

# Espectrograma

## Análisis Tiempo–Frecuencia

IIE

<sup>1</sup>Facultad de Ingeniería Universidad de la República

October 2, 2023

- 1 El Espectrograma
- 2 Análisis de Fourier de tiempo corto
  - Efecto del enventanado

# Contenido

- 1 El Espectrograma
- 2 Análisis de Fourier de tiempo corto
  - Efecto del enventanado

# Elección de la Ventana

Criterios:

- Ancho del lóbulo principal.
- Caída del peso de los lóbulos secundarios.
- Peso máximo relativo de lóbulos secundarios.

Otras ventanas:

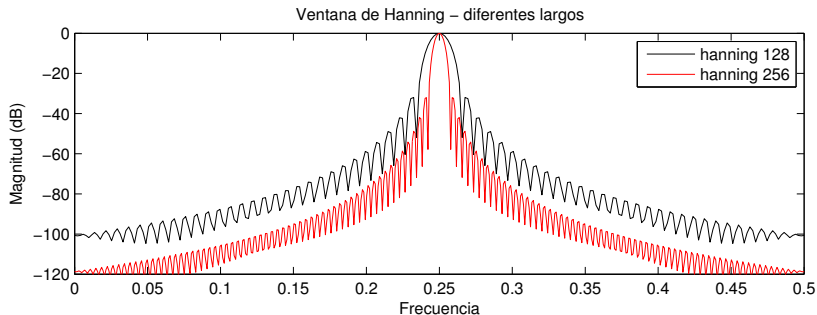
- Hamming
- Hann
- Blackman.

# Análisis de Fourier de tiempo corto (STFT)

## Efecto del enventanado

Propiedades de la ventana:

- ancho del lóbulo principal: inversamente proporcional al largo  $L$

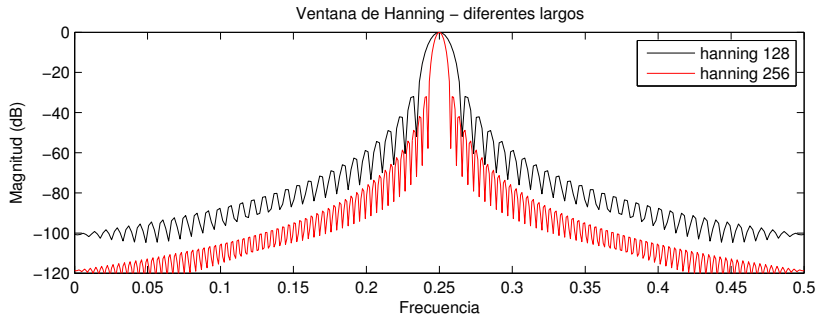


# Análisis de Fourier de tiempo corto (STFT)

## Efecto del enventanado

Propiedades de la ventana:

- ancho del lóbulo principal: inversamente proporcional al largo  $L$
- nivel de lóbulos secundarios: independiente del largo

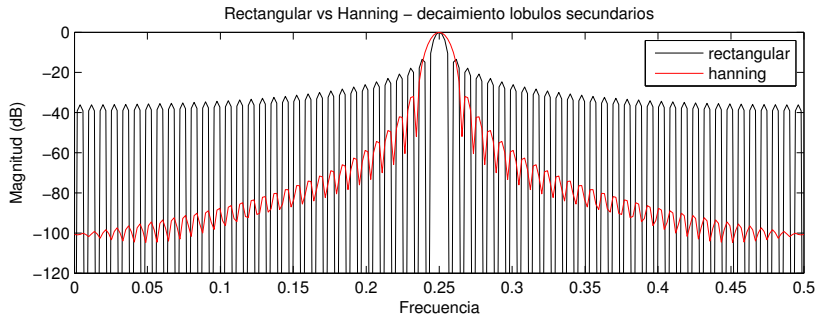


# Análisis de Fourier de tiempo corto (STFT)

## Efecto del enventanado

Propiedades de la ventana:

- ancho del lóbulo principal: inversamente proporcional al largo  $L$
- nivel de lóbulos secundarios: independiente del largo depende del tipo de ventana

