

Medidas de maniquís acústicos binaurales y espacialización de sonidos

Autor: **Javi García Morant**

Tutor: **José Javier López Monfort**

Cotutor: **Pablo Gutiérrez Parera**



Instituto de Telecomunicaciones y Aplicaciones Multimedia

Grupo de Tratamiento de Audio y Comunicaciones

Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación

Curso académico: 2016-2017

Resumen

El sonido binaural es aquel que se reproduce mediante auriculares y que proporciona la sensación de estar inmerso en el entorno en el que fue grabado, creando una “ilusión” sonora espacial al oyente. Este tipo de sonido se está empleando cada vez más en el mundo de la realidad virtual, sobre todo en videojuegos, donde experimentar la posición de una fuente de sonido aporta grandes ventajas.

En este trabajo se expone el procedimiento y los resultados del análisis y medida de dos maniquís acústicos (también denominados cabezas binaurales) en una cámara no anecoica. Para ello se han medido las respuestas al impulso binaurales (*binaural impulse response*, BIR) de los maniquís desde múltiples posiciones en 3D alrededor del mismo, hasta completar casi una esfera. Así mismo, se han construido altavoces especiales ex profeso que reunieran las características deseadas para esta tarea.

La aplicación más relevante de este trabajo reside en la posibilidad de auralizar cualquier sonido con las respuestas binaurales medidas, proporcionando la sensación de inmersión al oyente y logrando la binauralización del sonido.

Palabras clave: Sonido binaural, medidas binaurales, HRTF, sonido espacial, cabezas binaurales.

Abstract

Binaural sound is the one that is played through headphones and provides the sensation of being immersed in the environment where the sound was recorded, creating a spatial sound illusion to the listener. This type of sound is being widely used in the virtual reality field, mostly in videogames, where knowing the sound source's position gives advantages to the user.

In this study, the procedure and results of the analysis and measurement of two acoustic dummy heads in a non-anechoic chamber are exposed. In order to do it, binaural impulse responses (BIR) of the dummy heads have been measured from multiple 3D positions around them, covering almost the entire surface of a sphere. In addition, speakers with customized features have been built to make these measurements.

The most relevant application of this study is the possibility to create a new binaural sound just by using the BIR measurements, giving the immersed sensation to the listener and achieving sound's binauralization.

Keywords: Binaural sound, binaural measurements, HRTF, spatial sound, dummy heads.

Resum

El so binaural és aquell que es reproduïx mitjançant auriculars i que proporciona la sensació d'estar immers en l'entorn en el que fou gravat, creant una "il·lusió" sonora espacial a l'oient. Aquest tipus de so s'està emprant cada volta més en el món de la realitat virtual, sobre tot en videojocs, on experimentar la posició d'una font de so aporta grans avantatges.

En aquest treball s'exposa el procediment i els resultats del anàlisi i mesura de dos maniquís acústics (també denominats caps binaurals) en una càmera no-anechoica. Per a açò, s'han mesurat les respostes al impuls binaurals (*binaural impulse response*, BIR) dels maniquís desde múltiples posicions en 3D al voltant del mateix, fins completar quasi una esfera. Així mateix, s'han construït altaveus especials que reunien les característiques desitjades per a esta tasca

L'aplicació més relevant d'este treball resideix en la possibilitat d'auralitzar qualsevol so amb les respostes binaurals gravades, proporcionant la sensació d'immersió a l'oient i aconseguint la binauralització del so.

Paraules clau: So binaural, mesures binaurals, HRTF, so espacial.

Índice

	pág.
0. Introducción y objetivos	1
1. Respuestas binaurales	2
1.1. Principios físicos de la audición binaural	2
1.1.1. Percepción en azimut: <i>Duplex theory</i>	2
1.1.2. Percepción en elevación: <i>Pinna filtering effect</i>	4
1.1.3. Percepción de la distancia	5
1.2. Técnica de medida: <i>Sweep</i>	6
1.3. Tipos de respuestas binaurales	7
1.4. Aplicaciones	10
2. <i>Hardware</i> empleado en las medidas	11
2.1. Cabezas de medida	11
2.2. Medidas en elevación	12
2.2.1. Estructura de soporte de los altavoces	12
2.2.2. Altavoces	12
2.3. Medidas en azimut	14
2.4. Amplificadores	15
2.5. Interfaz de audio	15
2.6. Micrófono para medir la respuesta de los altavoces	16
3. <i>Software</i> empleado en las medidas	18
3.1. <i>Adobe Audition 3</i>	18
3.2. <i>MegaIR</i>	18
3.3. <i>Matlab</i>	19
4. Metodología y medidas realizadas	20
4.1. Medida de las cabezas	20
4.2. Corrección de la respuesta de los altavoces	23
4.3. Procesado de las medidas	28
4.3.1. Corrección de las medidas	28
4.3.2. Normalización y almacenamiento en formato *.wav	29
5. Resultados	31
6. Conclusión	34
7. Referencias	35
Anexo I	36
Anexo II	38
Anexo III	40

0. Introducción y objetivos

El sonido binaural consiste en la reproducción de señales captadas por un maniquí acústico sobre los oídos del oyente mediante auriculares. Los maniqués acústicos o cabezas binaurales son dispositivos empleados para simular la audición humana, disponiendo de un micrófono en cada canal auditivo para este fin (fig. 12).

Situando el maniquí en un lugar concreto podemos grabar los sonidos que percibe en un determinado momento y reproducirlos en cualquier persona. Esta técnica se conoce como grabación binaural. La limitación de la grabación binaural es que debe realizarse siempre in-situ, con el maniquí situado físicamente en el lugar de la grabación. Un paso más avanzado en esta técnica se denomina síntesis binaural y consiste en sintetizar los sonidos binaurales a partir del sonido monofónico de una fuente sonora puntual en el espacio.

Esto se consigue mediante técnicas de convolución. Para ello, se deben capturar las respuestas al impulso binaurales (*Binaural Impulse Response*, BIR) del maniquí desde diversas posiciones del espacio. A partir de la BIR para la posición de la fuente que se desea simular, se realiza la convolución de dicha respuesta al impulso con el sonido, de forma independiente para cada oído. El resultado de esta convolución es el sonido que captaría el maniquí si la fuente estuviera en la posición desde la que se midió la BIR.

La utilidad de la síntesis binaural reside en la capacidad de procesar cualquier tipo de sonido para que, mediante los auriculares, se escuche en un punto concreto del espacio. Esto es realmente útil en realidad virtual, que no sólo existe en imagen, sino también en audio. El campo que más está empleando este tipo de procesamiento en la actualidad son los videojuegos de realidad virtual.

En el presente trabajo se compararán dos tipos de maniqués acústicos con los cuales se logrará una binauralización de cualquier tipo de sonido. Para ello, se grabarán las respuestas al impulso para una malla de 36 posiciones en el plano horizontal (equiespaciados a 10°), logrando cubrir los 360° de azimut, y otra de 10 posiciones en el plano medio (nuevamente, equiespaciados a 10°), cubriendo 90° de elevación, desde -20° hasta 70°. Por tanto, habrá un total de 360 posiciones, es decir, 360 medidas para cada maniquí acústico, cubriendo una superficie cercana a la de una esfera. Todas estas posiciones se encuentran a 1.5 m. de distancia de las fuentes sonoras. Además, se medirán la respuesta de los altavoces para poder corregir las medidas realizadas, evitando la influencia de la respuesta de los altavoces sobre las medidas.

Los objetivos de este estudio son los siguientes:

1. Dar a conocer aspectos básicos y fundamentales del sonido binaural.
2. Explicar la técnica de medida empleada, los diagramas de bloques utilizados y el *hardware* y *software* necesario para finalizar con la explicación del proceso de medida en sí mismo.
3. Discutir los resultados obtenidos, observar los fenómenos físicos de la audición binaural y comparar los maniqués medidos.

Para lograr dichos objetivos, en primer lugar, se explican conceptos teóricos sobre la audición binaural y los principales fenómenos físicos que afectan a la localización del sonido en el espacio, así como la técnica de medida empleada y los tipos y aplicaciones de las respuestas binaurales.

En segundo y tercer lugar, se presentan el *hardware* y el *software* empleado en la realización de las medidas, para pasar al cuarto punto donde se expone la forma en que han sido realizadas.

En quinto lugar, se discuten los resultados obtenidos, realizando diversas comparativas, y por último se realiza una breve conclusión en el sexto punto.

1. Respuestas binaurales

1.1. Principios físicos de la audición binaural

Los humanos escuchamos los sonidos con nuestro oído, empezando por nuestras orejas, también llamadas pabellones auriculares o *pinna*, y los interpretamos en nuestro cerebro. El rango de la audición humana se sitúa entre los 20 Hz y los 20 kHz.

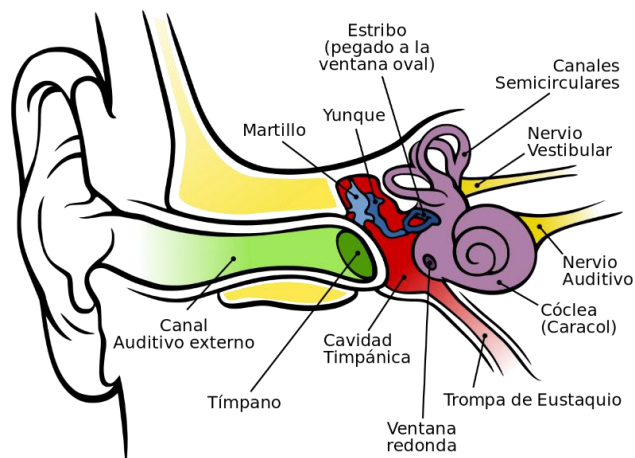


Fig. 1. Sistema auditivo humano

El ser humano percibe los cambios de presión de las ondas acústicas en el tímpano, que es donde la señal acústica se transforma en señal mecánica y se transmite por el martillo, yunque y estribo hasta llegar a la cóclea. El estribo genera ondas en el líquido del laberinto óseo, llamado endolinfa, y estas ondas provocan el movimiento en las células ciliadas que estimulan los nervios conectados al cerebro. [1]

Estableciendo un sistema de coordenadas esféricas, el ser humano es capaz de localizar el sonido de acuerdo a tres parámetros: distancia, elevación o colatitud y azimut.

1.1.1. Percepción en azimut: *Duplex theory* (teoría dúplex)

Esta teoría, propuesta por Lord Rayleigh (1907) al realizar experimentos con tonos puros [2], explica cómo se localizan los sonidos en el plano azimutal mediante dos fenómenos interaurales: diferencia de tiempos interaural (*interaural time difference*, ITD) y diferencia de intensidad interaural (*interaural intensity difference*, IID).

Una ITD se define como la diferencia de tiempos del mismo frente de ondas entre los oídos derecho e izquierdo. Por ejemplo, el frente de ondas de una fuente sonora situada a 45° a la derecha de un oyente llegaría antes al oído derecho de dicha persona, como se muestra en la figura 2.

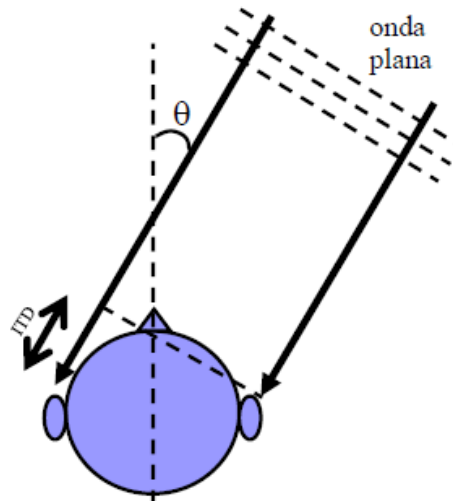


Fig. 2. Representación esquemática de ITD

De forma similar, una IID se define como la diferencia de amplitudes percibidas entre los oídos derecho e izquierdo para una misma fuente sonora en campo libre (fig. 5).

La ITD tiene dos limitaciones. Por una parte, existen puntos distintos que pueden provocar la misma ITD. Por ejemplo, una fuente situada a 45° provocaría la misma ITD que otra fuente situada a 135° , como se puede ver en la figura 3. A este fenómeno se le denomina confusión delante-atrás.

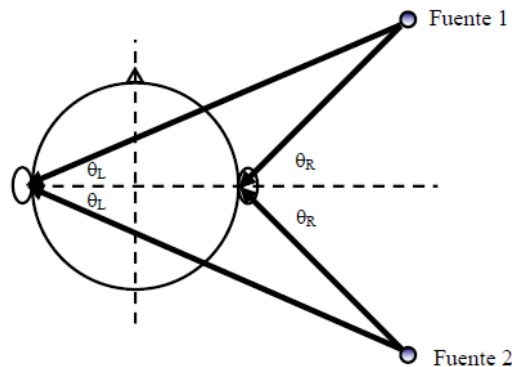


Fig. 3. Confusión delante-atrás por dos puntos con la misma ITD

Por otra parte, el sistema de audición humano percibe la ITD como una diferencia de fases más que como una diferencia de tiempos (para sinusoides puras), por ello, alrededor de 1500 Hz, la longitud de onda de una sinusoide es comparable a la distancia interaural y provoca que las ITD sean ambiguas. A partir de dicha frecuencia y superiores, la fase percibida por el oído puede ser errónea ya que la sinusoide podría haber girado más de una longitud de onda, provocando la percepción de ITDs menores, como se muestra en la figura 4.

