

Procesamiento Digital de Señales de Audio

Práctico 2

Santiago García Pose

4.595.400-6

19 de abril de 2020



Índice

| | |
|--|----------|
| 1. Ejercicio 1 | 1 |
| 1.1. Parte 1 | 1 |
| 1.1.1. Señales de Ruido | 2 |
| 1.1.2. Utilizando un banco con menos bandas | 3 |
| 2. Ejercicio 2 - Síntesis de cuerda pulsada | 4 |
| 2.1. Filtro peine | 4 |
| 2.2. Filtro peine aplicando pasabajos | 5 |
| 2.3. Filtro peine aplicando pasabajos y pasatodo | 6 |
| 3. Error en frecuencia experimental | 9 |

1. Ejercicio 1

1.1. Parte 1

Para este ejercicio se trabajó con un fragmento de 8 segundos de la canción 'Superstition' de Stevie Wonder, muestreada a una frecuencia de 8000 Hz. El objetivo era lograr una señal intermedia carente de información armónica o melódica pero que preserve las características rítmicas de la señal original, tal y como se muestra en la figura 1.1. El banco de filtros utilizado fue de tipo Butterworth, se testearon distintos órdenes, notando que para órdenes muy bajos las pendientes de los cortes eran poco pronunciadas por lo que el solapamiento era un problema, y en los filtros de bandas muy angostas al utilizar órdenes altos se producían inestabilidades y ripples muy pronunciados en la respuesta en frecuencia. Luego de algunos experimentos de ensayo y error se optó por trabajar con los órdenes descritos en la figura 1.2.

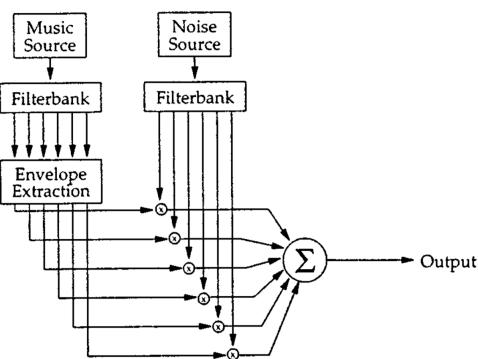


Figura 1.1: Algoritmo utilizado para el ejercicio.

El siguiente paso fue rectificar las señales resultantes de aplicar el banco de filtros, se hallaron las envelopantes de dos maneras distintas, la primera fue aplicar un filtro de media móvil de un largo de ventana de 320 muestras, equivalente a 40ms. La segunda forma fue aplicar directamente un filtro pasabajos también del tipo Butterworth de orden 6 y frecuencia de corte entre 15-20 Hz. Luego de comparar los resultados, si bien el Butterworth arrojaba un resultado con una curva de más simple y con transitorios más pronunciados, se notaba un cierto atraso en tiempo con respecto a la señal rectificada debido al componente de fase que agrega el filtro, mientras que el filtro de media móvil arrojaba una curva un tanto más sinuosa y con un pronunciamiento levemente menor de los transitorios, pero desfasaje mejor con respecto a la señal, por lo que se optó trabajar con esta última, ya que las diferencias entre amplitudes difícilmente afectaran el resultado final. Una ilustración de esta situación se puede encontrar en el Anexo (ver figura 3.1).

