Clase 1: Introducción al Análisis Espectral

Antes de comenzar a ver las herramientas complementarias al audio que nos ayudarán a analizar lo que escuchamos en forma objetiva, haremos una breve introducción a algunos temas relacionados con la matemática y la física del sonido, los cuales retomaremos en el módulo de Acústica. No es el objetivo del curso profundizar en complicados cálculos y algoritmos. Sin embargo, analizaremos algunos conceptos claves que nos serán de gran utilidad para comprender el comportamiento de algunos fenómenos físicos y su forma de medirlos.

Señales periódicas

Son aquellas que cada un cierto período 2T vuelven a repetir su patrón o comportamiento. La función senoidal (que responde a un seno o coseno) es un tipo de señal periódica, y será nuestro objeto de estudio de ahora en adelante.

Características de la Función Senoidal (Figura 1)

Amplitud: Máximo valor en módulo (osea, tanto en valor positivo como negativo) que puede alcanzar la función.

Período: es el tiempo medido en *t segundos* que demora en completar un ciclo.

Frecuencia: Es el número de veces por segundo que se repite un ciclo. Por ende, es la inversa del período.

Fase: corresponde al desplazamiento temporal del inicio del ciclo, expresado en radianes.

Función senoidal

Responde a la forma

$$a(t) = A_0 \cdot \sin(\omega t + \beta)$$

Donde A_0 es la amplitud máxima de la onda, **t** es el tiempo, **ω** es la pulsación $2\pi f$ y $\boldsymbol{\beta}$ es la fase inicial.

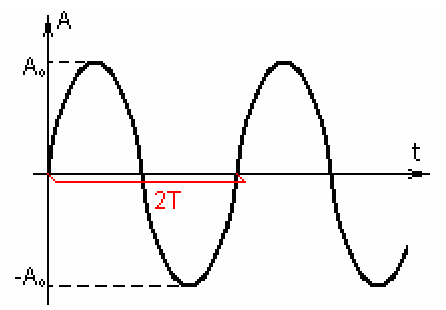


Fig. 1 – Onda Senoidal

Timbres

Los múltiplos de una onda senoidal también son conocidos como armónicos. Los timbres están formados por cantidades variables de armónicos o parciales que cambian a lo largo del tiempo con respecto a un tono o frecuencia fundamental. Los parciales son las ondas que complementan a la onda fundamental para crear un timbre, donde a las frecuencias de los parciales que son múltiplos enteros de la frecuencia fundamental se las llama *parciales armónicos*, y a las que no son múltiplos se las llama *no armónicos*. Gracias a los parciales y a la envolvente dinámica, logramos distinguir a un instrumento de otro, aún cuando están ejecutando una misma nota musical.

Espectro

El espectro de un sonido es la información sobre la distribución de las frecuencias que lo componen y sus respectivas amplitudes. En la Figura 2 podemos observar el espectro de una hinchada de fútbol gritando un gol.

