1 Síntesis mediante modulación de frecuencia

1.1 Introducción

A la hora de modelizar instrumentos musicales a través de algoritmos matemáticos, distinguimos dos grandes categorías: modelos de señales y modelos físicos. Los modelos de señales buscan reconstruir el sonido que se percibe sin analizar la fuente específica que provoca el sonido, mientras que los modelos físicos buscan simular el comportamiento de la fuente que emite dicho sonido.

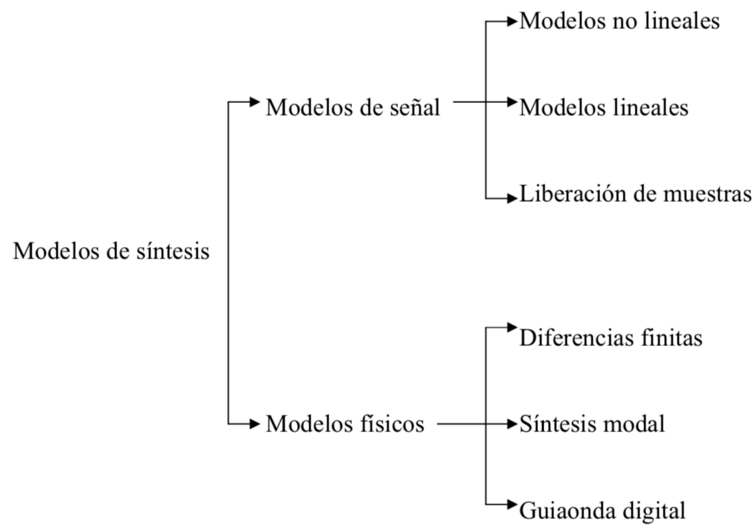


Figura 1: Esquema de los distintos modelos de algoritmos.

Dentro de los modelos de señal, podemos ver tres subconjuntos: modelos lineales, no lineales y liberación de muestras. En la sección de los métodos no lineales podemos distinguir distintos métodos de síntesis, a saber: de síntesis FM, síntesis AFM, síntesis DFM y síntesis PD entre varios métodos. En este trabajo práctico, nos vamos a centrar en la síntesis de FM.

1.2 Síntesis de FM

El método de síntesis de FM fue presentado por la marca de teclados Yamaha en 1980, basándose en los estudios de John Chowning sobre la teoría FM de radiofrecuencia a la síntesis de audio. La síntesis FM está basada, como su propio nombre indica, en la modulación en frecuencia. La frecuencia de una onda determinada se modula por otra onda de distinta frecuencia. El resultado contiene elementos de ambas frecuencias junto con nuevos armónicos relacionados matemáticamente con las frecuencias originales. Se puede demostrar teóricamente que cualquier sonido, por complejo que sea, puede ser modelado mediante una serie de modulaciones en frecuencia de ondas senoidales.

1.2.1 Análisis matemático

La ecuación general de una señal FM es la siguiente:

$$\begin{cases} y(t) = A(t) \cdot cos(\Phi(t)) \\ \Phi(t) = 2\pi f_c t + I(t) \cdot cos(2\pi f_m t + \phi_m) + \phi_c \end{cases}$$

donde A(t) y $\Phi(t)$ son la amplitud y fase de la señal (ambas variantes en el tiempo). Si combinamos ambas ecuaciones, asumiendo ademas que para nuestro caso el desplazamiento de fase es de $-\frac{\pi}{2}$, obtenemos la siguiente expresión:

$$x(t) = A(t) \cdot sen(2\pi f_c t + I(t) \cdot sen(2\pi f_m t))$$

Para sintetizar los distintos sonidos, basta con elegir adecuadamente A(t) e I(t), y proporcionar una adecuada relación entre f_c y f_m. La función de I(t), denominada envolvente del índice de modulación, es la de determinar el contenido armónico del sonido. Si derivamos la fase respecto del tiempo, podemos obtener la frecuencia instantanea

$$f_i = f_c - I(t) \cdot f_m sin(2\pi f_m t + \phi_m) + \frac{1}{2\pi}\frac{dI(t)}{dt} \cdot cos(2\pi f_m t + \phi_m)$$

con esa expresión, podemos ver que la relación precisa del índice de modulación I(t), podemos lograr que la función cambie la totalidad de sus contenidos armónicos. Normalmente tanto A(t) como I(t) suelen ser constantes la mayor parte del tiempo y dejan de serlo normalmente hacia el comienzo o final del sonido, para así tener en cuenta efectos transitorios como consecuencia del ataque de tecla. Las bandas laterales poseerán armónicos parciales a las frecuencias $f_c \pm nf_m$. Si f_c y f_m son ambos racionales, en una relación 1:N, el espectro resultante será armónico pero sin incluir los parciales que sean múltiplos de N. Esta propiedad es muy útil para sintetizar instrumentos de embocadura cilíndrica como el clarinete, los cuales se caracterizan por incluir en el espectro sólo los armónicos impares del tono fundamental. Si por el contrario f_c y f_m son irracionales entonces el espectro resultante será inarmónico y por lo tanto, se traduce una amplia paleta de timbres brillantes y vibrantes, incluyendo tañidos de campana y similares.

