EBU6018
Advanced Transform Methods

Short-Time Fourier Transform_2
Gabor Transform

Andy Watson

Gabor Transform

- This is a discrete version of the STFT that uses a Gaussian function as the window.
- Represent a signal in 2 dimensions, with time and frequency as coordinates.
- Expand signal as a series of elementary functions.
- Constructed from a single building block by
  - Translation (in the time-domain)
  - modulation (translation in the frequency domain)
- Transform sampled at regular intervals.
- Intervals are labelled T in time and Ω in frequency.
- The transform divides the signal into regular windows

$$s(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} h_{m,n}(t)$$

$$h_{m,n}(t) = h(t - mT) e^{jn\Omega t}$$

$h_{m,n}$(t) is the discrete basis function. It is the product of a Gaussian window (translated in time) and the Fourier complex exponential.
$c_{m,n}$ are the coefficients.

Choice of elementary function

- The discrete transform is derived from the continuous-time, infinite sum, version by replacing the sums by a finite length of input sequence of the sampled input waveform s[n].
- This is similar to the DFT
- Oversampling $T\Omega < 2\pi$
- Critical sampling $T\Omega = 2\pi$
- Undersampling $T\Omega > 2\pi$

- The only requirement is that $T\Omega \leq 2\pi$
- At critical sampling, the number of coefficients is the same as the number of samples.
- Oversampling would give redundancy That is, more coefficients than we need.
- Undersampling would give an insufficient number of coefficients.

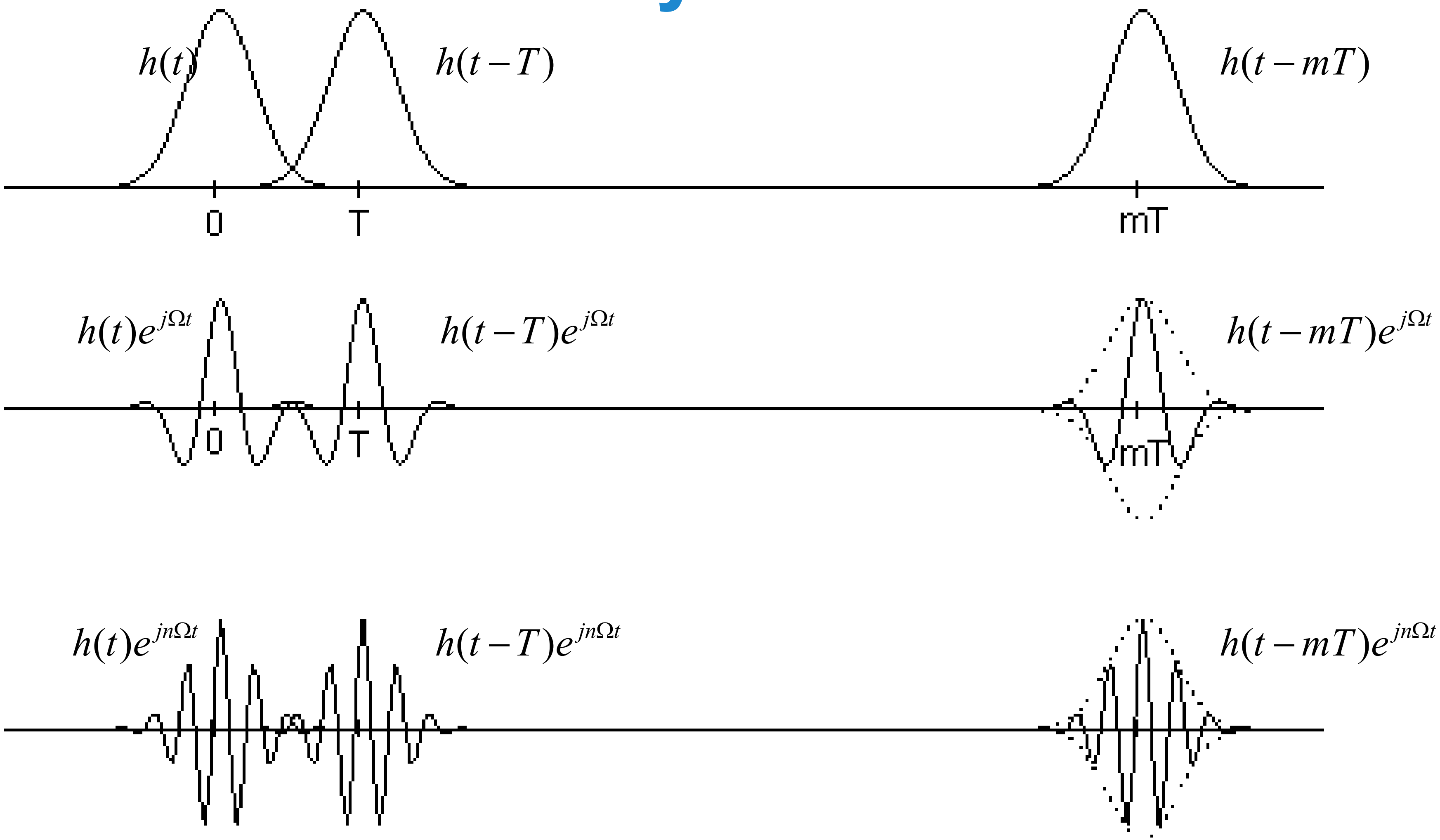
Choice of elementary function

- The most useful window function is a Gaussian window:

$$h(t) = g(t) = \sqrt[4]{\alpha / \pi}\, e^{-\alpha t^2 / 2} \qquad \Delta t \Delta \omega = \frac{1}{\sqrt{\alpha}} \frac{\sqrt{\alpha}}{2} = 1/2$$