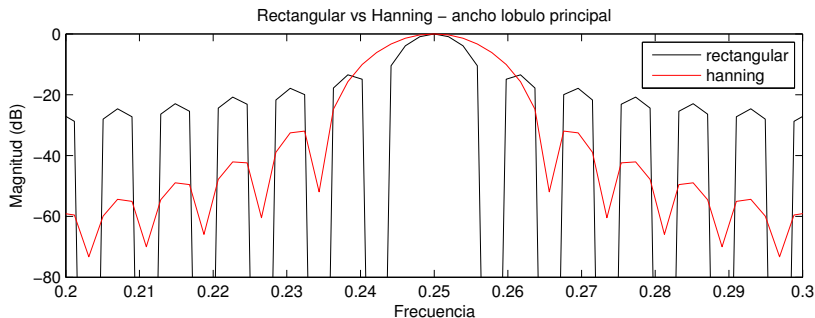


# Análisis de Fourier de tiempo corto (STFT)

## Efecto del enventanado

### Propiedades de la ventana:

- ancho del lóbulo principal: inversamente proporcional al largo  $L$
- nivel de lóbulos secundarios: independiente del largo depende del tipo de ventana
  - rectangular:  $-13\text{dB}$ ,  $2\frac{f_s}{L}$     hanning:  $-31\text{dB}$ ,  $4\frac{f_s}{L}$

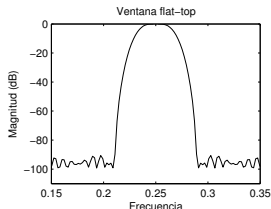
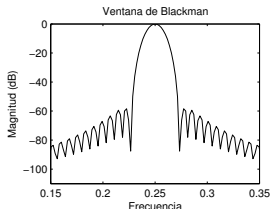
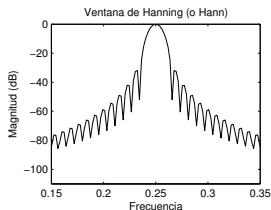
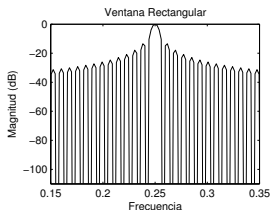




# Análisis de Fourier de tiempo corto (STFT)

## Efecto del enventanado

- compromiso entre ancho lóbulo principal y nivel lóbulos secundarios

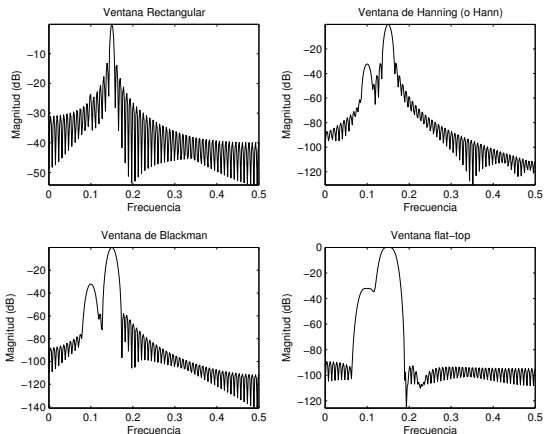


# Análisis de Fourier de tiempo corto (STFT)

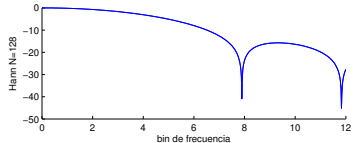
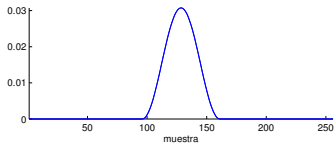
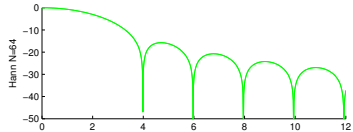
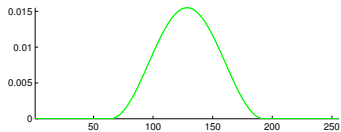
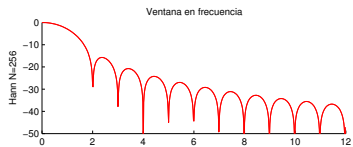
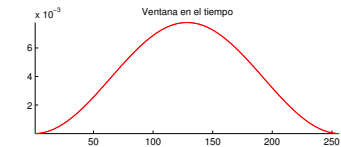
## Efecto del enventanado

