Spectrum Analysis Using the DFT

Jose Krause Perin

Stanford University

August 13, 2018

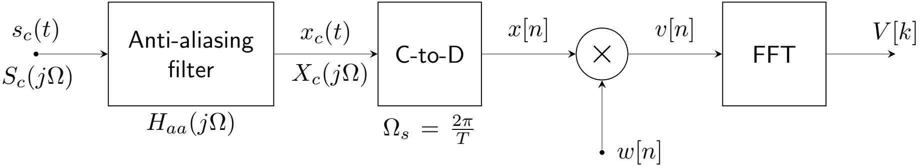
Last lecture

- Sampling the DTFT in frequency domain results in signal replicas in time domain
- The N-point DFT of $x[n]$ is equal to the DTFT of $x[n]$ sampled with period $2\pi/N$, only if $x[n]$ is time-limited with duration $\leq N$
- For sequences longer than N, the N-point DFT is equal to the samples of the windowed DTFT
- Fast Fourier transform (FFT) algorithms compute the DFT with complexity $\mathcal{O}(N \log N)$
- We can use DFT to perform linear convolution (filtering) efficiently using block convolution
- In the overlap and add method, blocks are non-overlapping and the result of circular convolution of each block is added to produce the output signal
- In the overlap and save method, blocks do overlap and we have to discard samples that are unusable due to the circular convolution not being equal to the linear convolution at all points

Outline

Discrete Fourier analysis of analog signals

Block diagram of the spectrum analyzer of an oscilloscope



- Anti-aliasing filter band-limits the analog signal
- Continuous-to-discrete time conversion

$$x[n] = x_c(nT) \Longleftrightarrow X(e^{j\omega}) = \frac{1}{T} X_c(j\omega/T) \quad |\omega| \le \pi \qquad \text{(no aliasing)}$$

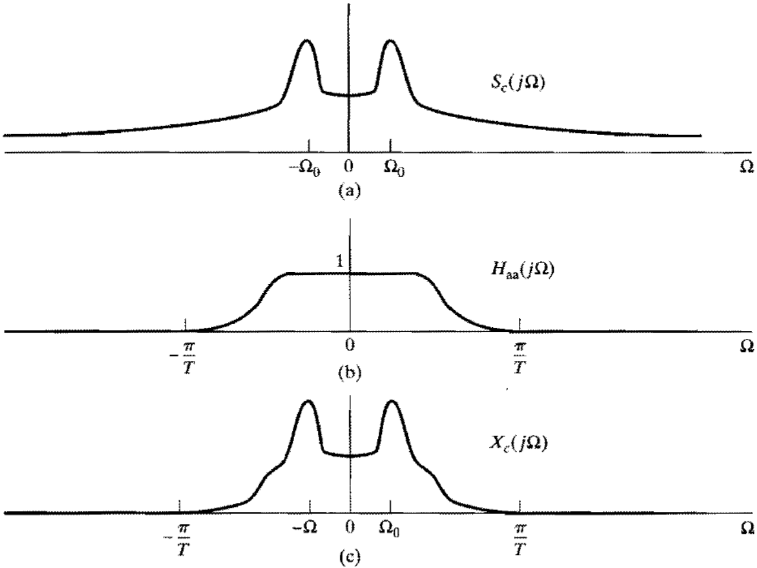
- Windowing time-limits the signal to N samples before FFT:

$$v[n] = x[n]w[n] \Longleftrightarrow V(e^{j\omega}) = \frac{1}{2\pi} X(e^{j\omega}) * W(e^{j\omega})$$

- DFT is a sampled version of the windowed DTFT:

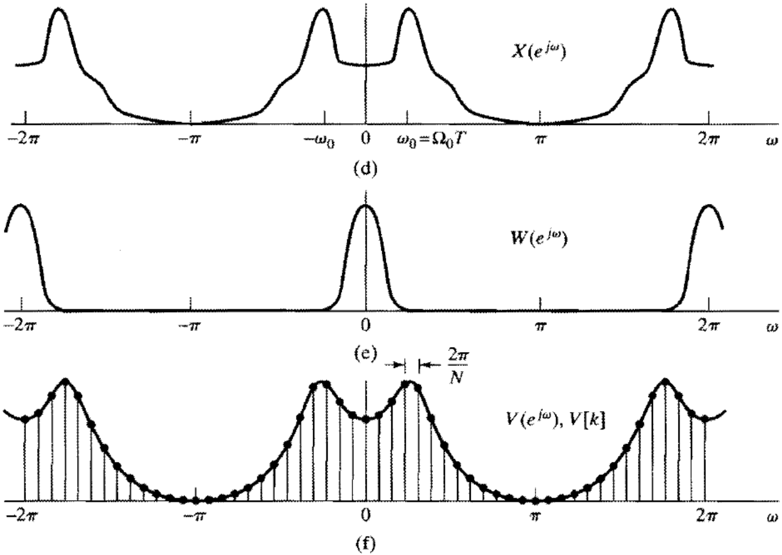
$$V[k] = V(e^{j2\pi/Nk}), \quad k = 0, \dots, N-1$$

Example



$S_c(j\Omega)$
$-\Omega_0$
0
Ω_0
Ω
(a)
$H_{aa}(j\Omega)$
1
$-\frac{\pi}{T}$
0
$\frac{\pi}{T}$
Ω
(b)
$X_c(j\Omega)$
$-\frac{\pi}{T}$
$-\Omega$
0
Ω_0
$\frac{\pi}{T}$
Ω
(c)

Example



$X(e^{j\omega})$
-2π
$-\pi$
$-\omega_0$
0
$\omega_0 = \Omega_0 T$
π
2π
ω
(d)
$W(e^{j\omega})$
-2π
$-\pi$
0
π
2π
ω
(e)
$\frac{2\pi}{N}$
$V(e^{j\omega})$, $V[k]$
-2π
$-\pi$
0
π
2π
ω
(f)

Discrete Fourier analysis of analog signals

- The goal is to estimate $S_c(j\Omega)$ by computing $V(e^{j\omega})$

$$S_c(j\Omega) = V(e^{j\Omega T}), \quad |\Omega| \leq \Omega_s/2 \qquad \text{(ideally)}$$

- However, anti-aliasing filtering and windowing cause disagreement between $S_c(j\Omega)$ and $V(e^{j\Omega T})$
- In particular, windowing limits the resolution. As a result, the peaks in $V(e^{j\omega})$ look broader than they actually are in $S_c(j\Omega)$
- Choosing good windows is crucial for spectrum analysis using the DFT

DFT analysis of sinusoidal signals

As another example, let's consider the sinusoidal signal

$$s_c(t) = \cos(\Omega_0 t) + 0.5\cos(\Omega_1 t).$$

Assuming ideal sampling with no aliasing, we obtain the discrete-time signal

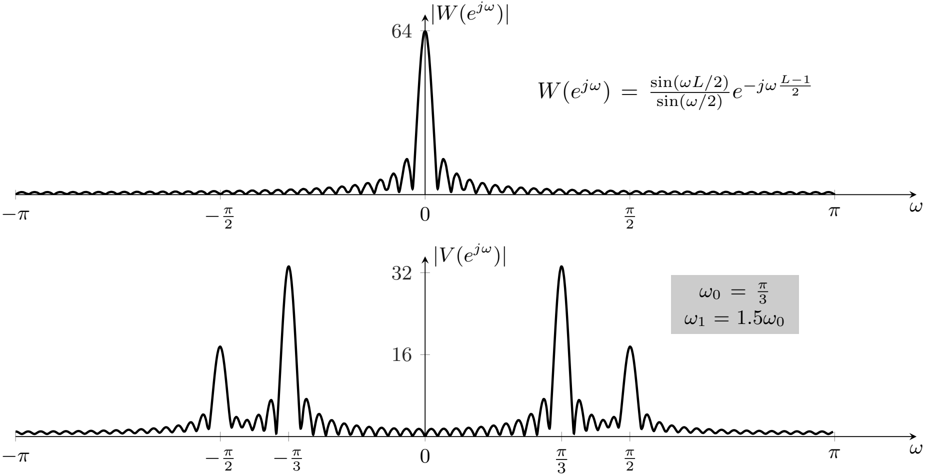
$$x[n] = \cos(\omega_0 n) + 0.5\cos(\omega_1 t),$$

