

Aumentación de Respuestas al Impulso

Las bases de datos de respuestas al impulso reales generalmente son pocas y no logran una buena cobertura de los parámetros acústicos como el T60 y el DRR. Es posible, mediante técnicas de procesamiento de señales, alterar los parámetros acústicos de una respuesta al impulso, generando otras con distintas características acústicas. De esta forma, se puede ampliar considerablemente la cantidad de respuestas al impulso del conjunto de datos, a partir de alteraciones a un conjunto pequeño de respuestas al impulso reales, logrando una mejor cobertura de los parámetros acústicos de interés. En este trabajo, se utilizan dos procesos para aumentar las respuestas al impulso reales: una alteración de amplitud en la parte temprana de la respuesta al impulso para controlar la relación directo-reverberado, y una alteración de envolvente de caída para controlar el tiempo de reverberación.

1.1 Aumentación de Relación Directo-Reverberado (DRR)

Para el primer proceso, a la parte correspondiente al sonido directo $h_e(t)$ ¹ se le aplica una ganancia definida por un factor α el cual se calcula para obtener el valor de *DRR* deseado generando una nueva señal $\tilde{h}_e(t)$. Para evitar generar discontinuidades durante el proceso, se aplican ventanas complementarias a la señal directa obteniendo una señal directa ventaneada y un residuo ventaneado. A partir de esto, la parte directa se puede definir según la ecuación 1.

$$h_e(t) = \alpha w_d(t) h_e(t) + [1 - w_d(t)] h_e(t) \quad (1)$$

En donde $w_d(t)$ corresponde a una ventana Hann de 5ms de longitud. De esta manera, partiendo de esta última definición junto con la expresión del parámetro *DRR* expresado en la ecuación ??, se plantea un sistema de ecuaciones a partir del cual se puede definir un valor pretendido de *DRR* y despejar el correspondiente valor de α . En la Figura 1 se puede observar una representación de una parte directa $h_e(t)$, las ventanas aplicadas, el efecto del factor de ganancia α y la nueva señal $\tilde{h}_e(t)$ generada. Finalmente, esta parte directa modificada se concatena con el resto de la respuesta al impulso completando así el proceso de aumentación referido a la relación directo-reverberado. Se debe tener en cuenta que para tiempos de reverberación cortos, puede ser un problema generar relaciones directo-reverberado demasiado bajas, ya que la energía de la parte tardía ya es de por sí muy baja. Para esos casos, se definen valores límites para la ganancia aplicada a la parte temprana.

¹Comúnmente se considera que los primeros 2,5 ms corresponden al sonido directo. Esto representa una diferencia de camino de 0,85 m para una velocidad del sonido de 340 ms⁻¹. Las normas de medición de respuestas al impulso establecen condiciones de posición de manera tal que la primera reflexión ocurre luego de los 2,5 ms

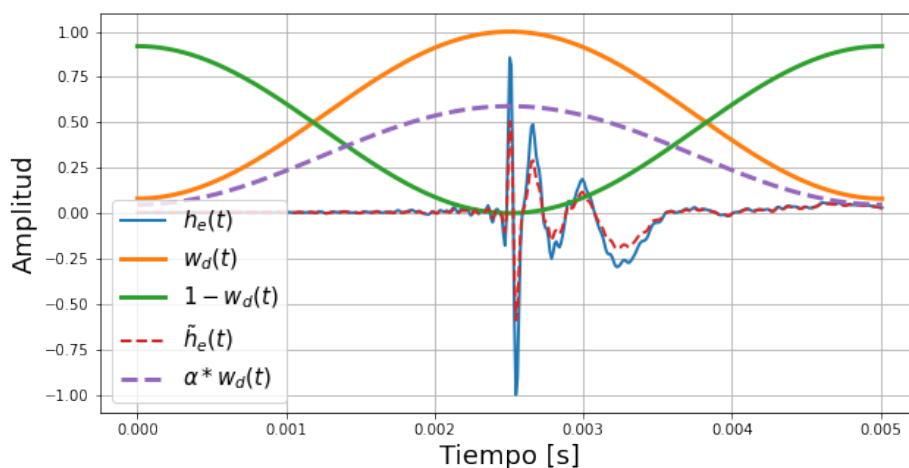


Figura 1: Señales involucradas en el proceso de aumentación de DRR.