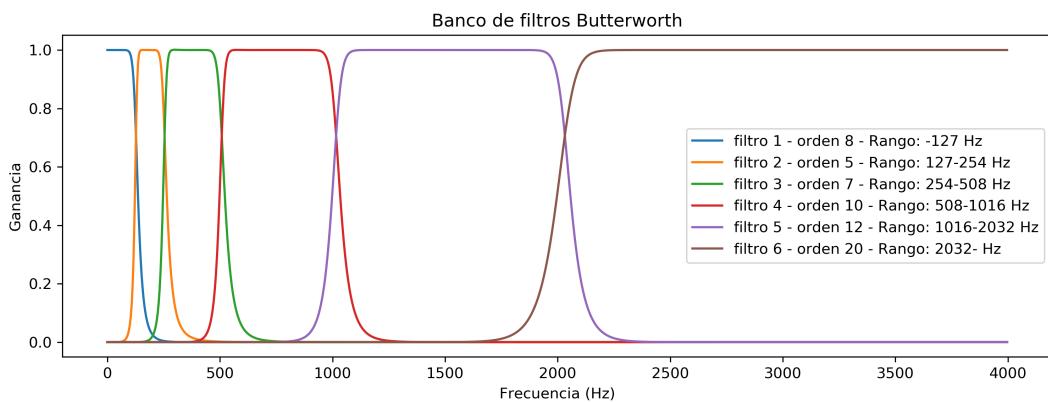


Figura 1.2: Banco de filtros Butterworth utilizado para filtrar las señales.

1.1.1. Señales de Ruido

Visto que la señal intermedia debía conservar solo las características rítmicas, se generó una señal de ruido blanco de la misma duración que el archivo original, y fue procesada por el banco de filtros, para luego aplicarle las envelopantes halladas. Al escucharlas por separado, se puede notar que las señales mas graves son las que llevan el beat de la canción y tienen transitorios mas parejos en cuanto a amplitud. Y las bandas de alturas mas agudas tienen componentes rítmicos que generan los contratiempos o síncopas. Lo cual es bastante lógico y esperable si se piensa en el rango de frecuencias fundamentales que ocupa cada instrumento del archivo original.

Como conclusión parcial, el algoritmo devuelve una señal que contrastada con la original preserva gran parte de las características rítmicas principales, por supuesto que tener en conocimiento cómo suena el audio original condiciona la escucha del resultado del algoritmo, ya que inevitablemente uno tiende a superponer la información melódica y armónica que tiene en la memoria y de alguna manera complementa 'mentalmente' lo que escucha a través del ruido. Pero una manera de verificar objetivamente podría ser mediante un espectrograma, como se ven en la figura 1.3 donde se puede constatar que hay una fuerte relación entre espectrogramas en los instantes temporales que la energía es grande, que precisamente coinciden con los momentos donde se producen los tiempos fuertes del beat.

