



Se os pide que hagáis las siguientes pruebas con la señal que habéis adquirido:

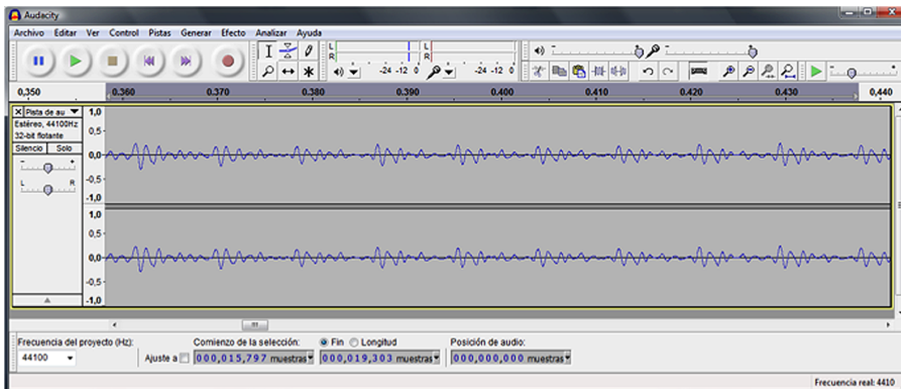
a) Seleccionad toda la señal completa (Ctrl+A) y haced el análisis de espectro. Tened en cuenta que lo que estáis visualizando es el promedio de todos los espectros de los fragmentos analizados para esta señal. El espectro que hemos obtenido en nuestro ejemplo se muestra en la figura 17.

Figura 17. Espectro promediado en toda la señal de voz



b) Seleccionad la parte de la /A/ y haced un zoom sobre la forma de onda. Deberíais ver que la forma de onda de las vocales es una señal más o menos compleja, pero que se repite de forma periódica. En la figura 18 vemos la forma de onda obtenida para nuestro ejemplo.

Figura 18. Forma de onda de la vocal /A/

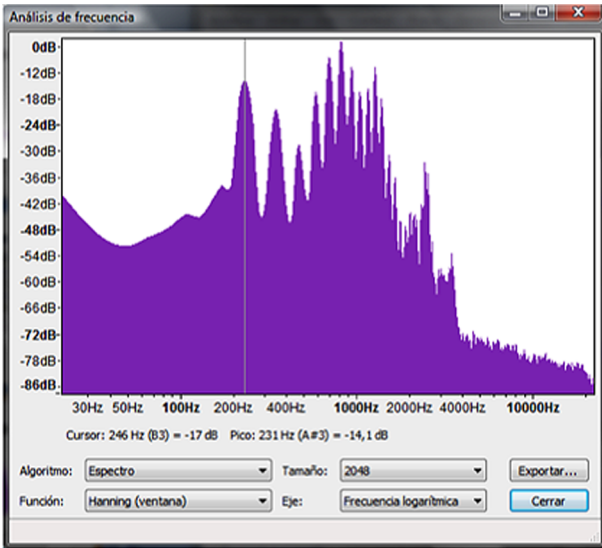


Pitch

Obsérvese, tal y como se marca en el recuadro en rojo, que existe una forma básica que se repite de manera periódica. El periodo de repetición se denomina periodo de *pitch*.

c) Seleccionad un fragmento suficientemente largo de la vocal /A/ y haced el análisis del espectro. El resultado que deberíais obtener es una señal armónica compleja, como lo muestra nuestro ejemplo de la figura 19. En nuestro caso detectamos un primer pico significativo en la frecuencia de 231 Hz y posteriormente, aparece una serie de armónicos que son los que le dan el matiz tímbrico de esta voz.

Figura 19. Espectro de la vocal /A/



d) Repetid la misma prueba que en los apartados b) y c) con la vocal /U/. Se proporcionan unas gráficas de ejemplo en las figura 20 y figura 21.

Figura 20. Forma de onda de la vocal /U/

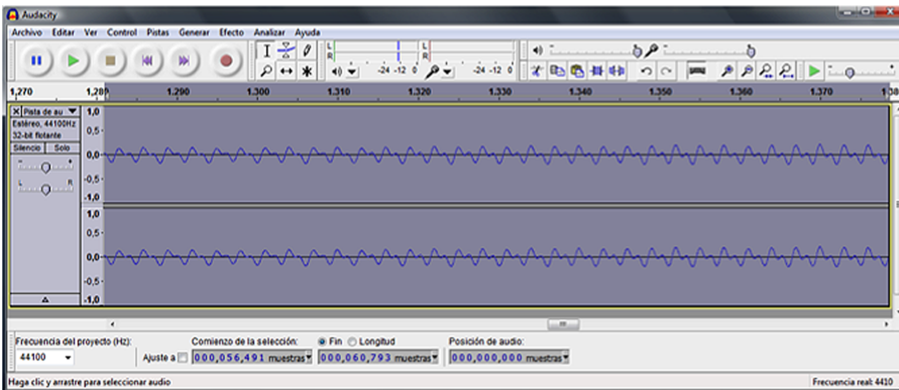
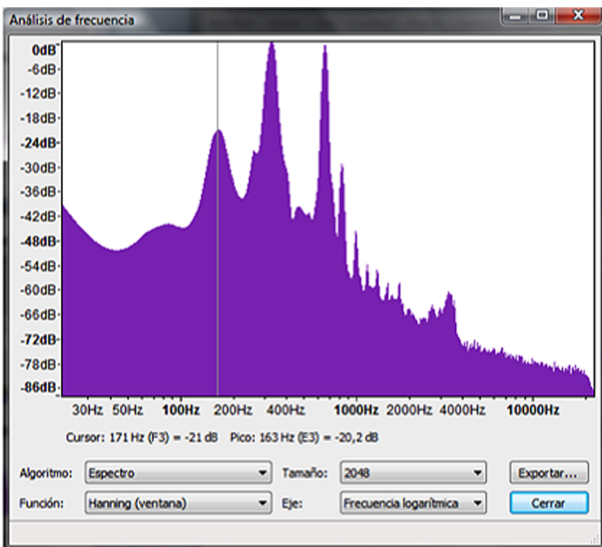


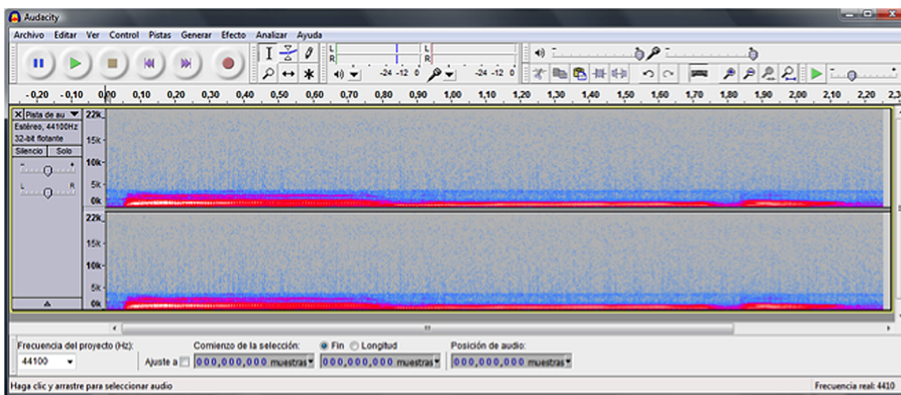
Figura 21. Espectro de la vocal /U/



e) Intentad ver la forma de onda de las consonantes /S/ y /D/ (Es más complejo debido a que la duración de las consonantes es más corta que la de las vocales). Comprobad que no se detectan periodicidades.

f) Finalmente, representad el espectrograma de la señal de audio completa y comprobad cómo evolucionan los espectros a lo largo del tiempo, modificándose la frecuencia fundamental de la voz y la importancia relativa de los armónicos. El espectrograma suele considerarse como una medida biométrica característica de la persona, como si se tratase de una huella dactilar. Hay métodos para identificar si una persona está en una determinada grabación a partir del análisis del espectrograma. En la figura 22 se muestra el resultado del espectrograma obtenido en nuestro ejemplo.

Figura 22. Espectrograma de la palabra *saludo*



1.3.10. Espectrograma de un silbido

En este ejercicio vamos a profundizar en el estudio del histograma mediante el análisis de uno sencillo, pero que puede resultar muy ilustrativo, como es un silbido.

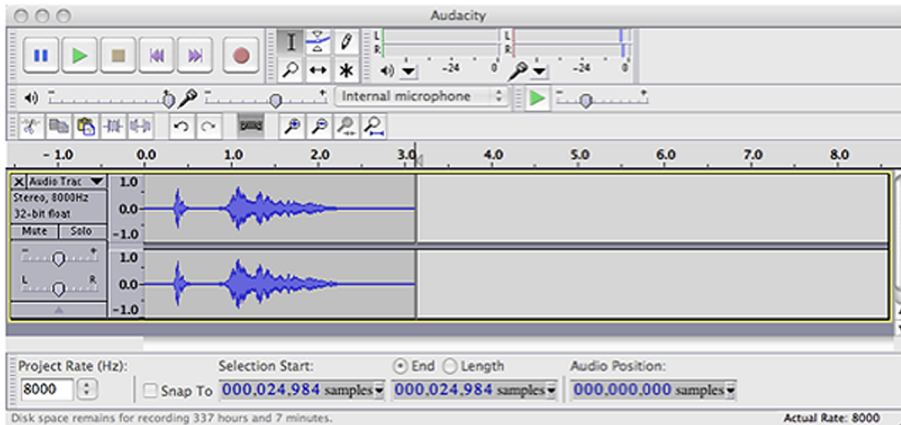
Se propone que se realice la captura mediante el programa Audacity de un silbido (un fragmento relativamente corto) donde se produzca un cambio de frecuencia significativo.

La idea principal de este ejercicio es comprobar cómo la representación del espectrograma nos ayuda a situar las componentes de frecuencia en el tiempo, mientras que las representaciones en el dominio temporal y en el dominio frecuencial no son tan claras.

1) En efecto, en la figura 23 se muestra el resultado de adquirir un silbido y su representación en forma de onda. Para que la representación en frecuencia y en espectrograma pueda visualizarse correctamente, sería conveniente que

modificara la frecuencia de muestreo a 8.000 Hz antes de adquirir la señal (la frecuencia de muestreo puede modificarse en la caja de selección situada en la parte inferior izquierda de la aplicación).

Figura 23. Representación de un silbido en el dominio temporal

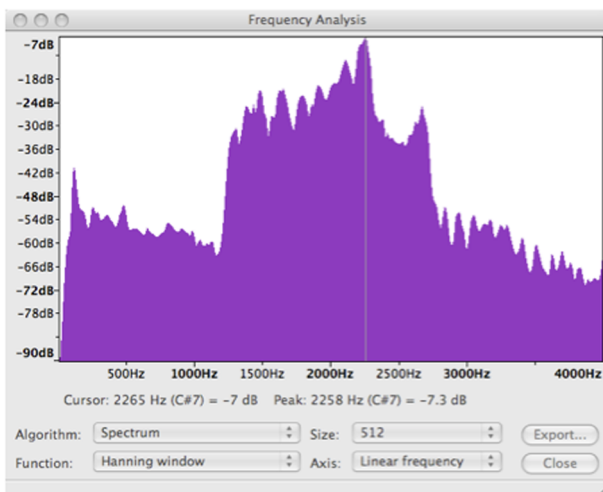


La frecuencia de muestreo está modificada a 8.000 Hz.

Vemos en la representación en tiempo del silbido que es posible ver que la señal se compone de dos fragmentos básicos, que podemos situarlos perfectamente en el tiempo, aunque no podemos determinar sus frecuencias. Así pues, la representación en tiempo nos permite tener mucha precisión para saber cuándo cambia la señal sus características.

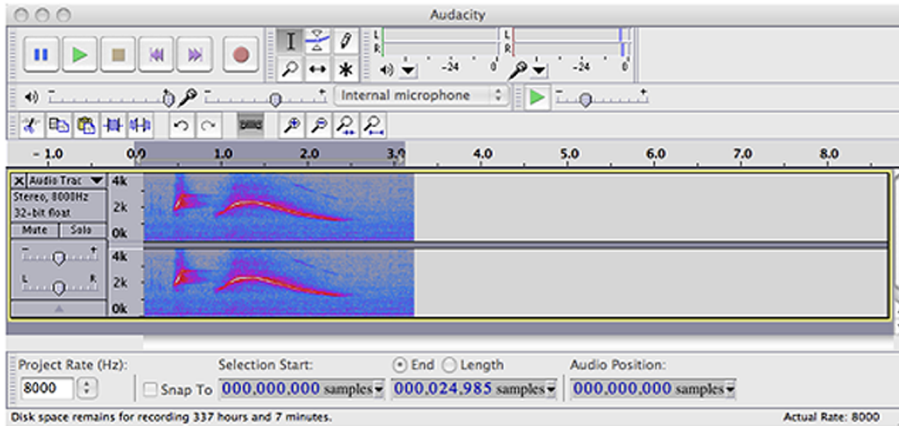
2) En la figura 24 representamos la estimación espectral que hemos obtenido seleccionando toda la señal y haciendo el análisis de espectro. Ahora vemos que la señal es espectralmente compleja, formada por muchas frecuencias que podríamos determinar con precisión. No obstante, en esta representación no sabemos los instantes de tiempo en los que se producen estos cambios, es decir, dónde se llevan a cabo los cambios de frecuencia y cómo están dispuestas las frecuencias en el tiempo.

Figura 24. Representación de la estimación de espectro del silbido mostrado en la figura anterior



3) Finalmente, en la figura 25, que corresponde al espectrograma, vemos cómo obtenemos un excelente compromiso entre **resolución temporal** y **frecuencial**, ya que es posible identificar la evolución del espectro a lo largo del tiempo.

Figura 25. Representación del silbido mediante el espectrograma



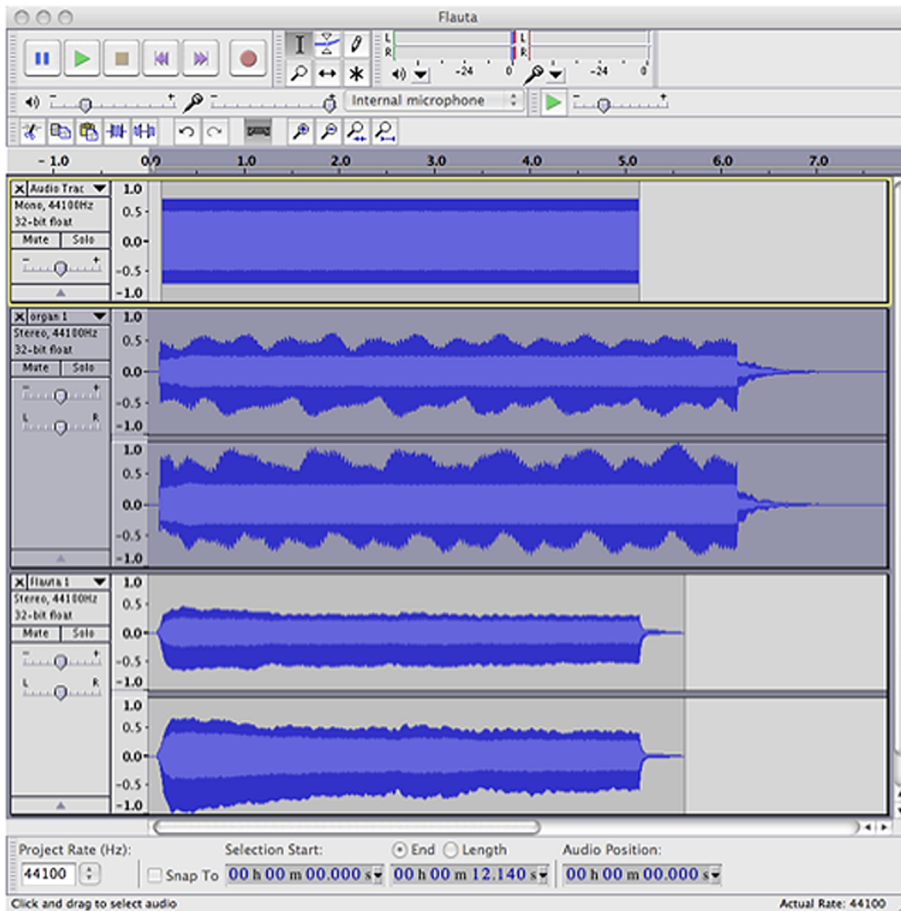
En esta gráfica podemos apreciar cómo las diferentes componentes frecuenciales evolucionan en el tiempo.

1.3.11. Timbre de los instrumentos musicales

Para finalizar con este apartado, vamos a analizar espectralmente instrumentos musicales que producen la misma nota.

- 1) Inicializad de nuevo Audacity y generad una señal sinusoidal de 330 Hz con una duración de unos 5 segundos.
- 2) Posteriormente, importad en una nueva pista la señal organo.wav que se proporciona con el material de esta práctica.
- 3) Finalmente, importad también el fichero flauta.wav. La pantalla de Audacity debería disponer de las tres señales tal y como se muestran en la figura 26.

Figura 26. Pantalla de Audacity con las tres pistas correspondientes al tono, el órgano y la flauta



Se pide que hagáis las siguientes pruebas:

- Reproducid las tres señales individualmente. Observad que los tres están produciendo la misma nota musical, aunque el oído puede identificarlos claramente como instrumentos distintos.
- Seleccionad el tono puro de 330 Hz y comprobad mediante el análisis espectral, que está compuesto por una única frecuencia.
- Seleccionad la señal proporcionada por el órgano y determinad su espectro. Observad la riqueza de armónicos a frecuencias que son múltiplos de la frecuencia fundamental.
- Repetid el apartado anterior para la flauta. En este caso también vemos que existe una importante riqueza armónica, aunque las amplitudes relativas de los armónicos son diferentes a la del caso del órgano. Son estas diferencias en las importancias relativas de cada armónico las que caracterizan cada uno de los instrumentos musicales, proporcionándoles un timbre característico a cada uno. En las figura 27 y 28 se muestran los espectros obtenidos para el órgano y la flauta respectivamente.

Figura 27. Análisis de espectro del órgano para una nota Mi3

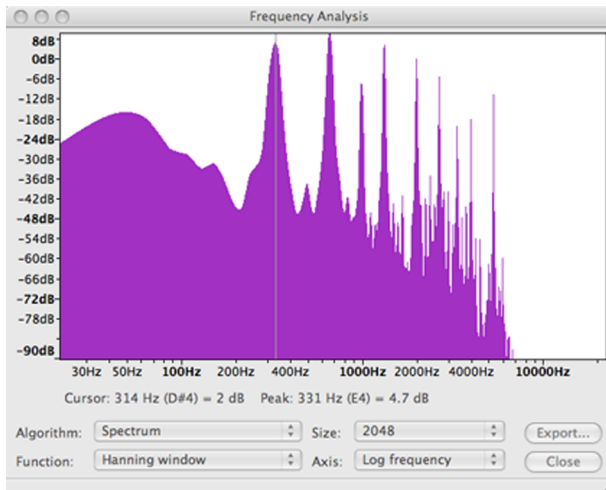
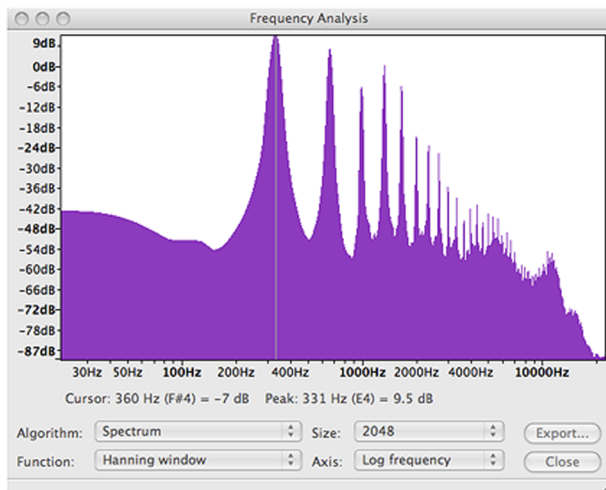


Figura 28. Análisis de espectro de la flauta para una nota Mi3



1.4. Problemas/preguntas

1) Se propone que capturéis la señal de audio de una pista de CD, tal y como ya habéis hecho en el primer ejercicio de esta actividad. Podéis utilizar el mismo tema musical u otro.

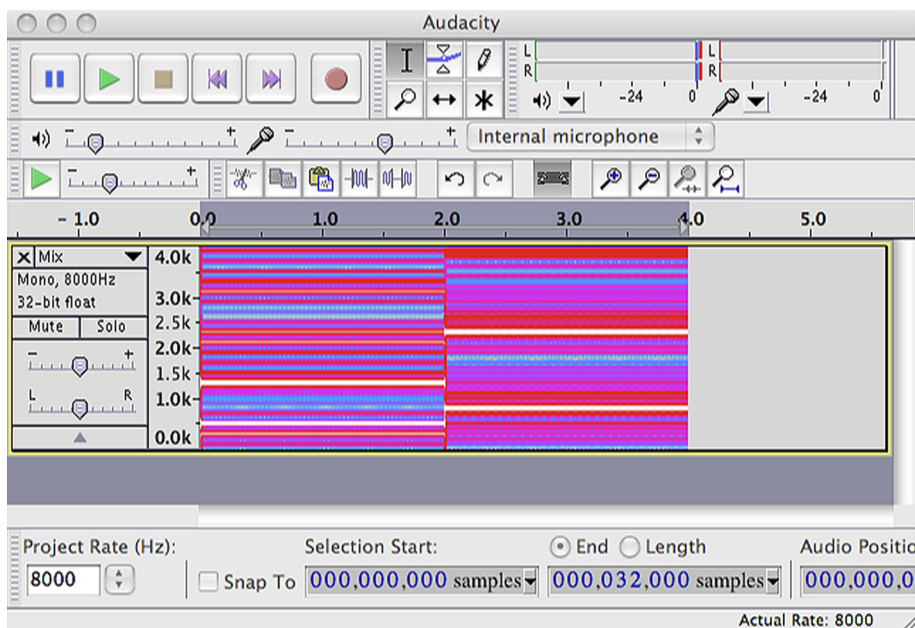
- Quitad los auriculares del ordenador, amplificad la señal por los altavoces hasta un nivel de conversación normal y activad el botón de grabación REC de Audacity. Con esto se grabará una nueva pista, capturada por el micrófono, sobre la señal que están reproduciendo los altavoces.
- Una vez finalizada la grabación, escuchad el resultado obtenido con la superposición de las dos pistas. Explicad qué ocurre cuando superponéis las dos pistas. ¿Podéis estimar cuál es el retardo entre las dos pistas?

2) Elaborad una lista, lo más completa posible, de aplicaciones de captura y edición de audio que encontréis por Internet. Seleccionad 3 de estas aplicaciones e intentad enumerar con el mayor detalle posible sus características incluyendo aspectos como:

- Número de pistas que pueden procesarse.
- Características de la interfaz gráfica (capturad las pantallas).
- Capacidades de cortar, pegar, copiar.
- Control de volumen y balance individual por pista.
- Representación de la forma de onda temporal de las señales de cada pista.
- Representación del espectro de las pistas, control de tono o ecualización de las pistas individuales.
- Capacidad de controlar tarjetas de captura de audio específicas.
- Disponibilidad de efectos de audio (enumerad los efectos más significativos, etc.).

3) Tenemos una señal de una duración de 30 segundos sobre la que hemos realizado un espectrograma que se muestra en la figura adjunta. ¿Qué características tiempo-frecuencia podéis indicarnos sobre esta señal?

Figura 29



Proporcionad los detalles sobre:

- Cuántos fragmentos estacionarios observáis.
- Qué duración tiene cada uno de los fragmentos estacionarios.
- Cuáles son las dos componentes frecuenciales más importantes de cada uno de los fragmentos estacionarios de la señal.

4) En el archivo **marcadoTelefónico.wav**, que se proporciona junto con el material de esta actividad, hemos registrado los tonos que produce un teléfono móvil al marcar un número. La calidad de la grabación no es excelente, de he-

cho, ha sido capturada por el micrófono incorporado en un ordenador portátil mientras se marcaban los números en el teléfono móvil. Se pide que determinéis los números que han sido marcados sabiendo que la tabla de frecuencias que utilizan los teléfonos para marcar los números es la de la figura adjunta.

Figura 30

	1.209 Hz	1.336 Hz	1.477 Hz	1.633 Hz
697hz	1	2	3	Un
770hz	4	5	6	B
852hz	7	8	9	C
941hz	*	0	#	D

2. Actividad 2. Digitalización de señales de audio: muestreo

2.1. Introducción

2.1.1. Contenido

La captura y digitalización de las señales de audio tienen un papel fundamental en todos los sistemas de edición, efectos y producción de sonido, ya que los parámetros con los que se ha hecho la adquisición determinan directamente la **calidad final del sonido**.

En esta actividad introduciremos el proceso de digitalización de la señal de audio como un problema de aproximación de las señales analógicas mediante una **retícula**.

Esta retícula nos permite identificar dos elementos clave en la digitalización:

- la frecuencia de muestreo (paso de la retícula en el eje horizontal) y
- el paso de cuantificación (paso en el eje vertical).

Los criterios para seleccionar la frecuencia de muestreo y su incidencia sobre la calidad final de la señal serán tratados en esta actividad. En la actividad siguiente se analizarán los efectos del paso de cuantificación y se evaluarán los efectos de ambos en la digitalización y el ancho de banda de las señales.

Para elaborar el estudio sobre la frecuencia de muestreo se proponen diferentes pruebas auditivas que intentarán familiarizarnos con la relación entre la frecuencia de muestreo y la calidad de la señal. Estas pruebas nos indicarán que existe un límite en la frecuencia de muestreo más allá del cual el oído humano no es capaz de apreciar diferencias entre las señales, por lo que no tendrá sentido aumentar excesivamente el número de muestras por segundo. Una vez introducidos estos conceptos de forma experimental, intuitiva y poco rigurosa, se enuncia el **teorema del muestreo** o **teorema de Nyquist**, que nos proporciona las condiciones necesarias para conocer la frecuencia a la que debe muestrearse una señal.

La parte final de la actividad se centra en el **problema del aliasing** definiéndolo de forma intuitiva y proporcionando diferentes experiencias auditivas para identificarlo.

2.1.2. Metodología

La actividad se centra en definir el concepto de muestreo de una señal y establecer los criterios para seleccionar el número de muestras por segundo que deben tomarse para su correcta representación.

La metodología utilizada es la realización de diferentes pruebas auditivas con ficheros de audio digitalizados con varias frecuencias de muestreo y la comparación de los resultados obtenidos.

Con ello, se pretende que nos familiaricemos con los efectos que la selección de la frecuencia de muestreo tiene sobre la calidad final de la señal de audio.

Para apoyar la audición también se sugiere que se analicen las señales desde el punto de vista de su espectro, con lo que es posible asociar el espectro de una señal con sus características audibles y ver cómo incide la frecuencia de muestreo en el espectro de la señal. Se aplican conceptos de frecuencia y espectro que han sido estudiados en actividades previas.

Una vez introducidos los conceptos básicos desde un punto de vista puramente experimental, se formalizan los conceptos de teorema de muestreo y aliasing, sugiriendo que el estudiante realice otras experiencias de digitalización y medida de espectros. El teorema del muestreo y el aliasing se presentan sin demostración, aunque se proporcionan muchos ejemplos auditivos para que queden consolidados.

2.1.3. Recursos

Los principales recursos utilizados en esta actividad son:

- Archivos de audio predigitalizados con diferentes frecuencias de muestreo para su audición.
- Software de captura y procesado de audio Audacity (licencia libre).
- Captura y análisis de voz a diferentes frecuencias de muestreo.
- Explicaciones de los fundamentos de la digitalización y la frecuencia de muestreo incluidas en el texto de la guía de la actividad.
- Propuesta de actividades de búsqueda de información en webs y recursos de Internet para profundizar en los conocimientos adquiridos en esta actividad (webs de fabricantes de tarjetas de captura, wiki, etc.).

- Problemas propuestos para verificar la adquisición de los conocimientos de esta actividad.

2.2. Objetivos

Los principales objetivos que se persiguen con esta actividad son:

- 1) Introducir el concepto de digitalización de una señal.
- 2) Interpretar la función de los dos elementos clave en el proceso de digitalización de una señal: la frecuencia de muestreo y el paso de cuantificación.
- 3) Relacionar, mediante experimentos dirigidos, la calidad subjetiva de una señal de audio con su frecuencia de muestreo.
- 4) Conocer el teorema del muestreo y los criterios de selección de la frecuencia de muestreo.
- 5) Identificar el fenómeno del aliasing en sus diferentes manifestaciones, tanto en audio como en otro tipo de señales.
- 6) Conocer las frecuencias de muestreo que se utilizan en los principales sistemas de audio y voz digitales.

Estos objetivos están relacionados con las siguientes **competencias de la asignatura**:

- 1) Capacidad de capturar audio e imágenes de forma eficiente y eficaz.
- 2) Capacidad de digitalizar eficiente y eficazmente contenidos de audio e imagen.
- 3) Capacidad de interpretar y modificar información auditiva de forma frecuencial.

Y con las siguientes **competencias generales del grado**:

- 1) Capturar, almacenar y modificar información de audio, imagen y vídeo digitales aplicando principios y métodos de realización y composición del lenguaje audiovisual.
- 2) Atender adecuadamente consultas sobre proyectos, tecnologías y mercado de productos multimedia, evaluando de forma precisa el entorno de aplicación, los recursos y las alternativas tecnológicas disponibles.

2.3. Guía

En esta actividad vamos a centrarnos en el proceso de digitalización de la señal de audio y muy especialmente en los criterios necesarios para elegir una frecuencia de muestreo adecuada.

El objetivo final que perseguimos es determinar el **número de muestras** que deben tomarse en cada segundo de una señal en función de sus características. Además, queremos familiarizarnos con este parámetro para mostrar cómo su elección resulta fundamental en la calidad final de la señal de audio.

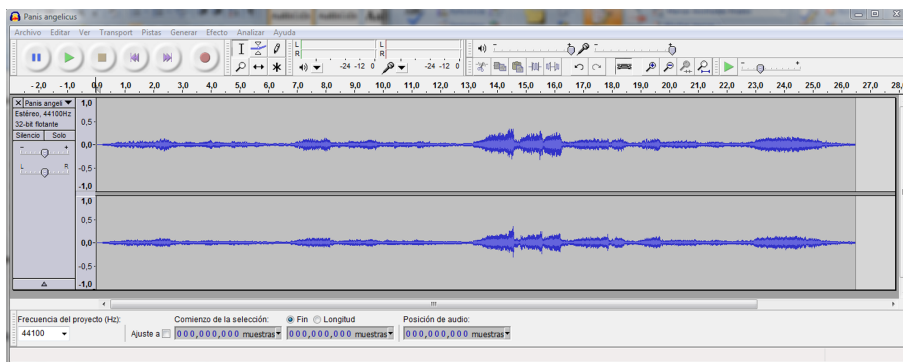
Como en la actividad anterior utilizaremos principalmente el programa Audacity para visualizar, analizar y escuchar las diferentes señales que nos servirán de guía para comprender los conceptos que queremos ilustrar.

2.3.1. La retícula de muestreo

1) Para empezar abriremos con el programa Audacity el fichero **panisAngelicus.wav**. Este fichero se proporciona como recurso adicional para esta actividad y consiste en un breve fragmento de una interpretación del *Panis Angelicus* por Kiri Te Kanawa. Al abrir el fichero observamos la forma de onda de los dos canales estereofónicos. Debido a que la escala del eje temporal incluye todo el fragmento musical, no pueden observarse muchos detalles de la forma de onda. No obstante, con esta resolución visual tenemos una idea general de cómo evoluciona en el tiempo el nivel de la señal.

2) Escuchad el audio para familiarizaros con su contenido y la localización en el tiempo de las frases musicales. En la figura 31 se muestra esta pantalla inicial.

Figura 31. Muestra de la pantalla inicial al abrir el fichero panisAngelicus.wav con el programa Audacity



Vamos a hacer un zoom de la señal para ver los detalles de su forma de onda.

Para activar el zoom debe seleccionarse un fragmento de audio y posteriormente ejecutar los menús **Ver/Ampliar la selección** (Ctrl E).

3) Obsérvese en pantalla el resultado y compruébese que aunque ahora el detalle con el que vemos la forma de onda de la señal ha mejorado considerablemente, todavía no podemos apreciar los cruces por cero y otras peculiaridades de la señal.

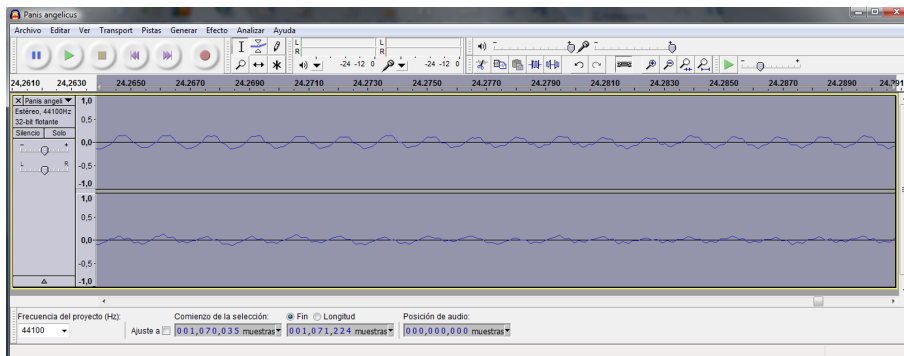
4) Ampliad el factor de zoom hasta que se pueda visualizar un fragmento de señal compuesto por unos 18-20 ciclos de una forma de onda cuasi periódica. En la figura 32 se muestra esta ampliación de la señal donde se ve que la forma

Sugerencia

Se sugiere hacer un zoom de la zona de audio entre el segundo 24 y el 25.

de onda es una repetición de una onda básica. En este ejemplo concreto se muestran un total de 18 repeticiones (ciclos) de la señal básica. Si probáis a ampliar otros fragmentos de la señal original, veréis que tienen características parecidas, cambiando la forma de onda del ciclo básico y el periodo de repetición.

Figura 32. Ampliación de un fragmento de la señal



En este ejemplo hemos reducido la escala de visualización hasta poder observar el carácter cuasi periódico de la forma de onda. En la señal superior podemos contar un total de 18 periodos fundamentales.

Obsérvese que con la resolución actual la señal tiene el aspecto de una señal analógica, es decir, una señal cuya amplitud es continua y que toma valores para todo instante de tiempo. En efecto, la señal que tenemos en pantalla puede representar perfectamente la evolución en el tiempo de la tensión obtenida en la salida de un micrófono. A su vez, esta tensión representa las variaciones del nivel de presión acústica existentes en la membrana del micrófono.

La característica principal de las señales analógicas es que pueden representarse mediante funciones que evolucionan de forma continua en el tiempo.

Señales analógicas

Un matemático diría que una señal analógica es una función real de variable real, aunque para nosotros será suficiente el concepto de **función continua** que evoluciona en el tiempo.

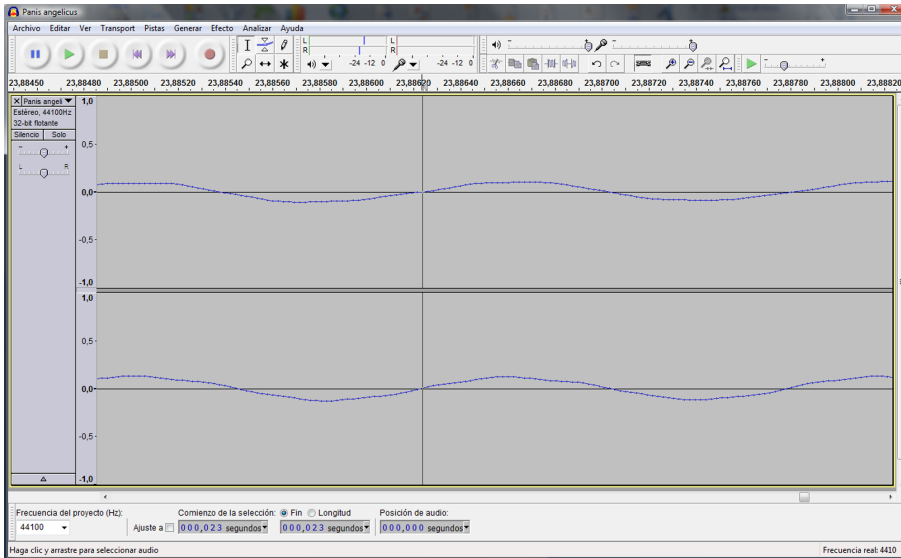
5) Siguiendo con nuestro ejemplo, ampliemos algo más el valor del zoom de la señal hasta visualizar sólo unos 2 o 3 ciclos de la misma. Observaremos ahora que en la forma de onda aparecen marcadas unas rayas que nos indican el carácter digital de la señal. En efecto, las rayas se corresponden con los valores de la señal que realmente están disponibles en el archivo WAV original que hemos introducido.

La consecuencia más importante de este resultado es que la información proporcionada en el archivo digital original sólo contiene los valores de la señal en unos instantes de tiempo determinados: los **instantes de muestreo**.

En la figura 33 se reproduce un resultado parecido al que deberíamos haber obtenido hasta este momento en la actividad. Es importante notar que las rayas están distribuidas uniformemente a lo largo del eje temporal (horizontal), lo que nos indica que el muestreo de la señal se realiza a intervalos de tiempo regulares y que denominaremos como el **periodo de muestreo**.

El inverso del periodo de muestreo es la **frecuencia de muestreo** y nos indica el número de muestras que se toman por segundo.

Figura 33. Representación de un pequeño fragmento del archivo donde se observa que la señal digital está formada por un conjunto de muestras tomadas a intervalos de tiempo constantes



Actividad

Determinad el número de muestras que se han tomado para un periodo de repetición básico de la señal de audio que tiene representada en la pantalla (para facilitar la cuenta, observad los cruces por cero de la señal).

Sabiendo que esta señal ha sido muestreada a una frecuencia de 44.100 Hz, determinad el periodo y la frecuencia de la forma de onda que tenéis representada en la pantalla.

Solución

El resultado depende del fragmento de señal que se haya ampliado. Si el fragmento corresponde al sonido entre el segundo 24 y 25, las muestras de un periodo fundamental deberían ser de unas 67 o 68 muestras.

Para el caso anterior (pongamos un periodo de 68 muestras) el periodo de repetición de la señal será:

$$P = 68 \times \frac{1}{f_m} = 68 \times \frac{1}{44.100} = 0,00154195 = 1,5419 \text{ ms}$$

La frecuencia fundamental de la señal en este fragmento será el inverso del periodo: $F = 648 \text{ Hz}$.