- compromiso entre ancho lóbulo principal y nivel lóbulos secundarios

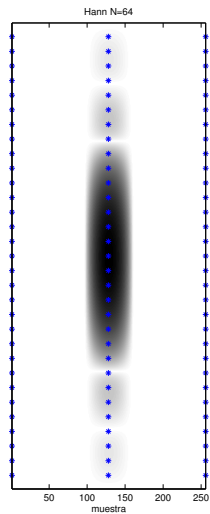
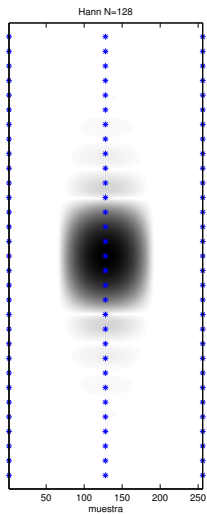
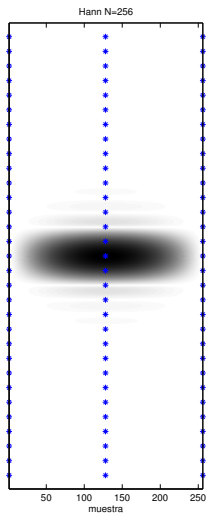
ejemplo: análisis de frecuencias cercanas (0.1 y 0.15)



# Elección de la Ventana



# Elección de la Ventana

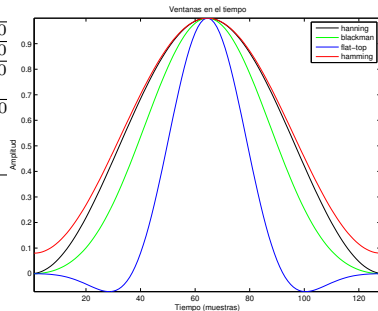
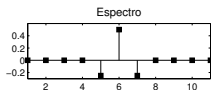
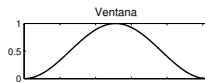


## Análisis de Fourier de tiempo corto (STFT)

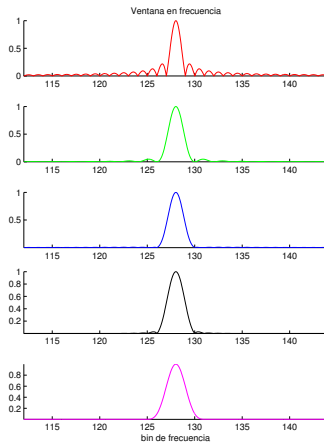
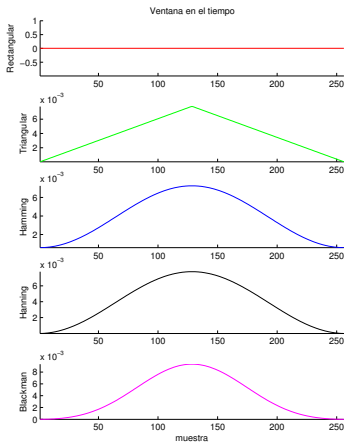
## Efecto del enventanado

ventanas típicas: unos pocos componentes en frecuencia no nulos

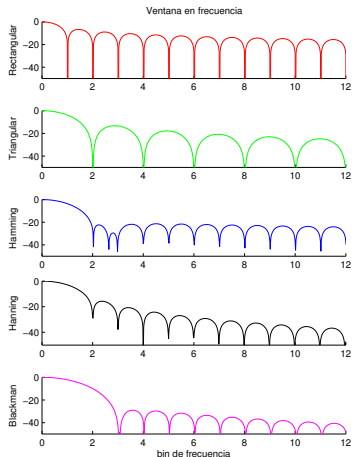
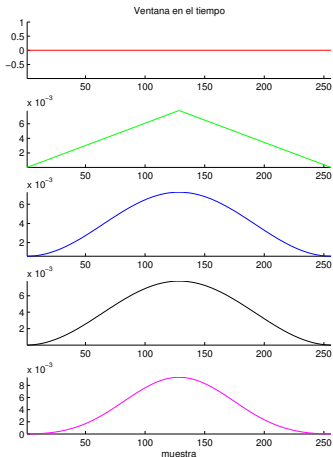
|          |  |   |
|----------|--|---|
| hann     | $a_0 - a_1 \cos(2\pi n/L)$   | $a_0 = 0.50, a_1 = 0$   |
| hamming  | $a_0 - a_1 \cos(2\pi n/L)$   | $a_0 = 0.54, a_1 = 0$   |
| blackman | $a_0 - a_1 \cos(2\pi n/L)$<br>$+ a_2 \cos(4\pi/L)$   | $a_0 = 0.42, a_1 = 0$<br>$a_2 = 0.08$   |
| flat-top | $a_0 - a_1 \cos(2\pi n/L)$<br>$+ a_2 \cos(4\pi/L)$<br>$- a_3 \cos(6\pi/L)$<br>$+ a_4 \cos(8\pi/L)$ | $a_0 \approx 0.22, a_1 \approx 0$<br>$a_2 \approx 0.28$<br>$a_3 \approx 0.08$<br>$a_4 \approx 0.01$ |