where $\omega_0 = \Omega_0 T$ and $\omega_1 = \Omega_1 T$.
After windowing

$$\begin{aligned} v[n] &= x[n]w[n] \\ V(e^{j\omega}) &= 0.5W(e^{j(\omega-\omega_0)}) + 0.25W(e^{j(\omega-\omega_1)}) \\ &\quad + 0.5W(e^{j(\omega+\omega_0)}) + 0.25W(e^{j(\omega+\omega_1)}), \quad |\omega| \leq \pi \end{aligned}$$

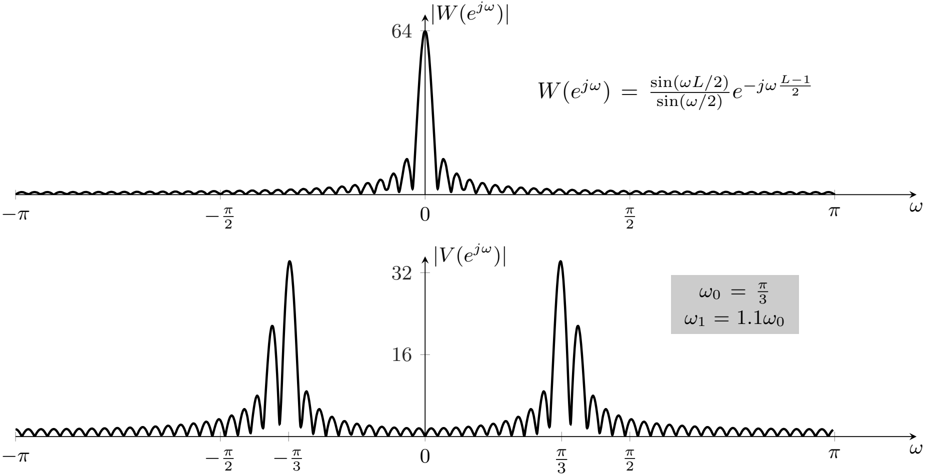
Rectangular window

Rectangular window of length $L = 64$



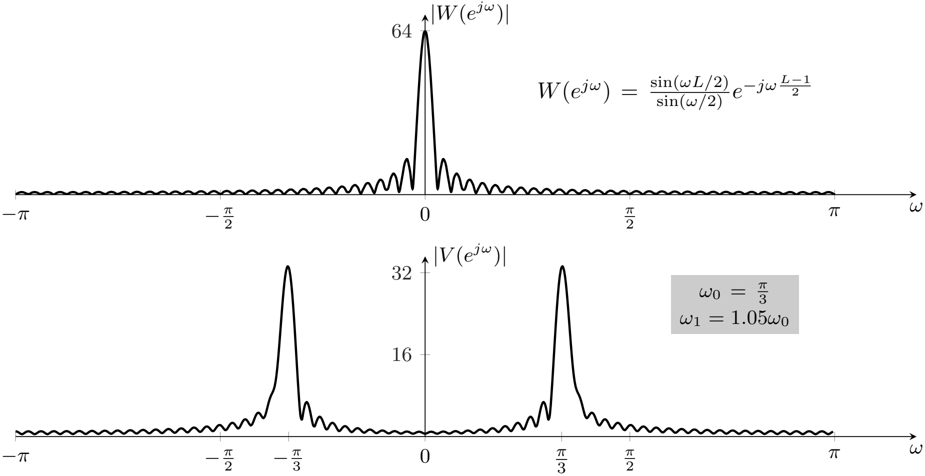
Rectangular window

Rectangular window of length $L = 64$



Rectangular window

Rectangular window of length $L = 64$



Problems:

Window characteristics

Main characteristics of rectangular window

- The main-lobe width is $4\pi/L$. This determines the resolution.
- The first side lobe is -13 dB below the main lobe.
- The higher the energy (area) under the side lobes compared to the main lobe, the greater the leakage will be.

