Short Time Fourier Transform (STFT)

Phan Duy Hùng

Fourier Transform

- Fourier Transform reveals which frequency components are present in a given function.

$$f(x) = \sum_{u=0}^{N-1} F(u) e^{\frac{j2\pi ux}{N}}, x = 0, 1, \ldots, N-1$$

(inverse DFT)

where:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{\frac{-j2\pi ux}{N}}, u = 0, 1, \ldots, N-1$$

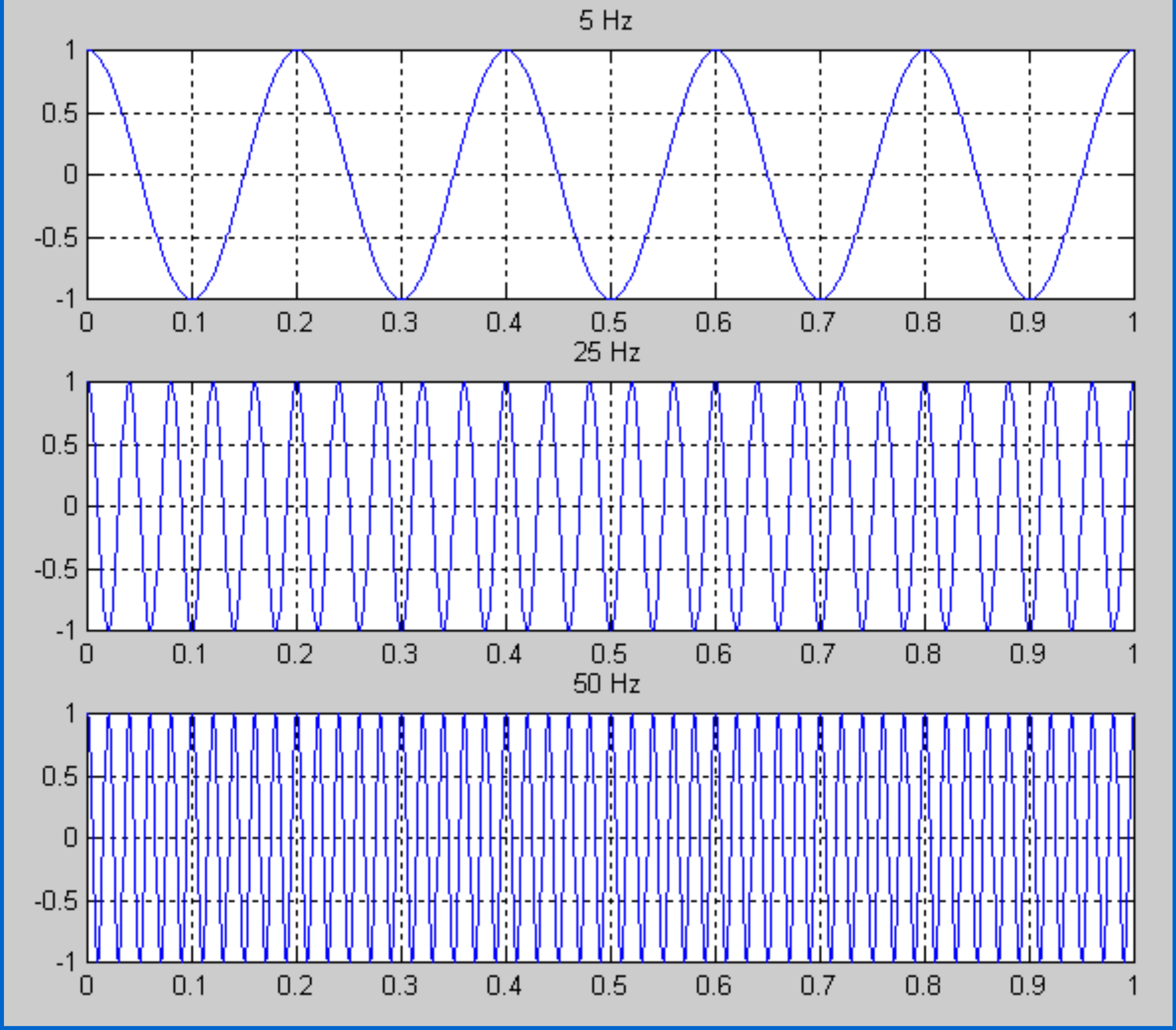
(forward DFT)

Examples

$$f_1(t) = \cos(2\pi \cdot 5 \cdot t)$$

$$f_2(t) = \cos(2\pi \cdot 25 \cdot t)$$

$$f_3(t) = \cos(2\pi \cdot 50 \cdot t)$$

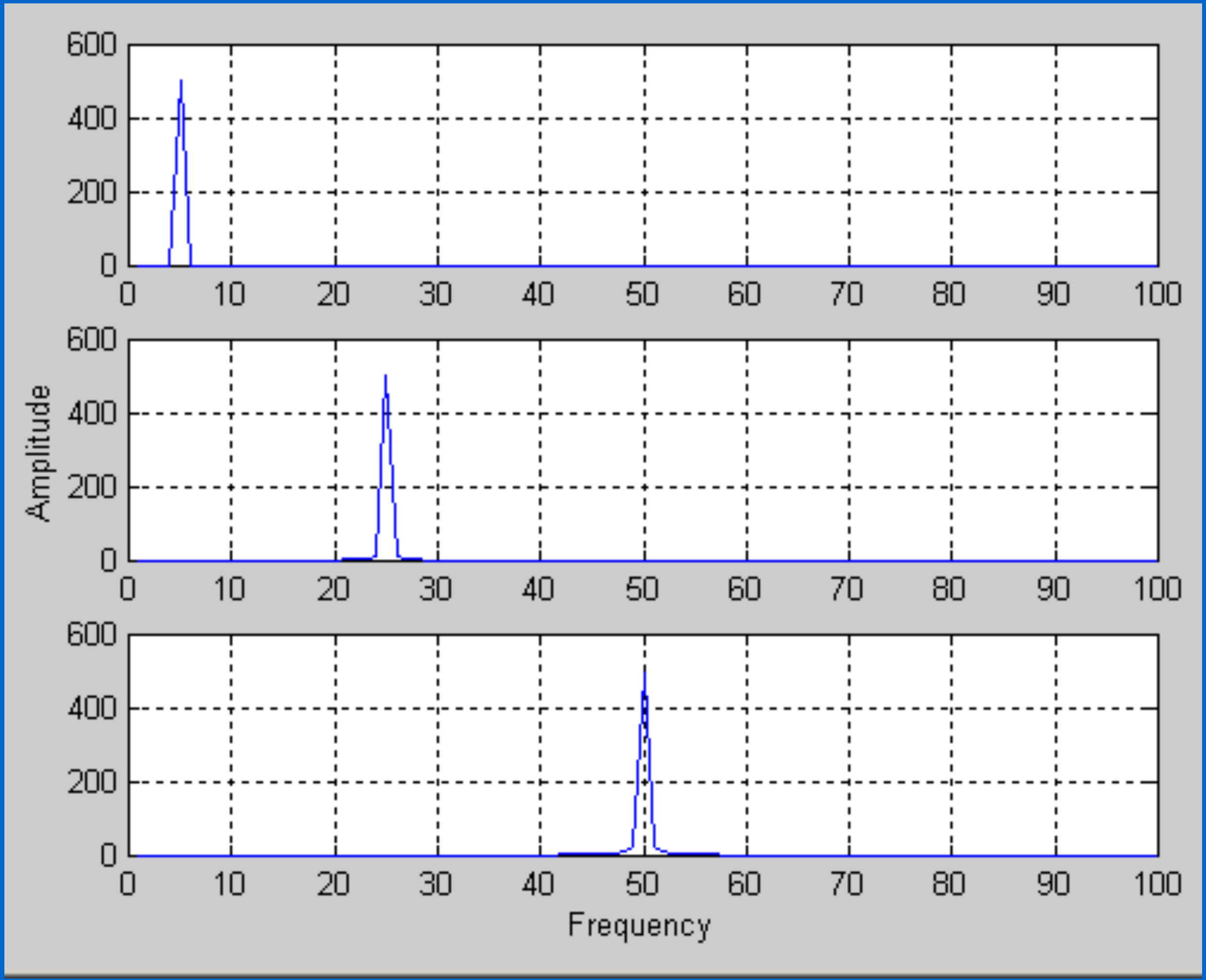


Examples (cont'd)

