Introduction to Audio and Music Engineering

Lecture 21

Topics:

- Frequency resolution of DFT
- Musical sound timbre analysis
- Short-time Fourier transform
- Spectrogram
- Time – frequency resolution of STFT

Frequency Resolution of the DFT

Two key points to remember …

- The maximum frequency in a spectrum is determined by R, the sampling rate. (f_{max} = R/2)
- The number of frequencies represented in the spectrum is N/2, where N is the total number of samples.

Therefore the frequency spacing (resolution) of the DFT is …

$$\frac{\text{Frequency Range}}{\text{\# of points}} = \frac{R/2}{N/2} = \frac{R}{N} \equiv \Delta f$$

But N/R = time length of the recording, we will call this T.

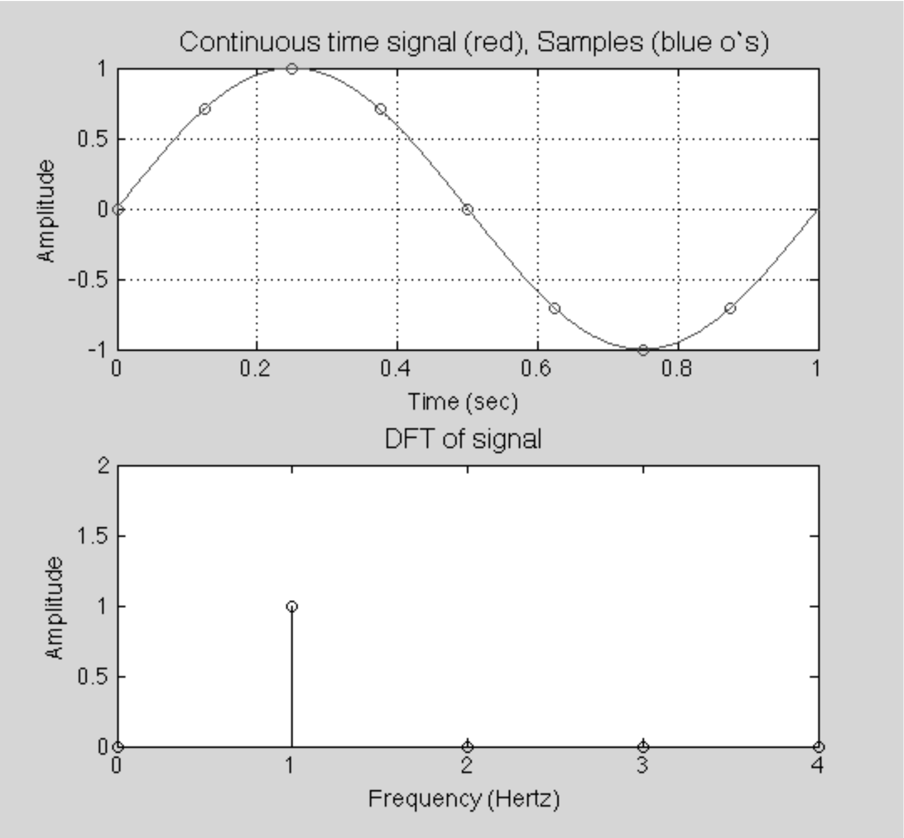
so … $\Delta f = 1/T$

Frequency resolution continued …

So for:

T = 1 sec → Δf = 1 Hz
T = 2 sec → Δf = ½ Hz
T = 4 sec → Δf = ¼ Hz
etc.

Example: R = 8, N = 8, (f_{signal} = 2 Hz)
→ T = 1 sec
→ f_{max} = 4 Hz (R/2)
→ Δf = 1 Hz



To increase f_{max} you increase R.

To decrease Δf you increase T.

Increasing R does not improve Δf!!!