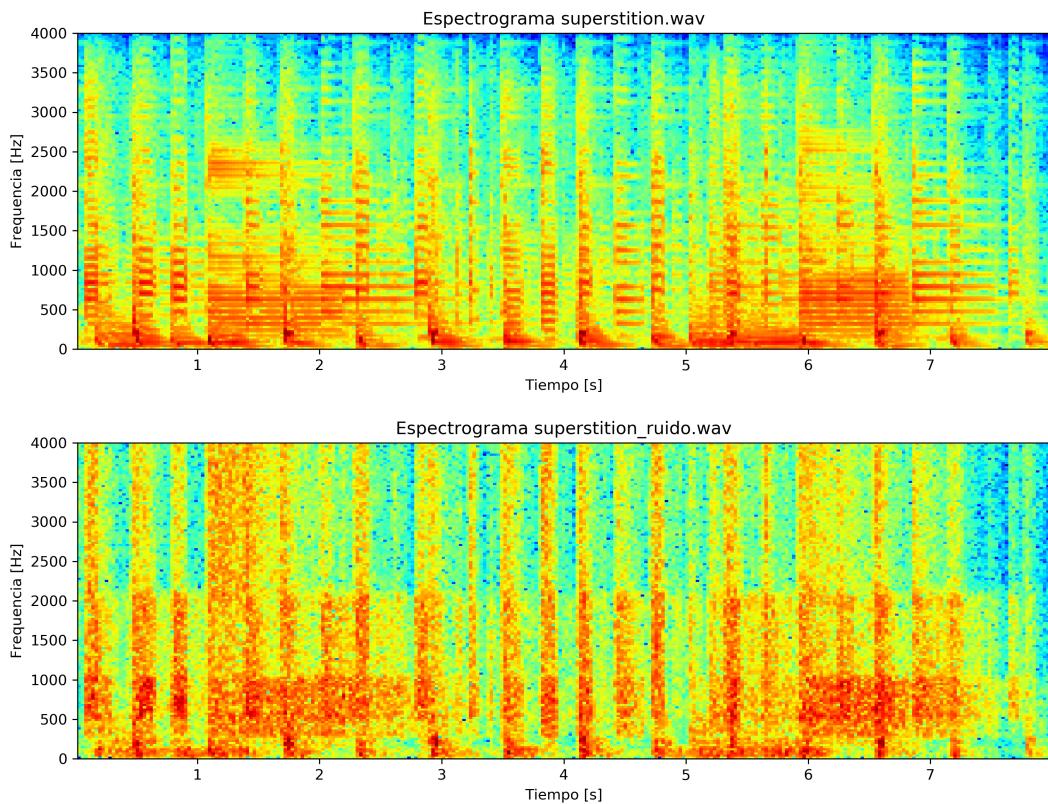


Figura 1.3: Espectrogramas de las señales original y procesada respectivamente.

1.1.2. Utilizando un banco con menos bandas

En esta parte se repitió exactamente el mismo algoritmo de la figura 1.1, pero ahora utilizando un banco de filtros como el de la figura 1.4. Los resultados obtenidos no fueron tan alejados del anterior en cuanto a lo rítmico, pero escuchando con atención se puede notar que se pierde definición en frecuencia en cuanto a transitorios, pasa a ser un audio más 'sucio' en términos de que no es tan claro en qué frecuencias se produce cada transitorio, al tener menos bandas lógicamente se discrimina menos en frecuencia a qué rango pertenece cada golpe dentro del beat. Esto también se puede observar en el espectrograma de la figura 1.5 donde se ve que hay más uniformidad en frecuencia que en el espectrograma correspondiente de la figura 1.3.

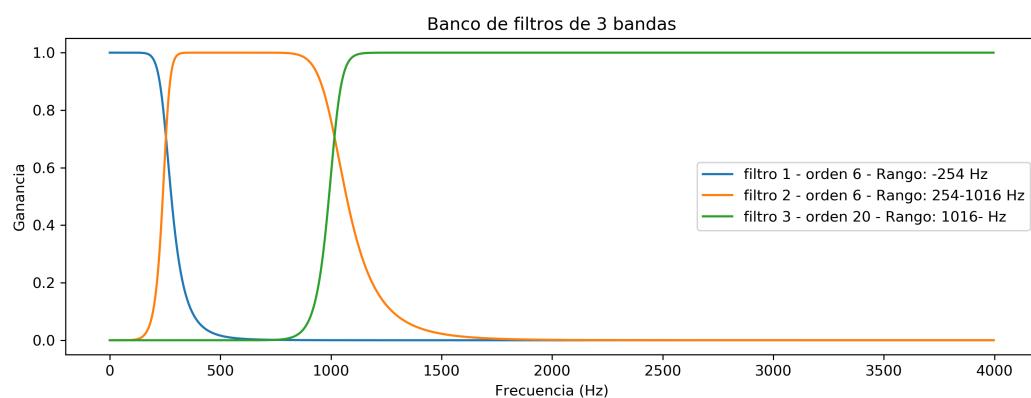


Figura 1.4: Banco de filtros de 3 bandas.

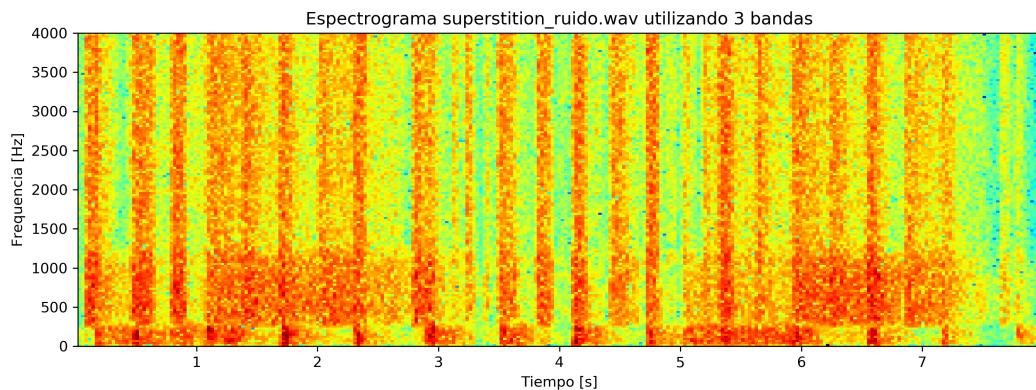


Figura 1.5: Espectrograma de la señal resultante con el banco de filtros de 3 bandas.