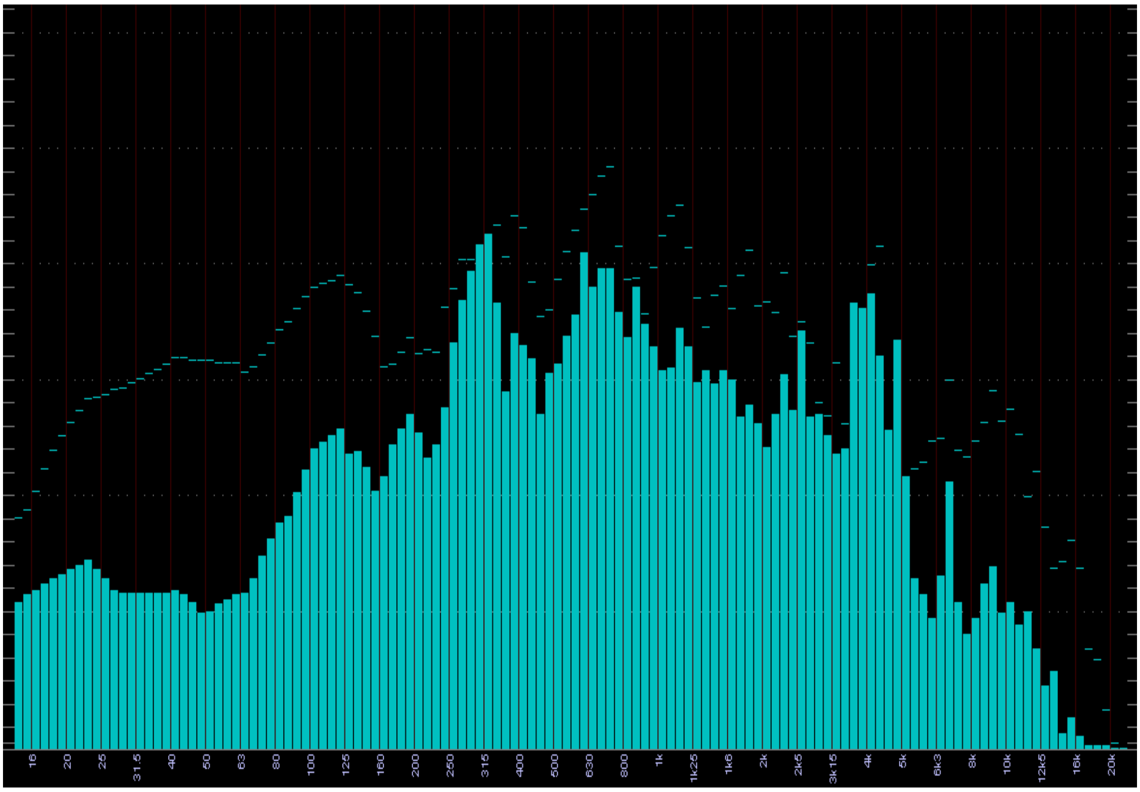


Fig. 2 – Análisis espectral de tipo RTA con AVG fast

Como era de esperarse, no hay una única frecuencia tonal, sino la superposición de una infinidad de frecuencias.

A continuación veremos la forma de poder descomponer a un sonido en cada una de esas frecuencias que lo constituyen.

Serie de Fourier

La *serie de Fourier* es una serie infinita que converge puntualmente a una función *continua y periódica*. Las series de Fourier constituyen la herramienta matemática básica del análisis de Fourier empleado para analizar funciones periódicas a través de la descomposición de dicha función en la suma de infinitas funciones senoidales.

Teorema de Fourier

"Toda función periódica de período 2T puede descomponerse en una suma de sinusoides armónicas, de amplitudes y fases adecuadas, cuya fundamental o primer armónico posea período 2T"

Para investigar un poco…

Las series de Fourier responden a la siguente forma:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos(nx) + b_n \sin(nx)\right]$$

Donde $\mathbf{a_n}$ y $\mathbf{b_n}$ se denominan *coeficientes de Fourier* de la serie de Fourier de la función **f(x)**.

Si ***f*** es una función (o señal) periódica y su período es **2*T***, la serie de Fourier puede expresarse como:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\frac{n\pi}{T}t + b_n \sin\frac{n\pi}{T}t\right]$$

Donde los coeficientes de Fourier toman los valores:

$$a_n = \frac{1}{T}\int_{-T}^{T} f(t)\cos(\frac{n\pi}{T}t)dt, \qquad b_n = \frac{1}{T}\int_{-T}^{T} f(t)\sin(\frac{n\pi}{T}t)dt, \qquad a_0 = \frac{1}{T}\int_{-T}^{T} f(t)dt$$

Si bien no vamos a profundizar mucho en el aspecto matemático, es importante saber que mediante la Serie de Fourier podemos estudiar el comportamiento de cualquier función periódica, pero rara vez en nuestro campo tengamos como objeto de estudio a este tipo se señales. Es por ello que a la música no es posible estudiarla mediante su desarrollo en Series de Fourier.

Transformada de Fourier

Para las señales no periódicas, la música, etc., poseemos otras poderosas herramientas, como las Integrales de Fourier. De ellas deriva la Transformada de Fourier, la cual responde a la fórmula

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{-i\omega x}$$

Mediante la Transformada de Fourier podemos descomponer a las señales en cada una de sus frecuencias que la componen.