1.2 Aumentación de Tiempo de reverberación

El proceso de aumentación de respuestas al impulso puede realizarse tomando como parámetros del proceso ciertos descriptores acústicos como el tiempo de reverberación TR_{60} . Esto es, partiendo de una respuesta al impulso con un TR_{60} determinado, se busca generar nuevas respuestas al impulso cuyo TR_{60} pueda ser controlado paramétricamente para ocupar de manera homogénea y balanceada un rango de valores de interés. El TR_{60} se relaciona directamente con la forma de la envolvente de caída de nivel exponencial presente en la parte tardía de las respuestas al impulso. El proceso de aumentación equivale a modificar esta pendiente de caída, multiplicando la señal original por una nueva envolvente que produzca el efecto deseado en la envolvente resultante. Los pasos a seguir para realizar este proceso son:

- Acondicionamiento de la respuesta al impulso de entrada.
- Filtrado por bandas de octava o bandas de tercio de octava.
- Estimación de piso de ruido.
- Estimación de envolvente de caída.
- Sintetizar una señal aplicando la envolvente estimada con piso de ruido cero a una señal de ruido Gaussiano.
- Realizar el cross-fade entre la señal sintetizada y la señal original en el punto inicial del piso de ruido.
- Aumentación de la envolvente de caída multiplicando la señal por la correspondiente envolvente exponencial creciente/decreciente.
- Suma de las sub-bandas para obtener la señal resultante en su espectro completo.

- Integración de la parte tardía aumentada con la parte directa de la respuesta al impulso inicial.

A continuación se explica cada paso del algoritmo en mayor profundidad. En primer lugar, se define una determinada frecuencia de muestreo y profundidad de bits para trabajar con la señal de entrada, en este caso una respuesta al impulso real. Una vez asegurada la homogeneidad de estas características, la señal se normaliza para trabajar en un rango de amplitud acotado en el intervalo $[-1, 1]$ y luego se separa la parte directa de la parte tardía de la respuesta al impulso. Esto último se realiza utilizando una ventana de tolerancia de $t_0 = 2,5ms$. Para los pasos siguientes se trabaja únicamente modificando la parte tardía, y la parte directa se almacena para ser utilizada en el paso final a la hora de reconstruir la respuesta completa. En la Figura 2 se muestra la respuesta al impulso desde la que se parte, distinguiendo la descomposición temporal de la misma y luego la parte tardía de la respuesta al impulso aislada con la que se va a trabajar durante el proceso.

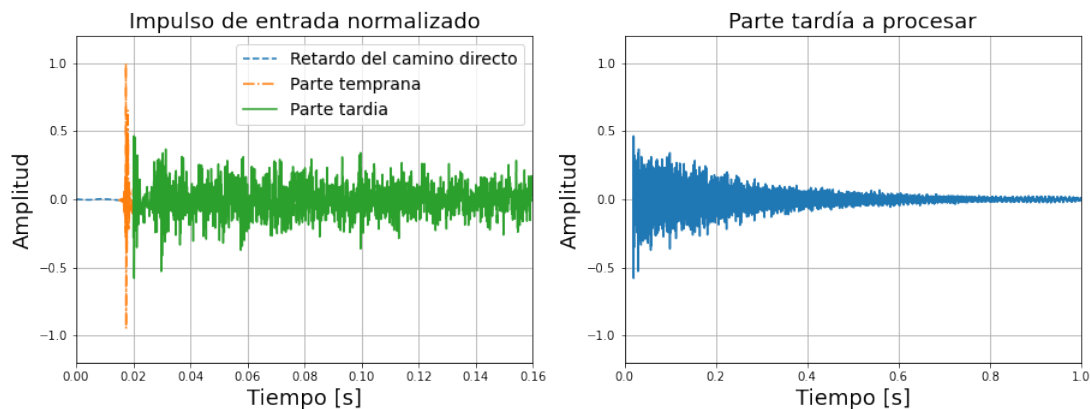


Figura 2: Descomposición temporal de la respuesta al impulso a procesar durante la aumentación.

