

Instituto Tecnológico de Buenos Aires

22.05 ANÁLISIS DE SEÑALES Y SISTEMAS DIGITALES

Trabajo práctico N°2

Grupo 3

MECHOULAM, Alan	58438
LAMBERTUCCI, Guido Enrique	58009
RODRIGUEZ TURCO, Martín Sebastian	56629
LONDERO BONAPARTE, Tomás Guillermo	58150

Profesores

Jacoby, Daniel Andres
Belaustegui Goitia, Carlos F.
Iribarren, Rodrigo Iñaki

Presentado: 14/05/20

Índice

1. Síntesis de sonidos mediante frecuencia modulada	2
1.1. Introducción al modelo	2
1.2. Instrumentos de viento	3
1.2.1. Clarinete	3
1.2.2. Trombón	4
1.2.3. Campana	5
1.2.4. Trompeta	6
1.3. Conclusiones	7
2. Efectos de Audio	8
3. Programas implementados	9
3.1. FFT	9
3.2. Programa Principal	9

1. Síntesis de sonidos mediante frecuencia modulada

1.1. Introducción al modelo

El siguiente método de síntesis se basa en modelar sonidos mediante señales moduladas en frecuencia. Es por eso que, dada la siguiente ecuación

$$x(t) = A(t) \cos(2\pi f_c + I(t) \cos(2\pi f_m t + \phi_m) + \phi_c) = A(t) \sin(2\pi f_c + I(t) \sin(2\pi f_m t)) \quad (1)$$

con $\phi_m = \phi_c = -\frac{\pi}{2}$, se busca elegir $A(t)$, $I(t)$, f_c y f_m adecuados para poder simular adecuadamente el sonido del instrumento deseado.

De la Ecuación (1) se observa que $I(t)$ es el índice de modulación, mientras que f_c y f_m son la frecuencia de la portadora y la moduladora, respectivamente. Cuando $I(t)$ es positivo, se observan frecuencias que se encuentran por encima y por debajo de la portadora en intervalos dados por la moduladora. La cantidad de frecuencias laterales que se observa está relacionada con dicho índice, es por ello que a mayor $I(t)$, mayor cantidad de picos en frecuencia. A su vez, la frecuencia central decrece con el aumento previamente mencionado ¹.

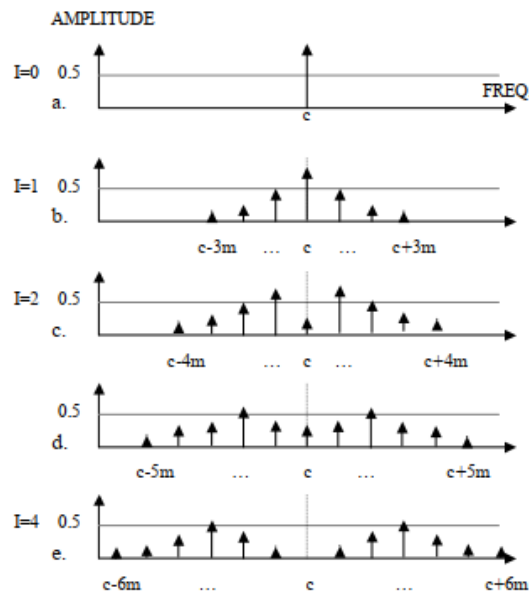


Figura 1: Aumento de $I(t)$ y como afecta las frecuencias laterales.

Las amplitudes de la portadora y las laterales son determinadas por las funciones de Bessel de primer y “n-esimo” orden $J_1(I(t))$ y $J_n(I(t))$, expresando $x(t)$ de la forma:

$$x(t) = A \sum_{k=0}^n J_k(I(t)) [\sin(\omega_c + k\omega_m)t - \sin(\omega_c - k\omega_m)t] \quad (2)$$

La principal ventaja de la síntesis mediante F.M. consiste justamente en el hecho de que se puede expresar con facilidad la relación entre la portadora y la moduladora, lo que produce las frecuencias laterales mencionadas previamente. Estas recaen en el lado negativo de frecuencias, reflejándose en el dominio positivo del espectro.

¹ John M. Chowning, [The Synthesis of Complex Audio Spectra by Means of Frequency Modulation](#). 1973.

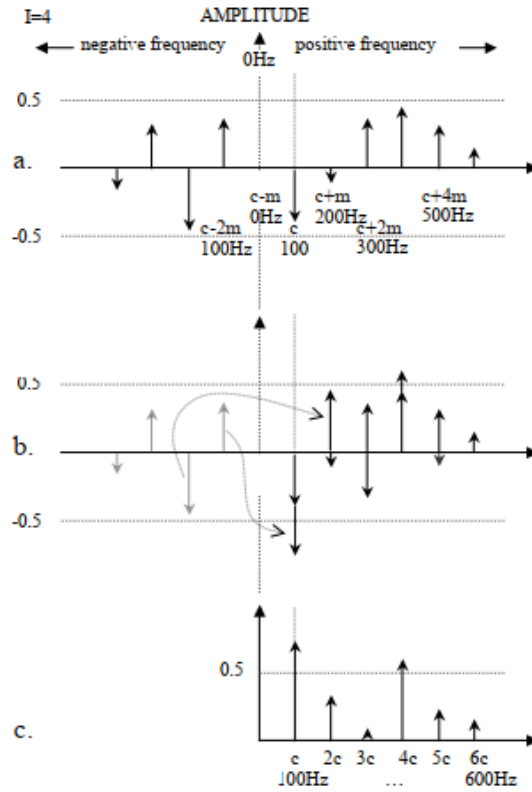


Figura 2: Representación gráfica de la proyección de frecuencias negativas.

Por otro lado, se puede expresar la relación de la frecuencia portadora y moduladora de la forma

$$\frac{f_c}{f_m} = \frac{N_1}{N_2} \quad (3)$$

lo que permite obtener una expresión para la frecuencia central

$$f_o = \frac{f_c}{N_1} = \frac{f_m}{N_2} \quad (4)$$

1.2. Instrumentos de viento

Este método puede ser empleado para sintetizar instrumentos de viento, fijando los parámetros $A(t)$, $I(t)$, f_c y f_m de una forma adecuada. Existen modelos dados para ciertos instrumentos, los cuales pueden ser modificados para conseguir un sonido más fidedigno, siendo el principal parámetro la relación presentada en la Ecuación (3).

1.2.1. Clarinete

Para la síntesis del clarinete se utilizó una relación $3f_c = 2f_m$. Además, las funciones $A(t)$ e $I(t)$ para una señal de una duración de 1 segundo son las presentadas a continuación.

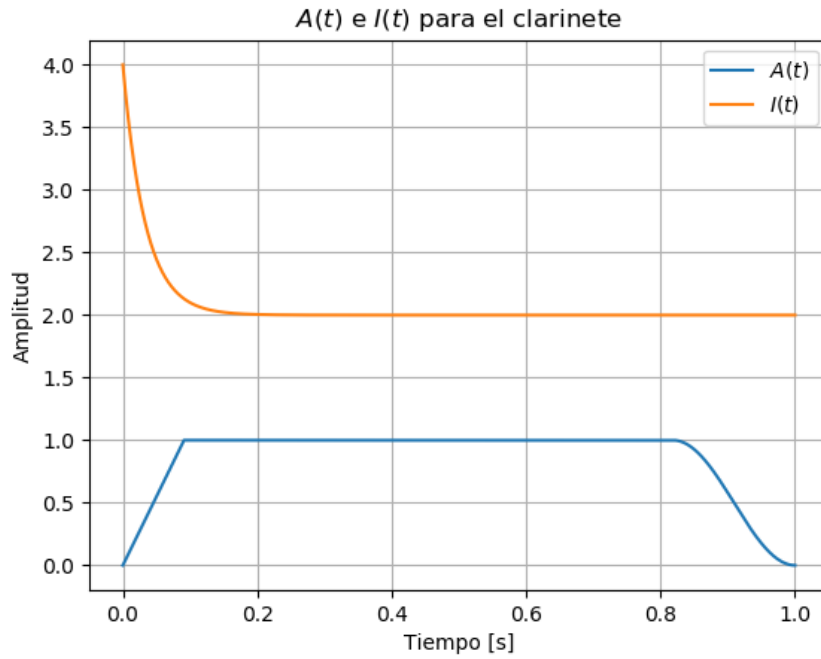


Figura 3: $A(t)$ e $I(t)$ para un clarinete.

Luego, se obtuvo un espectrograma para una nota a 440 Hz de dicho instrumento. Para dicho análisis se valió de una ventana de Hanning con 256 puntos para la $NFFT$ y una superposición de 128.