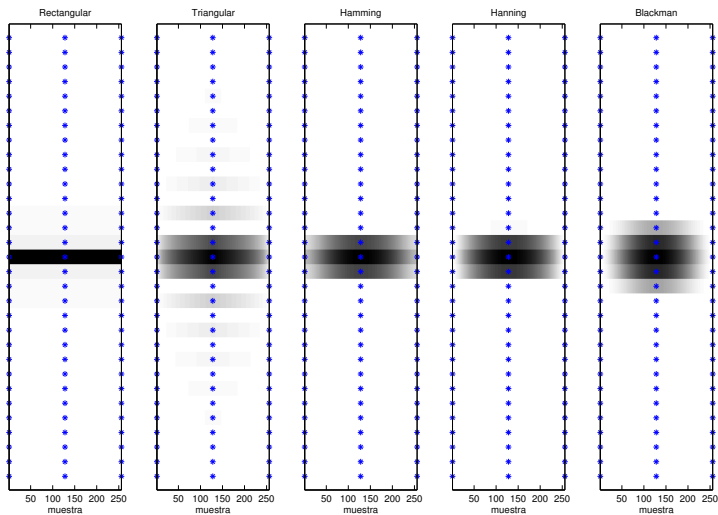
# Elección de la Ventana



# Elección de la Ventana

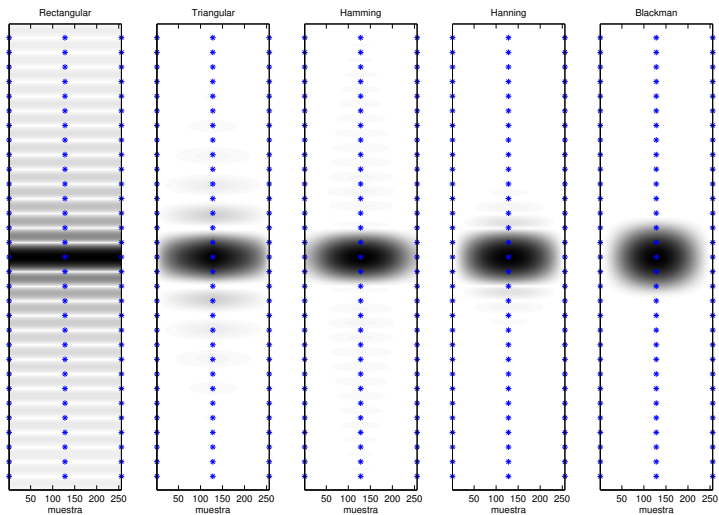


# Elección de la Ventana





# Elección de la Ventana



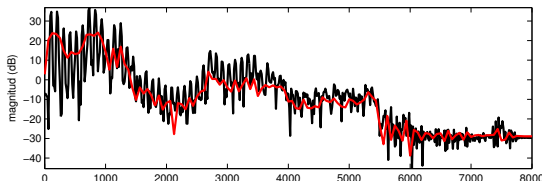
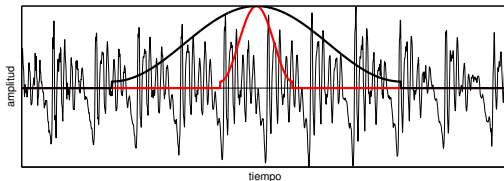
# Análisis de Fourier de tiempo corto (STFT)

## Efecto del inventanado

- Análisis de voz usando diferente largo de ventana

64 ms estructura armónica clara

16 ms sólo se distinguen las formantes



# Análisis de Fourier de tiempo corto (STFT)

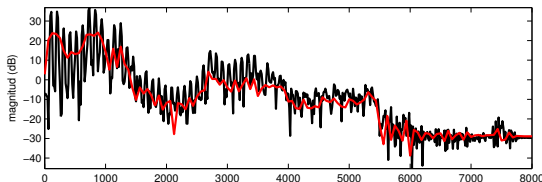
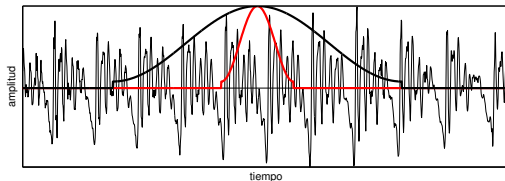
## Efecto del inventanado