Desired window characteristics:

- Small side-lobe energy (area) to reduce leakage
- Small main-lobe width to improve resolution

These are conflicting requirements.

Revisiting the Kaiser window

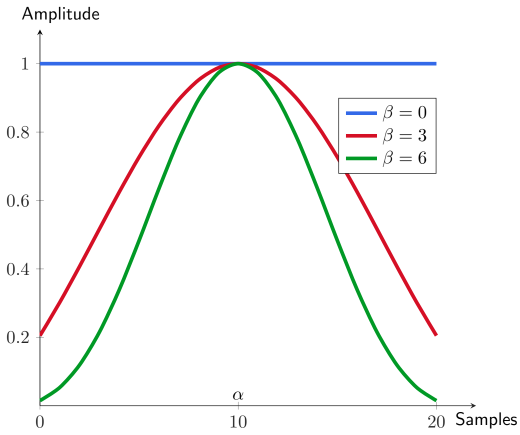
The **Kaiser window** offers a nearly optimal trade-off between main-lobe width and side-lobe area.

$$w[n] = \begin{cases} \dfrac{I_0\left(\beta\sqrt{1-(n-\alpha)^2/\alpha^2}\right)}{I_0(\beta)}, & 0 \leq n \leq L-1 \\ 0, & \text{otherwise} \end{cases},$$

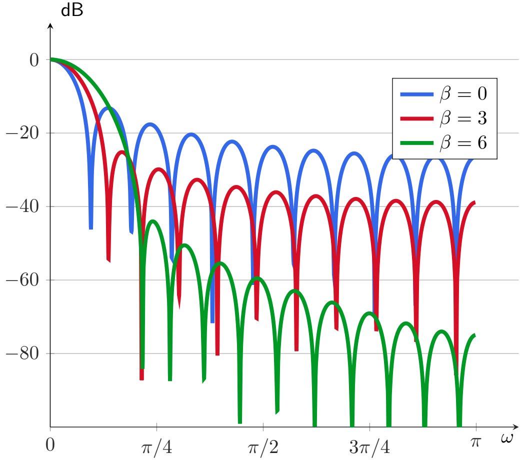
where $\alpha = (L-1)/2$, β is a design parameter, and $I_0(\cdot)$ is the **modified Bessel function of first kind and order 0**.

Kaiser window

Time domain



Frequency domain



For fixed window length L, β controls the trade-off between main-lobe width and side-lobe area.

Kaiser window

Define

- Δ_{ml} one-sided main-lobe width
- A_{sl} relative side-lobe level (in dB)

$$A_{sl} = \frac{\text{amplitude of main lobe}}{\text{amplitude of largest side lobe}} \quad \text{dB}$$

A_{sl} *approximately* only depends on β. See section 10.2.2 in textbook for analytical equation.

The following approximation was derived by Kaiser and Schafer:

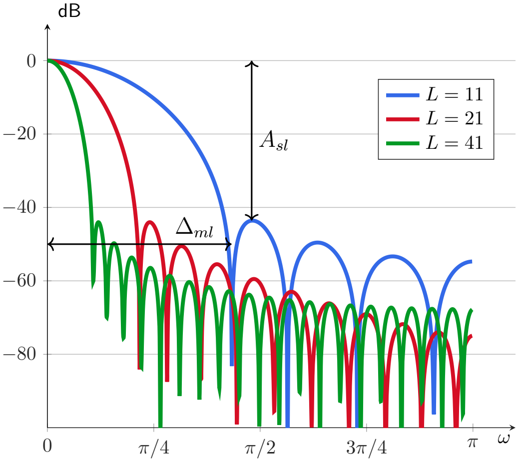
$$L - 1 \approx \frac{24\pi(A_{sl} + 12)}{155\Delta_{ml}}$$