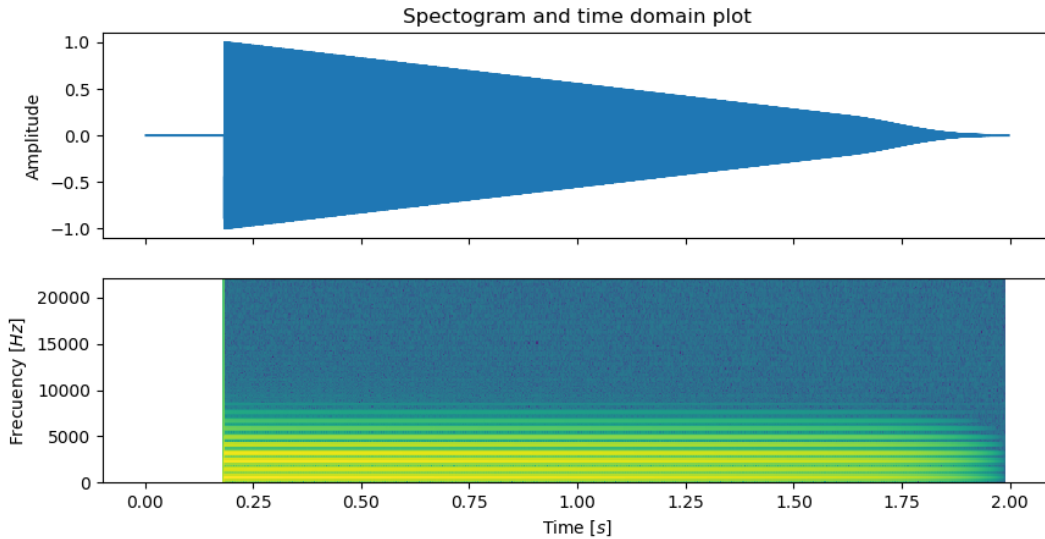


Figura 4: Espectrograma de una nota de clarinete a 440 Hz .

1.2.2. Trombón

De manera similar al caso anterior, para la síntesis del trombón se utilizó una relación $f_c = f_m$. Además, las funciones $A(t)$ e $I(t)$ para una señal de una duración de 1 segundo son las presentadas a continuación.

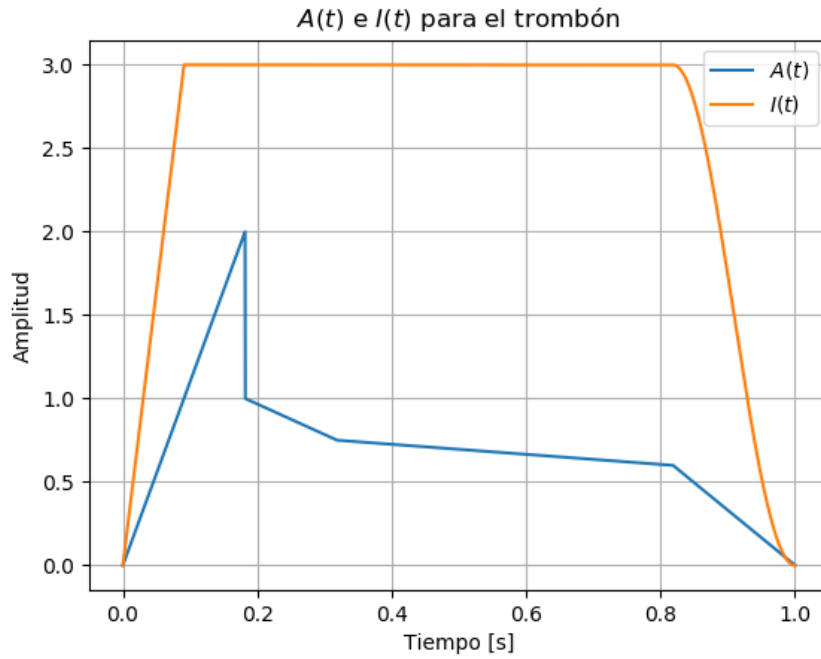


Figura 5: $A(t)$ e $I(t)$ para un trombón.

Después, se obtuvo un espectrograma para una nota a 440 Hz de dicho instrumento. Para dicho análisis se valió de una ventana de Hanning con 256 puntos para la $NFTT$ y una superposición de 128.

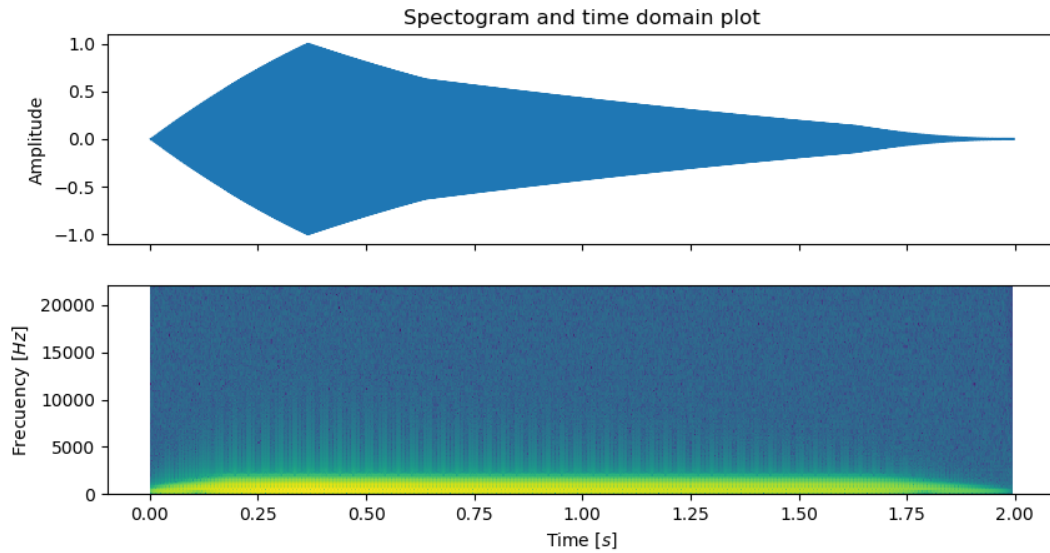


Figura 6: Espectrograma de una nota de trombón a 440 Hz .

1.2.3. Campana

Analogamente, para la campana se empleó una relación $f_c = 2f_m$. Además, las funciones $A(t)$ e $I(t)$ para una señal de una duración de 1 segundo son las presentadas a continuación.

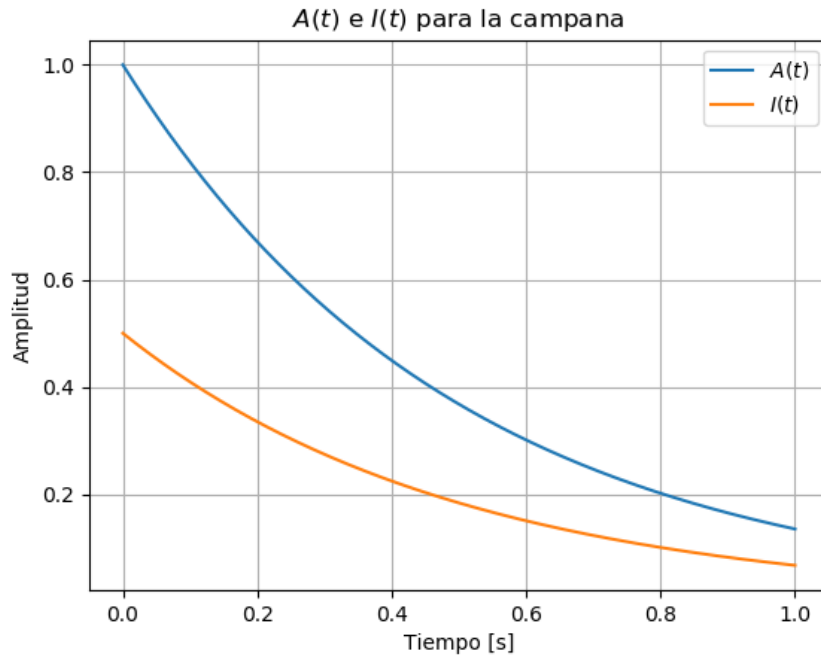


Figura 7: $A(t)$ e $I(t)$ para una campana.

Luego, se desarrollo un espectrograma para una nota a 440 Hz de dicho instrumento. Para dicho análisis se valió de una ventana de Hanning con 256 puntos para la $NFFT$ y una superposición de 128.

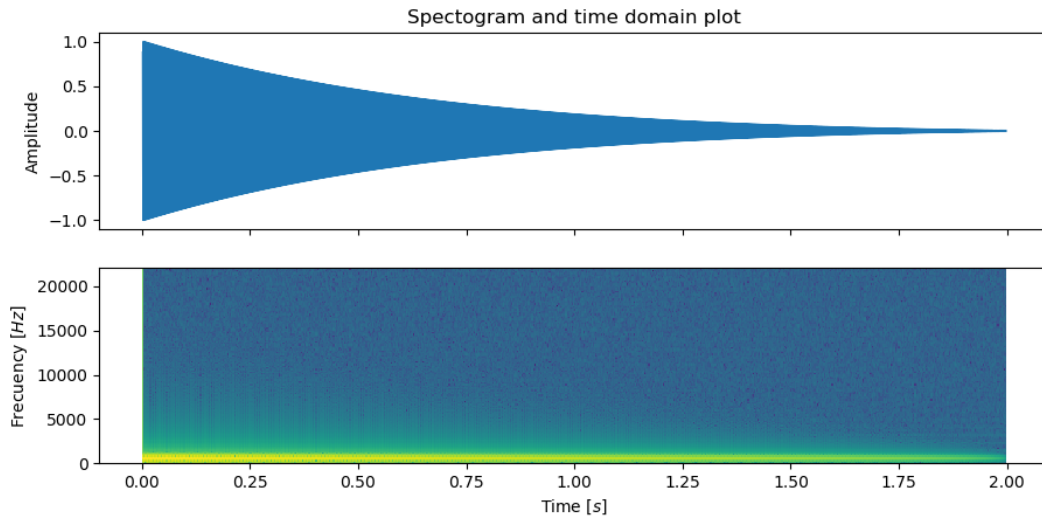


Figura 8: Espectrograma de una nota de campana a 440 Hz .