Examples

<u>Increase R, keep N the same.</u>

R = 16, N = 8, (f_{signal} = 2 Hz)

→ T = 0.5 sec

→ f_{max} = 8 Hz (R/2)

→ Δf = 2 Hz

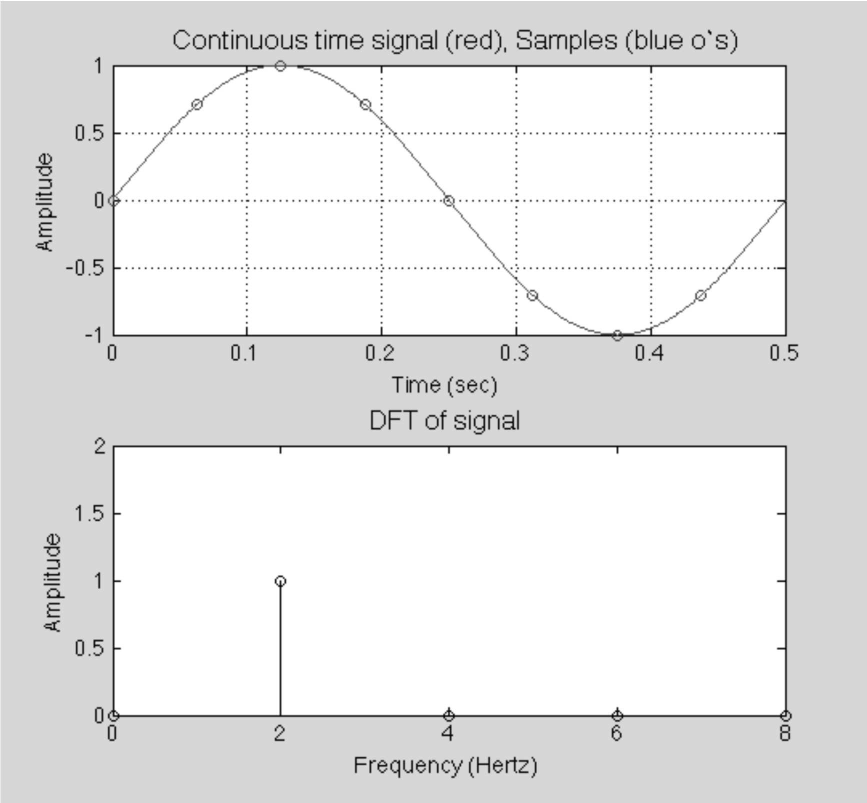
<u>Keep R the same, increase N</u>

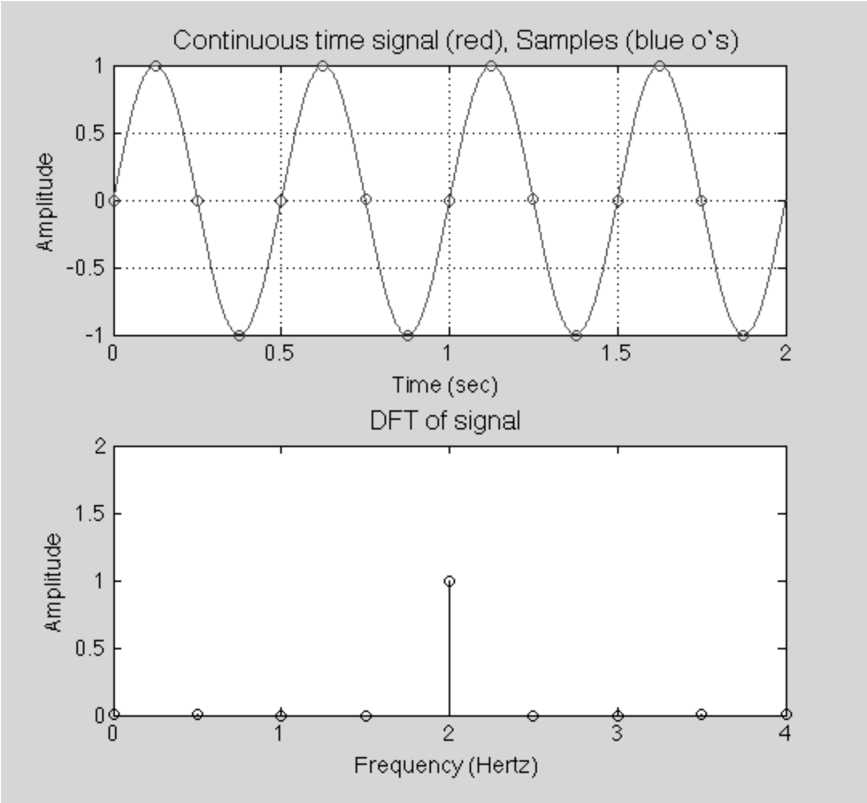
R = 8, N = 16, (f_{signal} = 2 Hz)