Luego, el siguiente paso consiste en descomponer la señal en bandas de octava o tercios de octava. En esta demostración se trabaja con bandas de octava desde 125 Hz hasta 4000 Hz teniendo en cuenta que se utiliza una frecuencia de muestreo de 16000 Hz . Es necesario trabajar en sub-bandas frecuenciales para contemplar la dependencia del tiempo de reverberación con la frecuencia, y mantener esa característica en las señales a generar en este proceso. Para conseguir esta descomposición en sub-bandas frecuenciales se crea un banco de filtros. El mismo se compone de filtros Butterworth que van siendo creados a partir de las frecuencias centrales que se quiera obtener en cada banda. El proceso consiste en crear filtros pasa-banda para generar una banda de paso alta y una banda de paso baja. Luego, se toma la banda de paso alta y se la vuelve a dividir en banda de paso alta y baja aplicando un nuevo par de filtros pasa-banda. Esto se repite hasta completar todas las frecuencias de corte necesarias. De esta forma se obtiene el prototipo IIR de cada filtro que compone el banco de filtros. Una vez obtenido esto, se crean respuestas de tipo FIR para cada filtro a través de pasar un impulso ideal por cada filtro. En la Figura 3 se puede observar la respuesta en frecuencia del banco de filtros. Un detalle importante a considerar es que la suma del efecto de todos los filtros es unitaria para todo el rango frecuencial analizado, siendo esta una característica necesaria para poder descomponer

la señal en bandas y luego re-componer la señal sumando las bandas sin generar ningún tipo de distorsión en el proceso. Para lograr esto, los coeficientes de los filtros deben ser elevados al cuadrado (lo que equivale a poner dos filtros en cascada) para lograr que en las intersecciones entre los filtros la suma de amplitud sea unitaria (en la frecuencia de corte se genera una caída de 6 dB en lugar de los 3 dB que presentaría un filtro Butterworth simple).

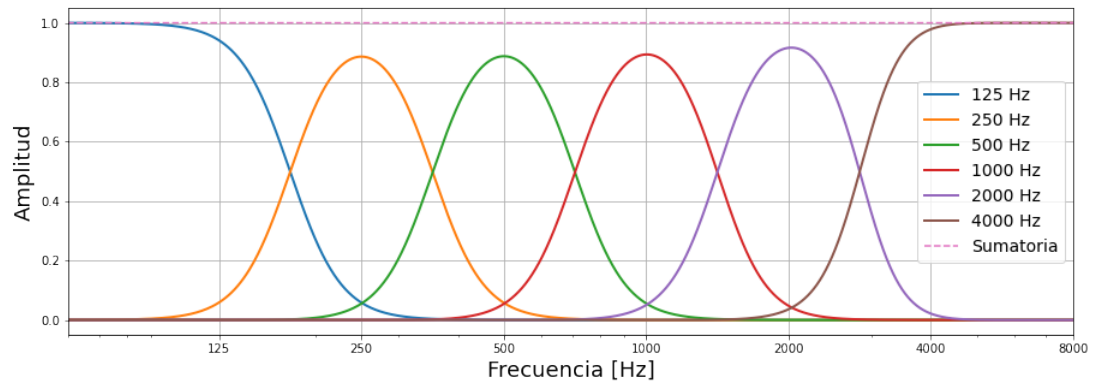


Figura 3: Banco de filtros Butterworth.

A partir de aplicar este banco de filtros, la señal de entrada se descompone en 6 sub-bandas como se muestra en la Figura 4. En este gráfico se puede apreciar como la pendiente de caída varía según la banda de frecuencia que se observa. Todos los procesos subsiguientes se aplican de manera independiente para cada banda, y luego al final las bandas se suman para volver a tener una señal correspondiente al espectro completo original. De ahora en adelante, por una cuestión de simplicidad se muestran los gráficos pertenecientes a la banda de 1000 Hz de manera ilustrativa.