1.2.4. Trompeta

Finalmente, para sintetizar una trompeta se utilizó la misma relación que para el trombón, $f_c = f_m$. Además, las funciones $A(t)$ e $I(t)$ para una señal de una duración de 1 segundo son las presentadas a continuación.

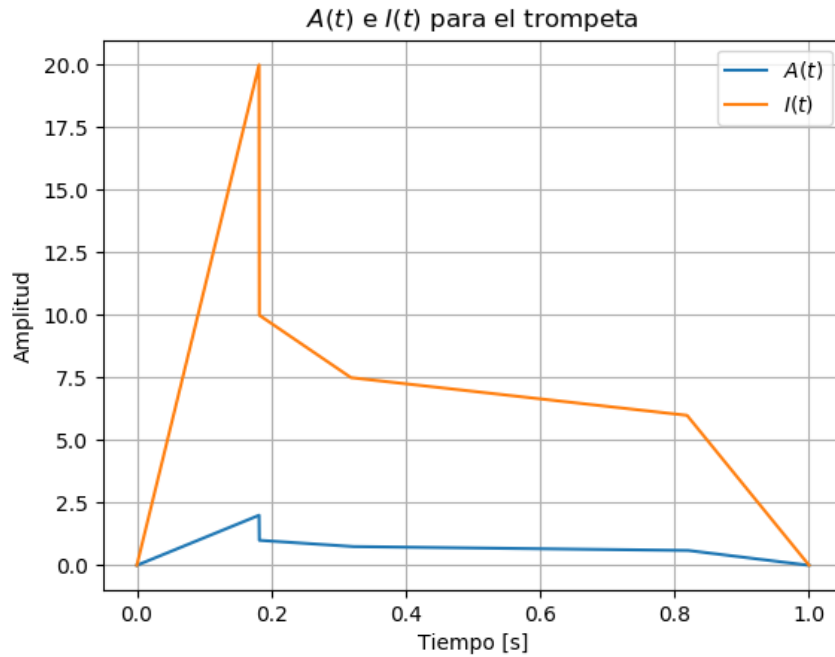


Figura 9: $A(t)$ e $I(t)$ para una trompeta.

Luego, se obtuvo un espectrograma para una nota a 440 Hz de dicho instrumento. Para dicho análisis se valió de una ventana de Hanning con 256 puntos para la $NFFT$ y una superposición de 128.

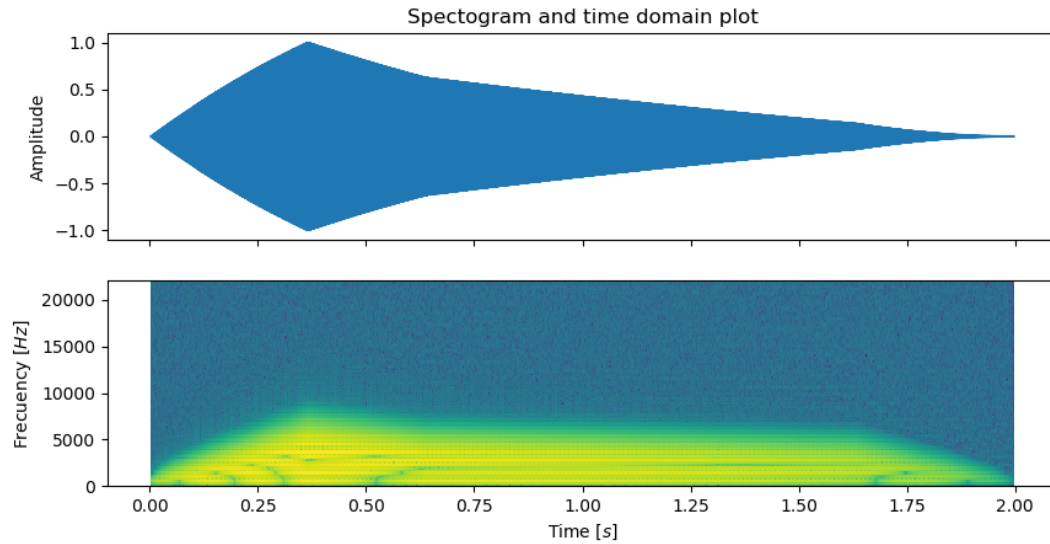


Figura 10: Espectrograma de una nota de trompeta a 440 Hz .

1.3. Conclusiones

No existe un acercamiento analítico para determinar el mejor conjunto de parámetros para la síntesis de un instrumento mediante el método de F.M., lo que dificulta la selección de estos ². Es por ello que existen bases de como tratar cada instrumento, las cuales pueden ser libremente modificadas, dependiendo del gusto de cada persona, justificando la existencia variaciones de cada “plantilla”. A pesar de ello, dichos cambios deben hacerse con cierto

²Andrew Horner y James W. Beauchamp, [Instrument Modeling and Synthesis](#). 2009.

criterio, ya que se desea mantener un sonido similar al real, y provocar cambios sin tener conocimiento pueden generar sonidos muy distantes a los deseados.

2. Introducción modelo Karplus-Strong

En esta sección se analiza el método de síntesis basado en el modelado físico, propuesto por Karplus-Strong.

2.1. Karplus-Strong básico

El modelo básico de Karplus-Strong consiste filtrar una forma de onda a través de una línea de retardo. Gracias a esto se logra simular el sonido de una cuerda de guitarra.

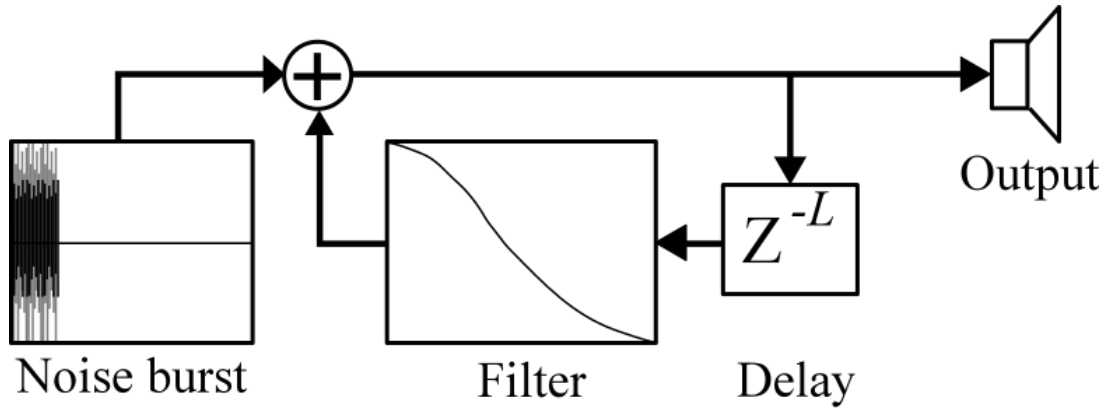


Figura 11: Modelo clásico Karplus-Strong.

2.1.1. Análisis teórico

Este algoritmo se puede describir por su diagrama en bloques como se ve a continuación.

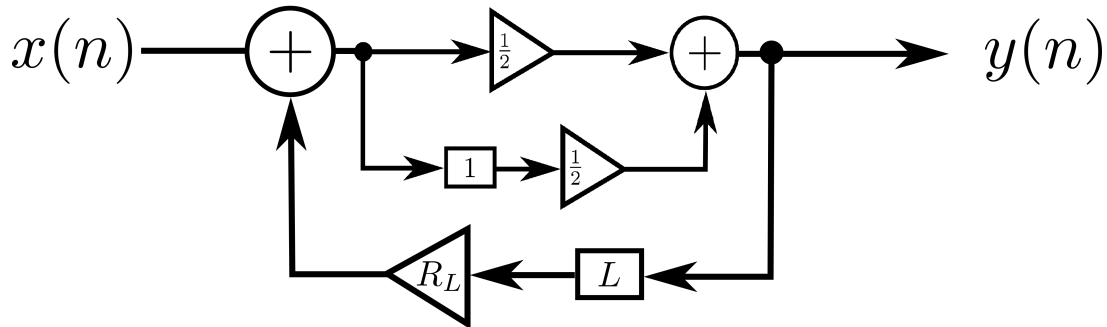


Figura 12: Algoritmo Karplus-Strong.