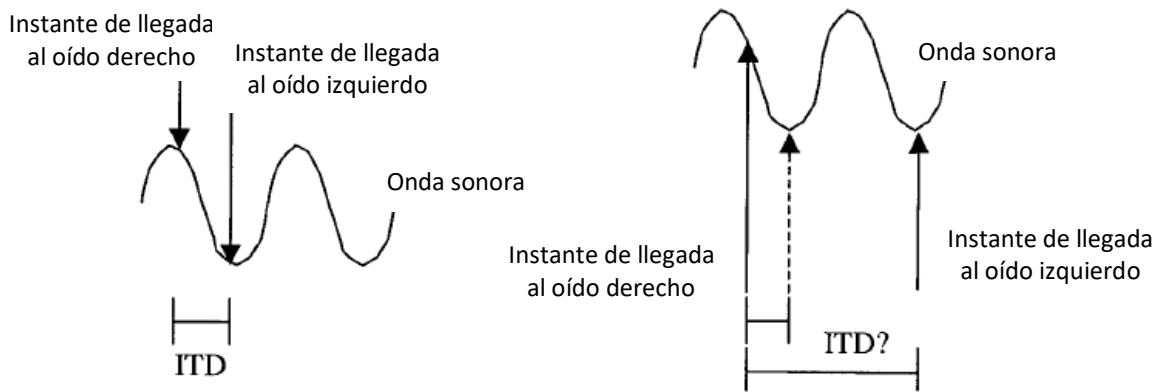


Fig. 4. Izquierda: ITD menor que una longitud de onda corresponde a un único valor de azimut
 Derecha: ITD mayor que una longitud de onda corresponde a uno o más valores de azimut [3]

Sin embargo, a partir de 1500 Hz el sistema de localización que prevalece es el de IID, ya que, debido al tamaño comparable entre cabeza y longitud de onda, la cabeza provoca una difracción en la onda, lo que produce un efecto de sombra en la oreja que percibe el frente de ondas difractado, recibiendo así un nivel menor (fig. 5). Además, como cualquier onda en campo libre, a mayor frecuencia, existe mayor atenuación causada por la cabeza. [3]

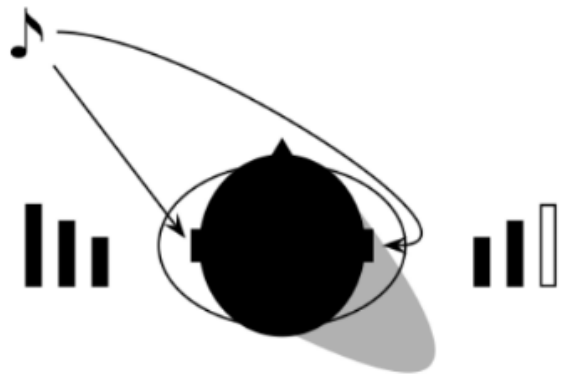


Fig. 5. Efecto sombra sobre IID por difracción de la cabeza [4]

Con todo ello, la percepción de los sonidos en el plano horizontal se realiza con dos mecanismos combinados: ITD desde frecuencias bajas hasta alrededor de 1500 Hz e IID desde 1500 Hz hasta las altas frecuencias de audio. La transición entre ambos se realiza de forma gradual siendo el cerebro el que procesa la información y calcula la dirección de llegada.

1.1.2. Percepción en elevación: *Pinna filtering effect* (efecto de filtrado de la pinna)

Hasta ahora se ha visto cómo se localizan sonidos en el plano azimutal, pero el ser humano también es capaz de localizar sonidos en distintos ángulos de elevación. Para ello, aunque existen varias teorías al respecto, la más importante es la teoría de las diferencias de timbre.

Esta teoría explica que cuando una onda acústica llega a la cabeza del receptor, sufre distorsiones lineales por difracción y efectos de sombra causados por la propia cabeza, la *pinna*, el canal auditivo externo [5] e incluso los hombros. Por ello, el espectro de la señal cuando llega al tímpano varía con el ángulo de incidencia, lo cual es advertido por el receptor como cambios en el timbre [6]. En la figura 6 se muestra un ejemplo del efecto que tendría la *pinna* sobre la respuesta de una señal que incidiese desde arriba o desde el frente.

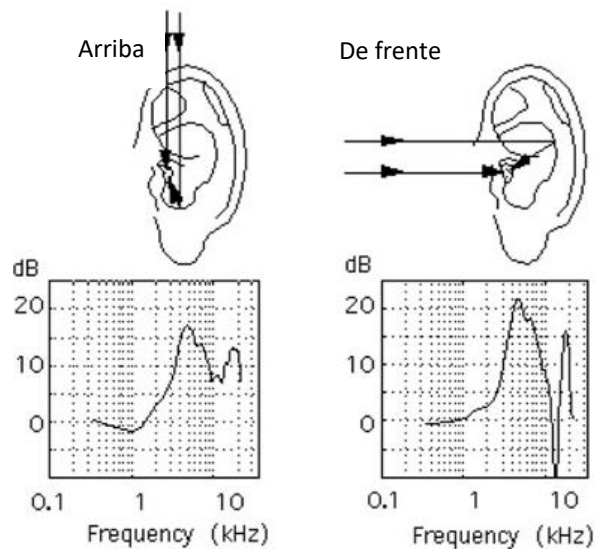


Fig. 6. *Pinna filtering effect*, alteración del espectro según ángulo de incidencia vertical (90°) u horizontal (0°)

1.1.3. Percepción de la distancia

Por último, la percepción de la distancia se puede determinar de cinco formas distintas o con combinaciones de ellas.

Relación sonido directo-sonido reverberante

Hace referencia al nivel de reverberación del entorno en el que nos encontremos. En general, a medida que una fuente se aleja del receptor, la proporción de energía de la onda directa que le llega disminuye, mientras que la energía de la onda reflejada en las diversas superficies del entorno (causando la reverberación) aumenta. Por ello, cuanto más alejada se sitúa una fuente sonora más lejos la percibiremos debido al aumento de la energía de las ondas reflejadas. [7] [8]

Intensidad

La intensidad o nivel de sonido es una de las formas más comunes de percepción de la distancia, siendo muy conocida la pérdida de 6 dB cada vez que se dobla la distancia de la fuente al receptor. Así mismo, también puede inducir a error ya que, si el nivel de la fuente aumenta o disminuye, el receptor lo podría situar más cerca o más lejos respectivamente, cuando en realidad la posición física de la fuente no habría cambiado. [7] [8]

Espectro

La forma del espectro de un sonido también se puede emplear para determinar la distancia de fuentes de sonido situadas a más de 15 m. del receptor [9] así como para sonidos cercanos al receptor [10]. En campo lejano, las altas frecuencias sufren una mayor atenuación que las bajas frecuencias cuando alcanzan el receptor, alterando así la forma del espectro y produciendo un sonido apagado y poco brillante. De la misma manera, puede ser relevante en campo cercano, debido a la forma en la que las difracciones producidas por la cabeza, *pinna* y hombros varían con la frecuencia y la distancia. [8]

IID

Además de la posición horizontal, la IID puede emplearse en campo cercano para la determinación de la distancia, ya que, al difractar la onda en la cabeza produciendo así el efecto sombra (fig. 5), la diferencia de nivel será mayor cuanto más cerca esté la fuente, ya que dicho efecto será mayor. [8]

Paralaje por movimiento

Se denomina paralaje por movimiento al cambio en el ángulo de dirección del sonido provocado por el movimiento relativo del receptor respecto de la fuente. Si dicha variación del ángulo es elevada, implica que la fuente está cerca, mientras que, si es baja o imperceptible, significa que la fuente está lejos. El fenómeno se explica visualmente en la figura 7. [8]

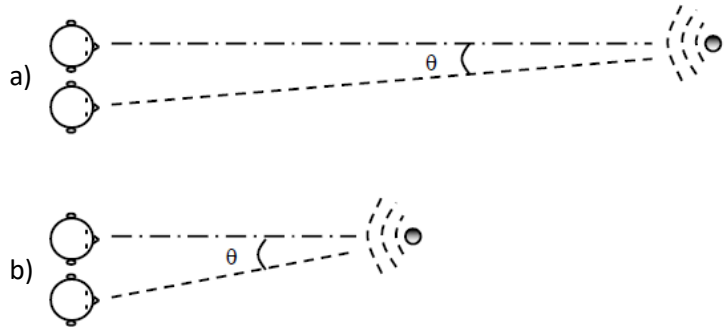


Fig. 7. Paralaje por movimiento. Ángulo θ menor en caso a) y mayor en caso b)

1.2. Técnica de medida: Sweep

Principalmente, existen dos tipos de técnicas para medir altavoces y maniquís binaurales: *Maximum Length Sequence* (secuencia de longitud máxima, MLS) y *sine sweep* (barrido senoidal). Para las diversas medidas que se han realizado se ha empleado la técnica *sine sweep* logarítmico. Esta técnica, desarrollada por Angelo Farina [11], se emplea para obtener la respuesta al impulso de un sistema, en este caso, un sistema acústico, como pueden ser las cabezas de medida o los altavoces.

Si se considera que tanto la fuente sonora como el receptor son puntuales y omnidireccionales, la información del entorno en el que se encuentren emisor y receptor vendrá caracterizada por su respuesta al impulso (*impulse response*, IR), asumiendo la hipótesis de que la acústica de una sala es lineal e invariante en el tiempo. Esta caracterización viene dada en los efectos sobre el tiempo (ecos y reflexiones) y sobre la frecuencia (respuesta en frecuencia y reverberación dependiente de la frecuencia).

Bajo estas hipótesis, una habitación puede verse como un sistema lineal invariante con una entrada, una salida y un ruido añadido por el propio sistema de medida, como se muestra en la figura 8.

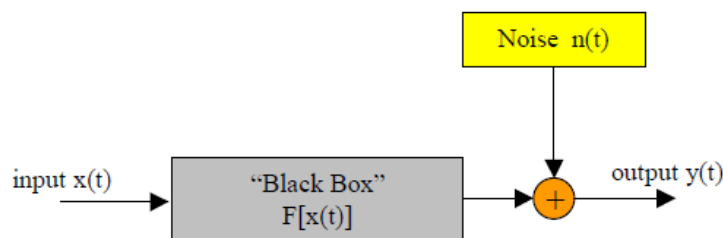


Fig. 8. Representación de una habitación (*black box*, caja negra) a medir [12]

La señal de entrada es generada por un ordenador y amplificada para transmitirse a través de un altavoz dentro de la sala. El receptor, un micrófono, capta dicha señal con todos los efectos de la sala y después de ser pre-amplificada, la señal se digitaliza en el mismo ordenador.

Sin embargo, la hipótesis de partida no es real, puesto que un altavoz puede presentar fenómenos no lineales, con lo cual la propagación por la sala no es invariante en el tiempo. Con

todo ello, se plantean tres problemas: el ruido del propio sistema de medida, los fenómenos no lineales del altavoz y, en consecuencia, la variabilidad temporal de la propagación en la sala.

El método de *sine sweep*, basado en un *sweep* exponencial con una de-convolución no periódica, proporciona solución a los tres problemas planteados: la reducción de ruido es mejor que con una señal MLS de la misma longitud, los efectos no lineales se separan de la respuesta lineal y el uso de un *sweep* largo evita cualquier problema en caso de que el sistema tenga alguna variabilidad temporal. [12]

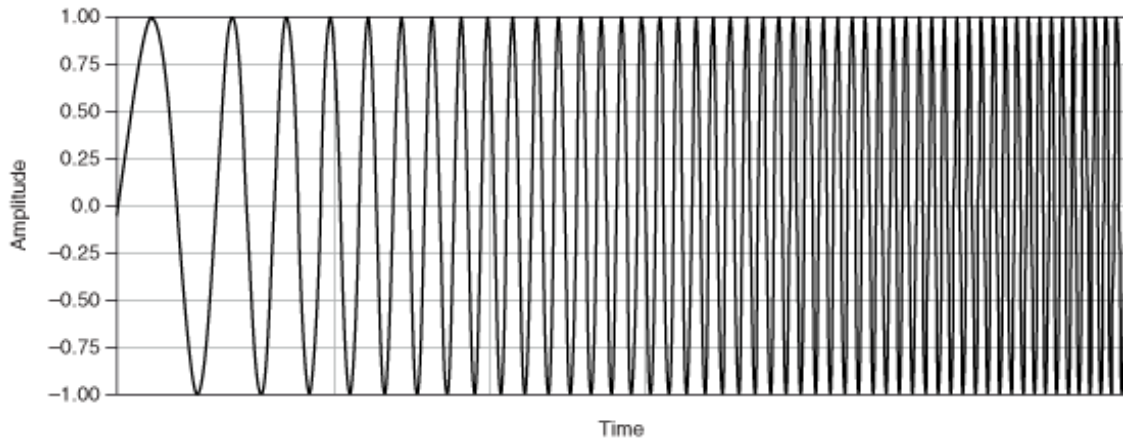


Fig. 9. Señal *sweep*

El *sine sweep* se define como:

$$x(t) = \sin \left[\frac{\omega_1 \cdot T}{\ln(\frac{\omega_2}{\omega_1})} \cdot \left(e^{\frac{t}{T} \cdot \ln(\frac{\omega_2}{\omega_1})} - 1 \right) \right] \quad [12]$$

Siendo ω_1 y ω_2 las frecuencias angulares de inicio y fin respectivamente y T la duración del *sweep*.

1.3. Tipos de respuestas binaurales

Existen dos tipos de respuestas binaurales: las respuestas al impulso relacionadas con la cabeza (*head-related impulse response*, HRIR) y las respuestas al impulso binaurales de sala (*binaural room impulse response*, BRIR). Tanto en cuanto BRIR y HRIR se graban por pares (un canal para cada oreja), una pareja de HRIRs o BRIRs determinan la sensación sonora de un punto en el espacio.

En primer lugar, las HRIR son respuestas al impulso binaurales resultantes de las distintas reflexiones y difracciones del sonido en el cuerpo humano antes de llegar al tímpano [13]. A diferencia de las BRIR, las HRIR se graban en una cámara anecoica y empleando de la misma forma una cabeza para medidas binaurales o una persona.