→ T = 2 sec

→ f_{max} = 4 Hz (R/2)

→ Δf = 0.5 Hz





Examples

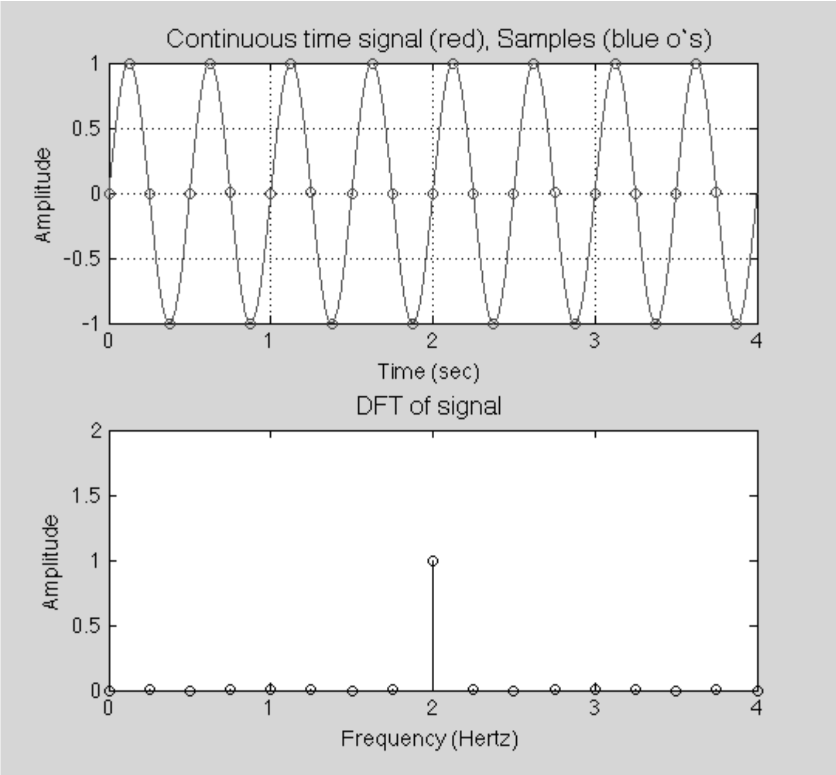
Hold R constant, keep increasing N

$R = 8$, $N = 32$, ($f_{signal} = 2$ Hz)

→ $T = 4$ sec

→ $f_{max} = 4$ Hz (R/2)