De este diagrama en bloques se puede obtener la ecuación en diferencias:

$$y(n) = \frac{1}{2} \cdot x(n) + \frac{1}{2} \cdot x(n-1) + \frac{1}{2} \cdot R_L \cdot y(n-L) + \frac{1}{2} \cdot R_L \cdot y(n-L-1) \quad (5)$$

A partir de esta expresión, se calcula su transformada Z y despeja para obtener la transferencia:

$$H(z) = \frac{\frac{1}{2} \cdot z^{L+1} + \frac{1}{2} \cdot z^L}{z^{L+1} - \frac{R_L}{2} \cdot z - \frac{R_L}{2}} \quad (6)$$

Vale la pena mencionar que de la Ecuación (??) es una ecuación en diferencias que cuenta como condiciones iniciales la “wavetable” suministrada por el ruido.

2.1.2. Análisis singularidades

Se observa que la Ecuación (??) cuenta con $L + 1$ polos y $L + 1$ ceros (de los cuales L de esos se encuentran en el origen). A continuación se muestra un diagrama de polos y ceros del sistema:

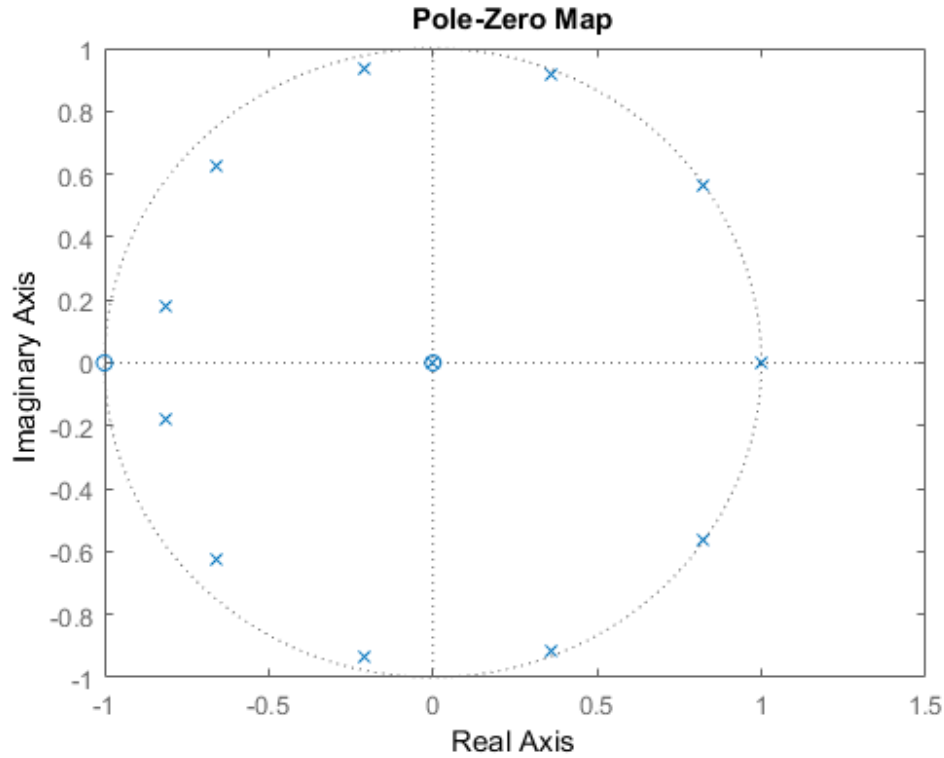


Figura 13: Diagrama de polos y ceros.

Adicionalmente, se graficó el diagrama de Bode del sistema.

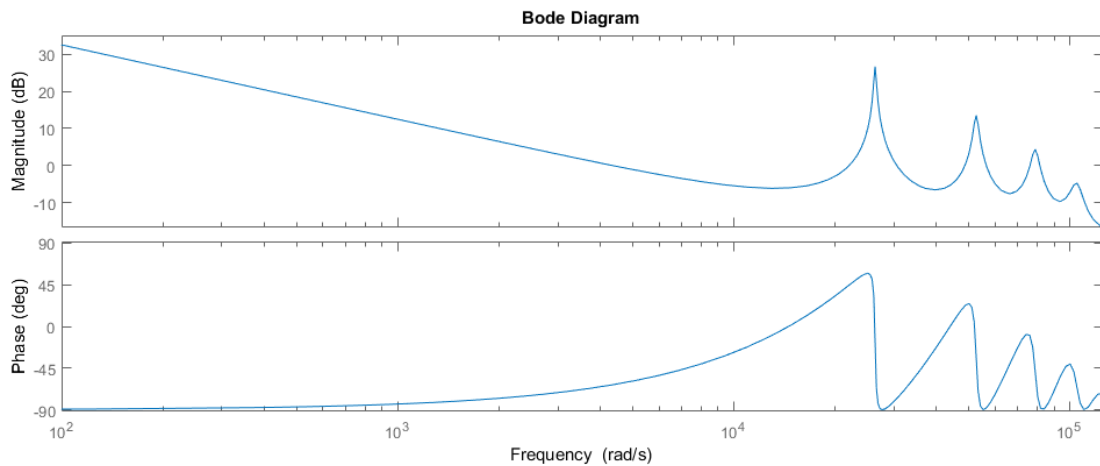


Figura 14: Diagrama de Bode.

Considerando que los parámetros eran: $f_s = 44.1 \text{ kHz}$, $L = 10$ y $R_L = 1$

2.1.3. Sintonización de frecuencia

En cuanto a la elección de una frecuencia de oscilación se puede observar en los últimos gráficos que existe un valor de frecuencia, la cual tiene mayor probabilidad de cumplir el criterio de Barkhausen, la cual corresponde a $f_r = \frac{f_s}{L+0.5}$. Esto se debe a que el sistema es la superposición de una línea de retraso L junto con otro sistema de retraso $L+1$. La señal al recorrer el lazo lo hace cada $\frac{L+L+1}{2}$ cambiando esto por frecuencia se obtiene:

$$f_r = \frac{f_s}{L + 0.5} \quad (7)$$

2.1.4. Tipos de ruido

Se propuso excitar el sistema con distintos tipos de ruido de entrada, siendo estos:

- Ruido Gaussiano
- Ruido Uniforme
- Ruido Binario

Primero se aplicó ruido gaussiano de longitud $L = 50$, y se obtuvo la siguiente salida:

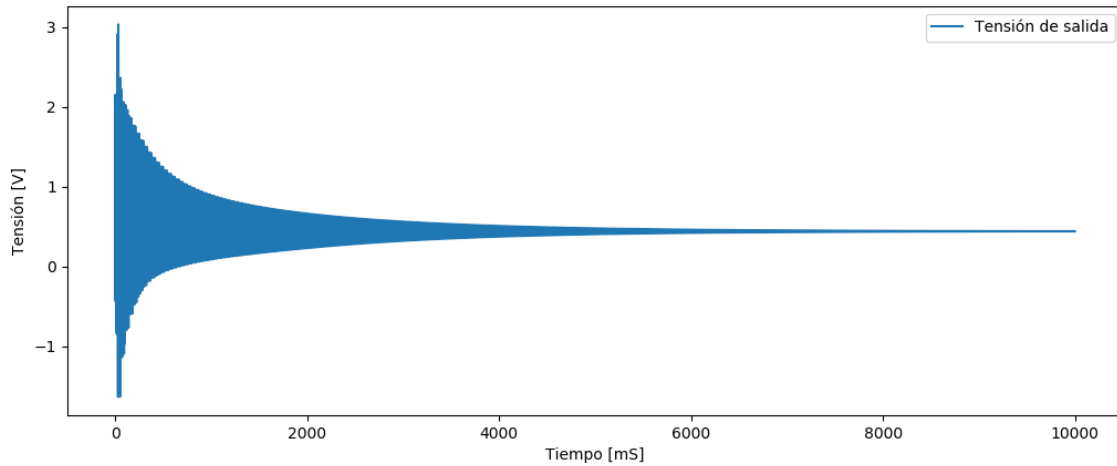


Figura 15: Respuesta a ruido gaussiano, $L = 50$.

Se puede apreciar que la respuesta a este tipo de ruido, parece tener cierta simetría respecto al eje. Algo notable de mencionar es como el sistema se estabiliza para un valor levemente superior a 0. Adicionalmente se realizó un detalle en la imagen:

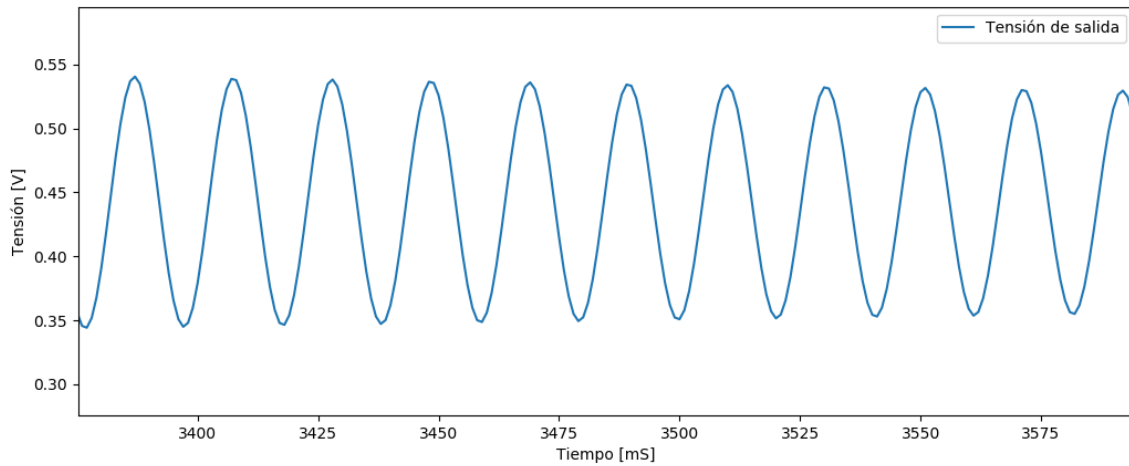


Figura 16: Respuesta a ruido gaussiano detalle, $L = 50$.