En segundo lugar, las BRIR son respuestas al impulso binaurales de una sala. Esto implica que dichas respuestas se graban en una sala con cierto grado de reverberación (es decir, una sala no anecoica) empleando un maniquí para medidas binaurales o una persona. [14]

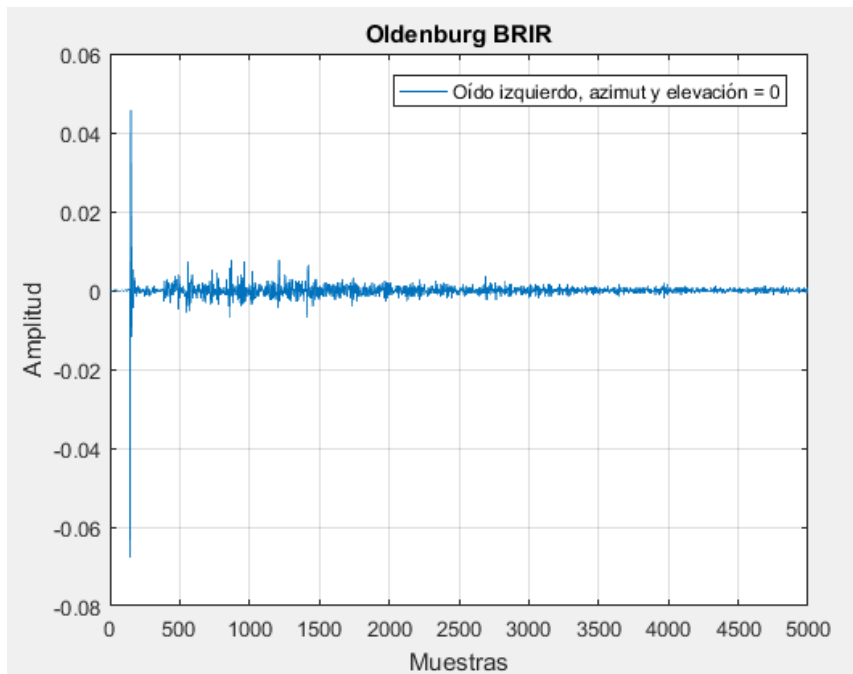


Fig. 10. BRIR de la librería libre de la universidad de Oldenburg [15]

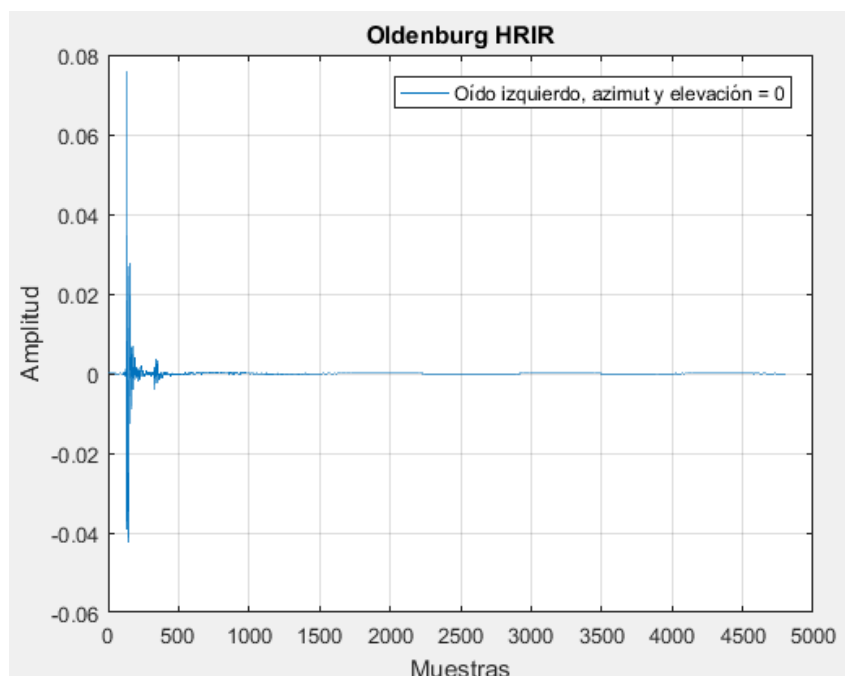


Fig. 11. HRIR de la librería libre de la universidad de Oldenburg [15]

Ambas respuestas son binaurales, pero mientras que las HRIR sólo tienen en cuenta las características de la cabeza/torso, las BRIR también tienen en cuenta el efecto que la sala produce en las medidas. Esto se puede observar en la gran cola de reverberación que se ve en la figura 10, mientras que en la HRIR (fig. 11) sólo se observa la primera reflexión a simple vista, para una escala muy similar.

HRIR y BRIR son personales y distintas para cada sujeto, puesto que ambas tienen en cuenta la respuesta binaural de la persona y/o maniquí que se mida, y estas características varían de una persona a otra, así como de un modelo de maniquí a otro, es decir, existe una individualización.

Las respuestas al impulso se caracterizan en el dominio temporal, sin embargo, la forma habitual de visualización de este tipo de datos es en el dominio frecuencial. Mediante la transformada directa de Fourier se puede obtener el espectro de las respuestas al impulso, es decir, la función de transferencia. Estas funciones de transferencia se denominan función de transferencia binaural de la sala (*binaural room transfer function*, BRTF) y función de transferencia binaural relacionada con la cabeza (*head-related impulse response*, HRTF). Se puede observar que la caracterización de la sala se encuentra en la alta variabilidad de la BRTF (fig. 12) provocada por las reflexiones de la sala, mientras que cuando sólo se pretende medir el efecto de la cabeza (HRTF), la respuesta en frecuencia es mucho más suave y limpia (fig. 13).

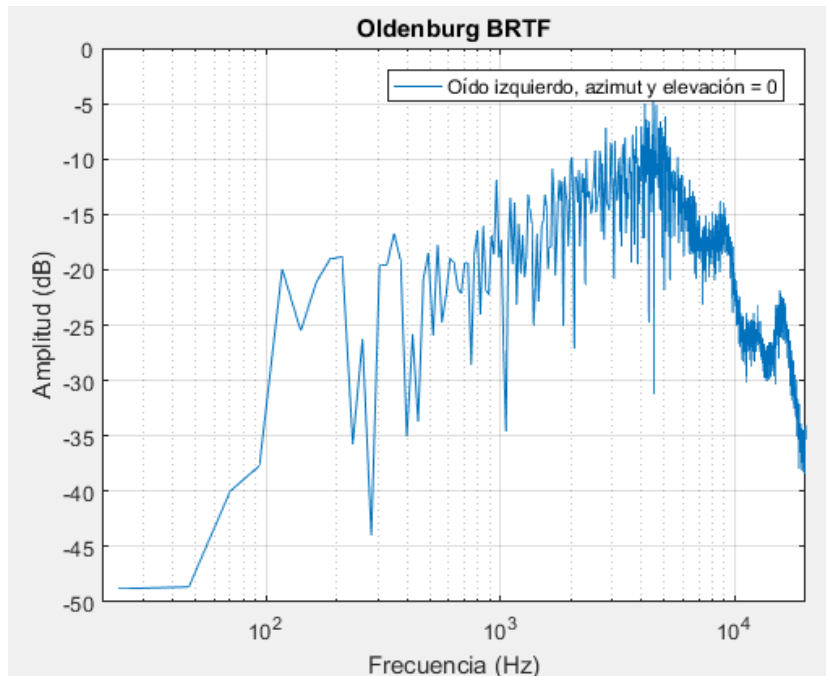


Fig. 12. Transformada de Fourier de BRIR (fig. 10)

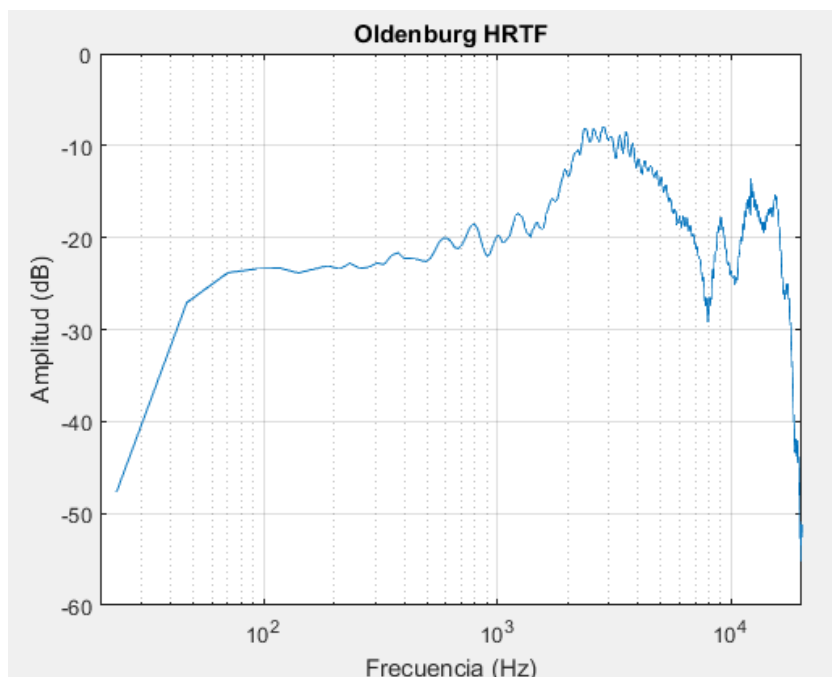


Fig. 13. Transformada de Fourier de HRIR (fig. 11)

En la bibliografía es habitual que la referencia a medidas binaurales en cámara anecoica se realice mediante "HRTF", aunque las medidas de las que se obtienen son HRIR.

1.4. Aplicaciones

Las medidas binaurales se utilizan para sintetizar sonidos binaurales, es decir, midiendo la HRIR o la BRIR de una persona o maniquí se obtienen las respuestas al impulso que, convolucionadas con cualquier sonido, proporcionan el mismo sonido como si lo escuchara la persona o maniquí al que se le midió su HRIR o BRIR, o sea, el sonido binauralizado. Con una malla suficientemente fina alrededor del sujeto (por ejemplo, midiendo HRIR o BRIR cada 10 o 5 grados), se podría situar el sonido donde se quisiera, logrando desplazar el sonido en todo el plano horizontal y en el plano medio, así como en distancia.

2. Hardware empleado en las medidas

Dado que uno de los objetivos de este trabajo es grabar las respuestas binaurales de dos cabezas, para ello se requieren diversos elementos *hardware* empleados para tal fin.

2.1. Cabezas de medida

Para medidas binaurales se emplean maniqués acústicos o cabezas construidas para este tipo de grabaciones (denominadas *dummy head recording* en inglés). Estos dispositivos se pueden dividir en dos tipos: El primer tipo modela en el mismo dispositivo la cabeza y los hombros o incluso medio torso de una persona humana a grandes rasgos, mientras que el segundo tipo sólo tiene en cuenta los efectos de la cabeza.

Ejemplos de estos dos tipos serían el maniqué de la empresa *Bruel & Kjaer*, con modelado de cabeza y hombros, y el de la empresa *Neumann*, sólo con modelado de cabeza.

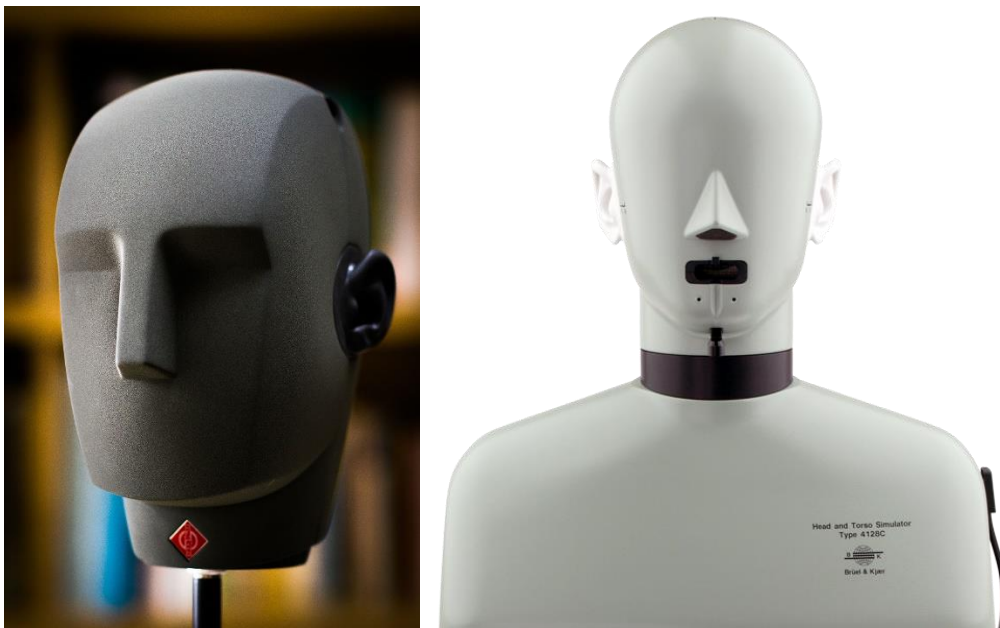


Fig. 14. Izquierda: *Neumann KU 100*
Derecha: *Bruel & Kjaer type 4128C*

Ambos tipos de maniqués se caracterizan por disponer de dos orejas hechas con un material maleable, emulando cartílago humano, que junto con la cabeza son las responsables de simular la forma en la que escucha una persona humana con un tamaño de cabeza y orejas estándar. Toda la influencia que los aspectos físicos de la cabeza tienen en el sonido se recoge con un micrófono situado en cada oreja.

En las medidas se emplearon dos maniqués diferentes: el modelo *Bruel & Kjaer type 4128C* y el modelo *Neumann KU 100*. Ambos se utilizan para medidas binaurales, con la diferencia de que el modelo de *Bruel & Kjaer* cuenta con torso, pudiendo así simular el efecto que tendría un torso humano sobre la escucha binaural, mientras que el modelo *Neumann* sólo tiene en cuenta el efecto de las orejas y la propia cabeza. Esta diferenciación hace aún más interesante la caracterización y comparación de ambas cabezas de medida.

2.2. Medidas en elevación

Para realizar las medidas en elevación se ha empleado una estructura de soporte de los altavoces, cajas para los altavoces y los propios altavoces.

2.2.1. Estructura de soporte de los altavoces

El objetivo de las medidas en elevación a realizar es cubrir 90° (desde -20° hasta 70°) en pasos de 10°. Para el estudio se disponía de un arco de 90°, conformado mediante varias tablas y listones de madera (fig. 15). Como las medidas a realizar en elevación son un total de diez, en la estructura se debían colocar diez altavoces que forman el *array* final.