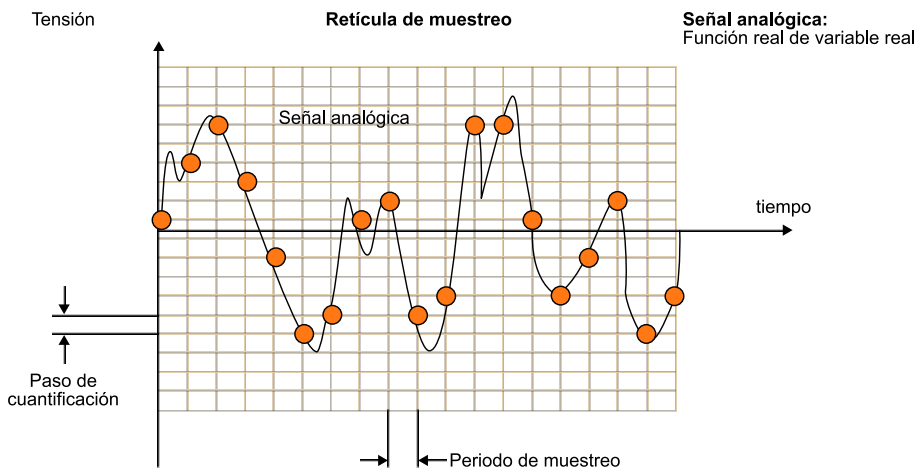
Una señal digital sólo contiene información en los instantes de muestreo. Estos instantes de muestreo se realizan a intervalos de tiempo constantes, que se conocen como el periodo de muestreo. La frecuencia de muestreo es el inverso del periodo de muestreo y nos indica el número total de muestras que se toman por segundo.

Por otra parte, si nos centramos en los valores de amplitud que toma la señal en la gráfica, veremos que no son posibles todos los valores sino que existen unos niveles predeterminados.

Esto no es sencillo de ver, ya que el número total de niveles es muy elevado y las diferencias entre un nivel y el siguiente pueden ser muy pequeñas. En definitiva, en el eje horizontal (tiempo) la señal sólo toma valores en unos instantes de tiempo concretos, mientras que en el eje vertical la señal sólo puede tomar unos niveles de amplitud predeterminados. Esta división de los ejes horizontal y vertical da lugar al concepto de **retícula de muestreo**.

Para clarificar estos conceptos representamos esquemáticamente la retícula de muestreo en la figura 34. En esta gráfica se representa una señal analógica y la retícula de muestreo que acabamos de describir. La retícula de muestreo se caracteriza por sus pasos, tanto en el eje horizontal como en el eje vertical. Los valores de estos pasos se conocen como **periodo de muestreo** (horizontal) y **paso de cuantificación** (vertical). La señal digital son los puntos de la retícula de muestreo que más se aproximan a la señal analógica y que han sido marcados en círculos rojos en la figura 34. Resulta evidente que la señal digital sólo es una aproximación a la señal analógica y que la calidad de esta aproximación depende de la densidad de la retícula de muestreo. Así pues, cuanto mayor sea la densidad de muestras y de pasos de cuantificación mejor será la aproximación digital a la señal analógica.

Figura 34. Aproximación de una señal analógica mediante una retícula de muestreo



La retícula de muestreo está caracterizada por el paso horizontal (periodo de muestreo) y el paso vertical (paso de cuantificación).

La observación anterior sugiere una pregunta:

¿Existe un número ideal para el periodo de muestreo y para el paso de cuantificación? Esta es la pregunta clave que intentaremos resolver en esta actividad y en la siguiente.

De momento, en esta actividad nos centraremos en analizar el periodo de muestreo o, lo que es lo mismo, el número de muestras por segundo que deben tomarse de una señal (frecuencia de muestreo). Veremos que la frecuencia de muestreo depende de las características en frecuencia de la señal y que son estas características las que determinan el número mínimo de muestras que deben considerarse. En la actividad siguiente estudiaremos los efectos del paso de cuantificación en la calidad de la señal digital.

2.3.2. La frecuencia de muestreo

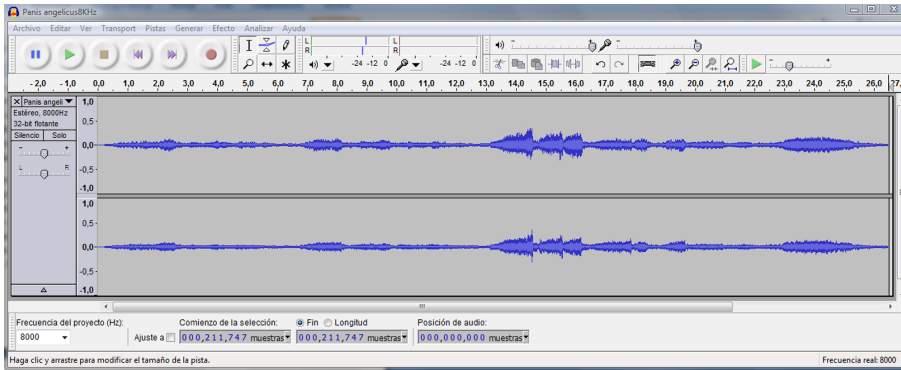
En este apartado vamos a experimentar con la frecuencia de muestreo y estudiar cuáles son los requisitos que deben cumplirse para realizar la correcta digitalización de una señal.

Llegaremos a establecer el teorema del muestreo que nos permitirá conocer el número mínimo de muestras por segundo que son necesarias para digitalizar correctamente una señal, sin que puedan observarse diferencias significativas con la señal analógica original. En todo este apartado y en el resto de esta actividad vamos a suponer que el número de niveles de cuantificación es suficientemente elevado como para que no influya en la calidad de la señal.

1) Empecemos cargando en Audacity las señales **panisAngelicus.wav** y **panisAngelicus8kHz.wav**.

- La primera es una señal en formato WAV (audio no comprimido, cuyo formato se estudiará en actividades posteriores), tomada directamente del CD original y muestreada a 44.100 Hz.
- La segunda señal se ha obtenido utilizando una frecuencia de muestreo de 8 kHz. Vemos que no es posible apreciar diferencias a simple vista entre las dos formas de onda. No obstante, si ajustamos los marcadores de posición del cursor a muestras (parte inferior del menú, ver figura 35) comprobaremos que mientras que uno de los fragmentos tiene un total de 1.168.000 muestras (26,5 s x 44.100 muestras/s) la otra señal tiene sólo unas 212.000 muestras (26,5 s x 8.000 muestras/s).

Figura 35. Forma de onda de la señal panisAngelicus muestreada a 8 kHz



En las ventanas inferiores se ha seleccionado el número de muestras para visualizar la selección y la posición del cursor.

Ahora vamos a escuchar con cierta atención el audio de cada una de las dos señales.

2) Escuchad los dos fragmentos y observad las diferencias de calidad entre las dos versiones.

Nota

Las frecuencias de muestreo se han elegido para que resulte bastante evidente la diferencia de calidad entre las dos señales. No obstante, en un ordenador portátil o en equipos con audio de baja calidad es posible que no resulte fácil discriminar entre las dos señales. En este caso, se sugiere que se utilicen auriculares.

3) Intentad expresar cualitativamente y anotad en un papel las diferencias de calidad que apreciáis entre las dos señales.

Vamos a repetir este experimento con otra señal de audio.

1) Ahora abrimos con Audacity los ficheros **sintoniaOpera-22050.wav**, **sintoniaOpera-16.000.wav** y **sintoniaOpera-8000.wav**. En este caso, se trata de la sintonía de un programa de radio difundido a través de Internet. La calidad de partida de estos archivos es considerablemente inferior a la grabación del primer fragmento.

2) Como en el caso anterior podemos comprobar que las frecuencias de muestreo son de 22.050 Hz, 16.000 Hz y 8.000 Hz y que el número total de muestras de los ficheros coincide con lo esperado. Escuchad atentamente las tres señales e intentad expresar cualitativamente las diferencias que observáis.

Los principales resultados que podemos extraer de las audiciones anteriores son:

1) La calidad de la señal **panisAngelicus** muestreada a 44.100 Hz es considerablemente superior a la muestreada a 8.000 Hz. En esta última se aprecia una pérdida considerable de agudos y brillantez.

- 2) Las dos señales **sintoniaOpera** muestreadas a 22.050 Hz y 16.000 Hz, en lo que se refiere a calidad, son indistinguibles.
- 3) La señal **sintoniaOpera** muestreada a 8.000 Hz representa una pérdida de calidad considerable respecto a las otras dos.

En resumen, uno de los efectos que puede tener la reducción de la frecuencia de muestreo es una pérdida de calidad que se percibe como una disminución de los agudos (casos 1 y 3).

No obstante, en algunas circunstancias no es posible observar ninguna mejora o pérdida cuando se realiza un cambio de frecuencia de muestreo (caso 2). Para cuantificar y analizar con mayor precisión qué está ocurriendo y por qué en algunos casos apreciamos diferencias y en otros casos no, es conveniente que nos centremos en el análisis en frecuencia de las señales.

2.3.3. Análisis en frecuencia de panisAngelicus

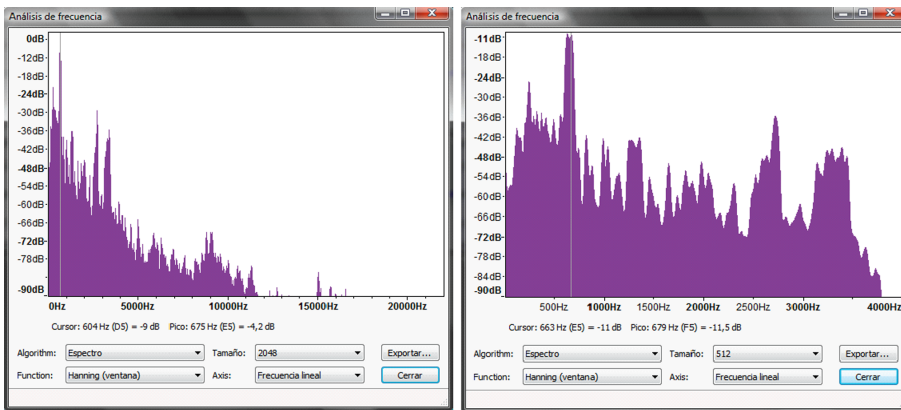
Vamos a comparar la representación en frecuencia de un fragmento de la señal panisAngelicus muestreada a 44.100 Hz con la representación del mismo fragmento de la señal muestreada a 8.000 Hz.

1) Para calcular la representación en frecuencia (transformada de Fourier) debemos seleccionar un fragmento de señal. En este caso, sugerimos que se tome el fragmento entre el segundo 23 y el segundo 24, que se corresponde con una nota de una frecuencia fundamental aproximada de 675 Hz.

2) Seleccionad el fragmento de 1 segundo de duración con el cursor y ejecutad las opciones **Analizar** y **Análisis de espectro**. Elegid una representación del espectro, con una ventana de **Hanning** de **2.048** muestras. Observad el resultado obtenido y repetid las mismas operaciones para el fichero muestreado a 8.000 Hz. En este caso elegid un tamaño de la ventana de 512 muestras⁴. Los resultados obtenidos se representan en la figura 36, donde se observa un cambio muy significativo del espectro de las dos señales. En efecto, el espectro de la señal muestreada a 44.100 Hz se extiende más allá de los 10.000 Hz, incluso con algunas componentes visibles alrededor de los 15.000 Hz. En cambio, el espectro de la señal de 8.000 Hz queda limitado a componentes inferiores a los 4.000Hz, por lo que ha perdido gran parte de las componentes de alta frecuencia de la primera señal. También es importante observar que el contenido frecuencial de las dos señales entre 0 y 4 kHz es aproximadamente igual.

⁽⁴⁾Los tamaños de las dos ventanas son diferentes debido a que queremos que los intervalos temporales en el que se calculan las dos transformadas sean aproximadamente iguales. Para que el periodo de tiempo considerado sea aproximadamente el mismo, tendremos que tomar más muestras de la señal con una frecuencia de muestreo alta que de la señal con una frecuencia de muestreo baja. Ved la actividad de audio 1 para interpretar el significado de la duración de las ventanas.

Figura 36. Comparativa entre los espectros de la señal panisAngelicus muestreada a 44.100 Hz y la muestreada a 8.000 Hz



En la imagen de la derecha el contenido espectral está limitado a frecuencias inferiores a los 4 Hz.

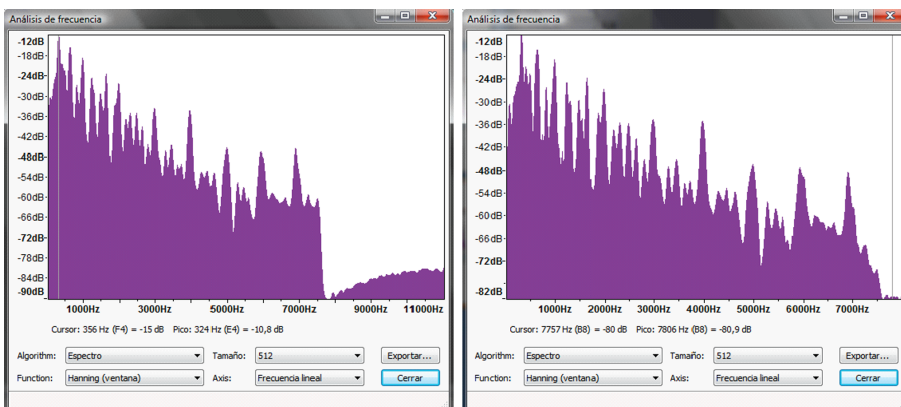
2.3.4. Análisis en frecuencia de sintoniaOpera

En este apartado repetimos las pruebas anteriores pero ahora con la otra señal. Hemos tomado el fragmento entre 1 s y 3 s en todos los casos para realizar el análisis en frecuencia.

En la figura 37 puede verse una comparativa entre la señal muestreada a 22.050 Hz y la muestreada a 16.000 Hz. En este caso, resulta evidente que las dos señales tienen prácticamente el mismo contenido en frecuencia. En efecto, el contenido entre 0 y 8 kHz de las dos señales es exactamente igual. La señal de 22.050 Hz tiene unas componentes frecuenciales muy bajas (muchos dB por debajo del resto), centradas alrededor de los 11 kHz. No obstante, debido a su reducida potencia, estas componentes no son audibles⁵.

⁽⁵⁾Seguramente, estas componentes de reducida potencia proceden de algún procesado realizado en la emisora de radio (o en otro sistema) para cambiar la frecuencia de muestreo a 22.050 Hz.

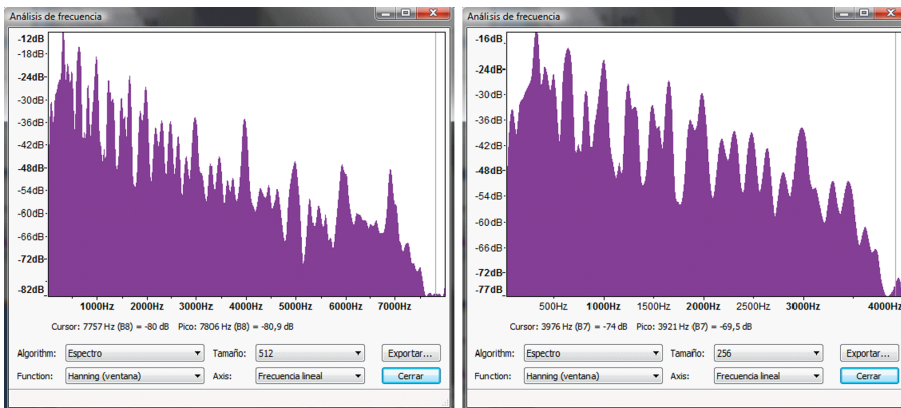
Figura 37. Espectros de las señales sintoniaOpera muestreadas a 22.050 y 16.000 Hz (fragmento de 2 segundos de duración)



Los espectros de las dos señales son idénticos en la banda entre 0 y 8.000 Hz. Se observan algunas componentes adicionales, entre 9.000 Hz y 11.000 Hz, para la señal de mayor frecuencia de muestreo. No obstante, debido a su baja potencia estas componentes no resultan audibles.

Finalmente, en la figura 38 también se comparan los espectros entre la señal sintoniaOpera muestreada a 16.000 Hz y la muestreada a 8.000 Hz. En este caso, se observa que la última señal sólo tiene componentes frecuenciales hasta 4.000 Hz, por lo que es lógico que desde un punto de vista auditivo represente una pérdida de brillantez y de componentes de alta frecuencia.

Figura 38. Comparativa entre los espectros de la señal sintoníaOpera muestreada a 16.000 Hz (izquierda) y 8.000 Hz (derecha)



La señal de la derecha ha perdido todas las componentes frecuenciales a partir de los 4.000 Hz.

Así pues, desde un punto de vista experimental e intuitivo parece que podemos obtener las siguientes conclusiones sobre la frecuencia de muestreo de una señal.

Cuando una señal se muestrea a F muestras por segundo, entonces todas las componentes frecuenciales superiores a $F/2$ no están presentes.

Así, por ejemplo, cuando la frecuencia de muestreo sea de 32.000 Hz, todas las componentes frecuenciales superiores a 16.000 Hz no estarán presentes.

Si una señal sólo tiene componentes frecuenciales significativas hasta una frecuencia W , no obtenemos ninguna mejora auditiva apreciable si muestreamos a frecuencias superiores a $2W$.

Por ejemplo, si una señal solo tiene componentes en frecuencias hasta los 4.000 Hz, entonces no tiene sentido muestrearla a más de 8 kHz. La calidad que obtendremos a 16 kHz, 30 kHz o 8 kHz es la misma.

2.3.5. Frecuencias de muestreo para señales de voz

Antes de empezar a formalizar los conceptos del teorema del muestreo se propone que realice la siguiente experiencia:

1) Ajustad el volumen de grabación del micrófono en el programa Audacity y proceded a grabar vuestra voz a diferentes frecuencias de muestreo. Para realizar el registro se recomienda repetir constantemente una misma palabra (p. ej. *probando, probando*), durante unas 4 o 5 veces y posteriormente comparad la calidad de la audición utilizando diferentes frecuencias de muestreo.

2) Ajustad los niveles en la forma correcta para que el nivel de grabación sea el correcto, sin que se produzca distorsión por saturación de los amplificadores.

Repetid la grabación para varias frecuencias de muestreo, empezando por 48.000 Hz y reduciendo a 44.100 Hz, 22.050 Hz, 16.000 Hz y 8.000Hz. Para cada uno de los registros, seleccionad un fragmento de audio (la palabra que habéis repetido) y determinad su espectro.

Justificad razonadamente cuál es la frecuencia de muestreo idónea para la señal de voz⁶.

En las secciones siguientes vamos a formalizar algunos conceptos relacionados con el muestreo de señales. En concreto, vamos a definir el concepto de señales de banda limitada y establecer el teorema del muestreo. También introduciremos el concepto de aliasing.

⁽⁶⁾El resultado final depende de la calidad del sistema de grabación: micrófono incorporado a la estación de trabajo, micrófono externo, calidad del micrófono, calidad de la tarjeta de adquisición.

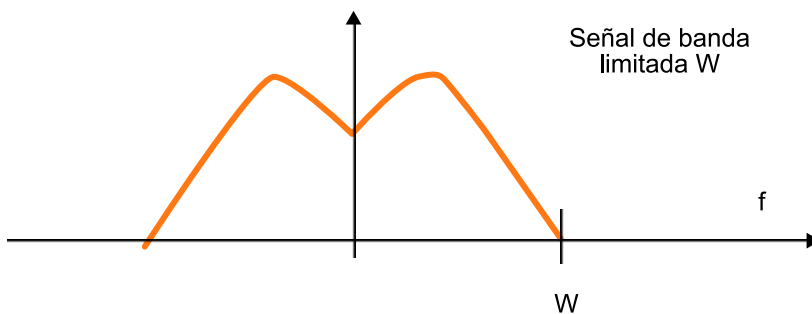
2.3.6. Señales de banda limitada

Se dice que una señal es de **banda limitada** cuando su contenido espectral es nulo a partir de una determinada frecuencia W .

En la figura 39 se representa esquemáticamente el espectro de una señal de banda limitada.

La frecuencia que actúa como límite del contenido espectral de la señal se denomina **ancho de banda**.

Figura 39. Señal de banda limitada con un ancho de banda W



Las señales de banda limitada son muy importantes en los sistemas digitales; como veremos, el teorema del muestreo establece que sólo puede realizarse un muestreo sin pérdida de información cuando la señal es de banda limitada. Las señales de audio y de vídeo son de banda limitada, por lo que admiten representaciones digitales.

En el caso de las **señales de audio**, es el propio oído humano el que establece el límite de ancho de banda de la señal. La capacidad para la detección de **frecuencias** y **tonos** depende de cada individuo y de la edad, pero está establecido entre los 15 KHz y los 20 KHz. Las señales con una frecuencia superior

a los 20 KHz pueden considerarse inaudibles para el oído. Por ello, la señal de audio únicamente requiere disponer de contenido espectral hasta esta frecuencia máxima para que pueda ser reproducida con total fidelidad.

En telefonía analógica convencional se considera que el ancho de banda de la señal de voz es de unos **3,4 KHz**. Este ancho de banda proporciona suficiente información como para que los mensajes telefónicos puedan ser correctamente interpretados y pueda reconocerse al interlocutor. El hecho de que en telefonía no se utilice todo el ancho de banda de la señal de audio está ligado a un problema tecnológico y económico para la optimización de los recursos de ancho de banda. En efecto, cuanto menor es el ancho de banda de cada una de las señales individuales que deben encaminarse a lo largo de la red telefónica mejor podrán aprovecharse los recursos, transmitiendo más canales en paralelo (multiplexación) por las mismas conexiones físicas.

En la práctica no resulta evidente determinar el ancho de banda de una señal. Cuando analizamos el espectro siempre aparece un **ruido de fondo** que dificulta la identificación de las componentes de espectro que corresponden a la señal y las que corresponden al ruido.

Pueden elaborarse diferentes reglas genéricas para calcular el ancho de banda de una señal a partir de la visualización de su espectro aunque todas estas reglas son empíricas y tienen excepciones.

Una regla utilizada con cierta frecuencia es dar como ancho de banda de la señal la frecuencia a partir de la cual todo el espectro está X dBs por **debajo de su valor máximo**. El valor de X depende de las características de nuestro sistema de adquisición.

Actividad

Teniendo en cuenta la regla anterior, determinad el ancho de banda de las señales de voz digitalizadas en el apartado anterior (considerando que habéis utilizado un sistema de captura doméstico, tomad el valor de $X = 60$ dBs).

2.3.7. Teorema del muestreo

El teorema del muestreo establece que cuando una señal es de banda limitada puede muestrearse sin que se produzcan pérdidas de información utilizando una frecuencia de muestreo mayor que el doble de su ancho de banda.

$$f_{\text{muestreo}} = \frac{1}{T_{\text{muestreo}}} > 2 \cdot W$$

Variación de X dBs

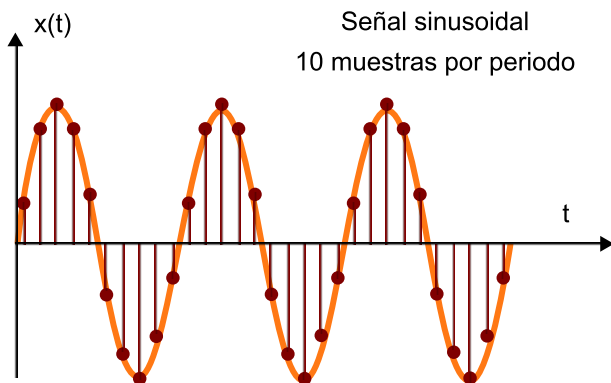
Si tenemos un sistema de adquisición económico con bastante ruido, X puede tomar el valor de 50 o 60 dBs. En sistemas de adquisición de estudios profesionales podemos estar alrededor de los 85-100 dBs.

Esta relación entre la frecuencia de muestreo y el ancho de banda de la señal también se conoce con el nombre de **criterio de Nyquist**.

Así, por ejemplo, si una señal tiene un ancho de banda de 15 kHz, la frecuencia mínima a la que debe muestrearse es el doble, es decir, 30 kHz.

De acuerdo con este teorema, la frecuencia de muestreo mínima para poder trabajar con señales de audio de alta fidelidad estaría situada por encima de los 40 KHz (20 KHz de ancho de banda). En la figura 40 se ilustra el proceso de muestreo sobre una componente sinusoidal.

Figura 40. Muestreo de una componente sinusoidal con una frecuencia de muestreo que verifica el criterio de Nyquist



En este ejemplo se toman un total de 10 muestras por periodo, por lo que se verifica sin ningún tipo de problema el teorema del muestreo. En efecto, al tomar 10 muestras por periodo, la frecuencia de muestreo es 10 veces mayor que la frecuencia de la señal, por lo que cumplimos perfectamente con el teorema de Nyquist. En la figura 40 vemos de forma intuitiva que el número de muestras es suficientemente elevado como para poder reconstruir la forma de la señal original a partir de sus muestras. Esencialmente, el teorema de Nyquist nos dice que podremos realizar esta reconstrucción siempre que tomemos más de dos muestras por periodo (frecuencia de muestreo superior al doble del ancho de banda de la señal).

2.3.8. Aliasing y filtros antialiasing

Para garantizar que el proceso de muestreo de una señal analógica es correcto, debemos asegurar que se cumple el criterio de Nyquist, esto es, que la frecuencia de muestreo que utilizamos es **mayor que el doble del ancho de banda**. En la práctica, para garantizar que esto se cumple se hace una especie de doble chequeo.

- En efecto, si queremos muestrear por ejemplo una señal de audio de alta fidelidad (20 kHz de ancho de banda) elegiremos una frecuencia de muestreo que sea como mínimo 2 veces mayor al ancho de banda.
- La segunda parte de este doble chequeo es que debemos garantizar que bajo ninguna circunstancia tenemos señales que no cumplen con el cri-

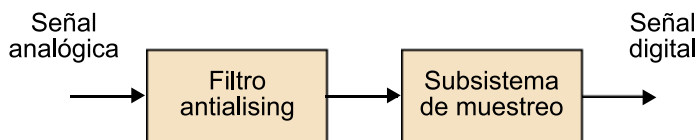
44.100 Hz

En audio digital, una frecuencia de muestreo muy utilizada es **44.100 Hz**, que cumple con el teorema de muestreo.

terio de Nyquist. Para ello, una vez hemos seleccionado la frecuencia de muestreo, pondremos un sistema que se encarga de eliminar en la señal las posibles componentes frecuenciales que no cumplan con el teorema de muestreo. Este sistema se denomina **filtro antialiasing** y su función es básicamente la de eliminar las posibles componentes de señal cuya frecuencia sea mayor que la mitad de la frecuencia de muestreo (22.050 Hz en nuestro ejemplo).

En la práctica, si nuestra primera hipótesis sobre las características de la señal es correcta, el filtro antialiasing no actuará. No obstante, siempre es conveniente ponerlo debido a que en un sistema real podemos tener, además de la señal deseada, interferencias o ruidos no deseados cuyas frecuencias pudieran ser superiores al límite marcado por el teorema de muestreo. En la figura 41 se representa un diagrama de bloques en el que se indica la presencia del filtro antialiasing precediendo al subsistema de muestreo de la señal.

Figura 41. Filtrado paso bajo previo al muestreo de una señal analógica

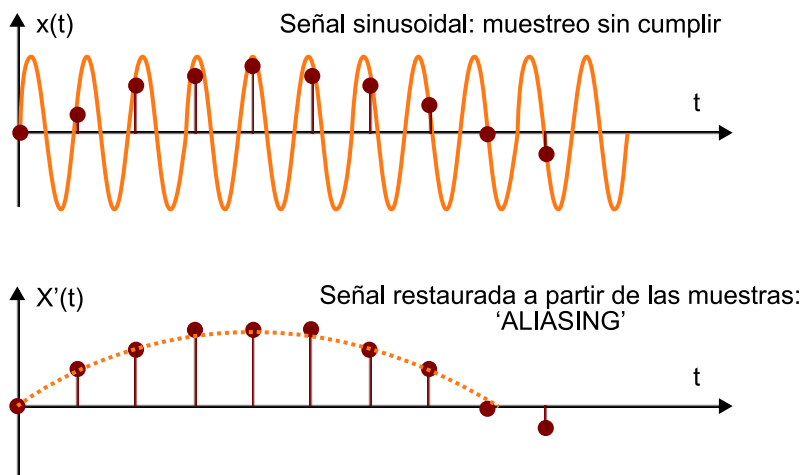


El filtro antialiasing impide la aparición del fenómeno de **aliasing**, que consiste en que las componentes de la señal con una frecuencia superior a la mitad de la frecuencia de muestreo (las que no cumplen el criterio de Nyquist) aparecen, una vez muestreadas, como componentes de señal con frecuencias falsas, dentro del ancho de banda de la señal útil.

El fenómeno del aliasing se ilustra de forma muy simplificada en la figura 42. En este ejemplo se representa una componente sinusoidal y sus muestras, utilizando una frecuencia de muestreo que no verifica el criterio de Nyquist (hemos visto que para sinusoides deberíamos tomar un mínimo de dos muestras por periodo). De forma intuitiva, la señal que podremos reconstruir a partir de las muestras es la que obtenemos a partir de la unión de los valores (puntos) de la señal en los instantes de muestreo.

Esto da lugar a la aparición de una componente sinusoidal, con una frecuencia mucho más baja, que no existía en la señal original y que, por ello, recibe el nombre de **alias**. El problema es más importante de lo que en un principio puede parecer debido a este cambio de frecuencia en las componentes de la señal, que pueden aparecer como interferencias, es decir, como señales que previamente no existían.

Figura 42. Aparición del fenómeno del aliasing cuando la frecuencia de muestreo no cumple el criterio de Nyquist



Podríamos pensar que la estructura de muestreo de la figura 42 sólo es necesaria cuando no podamos asegurar que la señal es de banda limitada. Así, en el caso de señales de audio podríamos prescindir del filtro antialiasing siempre que garantizáramos que la frecuencia de muestreo supera los 40 KHz, con lo que garantizamos que nuestras señales verifican el criterio de Nyquist. Sin embargo, el problema no es tan sencillo puesto que nada impide asegurar que la fuente de señal o los propios amplificadores reciban interferencias de frecuencias superiores a los 20 KHz.

Es importante que nos demos cuenta de que estas frecuencias no resultan audibles mientras las señales se mantengan en el dominio analógico pero que cuando realizamos un muestreo, debido al aliasing, pueden cambiar de frecuencia entrando en la banda audible.

Por ello, el filtro antialiasing es estrictamente necesario en todos los casos antes de proceder al muestreo de la señal.

2.3.9. Algunos ejemplos de aliasing

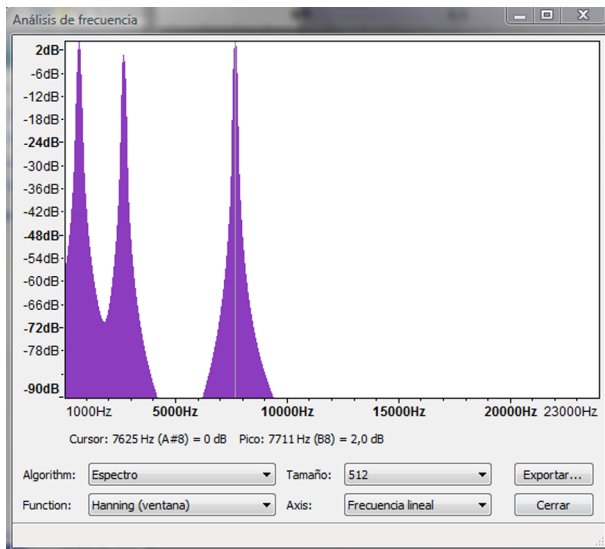
Para comprender mejor el problema y los efectos del aliasing en señales de audio parece aconsejable realizar algunas audiciones que nos evidencien sus efectos.

En esta sección vamos a escuchar y analizar con el programa Audacity algunos ficheros de audio que nos intentarán mostrar los efectos del aliasing. Empezaremos las pruebas con señales generadas artificialmente, utilizando tonos y modulaciones, en las que el aliasing se presenta de forma muy evidente. Posteriormente, utilizaremos archivos musicales en los que el aliasing se presenta de una forma menos clara.

1) Empezaremos con una señal sintética formada por tonos de diferentes frecuencias y que ha sido digitalizada con diferentes frecuencias de muestreo. La señal original es una superposición de tonos con modulación de amplitud y con una frecuencia máxima de 7.700 Hz. La señal **Orig_Tones.wav** está muestreada a **48.000 Hz** por lo que cumple el criterio de Nyquist.

2) Seleccionad un fragmento de un segundo de la señal y elaborad su análisis en frecuencia. El espectro de la señal se muestra en la figura 43, donde se observa que está compuesto por tres tonos en las frecuencias aproximadas de 700 Hz, 2.700 Hz y 7.700 Hz. Como la frecuencia de muestreo es considerablemente superior al doble de la máxima frecuencia, no observaríamos diferencias significativas entre una señal analógica compuesta por estos tres tonos y la señal digital correspondiente.

Figura 43. Espectro de la señal sintética **Orig_Tones.wav**, compuesta por 3 tonos con frecuencias de 700 Hz, 2.700 Hz y 7.700 Hz muestreada a 48.000 Hz, cumpliendo por tanto con el teorema de muestreo



El siguiente paso es analizar cuál es la señal que obtendríamos si se digitalizase la secuencia original con una frecuencia inferior a la establecida por el criterio de Nyquist. En este caso, vamos a suponer que muestreamos la señal anterior a una frecuencia de **8.000 Hz** y utilizando el esquema representado en la figura 41 con el filtro antialiasing. En este caso, como la frecuencia de muestreo es de 8.000 Hz, el filtro antialiasing eliminará todas las componentes de señal que estén situadas por encima de 4.000 Hz. La señal resultante está en el fichero **Tones_8K_antialias.wav**.

1) Abrid la señal con Audacity, escuchadla, seleccionad un fragmento de aproximadamente 1 segundo y representad su espectro (que hemos representado en la figura 44). A partir de este experimento, es evidente que:

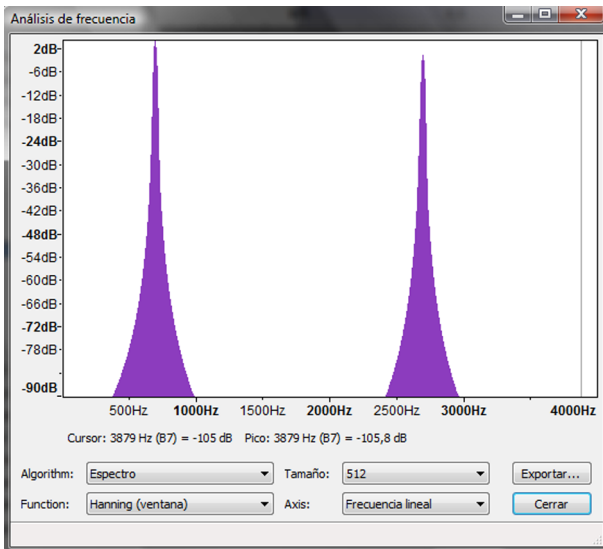
a) La señal muestreada a 8 KHz, con filtro antialiasing, carece de la componente a 7.700 Hz, ya que ha sido filtrada.

Precaución

Puede escuchar esta señal con Audacity pero tenga la precaución de reducir el volumen de los auriculares hasta un nivel adecuado ya que la señal es una superposición de tonos puros modulados y puede resultar muy molesta al oído.

b) Al escuchar esta señal sólo oímos la superposición de los tonos a 700 Hz y 2.700 Hz. Los dos tonos estaban presentes en la señal original. Por tanto, esta señal es la reproducción más fiel posible de la señal original, ya que contiene toda la parte del espectro que puede reproducirse correctamente con la frecuencia de muestreo actual y carece de las componentes que no pueden ser reproducidas correctamente.

Figura 44. Espectro de la señal sintética muestreada a 8 KHz utilizando un filtro antialiasing



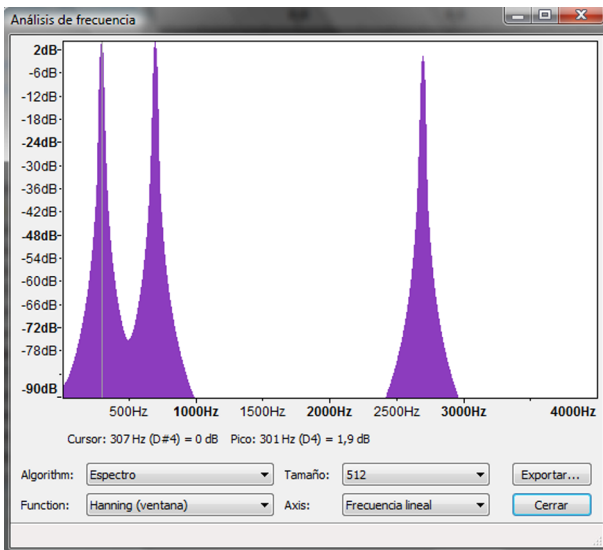
La componente a 7.700 Hz ha sido eliminada por el filtro.

Finalmente, vamos a analizar el resultado que se obtiene cuando la señal original se muestrea a 8 KHz **sin utilizar el filtro antialiasing**.

- 1) Cargad la señal **Tones_8K_NOAntialias.wav** en Audacity,
- 2) Seleccionad un fragmento de aproximadamente 1s y determinad el espectro de la señal, cuyo resultado se representa en la figura 45.
- 3) Escuchad esta señal y comparadla con la original y con la que se obtiene utilizando el filtro antialiasing.

Está claro que si no utilizamos filtro antialiasing aparece una nueva componente espectral en la señal que no estaba presente en la señal original. En nuestro caso, la nueva componente espectral aparece en la frecuencia de 300 Hz. Esta componente es debida al muestreo incorrecto de la componente de 7.700 Hz. Observad que la nueva frecuencia depende tanto de la frecuencia del tono original como de la frecuencia de muestreo de la señal: $300 \text{ Hz} + 7.700 \text{ Hz} = 8.000 \text{ Hz}$. Desde el punto de vista auditivo, la nueva componente produce un sonido que no tiene ninguna relación con el original.

Figura 45. Espectro de la señal sintética muestreada a 8 KHz sin utilizar un filtro antialiasing



Aparece una componente en la frecuencia de 300 Hz que no estaba presente en la señal original.

Como consecuencia directa de esta experiencia, deberíamos concluir que siempre que se digitaliza una señal a una frecuencia menor que el doble de su ancho de banda (criterio de Nyquist) se aprecia una **pérdida de calidad auditiva**.

No obstante, esta pérdida de calidad es especialmente grave si no se utiliza un filtro antialiasing, ya que pueden aparecer componentes frecuenciales que no estaban presentes en la señal original.

El efecto auditivo del aliasing en señales sintéticas es claro, pero resulta bastante más complejo de apreciar en **señales musicales**. Vamos a poner algunos ejemplos de aliasing con señales musicales que pueden escucharse y analizarse utilizando Audacity.

El primer ejemplo es con el fragmento de audio del *Panis Angelicus* que ya hemos utilizado en esta actividad.

1) Cargad los archivos **PA_48K.wav**, **PA_8K.wav** y **PA_8K_NOAntialias.wav**. El primero es el fragmento musical muestreado a 48.000 Hz lo que sirve como referencia de calidad. Los otros dos fragmentos están muestreados a 8.000 Hz. El primero de ellos ha sido muestreado utilizando un filtro antialiasing mientras que para el segundo no se ha utilizado ningún filtro.

2) Escuchad atentamente las dos señales. Apreciar el aliasing que se produce en la segunda señal requiere bastante atención. Debería apreciarse una especie de 'rasqueo', de baja frecuencia que aparece en algunos fragmentos. Se trata de un ruido que no existía en la señal original y cuyo contenido espectral varía en función de las propias características espectrales de la señal original, por lo

que en algunos fragmentos aparece de forma muy clara mientras que en otros no puede apreciarse. Es importante observar que la señal muestreada con el filtro antialiasing no tiene este ruido y se oye con mayor claridad.

Evidentemente, representa una pérdida considerable respecto a la señal original debido a que buena parte de la alta frecuencia ha sido eliminada.