Conclusions:

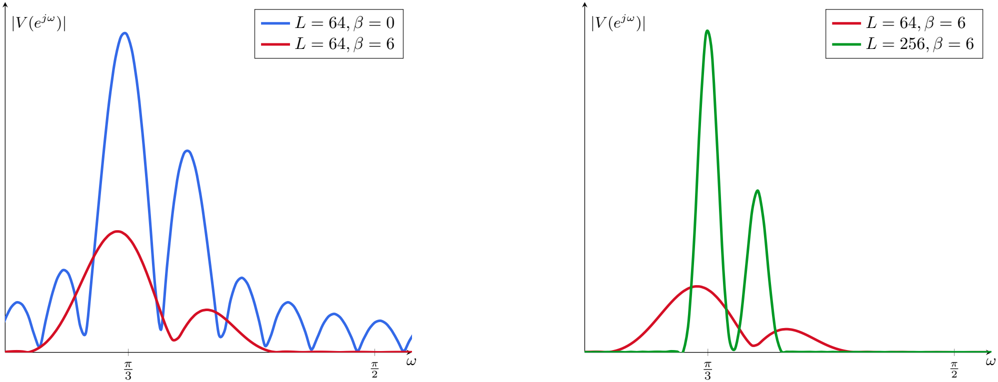
- β controls the side-lobe area A_{sl} (leakage)
- Main-lobe width Δ_{ml} is directly proportional to $L-1$ i.e., increasing the window length L improves resolution

Kaiser window

Spectrum of Kaiser window for fixed $\beta = 6$.
Amplitude of largest side-lobe A_{sl} remains approximately constant (fixed β), while main-lobe width Δ_{ml} is inversely proportional to $L-1$



Kaiser windowing of sinusoidal signals



See script Canvas/Files/Matlab/`spectrum_analysis_of_sinusoid.m`

Main conclusions

- β controls the side-lobe area A_{sl} (leakage)
- Increasing the window length L improves resolution
- Increasing the FFT size does not improve resolution

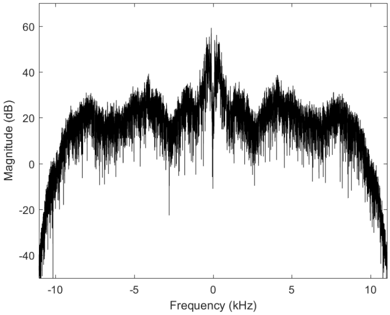
Outline

Spectrum Analysis Using the DFT

Time-Dependent Fourier Transform

DFT of a long signal

The DFT of the speech signal Canvas/Files/Matlab/`dft_speech.wav`



- It has duration of about 4 seconds
- 110033 samples
- Not very informative. What are the main frequency components at the beginning of the third second?

Time-dependent Fourier transform (TDFT)

Time-dependent Fourier transform (TDFT) or **short-time Fourier transform (STFT)** is defined as

$$X[n, \lambda) = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m}$$

Notation: n is discrete, while λ is continuous. This is why n appears with a square bracket $[$ and λ appears with parenthesis $)$.

If the window $w[n]$ has finite length L

$$X[n, \lambda) = \sum_{m=0}^{L-1} x[n+m]w[m]e^{-j\lambda m}$$

Time-dependent Fourier transform (TDFT)

Two interpretations for $X[n, \lambda)$:

1. For fixed n, $X[n, \lambda)$ is the DTFT of $x[n+m]w[m]$.

$$X[n, \lambda) = \mathcal{F}_\lambda\{x[n+m]w[m]\} \qquad \text{(fixed } n\text{)}$$

   Hence, at fixed n, $X[n, \lambda)$ has all the properties of the DTFT.