Sin profundizar mucho más, diremos que de la Transformada de Fourier llegamos a la Transformada Discreta (conocida como DFT, Discrete Fourier Transform), descrita a continuación:

$$f_j = \sum_{k=0}^{n-1} x_k e^{-\frac{2\pi i}{n} jk} \qquad j=0,\ldots, n-1$$

Transformada Rápida de Fourier

Finalmente estamos en condiciones de hablar del análisis espectral propiamente dicho. La herramienta fundamental del análisis espectral es la Transformada Rápida de Fourier, más conocida como FFT (Fast Fourier Transform). La FFT es un algoritmo que nos permite calcular en forma muy eficiente la DFT y su inversa. Es el arma principal del manejo de señales digitales y filtros de este tipo. Hay muchos algoritmos que permiten calcularla (no vamos a profundizar al respecto, pero se puede encontrar una rápida y didáctica explicación de uno de ellos en http://www.dspguide.com/ch12/2.htm), pero todos coinciden en que permiten realizar un número N de muestras donde N es potencia de 2 (512, 1024, 2048, 4096 u 8192 muestras generalmente).

Aplicaciones

Vamos a analizar una señal periódica muy común, conocida como *diente de sierra descendente* (Figura 3).

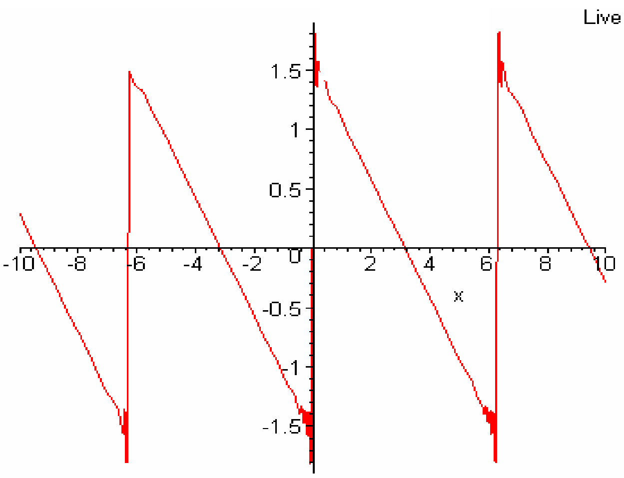


Fig. 3 – Diente de Sierra descendente

La misma corresponde a la siguiente sumatoria $\sum_{i=1}^{\infty}\frac{\sin(x\cdot i)}{i}$. Eso significa que es la sumatoria de infinitos senos aumentan su frecuencia y disminuyen su amplitud a medida que ***i*** crece. Esa descomposición en senos corresponde a la serie de Fourier de esta función. Para ver más claramente el aporte de cada término, vamos a tomar solamente los primeros tres senos presentes (Figura 4):

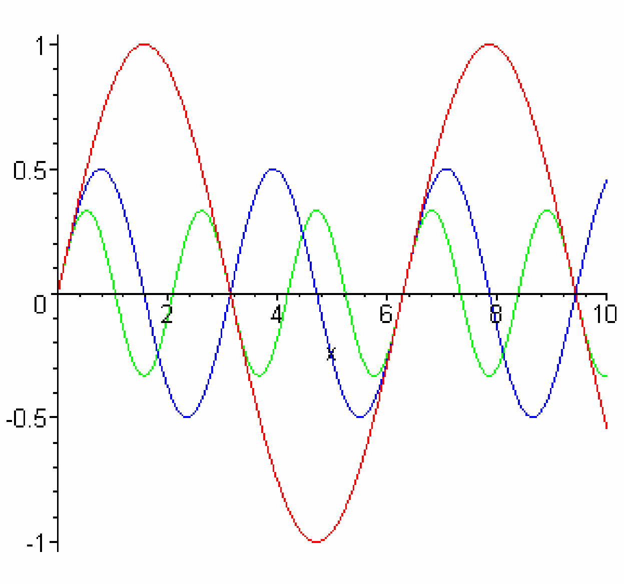


Fig. 4 – Primeras tres sinusoides

Si sumamos estas tres sinusoides, obtendremos como resultado:

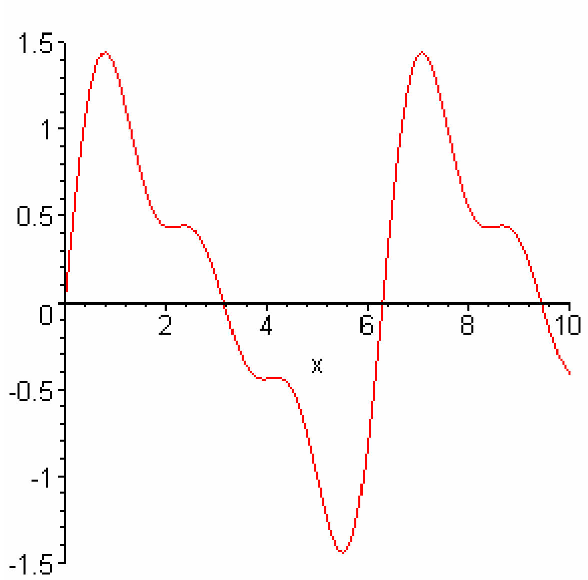


Fig. 5 – Suma de las primeras tres sinusoides

Vemos que si seguimos con la sumatoria, convergerá a la señal diente de sierra descendente.

Otra forma de ver la descomposición en sinusoides del análisis por series de Fourier es aplicando la transformada de Fourier, donde obtendremos la frecuencia de cada una de estas sinusoides y sus respectivas amplitudes (Figura 6). Finalmente, hemos hecho un análisis espectral de la señal.

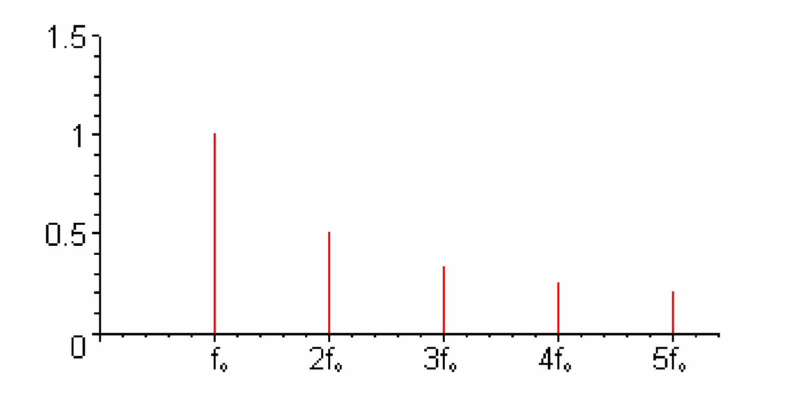


Fig. 6 – FFT de la función diente de sierra descendente

Funciones Ventana

Son funciones matemáticas utilizadas generalmente en el análisis y procesamiento de señales, que nos permite aislar una porción de la señal a analizar y desechar el resto en mayor o menor medida. Serán de gran importancia para los analizadores de espectro, ya que cada banda a estudiar estará definida por una ventana distinta que dejará pasar sólo las frecuencias comprendidas dentro de esa banda en cuestión. Para obtener la información en dicha banda, simplemente multiplicamos a la señal por la función ventana. Deberemos tener en cuenta que una señal real dura un determinado tiempo t. Para realizar esta multiplicación, la señal deberá durar un tiempo finito, y el cálculo además será posible a partir de un número finito de puntos. De ahí radican las limitaciones de los analizadores de espectro, que para anchos de banda muy reducidos necesitarán más tiempo para procesar con mayor precisión, dada poca cantidad de puntos. Eso es un problema a la hora de trabajar a tiempo real, como veremos más adelante.

Cuando multiplicamos a una señal $s(t)$ por una ventana $h(t)$, se obtienen solamente los T primeros segundos (el intervalo [0,T]). Considerando entonces a la fución $s_h(t) = s(t) \cdot h(t)$, pasamos al dominio de la frecuencia mediante la transformada de Fourier, y obtenemos la convolución $S_h(f) = S(f) * H(f)$, donde $H(f)$ es la transformada de Fourier de la ventana (Figura 7).