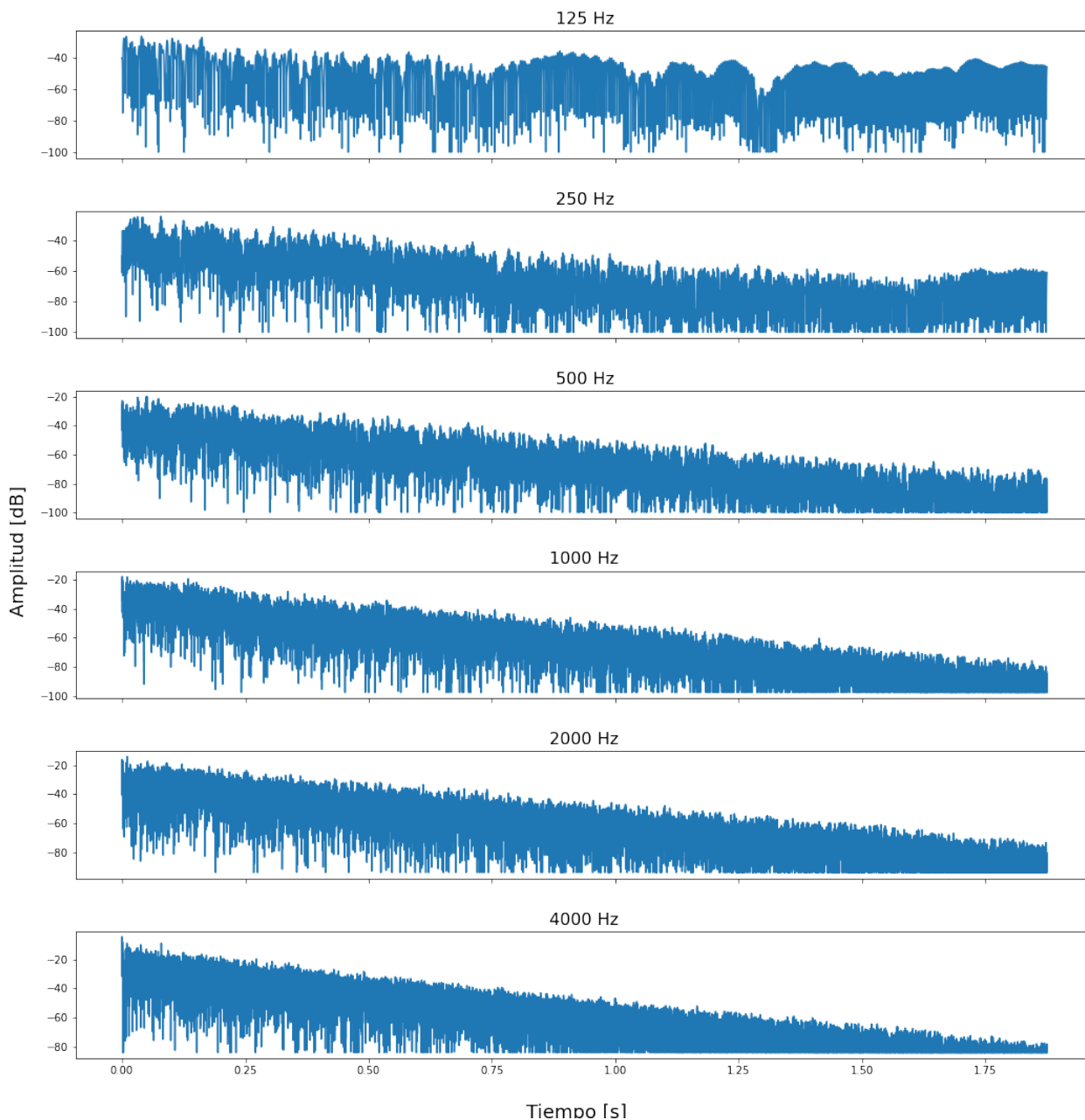


Figura 4: Sub-bandas obtenidas luego de aplicar el banco de filtros.

El paso siguiente consiste en determinar el piso de ruido de la señal. Detectar el piso de ruido permite asegurar que el método de aumentación no amplifique ruido cuando se busca obtener un tiempo de reverberación mayor al inicial propio de la respuesta al impulso de entrada. Para determinar el punto donde predomina el ruido en la respuesta al impulso se utiliza el método iterativo de Lundeby **Lundeby**. El mismo consta de los siguientes pasos:

1. La respuesta al impulso al cuadrado es promediada en intervalos de tiempo locales de entre 10 ms y 50 ms para obtener una curva 'suavizada', es decir, disminuir las variaciones instantáneas sin perder las pendientes cortas.