Se presenta otro ejemplo en los ficheros `srv.wav`, `srv_noanti.wav`, `srv_anti.wav`, que se corresponden con un fragmento de un **solo de guitarra eléctrica**. La primera señal está muestreada a 22.050 Hz y puede considerarse como la calidad de referencia. Las otras dos señales se corresponden con frecuencias de muestreo de 4.400 Hz sin filtro antialiasing y con filtro antialiasing. La calidad de estas dos últimas señales es muy baja debido a la drástica reducción de la frecuencia de muestreo. No obstante, tras repetidas audiciones, debería ser capaz de escuchar un ruido de fondo incoherente en la señal obtenida sin el uso del filtro antialiasing y que no está presente en la señal en la que sí se ha aplicado el filtro.

Ejemplo

En los archivos `gg_noanti.wav` y `gg_anti.wav` se proporcionan otros dos ejemplos de muestreo a frecuencias reducidas con el filtro y sin el filtro.

2.3.10. El aliasing en vídeo y cine

El aliasing puede producirse siempre que se realice el muestreo de una señal, por lo tanto, no es exclusivo de las señales de audio. En cine y televisión, la escena también se muestrea en el tiempo a razón de **24 fotogramas por segundo** (cine) o **25 fotogramas por segundo** (televisión en sistema PAL). El aliasing debería ser pues un fenómeno al que estamos bastante acostumbrados.

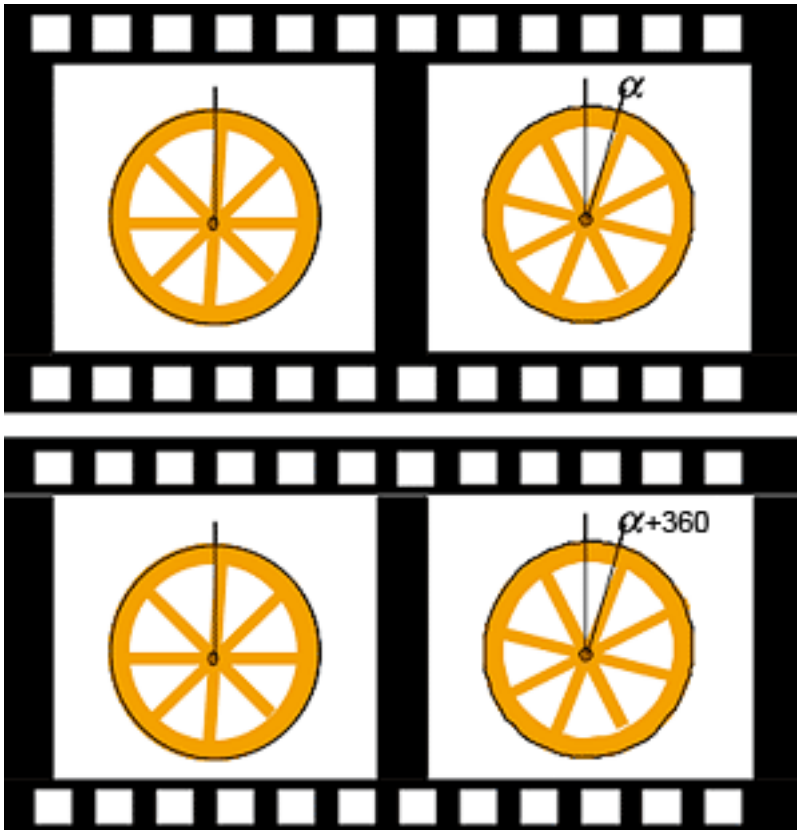
En efecto, uno de los fenómenos que puede introducir el muestreo temporal de la escena es la aparición de **efectos visuales** que falseen o desvirtúen el movimiento real de los objetos. Estos efectos son debidos a que la frecuencia de muestreo elegida resulta ser insuficiente para una correcta interpretación del movimiento. Así, tal y como establece el teorema del muestreo de señales, si los movimientos de los objetos son excesivamente rápidos, pueden existir componentes de alta frecuencia que, al muestrear la escena, se solapan con otras componentes desvirtuando el movimiento real de los objetos.

Uno de los ejemplos de este fenómeno aparece en muchas películas de acción y en ocasiones incluso es provocado por el director para producir en el espectador la sensación de movimiento. La escena típica consiste en centrar la acción de la cámara en la rueda de un carro (en películas del oeste) o de un automóvil. Cuando el vehículo arranca, el espectador observa e interpreta correctamente el movimiento de la rueda, sin embargo, a medida que la velocidad aumenta se produce la sensación de que los radios de la rueda se paran o incluso giran en sentido contrario al de la marcha del vehículo.

Este efecto es debido, como se ilustra en la figura 46, a que el ojo interpreta el movimiento de la forma más simple posible a partir de los fotogramas que se le presentan. En la figura se representan dos fotogramas consecutivos de la rueda del carro que corresponden a dos situaciones distintas pero que, inevitablemente, serán interpretadas del mismo modo, ya que las imágenes que las definen son idénticas. En la secuencia superior, la rueda del carro experimenta

un giro de α grados entre dos fotogramas consecutivos, mientras que en la inferior el giro real de la rueda es de $\alpha+360$ grados. El sistema visual interpretará en ambos casos el mismo movimiento aparente de α grados, ya que la simetría de la rueda no le permite distinguir entre las dos secuencias. En este ejemplo, si la velocidad de giro de la rueda fuera tal que entre dos fotogramas consecutivos se produjera un giro de 360 grados, el espectador no apreciaría ningún movimiento, ya que todas las imágenes serían idénticas. Se produce un caso curioso cuando el ángulo de giro entre fotogramas es ligeramente inferior a los 360 grados, ya que en este caso se interpreta que el movimiento de la rueda es en sentido contrario al real. En este último supuesto, es evidente que para el sistema visual resulta mucho más factible interpretar que se ha producido un pequeño movimiento de la rueda en sentido contrario al real que un movimiento considerable en el sentido correcto.

Figura 46. Aliasing temporal de las imágenes. En ambas situaciones, las secuencias de imágenes obtenidas son idénticas, por lo que el sistema visual interpretará el mismo movimiento



Este efecto visual se conoce como el **efecto estroboscópico** y a veces es provocado para poder observar movimientos de naturaleza periódica de algunos fenómenos físicos cuya rapidez no permite observarlos en condiciones normales.

El instrumento utilizado para ello se denomina **estroboscopio** y consiste en una lámpara (flash) que se ilumina a intervalos regulares y cuya frecuencia y fase de disparo puede ser controlada externamente.

El estroboscopio puede utilizarse, por ejemplo, para observar detalladamente la deformación que experimenta la membrana de un altavoz cuando es sometido a distintas excitaciones periódicas.

Para realizar el experimento debemos alimentar el altavoz con una señal sinusoidal a la frecuencia en la que deseemos hacer el análisis.

El sistema debe situarse en un entorno con poca iluminación, de modo que al disparar el estroboscopio a la misma frecuencia a la que está vibrando el altavoz podremos observar una imagen estacionaria de la deformación de la membrana. Si modificamos la fase de los disparos, veremos las deformaciones progresivas que va tomando la membrana del altavoz en las distintas fases de la vibración. Incluso si seleccionamos una frecuencia de disparo del flash ligeramente distinta a la de la vibración del altavoz podremos ralentizar el movimiento de la membrana a la velocidad deseada.

Se propone que el lector se plantee qué posibles frecuencias del estroboscopio utilizaría para poder observar el movimiento de una membrana que oscila a 300 Hz con una velocidad aparente de 0.1 Hz. Téngase en cuenta que por una parte es conveniente que la frecuencia del estroboscopio permita observar el movimiento sin parpadeo y por otra, que la frecuencia de disparo no sea excesivamente elevada por la limitación propia del flash.

Ejemplo

Experiencias parecidas pueden observarse en diversos museos científicos en los que se demuestran algunas aplicaciones y efectos visuales obtenidos mediante la ayuda del estroboscopio. Una de las más espectaculares consiste en ralentizar la caída libre de las gotas de agua. En este caso, se utiliza una cámara cerrada en la que se provoca un goteo constante de agua. Al iluminar el flujo de gotas de agua mediante un estroboscopio podemos provocar que éste se ilumine siempre que una gota de agua pase por una posición determinada, por lo que observaremos que la gota permanece suspendida en el aire. Una ligera disminución o aumento de la frecuencia de disparo producirá la sensación de que la gota cae o sube lentamente en el medio. Pueden realizarse experimentos similares para visualizar el vuelo de insectos, analizar las vibraciones de máquinas, etc.

2.3.11. Selección de la frecuencia de muestreo en sistemas de audio

El valor de la frecuencia de muestreo en un sistema de audio o voz depende de la calidad final de la señal que desee reproducirse o almacenarse y de la aplicación a la que se destina el sistema. Cuando se definió el sistema **CD-Digital Audio** se tomó como estándar una frecuencia de **44,1 kHz** que permite trabajar con todo el espectro audible (20 Hz a 20 kHz). El valor elegido verifica el criterio de Nyquist para la banda de audio, aunque, como ya hemos comentado, siempre es necesario hacer un filtrado previo de las señales antes

de proceder a su muestreo. Los filtros antialiasing están incorporados en los propios sistemas de digitalización. Así, cualquier grabador multipista para audio profesional incorpora unos filtros antialiasing que cumplen con los requisitos marcados por el estándar de audio digital utilizado.

44.1 kHz

La selección de la frecuencia de 44,1 kHz no es casual y se debe a los primeros sistemas de registro de señales de audio en formato digital que aparecieron en la década de los sesenta. Estos primeros sistemas fueron desarrollados por la BBC, utilizando sistemas de vídeo modificados para almacenar la información de audio en formato digital. El sistema elegido almacenaba 3 muestras de audio digital en el espacio de cinta que debería haber correspondido a una línea de la señal de televisión. Esta tasa de registro de datos da lugar, para el sistema de vídeo utilizado en aquella época en Gran Bretaña, a una frecuencia de muestreo de 44,1 KHz. Cuando se desarrolló el sistema CD Digital Audio, se tomó esta misma frecuencia de muestreo como un tributo a los primeros sistemas de audio digital.

La frecuencia de 44.1 KHz no es la única frecuencia de muestreo utilizada en audio digital. Cuando se definió el sistema DAT (*digital audio tape*) para el registro de audio en cintas magnéticas se definieron varias frecuencias de muestreo alternativas (32 KHz, 44,1 KHz y 48 KHz). El objetivo original de estos cambios en la frecuencia de muestreo era dificultar las **copias piratas** de CD en formatos de cintas magnéticas.

Los primeros **Minidisc** que salieron al mercado registraban la información utilizando frecuencias de muestreo de **32 KHz**. Estas frecuencias de muestreo exigen utilizar un filtro antialiasing que recorte la banda de audio útil entre los 15 KHz y los 16 KHz. Evidentemente, la calidad se reduce ligeramente aunque sólo un número reducido de melómanos son capaces de apreciarlo.

Actualmente, los sistemas portátiles **MP3** o **MP4** suelen utilizar una frecuencia de muestreo de **44.100 Hz**, ya que los registros suelen proceder de grabaciones para CD, aunque el estándar permite utilizar también otras frecuencias (32 kHz, 48 kHz, etc.).

El sistema **DVD-Vídeo** utiliza la frecuencia de **48 kHz** para el registro de la señal de audio. Esto significa que cuando se produce un DVD-Vídeo, partiendo de una banda musical obtenida en CD deber realizarse un cambio en la frecuencia de muestreo.

El sistema **DVD-Audio**, que no ha tenido demasiado éxito comercial, puede utilizar registros de audio con frecuencias de muestreo de **96 KHz** y **192 KHz**. Con ello, la banda de audio queda ampliamente cubierta con lo que se espera acabar para siempre con la polémica, promovida por algunos melómanos amantes de los viejos discos de vinilo, de que las frecuencias de muestreo del CD no permiten reproducir algunos matices sonoros que sí se reproducían con los discos de vinilo.

Los sistemas de **telefonía digital** utilizan frecuencias de muestreo más reducidas puesto que la banda de la señal de voz es más reducida que la musical y tampoco se pretende obtener calidades de reproducción de audio digital.

Recordemos que en telefonía el objetivo es la inteligibilidad del mensaje y el reconocimiento del locutor. No es necesaria la alta fidelidad. En este caso, la frecuencia de muestreo utilizada en telefonía fija para la voz es de **8 KHz**.

Veremos que esta frecuencia de muestreo exige un flujo de datos considerable (64 kbps–transmitir 64.000 bits cada segundo) que no es posible acomodar en los sistemas de **telefonía móvil digital** ni en las comunicaciones de audio a través de Internet. Por ello, se han introducido diversos estándares para la **compresión/descompresión** de la información de audio que intentan reducir el flujo total de datos con el que se trabaja sin afectar excesivamente a la calidad de la señal. Las frecuencias de muestreo de partida de estos sistemas se sitúan entre los **6 KHz** y los **20 KHz**.

2.3.12. Frecuencia de reproducción de la señal de audio

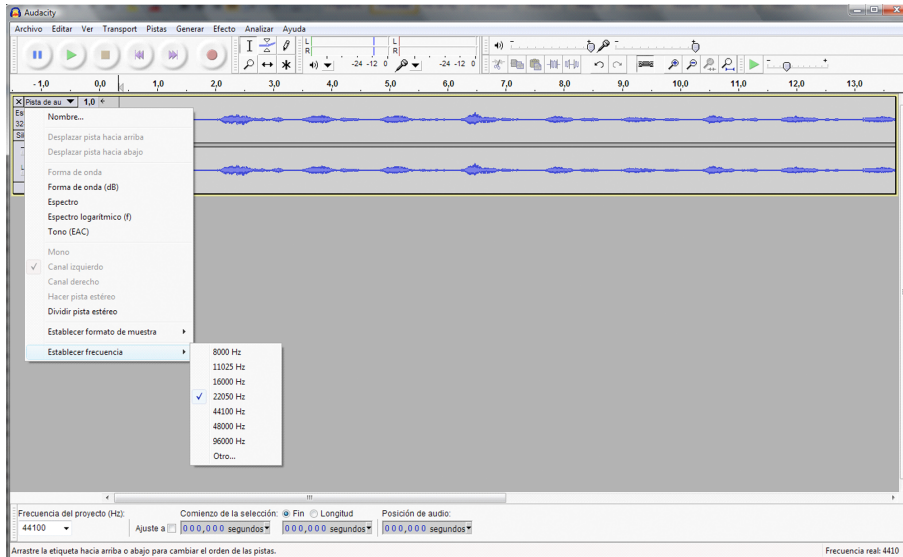
Es muy importante que la frecuencia de muestreo utilizada para la reproducción de la señal de audio sea **la misma** que la utilizada para su captura. Este hecho es especialmente importante y debe tenerse en cuenta debido a la multitud de frecuencias de muestreo que pueden utilizarse en los sistemas de audio digital. Es habitual, como ya hemos comentado, disponer de una señal que ha sido adquirida a una frecuencia de muestreo determinada y querer reproducirla en un sistema que no admite esta frecuencia.

Ejemplo

Un ejemplo muy típico es la captura de una señal de audio procedente de un CD-audio (44.100 Hz), editarla como pista de audio de un vídeo-clip y reproducirla en un DVD-Vídeo (solo admite reproducir audio a 48.000 Hz). Podemos encontrar problemas similares para pasar de Minidisc, DAT, MP3 a otros sistemas de audio.

- 1) Como ejemplo para ver los efectos de reproducir una señal con una frecuencia de muestreo diferente a la que ha sido adquirida podéis utilizar Audacity y cargar una de las señales utilizadas en esta actividad (por ejemplo, **panisAngelicus.wav**).
- 2) Podéis modificar la frecuencia de muestreo de la reproducción abriendo el desplegable de **opciones** de la pista de audio y seleccionando la nueva frecuencia de muestreo (ver figura 47).
- 3) Realizad el cambio de frecuencia de 44.100 Hz a 48.000 Hz y observad los efectos (aumento de la frecuencia fundamental de la voz y disminución del tiempo total de reproducción). El efecto producido es comparable al que tendríamos en analógico al pasar una cinta de audio a una velocidad mayor a la que ha sido grabada.

Figura 47. Opciones para el cambio de la frecuencia de muestreo del reproductor en Audacity



En todos estos casos, la señal debe ser procesada y remuestreada a la frecuencia de muestreo final del reproductor. Este procesado se conoce con el nombre de **cambio de frecuencia de muestreo** y es uno de los aspectos que incluyen todos los sistemas de edición de audio.

El cambio de la frecuencia de muestreo se lleva a cabo concatenando las operaciones de interpolación (aumento de la frecuencia de muestreo en un número entero) con las operaciones de diezmado (disminución de la frecuencia de muestreo en un número entero).

Así, por ejemplo, si queremos pasar de una frecuencia de muestreo de 44.100 Hz a 48.000 Hz deberíamos interpolar la señal en un factor 480 y diezmarla en un factor 441.

En la práctica los valores de interpolación y diezmado de la frecuencia de muestreo de una señal deben mantenerse bajos para que la complejidad computacional del procesado de audio no sea excesiva por lo que valores de 480 o 441 son absolutamente impracticables. Así pues, en la práctica, para pasar de 44.100 Hz a 48.000 Hz podríamos aproximarlos con una interpolación en un factor 12 y un diezmado en un factor 11. Con estos valores, si partimos de 44.100 muestras por segundo obtendremos un total de 48.109. Nos sobran por tanto 109 muestras cada segundo, que deberemos descartarlas. La forma más habitual de descartarlas es eliminar estas 109 muestras distribuyéndolas lo más uniformemente posible entre las 48.109 calculadas y obtener solo 48.000 Hz cada segundo.

2.4. Problemas/preguntas

1) Identificad las frecuencias de muestreo que se utilizan en los diferentes sistemas de audio digital comercializados. Entre los sistemas de audio estudiados incluid, como mínimo, los siguientes: CD-audio, DVD-Vídeo, DAT, Minidisc, MP3, MP4, DAB, DVD-Audio, SA-CD, DCC, NT, DAB+, DMB, MPEG-2 Layer 2, MPEG-2 LSF, NICAM 728, Dolby AC3, DTS, etc. Buscad la documentación en bibliografía relacionada, páginas de características técnicas de fabricantes o documentación genérica de Internet.

2) Identificad 2 modelos de tarjetas de captura, procesado y reproducción de sonido para bus PCI o PCI Express para PC. Determinad las frecuencias de muestreo con las que pueden trabajar las tarjetas y el número máximo de canales de audio que se pueden digitalizar. Si es posible intentad que los dos modelos sean de distintos fabricantes.

3) Identificad 2 modelos de tarjetas de captura, procesado y reproducción de audio con formato ExpressCard. Comparad las prestaciones de estos sistemas con los identificados en la cuestión anterior.

4) Digitalizad una señal musical procedente de un sistema de reproducción Compact Cassette analógico y mediante el programa Audacity, estudiad el espectro y ancho de banda de la señal obtenida. Probad con distintas frecuencias de muestreo y determinad cuál es desde el punto de vista auditivo la más adecuada.

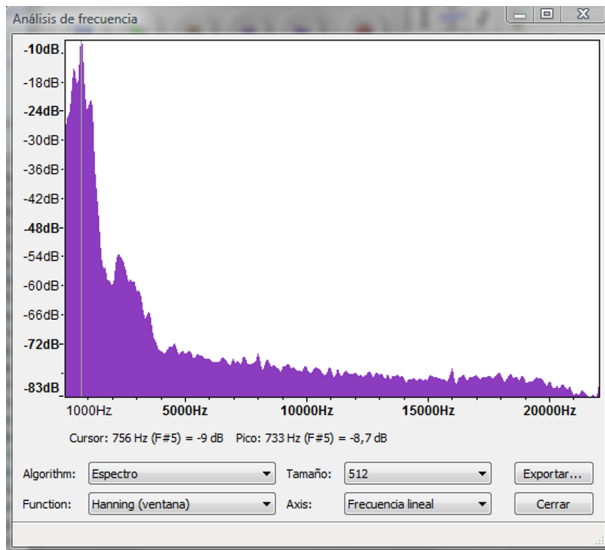
5) Una imagen digital puede considerarse como el muestreo en el espacio (filas y columnas) de una imagen analógica. Por lo tanto, como estamos realizando un muestreo, es posible que aparezca el fenómeno del aliasing. Proporcionad ejemplos en los que puede aparecer el fenómeno del aliasing en imágenes digitales. Indicad posibles maneras de realizar/aplicar un filtro antialiasing.

2.5. Evaluación

1) Considerad el espectro de la señal de audio que se muestra en la figura 48 y que ha sido obtenido digitalizando audio con Audacity. Suponed que para especificar un ancho de banda tomamos como criterio la frecuencia a partir de la cual el contenido espectral de la señal está 60 dBs por debajo del máximo. Se pide:

- a) Determinad la frecuencia de muestreo con la que ha sido adquirida esta señal.
- b) Calculad el ancho de banda de la señal teniendo en cuenta el criterio anterior.
- c) Determinad la frecuencia de muestreo mínima a la que considera que puede oírse esta señal sin pérdidas de calidad apreciable.

Figura 48



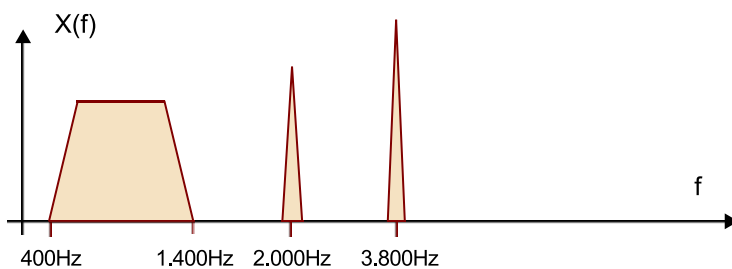
Solución

- La señal ha sido muestreada a 44.100 Hz, ya que la gráfica muestra hasta la frecuencia mitad (22.050 Hz).
- Teniendo en cuenta el criterio de los -60 dBs respecto al máximo (-10 dBs) deberemos ir aproximadamente a los -70 dBs que da aproximadamente unos 3.500 Hz de ancho de banda.
- Para el ancho de banda estimado en el apartado anterior, una frecuencia de muestreo de 7.000 Hz es suficiente y proporciona la misma calidad que la señal original.

2) Considerad una señal cuyo espectro puede aproximarse al de la figura 49.

- Determinad el espectro resultante si la señal se muestrea a 10 kHz utilizando un filtro antialiasing.
- Determinad el espectro resultante si la señal se muestrea a 4,2 kHz utilizando un filtro antialiasing.
- Determinad el espectro resultante si la señal se muestrea a 4,2 kHz sin utilizar un filtro antialiasing.

Figura 49



Solución

- Con una frecuencia de muestreo de 10 KHz el espectro coincide con el de la figura.
- El filtro antialiasing deberá filtrar a 2,1 kHz por lo que el tono a 3800 Hz desaparecerá.
- Con esta frecuencia, si no utilizamos el filtro antialiasing el tono que originalmente está en 3800 Hz aparecerá en 400 Hz, superponiéndose con la señal original.

3) Para realizar un cambio de frecuencia de muestreo desde 48.000 Hz a 44.100 Hz se propone utilizar una interpolación en un factor 11 y un diezmado en factor 12.

- a) Determinad la frecuencia de muestreo final que obtendríais.
- b) Proponed un procedimiento para obtener exactamente 44.100 muestras por segundo manteniendo los factores de interpolación y diezmado del enunciado.

Solución

- a) La frecuencia de muestreo que obtendremos es de 44.000 Hz.
- b) Debemos crear 100 muestras más cada segundo de las que hemos calculado. Una posibilidad es repetir una de las muestras cada 10 ms. El oído no apreciará la diferencia y obtendremos una señal real de 44.100 muestras por segundo.

4) Suponed una señal que ha sido muestreada a 44.000 Hz y que desea reproducirse a una frecuencia de muestreo de 44.100 Hz. Se pide:

- a) Si la señal original tiene una duración de 50 minutos, ¿cuánto durará la señal reproducida a 44.100 Hz?
- b) Suponed que la señal original contiene un tono de 800 Hz, ¿cuál será la frecuencia que percibiremos al reproducirla a 44.100 Hz?

Solución

- a) El número total de muestras que tenemos es $44.000 \times 50 \times 60 = 132.000.000$. Si se reproducen a una frecuencia de 44.100 muestras por segundo, tendremos una duración de 2993,1 segundos, es decir, 49 minutos y 53,1 segundos.
- b) La señal de 800 Hz tendrá un periodo de $44.000/800$ muestras = 55 muestras. Cuando estas muestras se reproducen a una velocidad de 44.100 obtenemos un periodo de 1,24717 min, que representa una frecuencia de 801,81 Hz.

5) Se desea aprovechar el fenómeno del aliasing para visualizar la deformación de la membrana de un altavoz cuando se excita a frecuencias de 150 Hz y con elevada potencia. Para ello, se desea ralentizar la observación de la deformación de la membrana utilizando un estroboscopio. Deseamos observar la membrana vibrando a 0.5 Hz (es decir, una oscilación completa cada dos segundos). Determinad cuál deberá ser la frecuencia de disparo del flash del estroboscopio.

3. Actividad 3. Digitalización de señales de audio: cuantificación

3.1. Introducción

3.1.1. Contenido

Esta actividad se centra en el estudio y análisis de los efectos de la codificación de las muestras en un número finito de bits. Para ello se propone que el estudiante realice diferentes audiciones de una señal digitalizada con diferentes calidades (número de bits) y evalúe el efecto sobre la calidad subjetiva del número de bits. A su vez, se propone que se analicen los efectos de la cuantificación, tanto en la forma de onda de la señal, comprobando la aproximación de nivel que se produce, como en los efectos sobre el espectro. Este análisis se realiza utilizando señales previamente muestreadas y con ayuda de un software genérico de edición y análisis de señales de audio.

Una vez se han realizado estas experiencias básicas, se formalizan los conceptos que se han introducido de manera intuitiva.

- En primer lugar se introduce el concepto de error de cuantificación y se interpreta como un ruido que se añade a la señal. Para valorar el nivel de ruido de una forma objetiva y numérica, se introduce el concepto de relación señal ruido y se proporcionan fórmulas que permitan calcularlo.

El objetivo principal de esta parte es que el estudiante sea capaz de asociar números de SNR en decibelios, que habitualmente se describen en las características técnicas de los equipos de adquisición, con la calidad perceptual del audio que se ha escuchado en la parte experimental.

- Seguidamente, se define la codificación de la señal digital en formato PCM, considerando la codificación en binario natural y la codificación en complemento a 2. Esta parte de la actividad se deja como opcional y pretende justificar por qué la mayor parte de los sistemas de adquisición utilizan el complemento a 2 como formato de representación de las muestras.
- Posteriormente, se introduce el concepto de tasa de bit y se proporcionan varios ejemplos para calcularlo en función de los parámetros de digitalización de la señal. Se da especial importancia a relacionar la tasa de bit con la capacidad de almacenamiento y la capacidad de difusión de contenidos

a través de redes de datos. Finalmente, se proporcionan algunos ejemplos de sistemas comerciales y se comentan las ventajas o inconvenientes de utilizar un número de bits mayor o menor. También, como conclusión de esta actividad, se presentan las ventajas principales de los sistemas de audio digital frente a los analógicos.

3.1.2. Metodología

En primer lugar se propone que el estudiante experimente con la calidad auditiva de una misma señal de audio codificada con diferentes niveles de cuantificación. El objetivo es que de forma natural se haga una asociación entre el número de bits de codificación y la calidad de la señal, comprobando de forma auditiva estos efectos y su importancia. También se pretende que se compruebe cómo existe un límite de las capacidades del sistema auditivo justificando que no tiene sentido aumentar la resolución de los equipos tecnológicos más allá de ese límite.

En paralelo y con ayuda de un software genérico de visualización, edición y procesado de señales de audio, se observan los efectos de la cuantificación sobre la forma de onda y sobre el espectro. Después de estos experimentos, el estudiante deberá tener una idea clara del número de bits que necesita para una correcta audición y de la caracterización en tiempo y en frecuencia del ruido de cuantificación.

En una segunda parte de la actividad se formalizan los conceptos de ruido de cuantificación, relación señal ruido, codificación PCM, tasa de bit y ventajas de los sistemas digitales frente a los analógicos. Todos estos conceptos se exponen de forma autocontenida en el texto de la actividad. Además, en la sección de problemas se proponen experiencias alternativas que fomentan el uso de bibliografía adicional o la búsqueda de información en Internet sobre algunos productos comerciales o sistemas de audio. A lo largo del texto se profundiza en algunos conceptos específicos, que sin ser fundamentales pueden tener cierto interés para aquellos estudiantes más motivados desde un punto de vista tecnológico. Estos conceptos se indican claramente como opcionales aunque se anima a los estudiantes a intentar seguirlos.

Finalmente, se proponen varios problemas parecidos a los que han ido apareciendo como ejemplos a lo largo del texto y se incluyen las soluciones.

3.1.3. Recursos

Los principales recursos utilizados en esta actividad son:

- Archivos de audio predigitalizados con diferentes resoluciones (número de bits).

- Software de captura y procesado de audio Audacity (licencia libre).
- Análisis de forma de onda y análisis de frecuencia de las señales.
- Explicaciones de los fundamentos de la cuantificación, el ruido, la relación señal ruido, la codificación PCM y la tasa de bits de los sistemas de audio digital.
- Propuesta de actividades de búsqueda de información en webs y recursos de Internet para profundizar en los conocimientos adquiridos en esta actividad (webs de fabricantes de tarjetas de captura, wiki, etc.).
- Problemas propuestos para verificar la adquisición de los conocimientos de esta actividad.

3.2. Objetivos

Los principales objetivos que se persiguen en esta actividad son:

- 1) Interpretar la función de los dos elementos clave en el proceso de digitalización de una señal: la frecuencia de muestreo y el paso de cuantificación, profundizando en este último.
- 2) Relacionar, mediante experimentos dirigidos, la calidad subjetiva de una señal de audio con el paso de cuantificación.
- 3) Relacionar el paso de cuantificación con el número de bits del cuantificador.
- 4) Conocer las características espectrales del ruido de cuantificación.
- 5) Comprender los efectos de la cuantificación como una aproximación a la señal original que introduce ruido en la misma.
- 6) Introducir el concepto de relación señal ruido y saber relacionar los valores obtenidos con la calidad subjetiva.
- 7) Comprender el procedimiento de codificación de una señal en el formato PCM (*pulse code modulation*).
- 8) Conocer las ventajas de la codificación en complemento a 2 respecto a la codificación en binario convencional para la asignación de códigos a la señal de audio.
- 9) Conocer la definición de tasa de bits para un sistema digital y en concreto para los sistemas de audio digital.

- 10) Distinguir entre los conceptos de cuantificador, codificador y convertidor A/D.
- 11) Saber calcular la tasa de bits de un sistema de audio digital en función de las características de los parámetros de muestreo.
- 12) Saber determinar el ancho de banda y las capacidades de almacenamiento de sistemas para la difusión y almacenamiento de audio digital.
- 13) Conocer las resoluciones y número de bits que se utilizan en los sistemas de audio comerciales más habituales.
- 14) Comprender la necesidad de comprimir la información PCM para optimizar la transmisión y el almacenamiento de la señal de audio digital.
- 15) Conocer las tecnologías de captura y digitalización y los equipos comerciales que pueden utilizarse en entornos de edición tipo PC.
- 16) Conocer las principales ventajas tecnológicas de los sistemas digitales frente a los analógicos.

Estos objetivos están relacionados con las siguientes **competencias de la asignatura**:

- 1) Capacidad de capturar audio e imágenes de forma eficiente y eficaz.
- 2) Capacidad de digitalizar eficiente y eficazmente contenidos de audio e imagen.
- 3) Capacidad de interpretar y modificar información auditiva de forma frecuencial.

Y con las siguientes **competencias generales del grado**:

- 1) Capturar, almacenar y modificar información de audio, imagen y vídeo digitales aplicando principios y métodos de realización y composición del lenguaje audiovisual.
- 2) Integrar y gestionar contenidos digitales en aplicaciones multimodales de acuerdo con criterios estéticos, técnicos y funcionales.
- 3) Atender adecuadamente consultas sobre proyectos, tecnologías y mercado de productos multimedia evaluando de forma precisa el entorno de aplicación, los recursos y las alternativas tecnológicas disponibles.

3.3. Guía

Generalmente los fenómenos físicos naturales se manifiestan como **señales analógicas**.

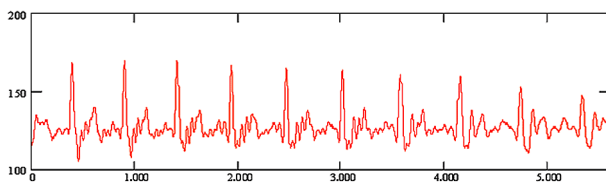
Podemos citar un gran número de ejemplos en los que el fenómeno físico puede representarse como una función matemática, es decir, como una señal analógica.

Ejemplo

Entre estos ejemplos: la evolución de la temperatura en un punto de observación, el caudal de un río, el crecimiento de una planta, las variaciones de presión que forman las ondas acústicas o la tensión recogida por un micrófono como consecuencia de estas ondas acústicas. La señal de audio es pues, en su origen, analógica.

En la figura 50 hemos representado de forma esquemática la señal de tensión recogida por un micrófono y que se corresponde a una vocal (la *o*).

Figura 50. Evolución en el tiempo de la tensión recogida por un micrófono al pronunciar la vocal *o*



La señal es analógica, es decir, es continua en el tiempo y amplitud.

Digitalizar esta señal analógica es un proceso de aproximación que, como ya vimos en el apartado anterior, utiliza una **retícula de muestreo**. La calidad de esta aproximación depende de la densidad de la retícula de muestreo, tanto en el eje horizontal (frecuencia de muestreo) como en el eje vertical (paso de cuantificación).

En el apartado anterior hemos considerado los efectos de la **frecuencia de muestreo** en el proceso de digitalización de la señal. Para valorar la incidencia de la frecuencia de muestreo, se supuso que el paso de cuantificación era suficientemente pequeño como para poder despreciar sus efectos en la pérdida de calidad de la señal.

Ahora vamos a realizar el proceso inverso: para analizar los efectos del paso de cuantificación vamos a suponer que la frecuencia de muestreo es suficientemente alta y que su valor no condiciona la calidad final de la señal en formato digital.

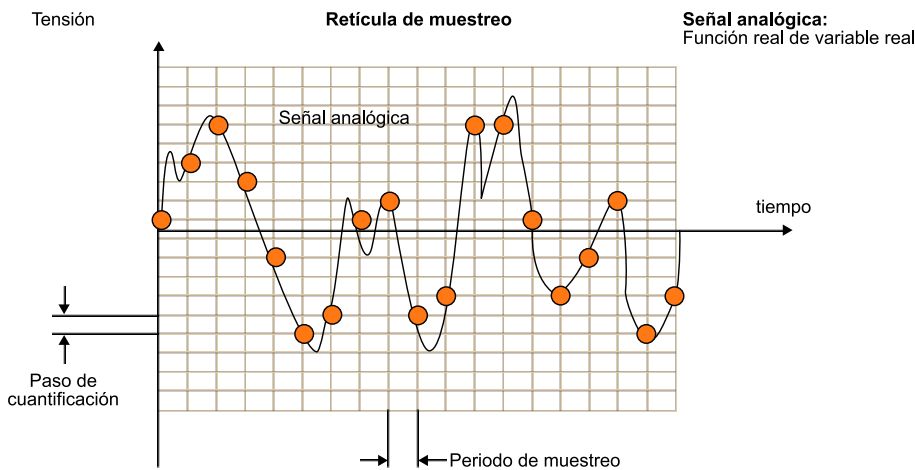
En esta actividad haremos diferentes pruebas auditivas para valorar los efectos del paso de cuantificación en la señal digital. Con ello, queremos determinar los criterios para fijar este parámetro en función de la aplicación y de la calidad final que deseemos obtener. Posteriormente describiremos el formato de audio digital no comprimido PCM, analizando sus ventajas y sus inconvenientes.

3.3.1. La retícula de muestreo y el paso de cuantificación

En la actividad anterior se introdujo el concepto de retícula de muestreo como una estrategia para aproximar una señal analógica mediante un conjunto numerable de puntos. En la figura 34 se reproduce esta retícula de muestreo en la que definíamos la frecuencia de muestreo y el paso de cuantificación como la densidad de la retícula en los ejes horizontal y vertical respectivamente.

Es intuitivo que cuanto más densa sea esta retícula mejor será la aproximación digital a la señal analógica. No obstante, también está claro que aumentar la densidad de la retícula aumentará el volumen de información asociado a la señal digital, afectando a su almacenamiento y a la eficiencia de su transmisión. Por tanto, es muy importante que la selección de estos parámetros se haga atendiendo a los criterios de calidad requeridos para una determinada aplicación. En efecto, no tiene sentido aumentar la frecuencia de muestreo o disminuir el paso de cuantificación más allá de lo que puede percibir el oído humano.

Figura 51. Aproximación de una señal analógica mediante una retícula de muestreo



La retícula de muestreo está caracterizada por el paso horizontal (periodo de muestreo) y el paso vertical (paso de cuantificación).

En consecuencia, el sistema auditivo se convierte en el juez último que determina cuál es la densidad adecuada de la retícula de muestreo en cada uno de los ejes. Sin embargo, como veremos, esta decisión no resulta trivial, ya que no todos los sistemas auditivos son igual de críticos y además el conjunto de pruebas que se realizan para fijar los parámetros puede ser que no contemple algunos casos especiales donde sí podrían apreciarse diferencias entre el original analógico y la aproximación digital.

Una de las características que hemos estudiado en la actividad anterior es el **teorema de Nyquist**, que establece un límite teórico para la frecuencia de muestreo. Según este teorema si una señal se muestrea a una frecuencia mayor que el doble de su ancho de banda, entonces la señal analógica puede recuperarse exactamente a partir de sus muestras.

En otras palabras, no se producen pérdidas por el hecho de muestrear la señal siempre que se cumpla el criterio de Nyquist.

Lamentablemente, no ocurre lo mismo en el eje vertical. Cuando fijamos un número finito de posibles niveles, el valor original de la señal ya no puede ser recuperado. En la figura 34 se observa que a partir de las muestras digitales (puntos) sólo podremos obtener una versión aproximada de la señal analógica original.

Para analizar el efecto del paso de cuantificación en la calidad de señal de audio, es conveniente realizar pruebas auditivas de calidad. En la práctica, la elección del paso de cuantificación se realiza mediante la realización de **pruebas subjetivas** a múltiples usuarios.

Estas pruebas consisten en la audición de varios fragmentos musicales, cuantificados con diferentes pasos y que deben compararse con la señal analógica original.

El paso de cuantificación elegido deberá ser aquel para el cual una gran mayoría de los usuarios (a ser posible todos) no puedan distinguir entre la señal original y la señal digitalizada.

Evidentemente, las pruebas deben realizarse con una frecuencia de muestreo suficientemente alta para que ésta no afecte a la calidad de la señal.

En esta actividad vamos a reproducir de forma simplificada estas pruebas subjetivas para ver el efecto del paso de cuantificación en la calidad final de la señal.

Tomaremos como señal de referencia la señal original de una grabación para CD-audio, que tiene una frecuencia de muestreo de 44.100 Hz y un total de 65.536 pasos de cuantificación. En realidad se trata de una señal digital, pero si aceptamos que las pruebas realizadas para la definición del formato CD-audio fueron correctas, entonces no debería ser posible apreciar diferencias entre esta señal digital y la original analógica⁷. El archivo original con el que vamos a trabajar es **NessunDorma.wav** y se adjunta con el material adicional de esta actividad.