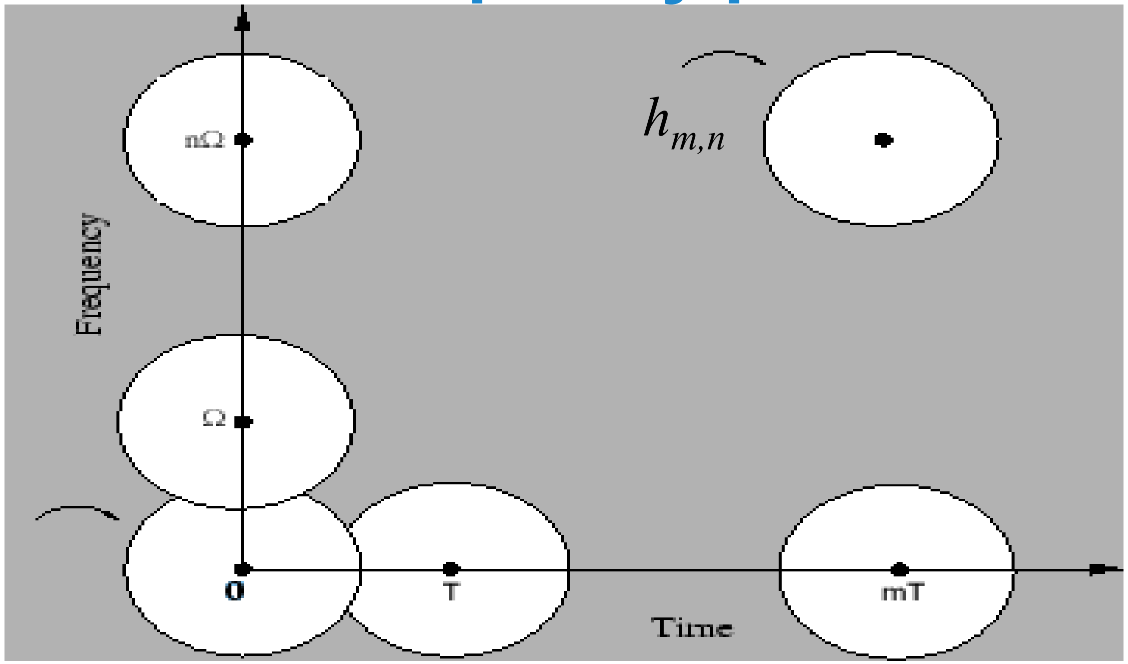
- Optimally concentrated in the time frequency domain.
- This function gives the best resolution
- However, this does not produce orthogonal basis functions.
- Unlike Fourier transform, Gabor transform does not have an orthogonal basis (and often not a tight frame).
- Finding the dual functions is typically not straightforward!

Gabor Elementary Functions



$h_{m,n}$ are shifted and modulated copies of a single building block h.
T denotes the time shift parameter.
Ω denotes the frequency shift parameter.
Each $h_{m,n}$ has an envelope of the shape of h.
Only the real part of the functions is shown.

Time-frequency plane



- $h_{m,n}$ are obtained by shifting h along a lattice in the *time-frequency plane*.
- If h and its Fourier transform are centred at the origin, then $h_{m,n}$ is centred at $(mT,n\Omega)$ in the time-frequency plane.
- Each $h_{m,n}$ essentially occupies a certain area in the time-frequency plane.
- Each expansion coefficient $c_{m,n}$, associated to a certain area of the time-frequency plane via $h_{m,n}$ represents one *quantum of information*.
- For properly chosen shift parameters the $h_{m,n}$ cover the time-frequency plane.

Gabor coefficients

- $c_{m,n}$ are the Gabor coefficients. They can be found from:

$$c_{m,n} = \int_{-\infty}^{\infty} s(t)\gamma_{m,n}^{\;*}(t)dt = STFT[mT, n\Omega] = \langle s, \gamma_{m,n} \rangle$$

- Reconstruction / synthesis window $\gamma(t)$ is not necessarily the same as the analysis window $h(t)$.
- These windows are only identical if the functions are orthogonal.
- A system $\gamma(t)$ is a Gabor frame if $A\|s\|^2 \leq \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left|\langle f, g_{m,n} \rangle\right|^2 \leq B\|s\|^2$

For a Gabor frame $f = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, g_{m,n} \rangle \gamma_{m,n} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \gamma_{m,n} \rangle g_{m,n}$

In general, $\gamma(t)$ and $h(t)$ are a Gabor frame, However, they must form a Dual Basis and satisfy the Kroneker Delta Function.