2. Ejercicio 2 - Síntesis de cuerda pulsada

Se trabajará ahora con el algoritmo de Karplus y Strong, para generar síntesis de cuerda pulsada utilizando tres tipos de filtros peine con retroalimentación positiva:

- Filtro peine
- Filtro peine aplicando un pasabajos
- Filtro peine aplicando un pasabajos y un pasatodo

2.1. Filtro peine

A partir del bloque de filtro peine que se observa en la figura 2.1 se deduce la siguiente ecuación de recurrencia

$$y[n] = x[n] + R^L y[n - L] \quad (2.1)$$

A partir de la ecuación (2.1) se deduce que ante un impulso de entrada (o semejante), la salida es una copia retardada L muestras y atenuada por un factor R^L que se repite periódicamente, por lo que la señal resultante tiene una frecuencia fundamental dada por

$$f_0 = \frac{f_s}{L} \quad (2.2)$$

se deduce de la ecuación anterior, y contemplando que la variable L solamente puede tomar valores enteros, que las frecuencias fundamentales que se pueden generar con este tipo de filtros son acotadas, es decir, son solo aquellas que sean una fracción entera de f_s .

Esto trae ciertas implicancias obvias y es que si quisiéramos generar una nota de frecuencia arbitraria como podría ser una nota dentro de la escala temperada, no sería posible a menos que la frecuencia de muestreo sea un múltiplo entero de dicha nota. Lo cual no ocurre prácticamente en ninguna situación.

Si el factor R^L no estuviera presente o $R^L = 1$, la salida sería simplemente una copia del impulso con su módulo invariante, la inclusión de este factor hace que la señal cada vez que se introduce en el lazo realimentado atenúe su amplitud por el factor, se asume que $R < 1$ ya que de lo contrario se obtendría una divergencia y por ende el filtro se volvería inestable. Se ilustra en la figura 2.2 el efecto de alterar el valor de R de manera muy sutil.

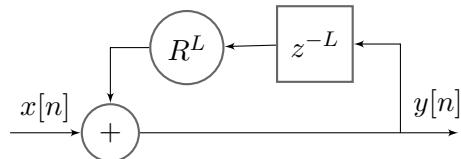


Figura 2.1: Filtro peine.

Resultado

El resultado obtenido al aplicar este filtro difiere bastante con el sonido de una cuerda real pulsada, para empezar es un sonido que da la sensación que está distorsionado pero no es específicamente una distorsión, sino que se debe principalmente a que todos los armónicos se escuchan con la misma sonoridad, algo que en una cuerda real claramente no ocurre, los armónicos de frecuencias más elevadas se atenúan bastante más rápido que los de frecuencias bajas.

Esto se revela en los espectrogramas, para todo el rango de frecuencias el degrade que se genera debido a la amplitud es uniforme verticalmente para cada bin calculado, lo que quiere decir que cada armónico pesa de la misma manera en el espectro.