⁽⁷⁾ Generalmente, por abuso del lenguaje, hacemos referencia a una **señal original analógica** cuando en muchas grabaciones la señal original que se toma en el máster ya es digital, por lo que no existe una versión analógica. No obstante, en este texto, entendemos como señal original analógica aquella que escucharíamos en condiciones ideales sin haber realizado ninguna aproximación de digitalización.

1) Abrid este archivo con el programa Audacity y escuchadlo. La calidad es buena, ya que se corresponde con las muestras originales del CD a 44.100 Hz y utilizando un total de 65.536 pasos de cuantificación⁸, es decir, 16 bits.

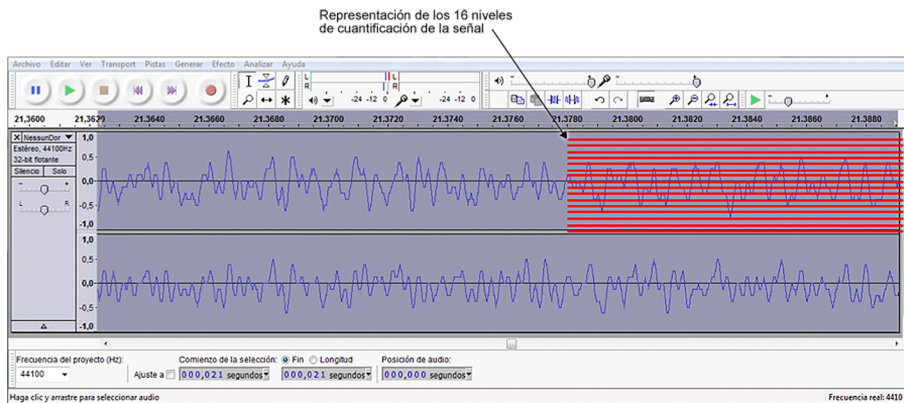
⁽⁸⁾ El número de pasos de cuantificación suele ser siempre una potencia de 2 debido a que, como veremos más adelante, el nivel de cuantificación se codifica mediante una palabra de un número fijo de bits. Así, con 16 bits podemos codificar hasta 65.532 niveles = 2^{16} niveles, mientras que con 15 bits podremos codificar 2^{15} niveles = 32.768. Obsérvese que no tiene mucho sentido elegir un número de niveles de cuantificación igual a 50.000, ya que para codificarlos necesitaremos utilizar 16 bits por muestra y, por este mismo coste, podemos codificar ¡hasta 65.532!

2) Abrid también el archivo **NessunDorma_4bit.wav**. Este archivo ha sido obtenido a partir del original pero modificando el número total de niveles de cuantificación a 16 (4 bits – 2^4 niveles).

3) Escuchad atentamente este archivo y comprobad que el efecto es que se ha sumado un ruido considerable sobre la señal original. Este ruido se denomina **ruido de cuantificación** y se debe a que el paso de la retícula de muestreo en el eje vertical es insuficiente para aproximar correctamente la señal. Es indudable que el oído humano es capaz de discriminar entre la señal original y una burda cuantificación de la misma que utiliza sólo 16 niveles. Veremos que a medida que aumentemos el número de niveles de cuantificación este ruido disminuirá y llegará un punto en el que el oído será incapaz de discernir entre la señal original y la aproximada.

4) Para observar con mayor claridad el efecto de la cuantificación, seleccionad un fragmento de corta duración de la señal de 4 bits y ampliad la ventana de zoom hasta que podáis ver los detalles de la forma de onda. En la figura 52 se muestra un fragmento de la señal donde se han marcado los 16 posibles niveles de cuantificación que se están utilizando. En este gráfico se observa que el efecto de la cuantificación es una especie de aproximación en escalera, a través de la retícula de muestreo, a la señal original.

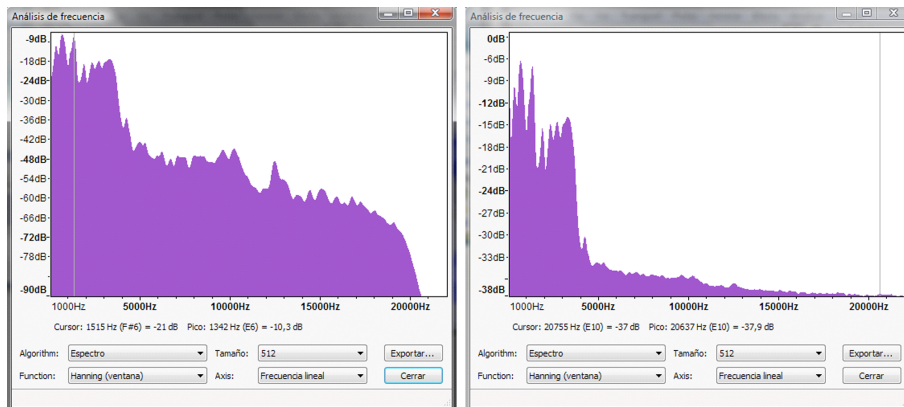
Figura 52. Representación de la forma de onda de un fragmento de la señal NessunDorma cuantificado a 4 bits



Se muestran en rojo los 16 niveles de cuantificación de la señal.

Es muy interesante analizar las señales desde el punto de vista espectral. Para ello, tomad un fragmento de señal relativamente largo (de unos 10 o 20 segundos) y comparad el espectro de la señal original con el de la señal cuantificada con 16 niveles. Analizamos un fragmento relativamente largo para que queden promediados distintos espectros y poder visualizar el valor promedio del ruido de cuantificación. Tened cuidado de seleccionar la misma ventana para las dos señales. En la figura 52 mostramos los resultados que hemos obtenido en el fragmento que va desde el segundo 3 hasta el 23.

Figura 53. Espectro de un fragmento de la señal NessunDorma original (16 bits-izquierda) y la señal cuantificada (4 bits-derecha)



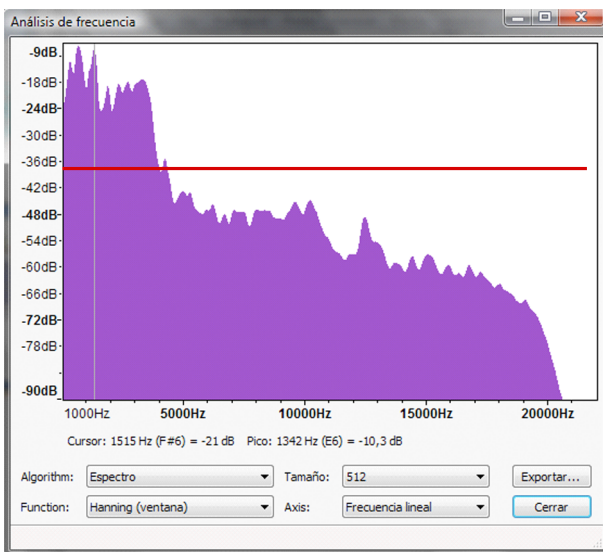
Los dos espectros deben compararse con cierto cuidado, observando que las escalas de los ejes verticales son distintas y que el aspecto general de los gráficos no debería confundirnos.

Esencialmente, estas gráficas nos muestran que los dos resultados son prácticamente idénticos en la región de baja frecuencia pero que en la zona de alta frecuencia las diferencias son significativas. Efectivamente, hasta los 4.000 Hz podemos considerar, de forma aproximada, que los dos espectros tienen las mismas componentes, en las mismas frecuencias y con los mismos niveles de amplitud. A partir de los 4.500 Hz las diferencias empiezan a ser considerables, ya que para la señal de 4 bits (16 niveles) predomina una componente de rui-

do, de espectro aproximadamente plano y que se sitúa alrededor de los -38 dBs. Esta componente de ruido enmascara las componentes de alta frecuencia de la señal original que pueden observarse en la gráfica de la izquierda.

En la figura 54 hemos representado el espectro de la señal con 16 bits y una línea roja que indica el espectro del ruido. El espectro de la señal de 4 bits puede considerarse como la suma entre este ruido espectralmente plano y el espectro de la señal original. Obsérvese en la figura 53 que los niveles en la zona de baja frecuencia para la gráfica de la derecha son algo mayores que los del espectro original debido a la suma de la componente de ruido.

Figura 54



Espectro de la señal cuantificada a 16 bits (considerada como original) con la superposición del ruido (en rojo) de espectro plano correspondiente a la cuantificación de 4 bits.

Analicemos ahora el efecto sobre la calidad de la señal y su espectro al aumentar o disminuir el número de pasos de cuantificación.

1) Abrid los archivos **NessunDorma_3bits.wav** y **NessunDorma_5bits.wav** que corresponden con la cuantificación de la señal con 8 y 32 pasos de cuantificación respectivamente y comparadlos con los resultados obtenidos con la cuantificación a 16 niveles.

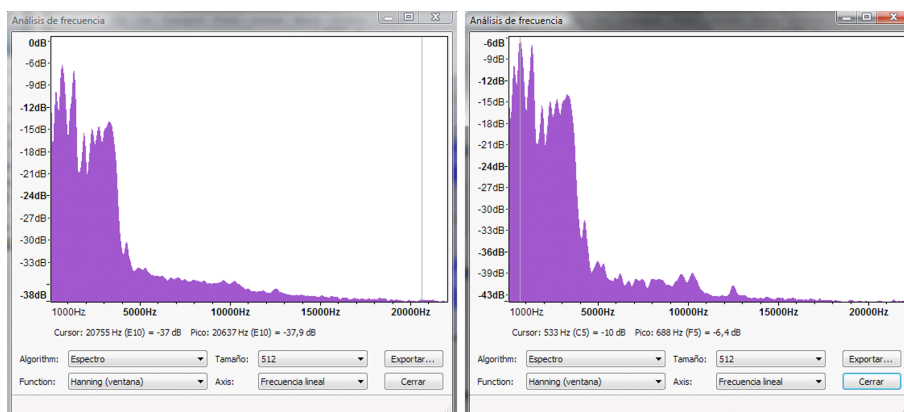
2) Desde el punto de vista auditivo resulta evidente cómo el nivel de ruido disminuye progresivamente a medida que aumenta el número de bits del cuantificador. En todos los casos la calidad final es todavía muy deficiente, pero la mejora progresiva al aumentar los bits indica que con algunos bits adicionales llegaremos a obtener calidades aceptables.

3) Respecto al análisis en frecuencia de las señales, es interesante observar cómo el nivel de ruido se va reduciendo cada vez que aumentamos en un bit el cuantificador.

Nótese que aumentar en un bit el número de pasos de cuantificación supone doblar el número total de niveles del cuantificador. Así, al pasar de 4 bits a 5 bits, aumentamos de 16 niveles a 32 niveles. Esto significa que la aproximación de la retícula de muestreo mejora, ya que el error de la aproximación se reducirá a la mitad. La reducción del error en la aproximación supone una reducción del ruido de cuantificación que se hace evidente en el análisis espectral de la señal (y por supuesto, en el análisis auditivo). En la figura 55 se compara el espectro obtenido con la señal cuantificada a 16 niveles (izquierda) y la cuantificada a 32 niveles (derecha). Obsérvese la reducción de nivel en el espectro del ruido que va desde unos -38 dBs en la gráfica de la izquierda hasta algo menos de -43 dBs en la gráfica de la derecha⁹. En el siguiente apartado vamos a formalizar estos resultados, proporcionando fórmulas matemáticas que nos relacionen directamente la calidad de la audición con el número de bits del cuantificador.

⁽⁹⁾En estas gráficas no resulta fácil medir los niveles del ruido. Veremos más adelante que la diferencia de nivel teórica debería ser exactamente de 6 dBs.

Figura 55. Comparativa entre los espectros obtenidos con un fragmento de señal cuantificado con 4 bits (izquierda) y el mismo fragmento cuantificado con 5 bits (derecha)



Las partes de baja frecuencia son prácticamente iguales pero aparece una diferencia de nivel de ruido entre los dos espectros de aproximadamente 6 dBs.

4) Analizad los espectros de las señales cuantificadas a 6, 7 y 8 bits y comprobad que el nivel de ruido se reduce progresivamente en una cantidad constante. Recordad que para poder comparar los espectros es conveniente que toméis un fragmento de señal de unos 10-20 segundos y que siempre se tome el mismo. Comparad también de forma auditiva la calidad de las señales obtenidas.

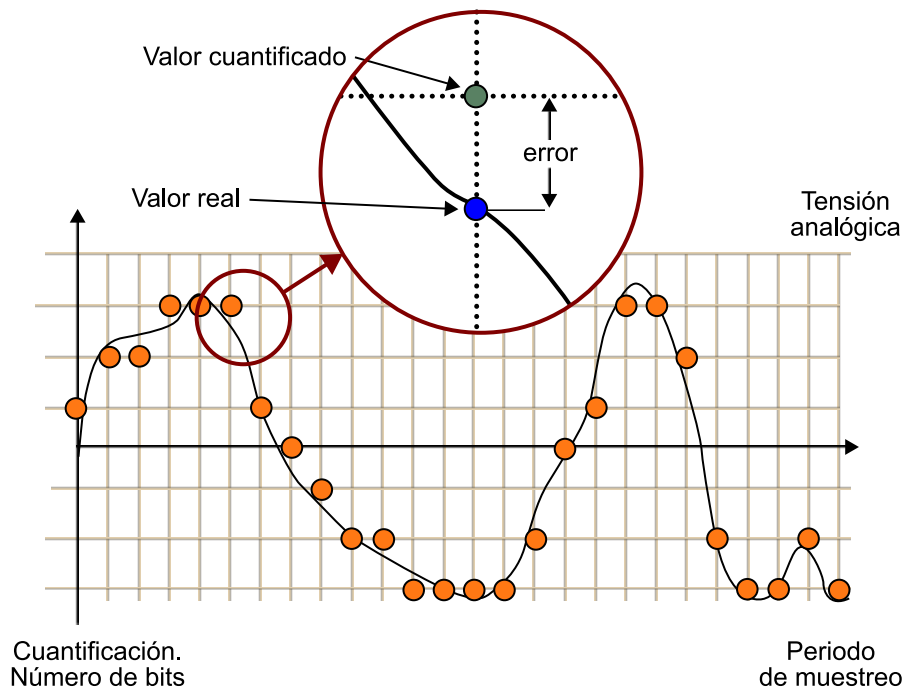
5) Escuchad las señales cuantificadas con 1 bit y 2 bits y analizad sus espectros. Tened en cuenta que en estas dos señales el ruido es muy importante. Bajad el volumen del altavoz o los auriculares para no dañar el equipo ni el oído. La audición de la señal de 1 bit es auténticamente espectacular. Notad que la forma de onda de la señal sólo toma dos posibles valores: positivo y negativo. Aun así, el sistema auditivo es capaz de escuchar el pasaje musical.

6) Aumentad progresivamente el número de bits: 9, 10, etc. y comparad auditivamente estas señales con la de 16 bits. Determinad el número mínimo de bits en el que es incapaz de apreciar diferencias con la señal del CD original (**NessunDorma.wav**).

3.3.2. Ruido de cuantificación y relación señal ruido

Al aproximar la señal analógica mediante una retícula de muestreo se introduce un error en el nivel de amplitud de la señal. En la figura 56 se muestra cómo el valor real de la señal en el instante de muestreo debe ser aproximado por uno de los pasos de cuantificación existentes en la retícula de muestreo. El error que se produce depende del nivel de la señal. En ocasiones, es posible que el error sea muy pequeño debido a que la señal pasa muy próxima a un nivel de cuantificación mientras que en otros casos el error puede ser más importante. No obstante, sí que está claro que el error máximo nunca superará la mitad de la amplitud del paso de cuantificación. Efectivamente, el nivel de cuantificación más próximo a la señal estará siempre, como muy lejos, a la mitad del paso de cuantificación, ya que siempre podemos aproximar la señal por el nivel de cuantificación más próximo por arriba o por abajo.

Figura 56. Representación gráfica del error que se produce al aproximar la señal por un nivel de cuantificación



Como los niveles de señal no se conocen a priori, podemos asumir que el error será una variable aleatoria cuyos valores están comprendidos entre:

$$-\frac{\Delta}{2} \leq error \leq \frac{\Delta}{2}$$

donde Δ representa el paso de cuantificación.

Además, también parece razonable suponer que todos los valores de error pueden producirse con la misma probabilidad.

El paso de cuantificación está relacionado con el número de bits (número de niveles de cuantificación) y los valores máximo y mínimo entre los que se define la retícula de muestreo.

En efecto, observemos que en la figura 34 hemos representado una retícula de muestreo que acotaba las posibles amplitudes de la señal entre unos valores máximo y mínimo. De esta forma, el número de niveles de cuantificación será finito. Por otra parte, esta restricción no es importante, ya que en la práctica los niveles de la señal siempre estarán acotados por el mismo equipo de medida o captura (micrófono, amplificador), que seguramente entrará en saturación si el fenómeno físico (volumen acústico) supera ciertos umbrales.

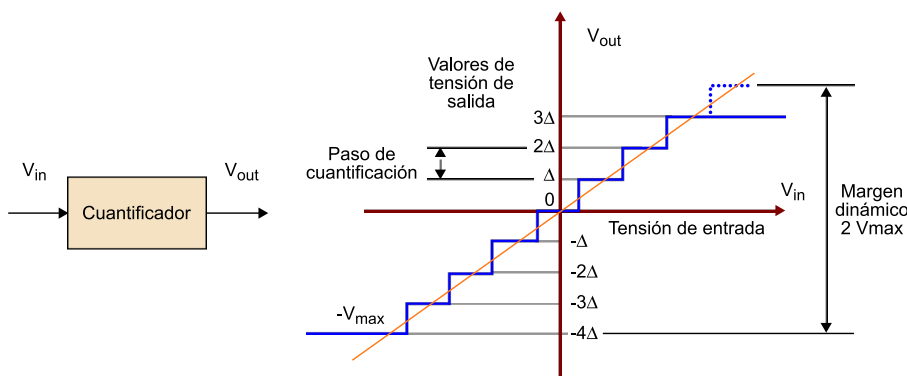
Una forma alternativa de representar la retícula de muestreo es definir los **cuantificadores**.

Un cuantificador es un sistema que aproxima el nivel de la señal de entrada en un número finito de posibles niveles de salida.

En la figura 58 se representa un cuantificador de 8 niveles (3 bits). La tensión de entrada se representa en el eje horizontal y puede tomar cualquier valor. La tensión en la salida¹⁰ sólo puede tomar uno de los 8 posibles niveles de cuantificación indicados en la gráfica.

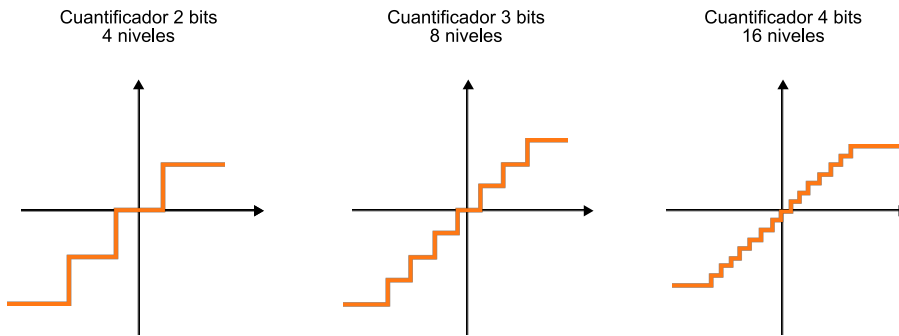
⁽¹⁰⁾En la mayoría de sistemas prácticos los niveles de salida no son simétricos respecto al cero. En la figura 58 puede comprobarse que el nivel mínimo alcanza hasta la tensión $-V_{ref}$ pero que el valor máximo no llega a alcanzar $+V_{ref}$. Describiremos con mayor profundidad esta asimetría más tarde, aunque por el momento vale la pena comentar que cuando el número de bits es suficientemente elevado los niveles máximo y mínimo pueden considerarse aproximadamente simétricos.

Figura 58. Relación entrada-salida de un cuantificador de 8 niveles (3 bits)



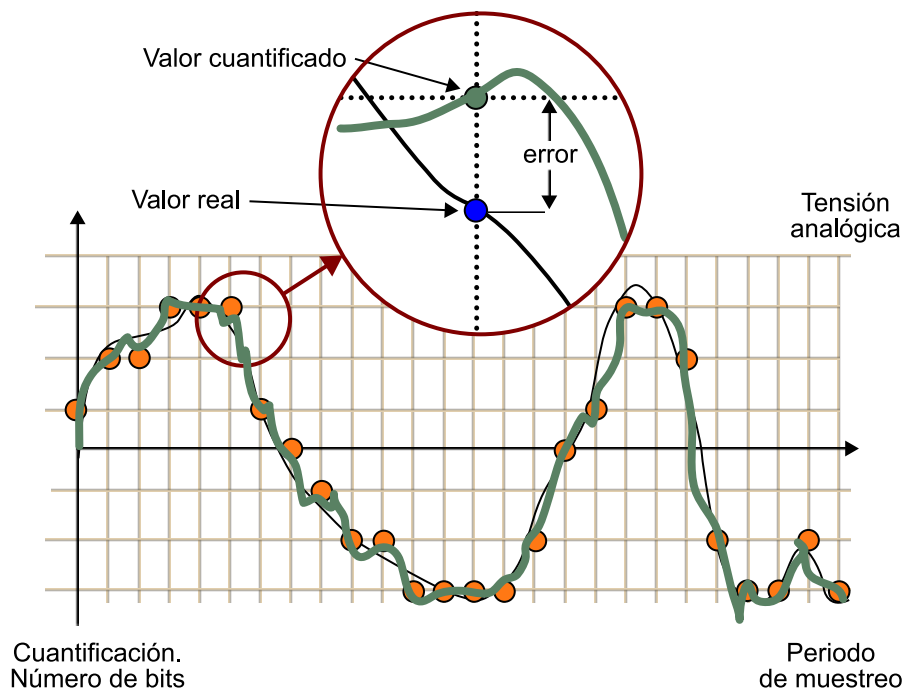
En la figura 59 se muestran algunos ejemplos de cuantificadores para diferentes números de bits. Se observa que al aumentar el número de bits tendremos más niveles de cuantificación, por lo que la aproximación entre la señal de entrada y la de salida será mejor.

Figura 59. Curvas entrada-salida para cuantificadores de 2, 3 y 4 bits



Hemos comprobado que el error de cuantificación es equivalente a un ruido y que por ello muchas veces se conoce con el nombre de **ruido de cuantificación**. En la figura 60 se intenta representar esta equivalencia entre el error y el ruido. En efecto, el proceso de cuantificación puede interpretarse como que se ha sumado un ruido a la señal y que este ruido es el que ha provocado que la muestra tome un valor de tensión que no era el original (sino la aproximación de nuestra retícula de muestreo).

Figura 60. Interpretación del error de cuantificación como un ruido que se suma a la señal (ruido de cuantificación)



3.3.3. La relación señal ruido

Para valorar desde un punto de vista cuantitativo la calidad de un sistema de audio suele utilizarse la relación señal ruido (SNR: *signal to noise ratio*) que se define como el cociente entre la potencia de la señal y la potencia de ruido y suele expresarse en decibelios mediante una relación logarítmica.

Así tenemos:

$$SNR = 10 \cdot \log_{10} \frac{P_{signal}}{P_{noise}}$$

Como podemos ver en esta fórmula, la relación señal ruido representa el cociente entre la potencia de la señal y la potencia del ruido. La calidad perceptual del audio está relacionada con este cociente. Evidentemente, cuanto menor sea la potencia de ruido mejor será la calidad perceptual. A su vez, si la potencia de ruido disminuye el resultado del cociente aumenta, proporcionando unos valores numéricos de SNR mayores.

La conclusión es que a mayores valores numéricos de SNR mejor es la calidad perceptual de la señal de audio.

Por otra parte, puede parecer algo extraño que se tome el logaritmo del cociente de potencias en vez del cociente de potencias propiamente dicho. Esto se debe, en parte, al comportamiento logarítmico del oído humano en cuanto a la apreciación de potencias de ruido. Cada vez que la potencia de ruido se reduce a la mitad, apreciamos una reducción progresiva del mismo, si volvemos a reducirlo a la mitad, volvemos a apreciar la misma reducción, etc. En otras palabras, al tomar el logaritmo del cociente de potencias los números de SNR que obtenemos expresan en una escala aproximadamente lineal la calidad perceptual de la señal.

Cuando se toma el logaritmo del cociente de potencias y se multiplica por 10, las unidades obtenidas reciben el nombre de **decibelios**.

En audio digital el cálculo de la relación señal ruido sólo puede hacerse mediante aproximaciones genéricas, ya que la potencia de la señal depende de la propia señal y posiblemente cambiará a lo largo de una grabación función del contenido musical. En efecto, en pasajes de bajo volumen o en pianísimo, la potencia de señal será mucho menor que en los fortísimos. Por lo tanto, en la práctica, lo que se suele tomar como potencia de señal en la expresión de la SNR es un valor promedio, válido para una amplia gama de señales musicales.

En cambio, la potencia de ruido es más fácil de calcular, ya que conocemos las características de error en la aproximación de la señal y sabemos que su amplitud depende del paso de cuantificación, que a su vez depende del número de bits del cuantificador.

A partir de la figura 58 podemos establecer la siguiente relación entre el paso de cuantificación, los límites de la retícula (tensión máxima) y el número de bits:

$$\Delta = \frac{2 \cdot V_{\max}}{2^N}$$

donde V_{\max} representan las tensiones límite del cuantificador (retícula) y N el número de bits del cuantificador. La expresión se obtiene dividiendo la excursión máxima de los niveles de entrada $V_{\max} - (-V_{\max}) = 2 V_{\max}$ por el número total de niveles de cuantificación 2^N .

A partir de esta última expresión, queda claro que al aumentar el número de bits del cuantificador se reduce el paso de cuantificación Δ y por lo tanto se reduce la amplitud máxima y la potencia del ruido.

En concreto, la expresión muestra que la amplitud del ruido se reducirá en la mitad cada vez que aumentemos en un bit el cuantificador.

Este resultado es muy importante y debe recordarse. De hecho, es un resultado muy intuitivo al que podemos llegar sin necesidad de fórmulas. En efecto, cuando el cuantificador se aumenta en un bit, el número de niveles se doblará, por lo que la amplitud máxima del error de cuantificación se reducirá a la mitad.

Si la amplitud del ruido se reduce a la mitad la potencia (que se calcula como el promedio de las amplitudes al cuadrado) se reducirá en una cuarta parte. Podemos expresar este resultado matemáticamente como:

$$\begin{aligned} SNR_{N+1bit} &= 10 \cdot \log_{10} \frac{P_{signal}}{P_{Noise\ N+1}} = 10 \cdot \log_{10} \frac{P_{signal}}{(P_{Noise\ N}/4)} = \\ &= 10 \cdot \log_{10} \frac{P_{signal}}{P_{Noise\ N}} + 10 \cdot \log_{10} 4 = SNR_{Nbit} + 6,02dB \end{aligned}$$

En definitiva, al aumentar en un bit el cuantificador, la relación señal ruido mejora en 6 dBs.

En la práctica puede llegarse a obtener una expresión genérica para la relación señal ruido como una fórmula simple que sólo depende del número de bits del cuantificador:

$$SNR = A + 6,02 \cdot N$$

donde el valor de la constante A depende de la potencia de la señal.

Generalmente, se suele tomar un valor de **A igual a 1,76**, que se corresponde con la potencia máxima de la componente de señal. De esta forma, la ecuación anterior se transforma en:

$$SNR = 1,76 + 6,02 \cdot N$$

Problemas

Vamos a ver algunos ejemplos utilizando esta fórmula.

1) Determinad la relación señal ruido del sistema CD-audio y calculad la potencia del ruido de cuantificación de este sistema como una fracción de la potencia de señal.

Solución

Utilizando que el número de bits del sistema CD-audio es 16 y utilizando la ecuación anterior obtenemos:

$$SNR = 1.76 + 6.02 \cdot 16 = 98 \text{ dBs}$$

Con 98 decibelios la relación entre la potencia de señal y la potencia de ruido de cuantificación puede expresarse mediante:

$$98 = 10 \cdot \log_{10} \frac{P_{\text{signal}}}{P_{\text{Noise}}} \Rightarrow P_{\text{Noise}} = \frac{P_{\text{Signal}}}{10^{9,8}} = 1,58 \cdot 10^{-10} P_{\text{signal}}$$

2) Determinad la relación señal ruido de un sistema de telefonía digital que utiliza 8 bits por muestra.

Solución

Particularizando la fórmula anterior para 8 bits obtenemos que la relación señal ruido de un sistema de telefonía digital será de 49,7 dBs.

3.3.4. Codificación PCM (*pulse code modulation*) de la señal de audio

Hemos visto que digitalizar una señal analógica consiste, esencialmente, en aproximarla mediante una retícula de muestreo y hemos analizado las limitaciones que introducen la frecuencia de muestreo y el paso de cuantificación en la calidad de esta aproximación. No obstante, todavía nos queda por definir cómo se construye la señal digital, es decir, la secuencia de bits que representa la información de la señal.

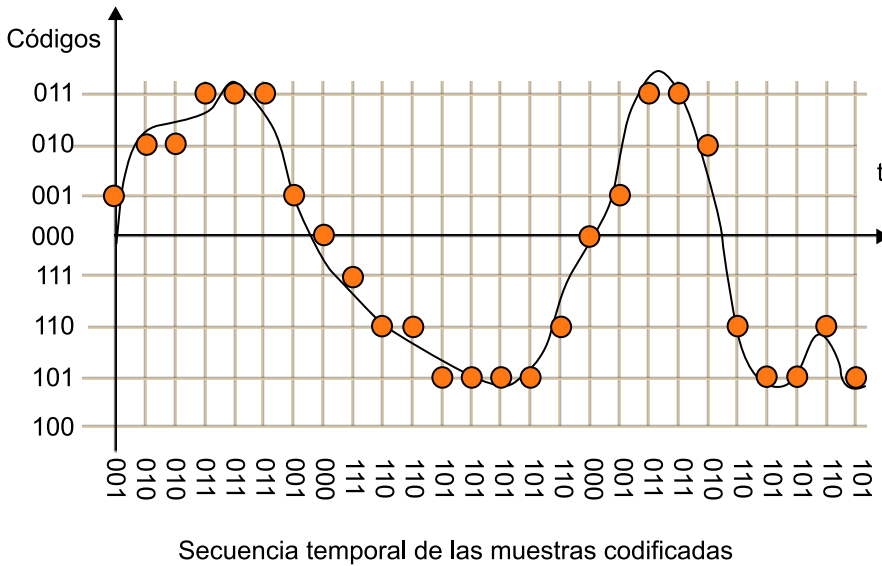
El proceso más directo para convertir la señal a una secuencia de bits es asignar un código binario a cada nivel de cuantificación. De esta forma, para indicar el valor de cada muestra bastará con proporcionar la palabra binaria asociada.

Este procedimiento se ilustra de forma esquemática en la figura 61 donde se muestra un cuantificador de 8 niveles en el que cada uno de los niveles tiene asignada una palabra código de 3 bits distinta. Para transmitir o almacenar la información asociada a la señal, bastará con enviar o almacenar de forma secuencial las palabras código de cada una de las muestras. En la figura 61 se muestra una secuencia de ejemplo 001 010 010 011 011 011 001 000 111 110 110 101 101 101... que se correspondería con la aproximación a la forma de onda representada.

PCM

Este tipo de codificación en el que asociamos una palabra binaria de longitud fija a cada muestra se la conoce con el nombre de **PCM** (*pulse code modulation*) y fue patentado en 1937 por Alec H. Reeves.

Figura 61. Ejemplo de la codificación PCM de una señal analógica como una secuencia de bits



001 010 010 011 011 011 001 000 111 110 110 101 101 101 ...

A pesar de su simplicidad, esta codificación es, sin duda, la más utilizada. Entre otras razones, debido a que es fácilmente realizable en hardware, tanto en aplicaciones de audio como en aplicaciones de imagen o vídeo. No requiere ningún tipo de procesamiento especial.

Ejemplo

Podemos encontrar ejemplos de este tipo de codificación en el formato WAV de audio, el *compact disc-audio*, el DAT (*digital audio tape*) y la grabación de audio PCM existente en muchos DVD-Vídeo.

En imágenes digitales los formatos BITMAP y RAW son diferentes versiones de la codificación PCM. En vídeo, los formatos 4:2:2 y 4:2:0 estandarizados en la recomendación ITU-601 son también una codificación PCM. Pero el formato PCM también se encuentra presente en muchos equipos que utilizan como base otros codificadores para almacenar o transmitir la información. Así, un grabador o reproductor portátil en MP3 captura la información de audio en PCM y posteriormente la procesa, para comprimirla y almacenarla de forma más eficiente en un formato comprimido como MP3, MP4 o WMA. A su vez, durante la reproducción, la información codificada en formato comprimido se procesa y se convierte a PCM antes de aplicarla a los circuitos electrónicos que se encargarán de su conversión a analógico y su amplificación.

La realización hardware de la codificación PCM se realiza mediante **convertidores A/D** (analógico/digital). Un convertidor A/D es un circuito integrado que se encarga de la digitalización de las señales y que suele incluir el subsistema de muestreo y cuantificación de la señal.

En esencia, de forma conceptual y muy simplificada, un convertidor A/D sería un circuito que tiene como entradas la señal analógica y la frecuencia de muestreo y que en su salida proporciona la secuencia de bits binaria asociada a la codificación PCM de la señal.

Un elemento que generalmente está asociado al convertidor A/D es el convertidor D/A que realiza la función inversa, es decir, convierte la secuencia de bits PCM a los valores de tensión del cuantificador. Generalmente el convertidor D/A también incorpora un filtro paso bajo para pasar a la señal analógica que será aplicada a los amplificadores.

Es importante distinguir entre los tres conceptos, muy próximos, que han aparecido en este apartado: el **cuantificador**, el **codificador** y el **convertidor A/D**.

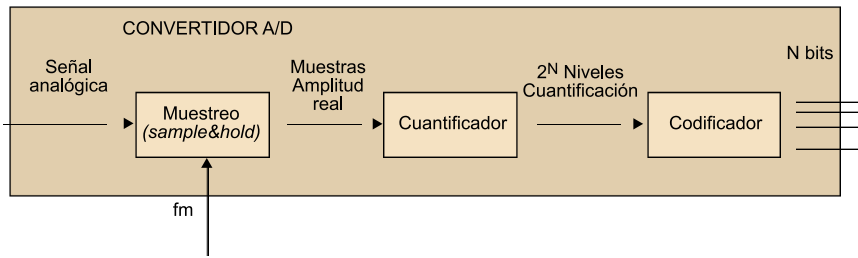
- El **cuantificador** aproxima la tensión de la señal de entrada por un nivel elegido entre un número finito de posibilidades. Un cuantificador de 16 niveles sólo admite 16 posibles valores de tensión en la salida.
- El **codificador** (PCM) asocia una palabra binaria a cada uno de estos niveles finitos del cuantificador. En el caso de que el cuantificador sea de 16 niveles se asignará una palabra de 4 bits a cada uno de los niveles¹¹.
- El **convertidor A/D** es una realización práctica de un sistema de digitalización que convierte la señal analógica en una secuencia de bits codificada en PCM. Internamente, el convertidor aproxima la señal de entrada a uno de los niveles de cuantificación pero en la salida nos proporciona directamente el código PCM asociado. Si conectamos un convertidor D/A¹² directamente en la salida de un convertidor A/D obtendremos los niveles de cuantificación a los que ha sido aproximada la señal de entrada.

⁽¹¹⁾Generalmente, el número de niveles del cuantificador es una potencia de 2 debido a que el codificador utiliza palabras código con un número de bits fijo. En efecto, pueden asociarse hasta 2^N niveles de cuantificación a una palabra de N bits.

⁽¹²⁾Muchos circuitos integrados D/A incluyen un filtro analógico paso bajo para reconstruir la señal analógica, por lo que en estos casos no es posible obtener los niveles de cuantificación originales.

La figura 62 intenta relacionar de forma esquemática estos elementos. El subsistema de muestreo se identifica con los circuitos de muestreo y mantenimiento que se encargan de congelar el nivel de tensión de la señal de entrada en el instante de muestreo.

Figura 62



Representación gráfica de la relación entre el cuantificador, el codificador y el convertidor A/D.

En la práctica, la secuencia de bits asociada a una señal PCM requiere algún tipo de señalización adicional que ayude a interpretar cuándo empieza o acaba un paquete de muestras. Esta señalización depende de la aplicación y puede variar considerablemente en función de ésta.

- Un posible ejemplo sería el formato CD-audio donde las muestras PCM se estructuran en paquetes de 33 bytes. Estos 33 bytes contienen un total de seis muestras PCM de cada uno de los dos canales (6 muestras x 2 canales x 2 bytes/muestra = 24 bytes) más 9 bytes adicionales que se utilizan para sincronismos, subcódigos de control y de display y bits de redundancia para la corrección de eventuales errores.
- Otro ejemplo son los formatos WAV, AUD, AIFF, etc., que especifican las cabeceras y las agrupaciones de las muestras PCM en ficheros de audio.

3.3.5. La selección de las palabras código en PCM (opcional)

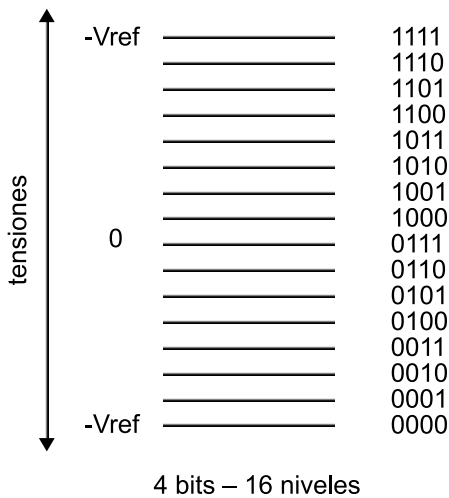
Hemos visto que un codificador PCM asigna una palabra código de N bits a cada uno de los 2^N niveles del cuantificador. En principio, la palabra código asignada a cada nivel puede ser arbitraria siempre que cada nivel tenga una y sólo una palabra código asociada, o lo que es lo mismo, pueda establecerse una **relación biunívoca** entre palabras código y niveles de cuantificación.

La pregunta que nos formulamos en este apartado es si existe un criterio para la asignación de las palabras código que resulte más conveniente que otros.

Vamos a considerar como posible asignación de palabras código la que se muestra en la figura 63 y que consiste en numerar consecutivamente cada uno de los niveles, desde el más bajo (palabra todo ceros) hasta el más alto (palabra todo unos). Parece ésta una asignación de palabras código bastante natural y que no carece de sentido práctico pues puede establecerse una fórmula sencilla que relacione el nivel de cuantificación con el número binario.

Figura 63

Codificación de los niveles



Asignación lineal de los códigos binarios a los niveles de cuantificación. Ejemplo con 16 niveles/4 bits.

En efecto, la fórmula que relaciona el nivel de cuantificación con la palabra binaria viene dada por:

$$V_{in} \cong 2V_{REF} (b_1 2^{-1} + b_2 2^{-2} + b_3 2^{-3} + b_4 2^{-4}) - V_{REF}$$

donde resulta fácil comprobar que la palabra todo ceros quedará asignada con la tensión $-V_{REF}$.