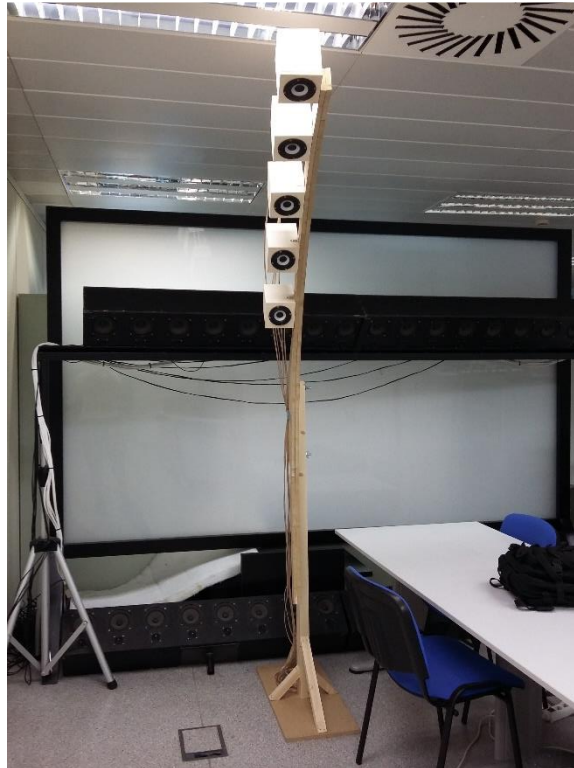


Fig. 15. Estructura de soporte con cinco altavoces ya montados

Para ello, se mide la longitud total del arco y se divide entre los nueve espacios que dejarán los altavoces.

$$\frac{L}{9} = \frac{235.5 \text{ cm}}{9} \cong 26.2 \text{ cm}$$

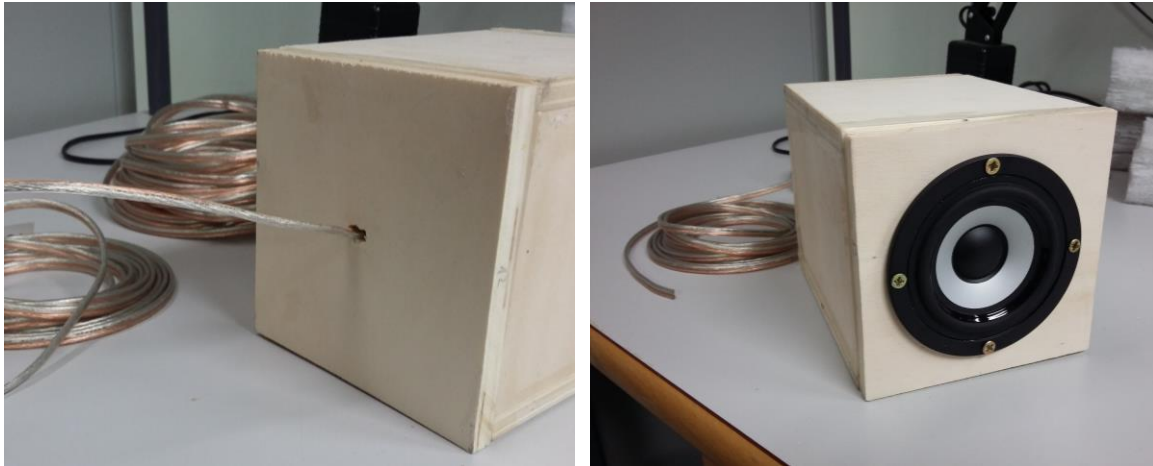
De esta forma, determinamos que las distancias entre los centros de los altavoces deberán estar a 26.2 cm de separación medida en el arco, sabiendo que los límites superior e inferior del arco corresponden a 70° y -20° respectivamente.

2.2.2. Altavoces

Como se disponía de seis altavoces ya contruidos y son necesarios diez, se construyeron cuatro más. Para la construcción de las cajas de los altavoces se han empleado tablas de madera pegadas entre sí. El resultado final son cuatro paralelepípedos de 12 cm. de longitud, 12 cm. de altura y 14 cm. de profundidad.

Dichas maderas se pegaron entre sí con cola blanca y se comprobó numerosas veces que no hubiera huecos entre las mismas, para evitar cambios en la respuesta en frecuencia del altavoz

provocadas por filtraciones del aire por los cambios de presión. El proceso de comprobación y secado se repitió hasta tres veces.

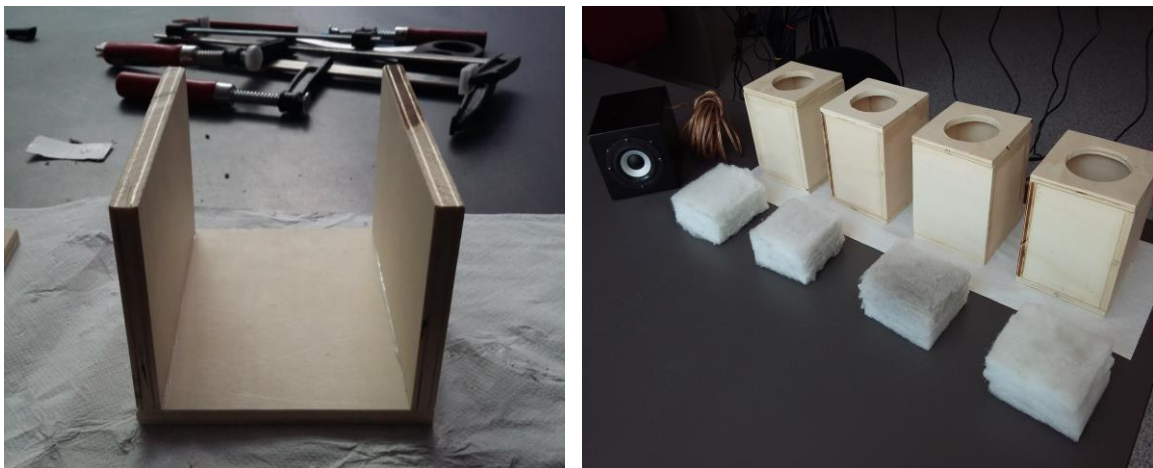


**Fig. 16. Izquierda: Proceso de pegado de las maderas
Derecha: Altavoz modelo, cuatro cajas terminadas y absorbentes**

Para poder insertar el altavoz, se les realizó un agujero circular en la parte frontal con una caladora. Además, para reducir en la mayor medida posible la reflexión del sonido en la parte trasera de la caja, se insertó un poco de absorbente (fig. 16 dcha.).

Antes de pegar las cajas, se realizó un agujero en la parte trasera para dejar hueco para el cable de pares (fig. 17 izqda.). Una vez introducido el absorbente, se procedió a pasar el cable por el agujero y soldar los altavoces para finalmente atornillarlos a la caja.

Los altavoces empleados son el modelo TB W3-315E.



**Fig. 17. Izquierda: Cable pasado por la parte trasera de la caja
Derecha: Altavoz finalizado**

Una vez soldados los altavoces, se colocaron conectores tipo banana en cada cable de pares para así facilitar la conexión a los amplificadores (fig. 18).



Fig. 18. Cables de pares finalizados con conectores banana

2.3. Medidas en azimut

El objetivo de las medidas en azimut a realizar es cubrir los 360° en el plano horizontal en pasos de 10° . Para ello, se ha empleado una plataforma giratoria dotada de un control mecánico (fig. 19) que permitía mover la plataforma con una precisión de hasta 0.1° . El ángulo de paso elegido ha sido el mismo que para las medidas en elevación, 10° , que corresponde a medio giro del control mecánico, con lo que son necesarios 36 medios giros para cubrir los 360° .

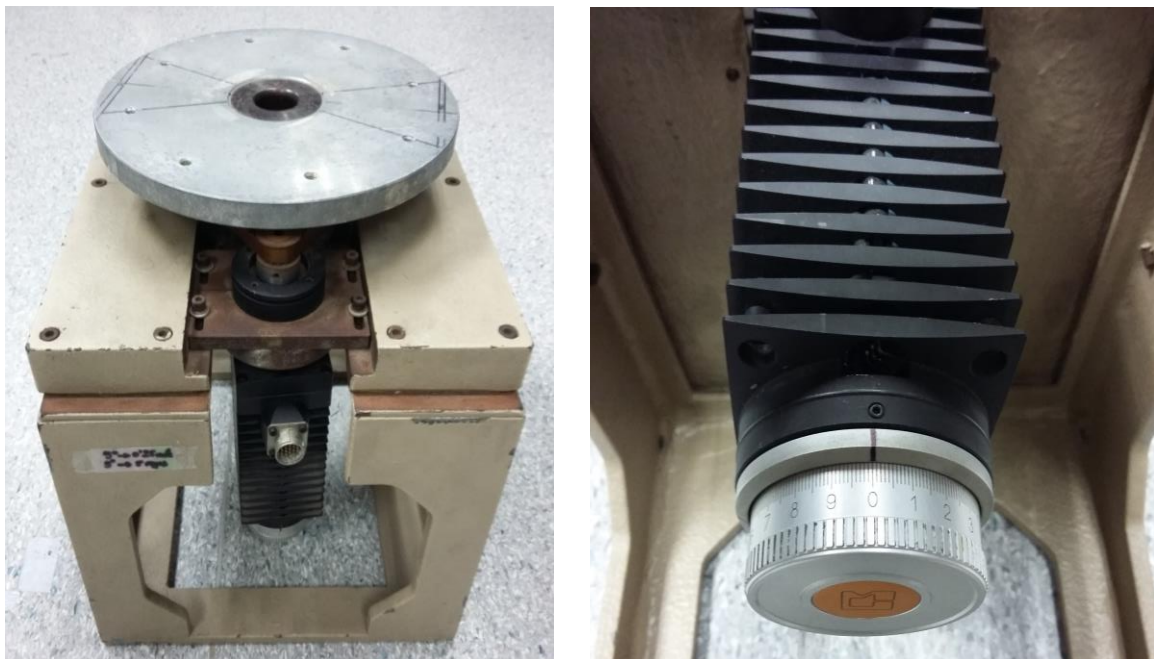


Fig. 19. Plataforma giratoria y control mecánico

2.4. Amplificadores

Los altavoces necesitan de un amplificador para poder reproducir sonido, por ello se han empleado dos amplificadores modelo *ONKYO TX-SR504E* (fig. 20). Cada uno cuenta con ocho salidas regulables de forma independiente, así como ocho entradas. Como se ha necesitado amplificar diez altavoces, se ha optado por emplear cinco salidas y entradas de un amplificador y cinco de otro. El ajuste del nivel de los canales se realizó y comprobó mediante el *software Adobe Audition 3*, que se comentará más adelante.



Fig. 20. Arriba: Panel delantero del amplificador
Abajo: Panel trasero del amplificador

2.5. Interfaz de audio

Para transmitir la señal de entrada de los micrófonos al ordenador y desde ahí enviar las salidas a las entradas del amplificador, es necesario una interfaz de audio. Para ello, se ha empleado la interfaz *MOTU Traveler* (fig. 21).



Fig. 21. Arriba: Panel delantero de la interfaz con la salida *phones* conectada
Abajo: Panel trasero de la interfaz con ocho salidas Jack conectadas

Esta interfaz cuenta con ocho salidas y ocho entradas en el panel trasero, siendo cuatro de las entradas balanceadas. Sin embargo, la señal de salida *phones* (auriculares) del panel delantero es independiente, y como es estéreo, se empleó como salidas nueve y diez para los

dos últimos altavoces. Además, en esta tarjeta se ha regulado el nivel de entrada (preamplificador, *preamp*) del micrófono de medida empleado (punto 2.6) y de los micrófonos de ambos maniqués acústicos.

2.6. Micrófono para medir la respuesta de los altavoces

Por último, la respuesta de los altavoces se midió con un micrófono de condensador modelo *Earthworks M30* (fig. 22).



Fig. 22. Micrófono de medida *Earthworks M30*

Este micrófono se caracteriza por ser muy adecuado para medidas de laboratorio, proporcionando un patrón polar muy omnidireccional y una respuesta en frecuencia plana, como se puede observar en la figura 23 y 24 respectivamente.

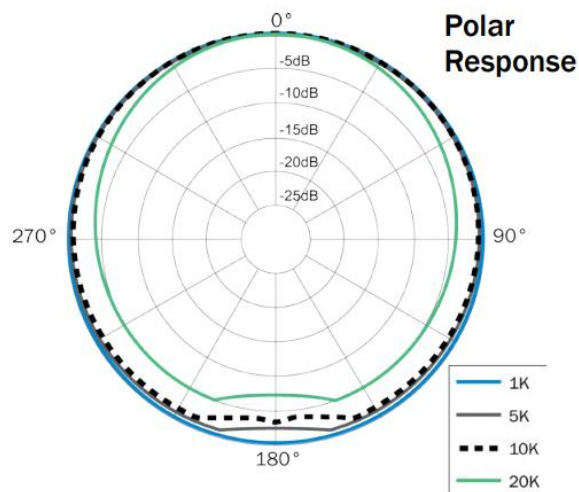


Fig. 23. Patrón polar del micrófono de medida *Earthworks M30*

Se observa que el error alcanza su máximo para 180° a una frecuencia de 20000 Hz, siendo este de poco más de 7 dB.

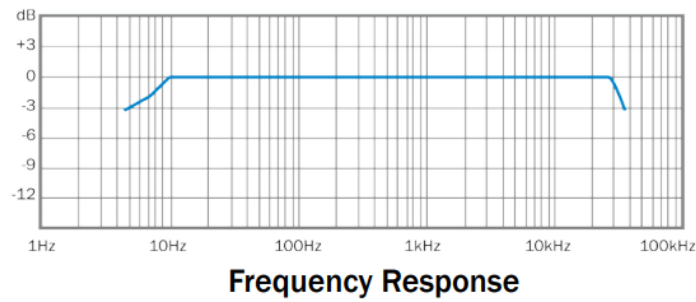


Fig. 24. Respuesta en frecuencia del micrófono de medida *Earthworks M30*

En cuanto a la respuesta en frecuencia, se puede ver que es completamente plana en todo el rango de frecuencias audibles por el ser humano (20 Hz a 20000 Hz).

Por estas razones, ha sido este modelo de micrófono el empleado en las medidas de los altavoces, que se explicarán con detalle más adelante.

3. Software empleado en las medidas

Durante el desarrollo de las pruebas, medidas y procesado, se han empleado diversos programas que se detallan a continuación.

3.1. Adobe Audition 3 [16]

Digital Audio Workstation (estación de audio digital, DAW) empleado para nivelar los altavoces mediante los amplificadores, ajustar las ganancias de los preamplificadores de la tarjeta y realizar diversas pruebas. El ajuste de niveles se ha realizado empleando ruido rosa, dado que es el tipo de ruido que proporciona un mismo nivel sonoro para todas las bandas.