2. For fixed λ, $X[n, \lambda)$ is the result of band-pass filtering the signal $x[n]$ by the time-reversed window $W(e^{-j\omega})$ centered at frequency λ

$$\begin{aligned}
X[n, \lambda) &= \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m} && \text{(definition)} \\
&= \sum_{l=-\infty}^{\infty} x[l]w[-(n-l)]e^{j\lambda(n-l)} && \text{(change of variables } l = n+m\text{)} \\
&= x[n] * h_\lambda[n] && \text{(fixed } \lambda\text{)}
\end{aligned}$$

   where

$$h_\lambda[n] = w[-n]e^{j\lambda n} \Longleftrightarrow H_\lambda(e^{j\omega}) = W(e^{-j(\omega-\lambda)})$$

Sampling and displaying the TDFT: spectrograms

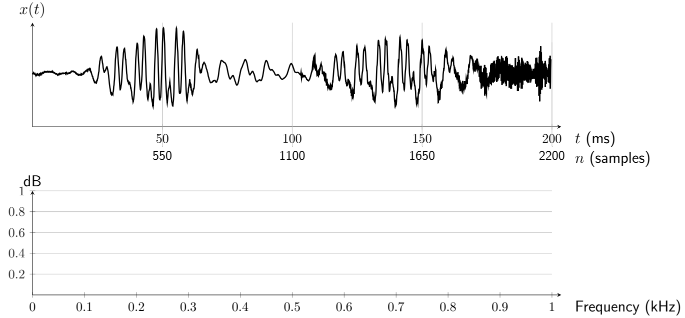
Spectrogram is a useful way of visualizing the TDFT
Sampling both in time and frequency. We sample with period R in time and with period $2\pi/N$ in frequency. R is the block spacing and N is the DFT length

$$\begin{aligned} X[rR,k] &= X[rR, 2\pi k/N), \quad k = 0, \ldots, N-1 && \text{(sample in time and frequency)} \\ &= \sum_{m=0}^{L-1} x[rR+m]w[m]e^{-j(2\pi/N)km} && \text{(finite-length window)} \end{aligned}$$

For fixed r, $X[rR,k]$ is the N-point DFT of $x[rR+m]w[m]$

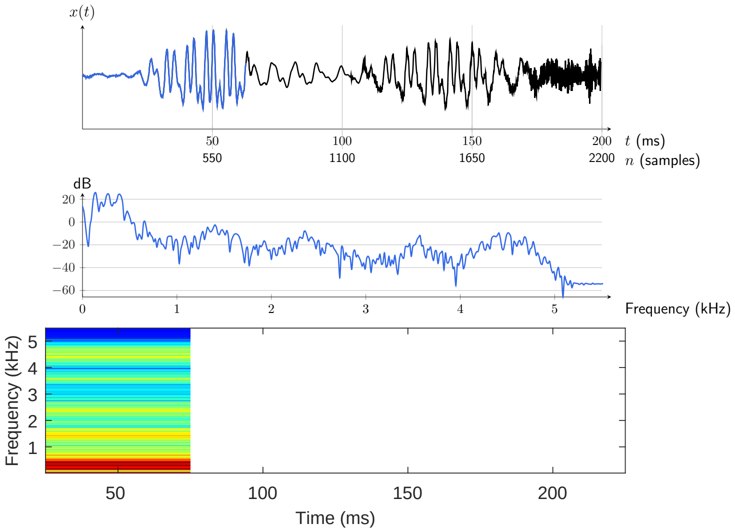
Sampling and displaying the TDFT: spectrograms

Example using Kaiser window of length $L = 700$, FFT size $N = 700$, and $R = 550$.