En esta, se puede apreciar que el sistema efectivamente se encuentra oscilando, mientras siendo atenuado por una envolvente.

Luego se excitó el sistema con ruido uniforme, obteniendo la siguiente respuesta:

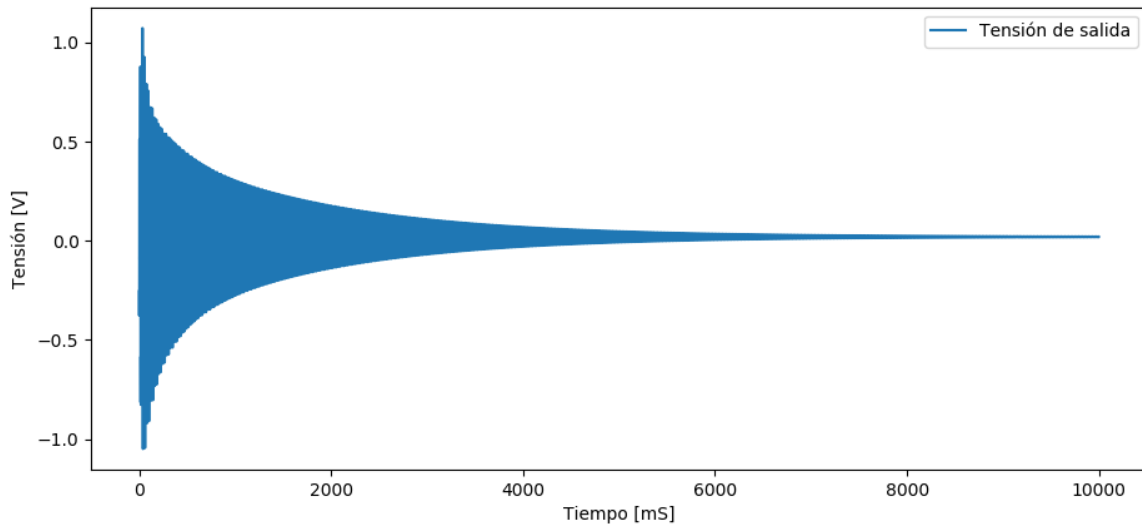


Figura 17: Respuesta a ruido uniforme, $L = 50$.

Se puede apreciar que la respuesta al ruido uniforme, cuenta con una simetría respecto al eje mas notable que el gaussiano, y un tiempo de decay más lento. Adicionalmente el valor al que tiende es mucho mas cercano a cero.

Luego, de la misma forma que para el gaussiano, se realizó un detalle en la figura:

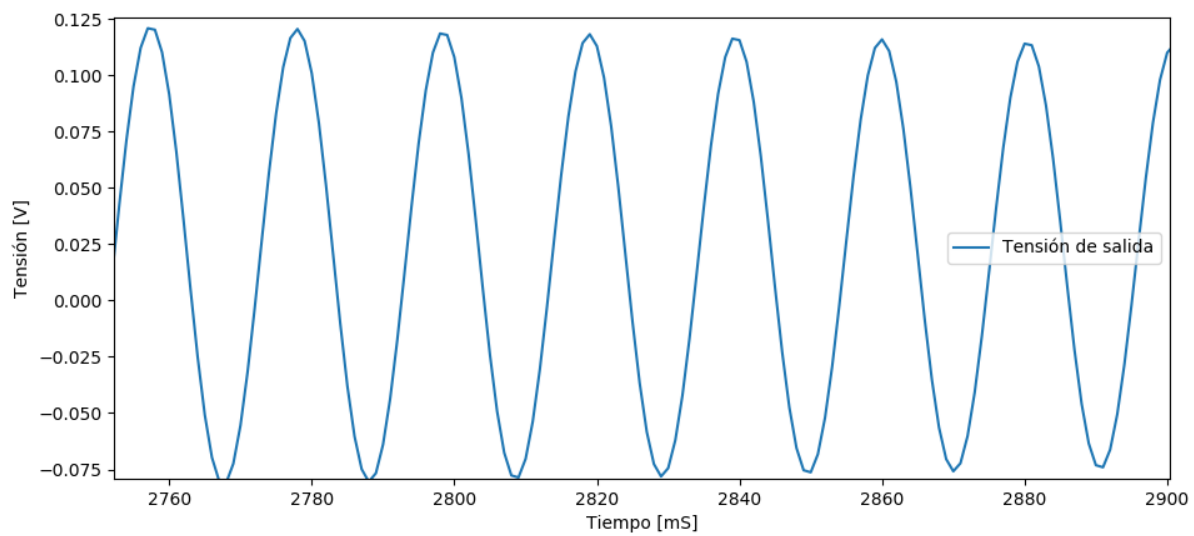


Figura 18: Respuesta a ruido uniforme detalle, $L = 50$.

Finalmente se ingresó al sistema con ruido Binario aleatorio.

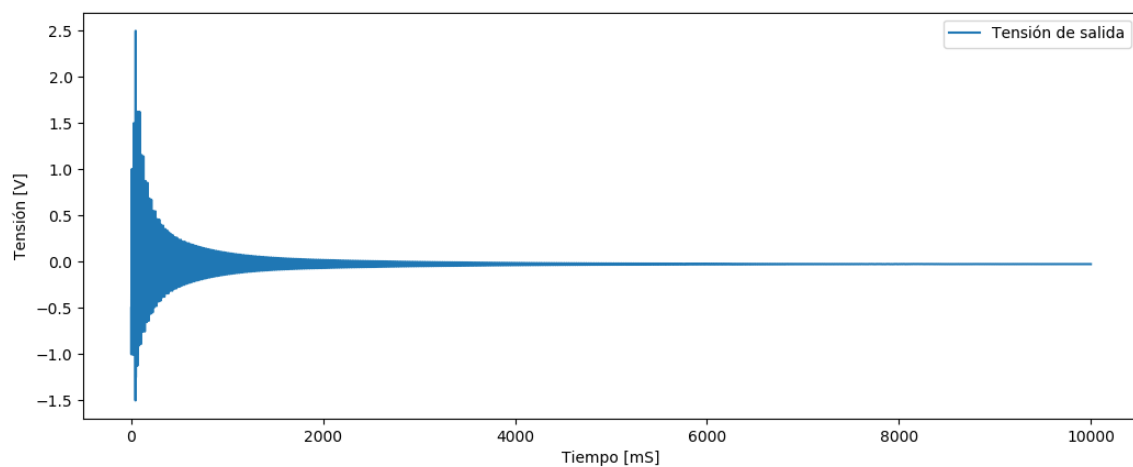


Figura 19: Respuesta a ruido Binario, $L=50$.

Se puede observar en la Figura (??) que existe una envolvente mucho más severa que las anteriores. Además, a medida que el tiempo aumenta, el valor de tensión tiende a cero, lo cual es un efecto deseado.

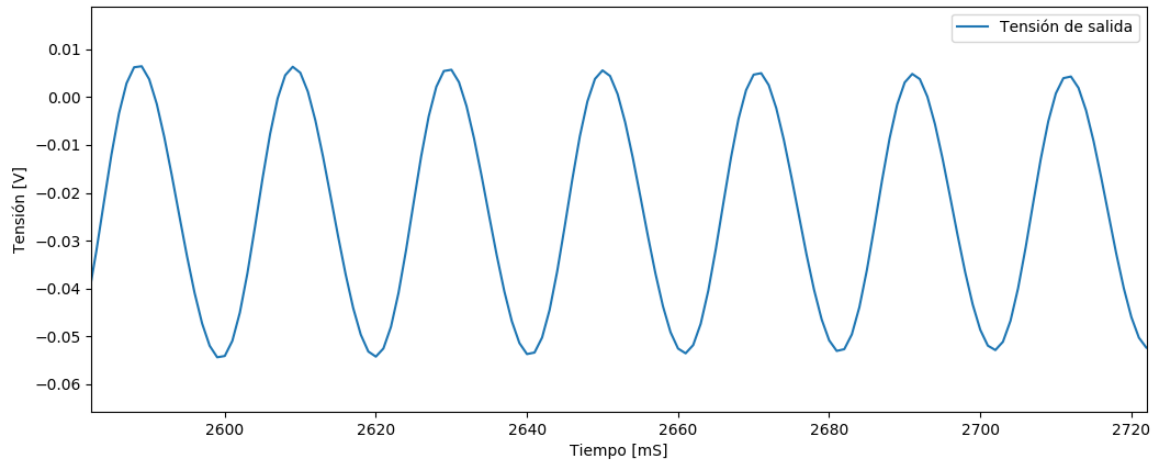


Figura 20: Respuesta a ruido Binario en detalle , $L = 50$.