$F_1(u)$

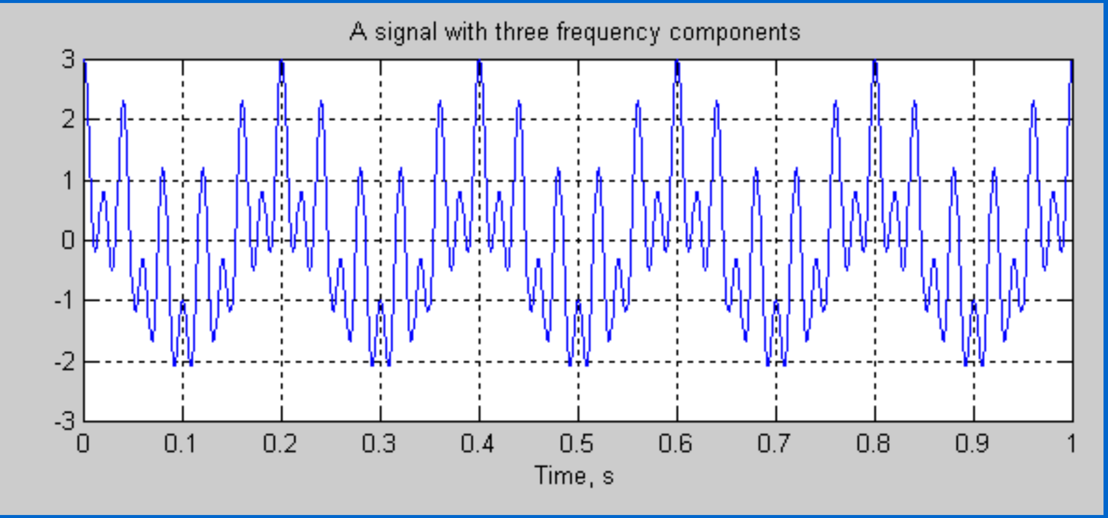
$F_2(u)$

$F_3(u)$

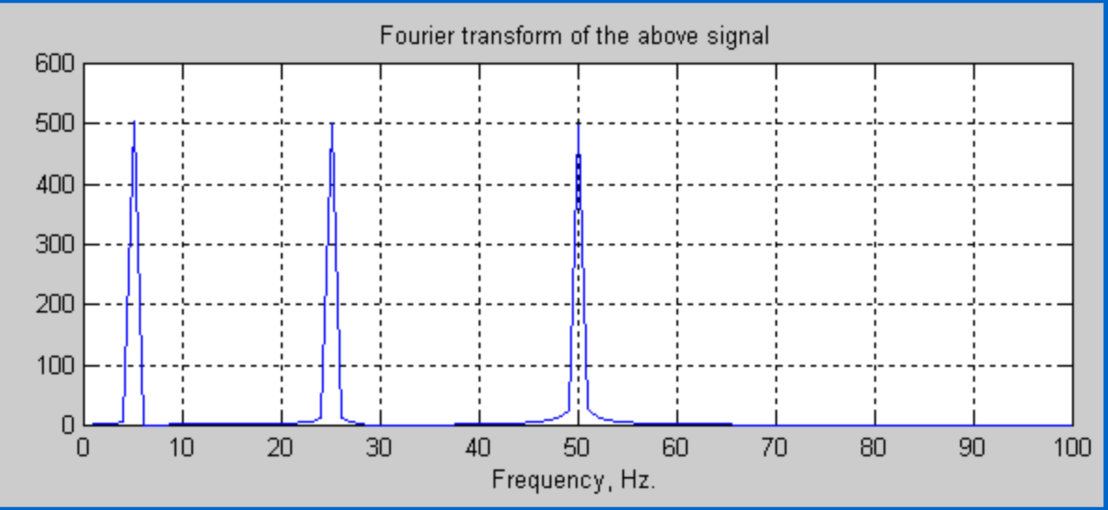


Fourier Analysis – Examples (cont'd)

$$f_4(t) = \cos(2\pi \cdot 5 \cdot t) + \cos(2\pi \cdot 25 \cdot t) + \cos(2\pi \cdot 50 \cdot t)$$

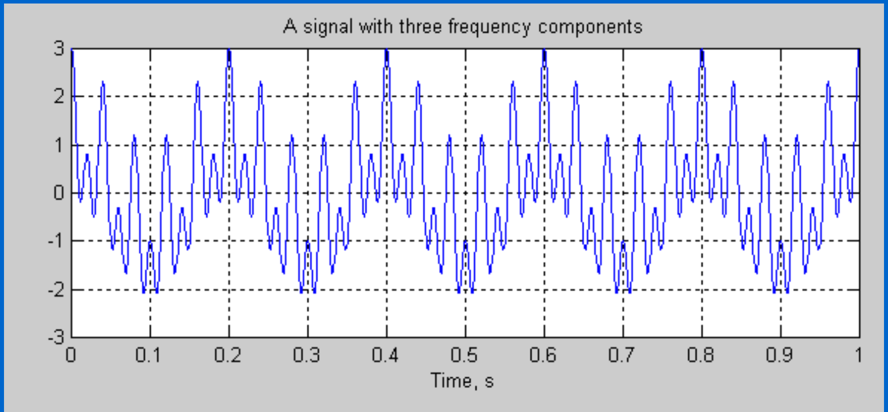


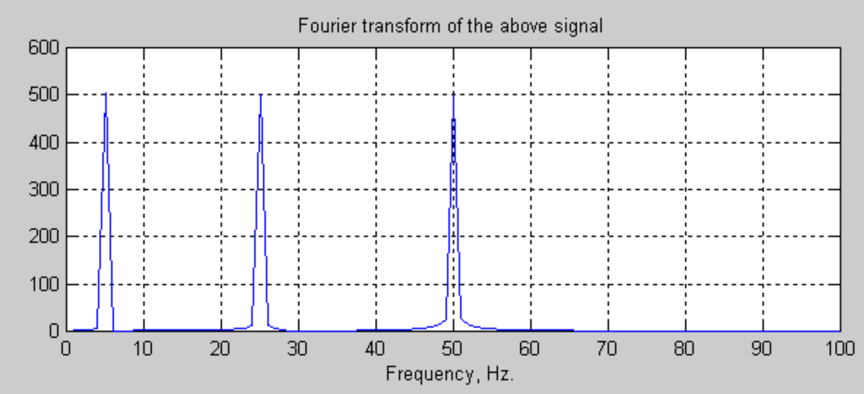
$F_4(u)$?



Limitations of Fourier Transform

1. Cannot not provide simultaneous time and frequency localization.





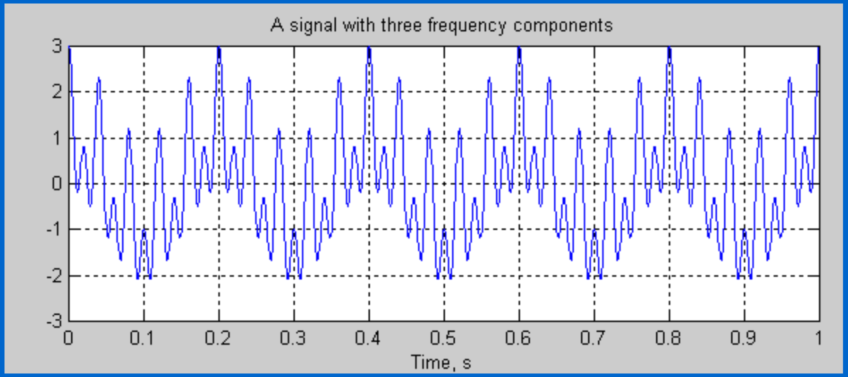
Poor localization in freq domain!

Great localization in freq domain!

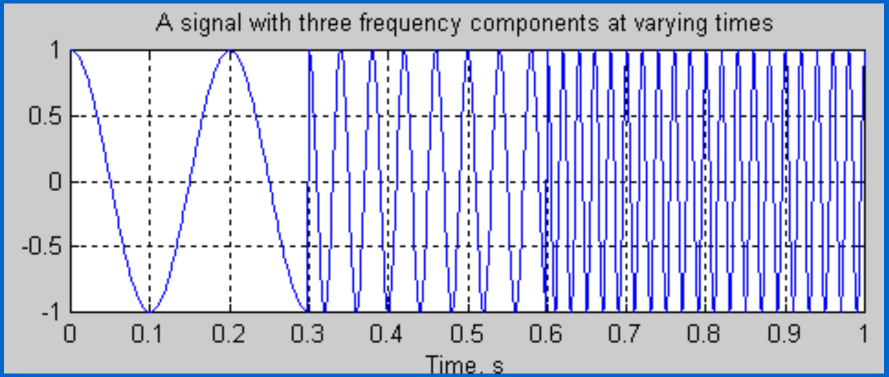
Limitations of Fourier Transform (cont'd)

2. Not very useful for analyzing time-variant, non-stationary signals.

$f_4(t)$ Stationary signal
(non-varying frequency)

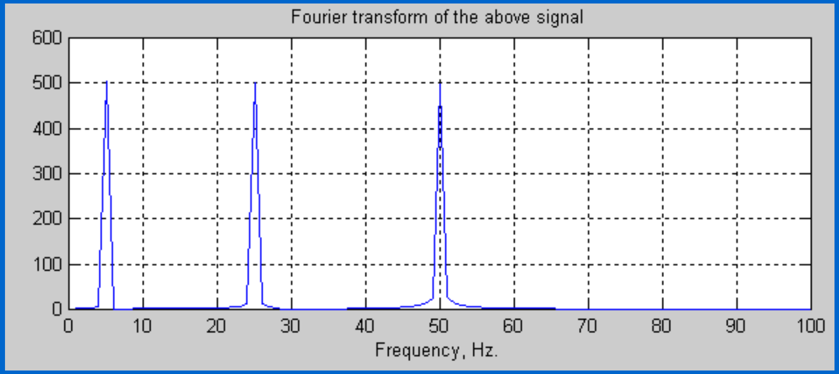


$f_5(t)$ Non-stationary signal
(varying frequency)

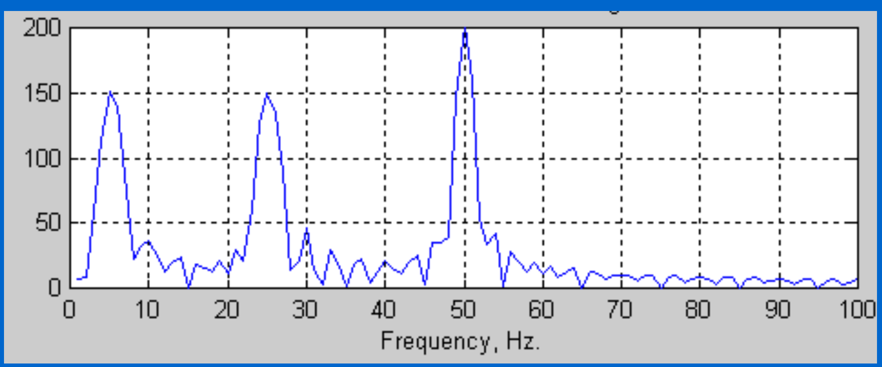


Limitations of Fourier Transform (cont'd)

$F_4(u)$ Three frequency components, present at all times!