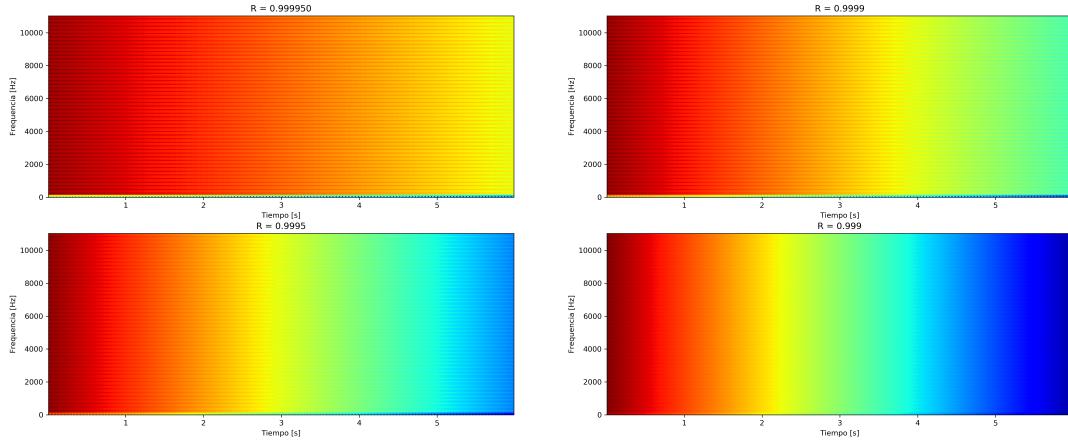


Figura 2.2: Respuesta al impulso de filtro peine ($L = 100$, $f_s = 22050\text{Hz}$)

2.2. Filtro peine aplicando pasabajos

Un primer refinamiento del algoritmo, para contrarrestar el efecto visto en la parte anterior, fue incluir un bloque **pasabajos** dentro del filtro, tal como indica la figura 2.3. Este bloque no es más que un filtro de media móvil de primer orden, que modela la atenuación de las altas frecuencias a lo largo del tiempo, cada vez que se produce una realimentación en el lazo el filtro se aplica de nuevo, dando el efecto de un filtro en cascada aplicado muchas veces.

La ecuación en recurrencia del filtro de media móvil esta dada por

$$y[n] = \frac{1}{2}x[n] + \frac{1}{2}x[n - 1]$$

de donde su transferencia es de la forma

$$H(z)_{m.mvil} = \frac{1}{2}[1 + z^{-1}]$$

que al evaluar en $z = e^{j\theta}$ se convierte en

$$H(e^{j\theta}) = \frac{1}{2}[1 + e^{-j\theta}] \frac{e^{\frac{j\theta}{2}}}{e^{\frac{j\theta}{2}}} = e^{\frac{-j\theta}{2}} \cos \frac{\theta}{2} \quad (2.3)$$

se observa que el filtro pasabajos introduce un retardo de fase lineal para todas las frecuencias, que coincide con un retardo de media muestra, esto sumado al retardo de L muestras del filtro peine por si solo, la nueva frecuencia fundamental que va a tener la salida será entonces

$$f_0 = \frac{f_s}{L + \frac{1}{2}} \quad (2.4)$$

La ecuación de recurrencia correspondiente al filtro nuevo está dada por

$$y[n] = \frac{1}{2}(x[n] + x[n - 1] + R^L(y[n - L] + y[n - L - 1])) \quad (2.5)$$

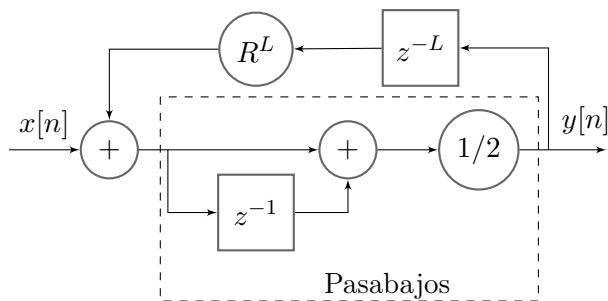


Figura 2.3: Filtro peine con pasabajos incorporado.

Resultado

Al agregar el pasabajo el cambio auditivo es muy notorio, se puede apreciar como ni bien pasan unos instantes de tiempo las frecuencias agudas se atenúan rápidamente producto de que a cada pasada de la señal realimentada corresponde a aplicar el filtro pasabajo y se produce una atenuación dada por $|H(e^{j\theta})|^m$. Lo que antes parecía un sonido más como un zumbido, ahora se percibe como un gran ataque y un decay más ameno, pero con un poco menos de cuerpo en el sonido.

