SHORT TIME FOURIER TRANSFORM

- Permite estudiar y analizar señales que varían en el tiempo.
- Transformaciones STFT pueden obtenerse para la CTFT, DTFT y la DFT. Sólo nos interesa la versions en tiempo discreto

$$\text{ST-DTFT}: \quad X(w, n) = \sum_{m=-\infty}^{\infty} x(m)\, w[n-m]\, e^{-jwm}$$

(DTFT que varía en el tiempo) → dado que $\text{length}(w) = N_w$

$w[n]$ es una ventana de largo N_w cuyos elementos se asume que son $\neq 0$

$x(m)\, w[n-m]$ es un porción de $x[m]$ truncada por la ventana (invertida y desplazada)

$X(w,n)$ es la DTFT de $x(m)\, w[n-m]$ para el tiempo n

OBS] • m es también tiempo! La variable n es independiente, i.e, $X(w, n)$ entrega una DTFT para cada n

- ¿Existe overlap entre la $X(w,n)$ y $X(w,n+1)$? Sí!! sólo difieren en 1 punto, es decir tienen overlap de $N-1$ puntos.

ST - DFT : $X(k,n) = \sum_{m=-\infty}^{\infty} x(m)\, w(n-m)\, e^{-j2\pi km/N}$

(DFT que varía en el tiempo) con $N \geq N_w$

$$X(k,n) = X(\omega,n)\Big|_{\omega = \frac{2\pi k}{N}}$$

OBS) $|X(k,n)|$ se grafica en 2D y se conoce como "espectrograma".

- si N_w es mayor que el periodo de la señal $x(n)$ $(N_w > T)$ → Espectrograma de banda angosta
  - ⇝ se pueden visualizar las Frecuencia Fundamental (F_0) y armónicos. ($2F_0, 3F_0, \dots$) (si n esta en el eje x se ven como lineas horizontales)
- si $N_w < T$ ⇒ Espectrograma de banda ancha.
  - ⇝ no se ven $F_0, 2F_0, 3F_0, \dots$ pero si se aprecian los máximos de cada ciclo (si n esta en el eje x se ven como lineas verticales)

- Es posible hacer saltos en las ventanas para evitar que estas tengan overlap en todos los puntos de modo que:

$$X(k, nL) = \sum_{m=-\infty}^{\infty} x[m]\, w[nL-m]\, e^{-j2\pi km/N}$$

- si $L = N_w$ no existe overlap entre cada cuadro.
- si $\displaystyle\sum_{n=-\infty}^{\infty} w[nL-m] = 1 \quad \forall m \in \mathbb{Z}$

entonces...

$$\sum_{n=-\infty}^{\infty} X(\omega, nL) = X(\omega)$$

$\Rightarrow$ es posible reconstruir la señal completa a partir de las DTFT de cada cuadro en el tiempo.

(obs. si N del DFT $\geq N_w$ se puede también con DFT)

Demostración

$$\sum_{n=-\infty}^{\infty} X(\omega, nL) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} x(m)\, w(nL-m)\, e^{-j\omega m}$$

$$= \sum_{m=-\infty}^{\infty} x[m]\, e^{-j\omega m} \underbrace{\sum_{n=-\infty}^{\infty} w(nL-m)}_{1}$$

$$= \sum_{m=-\infty}^{\infty} x(m)\, e^{-j\omega m} = X(\omega)$$

(síntesis en tiempo) (*)

OBS) Esta es una reconstrucción en tiempo de la señal

STFT como un banco de filtros

Es posible interpretar la STDTFT y la STDFT como bancos de filtros y, a su vez optimizar este tipo de operaciones mediante estas herramientas en tiempo / frecuencia.