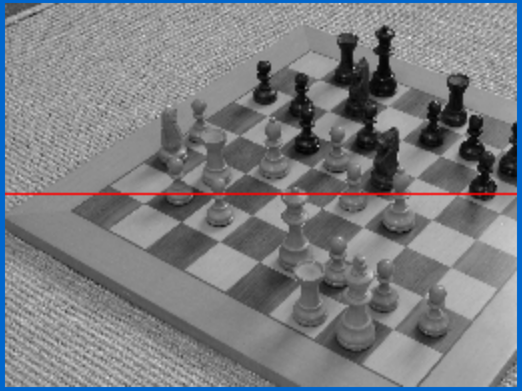
$F_5(u)$ Three frequency components, NOT present at all times!

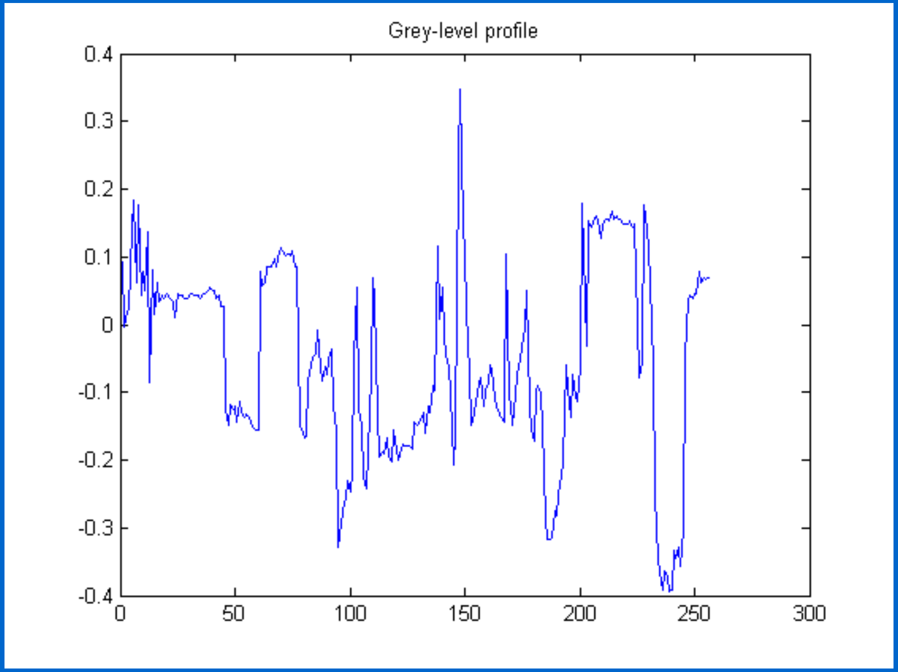


Perfect knowledge of what frequencies exist, but no information about where these frequencies are_located in time!

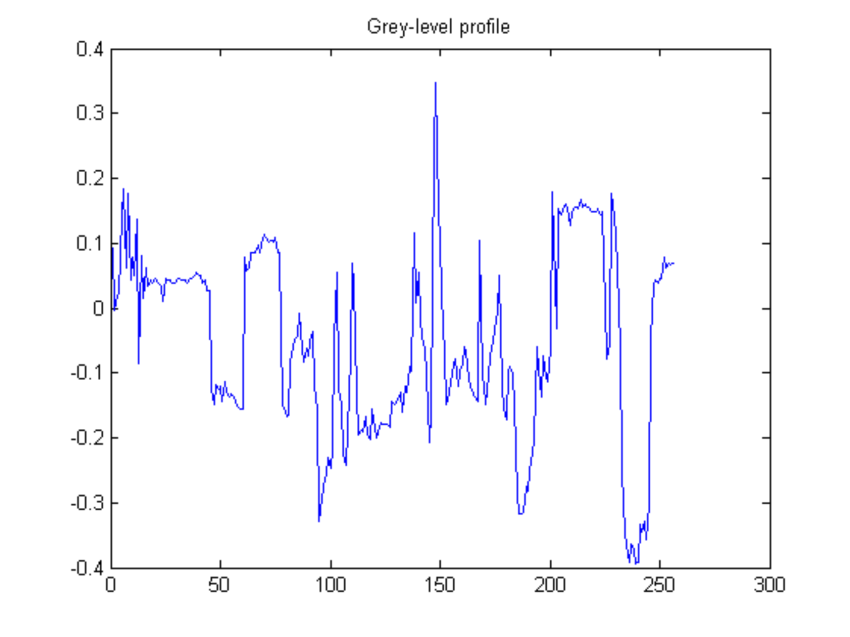
Limitations of Fourier Transform (cont'd)

3. Not efficient for representing non-smooth functions.

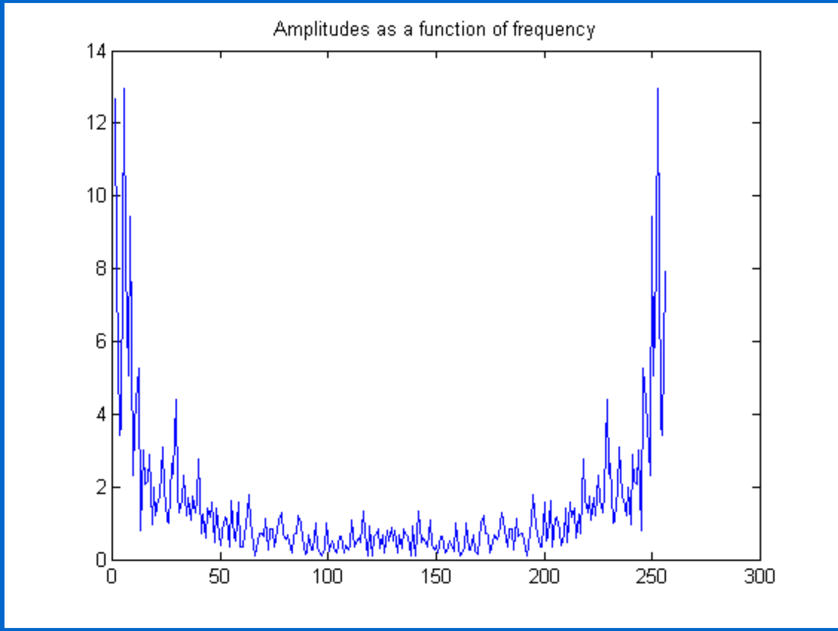




Representing discontinuities or sharp corners (cont'd)



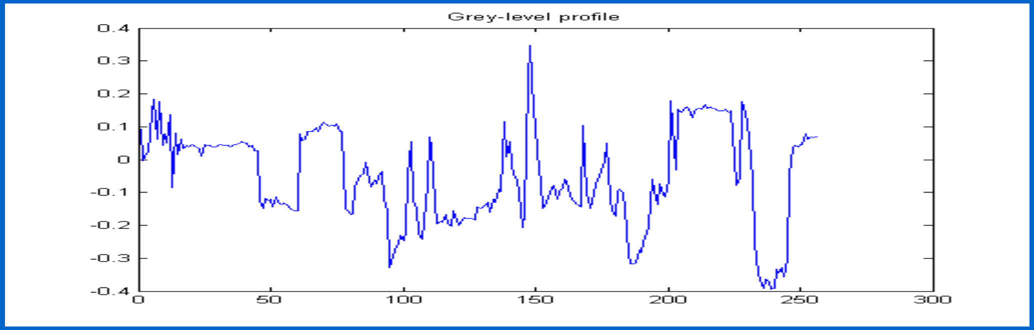
FT

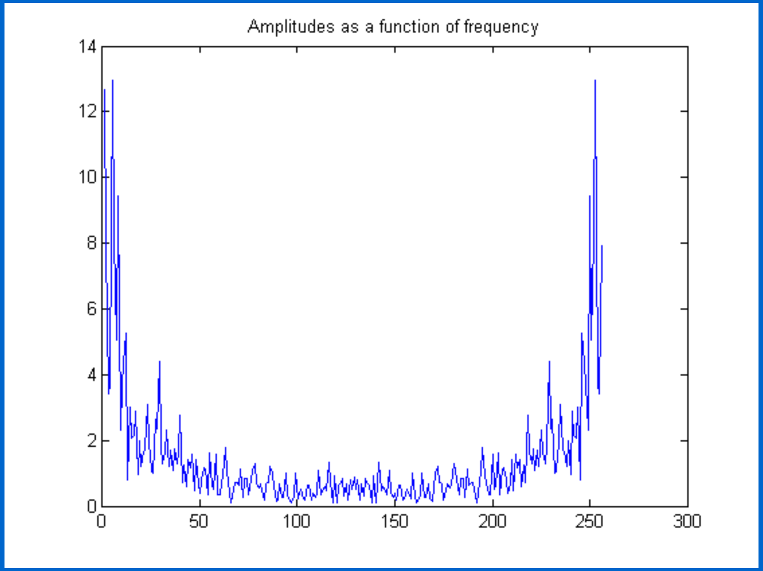


$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{\frac{-j2\pi ux}{N}}, u = 0, 1, \ldots, N-1$$