La amplitud de cada parcial $f_c \pm nf_m$ es $J_n(I)$ donde J_n es la función de Bessel de orden n, de modo que el espectro puede describirse analíticamente por:

$$Y(f) = A_c \cdot \sum_{n=-\infty}^{+\infty} J_n(I) \cdot \delta(f_c \pm nf_m)$$

Si aplicamos la transformada inversa de fourier, podemos escribir la función de la siguiente manera:

$$y(t) = A_c \cdot \sum_{n=-\infty}^{+\infty} J_n(I) \cdot sin(/\omega_c \pm n\omega_m) \cdot t)$$

Podemos ver como, para I =0, es decir sin modulación, la portadora tiene toda la energía y no hay parciales. Conforme I aumenta, la portadora pierde fuerza y aumenta la energía de los parciales. Es posible estimar la cantidad de parciales que serán audibles para un valor dado de I es I +1, donde I se redondea al entero más cercano. Para generalizar, podemos decir que a medida que I aumenta, mayor cantidad de frecuencias serán audibles.

1.3 Síntesis del clarinete

Aplicando lo explicado anteriormente, para sintetizar el sonido de un clarinete, basta con elegir adecuadamente sus funciones A(t) e I(t) y determinar la relación entre las frecuencias de la señal modulada y de la portadora.

Para el clarinete notamos que la mejor relación entre las frecuencias modulada y portadora es $\frac{f_c}{f_m} = \frac{3}{1}$.

La función A(t) es un poco más compleja. Si la dividimos en tres partes, podemos distinguir que la primera parte, más conocida como ataque, tiene forma de exponencial creciente. La segunda parte, conocida como sostén, es de un valor constante; y la tercera parte, conocida como soltar (release), es una exponencial decreciente. Esa combinación de funciones nos otorga la mejor A(t) posible para sintetizar un clarinete. Variando los parámetros de la exponencial y el valor de A_0 obtenemos distintos tipos de clarinetes.

La función I(t) por el contrario posee dos partes: la primera una exponencial creciente, y la segunda posee un valor constante I_0.

La frecuencia fundamental resultante, queda dada por el máximo común divisor entre la frecuencia modulada y la portadora.

Podemos ver la forma de la señal y su respectivo espectro en la siguiente figura:

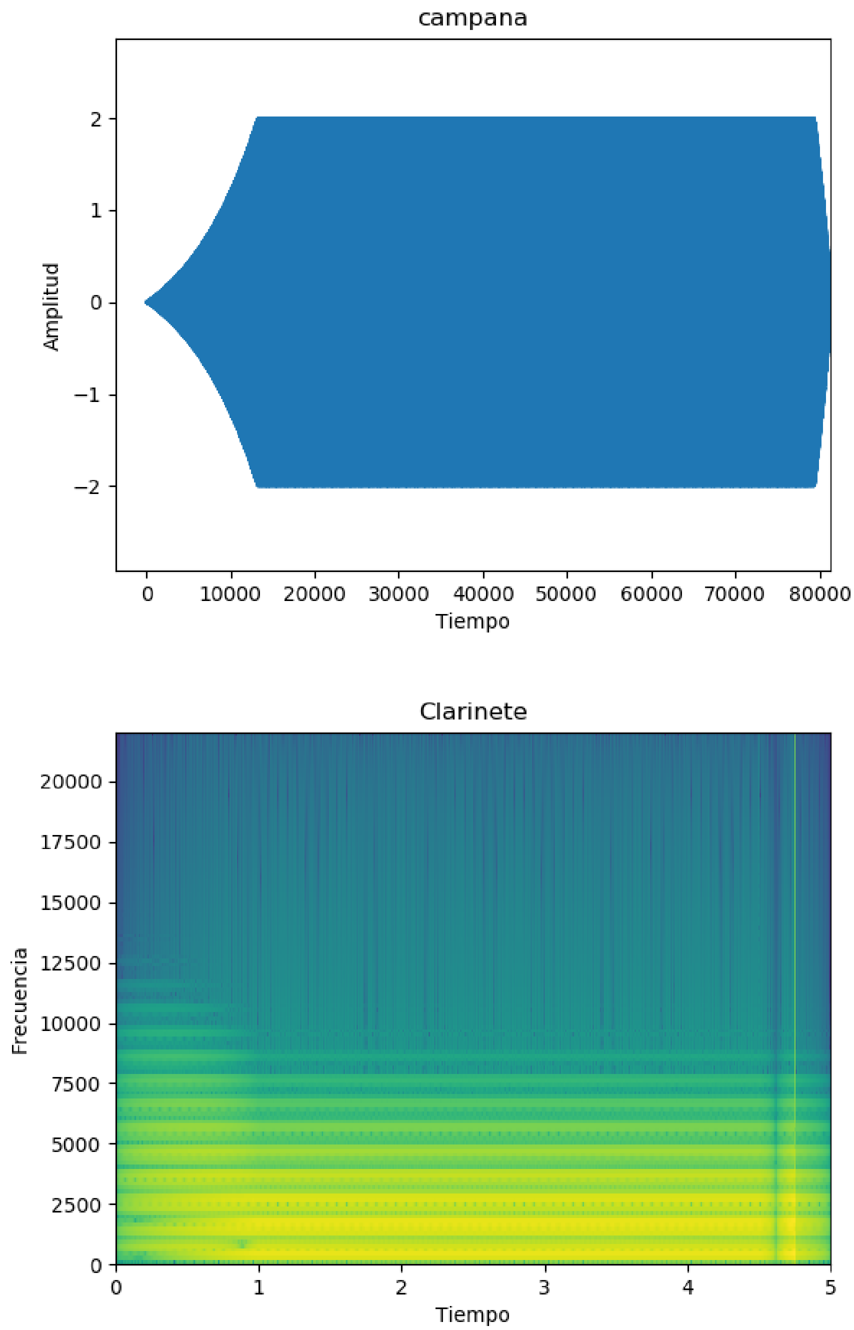


Figura 2: Gráfico en el tiempo y espectrograma del sonido generado por el clarinete.

1.4 Síntesis de la campana

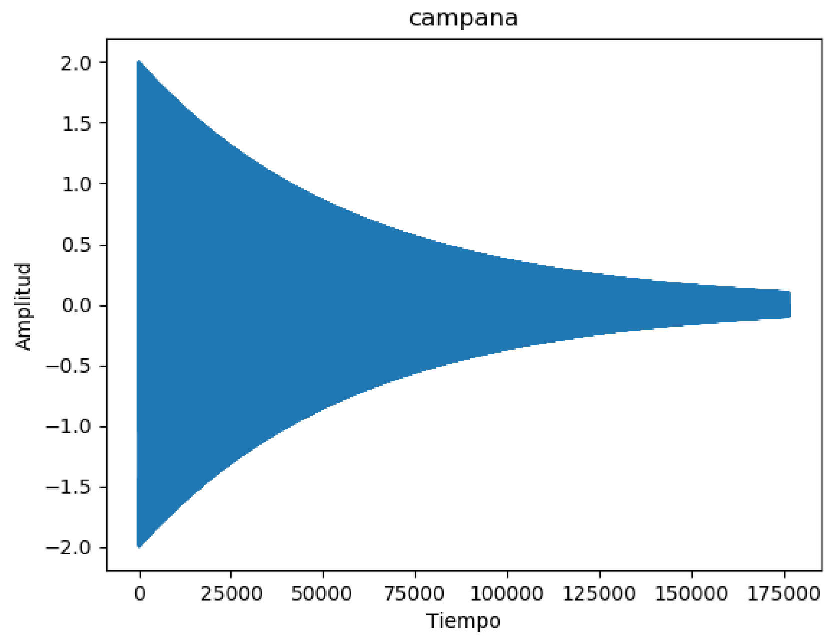
Para lograr el sonido que emite una campana, basta con establecer una relación de 5:7 entre frecuencias portadora y modulada. Las funciones A(t) e I(t) tienen la misma forma de exponencial negativa de la siguiente manera:

$A(t) = I(t) = C \cdot e^{-t/\tau}$

donde C es la A_0 o I_0 según el caso y τ es el parámetro que controla el tiempo de decaimiento. A distintos A_0, I_0 y τse obtienen distintos sonidos de campanas.

La frecuencia fundamental del sonido de la campana es $f_c/5$.