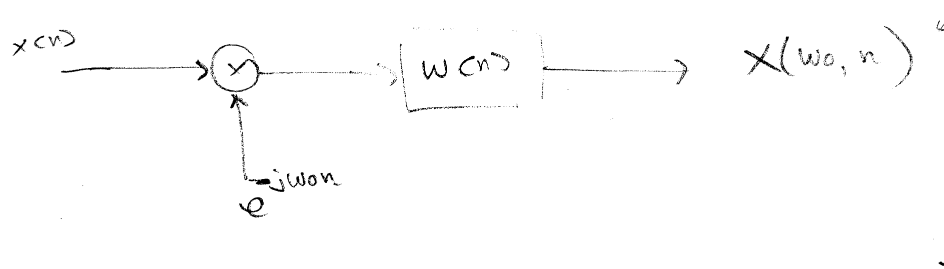
$$X(\omega, n) = \sum_{m=-\infty}^{\infty} x(m)\, w(n-m)\, e^{-j\omega m}$$

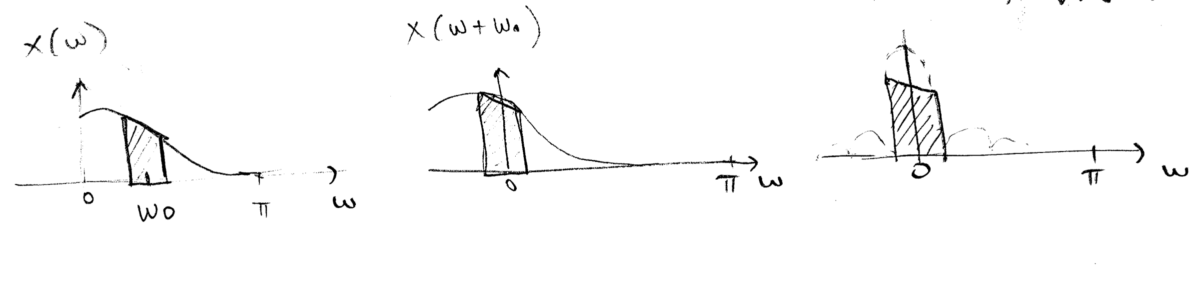
Si fijamos $\omega = \omega_0$

$$X(\omega_0, n) = \sum_{m=-\infty}^{\infty} \left(x(m)\, e^{-j\omega_0 m} \right) w(n-m)$$

$$X(\omega_0, n) = \left(x(n)\, e^{-j\omega_0 n} \right) * w(n) \quad \text{(señal en el tiempo)}$$

→ Interpretación: la energía de cada banda ω_0 se obtiene mediante:
- demodulación del espectro de $X(\omega)$ en ω_0
- Filtrado por un filtro centrado en DC $W(\omega)$ (ventana $w(n)$)





también

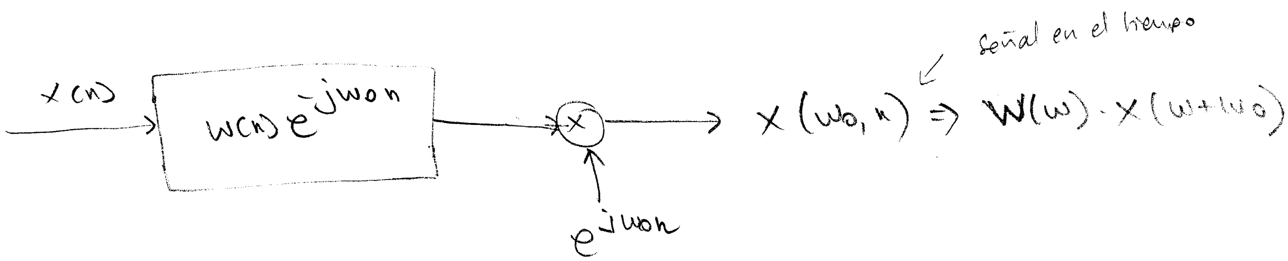
$$x(w_0, n) = x(n)\, e^{-jw_0 n} * w(n)$$