→ $\Delta f = 0.25$ Hz



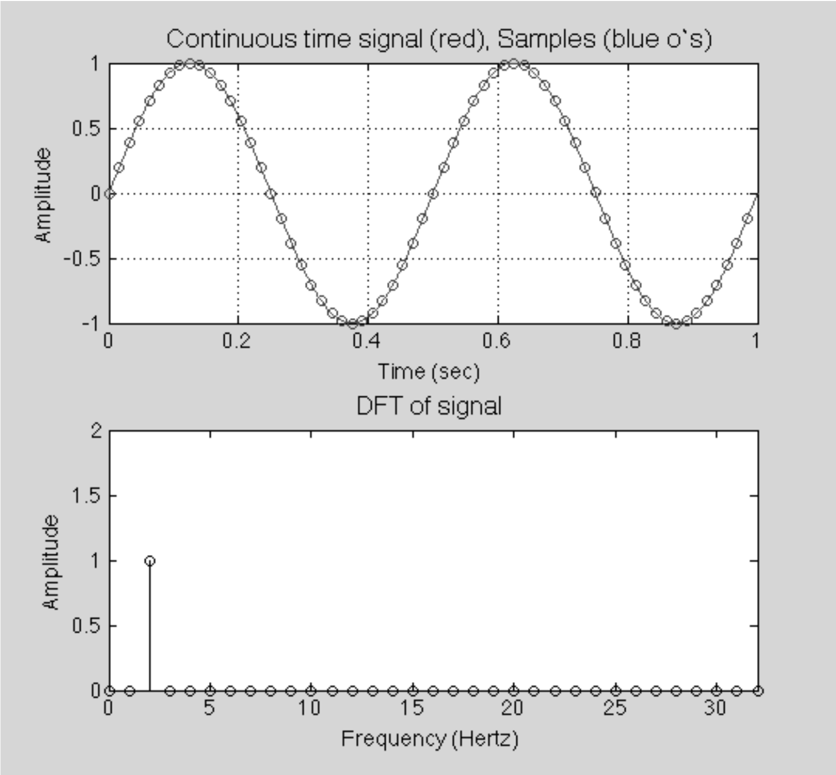
Increase both R and N a lot …

$R = 64$, $N = 64$, ($f_{signal} = 2$ Hz)

→ $T = 1$ sec

→ $f_{max} = 32$ Hz (R/2)

→ $\Delta f = 1$ Hz



Spectral Analysis of Music

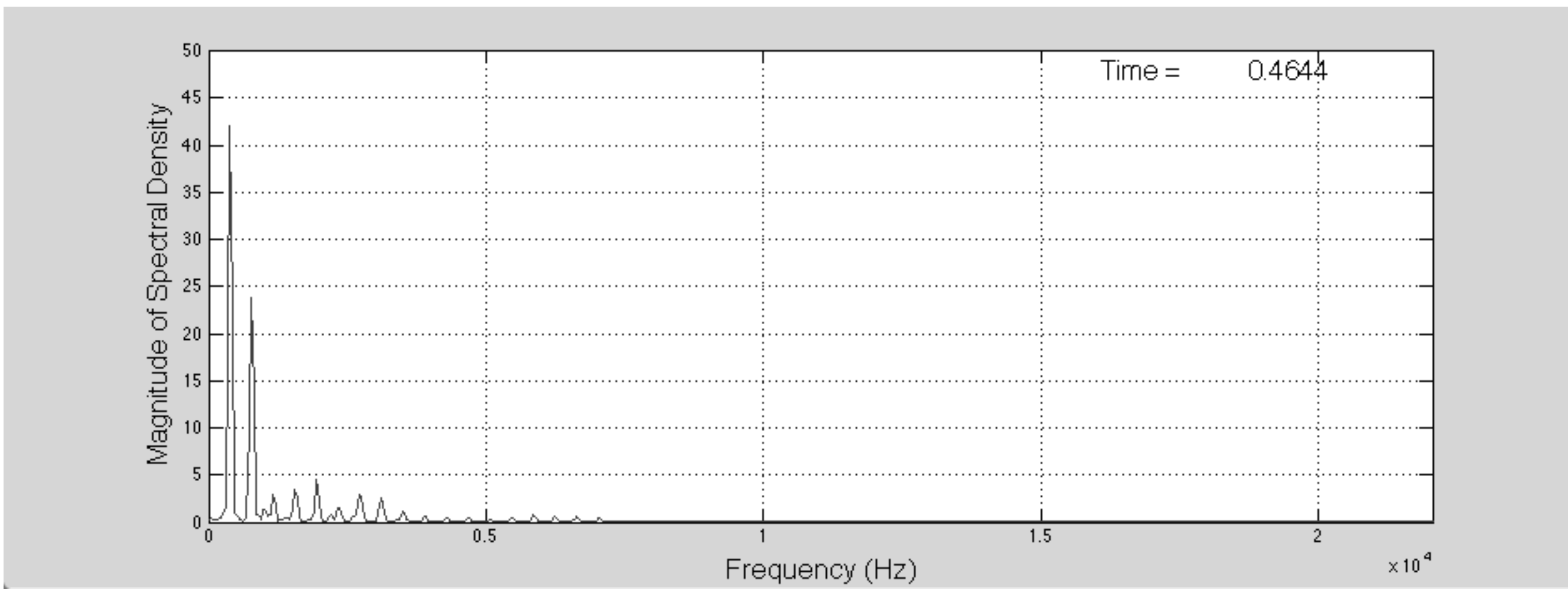
Increase R to 44,100 Hz → f_{max} = 22,050 Hz, set N = 1024

What is Δf?