Adicionalmente se computó la respuesta impulsiva del sistema, siendo esta:

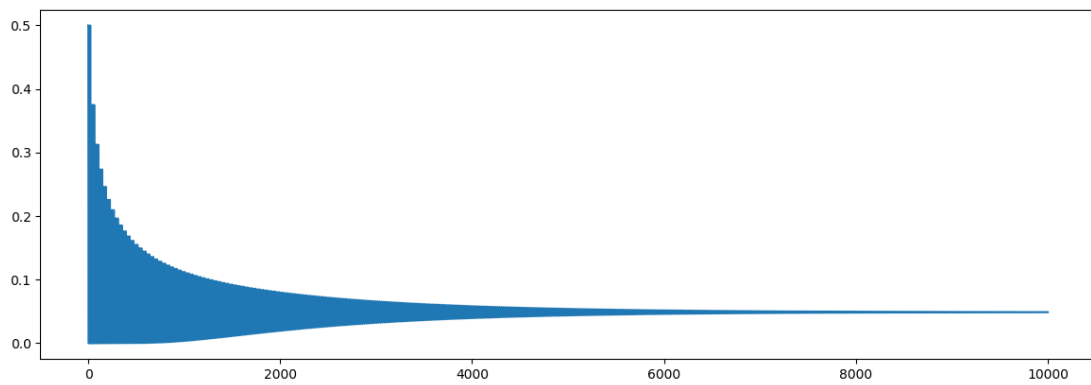


Figura 21: Respuesta al Impulso, $L = 50$.

Vale la pena mencionar que la salida nunca es negativa, esto se debe a que cuando la excitación es únicamente positiva, dado que la realimentación es positiva y que no existe ninguna inversión de fase, la salida siempre resulta positiva. Así también se muestra un detalle de la respuesta.

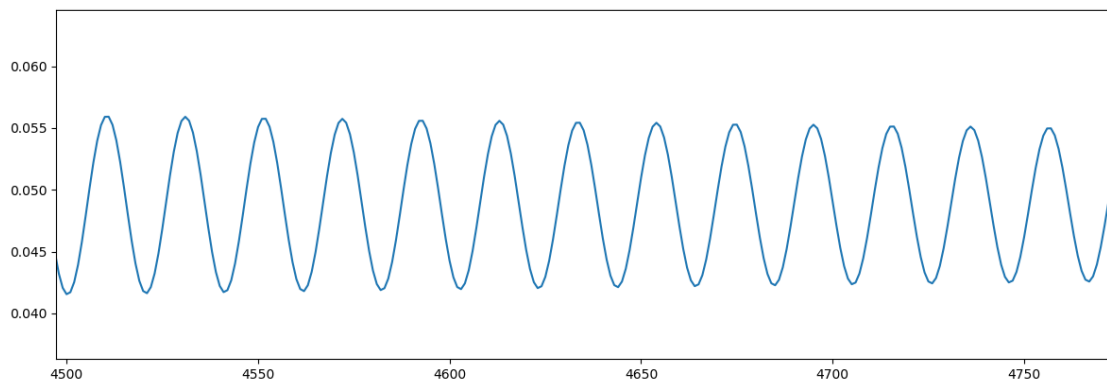


Figura 22: Respuesta al Impulso detalle, $L = 50$.

También cabe destacar que la “velocity” fue introducida no solo al final de la sintetización como modulador de volumen, sino también en el ruido del comienzo, dado que dicha propiedad, en la guitarra, simboliza la intensidad con la cual fue tocada la cuerda.

2.1.5. Estabilidad

La estabilidad del sistema es determinada por la Ecuación (??). Se puede observar que si RL es mayor o igual a la unidad, el sistema es inestable. Si bien, teóricamente es cierto, en la realidad se encuentra que si se da dicha condición, no solo no se provoca inestabilidad, sino que es preferible este valor dado, ya que logra extender las oscilaciones un mayor tiempo.

2.2. Mejora propuesta

El sistema anterior cuenta con algunas limitaciones. Un claro ejemplo de esto es la frecuencia, a cual para valores pequeños de L , la diferencia entre $f_r(L)$ y $f_r(L + 1)$ resulta grande, dejando una banda de frecuencias sin poder ser sintonizadas, como ejemplifica la siguiente imagen.

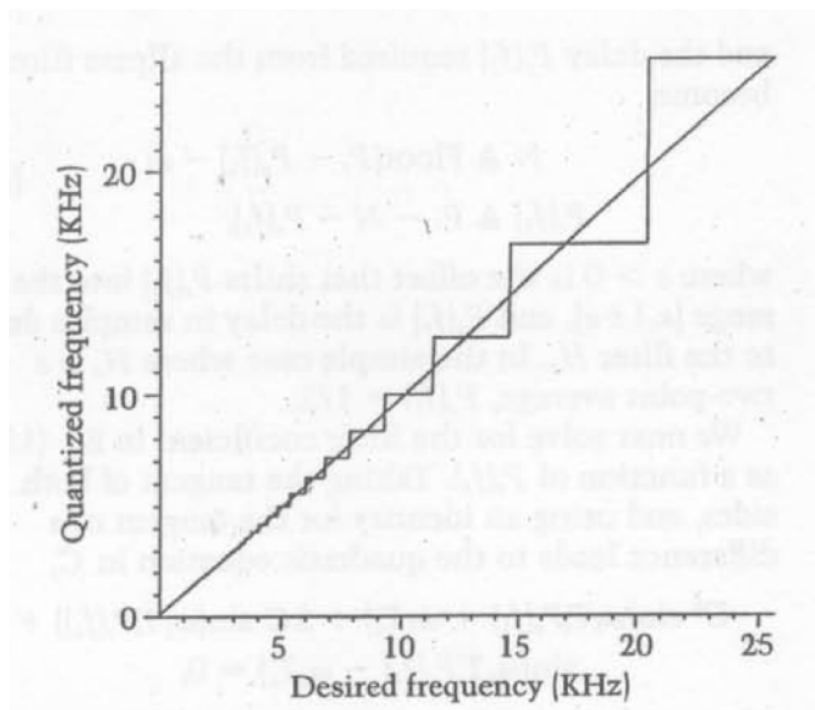


Figura 23: Dificultad en sintonización de frecuencia.

Otro defecto, es el final abrupto con el que cuentan algunas notas. Si se pide una duración reducida, las discontinuidades en el sonido provocan un efecto indeseado.

2.2.1. Sintonización de frecuencia

Para definir la frecuencia de resonancia del sistema, dado que la Ecuación (??) no deja grados de libertad, ya que la frecuencia de muestreo es fija, lo único que se puede hacer es realizar una modificación al circuito. El nuevo modelo propuesto es el siguiente:

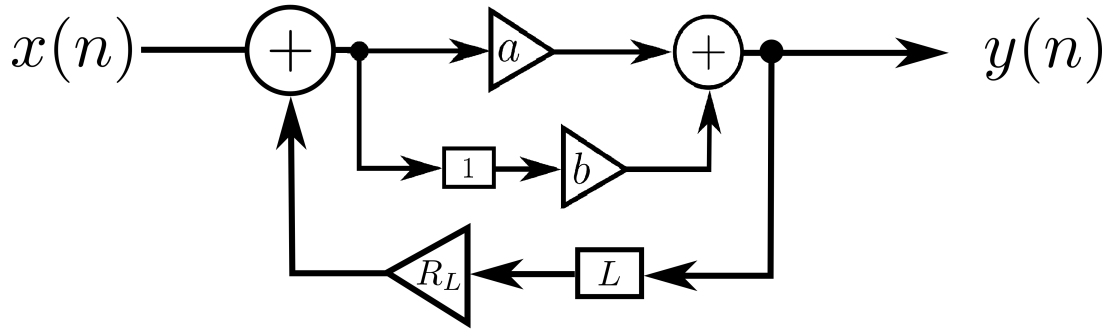


Figura 24: Algoritmo Karplus-Strong mejora propuesta.

Utilizando esta propuesta se puede obtener la ecuación en diferencias del sistema, siendo esta:

$$y(n) = a \cdot x(n) + b \cdot x(n-1) + a \cdot R_L \cdot y(n-L) + b \cdot R_L \cdot y(n-L-1) \quad (8)$$

Nótese, que si $a = b = \frac{1}{2}$ el sistema coincide con el original. Utilizando el mismo razonamiento, se llega a que la nueva frecuencia de resonancia es:

$$f_r = \frac{f_s}{L \cdot (a + b) + b} \quad (9)$$

Se debe elegir apropiadamente a y b , teniendo en cuenta que su suma debe de dar 1 para que el sistema sea el promedio ponderado realizado por estos sea realmente un promedio.

Un punto en contra de esta solución es que, dependiendo de los valores de a y b , se toma más en consideración una u otra línea del promedio.