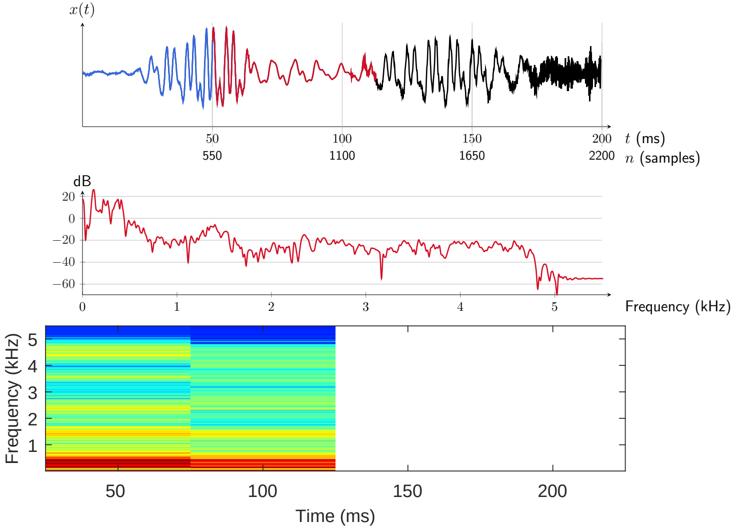
Sampling and displaying the TDFT: spectrograms

Example using Kaiser window of length $L = 700$, FFT size $N = 700$, and $R = 550$.



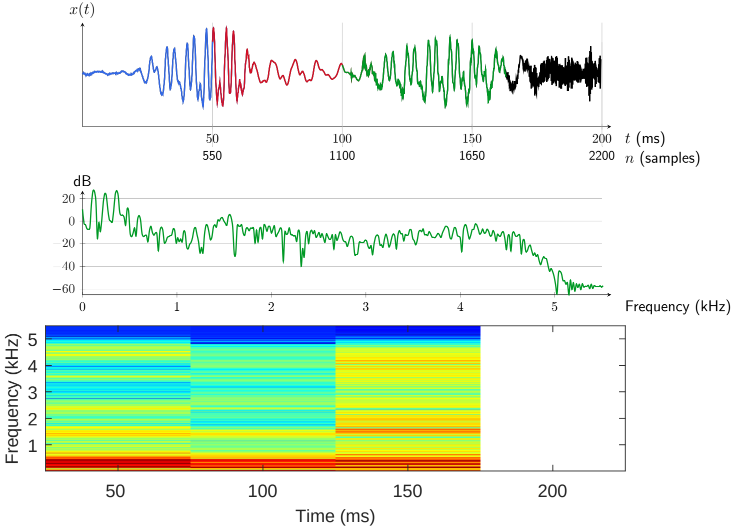
Sampling and displaying the TDFT: spectrograms

Example using Kaiser window of length $L = 700$, FFT size $N = 700$, and $R = 550$.



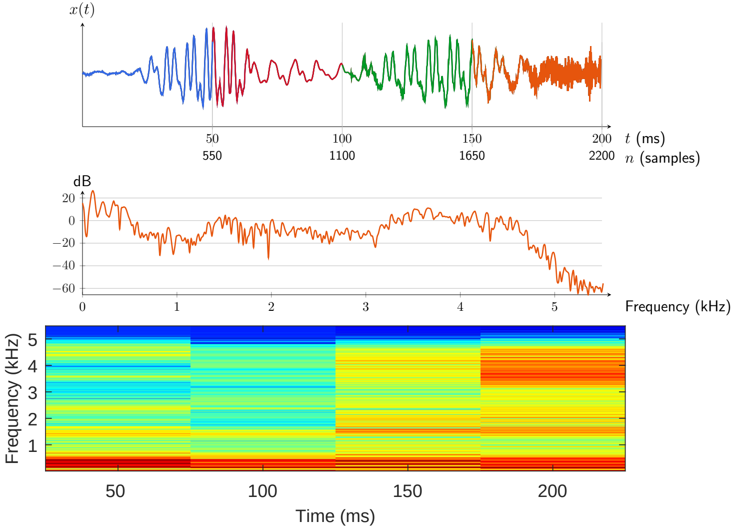
Sampling and displaying the TDFT: spectrograms

Example using Kaiser window of length $L = 700$, FFT size $N = 700$, and $R = 550$.



Sampling and displaying the TDFT: spectrograms

Example using Kaiser window of length $L = 700$, FFT size $N = 700$, and $R = 550$.