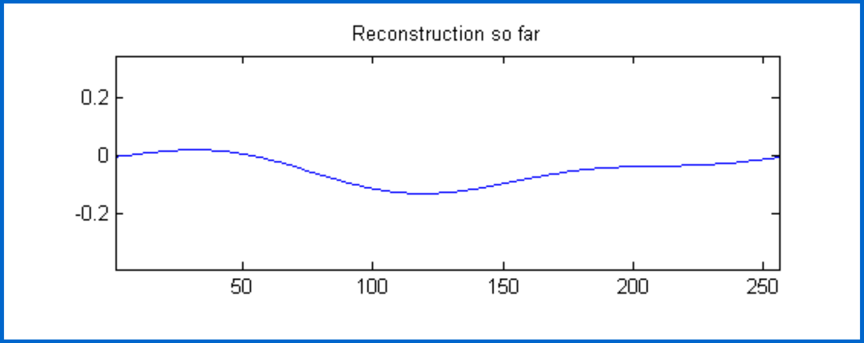
Representing discontinuities or sharp corners (cont'd)

Original





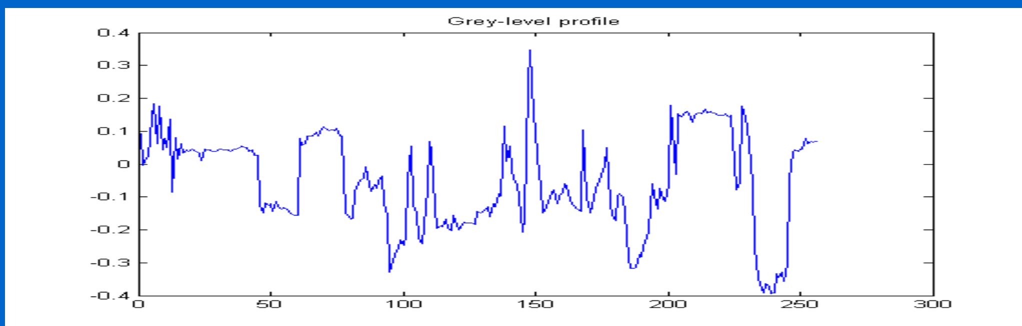
Reconstructed



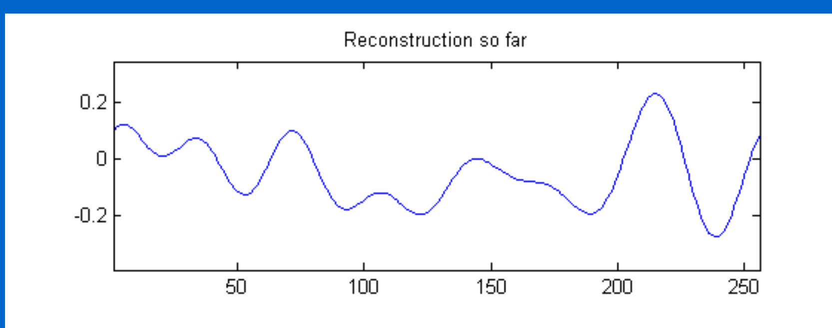
$$f(x) = \sum_{u=0}^{1} F(u) e^{\frac{j2\pi ux}{N}}, \; x = 0, 1, \ldots, N-1$$

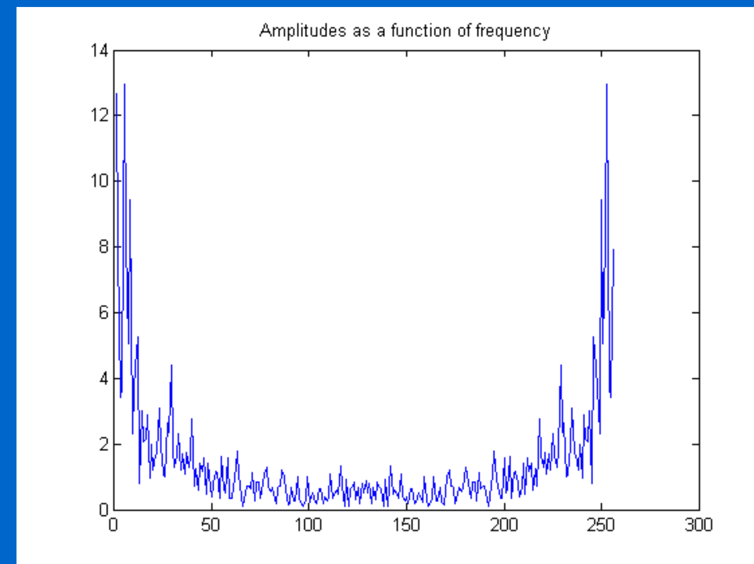
Representing discontinuities or sharp corners (cont'd)

Original



Reconstructed

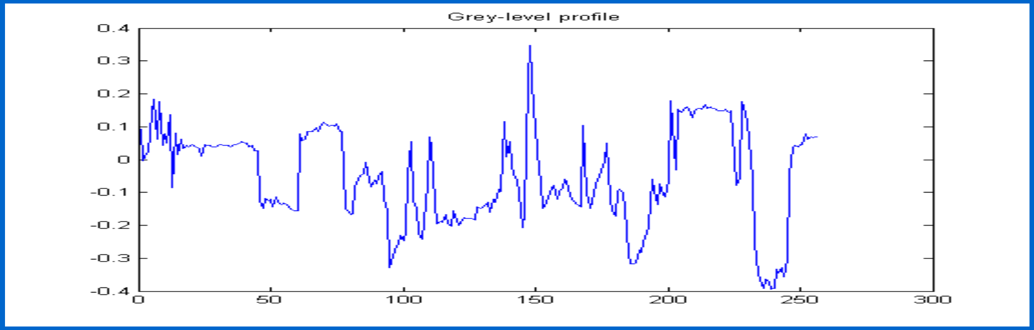


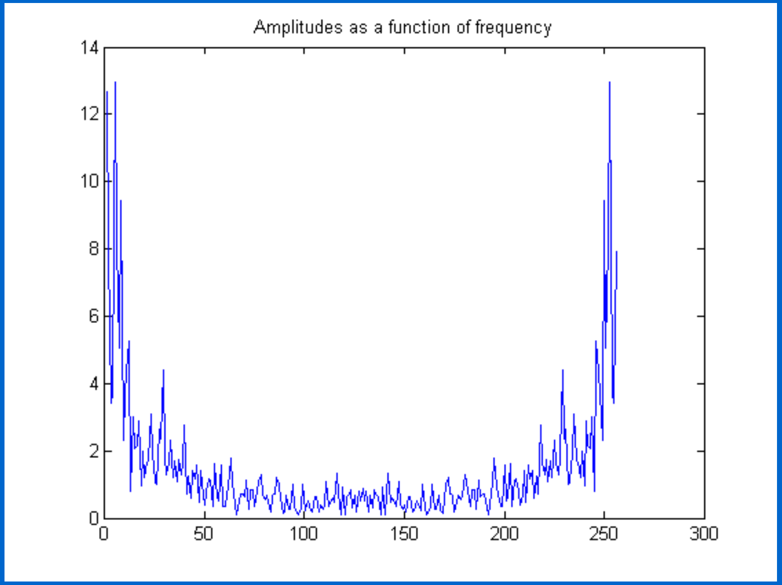


$$f(x) = \sum_{u=0}^{7} F(u) e^{\frac{j2\pi ux}{N}}, \; x = 0, 1, \ldots, N-1$$

Representing discontinuities or sharp corners (cont'd)