T = N/R = 1024/44,100 = 0.0232 sec

Δf = 1/T = 43.07 Hz

Typical spectrum of a musical signal …



Improving the frequency resolution

Increase N obviously …

$N = 1024 \rightarrow \Delta f = 43$ Hz

$N = 2048 \rightarrow \Delta f = 21.5$ Hz

$N = 4096 \rightarrow \Delta f = 10.25$ Hz

Musical example:

What frequency resolution do I need to tell A4 from A#4?

A4 $\rightarrow$ 440 Hz

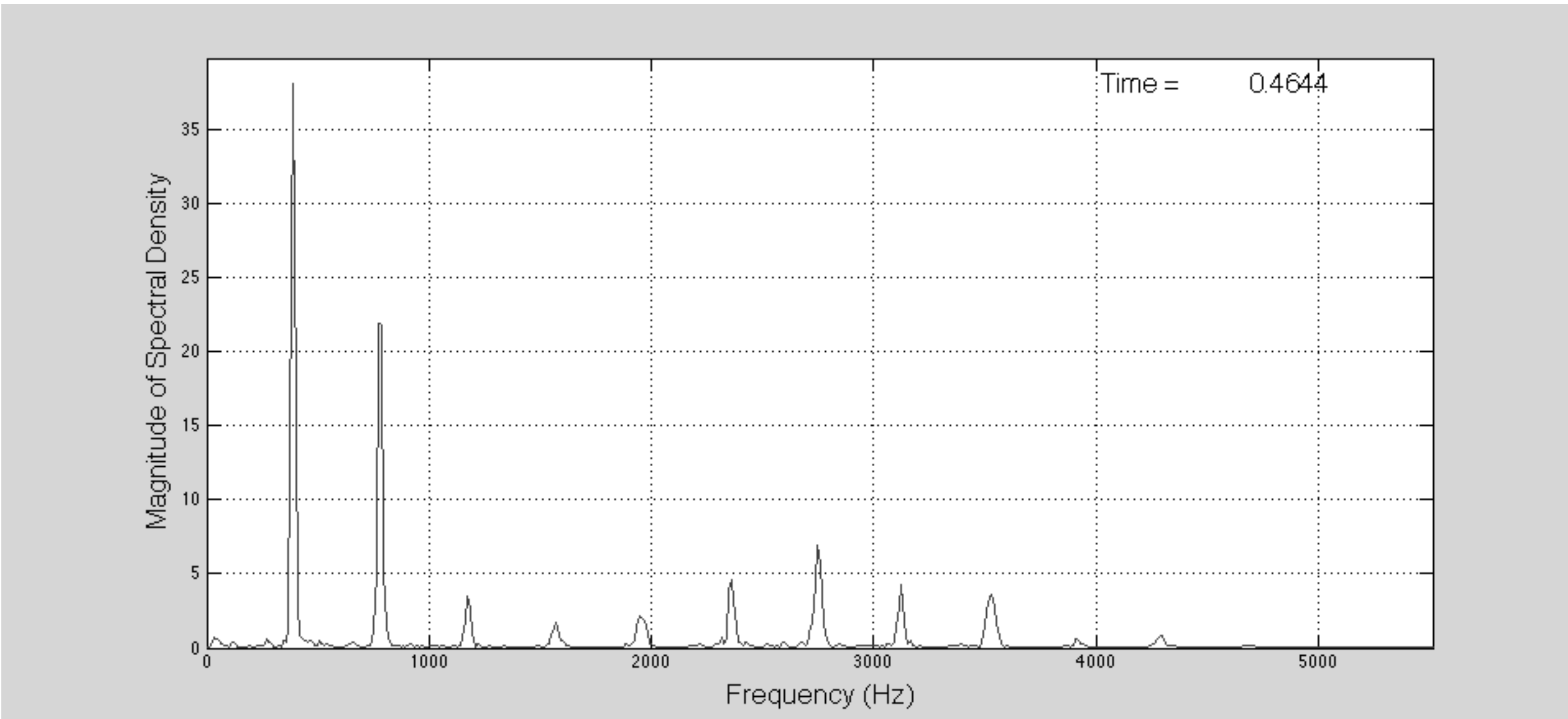
A#4 $\rightarrow$ $(2^{1/12})*440 = 1.059 \times 440 = 466$ Hz.