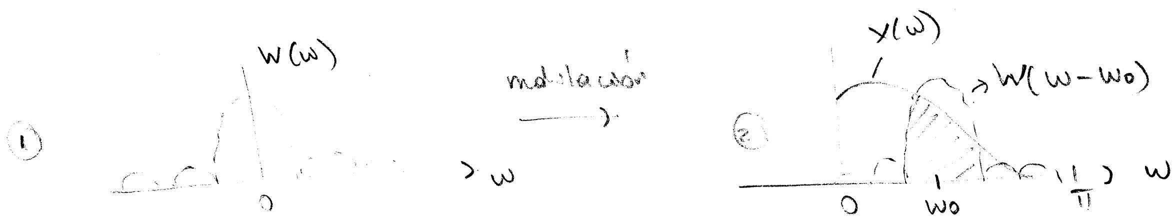
$$= e^{-jw_0 n} \left(x(n) * w(n)\, e^{jw_0 n} \right) \quad \text{(señal en el tiempo)}$$

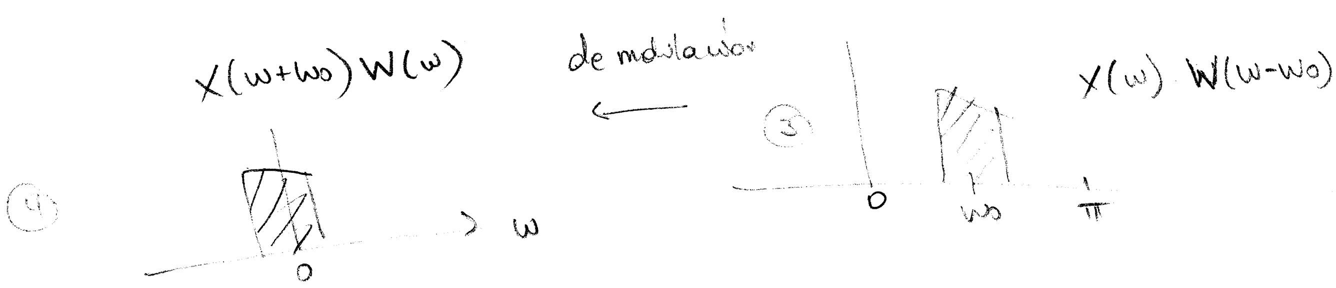
⤳ interpretación :

- la energía o el espectro de cada banda (la banda está dada por el ancho de banda del filtro pasabajo que se conforma con la ventana) se obtiene mediante.
- la señal pasa por un filtro que está modulado : $W(w - w_0)$
- la salida del filtro es luego demodulada

$$X(w)\, W(w - w_0) \longrightarrow X(w + w_0) \cdot W(w)$$







OBS) STDTFT se puede interpretar como un banco infinito de filtros para cada w_0 (variable continua !)

de igual forma para la STDFT,

$$X(K_0, n) = e^{-j\frac{2\pi}{N}K_0 n}\left(x[n] * w[n] \cdot e^{j\frac{2\pi}{N}K_0 n}\right)$$

esto se puede interpretar como un Banco de Filtros (Finito, de N bancos ⇒ # FT puntos) que procesa la secuencia de datos $x[n]$ en paralelo, donde K corresponde a la frecuencia central

OBS)
- Si length$(x) = N$, length$(w) = M$, entonces length$(X(K,n)) = N + M - 1$
- El ancho de banda de $X(K,n)$ es menor o igual al de la ventana $w[n]$
- la secuencia $X(K_0,n)$ está centrada en el origen para cada K_0
- Si se utilizan $X(K, nL)$ esto es equivalente a realizar downsampling de la salida de cada banco de filtros.
- Esto permite una evaluación más eficiente en una etapa de análisis de datos, y al mismo tiempo, si la ventana cumple con $\sum_{m=-\infty}^{\infty} w[n-Lm] = ct$ es posible reconstruir la señal perfectamente usando interpolación.