La palabra todo unos quedará asignada al **nivel de cuantificación máximo** que en este caso es $(7/8)V_{REF}$, como puede comprobarse si se sustituyen todos los bits por 1 en la ecuación anterior.

Aunque esta asignación parece bastante simple, directa e incluso práctica, adolece de un problema fundamental y es que no pueden aplicarse operaciones aritméticas directamente sobre los valores codificados. ¿Qué queremos decir con esto? y ¿por qué es tan importante?

Cuando digitalizamos una señal es habitual que queramos realizar operaciones digitales sobre la misma. Estas operaciones pueden ser un filtrado, una ecualización, efectos sonoros, etc. En cualquier caso, las muestras deben multiplicarse, sumarse, etc., por lo que sería de desear que los resultados de aplicar las operaciones a los niveles de cuantificación y a las palabras código fueran las mismas.

Pongamos un ejemplo para que quede más claro:

Con el sistema de codificación lineal anterior, si tenemos la muestra 0111 y la muestra 1000, el resultado de la suma de los códigos binarios es 1111. Hasta aquí todo parece correcto, pero ahora cuestionémonos que niveles de cuantificación representan los códigos anteriores y veremos que hemos cometido un error importante. En efecto, el código 0111 se corresponde con el nivel $-V_{REF}/8$ y el código 1000 con el nivel 0 V. Evidentemente, al sumar ambos niveles nos esperamos obtener como resultado $-V_{REF}/8$, es decir, el código 0111 y sin embargo, hemos obtenido el código 1111 que se corresponde con el nivel de $7V_{REF}/8$.

Parece que las propiedades de nuestro código no son tan buenas como nos esperábamos. No podemos aplicar directamente las operaciones aritméticas sobre las palabras código puesto que sus resultados no se corresponden con las operaciones aritméticas aplicadas sobre los niveles de tensión. Pensad que realmente la información importante está en los niveles de cuantificación, en las tensiones asociadas a las muestras, ya que al final de la cadena siempre hemos de volver a convertir la señal a analógica para poder escucharla. **¿Cómo podemos solucionar este problema?** Existen dos alternativas, una mucho más compleja que la otra:

1) **Mantener el código de asignación lineal** pero realizar las operaciones mediante tablas. Esto significa que sabemos que las operaciones sobre las palabras binarias no conservan las propiedades aditivas y multiplicativas y que nos construimos nuestras propias tablas. Por ejemplo, en la tabla de suma tendremos la regla siguiente: $1000 + 0111 = 0111$ (por coherencia con los resultados obtenidos en el párrafo anterior). Con esta solución sumar (y multiplicar) se reduce a buscar resultados en tablas. Hemos encontrado una solución al problema pero es una solución compleja y poco eficiente, ya que buscar los resultados en una tabla lleva mucho más tiempo que aplicar las reglas de la aritmética, para lo que existen procesadores especialmente diseñados. Además, necesitaremos una memoria adicional para almacenar las tablas de resultados.

2) La segunda solución consiste en **encontrar un sistema de codificación** que conserve las propiedades de la suma y el producto y sobre el que puedan aplicarse las operaciones directamente sobre las palabras código. Esta codificación existe, es la que más se utiliza en la práctica y se conoce con el nombre de complemento a 2.

3.3.6. PCM en complemento a 2 (CA2) (opcional)

La codificación en complemento a 2 (CA2) consiste en numerar consecutivamente en binario natural los niveles de cuantificación positivos.

Así, el nivel de cuantificación de 0 voltios recibe la palabra código todo ceros. El primer nivel de cuantificación positivo recibe la palabra 0001 (estamos suponiendo un cuantificador de 4 bits) y así sucesivamente. Los niveles de cuantificación positivos quedan pues codificados en binario natural.

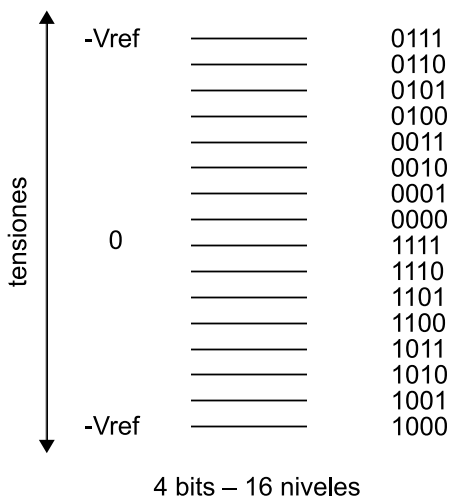
Para obtener la palabra código asociada a un nivel de cuantificación negativo debemos proceder de acuerdo con el siguiente algoritmo (ponemos un ejemplo para codificar el nivel -5):

- 1) Escribir la palabra código del nivel positivo asociado (+5 \rightarrow 0101)
- 2) Complementar a 1 la palabra anterior 1010 (cambiar los ceros por 1 y viceversa)
- 3) Sumar la unidad a la palabra obtenida en la etapa anterior 1011 (CA2).

La notación en CA2 es muy eficiente, ya que simplemente requiere complementar a 1 y sumar la unidad a los números negativos para obtener el código que los representa. Además, podemos identificar fácilmente todos los números negativos que son los que empezarán por 1. A partir del código en CA2 podemos obtener el valor absoluto del número aplicando el mismo algoritmo, es decir, volviendo a complementar a 2 el número.

En la figura 64 se muestra el resultado de asignar un código en CA2 a un cuantificador de 4 bits. Nótese que todos los niveles positivos empiezan por cero mientras que los negativos empiezan por 1.

Figura 64



Asignación de los códigos mediante complemento a 2 a un cuantificador de 4 bits.

Para comprobar que el CA2 mantiene las operaciones aritméticas, podemos poner un ejemplo con los niveles 6 (0110) y -3 (1101). En este caso la suma debería ser 3, que se corresponde con $0110 + 1101 = 0011$.

3.3.7. La tasa de bits y el almacenamiento de la codificación PCM

La digitalización de la señal y su codificación en PCM significan una forma alternativa para transportar o almacenar la señal de audio. En sistemas analógicos los mecanismos de transporte más utilizados han sido las modulaciones radio (AM o FM) o cable (telefonía), mientras que los sistemas de almacenamiento se han basado fundamentalmente en cinta magnética o discos de vinilo. En todos los casos, se envía o almacena una señal analógica que representa la información acústica que queremos reproducir a distancia o en el futuro. En digital los mecanismos para el transporte o el almacenamiento siguen sien-

do, en muchos casos, los mismos (radio, cable, cinta magnética, etc.), aunque en algunas aplicaciones han aparecido nuevas tecnologías de soporte (discos magnéticos, soportes ópticos, etc.).

El cambio fundamental de la transición al mundo digital es la naturaleza de la propia señal, constituida ahora por bits que soportan la información de los niveles de cada muestra.

En este apartado vamos a calcular uno de los parámetros más importantes de los sistemas de audio digital: la **tasa de bits asociada**. La tasa de bits nos permite calcular, entre otras cosas, cuánto espacio de almacenamiento necesitaremos para almacenar una canción, cuánto ancho de banda se requiere enviar un canal de radio o si es posible enviar un material audiovisual para ser recibido en tiempo real a través de Internet.

La tasa de bits es el número de bits por segundo que produce el proceso de digitalización de una señal.

En efecto, cuando digitalizamos una señal en PCM tomamos muestras de la misma de manera periódica y cada muestra la codificamos con un número de bits fijo. Por lo tanto, a medida que estamos digitalizando la señal se está produciendo un flujo constante de bits que la representan y que deberán ser almacenados o enviados a un hipotético receptor. La tasa de bit depende por tanto de los parámetros básicos de la digitalización: la frecuencia de muestreo y el número de niveles de cuantificación. Vamos a poner algunos ejemplos que nos ayuden a entender este concepto y relacionarlo con la capacidad de almacenamiento y la velocidad de transmisión.

Para calcular la tasa de bits por segundo de un determinado sistema digital, debe calcularse el producto entre el número de bits por muestra y el número de muestras por segundo. En función del sistema pueden incluirse varios canales de audio (estéreo, multicanal, etc.) que también deben tenerse en cuenta. Así, en general, tenemos:

$$R_T = M_{\text{canales}} \times f_{\text{muestreo}} \times N_{\text{bits}}$$

Ejemplo 1: Compact disc audio

Hemos visto que el sistema Compact disc audio utiliza una frecuencia de muestreo de 44.100 Hz y que cada muestra se codifica con 16 bits. Al tratarse de un sistema estereofónico deben codificarse dos canales de audio. Podemos calcular la tasa de bits multiplicando estos tres parámetros:

$$R_{CD} = 2 \text{ canales} \times 44.100 \frac{\text{muestras}}{\text{segundo}} \times 16 \frac{\text{bits}}{\text{muestra}} = 1.411.200 \text{ bps} = 1.411,2 \text{ kbps}$$

Es decir, la señal de audio digital de un CD produce un total de 1,4 Mbps, lo que representa una tasa ciertamente considerable que puede limitar su uso en algunas aplicacio-

nes. En efecto, en un reproductor de audio portátil con una capacidad de 4 GB pueden almacenarse un total de 70 horas de música codificada en MP3. En cambio, si almacenamos la música directamente en el formato de CD-audio tenemos una capacidad de:

$$\frac{4 \text{ GB}}{1,41 \text{ Mb/s}} = \frac{4.000 \times 8 \text{ Mb}}{1,41 \text{ Mb/s}} = 22.695 \text{ s} = 6,3 \text{ horas}$$

Existe por tanto una relación de aproximadamente un factor 10 respecto a la tasa de bits del formato CD al MP3. Esta reducción se hace aumentando considerablemente el procesado que debe realizar tanto el codificador como el decodificador de MP3. Además, la compresión es con pérdidas, lo que significa que la señal que reproducimos es sólo una aproximación a la señal original del CD. No obstante, los algoritmos de codificación utilizados actualmente utilizan modelos avanzados del sistema auditivo humano para lograr que éste no diferencie entre la señal original en CD y la señal comprimida.

Ejemplo 2. Voz digital para telefonía

En telefonía digital se elige una frecuencia de muestreo de 8 kHz y se utiliza un cuantificador de 8 bits. Con ello se obtiene una calidad aceptable que permite la inteligibilidad de la voz y el reconocimiento del locutor. Examinaremos con mayor detalle estos parámetros en una actividad posterior. La tasa de bits obtenida para estos parámetros de digitalización es:

$$R_T = 8 \frac{\text{bits}}{\text{muestra}} \times 8.000 \frac{\text{muestras}}{\text{s}} = 64.000 \text{ bits/s} = 64 \text{ kbps}$$

Históricamente, este valor de 64 kbps tiene cierta importancia ya que con los primeros sistemas digitales de telefonía, que se denominaron **Red digital de servicios integrados (RDSI)** se proporcionaban canales de 64 kbps al usuario, es decir, cuando un abonado contrataba una línea RDSI se le proporcionaba un ancho de banda de 64 kbps. El valor concreto de 64 kbps se debía a que en la línea digital debía de poder acomodarse una señal de voz digitalizada con la calidad de telefonía. Compárese este valor de 64 kbps con los anchos de banda que se ofrecen con las líneas ADSL (12 Mbps-20 Mbps) lo que supone un incremento en un factor de aproximadamente 300.

3.3.8. Niveles de cuantificación y sistemas de audio

El número de niveles de cuantificación utilizado en un sistema de audio depende de sus características y de la calidad final que deseemos obtener. Cuando se definió el sistema **CD-audio** el objetivo era obtener un sistema de alta fidelidad, cuya calidad fuera superior a todos los sistemas de audio domésticos existentes hasta la fecha. Se llevaron a cabo extensivas pruebas subjetivas para determinar el número de bits óptimo utilizando expertos en audio. Estas pruebas, de forma simplificada, han sido repetidas en esta actividad.

Los resultados obtenidos fueron que la gran mayoría de oyentes estaban bastante satisfechos con 13 bits por muestra y que no resultaba posible distinguir entre la señal original y la digital si se utilizaban 14 bits por muestra.

Durante la especificación del formato, realizada conjuntamente por Sony y Philips, se discutió profundamente si se especificaban 14 bits (Philips defendía esta opción debido a que tenía tecnología de muestreo de 14 bits en aquel momento) o se ampliaba a 16 bits (opción defendida por Sony bajo el argumento de que todas las memorias estaban estructuradas al byte, por lo que la tecnología, a largo plazo, sería más económica).

Finalmente, se eligió como **estándar** utilizar 16 bits por muestra, es decir, un total de **65.536 niveles** de cuantificación que, en principio, están más allá de la capacidad de discernir entre señales analógicas y digitales del oído humano.

14 bits

El sistema CD-audio se define a finales de los años setenta y el primer reproductor aparece en 1982 cuando la tecnología de digitalización es todavía incipiente. Por ello, la mayor parte de las grabaciones realizadas en esta primera época se hacen sólo con 14 bits dejando los dos bits menos significativos a cero.

Televisión NICAM 728

El sistema de audio digital para televisión NICAM 728 utiliza una frecuencia de muestreo de 32 kHz y transmite las muestras con un total de 10 bits. No obstante, estos 10 bits se obtienen a partir de compresión dinámica realizada sobre una señal que originalmente ha sido muestreada con 14 bits por muestra.

La compresión dinámica utilizada en el NICAM permite obtener una calidad muy parecida a los 14 bits utilizando únicamente 10 bits. ¿Cómo se consigue? El procedimiento es complejo y requiere cierto procesado de la señal. Esencialmente se trata de agrupar paquetes de 32 muestras y determinar la muestra que tiene el mayor valor absoluto. La elección de los 10 bits se hace teniendo en cuenta el valor de esta muestra de manera que sólo se enviarán a partir del primer bit que es diferente de cero. Así, por ejemplo, si el valor absoluto máximo del paquete de 32 muestras es 0010 0110 1101 11, no enviaremos los dos primeros bits, ya que toman el valor cero. Como el sistema envía solo 10 bits, tampoco se enviarán los dos últimos (que se considerarán poco significativos).

Resumiendo, si la amplitud supera un determinado umbral, los 10 bits que se transmiten son los 10 bits más significativos, descartando los 4 bits menos significativos de todas las muestras del paquete. En cambio, si el valor de la muestra es muy pequeño sólo se transmiten los 10 bits menos significativos. En general, siempre se transmiten 10 bits, que son seleccionados en función del nivel de la muestra de mayor valor absoluto. En paralelo, junto con la información de comprobación de errores, se envía información al receptor sobre los bits que han sido elegidos, para que pueda reconstruir la señal de forma aproximada. La calidad final que se obtiene con el sistema NICAM es aceptable y aunque en algunos pasajes musicales pueden apreciarse diferencias significativas respecto a la señal CD-audio, el resultado final es considerablemente superior al que obtendríamos muestreando la señal con sólo 10 bits por muestra.

Otro sistema especialmente interesante de analizar es el **DVD-audio**, en el que definen diferentes modos y calidades de audio, admitiendo cuantificadores de 16 bits, 20 y 24 bits. La pregunta que surge inmediatamente es si resulta necesario realizar un muestreo de 20 o 24 bits cuando hemos comentado (y comprobado experimentalmente en esta actividad) que el sistema auditivo humano no es capaz de apreciar diferencias entre la señal analógica original y

una señal cuantificada con 14 bits. Debe tenerse en cuenta que 24 bits representa una relación señal ruido muy importante y que incluso con el estado actual de la tecnología, existen pocos fabricantes que ofrezcan sistemas de adquisición fiables de 24 bits. **¿Cuál es pues el interés de muestrear una señal de audio a 20 o 24 bits?** Pues bien, el motivo es doble.

1) Por una parte, los sistemas de audio doméstico pueden realizar un procesamiento de la señal cada vez más avanzado. Hoy en día, es fácil encontrar equipos que acondicionan la ecualización de la señal de audio en función de las características de la sala o que añaden efectos de presencia, de énfasis, de reverberación, etc. a la señal de audio con el objeto de que su audición resulte más confortable y adecuada al entorno. Todos estos efectos se realizan aplicando operaciones a las muestras de entrada.

Estas operaciones pueden constituir un número importante de multiplicaciones y sumas con lo que se producen redondeos de los resultados parciales que se están obteniendo durante el tratamiento de la señal. Todos estos redondeos afectan a la señal de audio y serán tanto más importantes cuanto menor sea el número de bits originales con las que ha sido muestreada. El resultado final es que después de un procesado relativamente complejo, la calidad de la señal puede verse afectada debido a los efectos de redondeo de las muestras y que lo que inicialmente pensábamos que tenía una calidad comparable a los 16 bits haya sido redondeado en exceso y la relación señal ruido sea mucho menor que lo esperado. Para evitar estos efectos es conveniente efectuar las operaciones de procesado con la mayor precisión posible, por lo que es recomendable muestrear la señal con un número de bits superior a los 16.

En aplicaciones profesionales es muy importante que la señal original tenga el mayor número de bits posible, ya que durante las etapas de edición y postproducción se realiza un gran número de operaciones para el acondicionamiento de las señales que podría afectar a su calidad. Una vez la señal ha sido editada es posible reconvertirla a 16 bits para su registro en un CD-audio.

2) Otro de los aspectos con cierta importancia para considerar el aumento del número de bits es la gran dinámica que pueden tener algunas composiciones musicales, especialmente de música clásica.

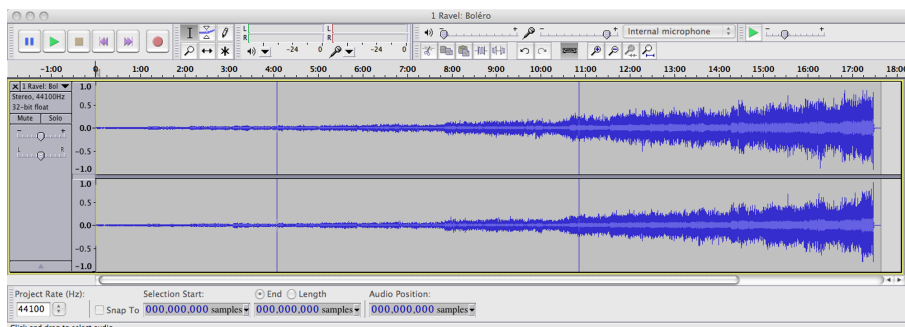
Para poner un ejemplo muy claro, consideremos el *Bolero de Ravel*, una pieza clásica, muy popular, que es una lección de orquestación de Maurice Ravel. En esta pieza Ravel utiliza una frase musical que se va repitiendo, entrando en ella, progresivamente, nuevos instrumentos de la orquesta. La composición se inicia con la percusión, con un volumen muy bajo y su potencia acústica va aumentando a medida que avanza. En una audición doméstica, con un equipo de alta fidelidad, es muy probable que pongamos el volumen inicial del equipo a un nivel excesivo y que tengamos que corregir el volumen (bajándolo) a medida que el tema vaya avanzando, ya sea por iniciativa propia o por protestas del vecino. En este caso, se trata de un ejemplo donde la señal que se ha registrado al comienzo del tema es una señal de un nivel muy bajo y que, por lo tanto, sólo utiliza un número muy reducido de los 65.536 niveles de que dispone el convertidor.

Bits

Cabe comentar en este punto que la mayor parte de los equipos profesionales son capaces de digitalizar las señales con 20, 22 o 24 bits.

Esta situación intenta ilustrarse de forma gráfica en la figura 65, donde se muestra una captura de la composición completa del *Bolero de Ravel* (la duración de la pieza musical es de unos 18 minutos).

Figura 65. Forma de onda del *Bolero de Ravel*



Obsérvese que el nivel de la señal durante los primeros minutos es muy reducido, por lo que no todos los niveles de cuantificación son utilizados.

Es importante observar en esta gráfica cómo el nivel de volumen va aumentando progresivamente, aprovechando toda la dinámica del convertidor en la parte final de la composición. No obstante, queda claro que durante los primeros minutos la amplitud de la señal es muy pequeña y sólo se utiliza una pequeña parte de todos los niveles de cuantificación.

Desde un punto de vista práctico, esto significa que durante esta primera fase de la composición la relación señal ruido que tenemos es considerablemente más baja que la que tenemos en la parte final.

La solución a este problema consiste en reducir también el ruido de cuantificación, de manera que la relación señal ruido se mantenga a unos niveles aceptables para el oído. La reducción de la potencia de ruido de cuantificación significa reducir su amplitud aumentando el número de bits del cuantificador. Es en estas circunstancias y en este tipo de pasajes y composiciones musicales donde puede tener cierta importancia el aumento del número de bits a 20 a 24.

Relación señal ruido

Obsérvese que la relación señal ruido disminuye debido a que baja la potencia de la señal, manteniendo la del ruido. No obstante, si este efecto es muy importante es posible que el ruido de cuantificación llegue a ser audible.

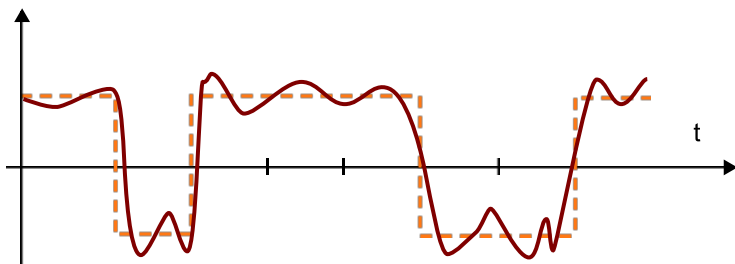
3.3.9. Ventajas e inconvenientes de la digitalización de la señal de audio

Son bastantes las ventajas que ofrecen los sistemas de audio digital respecto a los sistemas analógicos convencionales. El avance de la tecnología ha facilitado enormemente el desarrollo de los sistemas digitales, haciendo que en general sean, hoy por hoy, menos costosos que los analógicos y con unas ventajas muy significativas en cuanto a almacenamiento y posibilidades de difusión. Deberían destacarse, entre otras, las siguientes ventajas:

- **La calidad y la capacidad de reproducir el sonido en muy alta fidelidad** que ofrecen los sistemas digitales no tiene comparación con lo que podría obtenerse con sistemas analógicos. Hemos visto que si elegimos el número de bits del cuantificador suficientemente elevado no es posible distinguir entre una señal original y su aproximación digital. En estas con-

diciones, la principal virtud de los sistemas digitales es que la señal no se degrada debido a su almacenamiento o a su difusión. En efecto, una de las características principales de los sistemas digitales es su capacidad de regeneración, es decir, que si recibimos la señal con ruido, distorsiones, interferencias, rayadas del medio, etc., casi siempre es posible determinar el valor original de los bits (véase la figura 79). Esto significa que la calidad se mantiene prácticamente inalterada aunque las condiciones de transmisión o de almacenamiento no sean las ideales. Así, por ejemplo, un CD admite un cierto grado de suciedad y rayadas sin que se vea afectada la calidad de su reproducción. Con el mismo nivel de suciedad y rayadas en un disco de vinilo, la calidad quedaría gravemente afectada.

Figura 79. Comparativa entre una señal digital (línea discontinua) y la misma señal deteriorada por ruido, distorsiones, etc.



Si el grado de ruido no es excesivo, parece razonable que un procesador pueda recomponer la información original.

- **Inserción de metadatos.** El hecho de codificar las señales mediante una secuencia de bits repercute en que pueda incorporarse otro tipo de información diversa además del audio. Esta información puede contener los títulos de los temas musicales, los intérpretes, el género musical, etc.
- **Protección de los contenidos.** La información en formato digital puede encriptarse de forma muy eficiente, dificultando o impidiendo el acceso a la información a aquellos usuarios que no estén autorizados. Los contenidos pueden protegerse para preservar la confidencialidad, autenticar la autoría o proporcionar servicios de pago mediante acceso condicional.
- **Procesamiento digital de la información.** Al tratarse de información digital puede ser transformada directamente mediante procesadores digitales avanzados, que pueden ejecutar un elevado número de operaciones por segundo para adaptar las señales a las características del auditorio o generar los efectos deseados. Además, estos algoritmos son muy flexibles en el sentido de que pueden modificarse con facilidad, solo con cambiar el software o el firmware del procesador.
- **Capacidad de insertar códigos correctores.** Ya hemos comentado que los sistemas digitales son más robustos frente a los efectos no deseados que sufre la señal al ser transmitida por el canal. No obstante, si el nivel de ruido es suficientemente importante puede provocar errores en el proceso de regeneración de los bits. Estos errores pueden ser corregidos si se introducen **códigos redundantes** que protejan la señal frente a los errores. La

mayor parte de los sistemas digitales incluyen códigos redundantes para la protección de errores que les confieren una gran robustez.

- **Compresión de los datos.** Aunque la codificación PCM requiere una tasa de bits muy importante (1,4 Mbps para el caso del CD-audio) la información digital puede ser procesada y comprimida en unos factores muy importantes sin que el oído humano pueda apreciar las diferencias entre la información original y la comprimida. Los factores de compresión que se consiguen con los codificadores MP3 y MP4 son del orden 10 (el tamaño del fichero ocupa 10 veces menos).

Compresores de datos

Estos compresores permiten que puedan almacenarse señales de alta calidad en dispositivos de memoria de estado sólido de tamaño muy reducido, consiguiendo equipos portátiles muy sofisticados que de ningún modo hubieran podido obtenerse utilizando sistemas analógicos.

3.4. Problemas/preguntas

1) Buscad información y elaborad un breve resumen sobre las características básicas del formato WAV de audio. Intentad identificar si existen diferentes versiones, qué tipo de cabeceras se incluyen, qué información llevan las cabeceras sobre algunos parámetros básicos del PCM como la frecuencia de muestreo, el número de bits por muestra, el número de canales. Identificad cómo se empaquetan las muestras de audio.

2) Buscad información sobre el empaquetamiento de la trama de muestras en el sistema NICAM 728¹³. Elaborad un breve informe donde se proporcione respuesta a las siguientes cuestiones:

- ¿Cuánto dura una trama de NICAM 728?
- ¿Cuántas muestras se envían?
- ¿Cuántos canales se envían?
- ¿Qué modos pueden identificarse?
- ¿Cómo se realiza la selección de los 10 bits por muestra?
- ¿Cómo se señala, mediante la paridad, los 10 bits que han sido elegidos?
- ¿Por qué recibe el nombre de NICAM 728?

3) Buscad un par de ejemplos musicales, distintos al *Bolero de Ravel*, en el que la dinámica de la composición recomiende aumentar más allá de 16 el número de bits del cuantificador.

4) Construid una tabla resumen en la que se incluyan distintos sistemas de audio doméstico indicando la frecuencia de muestreo, el número de bits de cuantificación y si utilizan codificación PCM o realizan algún tipo de compresión de los datos. Si queréis, podéis ampliar esta tabla a sistemas de audio profesionales.

⁽¹³⁾No es necesario que este informe incluya información sobre la modulación utilizada para la transmisión de los bits junto con la señal de televisión. En esta experiencia sólo estamos interesados en la parte de la información digital de audio.

5) Identificad productos comerciales de adquisición de señales de audio que puedan trabajar con resoluciones de 16, 20 y 24 bits en un entorno de edición basado en ordenadores personales tipo PC o MAC. Identificad en las características técnicas de estos productos si las relaciones señal ruido que se proporcionan coinciden con las que esperamos obtener de productos que sean capaces de capturar a las resoluciones mencionadas.

6) Comparad las características de un sistema de adquisición de audio integrado en un ordenador portátil (elegid una marca/modelo determinado) con las características que pueden obtenerse de módulos de muestreo que se conectan en bus PCI express. Comparad las relaciones señal ruido, el número de bits, el número de canales que pueden muestrearse, etc.

7) Abrid el archivo **NessunDorma.wav** con el programa Audacity y determinad su tamaño a partir del número total de muestras, el número de canales y sabiendo que cada muestra está codificada con 16 bits. Anotad el número de bytes que obtiene. En paralelo, con el explorador de Windows o el Finder, determinad el tamaño del **archivo *.wav** que representa esta señal (para saber el número exacto de bits debéis entrar en propiedades del fichero). ¿Coinciden ambos tamaños? Si no coinciden, ¿podéis explicar a qué es debido? ¿Qué porcentaje de bits en un fichero WAV corresponden a información de cabeceras/sincronización?

8) Digitalizad una señal musical procedente de un sistema de reproducción Compact Cassette analógico y mediante el programa Audacity utilizando una frecuencia de muestreo de 44.100 Hz y 16 bits por muestra. Evaluad el ruido con el que ha sido capturada la señal. Posiblemente, si no se dispone de un sistema de captura profesional, el nivel de ruido existente en la señal será muy superior al ruido de cuantificación. Intentad medir experimentalmente este ruido a partir de la forma de onda de la señal o de su espectro. A partir de los resultados obtenidos, determinad el número de bits mínimo del cuantificador que mantiene el ruido de cuantificación por debajo del ruido de medida.

9) Una imagen digital puede considerarse como el muestreo en el espacio (filas y columnas) de una imagen analógica. Intentad generar una imagen en blanco y negro con varios niveles de gris (p. ej. con el Photoshop). En la medida en que sea posible, modificad el número de niveles de gris que se presentan por pantalla e intentad determinar cuántos niveles de gris es capaz de discriminar el sistema visual humano. Comparad el número de niveles de gris que puede discriminar el sistema visual humano con el número de niveles de cuantificación que se necesitan en la señal de audio.

3.5. Evaluación

1) Determinad la relación señal ruido de un sistema de DVD-audio que utiliza 24 bits por muestra. Calculad la fracción de la potencia de señal que representa el ruido de cuantificación en este sistema.

Solución

La relación señal ruido en un sistema de 24 bits vendrá determinada por:

$$SNR = 1,76 + 6 \cdot 24 = 145,76 \text{ dB}$$

A partir de este valor puede calcularse la relación entre la potencia de señal y la potencia de ruido, que vendrá dada por:

$$10 \cdot \log_{10} \frac{P_{\text{signal}}}{P_{\text{Noise}}} = 145,76 \Rightarrow \frac{P_{\text{signal}}}{P_{\text{Noise}}} = 10^{14,576} \Rightarrow P_{\text{Noise}} = \frac{P_{\text{Signal}}}{3,77 \cdot 10^{14}} = 2,65 \cdot 10^{-15} P_{\text{signal}}$$

2) Supongamos un cuantificador de 5 bits donde se utiliza una codificación en CA2 y que puede trabajar con unas tensiones máxima y mínima de +/-5 voltios respectivamente. Determinad:

- El paso de cuantificación.
- El error máximo que se cometerá al aproximar una tensión analógica por un nivel de cuantificación.
- Calculad la SNR del convertidor e indicad, a partir de las experiencias realizadas en esta actividad, si creéis que el oído humano será capaz de apreciar el ruido de cuantificación.
- Determinad los códigos asociados a los niveles +13 i -7.
- Comprobad que la suma de los códigos en CA2 del apartado anterior coincide con la suma de los niveles asociados.

Solución

a) El paso de cuantificación puede calcularse teniendo en cuenta que todo el margen de valores de entrada (10 voltios) debe ser repartido de forma uniforme entre los niveles de cuantificación.

$$\Delta = \frac{5V - (-5V)}{2^5} = \frac{10}{32} = 0,3125 \text{ V}$$

b) Hemos visto que el error máximo siempre es la mitad del paso de cuantificación, por lo tanto tendremos un error máximo de 0,15625 voltios.

c) La SNR será la correspondiente a un convertidor de 5 bits. Por lo tanto, aplicando la fórmula obtenemos una relación señal ruido de:

$$SNR = 1,76 + 6,02 \cdot 5 = 31,86 \text{ dB}$$

De acuerdo con las experiencias realizadas en esta actividad, el ruido de cuantificación¹⁴ que se obtiene con 5 bits es audible.

d) El código asociado al número 13 es directamente su código binario natural, es decir 01101. Para determinar el código asociado al -7 debemos expresar el número 7 en binario natural, complementarlo a 1 y sumarle la unidad. Obtenemos:

- Código binario natural 7: 00111

- Complemento a 1 (-7): 11000
- Complemento a 2 (-7): 11001

e) Si sumamos en binario natural los dos números anteriores obtenemos $01101+11001 = 00110^{15}$, que coincide con el código en CA2 del número 6, resultado de sumar 13 y -7.

⁽¹⁴⁾Debe observarse que la SNR no depende de los niveles de entrada del convertidor que son +/- 5 voltios en este caso. Esto es así porque suponemos que la señal está acondicionada para cubrir todo el margen de valores de entrada que admite el convertidor. El paso de cuantificación también es proporcional al margen de entrada, de manera que si se aumenta el margen de niveles de entrada también aumenta el nivel de ruido, manteniendo la relación señal ruido constante. Esta es una de las peculiaridades de la fórmula que estamos utilizando: la calidad (SNR) de un convertidor sólo depende del número de bits.

⁽¹⁵⁾Notar también en este punto que cuando se realiza la suma de $01101+11001$ obtenemos un *carry* igual a la unidad (es decir, el resultado de la suma es de 6 bits 100110). Este *carry* se ignora en todas las operaciones realizadas en CA2, ya que estamos trabajando con 5 bits. Esta circunstancia debe tenerse en cuenta cuando se definen los algoritmos de procesado, ya que pueden producirse desbordamientos en función de los valores de los números que se operan. No obstante, este problema no es específico de la notación en CA2 sino que también se produce en binario natural o en CA1.

3) Determinad la tasa de bit de un sistema de audio digital en el que se muestrean 5 canales de audio convencionales y un canal de subwoofer (audio multicanal). Los 5 canales de audio se muestrean a 96 kHz tomando un total de 20 bits por muestra. El canal de subwoofer se muestrea a 300 Hz tomando 16 bits por muestra.

Solución

Para determinar el número total de bits por segundo basta considerar:

- Número de bits por segundo canal convencional:

$$96.000 \text{ muestras/s} \times 20 \text{ bits/muestra} = 1.920.000,0 \text{ bps}$$

- Número de bits por segundo canal subwoofer:

$$300 \text{ muestras/s} \times 16 \text{ bits/muestra} = 4.800,0 \text{ bps}$$

Así, el total de bits por segundo es el resultado de sumar el canal de subwoofer por los 5 canales convencionales:

$$5 \times 1.920.000 + 4.800 = 9.604.800 \text{ bps}$$

4) Un estudio de grabación desea enviar urgentemente un total de 100 cintas de una hora de duración en formato DAT (digital audio tape). Las cintas son estereofónicas (2 canales) y están registradas a una frecuencia de muestreo de 48 kHz a 16 bits por muestra. Nos planteamos enviar esta información a través de una línea dedicada que puede transmitir 2 Mbps o bien enviarlo como un paquete a través de Courier, donde tenemos garantizada la entrega en 24 horas. ¿Cuál de los dos sistemas resulta más rápido? ¿Cuál es el número crítico de cintas a partir del cual un sistema resulta más rápido que el otro?

Solución

En primer lugar vamos a determinar el número total de bits que tenemos en las 100 cintas de DAT.

- El número de bits por segundo del DAT será:
 $2 \text{ canales} \times 48.000 \text{ muestras/s} \times 16 \text{ bits/muestra} = 1.536.000 \text{ bps}$
- El número de bits contenido en una cinta (1 hora = 3.600 s) será:

$$1.536.000 \text{ bps} \times 3.600 \text{ s} = 5,5296 \cdot 10^9 \text{ bits}$$

- Por lo que el número de bits que deben transmitirse para enviar las 100 cintas son:
Número de bits (100 cintas DAT); $5,5296 \cdot 10^{11}$ bits

Si la línea por la que realizamos la transmisión tiene una capacidad de 2 Mbps, necesitaremos un total de $5,5296 \cdot 10^{11} / 2 \cdot 10^6 = 276.480$ segundos, que se corresponden con 4.608 minutos o 76,8 horas. En definitiva, resulta más rápido enviar las cintas a través de un Courier por servicio de 24 horas.

Para determinar el número crítico de cintas a partir del cual un servicio resulta más rentable que el otro debemos calcular el número total de bits que pueden transmitirse en 24 horas mediante una línea de 2 Mbps:

$$2 \cdot 10^6 \text{ bits/s} \times 24 \text{ horas} \times 3600 \text{ s/hora} = 1,728 \cdot 10^{11} \text{ bits}$$

Como cada cinta tiene un total de $5,5296 \cdot 10^9$ bits, tenemos que el cociente entre ambos números nos da un total de 31,25 cintas. Es decir, si el número de cintas es 31, sale rentable enviarlas a través de nuestra línea de 2 Mbps.

5) Repetid el problema anterior pero ahora suponiendo que queremos transmitir 100 cintas de DV (digital vídeo) de una hora de duración. Suponed en este caso que la tasa de bits del sistema DV es de 25 Mbps.

Solución

El problema es prácticamente idéntico al anterior, pero ahora la capacidad de bits de una cinta de DV es francamente superior. En efecto, pasamos de los 1,536 Mbps que están registrados en un DAT a los 25 Mbps que están registrados en la cinta DV. La relación entre ambos factores es que las cintas DV tienen una capacidad 16,276 veces superior a una cinta de DAT.

Según esta relación podemos extrapolar los resultados:

- a) Se necesitan 1.250 horas para enviar las 100 cintas de DV a través de una línea de 2 Mbp.
- b) El límite para que resulte rentable enviar las cintas a través de la línea de 2 Mbps¹⁶ es de 1,92 cintas.

⁽¹⁶⁾ Observad en este problema que la tasa de bits contenida en la cinta DV es considerablemente superior a la velocidad de transmisión que estamos considerando. En efecto, suponemos que se transmiten bits a una velocidad de 2 Mbps y que la cinta contiene un total de 25 Mbps. Esto significa que para transmitir un fragmento de vídeo correspondiente a un segundo necesitaremos un total de 12,5 segundos. Por tanto, la transmisión no puede realizarse en tiempo real. Por la misma regla de tres, la transmisión de una cinta de 1 hora nos ocupará un total de 12,5 horas. Nótese que el caso es considerablemente distinto al del DAT, en el que la transmisión podía realizarse en tiempo real, ya que la capacidad del canal era superior a la capacidad de la cinta.

4. Actividad 4. Filtrado de señales de audio

4.1. Introducción

4.1.1. Contenido

En esta actividad se presentan los conceptos básicos sobre el filtrado y la ecualización de señales de audio, partes esenciales en el procesado digital de señales de audio. Debido a su importancia, esta actividad se ha dividido en 2 partes con el objetivo de facilitar la consolidación de los conocimientos.

En esta primera parte nos centraremos principalmente en los conceptos de **filtrado** y en la concatenación de varios filtros para la obtención de los efectos deseados.

La segunda parte se dedicará principalmente a los **ecualizadores paramétricos y gráficos**. En todo momento se intenta que los conceptos que se presentan estén acompañados de ejercicios prácticos que permitan mostrar de forma auditiva los efectos de un determinado tipo de filtrado.

La actividad se inicia con una presentación de los diferentes tipos de filtros que se usan en la práctica para el acondicionamiento de señales de audio y se intenta distinguir entre los denominados **filtros** y los denominados **ecualizadores**. A continuación se presentan los filtros paso bajo y paso alto que serán considerados como filtros básicos a lo largo de la actividad y sobre cuya base puede construirse cualquier otro tipo de filtro. Se experimenta con diferentes programas para ver los efectos de estos filtros sobre las señales de audio.

Posteriormente se presentan los principios de **conexión de filtros en serie y en paralelo**, lo que permite introducir los filtros paso banda como una conexión en serie de un filtro paso bajo y uno paso alto. También se experimentará sobre diferentes señales y programas para ver los efectos y aplicaciones de estos filtros.