2.2.2. Continuidad del sonido

Se tuvo en cuenta que el sonido sea una función suave, dado que cambios bruscos en el sonido no son placenteros ni esperados en la sintetización que es deseada implementar. Para lidiar con este problema lo que se realizó fue definir un factor de ventana, a partir de la cual se atenuó gradualmente el sonido hasta que sea nulo. Para esto se definió una ventana que tiene valor unitario para valores de tiempo menores al factor de ventana especificado, mientras que para valores superiores, decae gradualmente con una función cosenoidal, barriendo desde 0 a $\frac{\pi}{2}$ como se ilustra a continuación.

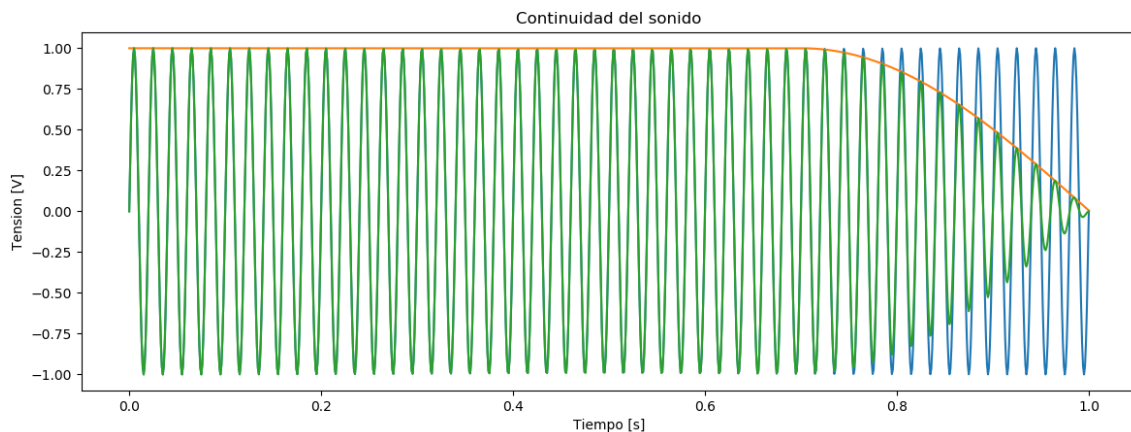


Figura 25: Muestra del factor de ventana.

De esta forma se aumenta la calidad de la síntesis en gran medida si se combina con la idea de que la duración de la nota y la del sonido son cosas distintas.

2.2.3. Caja de resonancia

La caja de resonancia es una parte primordial de la gran mayoría de instrumentos acústicos, principalmente de cuerda y percusión. Tiene la finalidad de amplificar o modular un sonido (en los instrumentos de cuerda generalmente a través de un puente). Los instrumentos que cubren rangos de sonidos más graves, como el contrabajo, el violonchelo o el bombo, necesitan una caja de resonancia bastante mayor que el resto.

La necesidad de utilizar una caja de resonancia en una guitarra se debe justamente a que la vibración de las cuerdas, que generan los frente de onda esféricos propios del sonido, tienen una gran dificultad para propagarse por el medio (el aire) debido a su reducida amplitud.

La manera en la cual este elemento sopesa dicho problema es capturando parte del frente de onda en su cavidad. Allí las frecuencias producidas por las cuerdas son amplificadas en igual magnitud y luego liberadas al medio con la intensidad suficiente para que pueda propagarse y que sea apreciable el sonido.

Finalmente cabe destacar que en el ámbito de señales realmente no es necesario dicha caja, al igual que tampoco un filtro que la emule, debido a que su función es amplificar, efecto que se puede hacer digitalmente sin ningún problema.

2.3. Karplus-Strong percusión

Se puede realizar una modificación al modelo inicial agregando un factor aleatorio como se observa a continuación:

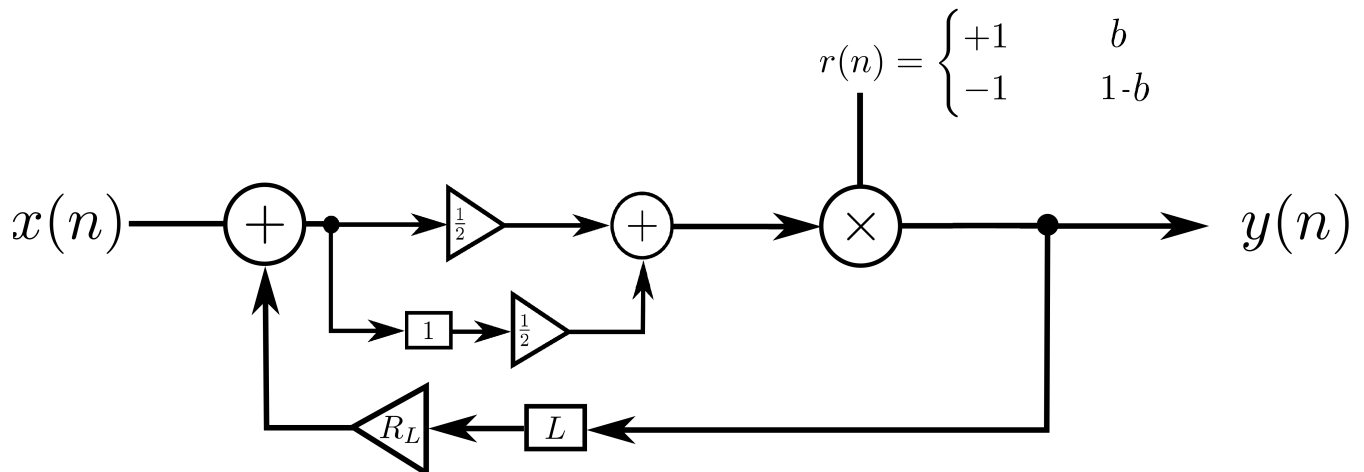


Figura 26: Karplus-Strong percusión.

El valor de b indica el nivel de cambio para el sonido final. Vale la pena notar que si $b = 1$, el algoritmo resulta igual al original. Para obtener sonidos de percusión se utiliza un valor de $b = 0,5$.

Una gran diferencia entre la guitarra sintetizada y un elemento de percusión es que estos últimos no cuentan con “notas”, sino que tienen un sonido característico, para obtener un sonido similar a un elemento de percusión se utilizan valores elevados de L .

2.4. Espectrogramas

Finalmente se realizaron espectrogramas de tanto la guitarra como el “redoblante” obteniendo los siguientes resultados.

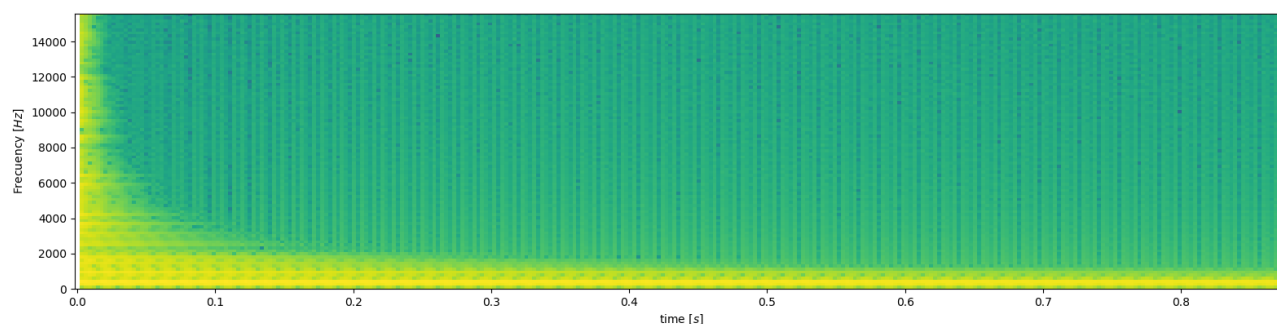


Figura 27: Espectrograma Guitarra.

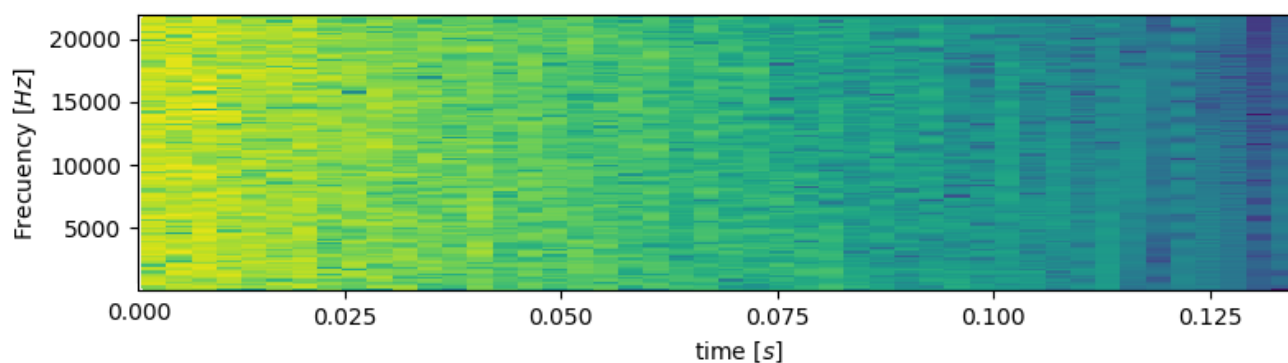


Figura 28: Espectrograma Redoblante.