- Análisis de voz usando diferente largo de ventana

64 ms estructura armónica clara

16 ms sólo se distinguen las formantes

pero las formantes pueden cambiar a lo largo de 50 ms



# Análisis de Fourier de tiempo corto (STFT)

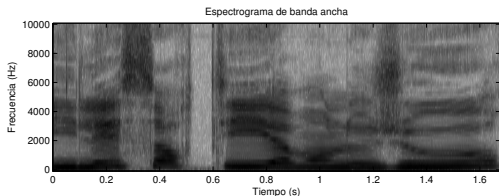
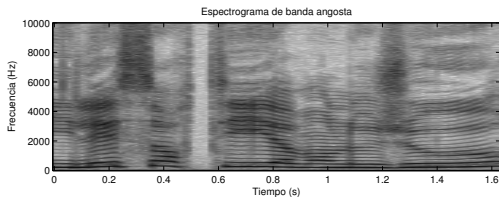
## Efecto del inventanado

- Análisis de voz usando diferente largo de ventana

64 ms estructura armónica clara

16 ms sólo se distinguen las formantes

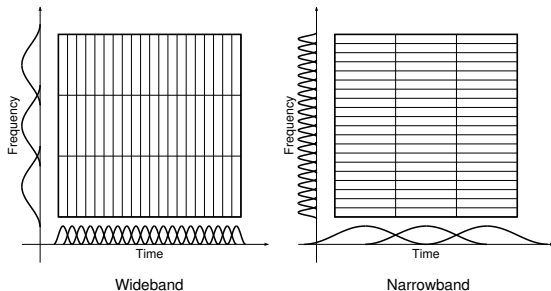
pero las formantes pueden cambiar a lo largo de 50 ms



# Análisis de Fourier de tiempo corto (STFT)

## Espectrograma

- largo  $L$  (y forma) de la ventana determina la resolución
  - espectrograma de *banda ancha* ( $L$  chico)
  - espectrograma de *banda angosta* ( $L$  grande)
- resolución constante en tiempo-frecuencia



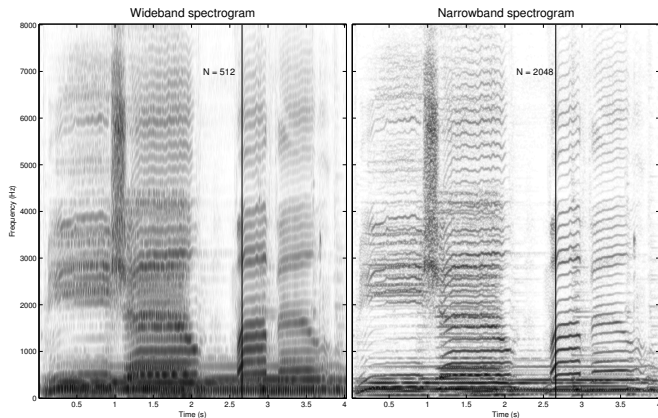
# Análisis de Fourier de tiempo corto (STFT)

## Banda ancha

- pobre resolución espectral
- buena resolución temporal

## Banda angosta

- buena resolución espectral
- pobre resolución temporal



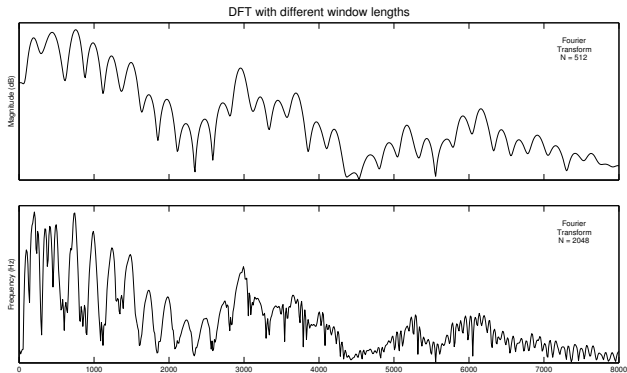
# Análisis de Fourier de tiempo corto (STFT)

## Banda ancha

- pobre resolución espectral
- buena resolución temporal

## Banda angosta

- buena resolución espectral
- pobre resolución temporal



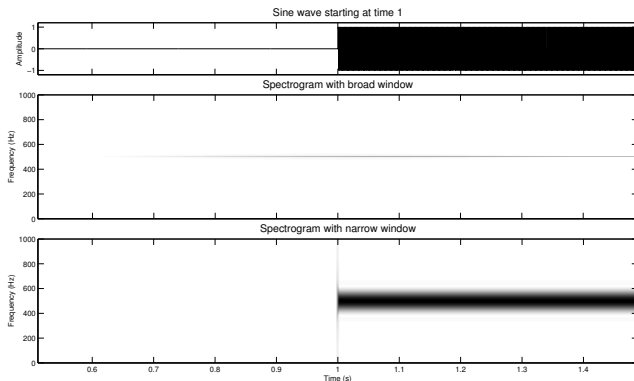
# Análisis de Fourier de tiempo corto (STFT)