A continuación observamos el diagrama en tiempo y espectrograma de la campana:



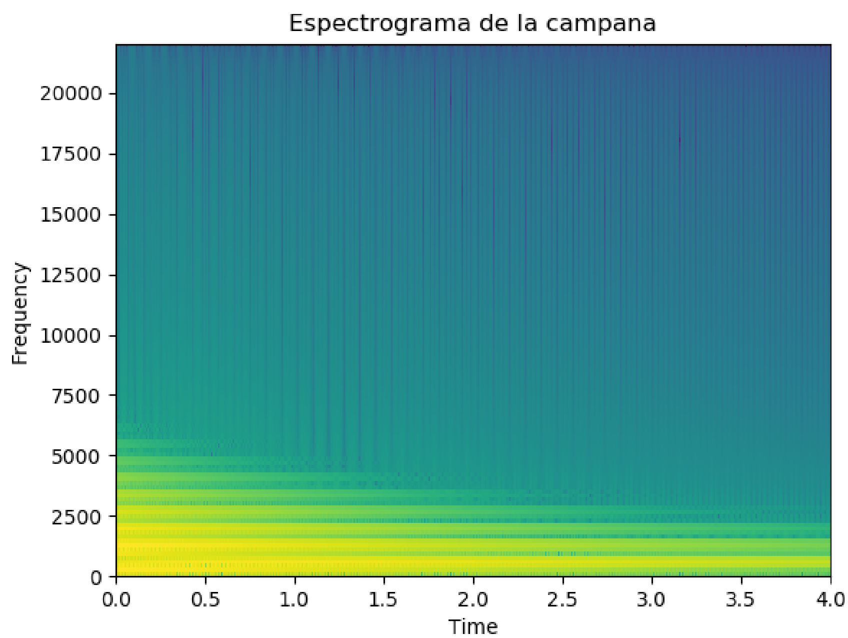


Figura 3: Gráfico en el tiempo y espectrograma del sonido generado por la campana

1.5 Síntesis de otros instrumentos

A partir de los dos instrumentos previamente explicados, cambiando sus parámetros y/o su rango de operación en frecuencias, es posible sintetizar otros instrumentos de viento. Sin embargo, el más llamativo de todos es el que se logra cuando la relación entre las frecuencias modulada y portadora del clarinete es de 2:3. Con esa configuración, es posible lograr el efecto de un violín. El espectrograma resultante es el siguiente:

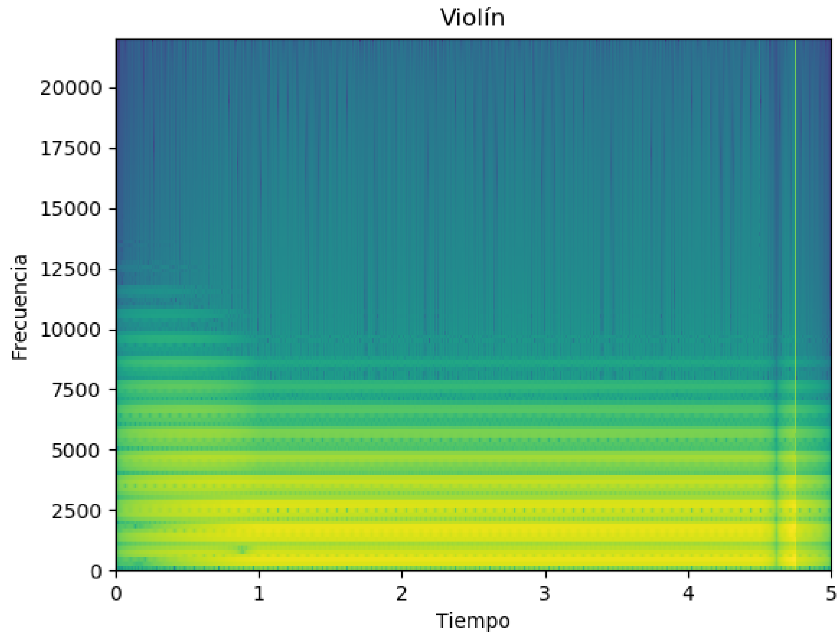


Figura 4: Gráfico en el tiempo y espectrograma del sonido generado por el violín

1.6 Conclusiones

La síntesis de instrumentos musicales mediante modulación de frecuencias es uno de los modelos más sencillos tanto de interpretar como de realizar. El sonido que se obtiene es claro y similar a su modelo real, aunque sufre alteraciones en términos de calidad de sonido cuando se pretende realizar sintetizaciones en todo el rango audible, ya que los instrumentos realizados no son fieles a su modelo real en todo ese rango.