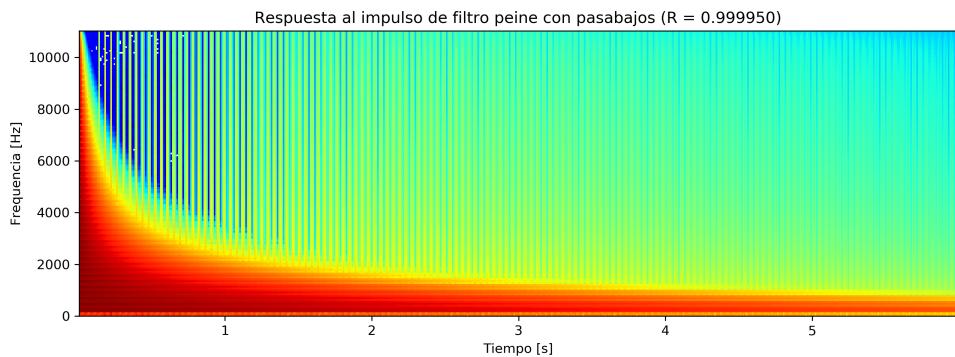


Figura 2.4: Espectrograma de nota sintetizada.

Como se observa en el espectrograma de la figura 2.4, en el instante inicial están presentes en el mismo orden de amplitud todos los componentes armónicos a lo largo del espectro, que es lo que da la sensación de un gran ataque como si se tratase de algo metálico. Esto en una cuerda real no sucede tan así, por mayores referencias en el anexo se puede ver un espectrograma de una cuerda real pulsada para contrastar.

Cabe destacar que para este caso se mantiene el problema que ya arrastrábamos de la sección anterior, y es que aún no podemos generar frecuencias arbitrarias, al agregar un retraso fijo en fase las nuevas frecuencias quedan limitadas a ser fracciones enteras de la frecuencia f_s , pero ahora con respecto a $(L + 1/2)$ en lugar de L .

2.3. Filtro peine aplicando pasabajos y pasatodo

Como último refinamiento y para solucionar el problema de generar frecuencias arbitrarias, se introdujo un filtro **pasatodo**. Idealmente quisiéramos un filtro de fase lineal que introduzca un retraso fraccionario δ que altere todas las frecuencias por igual pero que no altere su módulo,

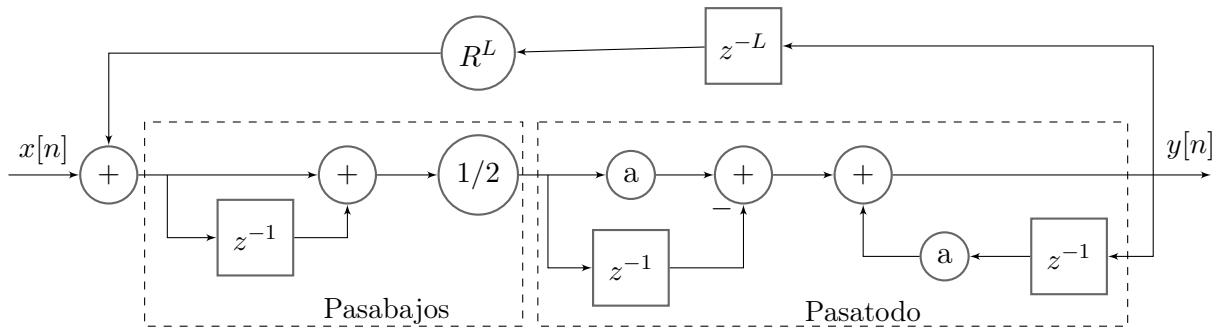


Figura 2.5: Filtro peine con afinación exacta.

de esta manera lograríamos generar cualquier frecuencia arbitraria a partir de f_s ya que cualquier cociente podríamos expresarlo en términos de la parte entera (L) y la parte decimal (δ). Pero la realidad es que el filtro pasatodo introducido en el diagrama de la figura 2.5 no introduce un retardo lineal, pero a efectos de trabajar en bajas frecuencias se aproximarán como si lo fuera.