## Banda ancha

- pobre resolución espectral
- buena resolución temporal

## Banda angosta

- buena resolución espectral
- pobre resolución temporal

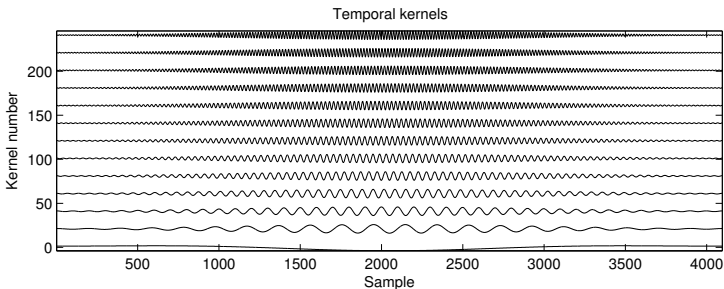




# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

DFT:

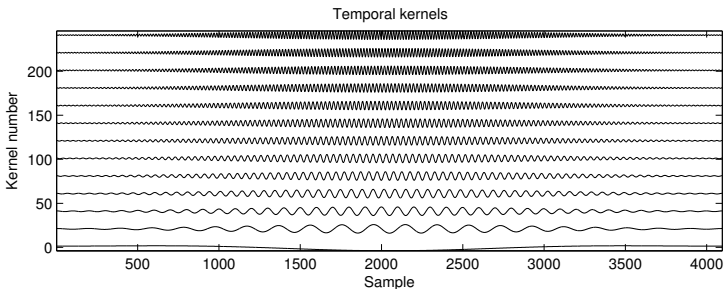


# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

DFT:

$$X[k] = \sum_{n=0}^{N-1} w[n]x[n]e^{-j2\pi kn/N}$$



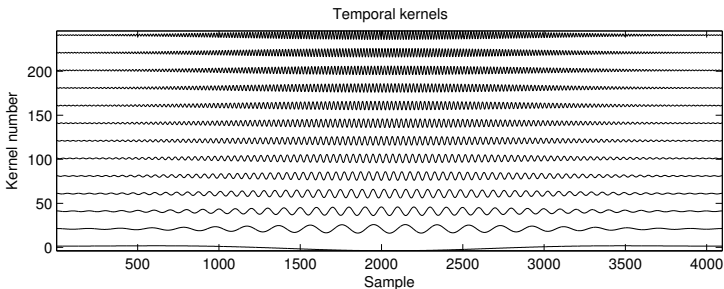
# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

DFT:

$$X[k] = \sum_{n=0}^{N-1} w[n]x[n]e^{-j2\pi kn/N}$$

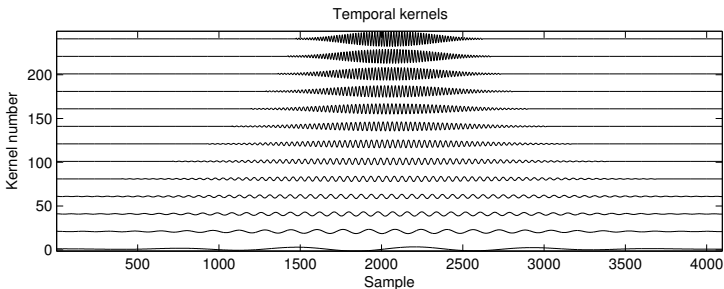
$$Q_k = f_k/\Delta f = k$$



# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

CQT:



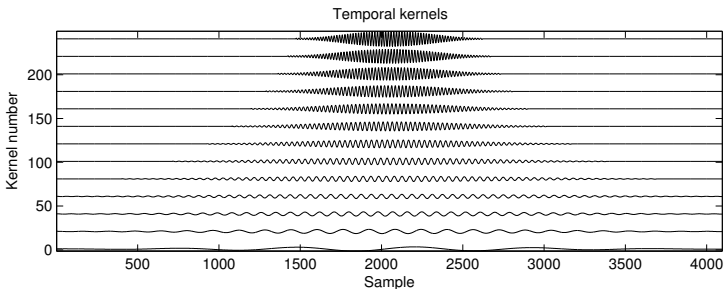
# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