2. Se hace una primera estimación del piso de ruido. Para hacerlo se toma el segmento correspondiente al último 10% de la respuesta al impulso.
3. La pendiente de caída se estima aplicando una regresión lineal sobre el intervalo de tiempo que contiene la respuesta entre el pico de 0 dB y el primer intervalo 5 – 10 dB por encima del ruido de fondo.
4. Se determina un punto de cruce provisorio en la intersección entre la pendiente de caída estimada y el nivel de piso de ruido.
5. Se obtiene un nuevo intervalo de tiempo de acuerdo a la pendiente calculada, de manera que haya entre 3 y 10 intervalos por cada 10 dB de caída.
6. Se vuelve a promediar localmente el impulso al cuadrado de acuerdo al nuevo intervalo temporal calculado previamente .
7. Se estima el ruido de fondo nuevamente. El segmento a evaluar debe corresponder a 5 – 10 dB luego del punto de cruce (siguiendo la curva estimada previamente), o bien, un mínimo del 10% de la señal total (en el caso de tener que optar por el 10% de nuevo, el resultado sería el mismo que antes, y el punto encontrado previamente sería el definitivo).
8. Se estima la pendiente de caída para un rango dinámico de entre 20 dB y 10 dB, empezando desde un punto 5 – 10 dB por encima del nivel de ruido.
9. Se encuentra un nuevo punto de cruce.
10. Los pasos 7-9 se repiten hasta que el valor del piso de ruido converja, tolerando un máximo de 6 iteraciones.

El paso siguiente es estimar la pendiente paramétrica que mejor se aproxime a la pendiente de caída real. La estimación se basa en el modelo de la ecuación ???. Por lo tanto, los parámetros que se busca estimar son la amplitud inicial, la tasa de caída y el nivel de piso de ruido. La estimación se realiza aplicando un algoritmo de ajuste no lineal por cuadrados mínimos. El resultado de la estimación para la banda de 1000 Hz se muestra en la Figura 5.

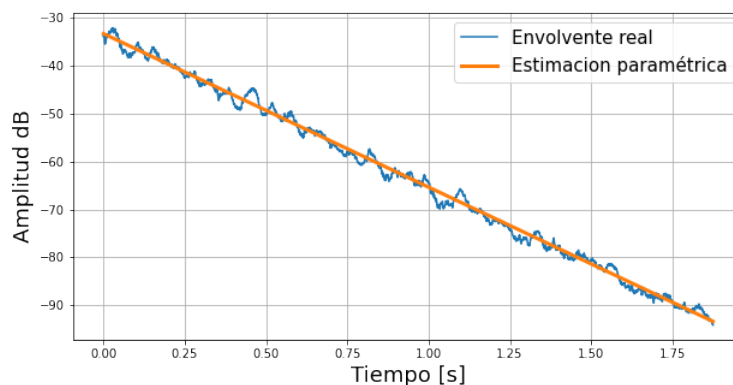


Figura 5: Estimación paramétrica de la pendiente de caída.

Con estos parámetros estimados se genera una nueva envolvente de caída pero llevando el nivel de piso de ruido a cero, y se aplica esta envolvente sobre una señal de ruido Gaussiano de media cero y desvío estándar unitario. Con esta señal sintética y la señal original se hace una transición cruzada en el punto estimado del piso de ruido. De esta manera se elimina el ruido de la señal original, reemplazándolo por la caída exponencial determinada por la envolvente paramétrica. En la Figura 6 se puede observar como se extiende la respuesta luego del punto de inicio del piso de ruido de la señal original.

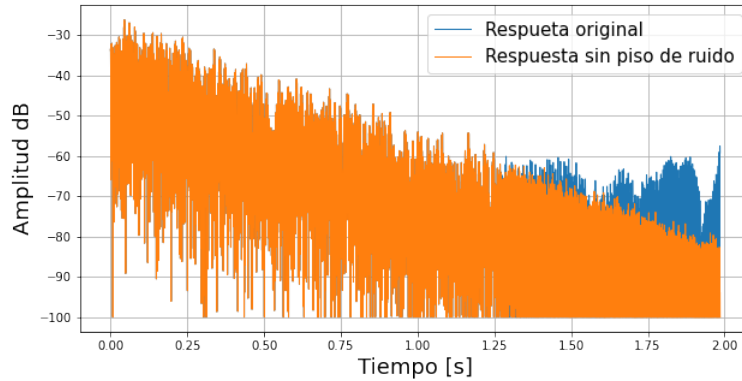


Figura 6: Respuesta original y extendida sin piso de ruido.