Finalmente, se introducen los **filtros de shelving**, que suelen considerarse como una primera versión de ecualizadores. Los ecualizadores más utilizados como el paramétrico y el gráfico se estudiarán con detalle en la segunda parte de la actividad.

4.1.2. Metodología

Se utiliza la misma metodología que se ha estado utilizando en actividades anteriores y que pretende que el alumno pueda ir comprobando los conceptos expuestos mediante la audición de los efectos del filtrado de las señales. En esta actividad se sigue utilizando el programa Audacity para la captura, edición y procesado de las señales de audio, pero también se introduce un **applet de Java** que implementa diferentes tipos de filtrado sobre la señal de audio.

La principal ventaja del applet de Java es que permite modificar los parámetros de los filtros en tiempo real observando su efecto de forma inmediata.

En cualquier caso, a medida que se introducen los conceptos se proponen una serie de ejercicios con estos programas. Los ejercicios son guiados y no deberían representar grandes dificultades para su realización. Lo más importante es que el alumno se centre sobre los resultados y los conceptos implícitos en los ejercicios.

Al final de la actividad se proponen algunos ejercicios adicionales con los que se pretende que el alumno resuelva con un mayor grado de autonomía y que interprete correctamente los resultados obtenidos. Finalmente, se proponen varios problemas parecidos a los que han ido apareciendo como ejemplos a lo largo del texto y se incluyen las soluciones.

4.1.3. Recursos

Los principales recursos utilizados en esta actividad son:

- Software de captura y procesado de audio Audacity (licencia libre).
- Software de demostración de filtros y ecualizadores basado en applets de Java.
- Análisis de forma de onda y análisis de frecuencia de las señales mediante el programa Audacity.
- Captura de señales de voz para su ecualización.
- Archivos de audio predigitalizados con diferentes características.
- Explicaciones de los fundamentos de la ecualización y sus diferencias con el filtrado.

- Explicaciones de las aplicaciones principales de los ecualizadores.
- Propuesta de actividades de búsqueda de información en webs y recursos de Internet para profundizar en los conocimientos adquiridos en esta actividad (aplicaciones software de ecualización, fabricantes de ecualizadores, etc.).
- Actividades propuestas para verificar la adquisición de los conocimientos de esta actividad.

4.2. Objetivos

Los principales objetivos que se persiguen en esta actividad son:

- 1) Introducir los conceptos clave para el procesado de señales de audio mediante filtros.
- 2) Introducir los conceptos básicos de filtros paso bajo y paso alto.
- 3) Relacionar mediante experimentos dirigidos los efectos del filtrado sobre la calidad y percepción sonora de una señal de audio.
- 4) Relacionar los efectos del filtrado sobre la representación en frecuencia de la señal de audio.
- 5) Comprobar que las realizaciones de los filtros que utilizamos en la práctica se alejan del comportamiento ideal, permitiendo el paso de residuos de señal que resultan audibles.
- 6) Conocer los principios básicos de asociación en serie y en paralelo de filtros.
- 7) Introducir los conceptos de filtrado paso banda y filtrado de banda eliminada mediante la concatenación en serie/paralelo de filtros básicos.
- 8) Introducción a los ecualizadores y controles de tono basados en los filtros de *shelving*.
- 9) Profundizar en el manejo de varias pistas de audio, la generación de efectos de filtrado y la mezcla de las mismas.

Estos objetivos están relacionados con las siguientes **competencias de la asignatura**:

- 1) Capacidad de capturar audio e imágenes de forma eficiente y eficaz.

- 2) Capacidad de digitalizar eficiente y eficazmente contenidos de audio e imagen.
- 3) Capacidad de operar de forma digital con señales digitales de audio e imagen.
- 4) Capacidad de interpretar y modificar información auditiva de forma frecuencial.

Y con las siguientes **competencias generales del grado**:

- 1) Ser capaces de analizar un problema en el nivel de abstracción adecuado a cada situación y aplicar habilidades y conocimientos adquiridos para abordarlo y resolverlo.
- 2) Capturar, almacenar y modificar información de audio, imagen y vídeo digitales aplicando principios y métodos de realización y composición del lenguaje audiovisual.
- 3) Integrar y gestionar contenidos digitales en aplicaciones multimodales de acuerdo con criterios estéticos, técnicos y funcionales.
- 4) Atender adecuadamente consultas sobre proyectos, tecnologías y mercado de productos multimedia evaluando de forma precisa el entorno de aplicación, los recursos y las alternativas tecnológicas disponibles.

4.3. Guía

El filtrado y la ecualización de una señal es uno de los procedimientos utilizados con mayor frecuencia en la captura, edición, producción y reproducción de señales de audio. En esta actividad definiremos los conceptos de filtrado y ecualización y veremos diferentes alternativas y sistemas para realizarla. La actividad está dividida en dos partes. En esta primera nos centramos principalmente en la introducción y experimentación sobre filtros, dejando la parte de ecualización y sus aplicaciones para la segunda parte.

Para comprender el proceso de filtrado y ecualización es importante tener en cuenta la representación de la señal de audio en el dominio de la frecuencia y que ya ha sido tratada con cierto detalle en la actividad 1.

El filtrado y ecualización consiste en enfatizar o atenuar algunas regiones frecuenciales respecto a otras, con el objeto de conseguir un determinado efecto.

Es posible modificar el balance entre bajas y altas frecuencias de un instrumento musical o de una voz humana ajustando las ganancias que proporciona el ecualizador a diferentes zonas del espectro.

No es fácil distinguir entre los conceptos de filtrado y ecualización, ya que a menudo suelen confundirse e intercambiarse. De hecho, el ecualizador suele estar constituido por uno o varios filtros, lo que aumenta la confusión.

En algunos textos, se toma como definición de filtro la de un sistema cuyo objetivo es dejar pasar una determinada banda de frecuencias y eliminar otras.

Los filtros paso bajo, paso banda y paso alto que estudiaremos en esta actividad se corresponden con esta definición de filtro.

En cambio, en un ecualizador el objetivo es **enfaticar/atenuar** unas bandas de frecuencias respecto a otras, pero sin necesidad de eliminarlas completamente.

Se entiende que un ecualizador es un sistema cuya función puede ser controlada por el usuario de una forma directa e intuitiva, sin necesidad de conocer a fondo los fundamentos del filtrado de señales.

Además, existen unos tipos de filtros, denominados de *shelving* (estanterías), que pueden considerarse como a medio camino entre un filtro convencional y un ecualizador.

En la figura 66 se pretende comparar de forma esquemática un filtro convencional, un filtro de *shelving* y un ecualizador. La representación del filtro se corresponde con la gráfica de la izquierda. En este ejemplo, se trata de un filtro paso bajo, es decir, un filtro que permite el paso de las componentes de baja frecuencia de la señal (hasta 800 Hz, en nuestro ejemplo) y que elimina las componentes de mayor frecuencia. Para comprender el efecto de un filtro paso bajo sobre una señal de audio pensemos que si se aplica este filtro, la parte de baja frecuencia de la señal se oirá igual mientras que las componentes que superan la frecuencia de 800 Hz dejarán de oírse.

Para cualquier filtro se definen tres zonas, que se denominan:

- la banda de paso,
- la banda atenuada, y
- la banda de transición.

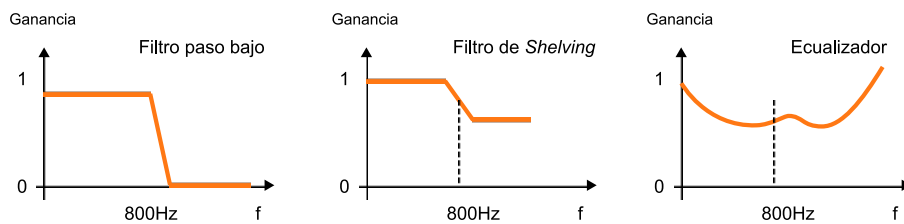
Filtros shelving

Estos filtros permiten que el usuario configure ganancias diferentes para las regiones de baja y alta frecuencia.

En el ejemplo de filtro paso bajo de la figura 66, la **banda de paso** es la zona que tiene ganancia unidad, es decir, todas aquellas componentes frecuenciales que pasarán inalteradas a través del filtro. La **banda atenuada** es la banda que tiene una ganancia cero, es decir, estas componentes de frecuencia serán eliminadas, no aparecerán en la salida. La **banda de transición** es la que se utiliza para pasar de la banda de paso a la banda atenuada. En ella, la ganancia varía de forma progresiva entre un valor unidad (no modificar la señal) y un valor cero (eliminarla).

En los filtros ideales la banda de transición no está definida, pasando directamente de la banda de paso a la atenuada. En los filtros reales tampoco es posible tener una banda de paso con un valor de ganancia exactamente igual a la unidad, ni una banda atenuada con un valor exactamente igual a cero. Entre la banda de paso y la banda de rechazo existe una banda de transición en la que la ganancia del filtro pasa del valor máximo al mínimo. De forma esquemática, el filtro se representa asignando un valor de ganancia unidad (señal sin modificar) en las frecuencias en las que permitimos el paso de la señal y un valor de ganancia cero en las frecuencias en las que no se dejan pasar, aunque en la práctica, las ganancias no serán exactamente iguales a la unidad y cero (es decir, ni se anularán completamente ni pasarán exactamente igual).

Figura 66. Representación comparativa entre un filtro convencional, un filtro de *shelving* y un ecualizador



Tipos de filtro

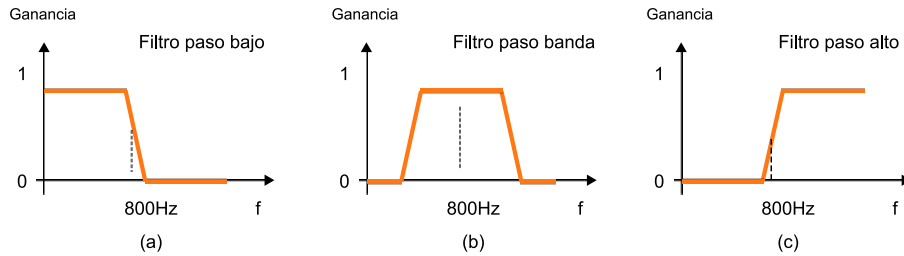
Puede haber varios tipos de filtros en función de las frecuencias que permiten pasar o eliminar. Así, por ejemplo, un filtro paso alto dejará pasar las altas frecuencias y eliminará las bajas, mientras que un filtro paso banda dejará pasar las frecuencias intermedias eliminando las componentes de alta y baja frecuencia.

En la figura 67 se representan esquemáticamente un filtro paso bajo, un filtro paso banda y un filtro paso alto (de izquierda a derecha respectivamente). El concepto de estos filtros convencionales siempre está ligado a dejar pasar unas determinadas componentes frecuenciales y eliminar el resto.

En cambio, el **filtro de shelving**, representado en la gráfica central de la figura 66 está caracterizado por tener dos ganancias distintas, una para la zona de frecuencias bajas y otra para la de las frecuencias altas. Generalmente el usuario puede controlar la frecuencia a la que se produce la transición y los valores de ganancia para cada una de las regiones. En nuestro ejemplo, la frecuencia

de transición está situada en los 800 Hz y las ganancias son de $G = 1$ para las componentes de frecuencia inferior a los 800 Hz y de $G = 0.6$ para las frecuencias superiores.

Figura 67. Ejemplos de filtros convencionales



a) paso bajo, b) paso banda, c) paso alto.

El ecualizador se representa en la parte de la derecha de la figura 66. La característica de estos filtros es que el usuario dispone de una gran flexibilidad para configurar la respuesta en frecuencia del sistema.

De hecho, el nombre de **ecualizador** procede de que es un sistema con el que trataremos de compensar alguna deficiencia o filtrado previo que ha sufrido el sonido, intentando igualar o restaurar con este sistema las condiciones originales.

Veremos que existen diferentes mecanismos para permitir que el usuario pueda configurar de forma fácil e intuitiva la forma de la curva de respuesta en frecuencia del ecualizador. En el caso concreto que está representado en la figura 66, el ecualizador aumentaría el volumen de las componentes de baja y de alta frecuencia

4.3.1. Ejemplos con filtrado paso bajo

Vamos a realizar algunos ejercicios prácticos para analizar y escuchar los efectos de un filtro paso bajo sobre una señal de audio. Para ello vamos a utilizar una aplicación Java desarrollada en la Universidad de Helmut-Schmidt, en Hamburgo (Alemania), concretamente por el profesor Udo Zölzer y que puede descargarse de la página web <http://ant.hsu-hh.de/dasp/>.

Esta página web incluye otros ejemplos de tratamiento digital de señales de audio que acompañan al libro: Udo Zölzer (agosto, 2008). *Digital Audio Signal Processing* (2.ª ed.). John Wiley.

Aunque este texto es mucho más avanzado que los contenidos que presentamos en este curso, si lo deseáis podéis descargaros otros ejemplos de cuantificación y muestreo o efectos de señales de audio. La página web incluye apli-

caciones escritas en Java y en MATLAB. En la versión Java se permite descargar la versión ejecutable o, si se desea, el código fuente¹⁷. La versión en MATLAB requiere tener instalado el MATLAB en el ordenador para su ejecución.

⁽¹⁷⁾La descarga del código fuente sólo se recomienda para aquellos alumnos que estén interesados en programar los algoritmos digitales que realizan el filtrado de las señales. Estos algoritmos son bastante complejos y no son el objetivo principal de este curso.

Alternativamente, puede descargarse la aplicación Java de la página de materiales adicionales de la asignatura. Se trata de un fichero comprimido (**filtros.zip**) que al descomprimirlo generará un directorio **Filters** en el que se incluyen varios ficheros. Descomprimid el fichero en el directorio que se desee, situadlo en ese directorio y cargad el fichero `eqApplet.html` en un navegador (Internet Explorer, Firefox, Safari). Deberéis tener instalado Java¹⁸ para que la aplicación se ejecute correctamente.

⁽¹⁸⁾Si no tenéis instalado el Java, podéis hacerlo desde: <http://www.java.com/es>. De todos modos, normalmente, si no está instalado, el propio navegador irá a la página de instalación de Java.

Además, debéis permitir que el navegador ejecute los *scripts* de esta página, opción que generalmente estará bloqueada en la configuración de seguridad por defecto del navegador.

Una vez permitáis la ejecución de *scripts* y del programa, aparece la aplicación Java de filtrado de señales que aparece en la figura 68 y que utilizaremos para diversas experiencias de esta actividad. Los botones Audio 1 y Audio 2 permiten cargar dos archivos de audio por defecto, con los que se muestran los efectos de los diferentes filtros. También aparece el botón **Load.wave file.**, que permite utilizar el fichero que desea el usuario y que usaremos más adelante para ver los efectos del filtrado respecto a diferentes capturas de audio.

La casilla de verificación **Linear Scale** se utiliza para seleccionar si la escala del eje horizontal de la gráfica es lineal o logarítmica. Debido a que el comportamiento del oído en cuanto a discriminación de frecuencias es logarítmico, para la mayor parte de las pruebas conviene no tener activada esta casilla de verificación. En la representación en escala lineal (casilla activada) las frecuencias se representan en una escala lineal entre **0 Hz** y **20 kHz**. En esta representación, los 10 kHz, una frecuencia extremadamente alta desde el punto de vista subjetivo queda en la mitad de la gráfica. Cuando la casilla no está activada la relación entre la escala horizontal y el efecto auditivo resulta más natural desde el punto de vista auditivo.

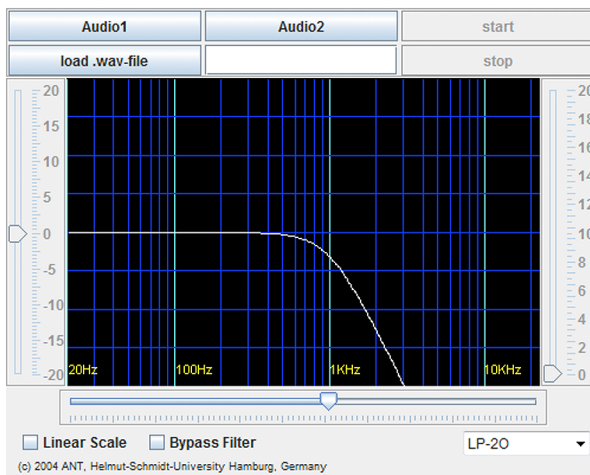
La otra casilla de verificación se denomina **Bypass Filter** y resulta muy útil para evaluar el efecto del filtro. Si la casilla no está activada, la aplicación Java procesa y modifica la señal de acuerdo con la respuesta del filtro mostrada en la gráfica.

Cuando se activa la casilla podremos oír la señal de audio original, sin ser procesada por el filtro. Resulta útil conmutar esta casilla de verificación para comprobar el efecto del filtrado sobre la señal. La barra de desplazamiento

situada por debajo del eje horizontal permite modificar la frecuencia de corte del filtro y la respuesta en frecuencia del filtro se muestra en la gráfica de forma dinámica al mover la barra de desplazamiento.

Obsérvese que, tal y como hemos definido anteriormente, el filtro intenta proporcionar una ganancia unidad en la banda de paso y una ganancia nula en la banda eliminada. Observe también que la escala en el eje horizontal es logarítmica, lo que significa que la ganancia unidad se corresponde con 0 dB's ($10 \log_{10}(1) = 0$) y que la ganancia nula de la banda atenuada se corresponderá con $-\infty$ dB's. En la caja de selección situada en la parte inferior izquierda de la aplicación podemos seleccionar distintos tipos de filtros. Para trabajar este apartado es conveniente tener seleccionado el filtro LP-2° (*low pass 2nd order*, paso bajo de segundo orden).

Figura 68. Aplicación Java eqApplet.html que muestra el efecto de un filtrado paso bajo sobre una señal de audio



1) Activad las señales de audio Audio 1 y Audio 2 y aplicad el filtro paso bajo con diferentes frecuencias de corte. Mover la barra de desplazamiento asociada a la frecuencia de corte y observad el efecto que tiene sobre el sonido. Activad y desactivad la casilla de Bypass para comparar la señal filtrada con la original. Comprobad cómo algunos instrumentos de percusión y los más agudos desaparecen al disminuir la frecuencia de corte.

2) Activad la escala lineal y modificad la frecuencia de corte del filtro paso bajo mediante la barra de desplazamiento. Anotad para qué valor de la frecuencia de corte empieza a notar cambios significativos entre la señal original y la señal filtrada. ¿Podéis dar una explicación a este fenómeno? ¿Cuál?

3) Desactivad la escala lineal (podéis mantenerla activada si lo deseáis). Observad cómo la potencia de la señal disminuye al reducir la frecuencia de corte del filtro. ¿Podéis explicar por qué?

Vamos a utilizar ahora el programa Audacity¹⁹ para capturar señal de voz y aplicar un filtrado paso bajo a la misma. El programa Audacity permite que el usuario configure las características de un filtro paso bajo, por lo que podremos analizar la señal en el dominio de la frecuencia y ver cuáles son los efectos del filtrado.

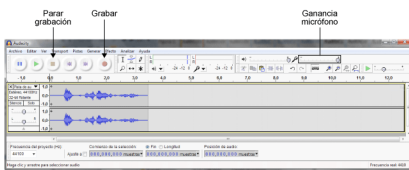
(19) Seguid el procedimiento descrito en la actividad 1 de audio para activar el programa Audacity.

1) Abrid la aplicación Audacity y capturad un fragmento de señal de voz hablando directamente delante del ordenador. Un fragmento de unos 3 o 4 segundos debería ser suficiente para realizar correctamente la experiencia, pero podéis tomar una mayor longitud si queréis. En cualquier caso, tratad de evitar silencios largos y hablad de forma constante durante la captura.

En la figura 69 se muestra una captura de la pantalla de Audacity con la adquisición de una señal de audio usando el micrófono integrado²⁰ en un ordenador. En la figura 69 también se indica cuál es la barra de desplazamiento para controlar la ganancia del micrófono, el botón de grabación de la señal y el botón de parada de grabación. Si se desea se pueden seleccionar los fragmentos previos y posteriores al registro y cortarlos (eliminarlos) utilizando el menú de **Editar + Cortar (Ctrl + X)**.

(20) Si no se dispone de micrófono integrado, podéis hacer la grabación con un micrófono externo o, si lo preferís, hacer la grabación de audio con un dispositivo externo y capturarlo posteriormente como señal de línea.

Figura 69. Captura de una pantalla de Audacity con la captura de la señal de audio procedente de un micrófono



2) Escuchad la señal de audio capturada por el micrófono varias veces. Después seleccionad toda la señal y calculad su representación en el dominio de la frecuencia. Recordad que para calcular la representación en el dominio de la frecuencia debéis ejecutar los menús **Analizar + Análisis de espectro**. El resultado para nuestro ejemplo se muestra en la figura 70. Resulta interesante notar que la señal muestra muy poco contenido en frecuencias altas, lo que es propio de la señal de voz, que tiene una energía muy baja más allá de los 3.000 Hz.

Figura 70. Análisis en frecuencia de la señal de voz digitalizada con un micrófono integrado

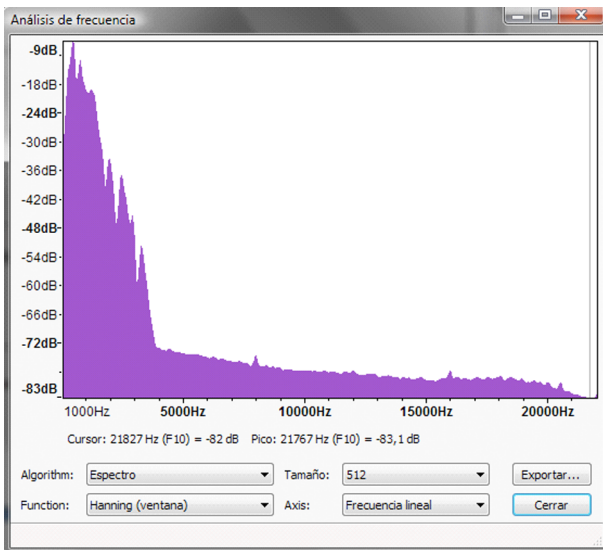
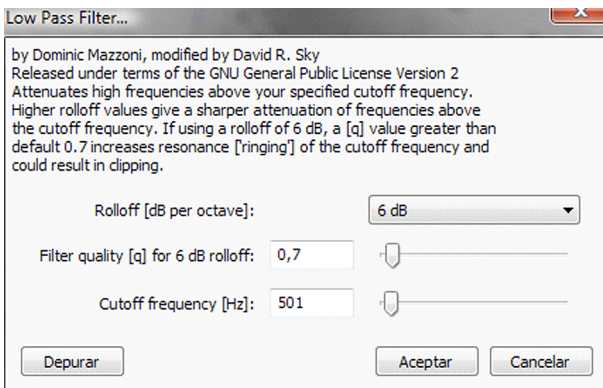


Figura 71. Selección de parámetros para el filtro paso bajo



3) Guardad la señal de audio anterior en un **fichero WAV**. Para ello, en el menú **Archivo** del programa Audacity seleccionad la opción **Exportar** y seleccionad como formato de salida WAV (Microsoft) *signed* 16 bit PCM.

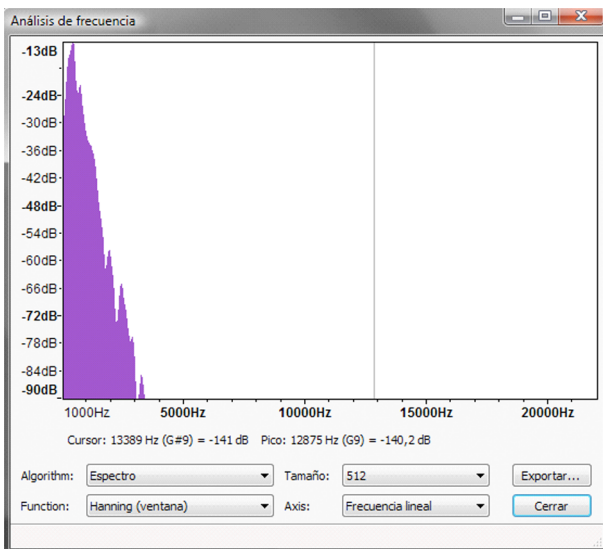
4) Realizad el filtrado de la señal de audio anterior mediante un filtro paso bajo con una frecuencia de corte de 500 Hz. Para ello, en el programa Audacity seleccionad la opción **Efecto/Filter/Low Pass** y modificad los parámetros del filtro para que coincidan con los que se muestran en la figura 71. Aplicad el filtro y escuchad la señal resultante. Para comparar la señal filtrada con la original, podéis ir al menú **Editar** y utilizar las opciones **Deshacer filtrado** y **Rehacer filtrado**.

5) Podéis comparar el resultado del filtro generado con Audacity y el filtro generado con la aplicación **eqApplet**. Para ello, reactivad la ventana de la aplicación Java con los filtros de audio, y con la barra de desplazamiento seleccionad una frecuencia de corte del filtro paso bajo de unos 500 Hz. Cargad el

fichero de audio que ha guardado en el fichero anterior mediante la opción **Load .wav file** y comprobad que los dos filtros generados tienen el mismo efecto desde el punto de vista auditivo.

6) Finalmente, seleccionad la señal de audio filtrada y realizad el análisis de espectro de la misma, siguiendo los mismos pasos utilizados en el punto 2). El resultado del análisis espectral se muestra en la figura 72, donde puede comprobarse cómo el efecto del filtro ha sido el de reducir las componentes de alta frecuencia de la señal de voz.

Figura 72. Resultado del análisis en frecuencia de la señal de voz procesada mediante un filtro paso bajo

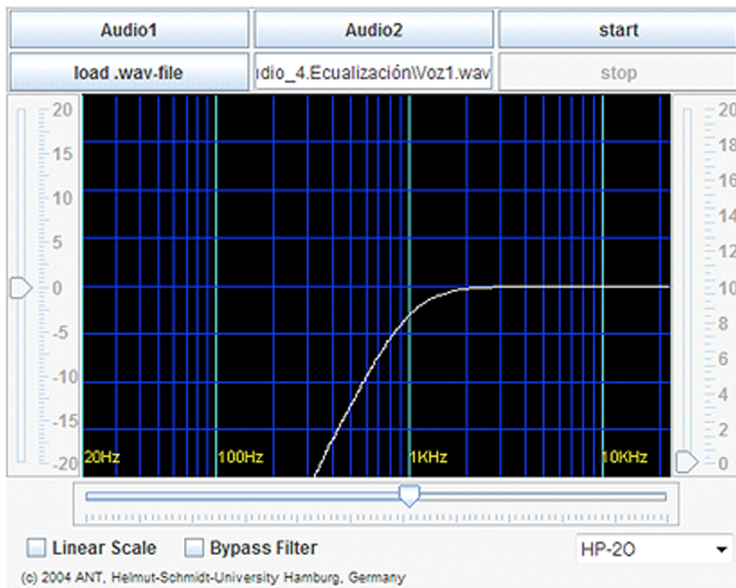


4.3.2. Ejemplos con filtrado paso alto

En este apartado vamos a repetir la misma experiencia anterior pero utilizando un filtro paso alto. Por tanto, en este caso, procesaremos la señal permitiendo el paso de las componentes de alta frecuencia y eliminando las bajas frecuencias. Desde el punto de vista auditivo, eliminaremos los sonidos graves y mantendremos los agudos.

Para empezar con la experiencia auditiva se sugiere cargar la aplicación Java de filtros de audio y seleccionar el filtro HP-2O (paso alto-2.º orden, *high pass-second order*). En la figura 73 se muestra la ventana de la aplicación, con la gráfica del filtro con una frecuencia de corte seleccionada de 1 kHz. Los botones, casillas y barras de desplazamiento tienen la misma interpretación que en el apartado anterior.

Figura 73. Ejemplo de la aplicación Java para el filtrado de audio utilizando un filtro paso alto de segundo orden



La frecuencia de corte está seleccionada en 1 kHz.

Así pues, vamos a repetir las pruebas que hemos realizado en el apartado anterior pero ahora con un filtro paso alto. En el siguiente ejercicio se resumen las pruebas que se recomienda realizar con este filtro.

1) Cargad las señales Audio1 y Audio2 que se proporcionan en el applet de Java y modificad la frecuencia de corte del filtro paso alto con la barra de desplazamiento. Activad y desactivad la casilla Bypass Filter para ver con mayor claridad el efecto del filtro. Observad cómo desaparecen los sonidos más graves de la señal mientras se mantienen los más agudos (platillos de las percusiones).

2) Indicad las frecuencias de corte en la que empieza a observar cambios significativos entre la señal original y la filtrada.

3) Observad el fenómeno de disminución de la potencia de la señal cuando se modifica la frecuencia del filtro.

4) Utilizad la misma señal de audio que habéis registrado en el punto 2) del anterior apartado (filtro paso bajo) y representad el análisis en frecuencia de la señal original y de la señal filtrada con el programa Audacity. Observad los efectos del filtrado sobre el espectro. Se sugiere tomar una frecuencia de corte de 1.000 Hz o superior, para que los efectos del filtrado resulten más claros y evidentes. Con este tipo de filtros, al desaparecer las componentes de baja frecuencia, que son las de mayor energía en la señal de audio, el efecto del filtrado sobre el espectro debería ser mucho más claro. Observad que, generalmente, obtendréis dos gráficas con escalas verticales distintas, ya que el espectro de la señal de voz filtrada tiene una energía mucho más baja. En la figura 74 se muestra el resultado que hemos obtenido con el filtrado paso alto de una voz humana con una frecuencia de corte de 1.000 Hz y un filtro de 6 dB por octava (segundo orden, equivalente a los filtros paso alto del applet de

Java). Para ver con mayor claridad el efecto del filtro sobre el espectro, puede ser conveniente realizar la presentación en escala logarítmica. Este resultado se muestra en la figura 75.

Figura 74. Representación en frecuencia de los espectros de una señal de voz original (izquierda) y filtrado paso alto con una frecuencia de corte de 1.000 Hz y un filtro de segundo orden

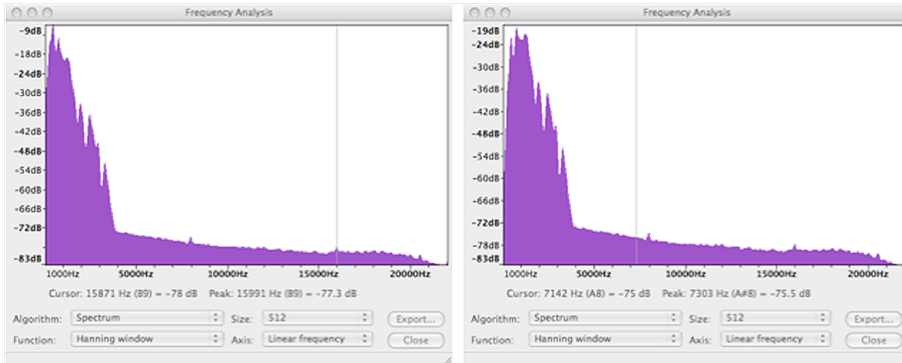
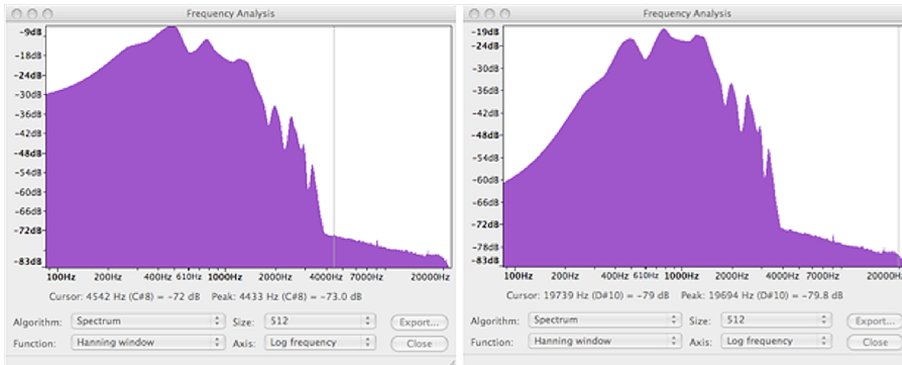


Figura 75. Comparativa entre los mismos espectros que en la figura anterior pero ahora representados utilizando el eje horizontal en escala logarítmica.



La representación logarítmica facilita la visualización del espectro en las regiones de baja frecuencia.

5) Filtrar con el applet de Java la misma señal del punto anterior, con la misma frecuencia de corte del filtro y comprobar que el efecto audible es el mismo.

4.3.3. Combinaciones de filtros en serie y paralelo

Los filtros paso bajo y paso alto estudiados en los apartados anteriores pueden considerarse como unos módulos básicos para construir otro tipo de filtros más complejos mediante la asociación en serie o en paralelo de diversos módulos de filtrado. En este apartado vamos a ver algunos ejemplos sencillos y en apartados posteriores extenderemos la asociación en serie y en paralelo para filtros más complejos.

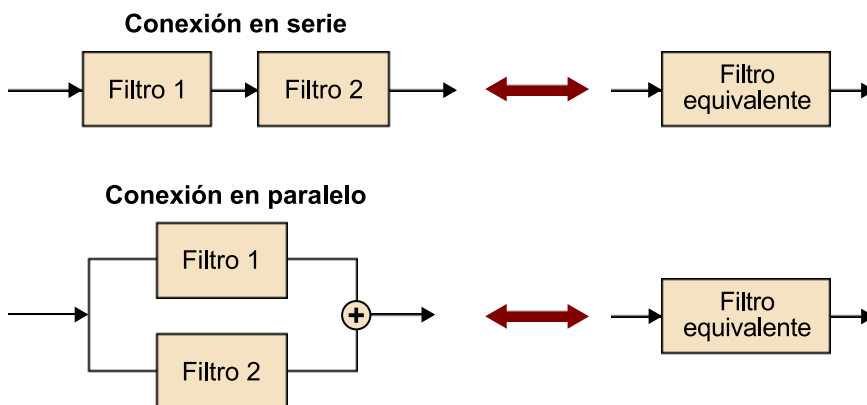
En la figura 76 se muestra el esquema básico de las conexiones en serie y en paralelo de dos filtros cualesquiera. Dado que el efecto final es que se obtiene una señal de salida para una entrada, podemos intentar buscar un **único filtro equivalente**, que realizará la misma función que la conexión en serie o en paralelo de los dos primeros.

En una conexión en serie los filtros se aplican uno detrás de otro de forma consecutiva. Así, la señal de entrada pasa primero a través del filtro 1 y su salida se utiliza como entrada del segundo filtro.

Para determinar el sistema equivalente deberemos tener en cuenta que el segundo sistema procesa las señales ya procesadas por el primero.

Así, si una componente en frecuencia ha sido atenuada por el primer sistema en un factor 0,9 y el segundo sistema le proporciona una atenuación de 0,5, entonces el efecto global sobre esta componente es la de una atenuación de $0,9 \times 0,5 = 0,45$. Por tanto, parece razonable que el efecto final sobre una componente es el resultado del producto de cada uno de los efectos de los filtros. Además, de acuerdo con este razonamiento es indiferente que una componente sea procesada primero por un filtro y luego por otro o al revés, pudiéndose intercambiar el orden de los filtros. Por otra parte, cuando los sistemas están conectados en paralelo cada uno realiza el procesamiento directamente sobre la señal de entrada y se suman las salidas. El sistema equivalente en este caso será la suma de los dos filtros.

Figura 76. Asociación en serie y en paralelo de filtros y filtros equivalentes



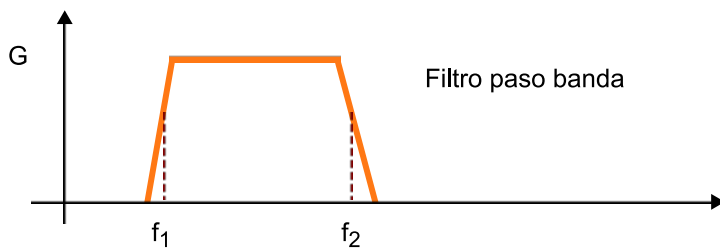
4.3.4. Filtros paso banda

Un filtro paso banda es un filtro que sólo deja pasar las frecuencias situadas dentro de una banda comprendida entre una frecuencia de corte inferior f_1 y una frecuencia de corte superior f_2 , tal como se representan en la figura 77. Las frecuencias bajas con una frecuencia menor que f_1 son eliminadas por el filtro. Análogamente, las frecuencias altas con una frecuencia superior a f_2 tampoco pasan a través del filtro.

Un filtro paso banda puede conseguirse como la conexión en cascada entre un filtro paso bajo y un filtro paso alto, tal como se muestra en la figura 78. En este caso, la frecuencia de corte del filtro paso bajo es f_2 , eliminando las frecuencias superiores a este límite. La frecuencia de corte del filtro paso alto es f_1 , por lo que se eliminan todas las frecuencias inferiores a este valor.

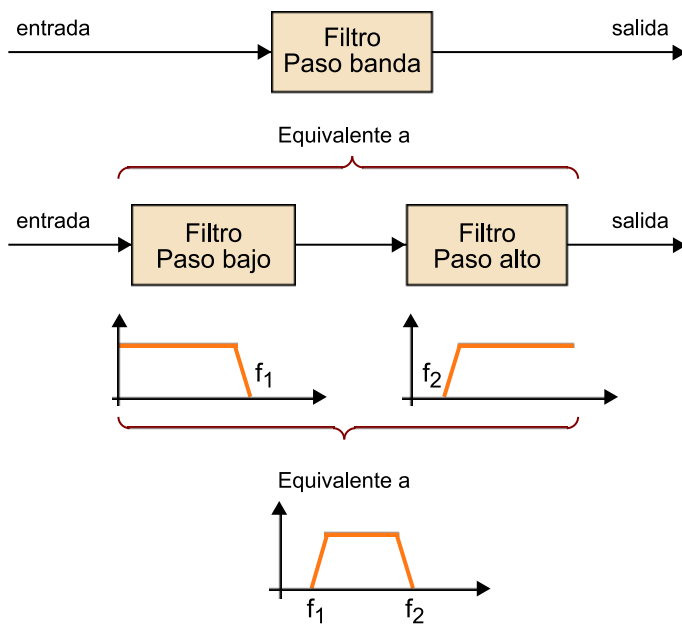
En resumen, el filtro ha eliminado las frecuencias superiores a f_2 y las inferiores a f_1 por lo que en la salida sólo tendremos las componentes que estén entre estas dos frecuencias. Los filtros paso banda tienen varias aplicaciones, sobretodo en el contexto de comunicaciones. En audio, pueden utilizarse para separar la voz humana de ruidos o instrumentos musicales. En efecto, las componentes frecuenciales de la voz humana están situadas entre unos 300 Hz y unos 3.400 Hz. De esta forma, si ponemos un filtro paso banda con estas frecuencias de corte podremos eliminar los ruidos que tengan frecuencias fuera de estos límites.

Figura 77. Ejemplo de respuesta de un filtro paso banda, caracterizado por dos frecuencias de corte



Sólo las componentes que estén comprendidas entre estas dos frecuencias podrán pasar a través del filtro.

Figura 78. Ejemplo de un filtro paso banda obtenido mediante la conexión en serie de un filtro paso bajo y un filtro paso banda.



Vamos a ver el efecto de un filtro paso banda sobre un tema musical con voz cantada. Si lo desea puede capturar con Audacity un fragmento cualquiera de su música favorita. Nosotros realizaremos el ejercicio con un fragmento del tema *El marido de la peluquera* de Pedro Guerra. Podéis encontrar este archivo en formato WAV con el material adicional asociado a esta actividad.