RELACIÓN TIEMPO - FRECUENCIA

Dado que $X(\omega, n) = \sum_{m=-\infty}^{\infty} x(m)\, w(n-m)\, e^{-j\omega m}$

$$w[n-m] \overset{DTFT}{\longleftrightarrow} W(-\omega)\, e^{+j\omega n}$$

$$x(m) \overset{DTFT}{\longleftrightarrow} X(\omega)$$

$$\Rightarrow \quad X(\omega, n) = X(\omega) * W(\omega) e^{j\omega n}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\theta) e^{j\theta n}\, X(\omega+\theta)\, d\theta$$

i) si $W(\omega) \longrightarrow \delta(\omega)$ banda angosta

$\Rightarrow X(\omega, n) \longrightarrow X(\omega)$ buena resolución en frecuencia

sin embargo para esto $w(n)$ tiene un largo muy grande y pierde la dependencia del tiempo

ii) si $W(\omega) \longrightarrow 1$ band ancha

$\Rightarrow X(\omega, n) \longrightarrow x(n)$ buena resolución en el tiempo

Ejemplos] chirp, señal de voz

Short-time synthesis

- Filter bank summation (FBS)

STFT : la señal de cada banda es demodulada a DC y filtrada en baja frecuencia por el filtro que constituye la ventana

Síntesis : Caso inverso !
lleva la banda en DC filtrada por la ventana a su banda original (modulación) y suma todas las bandas para reconstruir $x(n)$

$$\text{STDFT} = X(k,n) = \sum_{m=0}^{N-1} x[m]\, w[n-m]\, e^{-j 2\pi k m / N}$$

$$\text{FBS} : y(n) = \sum_{k=0}^{N-1} X(k,n)\, e^{j 2\pi k n / N}$$

modulación

suma de todas las bandas (síntesis en frecuencia) (*)

¿ bajo qué condiciones $y(n) = x(n)$?

Es posible demostrar que $x(n) = y(n)$ siempre que $\sum_{K=0}^{N-1} W\left(\omega - \frac{2\pi K}{N}\right) = 1$ (o una constante)

Considere

$$y(n) = \sum_{K=0}^{N-1} \underbrace{X(K,n)\, e^{j2\pi Kn/N}}_{y_K(n)}$$

$$= \sum_{K=0}^{N-1} y_K(n)$$

pero vimos que $X(\omega_K, n) = e^{-j\omega_K n}\left(x(n) * w(n) e^{j\omega_K n}\right)$

$$\Rightarrow \quad X(K,n) = e^{-j2\pi Kn}\left(x(n) * w(n) e^{j2\pi Kn/N}\right)$$

$$\Rightarrow \quad y_K(n) = x(n) * w(n) e^{j2\pi Kn/N}$$

$$\Rightarrow \quad Y_K(\omega) = X(\omega) \cdot W\left(\omega - \frac{2\pi K}{N}\right)$$

$$Y(\omega) = \sum_{K=0}^{N-1} Y_K(\omega) = \sum_{K=0}^{N-1} X(\omega) \cdot W\left(\omega - \frac{2\pi K}{N}\right)$$

$$= X(\omega) \cdot \sum_{K=0}^{N-1} W\left(\omega - \frac{2\pi K}{N}\right)$$

si $\sum_{K=0}^{N-1} W\left(\omega - \frac{2\pi K}{N}\right) = 1 \Rightarrow Y(\omega) = X(\omega)$!

Esta condición se cumple si $N_W \leq N$

OBS] • la condición

$$\sum_{k=0}^{N-1} W\left(\omega - \frac{2\pi k}{N}\right) = 1$$

indica que la suma de todas las bandas moduladas debe cubrir el espectro completo.

• Anteriormente vimos otra condición de ventana que permite reconstrucción de la señal $x(n)$, pero esta es una condición en el tiempo y donde

$$\sum_{n=-\infty}^{\infty} w[nL - m] = 1 \qquad \forall m \in \mathbb{Z}$$

Esta condición lleva a la reconstrucción en el tiempo $\sum_{n=-\infty}^{\infty} X(k, nL) = X(k)$

y no necesita la modulación por ser una suma temporal (en "n" y no en "k").

Esta condición es más estricta y permite saltar puntos en el calculo de la STFT manteniendo la reconstrucción perfecta.

$\Rightarrow$ más eficiente.