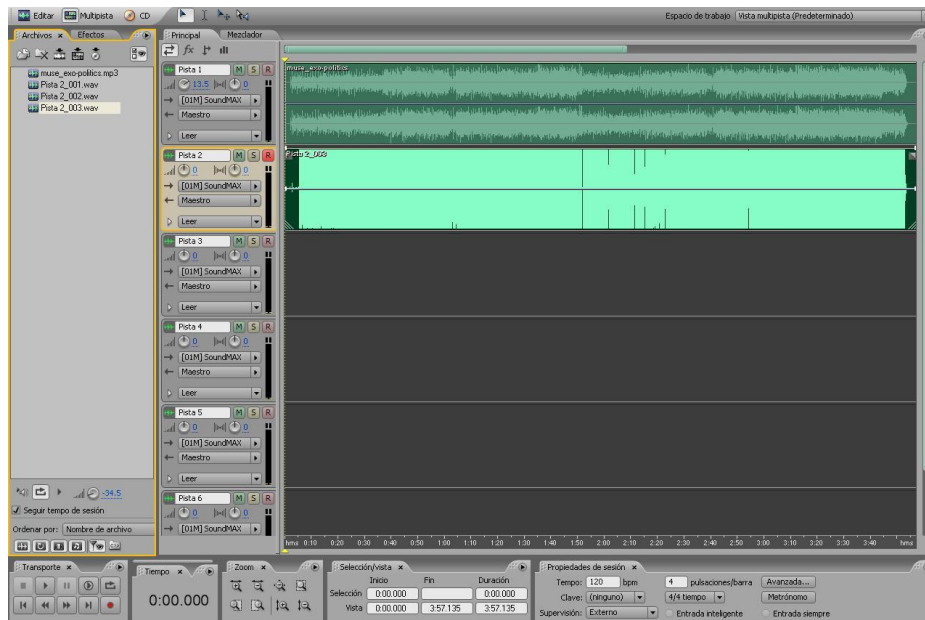


Fig. 25. Pantalla principal en modo *multitrack* de *Adobe Audition 3*

3.2. MegalR

Software desarrollado en el instituto de investigación donde se ha realizado el trabajo, empleado para reproducir los *sweeps* con los que se han medido las respuestas al impulso de los altavoces y de las cabezas (fig. 26).

Una vez reproducidos los *sweeps*, este programa guarda las medidas del número de altavoces especificado con un formato determinado por la cantidad de micros y de altavoces (*num. microphones* y *num. loudspeakers*), así como un nombre inicial (*IR File Name*). Estos nombres se modificaron para una mejor identificación de las medidas (explicado en el punto 4.3).

Seleccionando la tarjeta de salida, el programa reproduce los *sweeps* empezando por el altavoz especificado en *first loudspeaker* (primer altavoz) y sigue secuencialmente por el resto de salidas hasta reproducirlo el número de veces especificado en *num. loudspeakers* (número de altavoces). Como la tarjeta *MOTU Traveler* considera la salida de *phones* como salidas trece y catorce, el desarrollador del programa realizó una modificación opcional (casilla *avoid 9,10,11,12*, evitar 9,10,11,12) para que, después de terminar con la octava salida, pase directamente a reproducir el *sweep* por la salida trece.

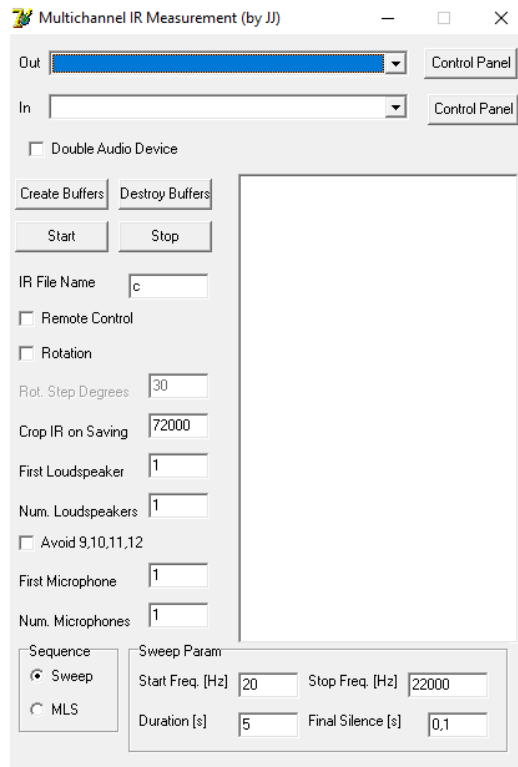
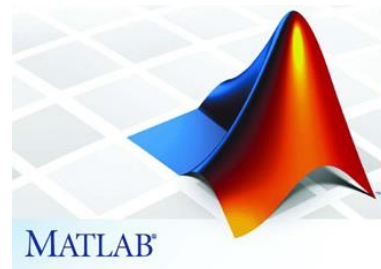


Fig. 26. Pantalla principal del programa MegalR

3.2. Matlab [17]

El conocido *software* de *Mathworks*, Matlab, ha sido empleado para diversas finalidades:

1. Renombrar las medidas guardadas por el programa MegalR para una mejor identificación de la posición desde donde se reprodujo el *sweep*.
2. Realizar y aplicar la corrección de los altavoces a las medidas.
3. Normalizar y guardar en formato *.wav las medidas finales.
4. Convolucionar las medidas finales con cualquier tipo de sonido para obtener el sonido binaural deseado.
5. Además, se han realizado diversas pruebas y gráficas a lo largo de todo el trabajo con el objetivo de corregir errores y comprobar resultados.



4. Metodología y medidas realizadas

4.1. Medida de las cabezas

Para la medida de las cabezas binaurales se ha empleado el siguiente diagrama de bloques.

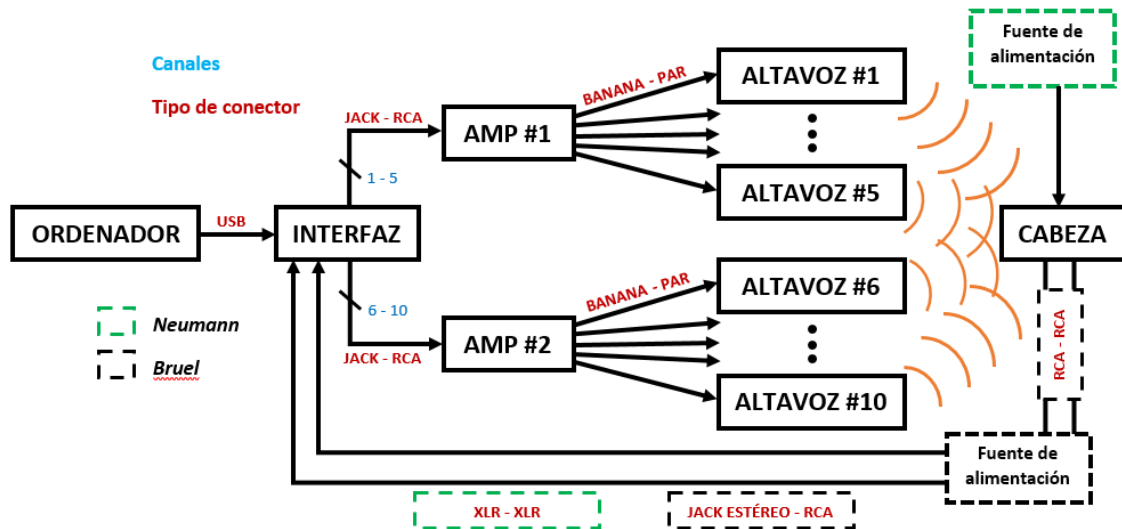


Fig. 27. Diagrama de bloques del análisis de las cabezas

En el caso del maniquí *Bruel* (color negro), es necesario incorporar un bloque de alimentación para los micrófonos, mientras que el maniquí *Neumann* (color verde) viene con una fuente de alimentación externa.

En ambos casos, los *sweeps* generados mediante el programa MegalR se envían por USB hasta la interfaz, la cual mediante diez cables jack-RCA envía las señales a los amplificadores. Una vez amplificada la señal, viaja a través de los cables de pares hasta reproducirse por los altavoces. La cabeza a analizar recibe las ondas acústicas en el micrófono situado en cada oreja y transmite las ondas eléctricas nuevamente a la tarjeta, que envía la señal al ordenador a través de USB (fig. 27).

El proceso de montaje para la realización de las medidas se empezó centrando la cabeza a medir en la plataforma giratoria, alineando el punto medio del eje interaural con el centro de giro (fig. 28).



Fig. 28. Cabeza Neumann centrada en la plataforma

Como se conoce el radio de la estructura de soporte, 1.5 m., y los oídos deben estar a una distancia equidistante de todos los centros de los altavoces, la cabeza se sitúa a 1.5 m. de distancia. A su vez, para garantizar dicha distancia, los oídos del maniquí deben situarse a la misma altura que el centro del altavoz a 0° , por esa razón se colocaron numerosos listones de madera para elevar la altura del maniquí hasta conseguir la deseada. Además, se colocaron paneles absorbentes para minimizar algunas reflexiones del suelo y paredes (fig. 29).



Fig. 29. Paneles absorbentes situados en el suelo y en la pared más cercana a la cabeza

Con todo ello, se pudo empezar a medir, situando el maniquí correspondiente a 0° de azimut. El proceso de medida seguía los siguientes pasos:

1. Se reproducen los *sweeps* desde el programa MegaIR para todos los ángulos de elevación y un mismo ángulo de azimut. Dicho programa almacena los datos grabados en formato *.txt.
2. Una vez guardados, mediante un *script* realizado en Matlab (“save_measurements”, anexo I) se cargan y se renombran para una mejor comprensión de los datos almacenados en los ficheros. El programa de renombre carga los *.txt guardados, junta los datos por pares para obtener en un mismo fichero la respuesta medida en cierto ángulo de azimut y elevación para los dos oídos, y los renombra de la forma az_[ángulo de azimut]_el_[ángulo de elevación].txt, con lo que se conoce la posición exacta de la medida.
3. Mientras dicho programa se ejecuta, mediante el control mecánico se gira la plataforma 10° hacia la derecha (siendo desde el sistema de referencia empleado en el campo de medidas binaurales –los altavoces- un giro de 10° hacia la izquierda).
4. Con el giro completado y los datos ya renombrados, se repite el proceso para el siguiente ángulo de azimut y todos los ángulos de elevación, hasta terminar en 350° de azimut.

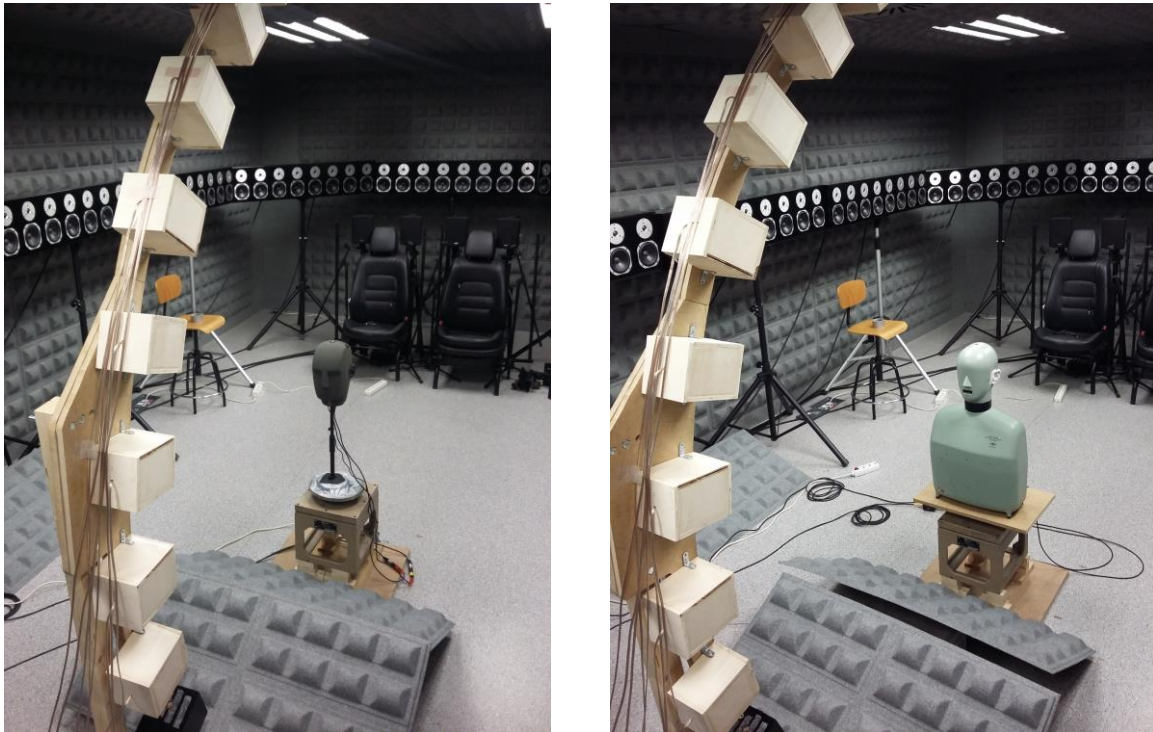


Fig. 30. Montaje para medida de la cabeza *Neumann* (izqda.) y *Bruel* (dcha.)

4.2. Corrección de la respuesta de los altavoces

Puesto que los altavoces empleados no son ideales, es necesario medir la respuesta en frecuencia de cada altavoz para poder luego corregir la medida binaural con la inversa de dicha respuesta, logrando así que las variaciones que sufre en su respuesta cada altavoz no afecten a las medidas.

Para ello, se ha empleado el siguiente diagrama de bloques.

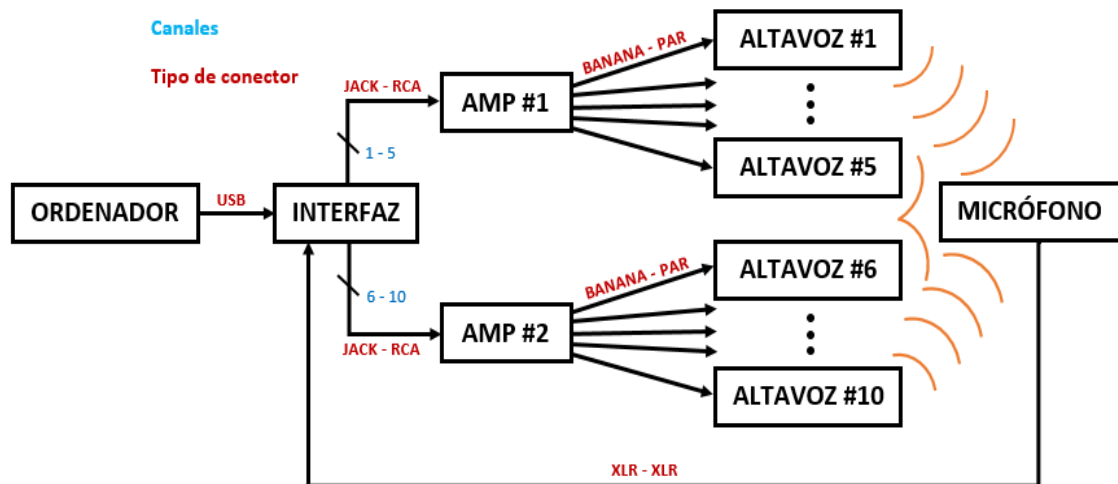


Fig. 31. Diagrama de bloques de la medida de los altavoces

Mediante el programa MegaIR se envían las señales *sweep* recorriendo las frecuencias de 20 Hz a 22000 Hz a través de un cable USB hasta la interfaz *MOTU*. Esta es la encargada de enviar dichas señales a los amplificadores a través de cables jack-RCA, los cuales las envían a los altavoces por el cable de pares. Finalmente, los barridos se graban con el micrófono conectado a una entrada balanceada de la interfaz mediante un cable XLR-XLR, habiendo ajustado la ganancia del pre-amplificador de entrada de la tarjeta con el programa *Audobe Audition 3*. El

programa MegaIR es el responsable de recoger los datos y obtener la respuesta al impulso de cada altavoz, que se guarda tanto en formato *.pcm como en formato *.txt (fig. 31).