Finalmente, teniendo las bandas extendidas y habiéndose eliminado el piso de ruido, se prosigue con el proceso de aumentación del tiempo de reverberación para cada sub-banda, multiplicando cada respuesta por la correspondiente envolvente exponencial creciente o decreciente según corresponda, para obtener la envolvente de caída necesaria para generar el tiempo de reverberación deseado, como se indica en la ecuación ???. Un ejemplo del resultado de la aumentación sobre una sub-banda de frecuencia se muestra en la Figura 7, en donde el tiempo de reverberación objetivo es menor que el tiempo de reverberación de la respuesta original, por lo cual la pendiente aumentada resulta mas atenuada que la pendiente original.

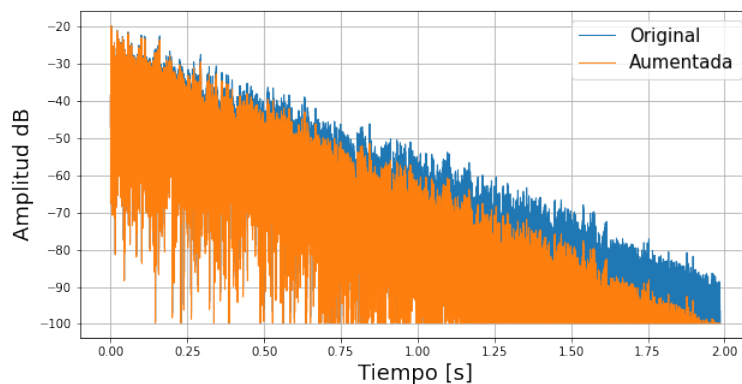


Figura 7: Aumentación del tiempo de reverberación alterando la envolvente de caída original.

Una vez realizada la alteración de la envolvente de caída para todas las bandas frecuencia-

les, estas se suman para conformar nuevamente la parte tardía de la respuesta al impulso de espectro completo. El paso final consiste en concatenar los segmentos de la respuesta al impulso original que no fueron alterados, es decir, el delay del camino directo y la parte directa de la respuesta. Con esto, el proceso de aumentación termina, y se obtiene como resultado una nueva respuesta al impulso con el tiempo de reverberación deseado. En la Figura 8 se muestra la respuesta al impulso original comparada con la nueva respuesta generada a partir del proceso anteriormente descrito.

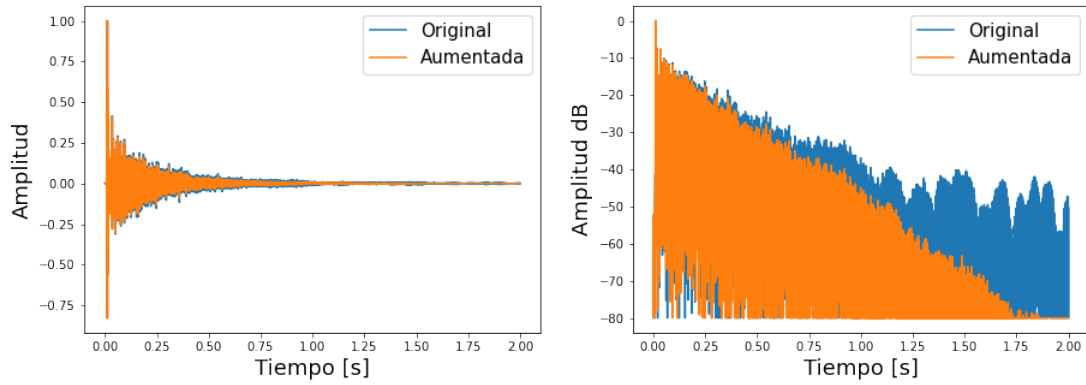


Figura 8: Resultado del proceso de aumentación del tiempo de reverberación de una respuesta al impulso.