$\Delta f = 26$ Hz

Not a lot of interesting information up at 22 kHz, so we can reduce the sampling rate.

There is not a lot of interesting information up at 22 kHz, so we can reduce the sampling rate … this applies for many audio signals.

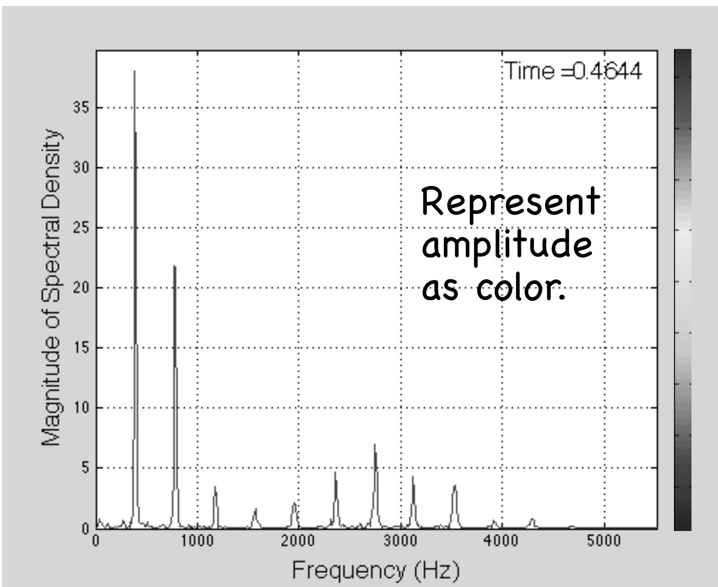
Reduce the sample rate to 10,025, so $f_{max} \approx 5{,}000$ Hz



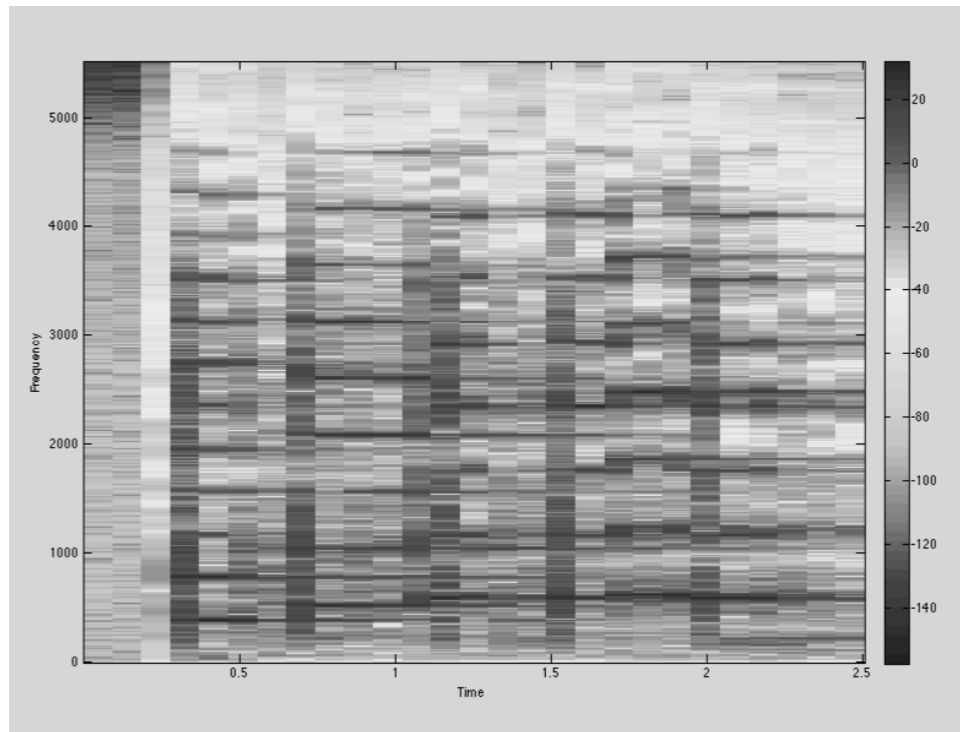
Spectrogram

When a signal is changing over time we use the spectrogram to view the signal in frequency and time.