<u>Dominio del Tiempo</u> <u>Dominio de la Frecuencia</u>

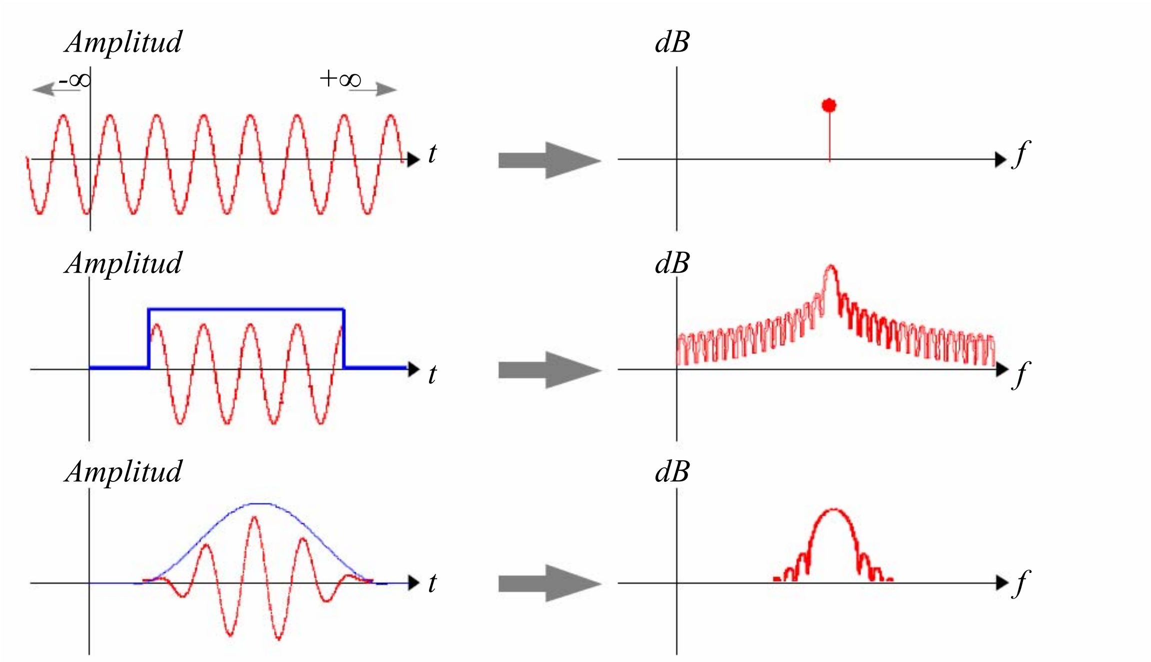
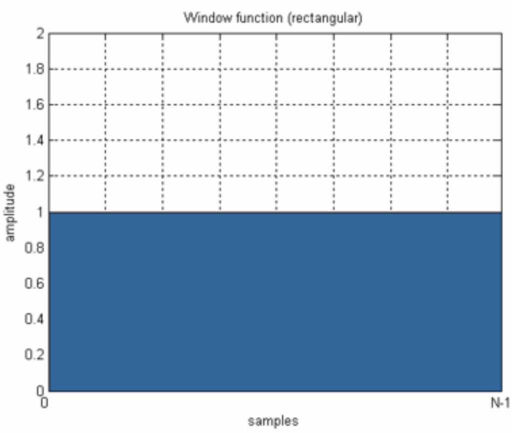


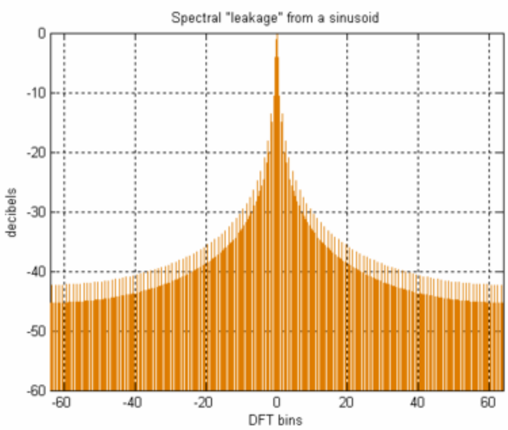
Fig. 7 – Señales senoidales en el dominio del tiempo y la frecuencia, convolucionadas con ventanas infinitas, rectangular y Hamming, respectivamente

Hay diferentes tipos de ventanas. En ellas, ***N*** simboliza el ancho, en muestras, de la función ventana en tiempo discreto. Generalmente son números enteros potencias de 2. n también es un número entero, el cual toma los valores $0 \le n \le N-1$. Así que estas son las formas del tiempo pasado de las ventanas: $v(n) = v_0(n - \frac{N-1}{2})$ donde ***v₀(n)*** es máximo cuando ***n=0***.