CQT:

$$Q = f_k / \Delta f_k \text{ constante}$$

$$N_k = Q / f_k \text{ variable}$$



# Análisis de Fourier de tiempo corto (STFT)

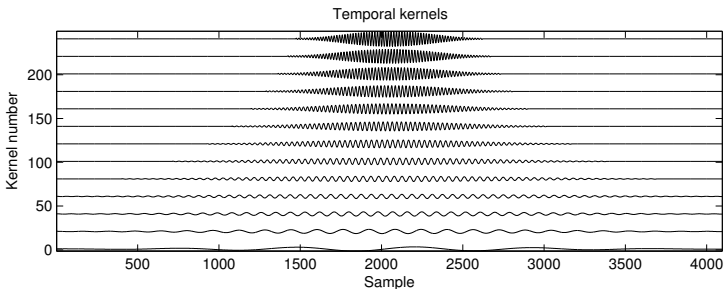
## Constant Q Transform

CQT:

$$Q = f_k / \Delta f_k \text{ constante}$$

$$N_k = Q / f_k \text{ variable}$$

$$X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n] x[n] e^{-j2\pi Qn/N_k}$$



# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT



# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $$X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n] x[n] e^{-j2\pi Qn/N_k}$$

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n]x[n]e^{-j2\pi Qn/N_k}$
- $X^{cq} = x \cdot T^*$ , multiplicación matricial con **kernels temporales**:

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n]x[n]e^{-j2\pi Qn/N_k}$
- $X^{cq} = x \cdot T^*$ , multiplicación matricial con **kernels temporales**:

$$T^*[n, k] = \begin{cases} \frac{1}{N_k} w_k[n] e^{-j2\pi Qn/N_k} & \text{si } n < N_k \\ 0 & \text{en otro caso} \end{cases}$$

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n]x[n]e^{-j2\pi Qn/N_k}$
- $X^{cq} = x \cdot T^*$ , multiplicación matricial con **kernels temporales**:

$$T^*[n, k] = \begin{cases} \frac{1}{N_k} w_k[n] e^{-j2\pi Qn/N_k} & \text{si } n < N_k \\ 0 & \text{en otro caso} \end{cases}$$

- usando la relación de Parseval para la DFT:

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n]x[n]e^{-j2\pi Qn/N_k}$
- $X^{cq} = x \cdot T^*$ , multiplicación matricial con **kernels temporales**:

$$T^*[n, k] = \begin{cases} \frac{1}{N_k} w_k[n] e^{-j2\pi Qn/N_k} & \text{si } n < N_k \\ 0 & \text{en otro caso} \end{cases}$$

- usando la relación de Parseval para la DFT:

$$X^{cq}[k] = \sum_{n=0}^{N-1} x[n]T^*[n, k] = \frac{1}{N} \sum_{k'=0}^{N-1} X[k']K^*[k', k]$$

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n]x[n]e^{-j2\pi Qn/N_k}$
- $X^{cq} = x \cdot T^*$ , multiplicación matricial con **kernels temporales**:

$$T^*[n, k] = \begin{cases} \frac{1}{N_k} w_k[n]e^{-j2\pi Qn/N_k} & \text{si } n < N_k \\ 0 & \text{en otro caso} \end{cases}$$

- usando la relación de Parseval para la DFT:

$$X^{cq}[k] = \sum_{n=0}^{N-1} x[n]T^*[n, k] = \frac{1}{N} \sum_{k'=0}^{N-1} X[k']K^*[k', k]$$

dónde  $X[k']$  y  $K[k', \cdot]$  son la DFT de  $x[n]$  y  $T[n, \cdot]$  respectivamente

# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

la evaluación directa de la CQT es computacionalmente costosa  
aproximación eficientemente usando la FFT

- $X^{cq}[k] = \frac{1}{N_k} \sum_{n=0}^{N_k-1} w_k[n]x[n]e^{-j2\pi Qn/N_k}$
- $X^{cq} = x \cdot T^*$ , multiplicación matricial con **kernels temporales**:

$$T^*[n, k] = \begin{cases} \frac{1}{N_k} w_k[n] e^{-j2\pi Qn/N_k} & \text{si } n < N_k \\ 0 & \text{en otro caso} \end{cases}$$