Spectrogram in Matlab

To produce spectrogram plot

```
>> spectrogram(x, window, noverlap, nfft)
```
(Matlab notation)

Using notation of this lecture notes:

```
>> spectrogram(x, kaiser(L, beta), L-R, N)
```
(this lecture's notation)

The time blocks overlap by $L - R$ samples. Kaiser window was just as an example, any other window would work.

We can also use:

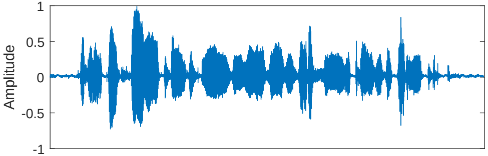
```
>> [s,f,t] = spectrogram(x, kaiser(L, beta), L-R, N, fs)
```

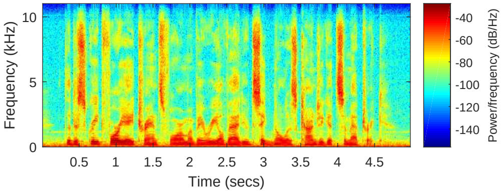
This will not produce the plot, but it'll return the magnitude `s` in dB, the frequency vector `f` in Hz, and the time vector `t` in s.

Spectrogram example

Spectrogram Canvas/Files/Matlab/`dft_speech.wav`

```
>> spectrogram(x, kaiser(L=441, beta=6), R=220,...
 N = 441, Fs=22050, 'yaxis')
```





Importance of window length

Consider the following signal

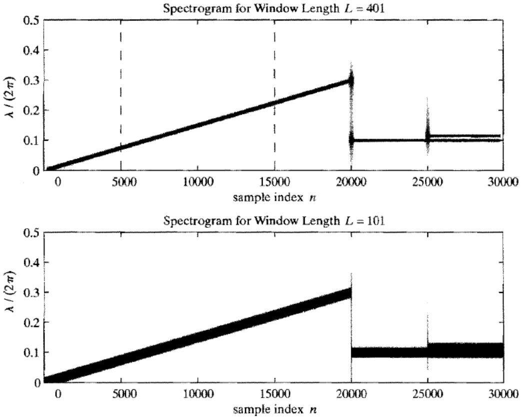
$$x[n] = \begin{cases} 0, & n < 0 \\ \cos(\alpha_0 n^2), & 0 \leq n \leq 20000 \\ \cos(0.2\pi n), & 20000 < n \leq 25000 \\ \cos(0.2\pi n) + \cos(0.23\pi n), & n > 25000 \end{cases}$$

This signal has three sections

1. linear chirp section. Frequency increases linearly
2. Single tone at frequency 0.2π
3. Two tones: one at frequency 0.2π and another at 0.23π

Importance of window length

Spectrogram of $x[n]$ using the Hamming window of length 401 and 101.



Important: window length determines the resolution.

Inverting the TDFT

We want to reconstruct $x[n]$ from its TDFT $X[n, \lambda)$:

$$x_r[n] = x[rR + n]w[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[rR, k] e^{j(2\pi/N)kn} \quad 0 \le n \le L - 1$$

If the window is non-zero, we can divide it out to recover the signal $x[rR + n]$ over the window interval. If the windows overlap, we can recover all the original samples. This requires

$$R \le L \le N$$

The particular case when $R = L \le N$ is called **maximally decimated condition**.

Summary

- Leakage and resolution are important considerations in spectrum analysis
- By properly choosing windows we can minimize these issues
- Kaiser window is a nearly optimal choice. Must choose correct β and window length L
- β controls the ratio between the amplitudes of the main-lobe and the largest side-lobe i.e., β controls the amount of leakage.
- The larger the main-lobe width, the smaller the resolution
- By increasing the window length we reduce the main-lobe width and consequently improve the resolution
- Time-dependent Fourier transform or short-time Fourier transform allows us to keep track of frequency variation in time
- Spectrogram is a commonly used way to display the TDFT
- In the spectrogram the TDFT is sampled both in time and in frequency
- The window length determines the resolution of the spectrogram