El micrófono se situó a medio metro de cada altavoz, apuntando hacia su centro (fig. 32). Pese a ser un micrófono omnidireccional, se ha realizado de esta forma para evitar que la respuesta de la sala (reverberaciones) afecte a la respuesta real medida del altavoz, minimizando el camino del sonido directo y la posible variación en la omnidireccionalidad del micrófono (fig. 23).



Fig. 32. Medida de la respuesta de los altavoces

Una vez se tiene la respuesta al impulso (fig. 33), se puede trabajar con la respuesta en frecuencia (fig. 34) realizando la transformada directa de Fourier (punto 1.3).

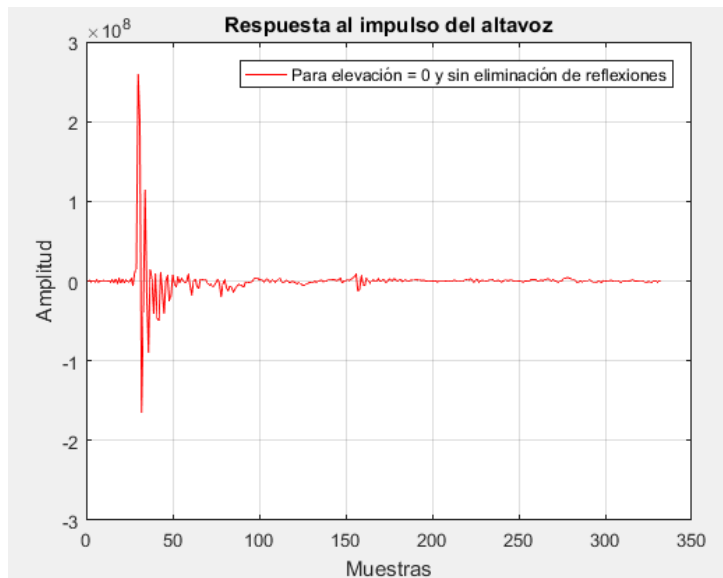


Fig. 33. Respuesta al impulso del altavoz a 0° sin corregir

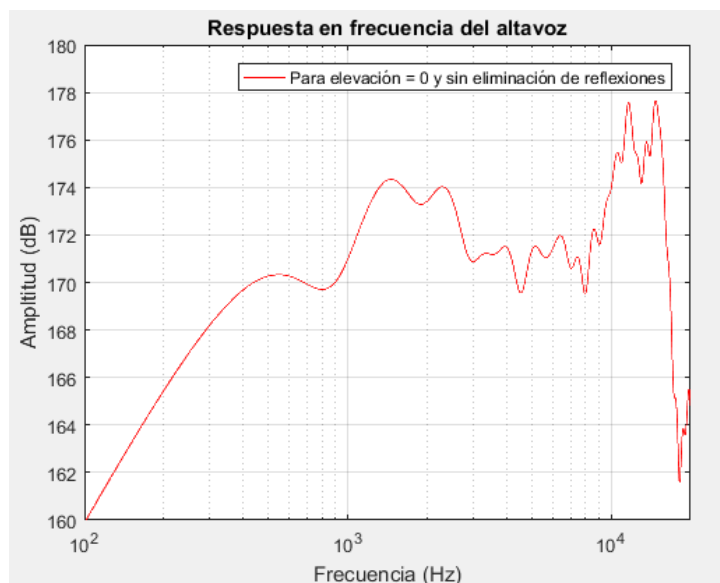


Fig. 34. Respuesta en frecuencia del altavoz a 0° sin corregir

Para corregir la respuesta del altavoz se han empleado dos procesados: la eliminación de las primeras reflexiones de la sala y la realización del filtro inverso de fase mínima.

Para evitar las reflexiones se puede cortar la respuesta al impulso antes de la llegada del primer eco de la sala. En nuestro caso el primer eco estaba presente alrededor de la muestra 1800. Sin embargo, acortar tanto la respuesta al impulso tiene una limitación: produce un efecto de eventanamiento que nos hace perder resolución en bajas frecuencias y altera la medida en las octavas más bajas.

Para minimizar el efecto de eventanamiento, se ha empleado un método de eventanado sub-banda con una resolución de 20 Hz (recortando la respuesta al impulso sin reflexiones a 2400 muestras para una frecuencia de muestreo de 48000 Hz). El método ya estaba desarrollado en el grupo de investigación donde se ha realizado este proyecto y se describe a continuación:

Método de enventanado sub-banda

- 1) Se definen:
 - a. Las N bandas de trabajo.
 - b. La frecuencia de corte f_c de la banda más alta.
 - c. El tamaño W de la ventana de la ventana más alta.
- 2) Se enventana la señal con un tamaño W y se calcula su respuesta en frecuencia mediante la transformada de Fourier.
- 3) Se dobla W y se divide por el doble de f_c y se repite el punto 2) hasta completar las N bandas.
- 4) Se combinan las respuestas en frecuencia con cada sub-banda para obtener la respuesta en frecuencia total.
- 5) Se obtiene la IR a partir de 4), realizando la transformada de Fourier inversa.

Este método está implementado en una función de Matlab denominada “irsubbanda”.

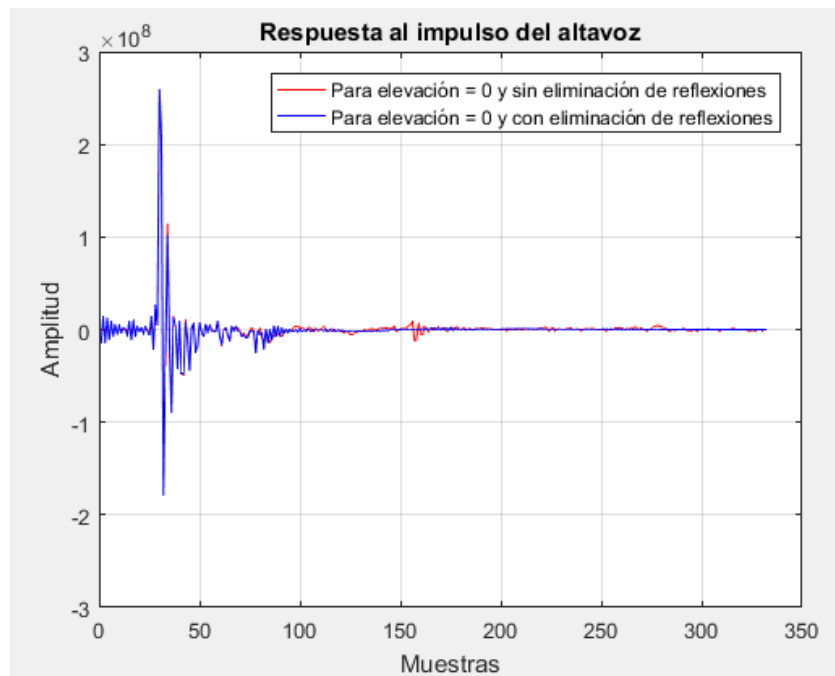


Fig. 35. Respuesta al impulso del altavoz a 0° con primeras reflexiones (rojo) y sin ellas (azul)

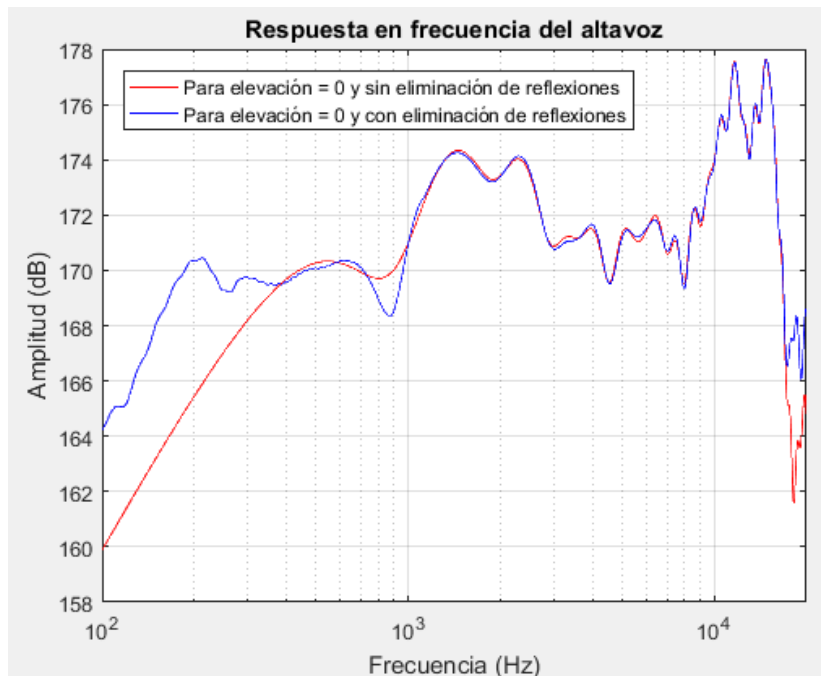


Fig. 36. Respuesta en frecuencia del altavoz a 0° con primeras reflexiones (rojo) y sin ellas (azul)

Como se puede observar en la respuesta al impulso (fig. 35), el método de eliminación de reflexiones funciona de forma muy eficaz, limpiando la señal y dejando sólo el impulso deseado. Además, en la respuesta en frecuencia (fig. 36) vemos que se ha obtenido mayor resolución a bajas frecuencias.

En segundo lugar, dado que el objetivo es corregir la respuesta del altavoz, es decir, obtener una respuesta en frecuencia que en el caso ideal sería plana, hay que aplicar un filtro que sea la inversa del espectro del altavoz. En el campo de la escucha binaural, este filtro debe ser de fase mínima, ya que un cambio en la fase puede suponer, para ciertas frecuencias, un cambio en el instante en el que se percibe el sonido, lo que puede llevar a confusiones (punto 1.1.1, fig. 4).

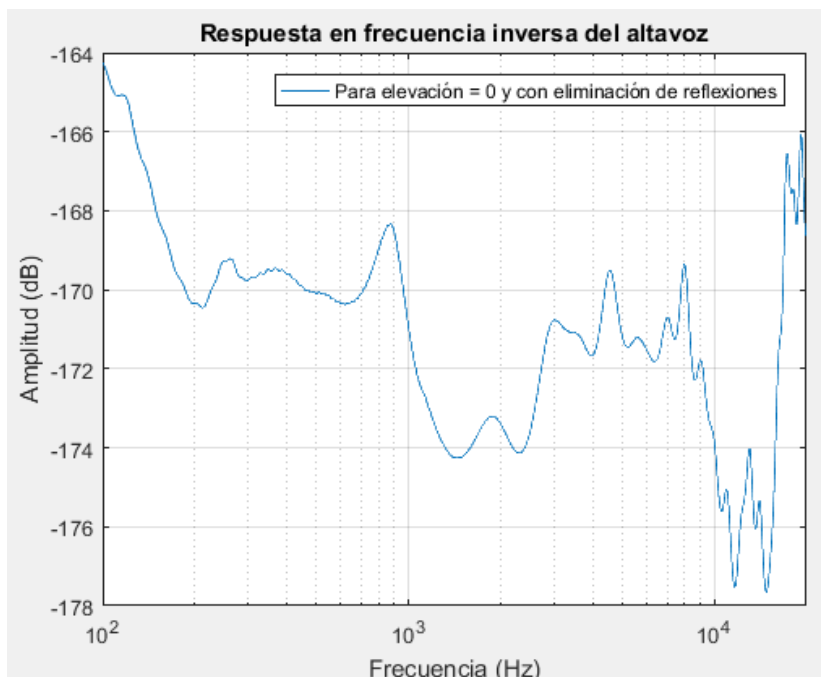


Fig. 37. Respuesta en frecuencia inversa del altavoz a 0°

A simple vista se puede observar que la respuesta en frecuencia inversa obtenida (fig. 37) es, efectivamente, la inversa de la visualizada previamente (fig. 36). No obstante, realizando la convolución entre ambos espectros se comprueba el correspondiente espectro plano del altavoz ideal mencionado anteriormente (fig. 38).

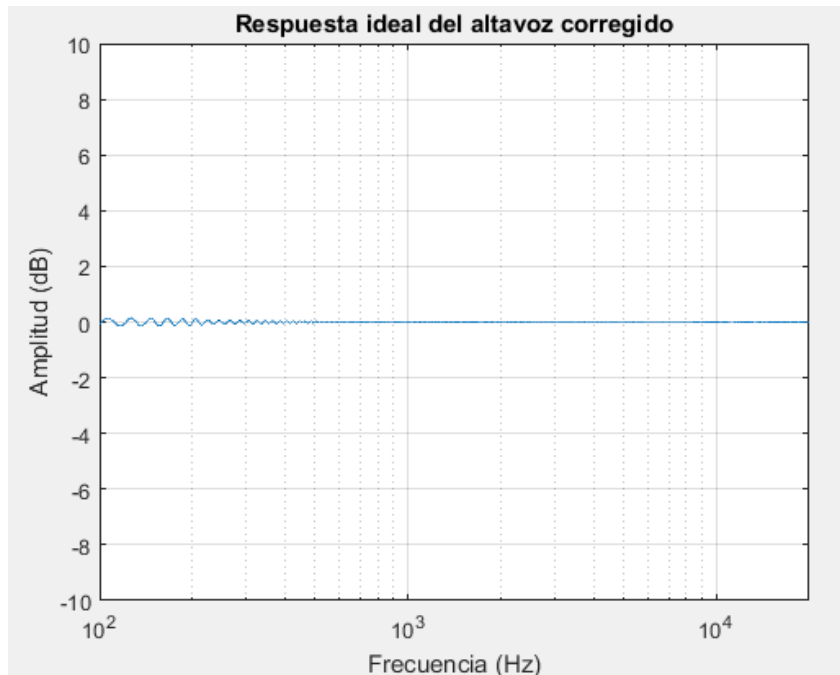


Fig. 38. Respuesta en frecuencia ideal del altavoz a 0°

El perfil obtenido es completamente plano a excepción de un ligero rizado (de ± 0.15 dB) en baja frecuencia, fruto de la baja resolución en este rango de frecuencias, ya que, al establecer una resolución de 20 Hz, las frecuencias no múltiplos de este valor sufren comportamientos que escapan del procesado realizado, dando lugar a dichos rizados.