Los siguientes cálculos muestran cuál es el retardo introducido por el filtro pasatodo¹. La transferencia del bloque pasatodo esta dada por

$$H(z) = \frac{z^{-1} - a}{1 - az^{-1}} \quad (2.6)$$

Evaluando en $e^{j\theta}$ y luego multiplicando numerador y denominador por $e^{\frac{j\theta}{2}}$

$$H(e^{j\theta}) = \frac{e^{-j\theta} - a}{1 - ae^{-j\theta}} = \frac{e^{\frac{-j\theta}{2}} - ae^{\frac{j\theta}{2}}}{e^{\frac{j\theta}{2}} - ae^{\frac{-j\theta}{2}}} \quad (2.7)$$

Observando que el denominador es el conjugado del numerador, o viceversa, es fácil ver que el argumento total está dado por el doble del argumento del numerador. De esta manera, escribiendo en coordenadas cartesianas el numerador del término a la derecha de la ecuación (2.7) y hallando el doble de su argumento

$$e^{\frac{-j\theta}{2}} - ae^{\frac{j\theta}{2}} = (a + 1) \cos\left(\frac{\theta}{2}\right) + j(a - 1) \sin\left(\frac{\theta}{2}\right) \quad (2.8)$$

$$\text{Arg}\left(H(e^{j\theta})\right) = -2 \arctan\left(\frac{1 - a}{1 + a} \tan\left(\frac{\theta}{2}\right)\right) \quad (2.9)$$

Se nota que claramente no es lineal el argumento de la transferencia, pero utilizando las aproximaciones

$$\theta \approx 0 \Rightarrow \tan\theta \approx \theta, \quad \arctan\theta \approx \theta \quad \text{y definiendo} \quad \delta = \frac{1 - a}{1 + a}$$

obtenemos aplicando en la ecuación (2.9)

$$\text{Arg}\left(H(e^{j\theta})\right) \approx -\delta\theta \quad \Rightarrow \quad \tau_{\text{Arg}}(\theta) \approx -\frac{\text{Arg}(\theta)}{\theta} \approx \delta \quad (2.10)$$

donde $\tau_{\text{Arg}}(\theta)$ corresponde a la cantidad de muestras que se retarda cada componente ².

¹Se omiten ciertos detalles del cálculo para no hacer pesada la lectura.

²Se destaca que esta aproximación solo sirve para bajas frecuencias, cerca de la frecuencia de nyquist el comportamiento de este bloque no se comporta para nada de manera lineal.

Para este caso la ecuación de recurrencia del filtro completo se puede escribir para que quede de una manera más legible como el aporte de la entrada más el aporte introducido por la propia salida retroalimentada. Las ecuaciones quedan de la siguiente manera

$$y_{\text{entrada}}[n] = \frac{-a}{2} (x[n] + x[n - 1]) + \frac{1}{2} (x[n - 1] + x[n - 2]) \quad (2.11)$$

$$y_{\text{retro}}[n] = ay[n - 1] + \frac{R^L}{2} (-a(y[n - L] + y[n - L - 1]) + y[n - L - 1] + y[n - L - 2]) \quad (2.12)$$

$$y[n] = y_{\text{entrada}}[n] + y_{\text{retro}}[n] \quad (2.13)$$

Por último, la frecuencia fundamental de la señal que devuelve este filtro va a tener la siguiente forma

$$f_0 = \frac{f_s}{L + \frac{1}{2} + \delta} \quad (2.14)$$

Siguiendo un análisis análogo al de los casos anteriores, la parte entera L del retardo la aporta el filtro peine estricto en si mismo, el retardo de media muestra lo aporta el bloque pasabajos, y por último se logra un filtro fraccionario arbitrario que está dado por el bloque pasatodo. De esta manera se logra a priori lo que se buscaba en un principio, que es poder generar frecuencias fundamentales arbitrarias sin depender de manera directa del valor de f_s al cual estemos trabajando.