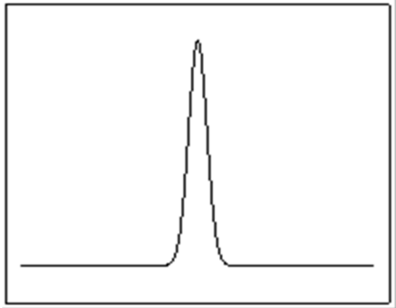
Similar to Nyquist criterion for sampling and reconstruction of bandlimited functions in Shannon's Sampling Theorem.

Classify Gabor systems according to the corresponding sampling frequency of the time-frequency lattice:

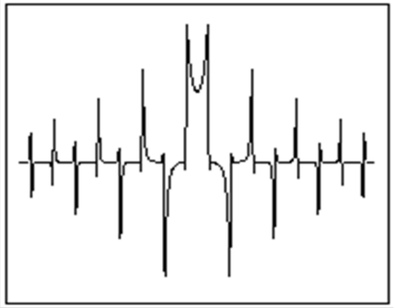
- ***Oversampling*- Frames with excellent time-frequency localization properties exist**
  - **Gaussian with appropriate oversampling rate**

- ***critical sampling*- Frames and orthonormal bases possible, but without good time-frequency localization**

- ***Undersampling*- Any Gabor family will be incomplete. Cannot have a frame.**

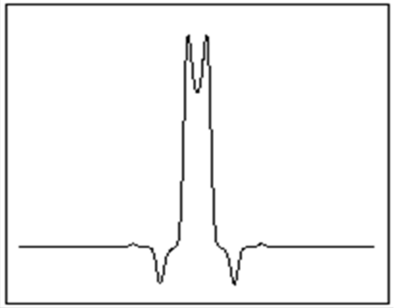


(e) Gabor function g

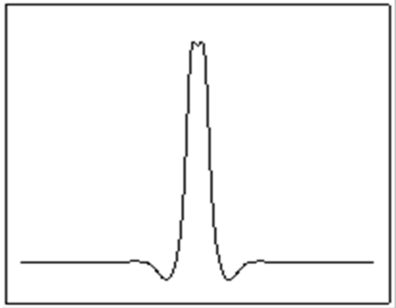


(f) Dual function for critical sampling

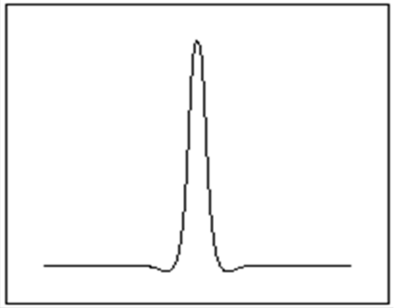
(g) Dual function for oversampling rate 1.06



(h) Dual function for oversampling rate 1.25



(i) Dual function for oversampling rate 1.67



(j) Dual function for oversampling rate 2.4

Gabor summary

- If we want to use a Gaussian function as the reconstruction function, we need to find the window function $h(t)$.
- Discrete form is difficult to find, but possible
- No fast transform
- *But…* Lots of applications
  - **Speech signal analysis:**
  - **Time-Varying Spectral Estimation:**
  - **Representation and Identification of Linear Systems:**
  - **Digital Communication:**
  - **Image representation and biological vision**
- Active area of research

Revision: Bases and Frames

- A set of vectors or functions $\{\psi_n\}$ spans a vector space if any element of that space can be expressed as a linear combination of members of that set
$$s = \sum_n c_n \psi_n$$
- $\{\psi_n\}$ is a *basis set* if the c_n are unique.
- The set is an orthogonal basis if $n \neq m \Rightarrow \langle \psi_n, \psi_m \rangle = 0$
- The set is an orthonormal basis if $n = m \Rightarrow \langle \psi_n, \psi_m \rangle = 1$

Revision: Dual basis

- A basis set $\{\hat{\psi}_i\}$ is said to be the dual basis of $\{\psi_i\}$ if the biorthogonality condition

$$\langle \psi_i, \hat{\psi}_j \rangle = \sum_k \psi_i(k)\hat{\psi}_j(k) = \delta_{ij}$$

is satisfied.

As a transform,
the dual basis set $\{\hat{\psi}_j\}$ is the inverse of the basis set $\{\psi_i\}$

- Example:

$$\{\psi_i\} = \{ \ (1,0),(1,1)\}$$

$$\{\hat{\psi}_i\} = \{ \ (1,-1),(0,1)\}$$

$$\langle \psi_1, \hat{\psi}_1 \rangle = \sum_k (1,0)\cdot(1,-1) = 1$$

$$\langle \psi_1, \hat{\psi}_1 \rangle = \sum_k (1,0)\cdot(1,-1) = 1$$

$$\langle \psi_1, \hat{\psi}_2 \rangle = \sum_k (1,0)\cdot(0,1) = 0$$

$$\langle \psi_2, \hat{\psi}_1 \rangle = \sum_k (1,1)\cdot(1,-1) = 0$$

$$\langle \psi_2, \hat{\psi}_2 \rangle = \sum_k (1,1)\cdot(0,1) = 1$$