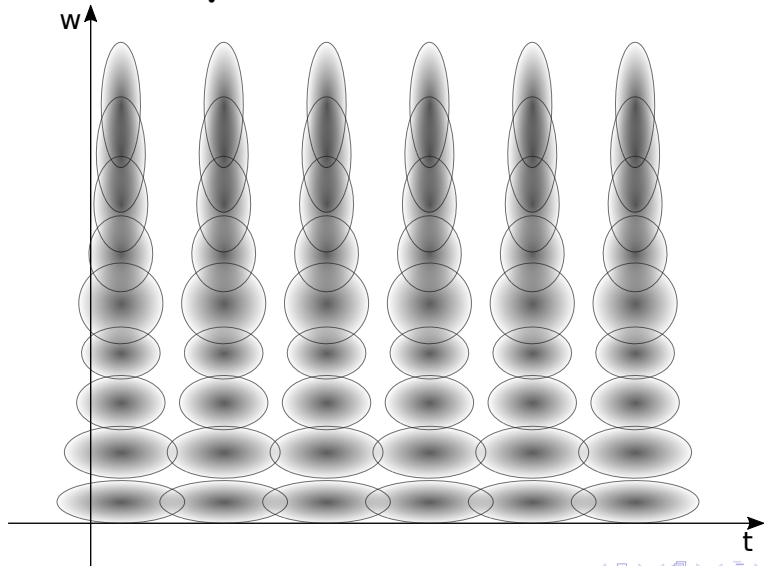
- usando la relación de Parseval para la DFT:

$$X^{cq}[k] = \sum_{n=0}^{N-1} x[n]T^*[n, k] = \frac{1}{N} \sum_{k'=0}^{N-1} X[k']K^*[k', k]$$

dónde  $X[k']$  y  $K[k', \cdot]$  son la DFT de  $x[n]$  y  $T[n, \cdot]$  respectivamente  
 $K[k', \cdot]$  denominados **kernels espectrales**

# Constant Q Transform

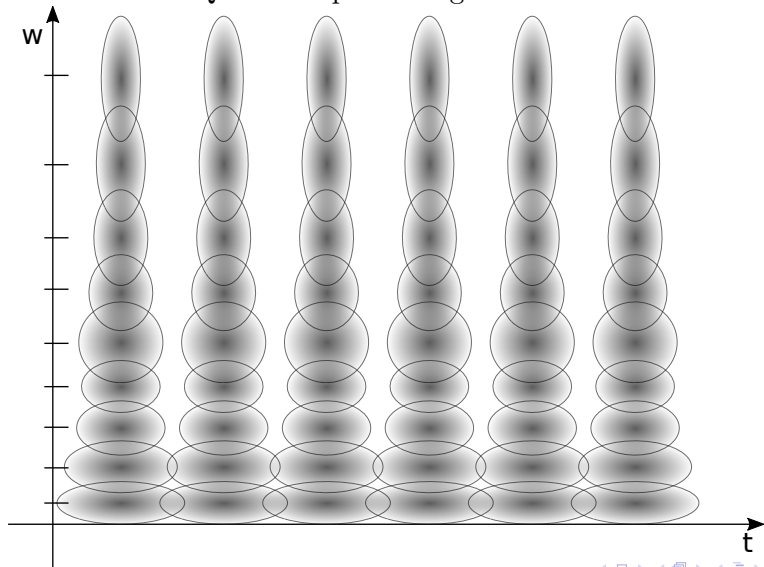
Transformada Q - bins lineales





# Constant Q Transform

Transformada Q - bins espaciados geoméricamente

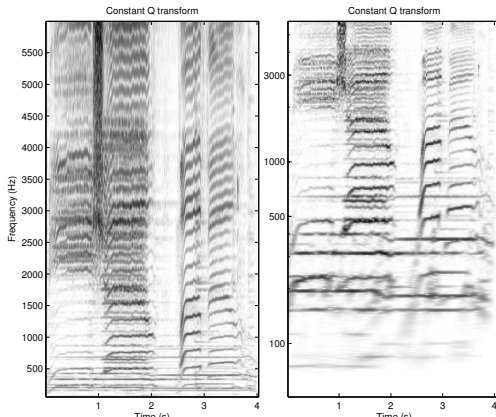


# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

distribución de los bins en frecuencia

- la formulación original de la CQT implica distribución geométrica



# Análisis de Fourier de tiempo corto (STFT)

## Constant Q Transform

distribución de los bins en frecuencia

- la formulación original de la CQT implica distribución geométrica
- se puede formular para cualquier otro espaciado, por ejemplo lineal