Para empezar abriremos el archivo **PedroGuerra_1.wav** con Audacity y lo escucharemos en su versión original. En la figura 79 se muestra la forma de onda del fichero que se representa en la pantalla de la aplicación. Después procederemos al filtrado. Conseguir un buen efecto de filtrado paso banda para aislar la voz humana es un proceso con bastante paciencia y que además sólo funcionará de forma aproximada, ya que, por una parte, los filtros no son ideales, por lo que no son capaces de eliminar completamente las frecuencias indeseadas.

Además, y no menos importante, gran parte de los instrumentos musicales que se utilizan en cualquier grabación también tienen componentes frecuenciales en la región espectral de la voz humana. Por ello, será totalmente imposible eliminar estas componentes musicales de la señal.

Como primera aproximación, se recomienda utilizar un filtro paso bajo con una frecuencia de corte de 700 Hz y 12 dB por octava (esto significa que la transición entre la banda de paso y la banda atenuada es más rápida que la que hemos venido utilizando hasta ahora). Para el filtro paso alto hemos tomado una frecuencia de corte de 300 Hz y también 12 dB por octava. Los parámetros que se han seleccionado para cada uno de los filtros se muestran en la figura 80. No obstante, estos valores son sólo una primera aproximación. Se recomienda probar otras frecuencias de corte para ver su influencia desde el punto de vista auditivo. Recordad que los filtros podéis encontrarlos en la opción **Efectos/Filtrado** y que podéis aplicar los filtros en el orden que queráis (son intercambiables). Recordad también que podéis cancelar los filtros aplicados a la señal mediante el menú **Edición/Deshacer**.

Figura 79. Representación de la forma de onda de la señal PedroGuerra1.wav

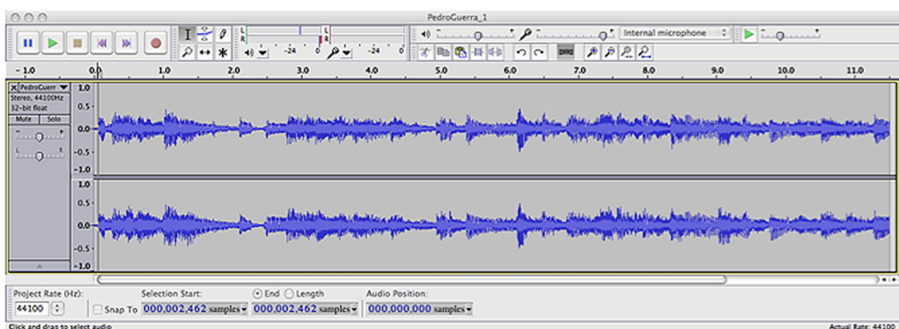
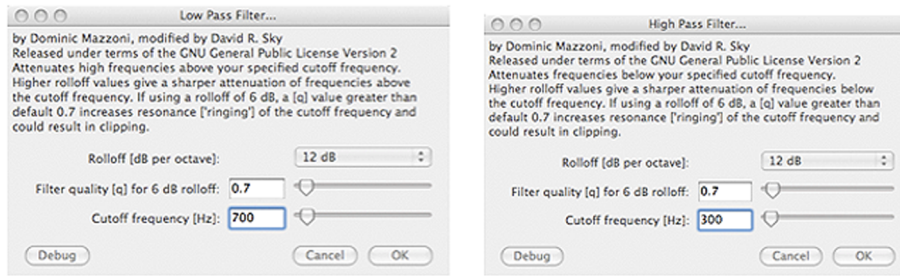
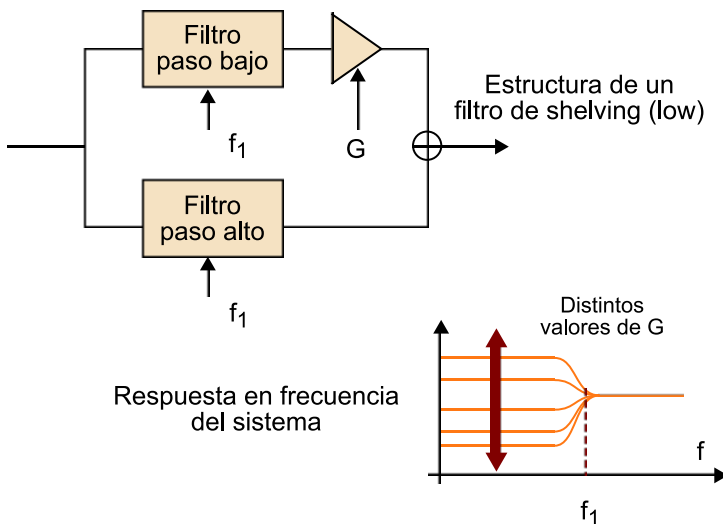


Figura 80. Ejemplo de los parámetros seleccionados para los filtros paso alto y paso bajo



4.3.5. Filtros de *shelving*

Los filtros de *shelving* (estantería) son unos filtros que pueden conseguirse mediante la conexión en paralelo de un filtro paso bajo con un filtro paso alto, ambos con la misma frecuencia de corte, que puede ser controlada por el usuario. Existen dos configuraciones típicas. En la figura 81 se muestra una configuración denominada *low shelving filter* (**filtro de *shelving* de baja frecuencia**), en la que se aplica una ganancia variable en la salida del filtro paso bajo mientras que la salida del filtro paso alto se mantiene constante. Esta configuración puede producir distintas respuestas en función del valor de la ganancia tal y como se muestra en la figura 81. Es interesante notar que el nombre de *shelving* (estantería) procede de los distintos niveles que pueden obtenerse en la banda de frecuencias bajas al variar el valor de la ganancia y que recuerda una estantería.

Figura 81. Representación de la estructura de un filtro de *shelving* como una configuración en paralelo de un filtro paso bajo y un filtro paso alto

La configuración alternativa es el filtro de *shelving* de alta frecuencia (*high shelving filter*) que se obtiene manteniendo fija la ganancia de la rama de baja frecuencia y modificando la de alta frecuencia.

Aplicación de los filtros de *shelving*

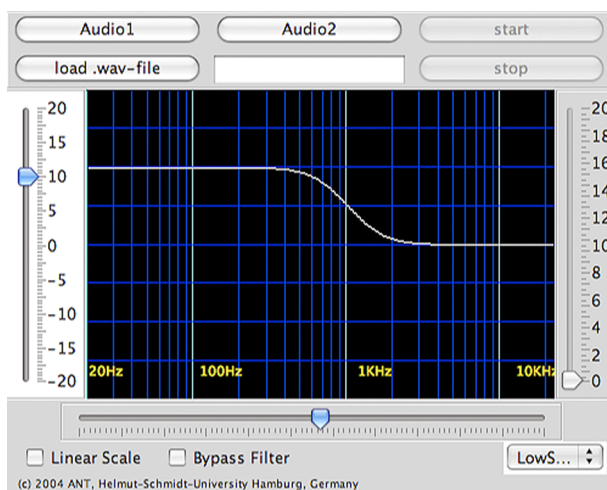
Los filtros de *shelving* se utilizan en muchas aplicaciones como por ejemplo para el control de tono mediante técnicas digitales. El filtro de *low shelving* se usa como control de tono para bajos (*bass*) mientras el de *high shelving* se usa para control de tonos agudo (*treble*). También se utilizan para filtrar el ruido de un micrófono en la captura de la voz humana.

En efecto, el filtro de *low shelving* es útil para eliminar los *pops* (explosiones que producen al hablar cerca de un micrófono) o para eliminar el *rumble* (ruido aleatorio producido por alguna fuente de sonido eléctrica) o los silbidos (*hiss*).

En este ejercicio vamos a realizar algunas pruebas con los filtros de *shelving* utilizando el applet de Java que ya hemos usado para los filtros paso bajo y paso alto. En este caso, tenemos que activar la aplicación `eqApplet.html` y seleccionar en la casilla de selección los filtros *lowshelvingfilter* y *highshelvingfilter*. El funcionamiento del applet es parecido al que ya hemos explicado para los filtros paso bajo y paso alto. En este caso se dispone de dos controles de desplazamiento: uno de ellos está en el eje horizontal de la gráfica y se utiliza para ajustar la frecuencia de corte de los filtros, el otro está en el eje vertical y se utiliza para controlar el valor de ganancia del filtro paso bajo (en el caso del *lowshelvingfilter*) o del filtro paso alto para el (*highshelvingfilter*). La figura 82 muestra el aspecto de la aplicación `eqFilter` cuando está seleccionado el *lowshelvingfilter* con un valor mayor que la unidad para la ganancia del filtro paso bajo.

Se propone que de forma más o menos autónoma se realicen varias pruebas de filtrado de señales con este tipo de filtros utilizando las muestras de audio que hemos estado usando en esta actividad. Los filtros de *shelving* suelen ser considerados como un primer tipo de filtros ecualizadores, muy simples y de los cuáles podemos controlar sólo dos parámetros que son la frecuencia de corte y la ganancia de uno de los filtros. En la siguiente actividad veremos otros tipos de ecualizadores más complejos.

Figura 82. Filtro de *shelving* paso bajo realizado con la aplicación `eqFilter` donde puede comprobarse que disponemos de control sobre la frecuencia de corte y la ganancia del bloque de paso bajo



4.4. Problemas/preguntas

1) En el ejercicio 1 ("Ejemplos con filtrado paso bajo") hemos capturado una señal de voz a la que hemos aplicado diversos tipos de filtros paso bajo y hemos visto los efectos del filtro en la representación frecuencial. No obstante, hemos visto que la señal de voz tiene muy poco contenido en alta frecuencia y por lo tanto no es necesario muestrearla a 44.100 Hz. En este primer ejercicio

se propone que se repitan las pruebas realizadas en este ejercicio utilizando una frecuencia de muestreo de 8.000 Hz (deberá modificar la frecuencia de muestreo del proyecto en la casilla inferior izquierda del programa Audacity). Representad y comparad los espectros que se obtienen con la señal filtrada y sin filtrar.

2) Generad una señal compuesta por 3 tonos con las frecuencias 400 Hz, 1.000 Hz, 1.700 Hz utilizando el programa Audacity. Recordad el procedimiento descrito en el apartado 1 para generar diferentes pistas y combinarlas en una única señal. Determinad las amplitudes de las 3 componentes para que sean aproximadamente iguales pero que no se produzca saturación de la señal. Realizad las pruebas siguientes con la señal obtenida:

- a) Escuchad la señal original con la superposición de los tres tonos.
- b) Representad el espectro de la señal original y observad la aparición de los tres tonos.
- c) Realizad un filtrado paso bajo de la señal con una frecuencia de corte de 900 Hz y una pendiente de 6 dB.
- d) Realizad el mismo filtrado anterior con otras pendientes (3 dB, 12 dB, etc.). Observad los efectos audibles al modificar la pendiente del filtro.
- e) Determinad el espectro de la señal obtenida en el apartado c y comprobad que la reducción en la potencia de las componentes en 1.000 Hz y 1.700 Hz.
- f) Realizad un filtro paso de la señal con una frecuencia de corte de 1.500 Hz y una pendiente de 6 dB.
- g) Representad el espectro de la señal obtenida en el apartado f.
- h) Generad una pista de audio con el contenido del filtrado del apartado c y otra con el contenido del apartado f. Combinad estas dos pistas de audio en una nueva pista y escuchad la señal resultante. ¿Qué tonos está oyendo?
- i) Representad el espectro obtenido en el apartado h y comprobad que se ha eliminado el tono central.

3) Los filtros banda eliminada son los complementarios de los filtros paso banda. También están definidos por dos frecuencias de corte f_1 y f_2 pero en este caso todas las componentes existentes entre estas dos frecuencias se eliminan. Así como los filtros paso banda se obtienen como la conexión en serie de un filtro paso alto y un filtro paso bajo, los filtros de banda eliminada pueden obtenerse mediante la conexión en paralelo de estos dos tipos de filtros. En el apartado h) del problema anterior hemos obtenido un filtro banda eliminada. En este problema pretendemos realizar el filtrado de una señal musical con voz humana, reduciendo la presencia del cantante mediante el uso de un filtro de banda eliminada. Para ello se propone que elijáis un fragmento musical con voz cantada (o utilizéis el archivo **PedroGuerra1.wav** que se proporciona con la práctica). Realizad las siguientes operaciones para crear un filtro banda eliminada:

- a) Cargad a Audacity el fichero que desea procesar.
- b) Generad dos pistas nuevas y copiad el archivo original a cada una de ellas.

- c) Aplicad un filtro paso bajo con una frecuencia de corte de unos 100 Hz -120 Hz y 12 dB por octava a la primera de las pistas nuevas que habéis generado.
- d) Aplicad un filtro paso alto a la otra pista, con una frecuencia de corte de unos 4.000 Hz y también con 12 dBs por octava.
- e) Poned en mute la pista original y escuchad la superposición de las otras 2.
- f) Determinad el espectro de la superposición de estas dos señales y comparadlo con el espectro de la señal original.
- g) Comparad la señal original con la obtenida en el apartado e).

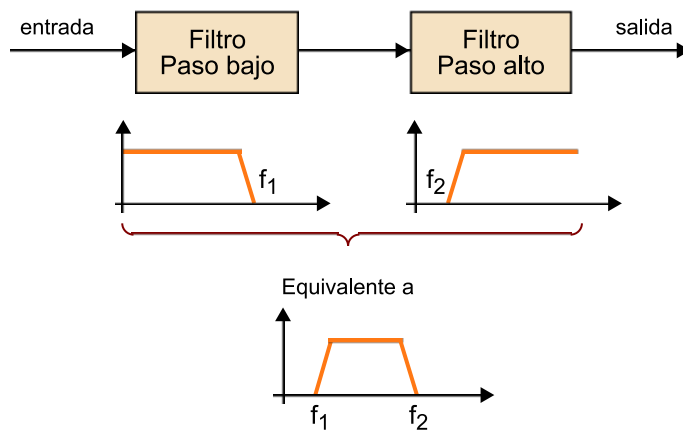
4.5. Evaluación

- 1) Determinad y representad de forma esquemática el filtro que obtenéis cuando se conectan en serie un filtro paso bajo con frecuencia de corte $f_1 = 3.000$ Hz con un filtro paso alto con frecuencia de corte $f_2 = 500$ Hz.

Solución

Se obtiene un filtro paso banda que permite el paso de las componentes frecuenciales que están situadas entre 500 Hz y 3.000 Hz. El resultado se muestra de forma esquemática en la figura 83.

Figura 83. Ejemplo de la concatenación de un filtro paso bajo y un filtro paso alto para la obtención de un filtro paso banda (problema 1)



- 2) Determinad y representad de forma esquemática el filtro que obtenéis cuando se conectan en serie un filtro paso bajo con frecuencia de corte $f_1 = 300$ Hz con un filtro paso alto con frecuencia de corte $f_2 = 1500$ Hz.

Solución

En principio, en este caso, si los filtros fueran ideales no se obtendría ninguna señal en la salida. En efecto, el primer filtro elimina todas las componentes que estén por encima de 300 Hz por lo que la entrada al segundo filtro no contiene ninguna componente superior a los 1.500 Hz, que son las que podrían pasar por el mismo. En la práctica, como los filtros no son ideales obtendremos un pequeño residuo de señal.

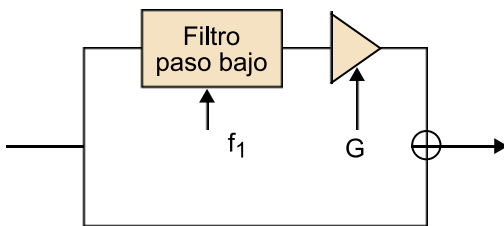
- 3) Determinad y representad de forma esquemática el filtro que obtenéis cuando se conectan en serie un filtro paso bajo con frecuencia de corte $f_1 = 800$ Hz con un filtro paso bajo con frecuencia de corte $f_2 = 500$ Hz.

Solución

En este caso obtenemos un filtro paso bajo con una frecuencia de corte de 500 Hz, la más pequeña de las dos. En efecto, las componentes con una frecuencia inferior a 500 Hz pueden pasar a través de los dos filtros, mientras que las componentes superiores son eliminadas por alguno o ambos filtros. Resumiendo, al conectar en serie dos filtros paso bajo, obtenemos un nuevo filtro paso bajo con una frecuencia de corte igual a la mínima de los dos filtros. Análogamente, podemos razonar que al conectar en serie dos filtros paso alto obtenemos un filtro paso alto con una frecuencia de corte igual al máximo de las dos.

4) Razonad qué tipo de filtro podemos obtener utilizando la estructura representada en la figura adjunta cuando los valores de ganancia pueden modificarse entre -1 y 9.

Figura 84



Solución

Se obtiene un filtro de *shelving* de baja frecuencia. Esta es una estructura alternativa para obtener filtros de *shelving*. En efecto, cuando la ganancia es -1, las componentes de baja frecuencia que provienen de la rama superior se cancelarán con las componentes que vienen de la rama inferior, obteniendo una señal nula en la región de baja frecuencia (situada por debajo de la frecuencia de corte). En cambio, cuando la ganancia sea de la rama superior sea de 9, la suma con la rama inferior dará una ganancia total de 10 en la zona de baja frecuencia. El filtro de *shelving* tendrá pues una curva de variación de ganancia de la parte de baja frecuencia que puede oscilar entre 0 y 10.

5. Actividad 5. Ecuación de señales de audio

5.1. Introducción

5.1.1. Contenido

Esta actividad debe considerarse como una segunda parte de la actividad de filtrado de señales en la que se profundiza en el concepto de ecualización y en la que se presentan los dos tipos de ecualizadores más utilizados: los **ecualizadores gráficos** y los **paramétricos**. Los dos tipos de ecualizadores son introducidos como modelos evolucionados de filtros, que permiten que el usuario controle la forma de la respuesta en frecuencia y que pueden ser interpretados como asociaciones en serie o en paralelo de filtros básicos. En ambos casos, el alumno se familiariza con el funcionamiento de estos sistemas utilizando herramientas software, la mayor parte de las cuales ya han sido utilizadas en otras actividades.

Para comprender los filtros paramétricos es necesario introducir los conceptos de frecuencia central, ganancia y ancho de banda como parámetros de control de un filtro paso banda. La dependencia del ancho de banda con la frecuencia central se explica mediante la definición del factor de calidad y se experimenta mediante el filtrado de señales de muestra. También se proporcionan diversos ejemplos de aplicación en los que se utilizan los ecualizadores para mejorar la calidad de la señal de audio y se introducen modelos comerciales de ecualizadores paramétricos.

Los ecualizadores gráficos son más intuitivos, ya que generalmente se incorporan en muchos equipos de audio doméstico o aplicaciones informáticas de reproducción musical. En este caso se presentan los conceptos básicos de descomposición del ecualizador como un banco de filtros y se determinan las frecuencias centrales de cada uno de los filtros en función del número de bandas del ecualizador.

5.1.2. Metodología

Se utiliza la misma metodología que se ha estado utilizando en actividades anteriores y que pretende que el alumno pueda ir comprobando los conceptos expuestos mediante la audición de los efectos del filtrado de las señales. En esta actividad se sigue utilizando el mismo applet de Java que ya se utilizó en la actividad anterior pero ahora con la opción de **PeakFilter**, que simula un ecualizador paramétrico. También se utiliza un nuevo applet de Java que simula un ecualizador gráfico de octavas y se sigue utilizando el programa

Audacity para la captura, edición y procesado de las señales de audio. A lo largo de la práctica se van proponiendo diversos ejercicios que deben realizarse utilizando los ficheros de audio que se proporcionan.

Cabe comentar que la mayor parte de estos ejercicios no tienen soluciones cerradas, ya que se trata de ajustar los ecualizadores para que la percepción de la señal de audio mejore, y esto, en definitiva, es un aspecto subjetivo. Estos ejercicios tan solo pretenden que el alumno se acostumbre al manejo de diferentes ecualizadores sin pretender convertirse en un experto en acondicionamiento de sonido.

Al final de la actividad se proponen algunos ejercicios adicionales que se intenta que el alumno resuelva con un mayor grado de autonomía y que interprete correctamente los resultados obtenidos.

Finalmente, se proponen varios problemas parecidos a los que han ido apareciendo como ejemplos a lo largo del texto y se incluyen las soluciones.

5.1.3. Recursos

Los principales recursos utilizados en esta actividad son:

- Software de captura y procesado de audio Audacity (licencia libre).
- Software de demostración de filtros y ecualizadores basado en applets de Java.
- Análisis de forma de onda y análisis de frecuencia de las señales mediante el programa Audacity.
- Captura de señales de voz para su ecualización.
- Archivos de audio predigitalizados con diferentes características.
- Explicaciones de los fundamentos de la ecualización y sus diferencias con el filtrado.
- Explicaciones de las aplicaciones principales de los ecualizadores.
- Propuesta de actividades de búsqueda de información en webs y recursos de Internet para profundizar en los conocimientos adquiridos en esta actividad (aplicaciones software de ecualización, fabricantes de ecualizadores, etc.).
- Actividades propuestas para verificar la adquisición de los conocimientos de esta actividad.

5.2. Objetivos

Los principales objetivos que se persiguen en esta actividad son:

- 1) Introducir los conceptos clave utilizados para la ecualización de señales de audio.
- 2) Definir el concepto de filtro paramétrico y familiarizar al estudiante con el control de los parámetros de ganancia, frecuencia central y ancho de banda.
- 3) Relacionar, mediante experimentos dirigidos, los efectos de un ecualizador paramétrico sobre la calidad y la percepción sonora de una señal de audio.
- 4) Conocer las aplicaciones de los ecualizadores en el contexto de la captura, reproducción y tratamiento de la señal de audio.
- 5) Introducir los diagramas de bloques de un ecualizador paramétrico en función de filtros paso banda.
- 6) Conocer productos y equipos de ecualización comerciales.
- 7) Introducir el concepto de los filtros de *notch*, como una solución para atenuar los efectos del acoplamiento de señales de audio.
- 8) Definir el factor de calidad de un filtro y comprender el concepto de filtros con factor de calidad constante.
- 9) Introducir el diagrama de bloques de un ecualizador gráfico como un banco de filtros paso banda.
- 10) Proporcionar ecuaciones necesarias para calcular las frecuencias centrales de un ecualizador gráfico para cualquier número de bandas.

Estos objetivos están relacionados con las siguientes **competencias de la asignatura**:

- 1) Capacidad de capturar audio e imágenes de forma eficiente y eficaz.
- 2) Capacidad de digitalizar eficiente y eficazmente contenidos de audio e imagen.
- 3) Capacidad de operar de forma digital con señales digitales de audio e imagen.
- 4) Capacidad de interpretar y modificar información auditiva de forma frecuencial.

Y con las siguientes **competencias generales del grado**:

- 1) Ser capaz de analizar un problema en el nivel de abstracción adecuado a cada situación y aplicar habilidades y conocimientos adquiridos para abordarlo y resolverlo.
- 2) Capturar, almacenar y modificar información de audio, imagen y vídeo digitales aplicando principios y métodos de realización y composición del lenguaje audiovisual.
- 3) Integrar y gestionar contenidos digitales en aplicaciones multimodales de acuerdo con criterios estéticos, técnicos y funcionales.

4) Atender adecuadamente consultas sobre proyectos, tecnologías y mercado de productos multimedia evaluando de forma precisa el entorno de aplicación, los recursos y las alternativas tecnológicas disponibles.

5.3. Guía

En la actividad anterior hemos estudiado diversos tipos de filtros. Los filtros constituyen uno de los métodos más importantes para **procesar y acondicionar las señales de audio**. Hemos visto que existen dos tipos de filtros básicos a partir de los cuales podemos, mediante las asociaciones en serie y en paralelo, conseguir otros tipos de filtros más complejos.

Estos filtros básicos son el paso bajo y el paso alto. Ambos están caracterizados por un único parámetro: la **frecuencia de corte**. El filtro paso bajo permite que todas las componentes de la señal de audio situadas por debajo de esta frecuencia de corte pasen a través del filtro, mientras que las componentes de frecuencia superior son eliminadas. El filtro paso alto realiza el efecto inverso, dejando pasar las componentes de alta frecuencia y eliminando las bajas.

Estos dos filtros básicos tienen la frecuencia de corte como único parámetro de control. No obstante, cuando asociamos estos filtros en serie o en paralelo²¹ podemos obtener más parámetros de control. Así, vimos que cuando conectamos en serie un filtro paso alto y uno paso bajo podemos obtener un **filtro paso banda**. El nuevo filtro paso banda tiene dos parámetros de control: la **frecuencia de corte inferior y la superior**. También vimos que los filtros de *shelving*, que se obtenían como la asociación en paralelo de un filtro paso bajo y uno paso alto tenían dos parámetros de control: la frecuencia de corte y la ganancia de una de las bandas (que llamábamos la estantería). Recuérdese que los controles de tono de bajos y agudos (*bass* y *treble*) que se proporcionan en la mayoría de equipos de audio pueden verse como dos filtros de *shelving*, en los que el usuario tiene el control sobre la ganancia de una de las bandas pudiendo aumentarla o atenuarla.

⁽²¹⁾Las asociaciones de filtros en serie y paralelo han sido explicadas en la actividad 4 de audio.

En esta actividad nos vamos a centrar en unos tipos de filtros algo más complejos, que proporcionan mayor flexibilidad para el acondicionamiento de las señales de audio a través de un número mayor de parámetros que pueden ser controlados por el usuario. Estos filtros reciben en general el nombre de **ecualizadores**. Tal y como ya hemos comentado, todos los ecualizadores pueden obtenerse mediante la asociación de filtros básicos.

Los dos tipos de ecualizadores más utilizados son los **paramétricos** y los **gráficos**, de los cuales analizaremos sus principios de funcionamiento, experimentaremos con ellos y veremos algunas de sus aplicaciones en los siguientes apartados. Los filtros de *shelving* y los filtros de *notch* también se denominan, en

algunos textos, ecualizadores. Los primeros ya fueron estudiados en la actividad anterior. En esta actividad veremos que los filtros de *notch* pueden considerarse como un caso particular de un filtro paramétrico.

5.3.1. Ecualizadores paramétricos

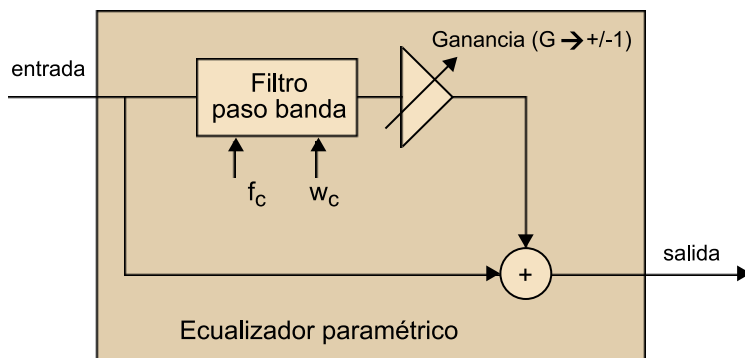
Un filtro paramétrico es un filtro que puede obtenerse a partir de un filtro paso banda utilizando una estructura como la indicada en la figura 85.

La particularidad del filtro es que permite que el usuario controle, de forma independiente, los siguientes parámetros del filtro:

- La ganancia (G)
- La frecuencia central del filtro paso banda (f_c)
- El ancho de banda del filtro paso banda (W)

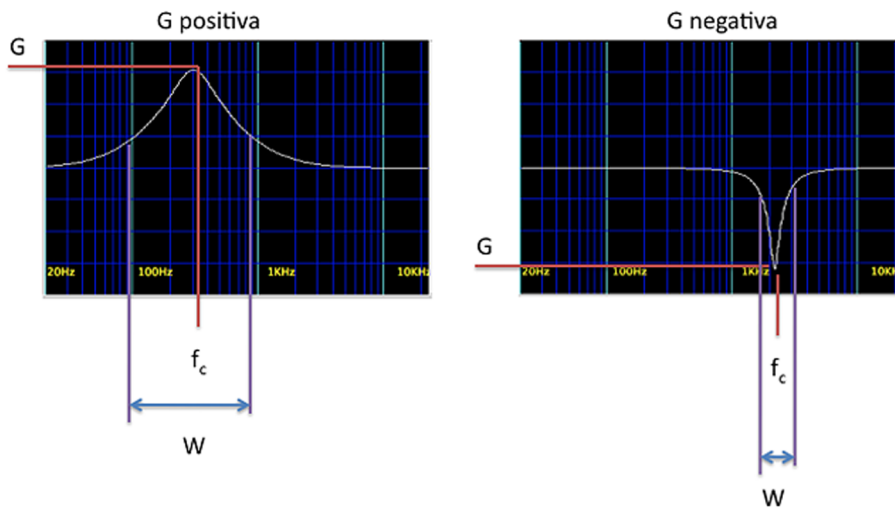
En la figura 86 se muestran algunos ejemplos de las curvas de respuesta de un filtro paramétrico para diferentes valores de los parámetros de diseño. En general, las curvas que obtenemos tienen una forma de campana, que puede ser hacia arriba o hacia abajo en función del valor del parámetro G . Cuando G es positivo, la campana es hacia arriba, mientras que cuando G toma valores negativos la campana será hacia abajo. Obsérvese que cuando el valor de G es igual a cero la rama del filtro paso banda en la configuración de la figura 85 no tiene ningún efecto por lo que el filtro tendrá una respuesta totalmente plana (no se modifica la señal). El control de la frecuencia central f_c modifica la posición de la campana en el eje horizontal, mientras que el control del ancho de banda W controla la anchura de la ventana. En la figura 86 se muestran dos posibles respuestas en frecuencia obtenidas para diferentes valores de los parámetros de control G , f_c y W .

Figura 85. Esquema de un filtro ecualizador paramétrico representado en función de un filtro paso banda



Los parámetros que el usuario puede controlar son la ganancia (G), la frecuencia central (f_c) y el ancho de banda (W).

Figura 86. Ejemplo de las respuestas en frecuencia de los filtros paramétricos para distintos valores de los parámetros



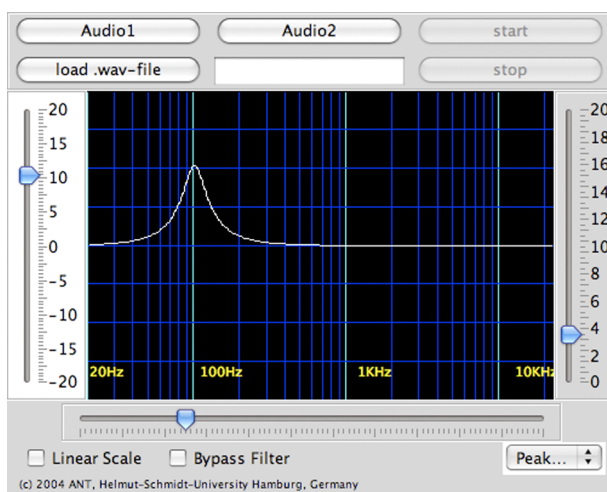
En el ejemplo de la izquierda tenemos un valor de G positiva, que da lugar a un filtro con forma de campana no invertida. La gráfica de la derecha se corresponde a una G negativa y da lugar a una forma de campana invertida.

En cualquier caso, posiblemente la mejor forma de comprender los filtros paramétricos es utilizándolos y escuchando los efectos que producen sobre diferentes señales de audio.

Para ver el efecto de estos filtros sobre ejemplos de audio vamos a usar el mismo applet²² de Java que hemos utilizado en la actividad anterior para ilustrar los efectos de los filtros paso bajo, paso alto y *shelving*. En este caso, ejecutamos de nuevo el applet `eqApplet.html` y seleccionamos el filtro `PeakFilter`. Veremos que aparece una pantalla como la mostrada en la figura 87 y que proporciona 3 controles independientes al usuario. Los controles están en las barras verticales izquierda (G, ganancia), derecha (W, ancho de banda) y en la barra de desplazamiento horizontal (f_c , frecuencia central).

⁽²²⁾Este applet ha sido desarrollado por el Grupo de Procesado de Señal de la Universidad Helmut-Schmidt de Hamburgo bajo la dirección del profesor Udo Zölzer.

Figura 87. Applet de Java con la implementación de un filtro paramétrico (PeakFilter)



Para familiarizarse con los filtros paramétricos se sugiere que se actúe sobre los diferentes controles mientras se está reproduciendo alguna de las señales de audio que se suministran por defecto (botones Audio 1 y Audio 2), con

ello se puede tener una primera aproximación de cómo afecta cada uno de los parámetros a la señal de audio. Mientras se realizan estas pruebas pueden observarse varios aspectos que ayudarán a comprender algunas características de los filtros paramétricos y que se resumen a continuación:

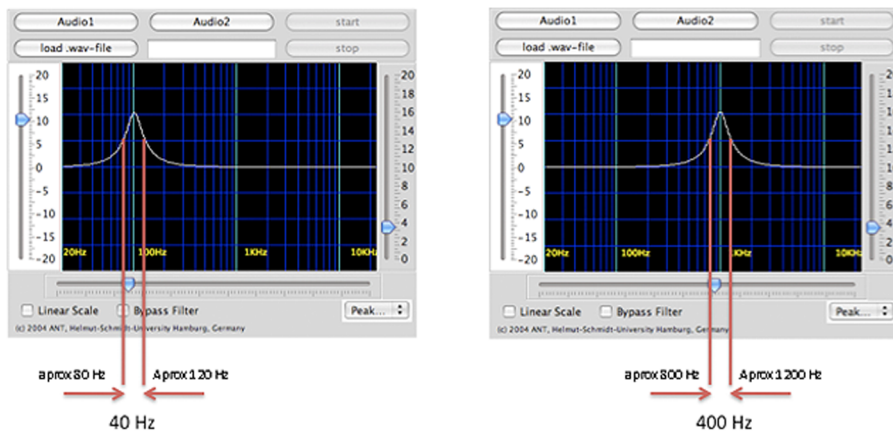
a) Cuando la ganancia del filtro está situada en un valor cero, la respuesta del filtro es completamente plana. En efecto, cuando ponemos esta ganancia a cero se desactiva la rama superior del filtro paso banda representado en la figura 85, por lo que el sistema deja pasar la señal de entrada sin realizar ninguna modificación de la misma. Obsérvese que aunque se modifiquen los valores de la frecuencia central (barra de desplazamiento inferior) o del ancho de banda (barra de desplazamiento vertical de la derecha) no obtenemos ningún efecto.

b) Cuando la ganancia del filtro es positiva, la respuesta del filtro tiene siempre una respuesta de campana. En cambio, cuando la ganancia es negativa obtenemos una campana invertida.

c) La amplitud de la campana se controla mediante la barra de desplazamiento vertical derecha. Los valores numéricos que aparecen junto a la barra vertical representan el **factor de calidad** del filtro, denominado Q . El factor de calidad es un parámetro directamente relacionado con el ancho de banda del filtro que definiremos posteriormente. Por el momento, basta con observar que al aumentar el factor de calidad disminuye el ancho de banda del filtro, obteniendo un filtro mucho más estrecho.

d) Cuando mantenemos fijos la ganancia (G) y el ancho de banda (factor de calidad) del filtro pero desplazamos la frecuencia central (f_c) observamos que, de acuerdo con lo esperado, la campana del filtro se desplaza en frecuencia. No obstante, deberíamos observar que el eje de frecuencias está representado en escala logarítmica, por lo que si la anchura de la campana se nos mantiene constante significa que el ancho de banda real del filtro aumenta con la frecuencia central. Veremos con más detalle este efecto cuando analicemos la definición del factor Q . Por ahora, observe que en la figura 88 se representan dos diagramas del filtro en los que la ganancia y el factor Q son constantes y en los que sólo hemos variado la frecuencia central del filtro. Tal y como se observa en esta última figura, el ancho de banda del filtro cuando la frecuencia está centrada en 100 Hz es aproximadamente de 40 Hz, mientras que cuando está centrado en 1.000 Hz es de 400 Hz (estos valores deben tomarse sólo como aproximaciones realizadas sobre las gráficas).

Figura 88. Comparativa entre los anchos de banda del filtro paramétrico manteniendo el factor Q constante y con usando diferentes frecuencias centrales.



e) Cuando ponemos la ganancia al valor mínimo (barra de desplazamiento vertical izquierda en el extremo inferior) y el factor de calidad al máximo (barra de desplazamiento vertical derecha en el extremo inferior), obtenemos lo que se conoce como un **filtro notch**, es decir, un filtro que elimina una frecuencia concreta. Podemos modificar el valor de esta frecuencia actuando sobre la frecuencia central del filtro.

1) En este ejercicio vamos a utilizar un ecualizador paramétrico para mejorar la calidad de una señal de voz. Para ello se propone:

a) Digitalizad un fragmento de vuestra propia voz utilizando el programa Audacity. Si es posible, utilizad directamente los micrófonos integrados en el ordenador. (En este ejercicio, no se trata de obtener un registro excelente sino de comprobar que si el registro es de mala calidad, con una correcta ecualización se puede mejorar algo el resultado.) El tiempo de registro debería ser aproximadamente de 1 minuto, con el objeto de disponer de suficiente tiempo para ajustar los parámetros del ecualizador y evaluar la calidad del sonido.

b) Exportad la señal registrada en el formato WAV (Microsoft) *signed* 16 bits.

c) Activad el applet de Java **eqApplet.html** y seleccionad el filtro PeakFilter. Seleccionad el botón **load .wav file** y seleccionad el archivo de audio que acabáis de exportar.

d) Escuchad repetidas veces el fragmento de audio y actuad con los diferentes parámetros del filtro hasta obtener una ecualización que os resulte satisfactoria.

Nota

Debido a la gran variabilidad de los equipos de captura y la subjetividad de la propia ecualización, no podemos proporcionar una solución a este ejercicio. Es una solución válida cualquier ecualizador paramétrico que introduzca una mejora subjetiva en la percepción de la señal de voz que ha registrado.

2) Ahora vamos a tomar un fragmento musical de guitarra clásica y vamos a intentar ecualizarlo. En este caso, el archivo original ha sido preprocesado para intentar simular un problema que tienen algunos instrumentos de cuerda y es que la caja de resonancia puede amplificar unas frecuencias más que otras, obteniendo un instrumento con un sonido que no está correctamente equilibrado. Para compensar este efecto suelen usarse filtros paramétricos.

El objetivo del filtro paramétrico es compensar el efecto de la caja de resonancia, atenuando aquellas frecuencias en las que se produce la resonancia del instrumento.

3) Cargad el archivo **Alhambra1.wav** que se proporciona junto con el material de esta actividad en el applet de Java e intentad, después de repetidas audiciones, determinar los parámetros de un filtro ecualizador que acondicione de forma correcta la señal de audio. Igual que con el ejercicio anterior, la solución final puede no resultar evidente para aquellos que no sean expertos musicales. No os preocupéis, el objetivo es que os familiaricéis con los efectos de los parámetros de este tipo de filtros y que veáis cómo afectan a la percepción del sonido. No pretendemos que con estos simples ejercicios los estudiantes se conviertan en prestigiosos ingenieros de sonido. Para dar una pista sobre cuál debería ser un posible resultado final se proporciona el archivo de audio **Alhambra2**, que realmente se corresponde con la señal original de *Recuerdos de la Alhambra* de Francisco Tárrega grabada por el músico Pepe Romero.