Original



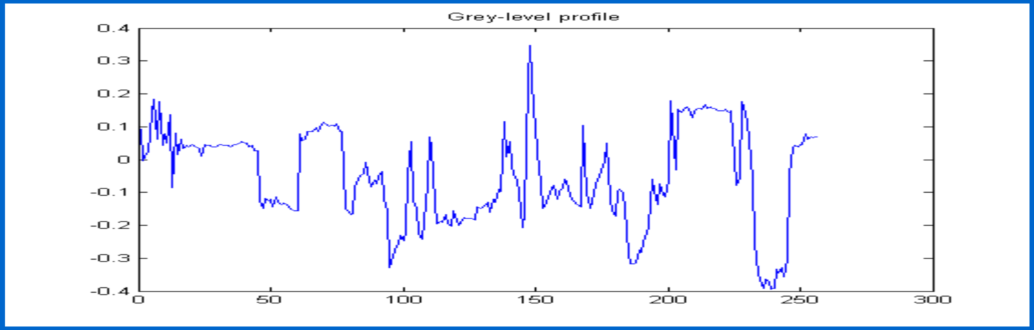


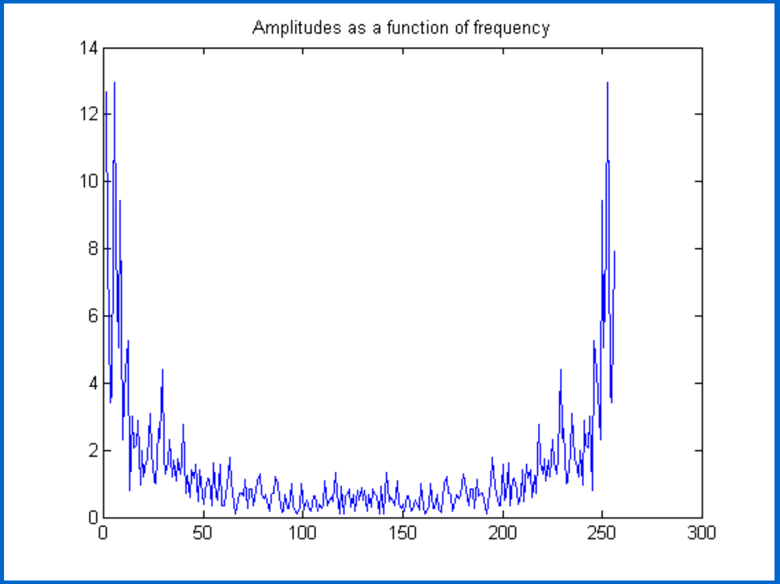
Reconstructed

$$f(x) = \sum_{u=0}^{23} F(u) e^{\frac{j2\pi ux}{N}}, \; x = 0, 1, \ldots, N-1$$

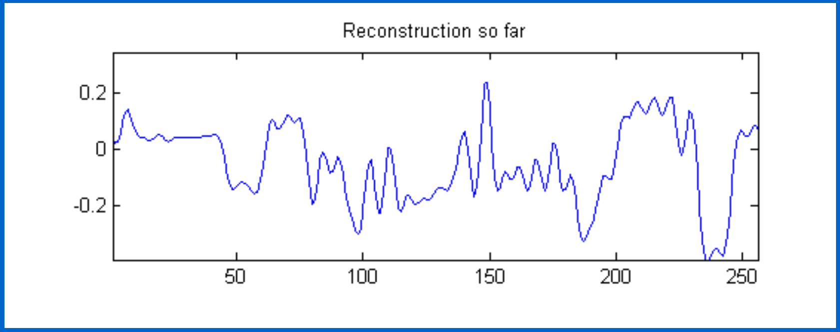
Representing discontinuities or sharp corners (cont'd)

Original





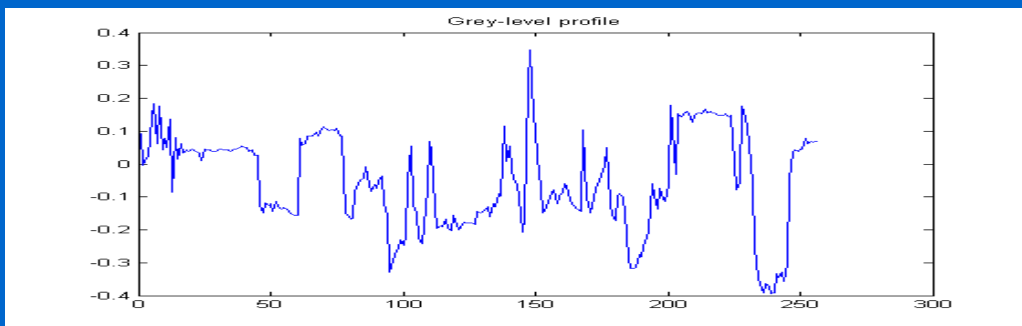
Reconstructed

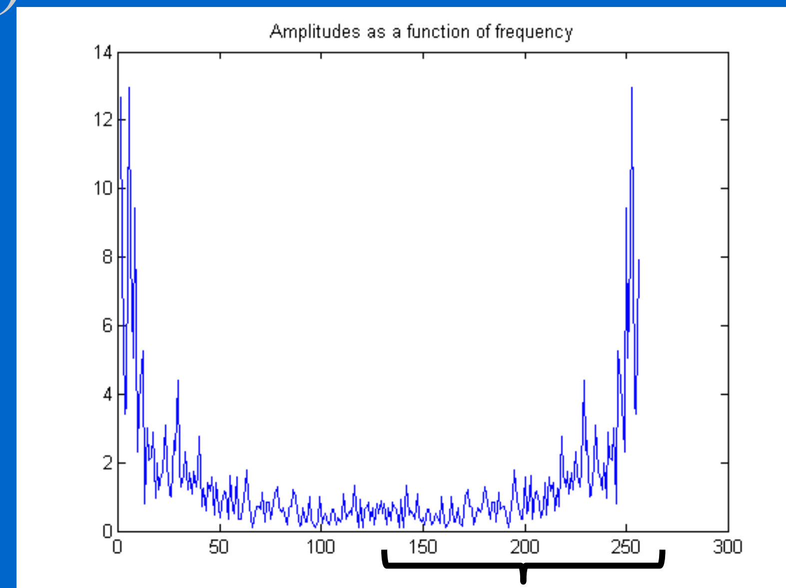


$$f(x) = \sum_{u=0}^{39} F(u) e^{\frac{j2\pi ux}{N}}, x = 0, 1, \ldots, N-1$$

Representing discontinuities or sharp corners (cont'd)

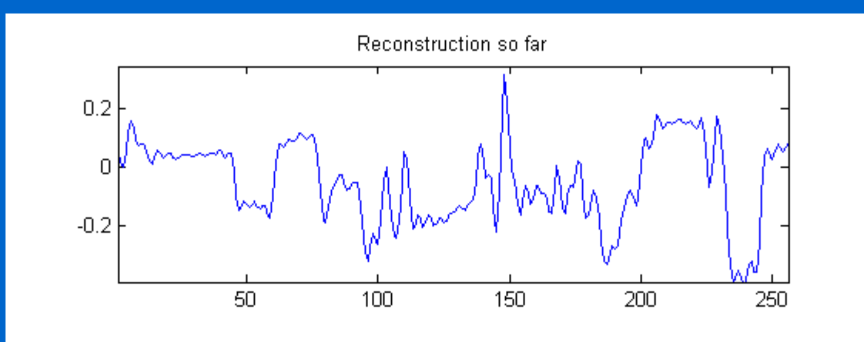
Original





128 coefficients

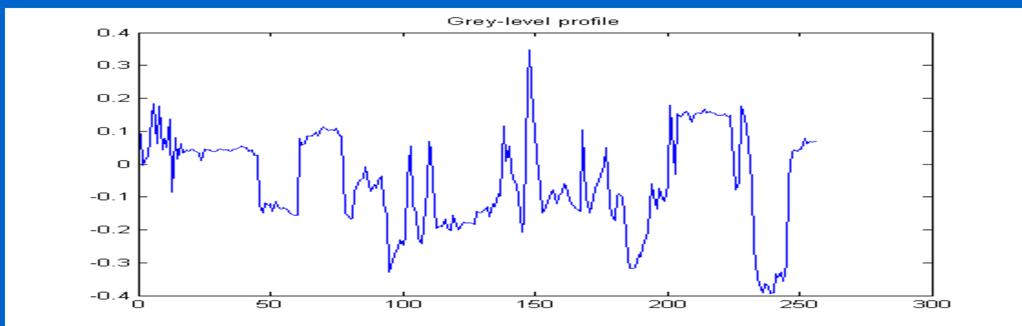
Reconstructed



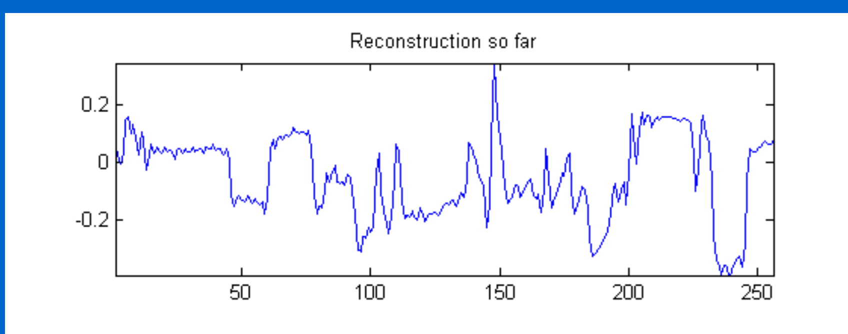
$$f(x) = \sum_{u=0}^{63} F(u) e^{\frac{j2\pi ux}{N}}, x = 0, 1, \ldots, N-1$$

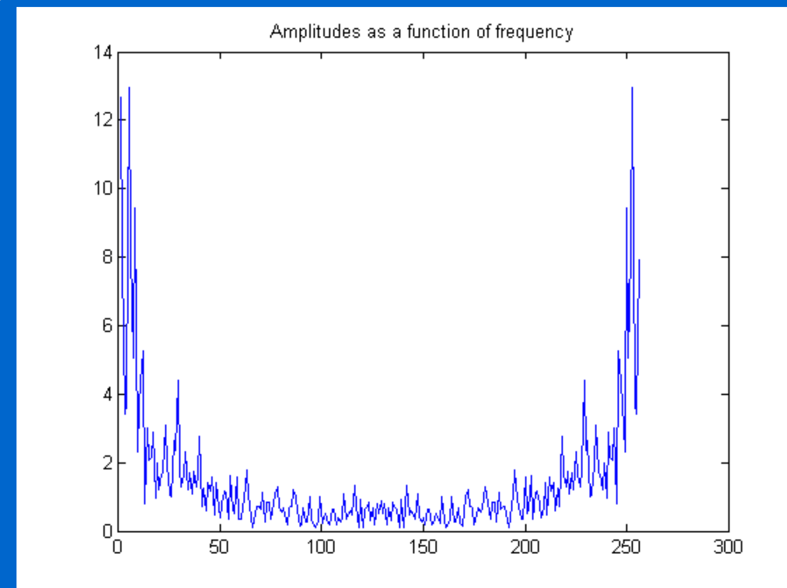
Representing discontinuities or sharp corners (cont'd)