Resultado

Lo destacable para este último refinamiento es corroborar que efectivamente existe un cambio en la altura de las notas generadas por el filtro completo a medida que se varía el valor de δ entre $[0, 1]$. Para visualizar esto se puede ver en el espectrograma de la figura siguiente un batido de frecuencia generado con una diferencia de 0.5Hz.

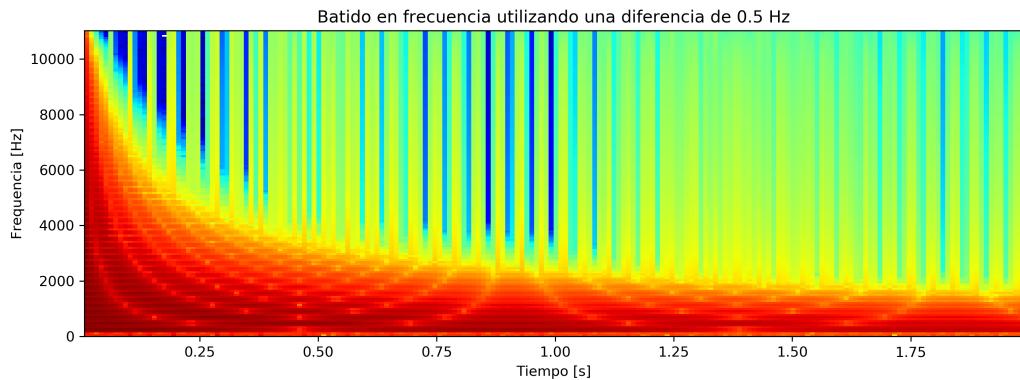


Figura 2.6: Batido de frecuencia generado a partir de sumar dos notas con diferencia de 0.5 Hz

3. Error en frecuencia experimental

Como era de esperar, a medida que el filtro se fue mejorando el error relativo fue disminuyendo, en particular se da una especie de simetría en el error del primer filtro con el segundo debido a que existe simplemente una diferencia de un factor de 1/2 entre el cálculo de un L y otro. Con el pasatodos se puede ver que el error relativo es bastante más acotado.

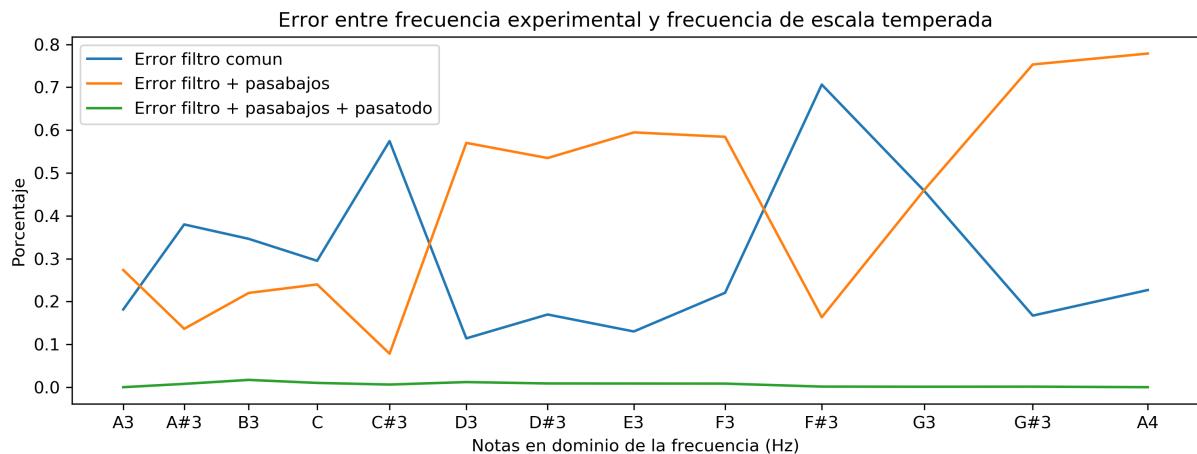


Figura 3.1: Error experimental entre la frecuencia teórica y la generada.