4.3. Procesado de las medidas

Una vez se dispuso de todas las medidas almacenadas y renombradas mediante el *script* “save_measurements” (anexo I), se prosiguió con su procesamiento para corregirlas, normalizarlas y guardarlas.

4.3.1. Corrección de las medidas con la respuesta inversa del altavoz

Para aplicar la corrección de los altavoces comentada en el punto 4.1, es necesario convolucionar la respuesta al impulso inversa del altavoz (que es la salida de la función de “irsubbanda”) con la medida realizada con el maniquí binaural. Esto se realiza en la función “conv_max_sp_brir” (anexo II).

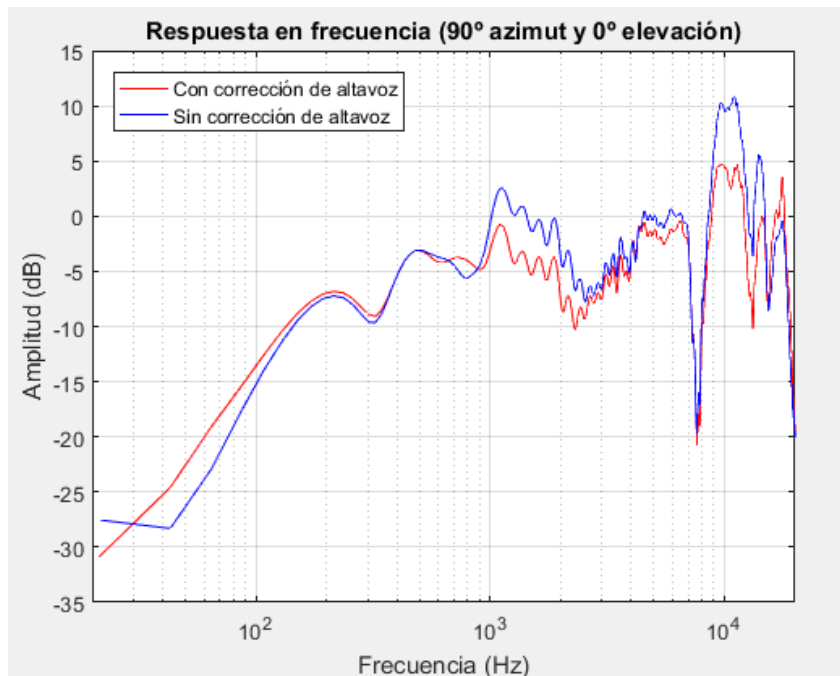


Fig. 39. Respuestas en frecuencia de la cabeza *Neumann* aplicando un offset a la no corregida para visualizar de forma óptima las diferencias en el espectro

En la figura 39 se puede apreciar el cambio sufrido en la respuesta en frecuencia con el altavoz sin corregir y corregido. Estos resultados se discutirán más adelante.

4.3.2. Normalización y almacenamiento en formato *.wav

Con todas las medidas corregidas (para ambas cabezas), se llega al último paso del procesado, que es normalizar las medidas para almacenarlas en formato *.wav. El objetivo de guardarlo en este formato radica en la facilidad posterior para realizar la convolución con otros ficheros de audio en *.wav, puesto que ambos archivos estarían en el mismo formato con la misma escala (± 1). Sabida la escala del formato, dividiendo las medidas entre el máximo de la convolución entre la respuesta al impulso inversa del altavoz correspondiente y las medidas binaurales para cada maniquí, obtendremos las mismas ya normalizadas entre ± 1 , con lo cual se puede guardar en formato *.wav aprovechando el máximo margen dinámico que proporciona. El cálculo del máximo se realiza en la función “conv_max_sp_brir” (anexo II) y la normalización y guardado en *.wav en “norm_wav_sp_brir” (anexo III).



Fig. 40. Respuesta al impulso corregida y sin normalizar

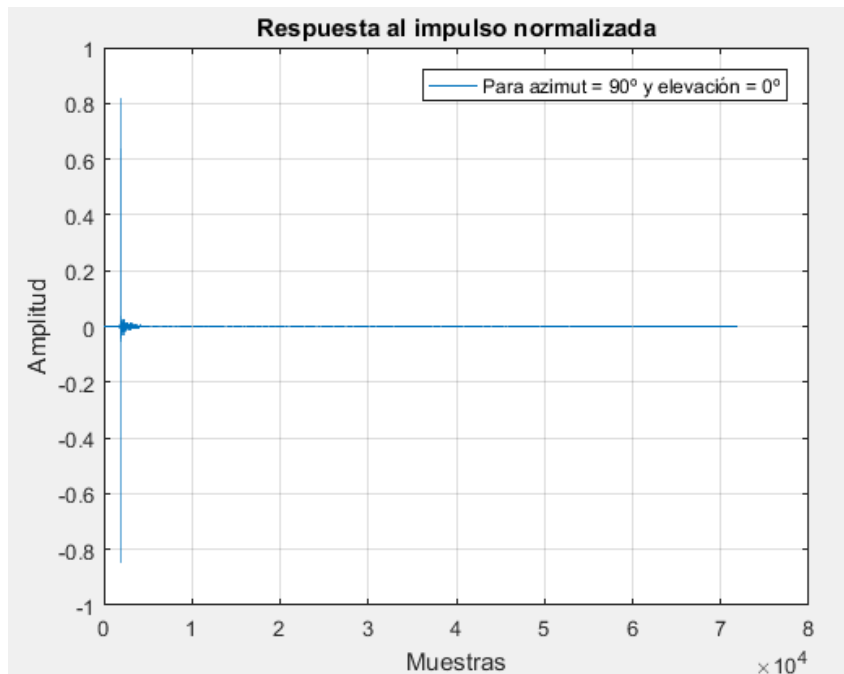


Fig. 41. Respuesta al impulso corregida y normalizada

Se observa en la figura 40 que la respuesta al impulso sin normalizar apenas alcanza los valores de ± 0.25 , mientras que al aplicar el proceso de normalización a ± 1 (fig. 41) los valores ya se acercan más a los esperados.

5. Resultados

Los resultados que se desprenden de este estudio son en sí mismo las medidas realizadas y procesadas, sobre las cuales se pueden discutir varios aspectos.

En primer lugar, se puede comparar una medida realizada a 90° de azimut y 0° de elevación (según el sistema de referencia del maniquí acústico).

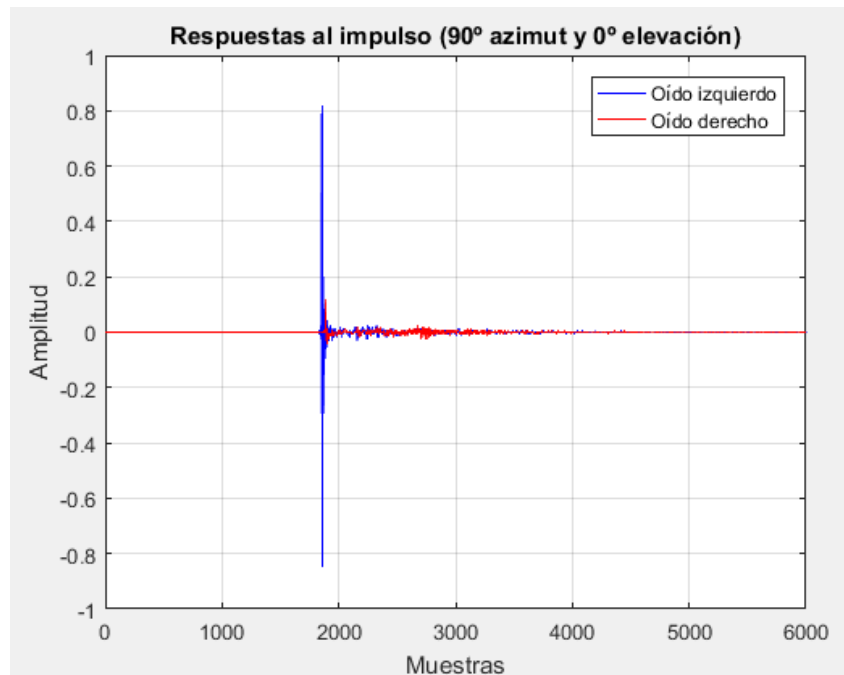


Fig. 42. Respuestas al impulso de los dos oídos del maniquí *Neumann*

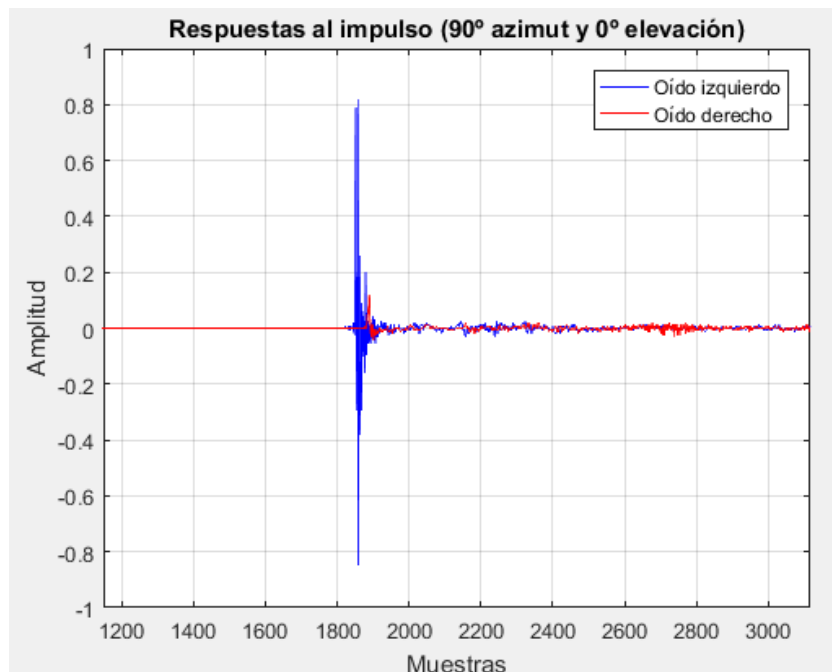


Fig. 43. Respuestas al impulso de los dos oídos ampliadas del maniquí *Neumann*

A simple vista (fig. 42), sin aumentar la representación, es complicado ver algún efecto, pero si se presta atención a la vista ampliada (fig. 43) podemos advertir dos efectos.

El primero de ellos es el retardo sufrido por la oreja contralateral (la situada en el lado opuesto a la fuente sonora) respecto a la oreja ipsilateral (la situada más cerca de la fuente

sonora). Este desplazamiento, que corresponde aproximadamente con 0.8 milisegundos para una frecuencia de muestreo de 48000 Hz, es el efecto conocido como ITD (punto 1.1.1, fig. 2). Pese a ser una diferencia de tiempos muy pequeña, el sistema auditivo es capaz de localizar perfectamente la fuente en el lado izquierdo de la cabeza.

El segundo de los efectos que se observan es la diferencia de niveles entre los oídos, lo que se conoce como IID (punto 1.1.1, fig. 5). Se puede observar que la amplitud del oído ipsilateral es más de cuatro veces mayor que la del oído contralateral (fig. 43). Este efecto también es empleado para la localización de la fuente.

En segundo lugar, se puede comparar el espectro de uno de los oídos obtenido para cada maniquí (fig. 44).

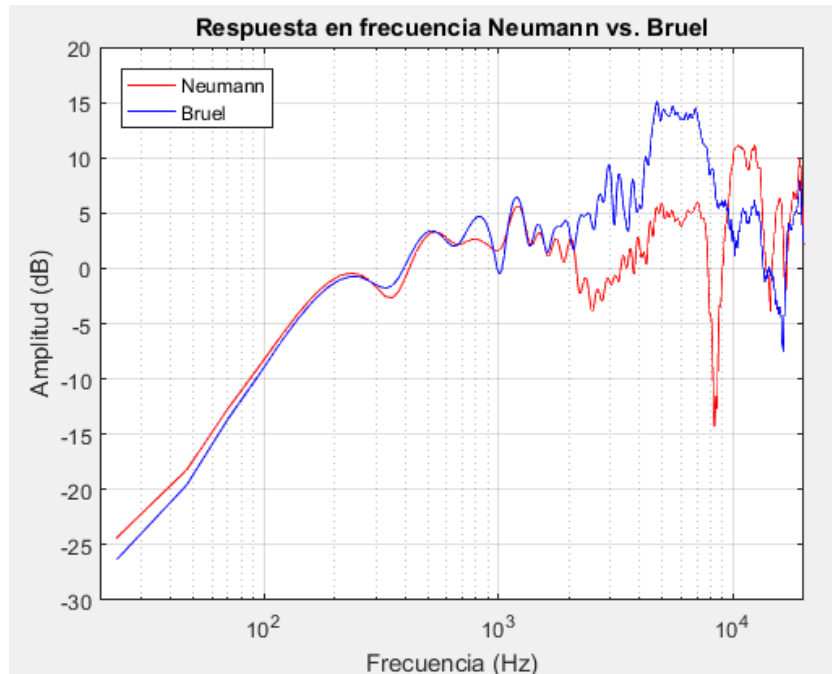


Fig. 44. Respuestas en frecuencia de la cabeza *Neumann* (rojo) y *Bruel* (azul) para 90° de azimut y 0° de elevación

En este caso, a simple vista sí que se pueden apreciar grandes diferencias. Una de ellas, que puede llamar más la atención, es la caída de nivel de la cabeza *Neumann* alrededor de los 7500 Hz respecto de la *Bruel*. Este hecho se debe principalmente a las diferencias entre la *pinna* de cada cabeza y al efecto que introduce el torso que tiene el maniquí *Bruel*, ya que la cabeza *Neumann* no cuenta con torso. En general, en alta frecuencia ambas cabezas se diferencian relevantemente debido a las propiedades físicas de la cabeza, sobre todo de la *pinna*.