Original



Reconstructed

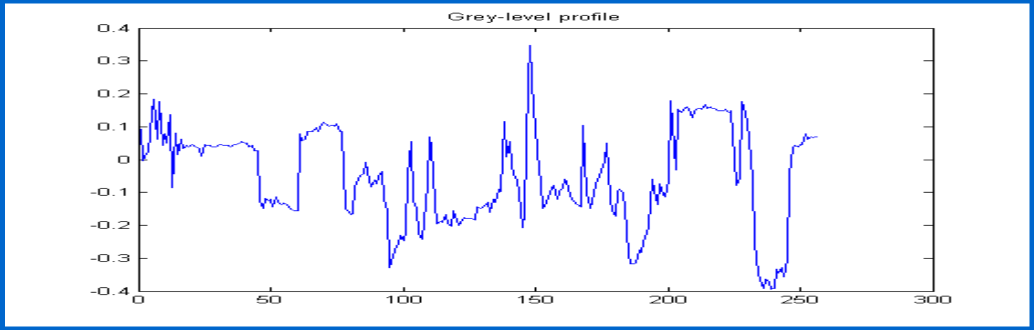


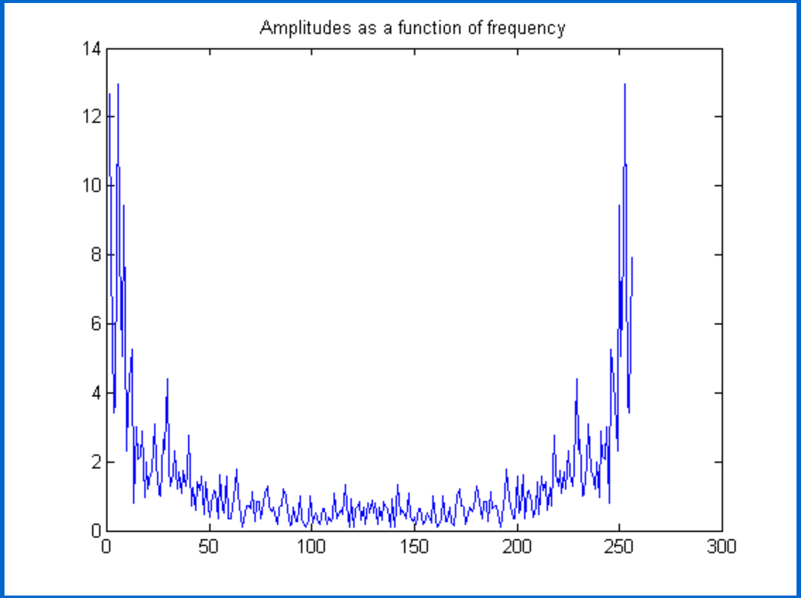


$$f(x) = \sum_{u=0}^{95} F(u) e^{\frac{j2\pi ux}{N}}, \quad x = 0, 1, \ldots, N-1$$

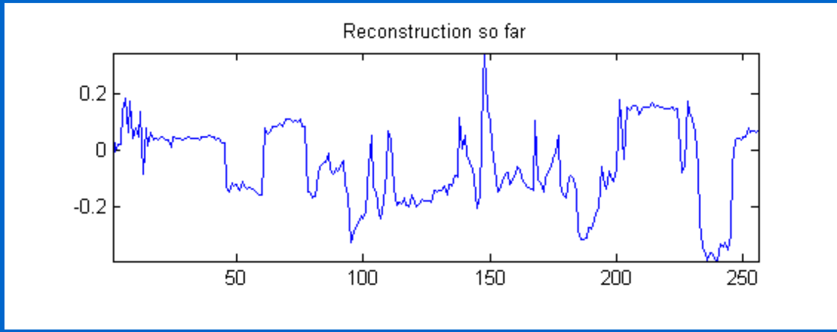
Representing discontinuities or sharp corners (cont'd)

Original





Reconstructed

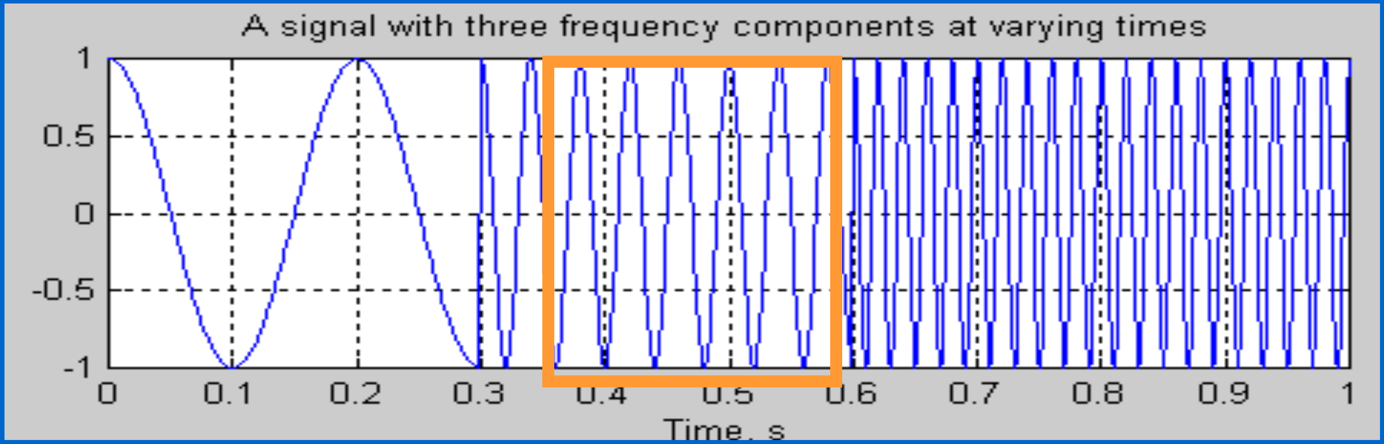


$$f(x) = \sum_{u=0}^{127} F(u) e^{\frac{j2\pi ux}{N}}, \; x = 0, 1, \ldots, N-1$$

A large number of Fourier components is needed to represent discontinuities.

Short Time Fourier Transform (STFT)

- Segment signal into narrow time intervals (i.e., narrow enough to be considered stationary) and take the FT of each segment.
- Each FT provides the spectral information of a separate time-slice of the signal, providing simultaneous time and frequency information.



STFT - Steps

(1) Choose a window of finite length

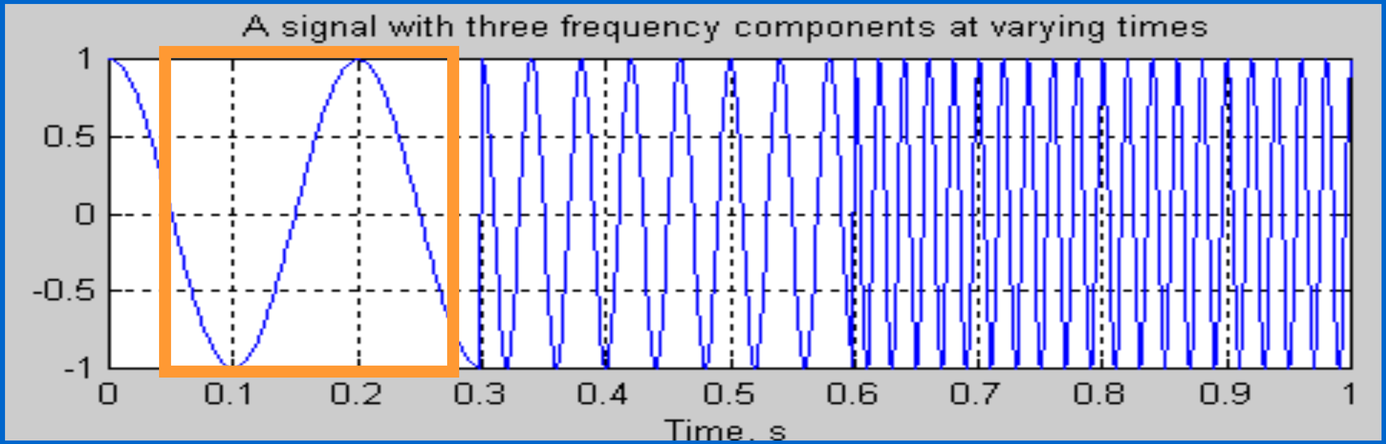
(2) Place the window on top of the signal at t=0

(3) Truncate the signal using this window

(4) Compute the FT of the truncated signal, save results.