- **Ventana Rectangular**

Es la más simple de todas, y también se la conoce como Ventana de Dirichlet

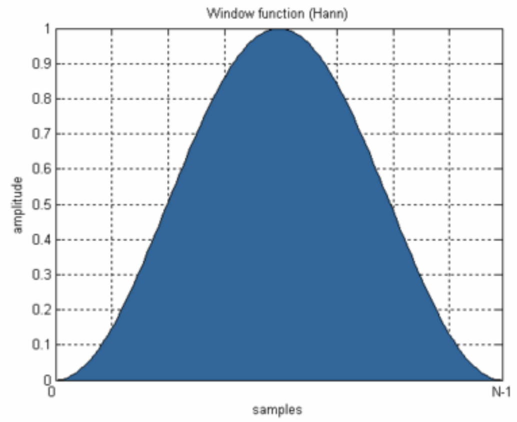


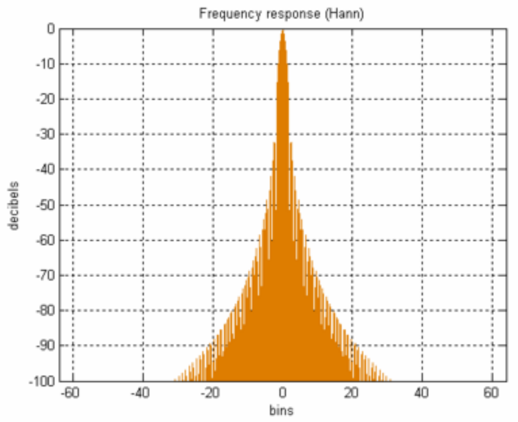


$v(n) = 1$

- **Hann**

Las ventanas de Hann y de Hamming son de la familia de las ventanas de "cosenos elevados", y comúnmente son confundidas una con la otra.

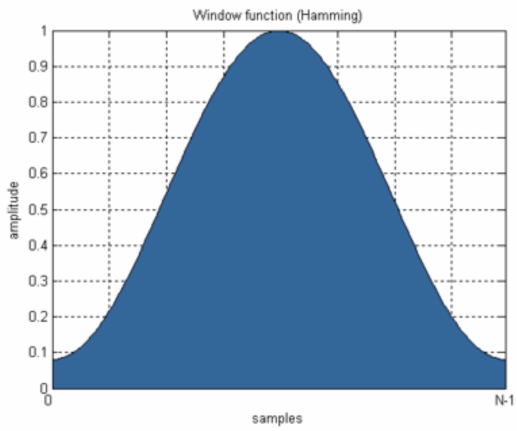


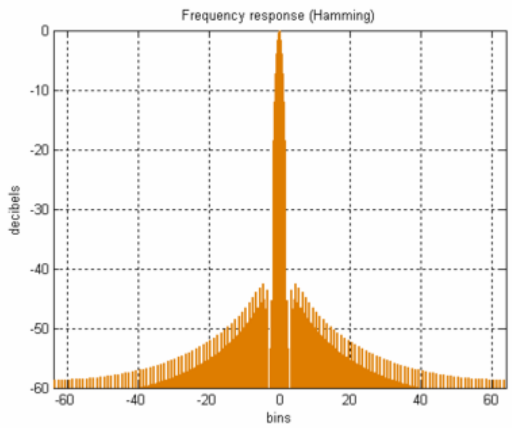


$$v(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{N-1}\right)$$

- **Hamming**

La altura máxima de los lóbulos laterales es de alrededor de un quinto del de la ventana de Hann, un coseno elevado con coeficientes más simples.