No obstante, puede llamar la atención que a baja frecuencia sean muy parecidas. Esto es debido a que, en la escucha binaural en dicho rango frecuencial, el efecto de la *pinna* y en definitiva de todos los elementos físicos (torso, cabeza, canal auditivo...) que pueden afectar al espectro es mínimo, ya que las longitudes de onda son tan grandes que las ondas no se ven alteradas.

Por último, se puede realizar una comparación entre una medida con la respuesta del altavoz corregida y sin corregir. Para ello se sigue empleado la medida realizada a 90° de azimut y 0° de elevación.

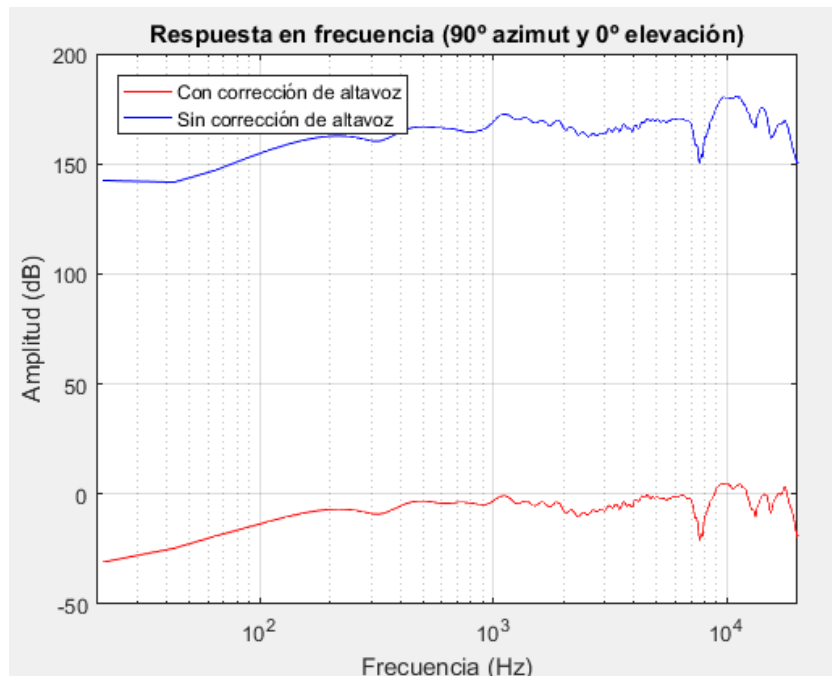


Fig. 45. Respuestas en frecuencia de la cabeza Neumann con (rojo) y sin corrección (azul)

En la fig. 45 se puede observar que el filtro ha realizado la inversa correctamente, puesto que ha dejado la respuesta resultante alrededor de 0 dB, que era el objetivo, junto con la corrección de las irregularidades del altavoz, previamente comprobado en la figura 38.

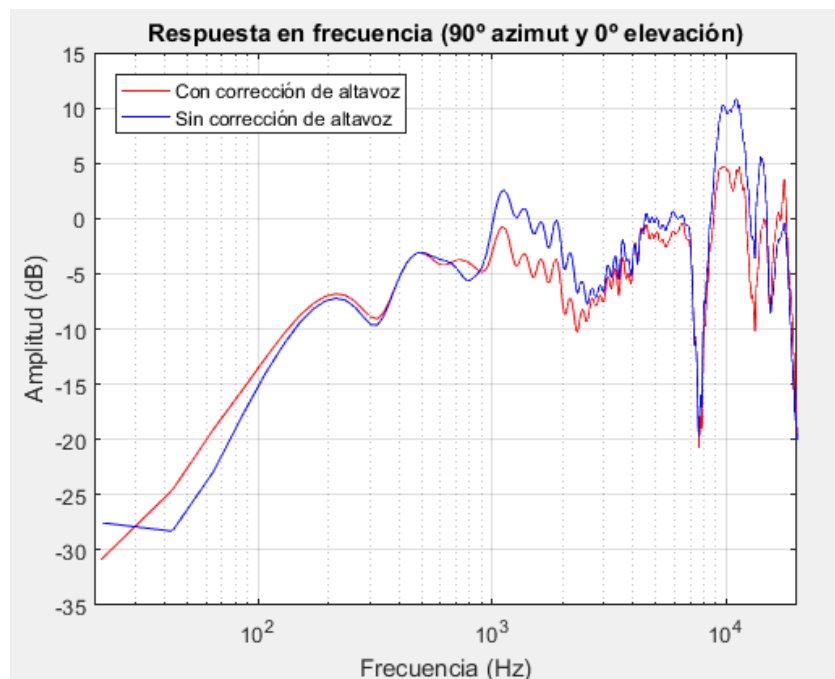


Fig. 46. Respuestas en frecuencia de la cabeza Neumann aplicando un offset a la no corregida para visualizar de forma óptima las diferencias en el espectro

Después de aplicar un *offset* a la señal sin corregir, podemos apreciar de mejor forma las diferencias en el espectro. Vemos que la corrección está actuando si se presta atención a la zona de alta frecuencia, entre 1000 Hz y 17000 Hz, donde se han compactado más los niveles en un rango alrededor de -10 dB a 5 dB, evitando grandes variaciones (a excepción de la caída de nivel alrededor de 7500 Hz a causa de la propia cabeza binaural).

6. Conclusión

En este trabajo de fin de grado se han explicado los conceptos fundamentales de la escucha binaural y se han medido dos maniqués acústicos. Para ello, ha sido necesario un montaje de altavoces y una estructura de soporte de los altavoces para las medidas en elevación y una plataforma giratoria para las medidas en azimut.

Con ello se han obtenido 720 medidas o 360 pares de medidas binaurales de tipo BRIR para cada maniqué binaural. Estas medidas han sido procesadas en diversos momentos del trabajo para: el renombre de las medidas, realizar las parejas de canales, compensar la respuesta de los altavoces y normalizar todas las medidas para guardarlas en formato *.wav, para, finalmente, poder utilizar dichos resultados para convolucionar las medidas finales con cualquier tipo de sonido, consiguiendo así la binauralización del sonido.

Para finalizar, los resultados han sido discutidos, valorando las diferencias entre las medidas realizadas con los distintos procesamientos. Con ello se ha logrado comprobar los fundamentos explicados de la audición binaural, así como exponer un procedimiento de medida válido para cualquier persona que esté interesada en adentrarse en el campo del sonido binaural.

7. Referencias

- [1] A. R. Moller, *Hearing: Anatomy, physiology, and disorders of the auditory system*, Plural Publishing Inc., 2012.
- [2] J. W. Strutt, «On our perception of the direction of sound,» *Philos. Magazine*, vol. 13, nº XII, pp. 214-232, 1907.
- [3] C. I. Cheng y G. H. Wakefield, «Introduction to head related transfer functions (HRTFs): Representations of HRTFs in Time, Frequency and Space,» de *AES 107th convention*, New York, 1999.
- [4] M. Cohen y J. Villegas, *Fundamentals of Wearable Computers and Augmented Reality*, CRC Press, 2000, pp. 309-330.
- [5] F. M. Weiner y D. A. Ross, «The Pressure Distribution in the Auditory Canal in a Progressive Sound Field,» *The Journal of the Acoustical Society of America*, vol. 18, nº 2, pp. 401-408, 1946.
- [6] J. Blauert, «Sound Localization in the Median Plane,» *Acta acustica*, vol. 22, nº 4, 1969.
- [7] D. H. Mershon y L. E. King, «Intensity and reverberation as factors in the auditory perception of egocentric distance,» *Perception & Psychophysics*, vol. 18, nº 6, pp. 409-4015, 1975.
- [8] A. J. Kolarik, B. C. J. Moore, P. Zahorik, S. Cirstea y S. Pardhan, «Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss,» *Atten Percept Psychophys*, vol. 78, pp. 373-395, 2016.
- [9] J. Blauert, *Spatial hearing: The psychophysics of human sound localization*, Cambridge, MA: Massachusetts Institute of Technology, 1997.
- [10] B. DS, «Auditory localization of nearby sources III. Stimulus effects,» *The Journal of the Acoustical Society of America*, vol. 106, pp. 3589-3602, 1999.
- [11] A. Farina, «Simultaneous measurement of impulse response and distortion with a swept sine technique,» de *108th AES Convention*, Paris, 2000.
- [12] A. Farina, «Advancements in impulse response measurements by sine sweeps,» de *122ed AES Convention*, Vienna, 2007.
- [13] Hugeng, D. Gunawan y W. Wahab, «Effective Preprocessing in Modeling Head-Related Impulse Responses Based on Principal Components Analysis,» *Signal Processing: An International Journal (SPIJ)*, vol. 4, nº 4, pp. 201-212, 2010.
- [14] H. Lehnert y J. Blauert, «Principles of binaural room simulation,» *Applied Acoustics*, vol. 36, nº 3-4, pp. 259-291, 1992.
- [15] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann y B. Kollmeier, «Database of Multichannel In-Ear and Behind-the-Ear Head-Related and Binaural Room Impulse Responses,» *EURASIP Journal on Advances in Signal Processing*, vol. 2009, 2009.
- [16] A. S. Incorporated, «Adobe Audition,» <http://www.adobe.com/es/products/audition.html>.
- [17] Mathworks, «Matlab,» <https://es.mathworks.com/products/matlab.html>.

Anexo I

Script "save_measurements"

```
% Renombra los archivos .txt generados por MegaIR.exe,
% para 10 posiciones en elevación (-20° a 70°) y 36 en azimut (0° a
350°),
% con el siguiente formato:
% L_az_[ángulo de azimut]_el_[ángulo de elevación] y
% R_az_[ángulo de azimut]_el_[ángulo de elevación]

% az = azimut
% el = elevation

fs = 48000;
az = 0;

head = input('Escribe N para la cabeza Neumann o B para la Bruel:
','s');

% Creación de directorios
if head == 'N'
    mkdir Neumann_txt;
else
    mkdir Bruel_txt;
end

while az <= 350

    display(['Pulsa cualquier tecla para renombrar azimut = ',...
            num2str((az)), ' grados.']);
    pause;

    for el = 1:10

        % Carga de ficheros
        dataL = load(strcat('c_1_',num2str(el),'.txt'));
        dataR = load(strcat('c_2_',num2str(el),'.txt'));

        % Cambio de directorio
        if head == 'N'
            cd Neumann_txt;
        else
            cd Bruel_txt;
        end

        % Renombre de los archivos
        L_newName = strcat('L_az_',num2str(az),'_el_',...
                            num2str(80-el*10),'.txt');

        fid = fopen(L_newName,'w');
        fprintf(fid,'%g\n',dataL);
        fclose(fid);

        R_newName = strcat('R_az_',num2str(az),'_el_',...
                            num2str(80-el*10),'.txt');

        fid = fopen(R_newName,'w');
```

```
fprintf(fid, '%g\n', dataR);  
fclose(fid);  
  
cd ..  
  
% Borrado de los archivos generados por MegaIR.exe  
delete(strcat('c_1_', num2str(e1), '.txt'));  
delete(strcat('c_2_', num2str(e1), '.txt'));  
end  
az = az + 10;  
end
```

Anexo II

Función “conv_max_sp_brir”

```
% Función que calcula el valor máximo de la convolución entre la IR
% inversa del altavoz y las BRIR medidas
% az = azimuth
% el = elevation
% ** SALIDAS ** %
% brir_sp contiene una medida binaural cada dos columnas, siendo los
10
% primeros pares las medidas correspondientes a az = 0 y el = 70:-10:-
20,
% siguiendo en los siguientes pares de columnas con az =
10,20,30...350.

% maxValue contiene el valor máximo encontrado entre todas las
% convoluciones

function [maxValue, brir_sp] = conv_max_sp_brir

head = input('Escribe N para la cabeza Neumann o B para la Bruel:
','s');

aux = 0;
maxValue = 0;
i = 1;

for az = 0:10:350

    for el = 70:-10:-20

        if head == 'N'
            cd Neumann_txt;
        else
            cd Bruel_txt;
        end

        dataL =
load(strcat('L_az_',num2str(az),'_el_',num2str(el),'.txt'));
        dataR =
load(strcat('R_az_',num2str(az),'_el_',num2str(el),'.txt'));

        cd ../sp_IR_05

        sp_IR = load(strcat('el_',num2str(el),'.txt'));

        cd ..

        % irsubbanda devuelve la IR inversa recortada a 2400 muestras
iir = irsubbanda(sp_IR);

        brir_sp_L = conv(iir,dataL);
        brir_sp_R = conv(iir,dataR);

        % Recorte a la más larga
L = length(dataL);
```

```

brir_sp_L = brir_sp_L(1:L);
brir_sp_R = brir_sp_R(1:L);

brir_sp(:,i:i+1) = [brir_sp_L, brir_sp_R];

i = i+2;

display(['Buscando máximo para az_', num2str(az), ...
        '_el_', num2str(e1)]);
display(['Azimut = ', num2str(az), ...
        'corresponde con al columna = ', num2str(i)]);

aux = max(max(abs(brir_sp)));

if aux > maxValue
    maxValue = aux;
end

end % for e1
end % for az

display(['El valor máximo es: ', num2str(maxValue)]);

```

Anexo III

Función "norm_wav_sp_brir"

```
% Función que normaliza el contenido de brir_sp con el máximo valor
% encontrado por la función "conv_max_sp_brir" y guarda los datos en
% formato *.wav en un nuevo directorio.
% az = azimut
% el = elevation

function norm_wav_sp_brir(maxValue,brir_sp)

head = input('Escribe N para la cabeza Neumann o B para la Bruel:
','s');

fs = 48000;
i = 1;

if head == 'N'
    mkdir Neumann_sp_brir_wav;
else
    mkdir Bruel_sp_brir_wav;
end

if head == 'N'
    cd Neumann_sp_brir_wav;
else
    cd Bruel_sp_brir_wav;
end

for az = 0:10:350

    for el = 70:-10:-20

        display(['Normalizando y guardando en .wav...
az_',num2str(az),...
                '_el_',num2str(el)]);

        norm_data = brir_sp(:,i:i+1)./maxValue; % Normalización +/- 1
        newName = strcat('az_',num2str(az),'_el_',num2str(el),'.wav');
        audiowrite(newName,norm_data,fs);

        i = i+2;

    end % for el
end % for az

cd ..
```