(5) Incrementally slide the window to the right

(6) Go to step 3, until window reaches the end of the signal



STFT - Definition

time parameter

frequency parameter

$$STFT_f^u(t',u) = \int_t \left[f(t) \cdot W(t-t') \right] \cdot e^{-j2\pi ut} dt$$

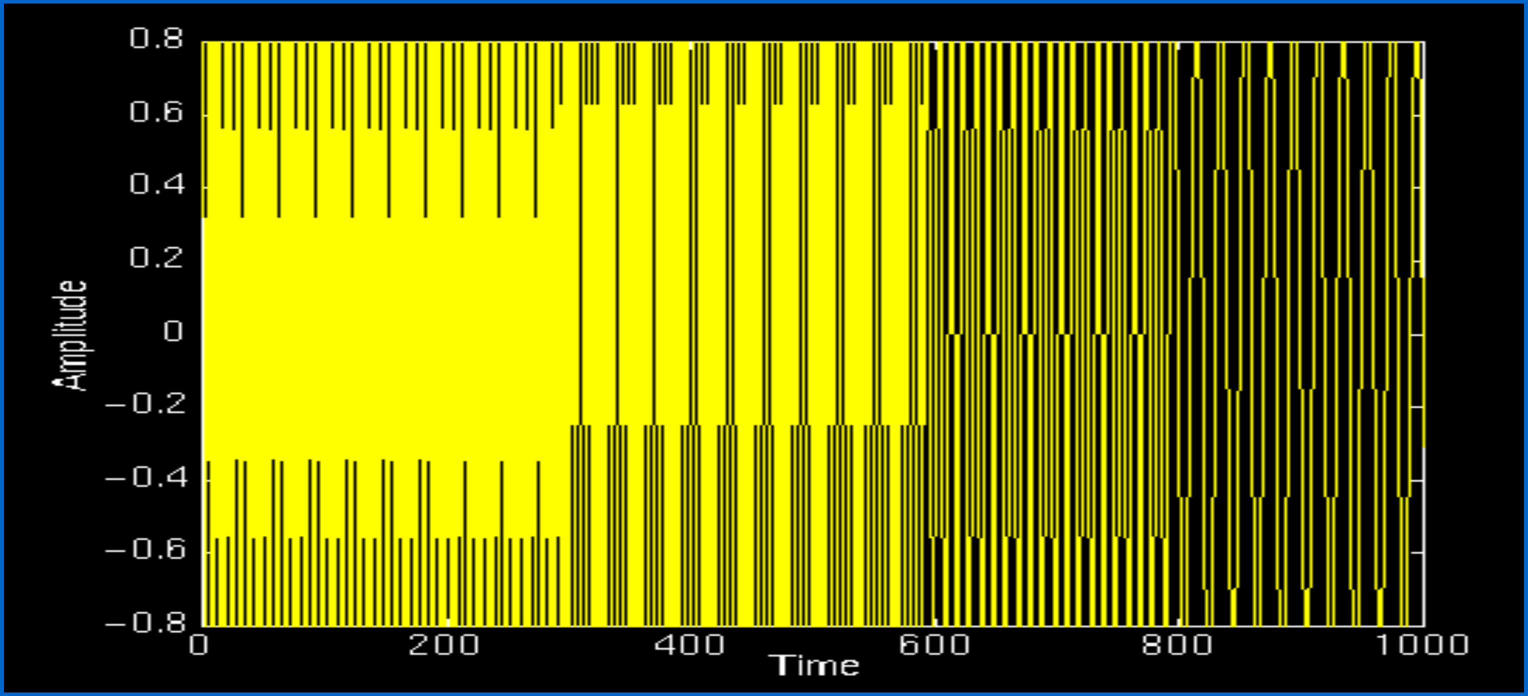
2D function

windowing function

centered at t=t'

Example

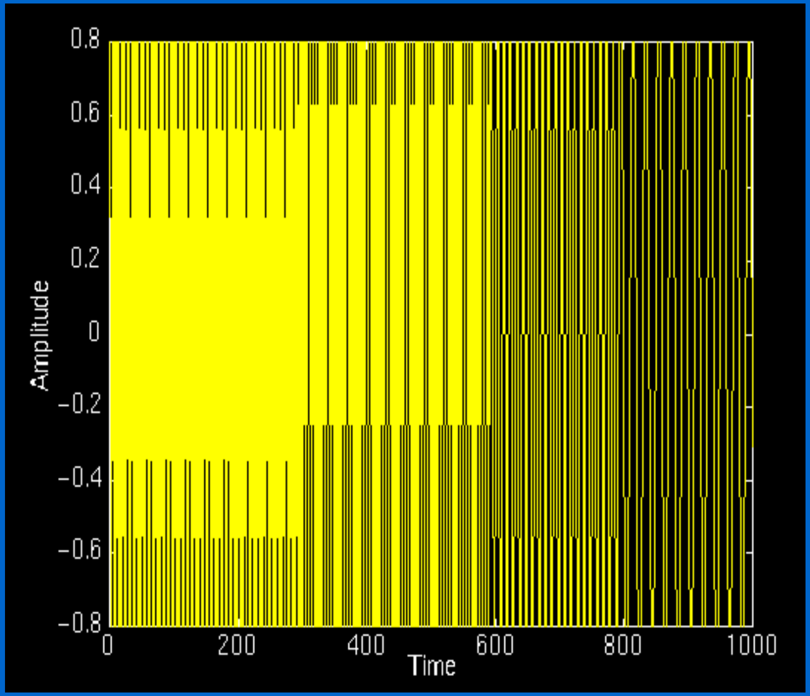
f(t)



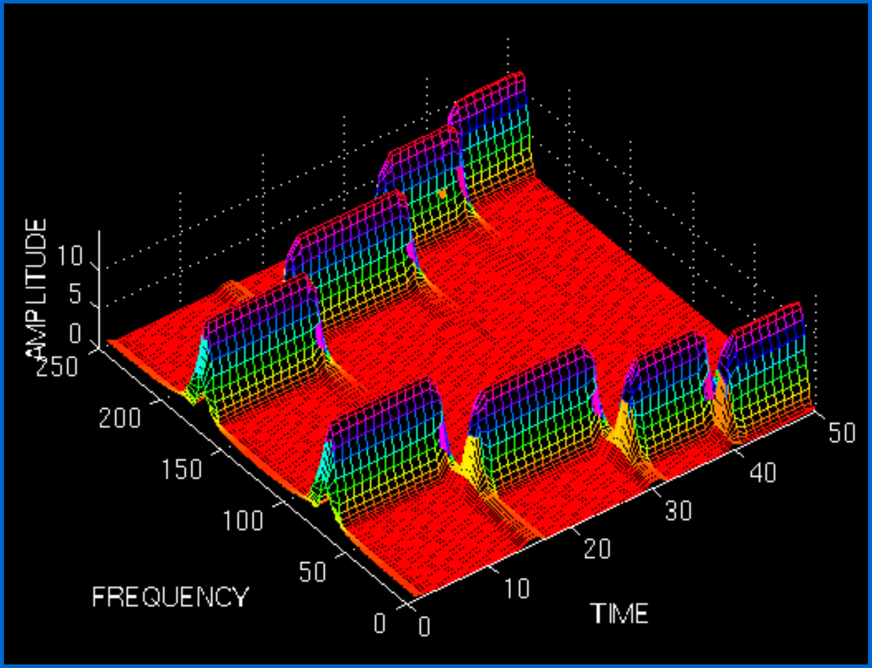
[0 – 300] ms → 75 Hz sinusoid
[300 – 600] ms → 50 Hz sinusoid
[600 – 800] ms → 25 Hz sinusoid
[800 – 1000] ms →10 Hz sinusoid

Example

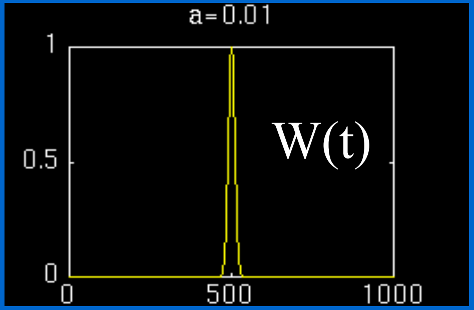
f(t)



$STFT_f^u(t', u)$



[0 – 300] ms → 75 Hz
[300 – 600] ms → 50 Hz
[600 – 800] ms → 25 Hz
[800 – 1000] ms → 10 Hz



scaled: t/20

Choosing Window W(t)

- What shape should it have?
  - Rectangular, Gaussian, Elliptic …
- How wide should it be?
  - Should be narrow enough to ensure that the portion of the signal falling within the window is stationary.
  - Very narrow windows, however, do not offer good localization in the frequency domain.

STFT Window Size

$$STFT_f^u(t',u) = \int_t \left[f(t) \cdot W(t-t') \right] \cdot e^{-j2\pi ut} dt$$

***W(t)* infinitely long**: $W(t) = u(t)$ → STFT turns into FT, providing excellent frequency localization, but no time localization.

***W(t)* infinitely short**: $W(t) = \delta(t)$ → results in the time signal (with a phase factor), providing excellent time localization but no frequency localization.

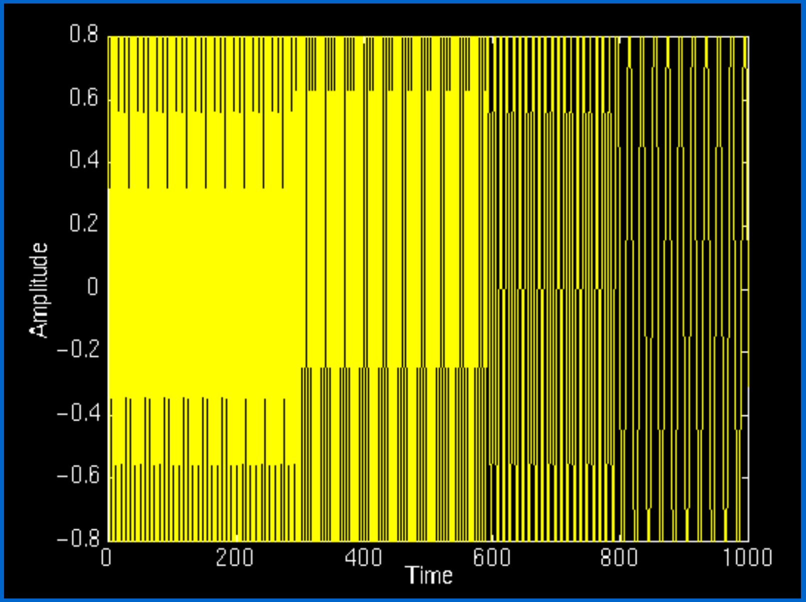
$$STFT_f^u(t',u) = \int_t \left[f(t) \cdot \delta(t-t') \right] \cdot e^{-j2\pi ut} dt = f(t') \cdot e^{-jut'}$$