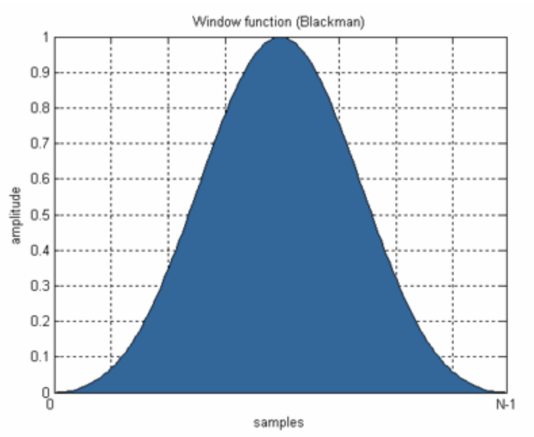


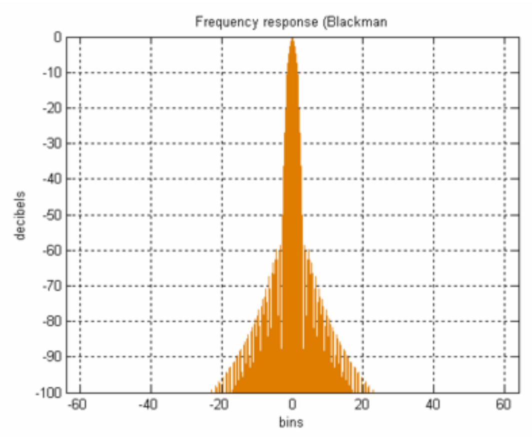


$$v(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{N-1}\right)$$

$$a_0 = 0{,}53836; \quad a_1 = 0{,}46164$$

- **Blackman**

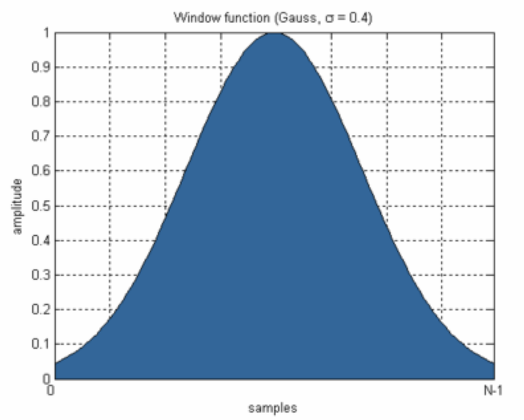


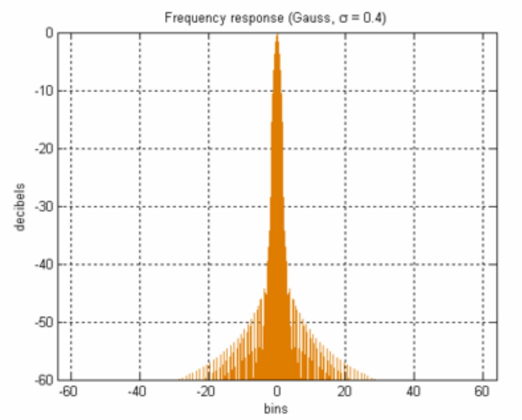


$$v(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{N-1}\right) + a_2 \cos\left(\frac{4\pi n}{N-1}\right)$$

$$a_0 = 0{,}42; \quad a_1 = 0{,}5; \quad a_2 = 0{,}08;$$

- **Gaussiana**





$$v(n) = e^{-\frac{1}{2}\left(\frac{n-(N-1)/2}{\sigma(N-1)/2}\right)^2}$$

$$\sigma \leq 0{,}5$$

Existen muchas otras ventanas, y cada una poses sus características particulares. Si bien para el audio, a grandes rasgos, cualquiera puede ser de utilidad, para realizar mediciones que precisan de gran exactitud, no dará lo mismo cualquier elección. La ventana de Hamming posee buena resolución tanto en frecuencia como en amplitud (aunque a bajas amplitudes puede dificultarse su interpretación en aplicaciones de audio). La ventana de Blackman posee buena resolución en frecuencia, pero regular resolución en amplitud, y se la suele utilizar para medir distorsiones.

Bibliografía

Basso, Gustavo. Análisis Espectral, La trasformada de Fourier en la Música, Ed. Al Margen, 2001.

Henderson, Pau D. FFT Fundamentals. Smaart Docs, 2004.

Miyara, Federico. Acústica y Sistemas de Sonido, UNR Editora, 2006.

Salvioli, Sahra. Matemáticas Especiales, Ed. CEILP, 2005.

Sorín, Saúl. Teoría y Práctica del Decibel, Ed. Bell S.A., 1980.

Wikipedia, Free Encyclopedia, 2009.