Se puede diferenciar que la guitarra cuenta con mayor densidad armónica de baja frecuencia al igual que se extiende considerablemente mas en el tiempo, mientras que el redoblante cuenta con mayor contenido de alta frecuencia y su duración es mucho menor.

3. Efectos de Audio

Se implementaron dos efectos, eco simple y reverberación, siguiendo el método descrito por M. Schroeder ³. Para el eco se utilizó una sola línea de delay mostrada en la Figura (11).

³M. Schroeder, "Natural Sounding Artificial Reverberation", Journal of the Audio Engineering Society, vol. 10, no. 10, 1962. [Accessed 9 May 2020].

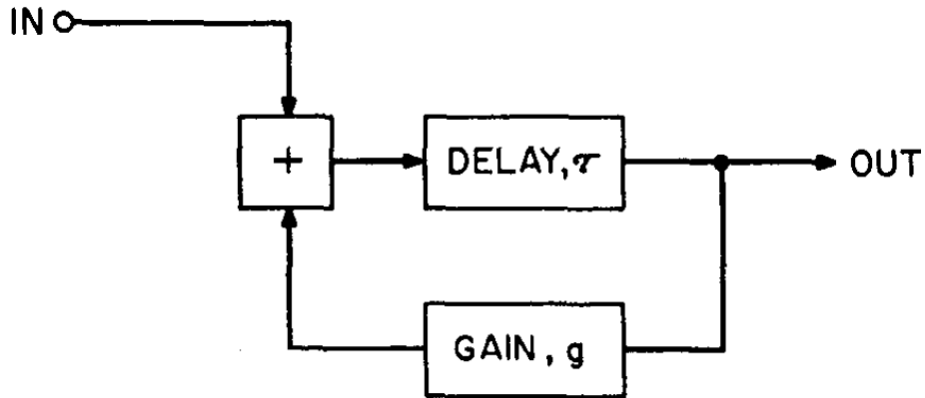


Figura 29: Línea de retardo.

Los parámetros para configurar este efecto son los de delay time τ , el cual denota el retardo de la línea; y decay factor g , el cual describe la ganancia del lazo. Se puede modelizar el tiempo en el cual el sonido decae por 60 dB, T , como

$$T = \frac{3 * \tau}{-\log_{10}|g|} \quad (10)$$

Luego, se utilizó el modelo propuesto por M. Schroeder mostrado en la Figura (12) para implementar la reverberación, con cuatro líneas de retardo en paralelo, seguidas de dos filtros pasa todo con línea de retardo, los cuales aumentan la densidad de los ecos sin afectar la ganancia del sistema en función de la frecuencia.

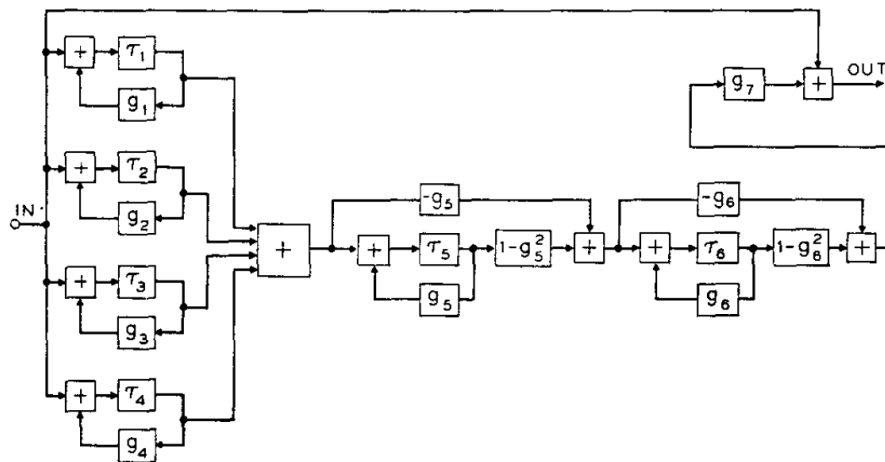


Figura 30: Filtro all-pass con línea de retardo.

Se utilizaron los parámetros sugeridos por Schroeder detallados a continuación⁴

⁴John M. Chowning, [The Synthesis of Complex Audio Spectra by Means of Frequency Modulation](#). 1973.

Parámetro	Sugerencia
τ_5	5 <i>ms</i>
τ_6	1.7 <i>ms</i>
g_5	0.7
g_6	0.7
τ_1	101.560 <i>ms</i>
τ_2	113.356 <i>ms</i>
τ_3	122.426 <i>ms</i>
τ_4	131.54 <i>ms</i>

Tabla 1: Parámetros utilizados según las referencias empleadas.

Luego, se ajustaron los valores de g_1 hasta g_4 según (5). Finalmente, para el eco se le pide al usuario que ingrese el delay time y el decay factor, denominados en el programa con las letras T y D . Para la reverberación, se le pide al usuario el tiempo de reverberación, el cual es el tiempo en el que el sonido decae 60 *dB*, denotado con la letra T ; y el mix factor, detallado en la Figura (12) como g_7 , denotado en el programa con la letra M .

4. Programas implementados

4.1. FFT

Se implementó la FFT utilizando el algoritmo de Cooley-Tukey de manera recursiva. Se probó con diversas entradas reales aleatorias, de tamaño 4096, con una media temporal de 40 μs .

4.2. Programa Principal

Se desarrolló un programa, el cual permite agregar midis de cualquier duración, para luego sintetizar cada track deseado dicho archivo. Permite seleccionar diversos instrumentos, varios de ellos con parámetros modificables. Además, es posible compilar dichos tracks en un archivo del tipo wav y hasta generar un archivo de “preview” de un track dado, es decir una pequeña muestra de como se escucha un track en particular.

También es posible elegir efectos a aplicar, tanto para el archivo wav final, como para cada track en particular.

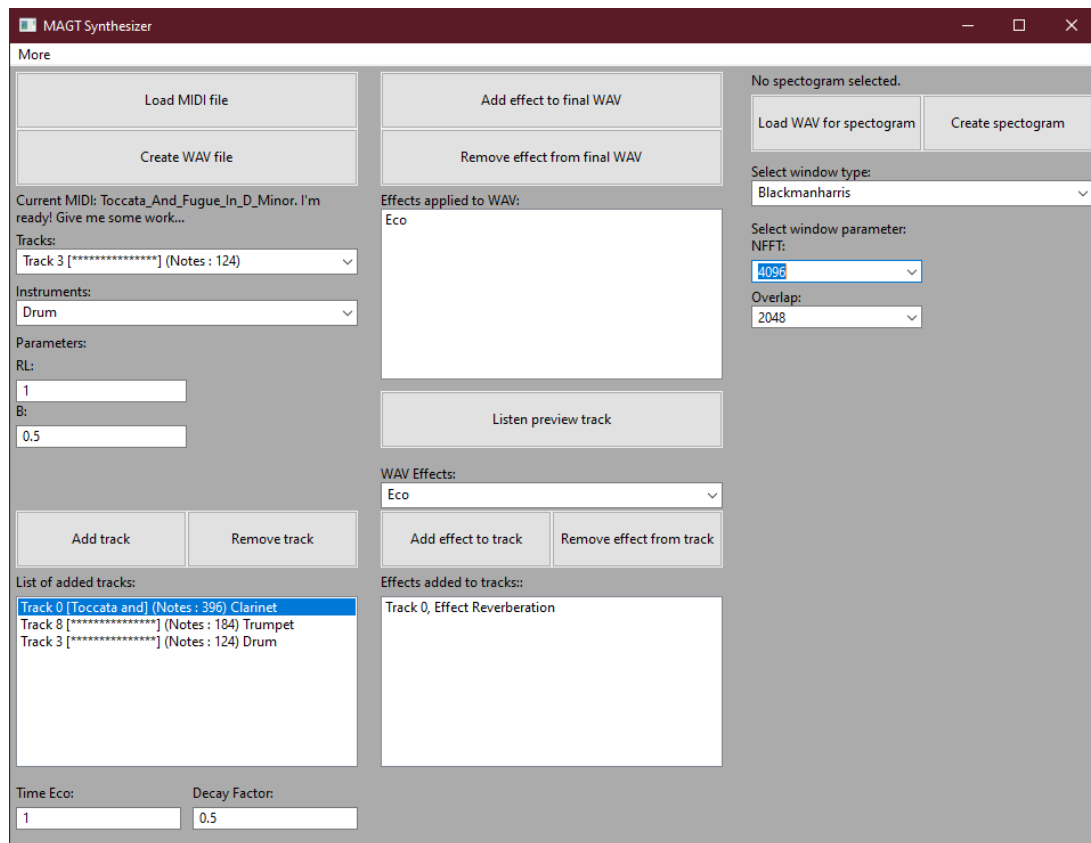


Figura 31: Captura de la GUI implementada.

Finalmente se destaca que es posible realizar espectrogramas de cualquier archivo wav, sea generado por este programa o no, pudiendo elegir el tipo de ventana, la cantidad de puntos de la FFT y el factor de overlap.

Cabe destacar que el programa fue desarrollado en su totalidad en C++, a excepción de los gráficos del espectrograma, los cuales son generados mediante un llamado a un código en Python. El front-end se implementó mediante el uso de la librería WxWidgets.