4) Ahora vamos a cambiar la guitarra clásica por la eléctrica y vamos a ver cómo pueden utilizarse los filtros paramétricos para generar efectos de audio. Cargad el fichero **SRV_Pride&Joy1.wav** en el applet de Java. El fichero corresponde con un fragmento introductorio del tema *Pride and Joy* interpretado por Stevie Ray Vaughan. Los efectos de filtrado sobre señales de audio se utilizan en rock y en música electrónica. Esencialmente consisten en modificar algún parámetro de un filtro en el tiempo, haciéndolo oscilar entre dos valores con la misma cadencia que la música. En este ejemplo, puede ser particularmente interesante seleccionar un filtro con una ganancia positiva y próxima al valor máximo y un factor de calidad de valor aproximadamente 4. Con estos parámetros fijos desplazad el cursor de la frecuencia central del filtro al ritmo de la música y ved cuál es el efecto de filtrado que se produce. Realizad otro tipo de efectos de modulación por filtrado actuando sobre los parámetros del filtro y modificando uno de ellos en el tiempo.

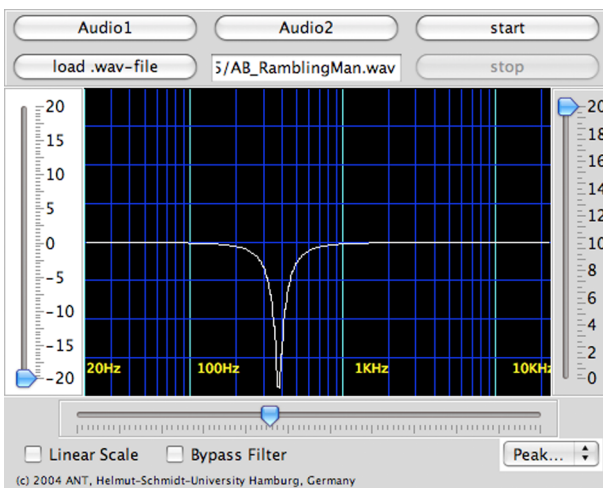
5) Finalmente, vamos a ver una de las principales aplicaciones de los filtros de *notch* que es la reducción del efecto Larssen. Este fenómeno fue descrito por primera vez por el ingeniero danés Søren Larssen que explicó las causas que lo producen.

Se trata de un efecto característico que se produce por un problema de realimentación entre la fuente de emisión y la fuente de reproducción del sonido.

Es habitual que se produzca en locales amplificados cuando el sonido emitido por los altavoces es capturado por el micrófono y este sonido es a su vez amplificado y vuelto a enviar a los altavoces. Informalmente, se conoce con el nombre de acoplamiento (¡el micrófono se acopla!).

El fenómeno también puede producirse con el *pickup* (sensor de vibración) de una guitarra eléctrica, que en este caso sustituye al micrófono. En cualquier caso, el resultado es la aparición de un tono molesto que suele resultar habitual en conciertos de rock, salas de fiestas o conferencias. La forma más habitual de reducir este efecto es utilizando un filtro de *notch*. Los filtros de *notch* son filtros que tienen una respuesta prácticamente plana en todas las frecuencias excepto en una frecuencia concreta, que es la que pretendemos eliminar. Podemos conseguir un filtro de *notch* con el applet de Java `eqApplet.html` poniendo la ganancia del filtro al mínimo (máximo valor negativo) y el ancho de banda del filtro también al mínimo. La frecuencia que pretendemos eliminar puede seleccionarse mediante el parámetro de selección de la frecuencia central. En la figura 89 se muestra la respuesta de un filtro de *notch* obtenida con el applet de Java.

Figura 89. Ejemplo de un filtro de notch



En este ejercicio vamos a cargar el fichero `AB_RamblingMan.wav` que contiene una grabación del tema musical *Rambling Man* interpretado por los Allman Brothers. Esta grabación ha sido manipulada y se ha insertado un tono molesto cuya amplitud debería ser reducida por el filtro de *notch*. Se sugiere que se escuche el fichero varias veces, activando y desactivando el filtro hasta encontrar la frecuencia que mejor reduce la amplitud del tono.

Es importante notar que en nuestro ejemplo el tono no desaparece completamente. Esto es debido a que los filtros no son nunca ideales y no eliminan de forma completa las señales que pretendemos. No obstante, en la práctica y en **directo**, este tipo de filtros suelen ser suficientes para eliminar de forma completa el efecto Larssen. Pensad que cuando utilizamos estos filtros en un directo, atenúamos el valor de la frecuencia capturada por el micrófono, por lo que el efecto de realimentación al que antes hacíamos referencia desaparece por completo.

5.3.2. El factor de calidad en los filtros paramétricos

Hemos comentado que los filtros paramétricos permiten controlar 3 parámetros, que son la **frecuencia central**, la **ganancia** y el **ancho de banda de la respuesta**.

El control del ancho de banda suele realizarse utilizando un parámetro que se denomina **factor de calidad** y que generalmente está representado con la letra **Q**. Así pues, en la práctica, el usuario especifica un factor de calidad para cada filtro. El factor de calidad relaciona la frecuencia central de la campana con su ancho de banda de acuerdo con la ecuación:

$$Q = \frac{f_c}{W}$$

donde f_c representa la frecuencia central del filtro y W el ancho de banda, tal como han sido representados en la figura 2.

Para medir el ancho de banda W , se utilizan las frecuencias en las que la campana del filtro atenúa la potencia de la señal a la mitad respecto al máximo.

Vamos a poner algunos ejemplos para clarificar la interpretación del factor de calidad. Supongamos que el usuario selecciona un factor de calidad de valor unidad. En este caso, cuando la frecuencia central del filtro esté en 100 Hz el ancho de banda del filtro también es de 100 Hz. A su vez, cuando la frecuencia central del filtro sea de 1.000 Hz, su ancho de banda también será de 1.000 Hz. Si el factor de calidad fuera 4, el ancho de banda del filtro siempre sería una cuarta parte de la frecuencia central. Así pues:

- a) Cuando el factor de calidad es alto el ancho de banda del filtro se reduce.
- b) Para un valor de Q fijo, el ancho de banda del filtro es proporcional a la frecuencia central del filtro. Si aumentamos la frecuencia central el ancho de banda del filtro aumenta de forma proporcional.

Esta relación de proporcionalidad entre la frecuencia central y el ancho de banda del filtro está especialmente adaptada a las características del sistema auditivo. En efecto, el sistema auditivo también tiene un comportamiento basado en este mismo principio ya que la diferencia perceptual de cambio de frecuencia entre dos señales de 200 Hz y 250 Hz es la misma que se produce

entre dos tonos de 2.000 Hz y 2.500 Hz. Nótese que en el primer caso las frecuencias están separadas por 50 Hz, mientras que en el segundo lo están por 500 Hz. Así, cuando aumentamos la frecuencia central de los tonos también tenemos que aumentar su separación (ancho de banda) de manera proporcional. Si utilizamos dos tonos de 2.000 Hz y 2.050 Hz (ahora están separados por 50 Hz) el oído prácticamente no apreciará diferencias significativas entre ellos. La proporcionalidad entre el ancho de banda del filtro y la frecuencia central cuando Q se mantiene constante justifica que, cuando se representa el filtro en una escala logarítmica de frecuencias, la anchura de la campana se mantenga constante.

Finalmente, cabe comentar que el nombre de factor de calidad no debe confundir al lector, ya que sólo es una medida del ancho de banda del filtro. En definitiva, una buena ecualización es aquella que corrige los problemas de un sistema de registro o de reproducción de audio. La corrección adecuada puede que tenga que realizarse con filtro con un Q elevado o con un Q bajo, dependiendo en gran parte de las causas que pueden deteriorar la señal.

5.3.3. Aplicaciones de los ecualizadores paramétricos

Los filtros paramétricos pueden utilizarse para acondicionar las señales de audio en diferentes circunstancias. El propio nombre de ecualizador ya nos indica cuáles son sus aplicaciones principales.

Ecualizar significa igualar y esencialmente consiste en modificar las características frecuenciales de un sonido para devolverle su equilibrio original.

Así, generalmente intentaremos ecualizar un sonido cuando sospechamos que los equipos de captura o reproducción han deteriorado de alguna forma sus características originales. En los siguientes párrafos pretendemos elaborar un breve resumen de los principales usos de los ecualizadores paramétricos.

1) Compensar respuesta en frecuencia de micrófonos. Generalmente, los micrófonos direccionales utilizados para la captura de voz aumentan su respuesta a bajas frecuencias a medida que se aproximan a la fuente de sonido. Este problema se conoce con el nombre de *efecto de proximidad* y resulta en un aumento desproporcionado de la cantidad de sonido grave a medida que nos acercamos al micrófono. Algunos vocalistas aprovechan el efecto de proximidad de forma deliberada para mejorar las características de su voz. Para reducir el efecto de proximidad puede utilizarse un filtro paso alto (que elimine las altas frecuencias) o un filtro paramétrico con la campana invertida y que elimi-

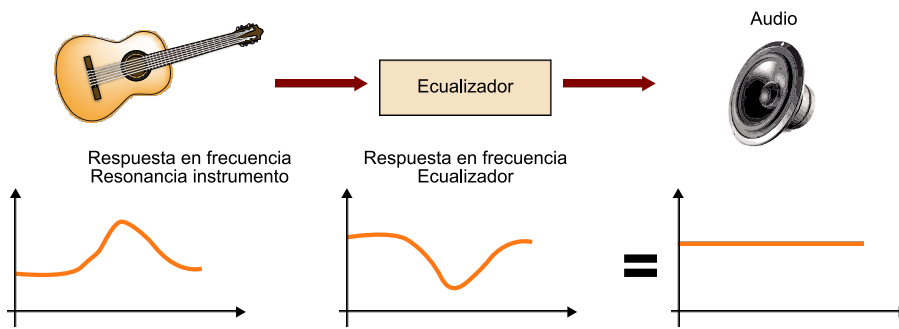
ne parte de estas bajas frecuencias. Aunque muchos micrófonos ya incorporan internamente un ecualizador, no suele ser suficiente y para obtener una máxima calidad suele tener que incorporarse un ecualizador de la mesa de mezclas.

2) Compensar las características direccionales de los micrófonos. La mayoría de los micrófonos presentan una respuesta direccional que depende de la frecuencia. En general, al aumentar la frecuencia los micrófonos se vuelven más directivos, capturando principalmente el sonido que procede de la dirección en la que están apuntando. Estas características no son importantes cuando el micrófono está situado cerca de la fuente de sonido (*close-miking*), pero pueden resultar muy significativas cuando nos alejamos de la fuente (registro de música clásica). En efecto, la pérdida de direccionalidad en las bajas frecuencias hace que se capture mejor el campo reverberante, es decir, el sonido que no procede directamente del instrumento sino de las reflexiones de la onda en las paredes del recinto. Como consecuencia, las altas frecuencias se ven perjudicadas y son capturadas con menor intensidad. La solución al problema suele ser un filtro paramétrico con una campana significativa en la zona de alta frecuencia.

3) Corrección de ruidos durante la captura de señal. Los ecualizadores paramétricos también se utilizan para reducir diversos ruidos desagradables que se producen durante la captura de señales. Entre estos ruidos debemos destacar las explosiones de aire que se producen al hablar cerca de un micrófono (*pops*), la captura de ruido procedente de aire o viento en un micrófono, el ruido de la púa de la guitarra al rasguear las cuerdas, el deslizamiento de los dedos de la mano izquierda en instrumentos de cuerda, el pedal del piano, etc. Evidentemente, todos estos ruidos tienen contribuciones en un margen de frecuencias muy amplio, por lo que no pueden eliminarse completamente mediante el filtrado, aunque sí que podemos reducirlos y que resulten menos molestos.

4) Compensar efectos de resonancia de instrumentos musicales. Una de las aplicaciones principales de los ecualizadores paramétricos es la de compensar las posibles deficiencias de resonancia en algunos instrumentos musicales. En efecto, en determinados instrumentos es posible que las cajas acústicas enfatizen o atenúen de forma excesiva un margen de frecuencias del instrumento respecto a otro. Estos problemas pueden compensarse aplicando un ecualizador paramétrico al sonido capturado. En la figura 90 se muestra la esencia de este principio. La respuesta en frecuencia del instrumento representa su naturaleza tímbrica, que en el ejemplo que nos ocupa tiene una cierta tendencia hacia el predominio de agudos. El ecualizador puede compensar esta respuesta del instrumento para dar una respuesta tímbrica más plana. En este sentido, el ecualizador iguala o equilibra las diferentes componentes frecuenciales del instrumento.

Figura 90. Ejemplo del uso de un ecualizador paramétrico para ajustar la resonancia de un instrumento musical o modificar el timbre



Evidentemente, el ecualizador también se puede utilizar para modificar o ajustar el timbre de un determinado instrumento a nuestros propios gustos. Modificar el timbre es modificar la contribución relativa de cada una de las componentes en frecuencia y esto puede conseguirse mediante el uso de un ecualizador. Los módulos ecualizadores paramétricos son de uso habitual en guitarra eléctrica y muchos guitarristas los utilizan para adaptar las características tímbricas del instrumento en función de las piezas musicales. En la figura 91 se muestra una fotografía de uno de estos módulos de ecualización para guitarra eléctrica.

Figura 91. Módulo de filtro paramétrico para guitarra. Observad los tres controles de ganancia, ancho de banda y frecuencia



El pedal se utiliza para activar y desactivar el filtro.

5) Reducción de efectos de realimentación de audio. Ya hemos comentado en el punto 4) del ejercicio 1 que una de las aplicaciones más importantes de los filtros de *notch* es eliminar los efectos de realimentación en sistemas de amplificación electrónica de audio (efecto Larsen). En este caso, el filtro se ajusta para atenuar la frecuencia a la que se produce la interferencia tonal, reduciendo la realimentación y eliminando la aparición del efecto.

6) Alteración del balance en la mezcla de instrumentos. Esta es una de las aplicaciones más importantes y justifica que las mesas de mezclas incorporen un considerable número de filtros para poder realizar la ecualización independiente de cada una de las pistas de audio. Para ilustrar el concepto básico utilizaremos la figura 8, en la que se supone que estamos intentando realizar la mezcla entre un instrumento solista y el resto de la orquesta. La gráfica de la izquierda muestra la superposición de los dos espectros y se observa que el espectro del solista estará enmascarado por la orquesta, esto significa que tendremos un sonido apagado por la propia orquesta, con un solista que parece situarse por detrás de la misma. Mediante el uso de filtros paramétricos podemos atenuar ligeramente una banda de la orquesta (con un filtro paramétrico

de campana invertida) y aumentar esta misma banda para el solista (con un filtro paramétrico de campana no invertida), tal como mostramos en la figura 8. Ahora, el solista ha ganado presencia respecto a la orquesta, por lo que el sonido estará más equilibrado. En el caso concreto de vocalistas se suelen utilizar filtros centrados en la frecuencia de 1.500 Hz, con unos factores de calidad entre 5 y 7.

7) Ecualizadores paramétricos multibanda. La importancia de los ecualizadores paramétricos en el acondicionamiento de las señales de audio hace que en muchos casos no tengamos suficiente con un único ecualizador paramétrico y tengamos que conectar varios de ellos en cascada. En la práctica existen filtros paramétricos multibanda que combinan diferentes ecualizadores en un único equipo. Así, un ecualizador paramétrico de 5 bandas no es más que una conexión en serie de 5 ecualizadores paramétricos. Si sólo queremos utilizar uno de ellos pondremos la ganancia de todos los demás a cero, obteniendo una ganancia plana. Cuando todos los ecualizadores están actuando, el resultado del conjunto es el producto de las respuestas en frecuencia de cada uno de ellos. Si cada uno de los ecualizadores se programa con una frecuencia central distinta y los anchos de banda no se superponen, los diferentes filtros pueden programarse de forma bastante independiente. En la figura 93 se muestra un esquema que interconecta 3 ecualizadores paramétricos en serie, y en la figura 94 se muestra un equipo profesional que dispone de 5 ecualizadores paramétricos y un filtro de *notch*. Muchos ecualizadores multibanda incorporan también filtros paso bajo y paso alto como el que se muestra en la figura 95.

Figura 92. Ejemplo de una ecualización para alterar el balance de una mezcla musical

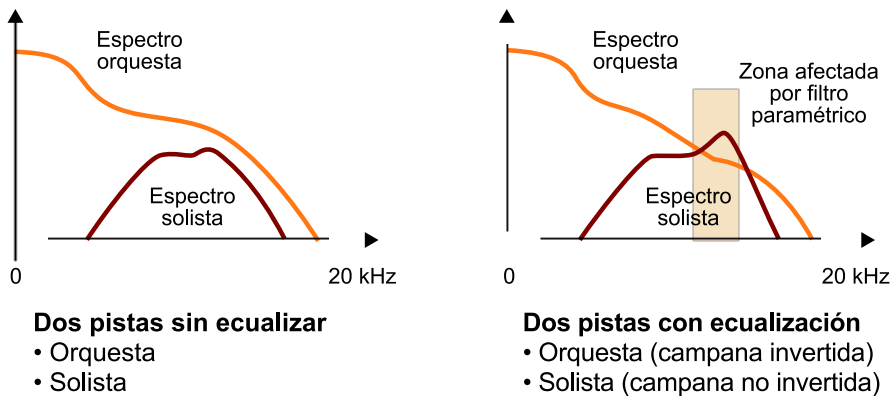


Figura 93. Ejemplo de un ecualizador paramétrico multibanda (3 bandas)

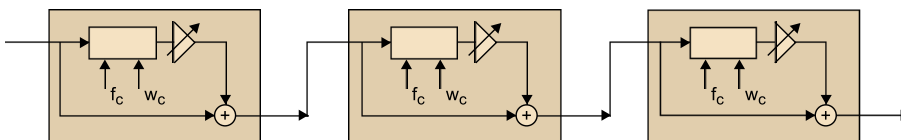


Figura 94. Ecuador paramétrico profesional de 5 bandas más un filtro de *notch*

Figura 95. Ecuador paramétrico multibanda que incorpora un filtro paso alto y uno paso bajo



5.3.4. Ecuadores gráficos

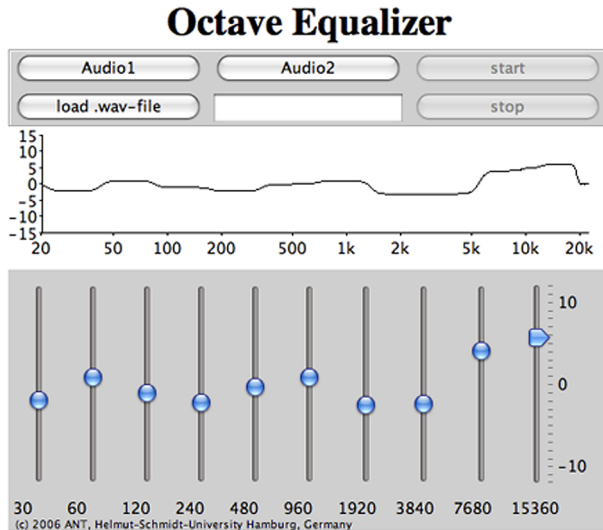
Los ecualizadores gráficos son una alternativa a los filtros paramétricos, proporcionando también una gran flexibilidad al usuario y resultando considerablemente más intuitivos de utilizar. De hecho, el ecualizador gráfico es un elemento habitual en muchas cadenas de audio y en muchos programas de reproducción musical para ordenador. Desde un punto de vista intuitivo, un ecualizador gráfico nos proporciona una serie de controles de desplazamiento que nos permiten configurar la curva de respuesta en frecuencia del filtro ecualizador. Posiblemente, la forma más simple de comprender su principio de funcionamiento es utilizando uno de ellos y comprobando el efecto que tiene sobre las señales de audio.

En este ejercicio vamos a utilizar un applet de Java para familiarizarnos con el uso de un ecualizador gráfico. En concreto, vamos a utilizar un applet desarrollado por el mismo centro²³ que ha desarrollado el `eqApplet.html`, que ya hemos utilizado para ilustrar los filtros y los ecualizadores paramétricos. Para iniciar la ejecución del applet debe descomprimirse el contenido

⁽²³⁾Recordemos que se trata del Departamento de Procesado de Señal de la Universidad Helmut-Schmidt de Hamburgo dirigido por el profesor Udo Zölzer.

del fichero **octaveEq.zip** en una carpeta del ordenador y ejecutar el archivo **octaveEqApplet.html**. Debemos dar permiso al navegador para ejecutar el código del applet y aparecerá una pantalla como la que se muestra en la figura 96

Figura 96. Ecuador gráfico de octavas proporcionado por el applet de Java octaveEqApplet.html



El manejo de un ecualizador gráfico es muy intuitivo. En nuestro caso disponemos de 10 controles deslizantes que cubren toda la banda de frecuencias de audio. El applet nos permite ver cómo se modifica la respuesta en frecuencia del ecualizador a medida que actuamos sobre estos controles. Se sugiere que se practique con distintas curvas de ecualización, observando los efectos que se producen sobre los dos temas musicales de ejemplo. Nótese que se trata de un filtro bastante efectivo, ya que cada uno de los controles puede aumentar la ganancia en 10 dBs (potencia 10 veces superior) o atenuarla en 10 dBs (potencia 10 veces inferior). Por tanto, cada una de las bandas admite actuar sobre la potencia en una relación 1 a 100. Evidentemente, cuando todos los controles están en la posición de 0 dBs tenemos una respuesta del filtro totalmente plana, equivalente a no realizar ningún filtrado.

El filtro anterior se denomina **filtro de octava** debido a que las frecuencias centrales de cada uno de los controles tienen una relación de octavas. En efecto, la primera banda está centrada en la frecuencia de 30 Hz, la segunda la encontramos una octava²⁴ por encima, es decir, a 60 Hz y así consecutivamente, doblando cada vez la frecuencia.

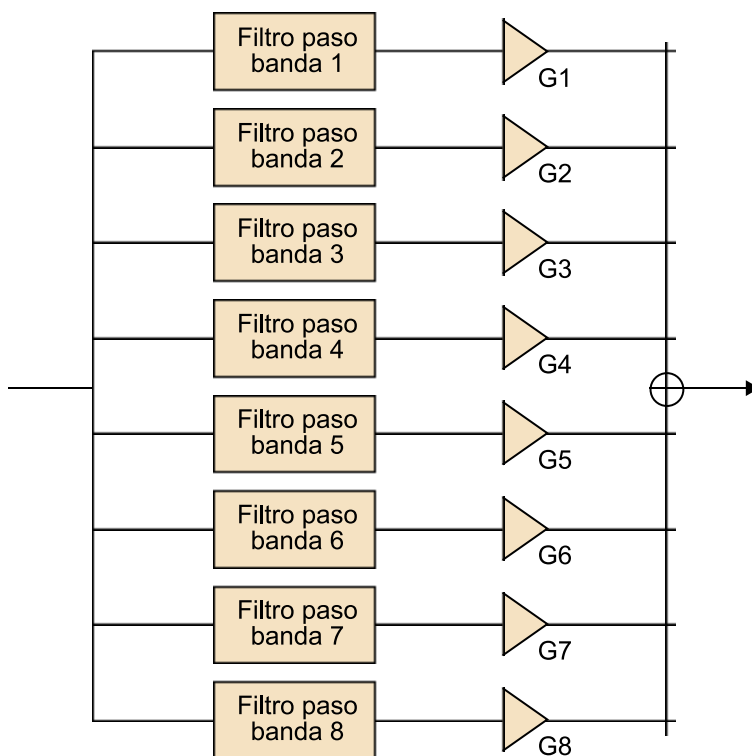
⁽²⁴⁾En terminología musical una octava equivale a doblar la frecuencia. El nombre se debe a que existen 8 notas musicales entre una nota determinada y la siguiente nota con frecuencia doble (do, re, mi, fa, sol, la, si, do).

Cada uno de los controles actúa sobre una frecuencia central y sobre su banda de frecuencias adyacentes. Es importante notar que estas bandas de frecuencia son más anchas a medida que aumenta la frecuencia central. Como en el caso de los filtros paramétricos, se trata de filtros en los que la relación entre la frecuencia central y el ancho de banda se mantiene constante.

5.3.5. Estructura de un ecualizador gráfico

En la figura 97 se muestra la estructura de un ecualizador gráfico que se ha descompuesto como una asociación en paralelo de varios filtros paso banda. La señal de entrada se descompone en un banco de filtros paso banda que son controlados de forma independiente mediante los controles de ganancia. El número de filtros que se utilizan son las bandas del ecualizador, que para nuestro ejemplo serán un total de 8 bandas. Comercialmente, pueden encontrarse ecualizadores gráficos con una gran variedad de bandas, desde 3 o 5 hasta 20 o más. El número total de bandas determina la flexibilidad del filtro para configurar la curva de respuesta en frecuencia.

Figura 97. Estructura de un ecualizador gráfico de 8 bandas



En la figura 98 se indica cómo las frecuencias centrales están distribuidas en la frecuencia de acuerdo con una regla logarítmica, es decir, la frecuencia central de la siguiente banda se obtiene multiplicando la frecuencia de la banda anterior por una constante. Esta constante depende del número de bandas, en concreto, normalmente, si tenemos que cubrir toda la banda desde los 20 Hz hasta los 20 kHz, toma el valor $C=10^{3/B}$, donde B representa el número de bandas⁽²⁵⁾. La frecuencia central del primer filtro puede calcularse también en función de este valor de C utilizando la fórmula:

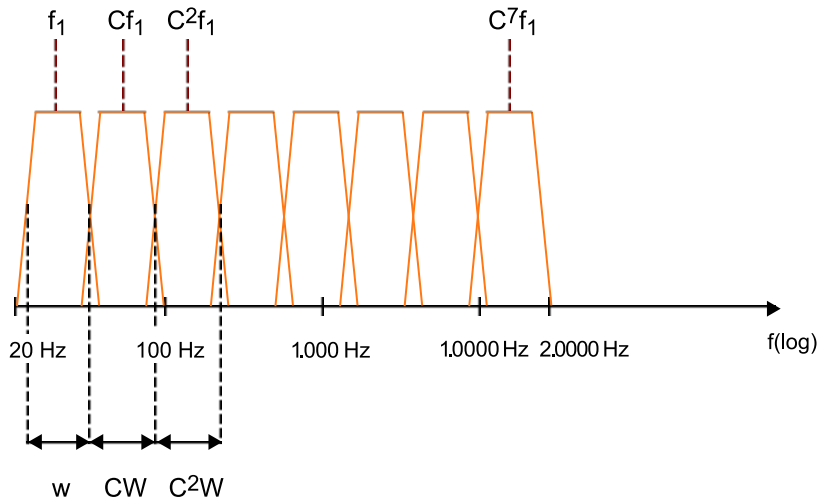
$$f_1 = 20 \cdot \sqrt{C}$$

⁽²⁵⁾ Observad que para un ecualizador de 10 bandas tendremos $C = 2$, es decir, la frecuencia central de la banda siguiente se obtiene duplicando la frecuencia central de la banda actual. Por lo tanto, un ecualizador de 10 bandas será siempre un ecualizador de octavas.

⁽²⁶⁾ En el caso concreto del ecualizador de octava del filtro del applet, si aplicamos directamente la fórmula anterior obtenemos una frecuencia de 28,28427 Hz, que ha sido aproximada por 30 Hz.

La mayoría de veces, si aplicamos directamente esta fórmula obtenemos frecuencias con muchos decimales, por lo que el resultado se suele aproximar a un número entero próximo²⁶.

Figura 98. Representación de los filtros paso banda que componen un ecualizador gráfico



En la misma figura 98 también podemos observar cómo el ancho de banda de cada uno de los filtros aumenta de acuerdo con la constante multiplicativa C . La relación entre la frecuencia central del filtro y su ancho de banda es por tanto una constante. Por ello, cuando los filtros se representan utilizando un eje de frecuencias logarítmico todos ellos tienen la misma anchura.

En este ejercicio se propone que se experimente utilizando el ecualizador gráfico de octavas proporcionado en el applet de Java que habéis utilizado en el ejercicio anterior utilizando los diferentes ejemplos de audio que han sido procesados mediante un ecualizador paramétrico en ejercicios anteriores. Para ello, se propone que realicéis las siguientes actividades:

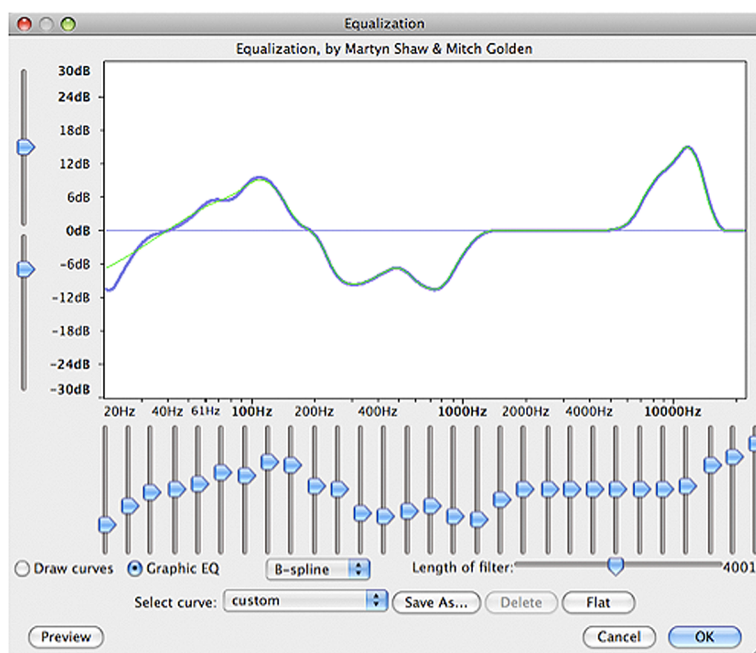
- 1) Escuchad repetidas veces la señal de voz que habéis grabado en el ejercicio 1 y ajustad los parámetros del ecualizador gráfico para obtener un sonido satisfactorio. Comparad las curvas de respuesta en frecuencia que habéis obtenido con las que obtuvisteis en el ejercicio 1.
- 2) Ecualizad la señal **Alhambra1.wav** hasta obtener el timbre deseado para el instrumento musical.
- 3) Intentad reducir la interferencia tonal que aparece en la grabación **AB_RamblingMan.wav** mediante este ecualizador gráfico.

El programa Audacity también dispone de un ecualizador gráfico excelente al que puede accederse desde el menú **Efectos/Filtros/Equalizador**. En este caso, tal como mostramos en la figura 15 se proporcionan un total de 30 bandas para ajustar la curva de respuesta, lo que da una gran flexibilidad al usuario. También se dispone de un modo que permite que el usuario dibuje directamente la

curva que desea (activar la casilla **Draw Curves**). Esta opción que permite que el técnico de sonido dibuje directamente la curva que desea obtener sobre un terminal gráfico es cada vez más utilizada en las modernas mesas de mezcla, basadas en gran parte en aplicaciones software genéricas que controlan sistemas de filtros programables. Se propone que en este ejercicio se experimente con varios de los archivos de audio de que dispone para ver la funcionalidad del ecualizador que se proporciona con Audacity.

El ecualizador de Audacity incorpora también diferentes **Presets** de ecualización que pueden obtenerse a partir del desplegable **Select Curve** y permite que el usuario pueda grabar sus propias curvas de ecualización para recuperarlas en otras sesiones.

Figura 99. Ecualizador gráfico de 29 bandas incluido en la aplicación Audacity



5.4. Problemas/preguntas

1) En los ejercicios 1 y 5 hemos capturado una señal de voz a la que hemos aplicado filtros de ecualización paramétricos y gráficos. No obstante, hemos visto que la señal de voz tiene muy poco contenido en alta frecuencia y por lo tanto no es necesario muestrearla a 44.100 Hz. En este primer ejercicio se propone que se repitan las pruebas realizadas en los anteriores ejercicios utilizando una frecuencia de muestreo de 8.000 Hz (deberéis modificar la frecuencia de muestreo del proyecto en la casilla inferior izquierda del programa Audacity). Representad y comparad los espectros que se obtienen con la señal ecualizada y sin ecualizar.

2) Generad una señal compuesta por 3 tonos con las frecuencias 400 Hz, 1.000 Hz, 1.700 Hz utilizando el programa Audacity (esta señal ya fue generada en el problema 2 del apartado 4, apartado Problemas/preguntas). Se pide que llevéis a cabo las pruebas siguientes con la señal obtenida:

- a) Escuchad la señal original con la superposición de los tres tonos.
- b) Mediante los ecualizadores paramétricos de la aplicación java **eqApplet.html** generad un filtro de *notch* que atenúe la señal de 1.700 Hz y escuchad el resultado.
- c) Mediante los ecualizadores paramétricos de la aplicación Java **eqApplet.html** generad un filtro con una campana de ganancia máxima centrada en la frecuencia de 1.000 Hz y escuchad sus efectos.
- d) Mediante el ecualizador gráfico de la aplicación **octaveEqApplet.html** generad un filtro que elimine las componentes de 400 Hz y 1.700 Hz y escuchad el resultado obtenido.
- e) Comparad el resultado anterior cuando realizáis la misma ecualización con el filtro ecualizador gráfico incorporado en Audacity.

3) Encontrad distintos fabricantes, modelos, precios y características de ecualizadores comerciales para las siguientes aplicaciones:

- a) Pedales de filtros paramétricos para guitarra.
- b) Racks de filtros paramétricos multibanda. Determinad sus características.
- c) Módulos de ecualizadores gráficos para equipos de alta fidelidad domésticos.
- d) Módulos de ecualizadores gráficos que puedan usarse para la ecualización de salas.
- e) Mesas de mezclas para disc-jockey. Identificad los ecualizadores que se proporcionan para cada canal y el tipo.
- f) Mesas de mezclas profesionales. Identificad los ecualizadores/filtros que incluyen en cada canal y su tipo.

4) Seleccionad una aplicación software de reproducción de señales de audio (winamp, itunes, ...) e identificad si permite que el usuario configure su propio ecualizador gráfico. ¿De cuántas bandas está compuesto? Comprobad si las frecuencias centrales coinciden con las fórmulas proporcionadas en la guía de esta actividad. Si el software de reproducción de audio contiene presets de ecualización, comprobad sus características y escuchad el efecto que tienen sobre distintos temas musicales.

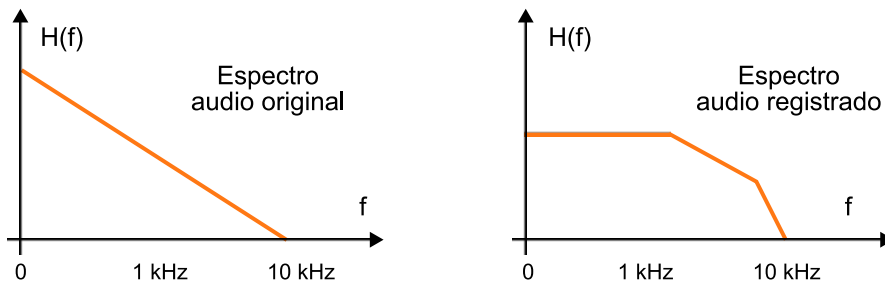
5) Si disponéis de un reproductor de audio portátil, comprobad si permite que el usuario configure un ecualizador gráfico. ¿De cuántas bandas está compuesto? ¿Cuáles son las frecuencias centrales? ¿Contiene presets? ¿Las curvas de los presets coinciden con las obtenidas en 4)?

6) Buscad información en Internet para determinar a qué hacen referencia las diferentes curvas de preset que se proporcionan en el ecualizador de Audacity (AES, Columbia 78, RCA Victor 1938, RCA Victor 1947, Decca FFRR LP, etc.).

5.5. Evaluación

1) Suponed que estamos registrando una fuente de audio cuyo espectro tiene una forma triangular como la que se muestra en la figura adjunta (izquierda). Una vez registrada la señal analizamos su espectro y vemos que tiene la forma representada en la gráfica de la derecha.

Figura 100

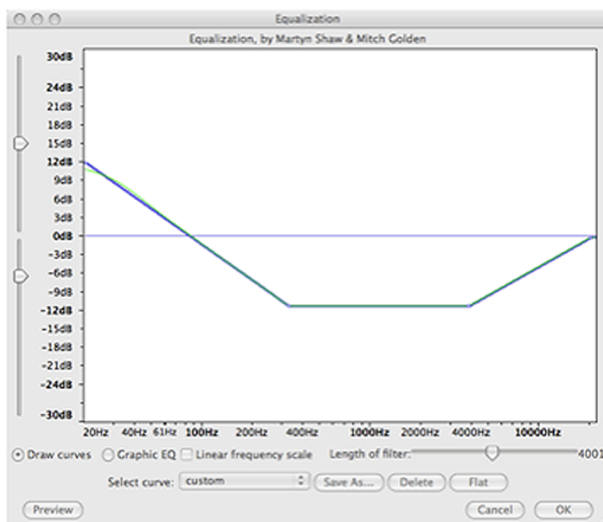


Determinad de forma aproximada cuál deberá ser la curva de ecualización de la señal.

Solución

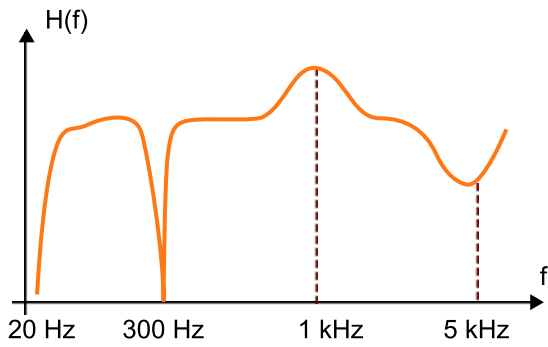
El filtro de ecualización deberá restaurar el espectro a su forma original, por lo tanto deberá tener una forma como la representada en la figura:

Figura 101



2) Considerad la siguiente curva de ecualización que sabemos que ha sido obtenida mediante un sistema de filtros multibanda. Indicad cuántos filtros han intervenido y qué características tienen.

Figura 102

**Solución**

El sistema está compuesto por 4 filtros activos:

- a) Un filtro paso bajo
- b) Un filtro de *notch* centrado en 300 Hz
- c) Un filtro paramétrico con ganancia positiva y centrado en 1 kHz
- d) Un filtro paramétrico con ganancia negativa y centrado en 5 kHz

3) Determinad las frecuencias centrales de un ecualizador gráfico de 5 bandas.

Solución

De acuerdo con las fórmulas proporcionadas $C = 10^{3/5} = 4$. Por lo tanto, la primera frecuencia tomará el valor $f_1 = 20 \times 2 = 40$ Hz. El resto de frecuencias serán: 160 Hz, 640 Hz, 2560 Hz, 10240 Hz.

4) Un filtro paramétrico tiene un $Q = 6$. Determinad el ancho de banda del filtro cuando la frecuencia central es de 150 Hz y cuando es de 3.000 Hz.

Solución

Utilizando la definición de $Q = f/W$ obtenemos que el ancho de banda vendrá determinado por el cociente entre la frecuencia central y el factor de calidad. Así, en este caso tendremos un ancho de banda $150 \text{ Hz}/6 = 25$ Hz y $3.000/6 = 500$ Hz en cada uno de los dos casos.

Bibliografía

Bibliografía básica recomendada

Howard, David M.; Angus, Jamie. *Acoustics and Psychoacoustics*. Third Edition. Focal Press.

Pohlman, Ken C. *Principios del Audio Digital*. McGraw-Hill. Interamericana de España.

Bibliografía complementaria

Chilitos Valenzuela, José. *Audio digital: Conceptos Básicos y Aplicaciones*. Backbeat books.

Watkinson, John. *An Introduction to Digital Audio*. Focal Press.

Zölzer, Udo (agosto, 2008). *Digital Audio Signal Processing* (2.^a ed.). John Wiley.