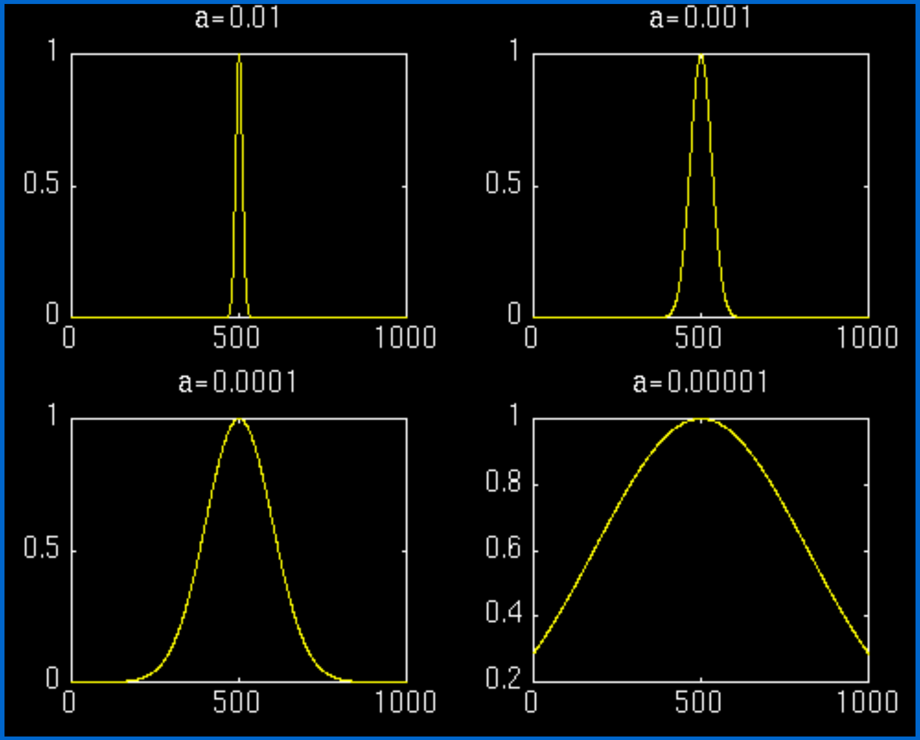
STFT Window Size (cont’d)

- Wide window → good frequency resolution, poor time resolution.
- Narrow window → good time resolution, poor frequency resolution.

Example

different size windows

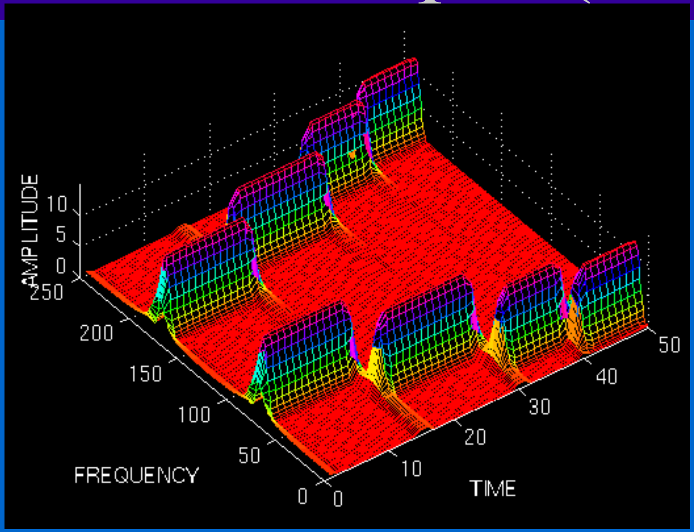


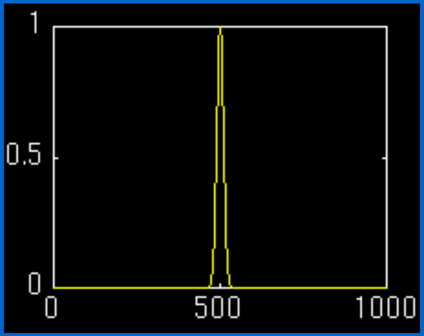


[0 – 300] ms → 75 Hz
[300 – 600] ms → 50 Hz
[600 – 800] ms → 25 Hz
[800 – 1000] ms →10 Hz

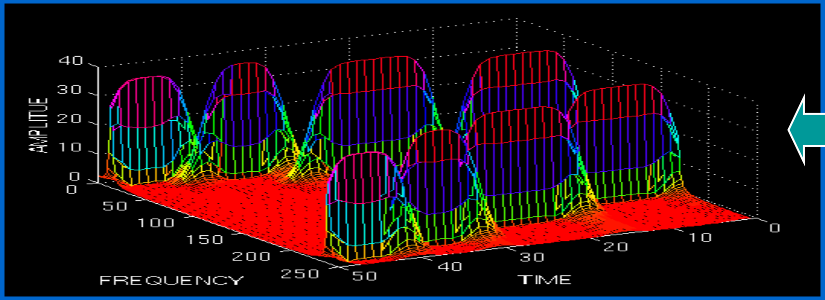
Example (cont'd)

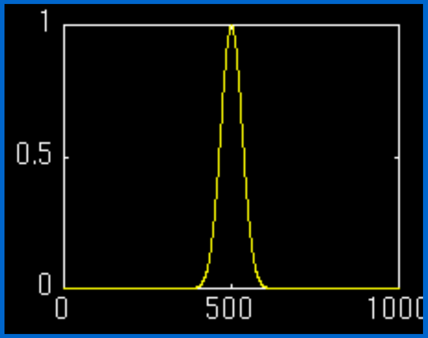
$STFT_f^u(t',u)$





$STFT_f^u(t',u)$

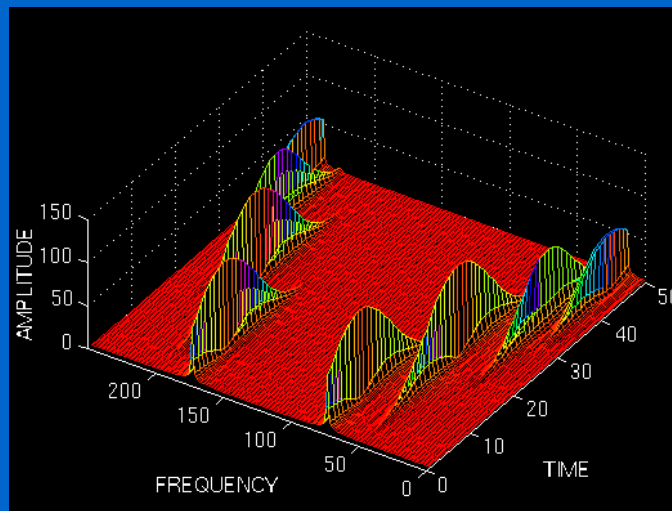


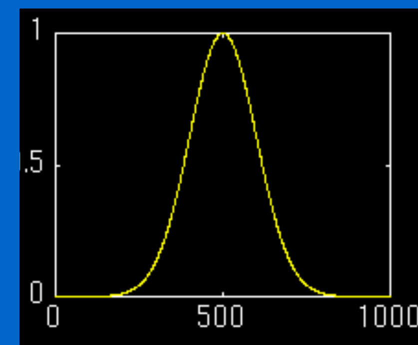


scaled: t/20

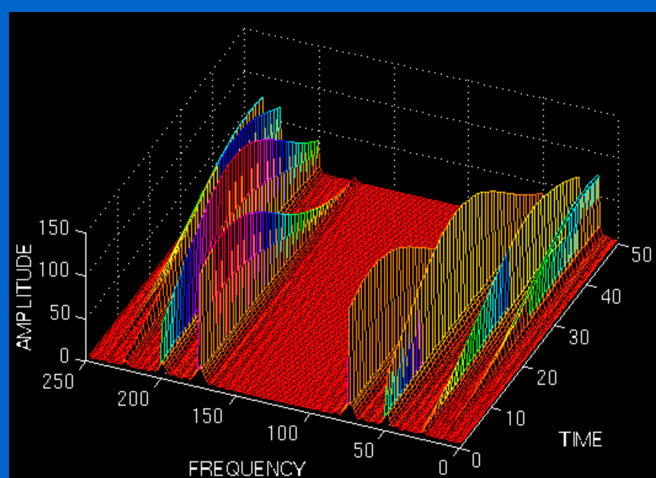
Example (cont'd)

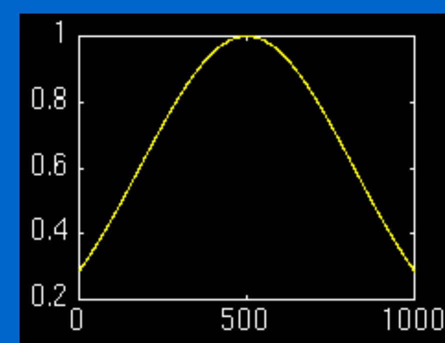
$STFT_f^u(t',u)$





$STFT_f^u(t',u)$





scaled: t/20

Heisenberg (or Uncertainty) Principle

$$\Delta t \cdot \Delta f \geq \frac{1}{4\pi}$$

Time resolution: How well two spikes in time can be separated from each other in the frequency domain.

Frequency resolution: How well two spectral components can be separated from each other in the time domain

$$\Delta t \text{ and } \Delta f \text{ cannot be made arbitrarily small!}$$

Heisenberg (or Uncertainty) Principle

- We cannot know the exact time-frequency representation of a signal.
- We can only know what *interval of frequencies* are present in which *time intervals*.