Single time "slice"

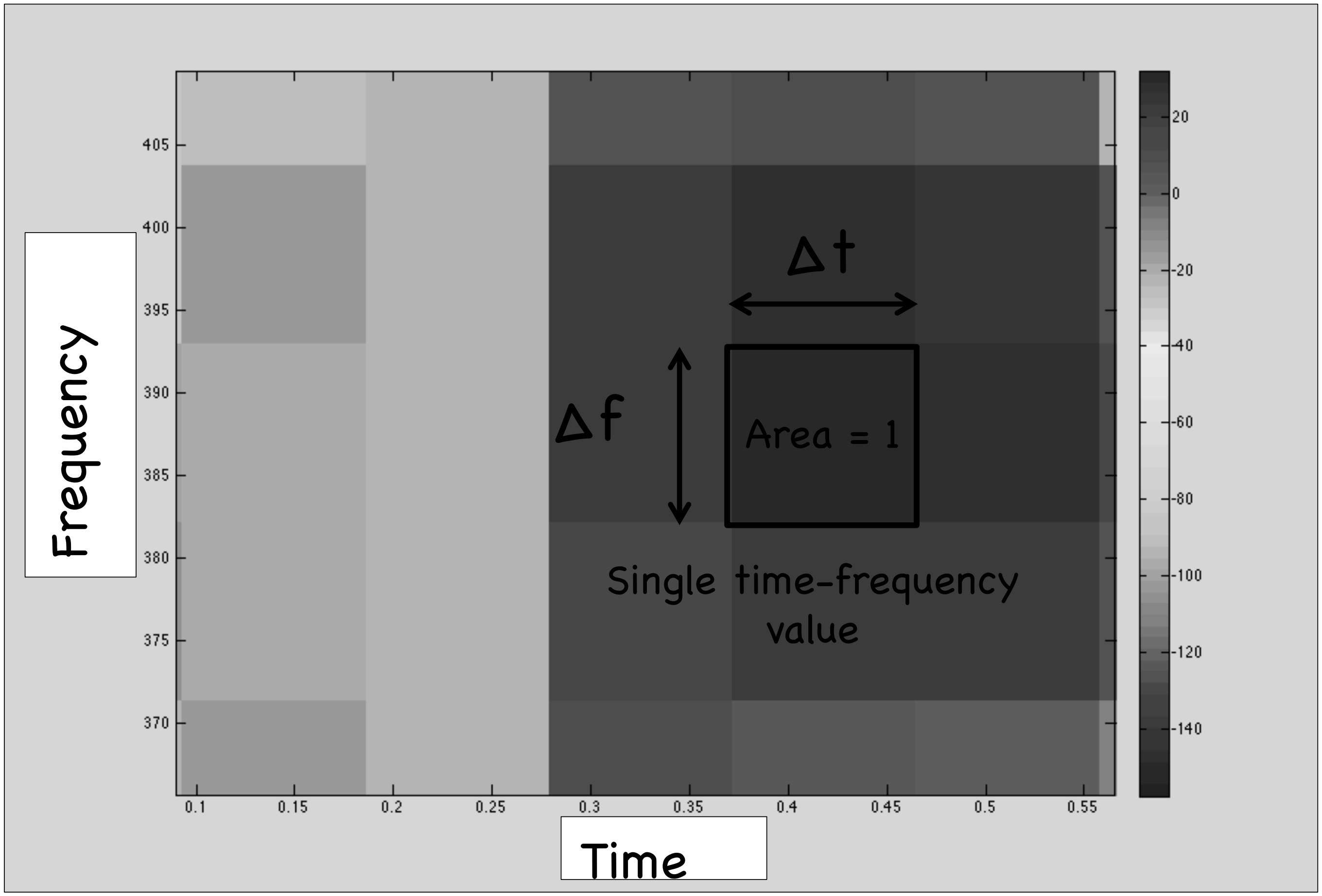


Spectrogram



Time – frequency resolution tradeoff in a spectrogram

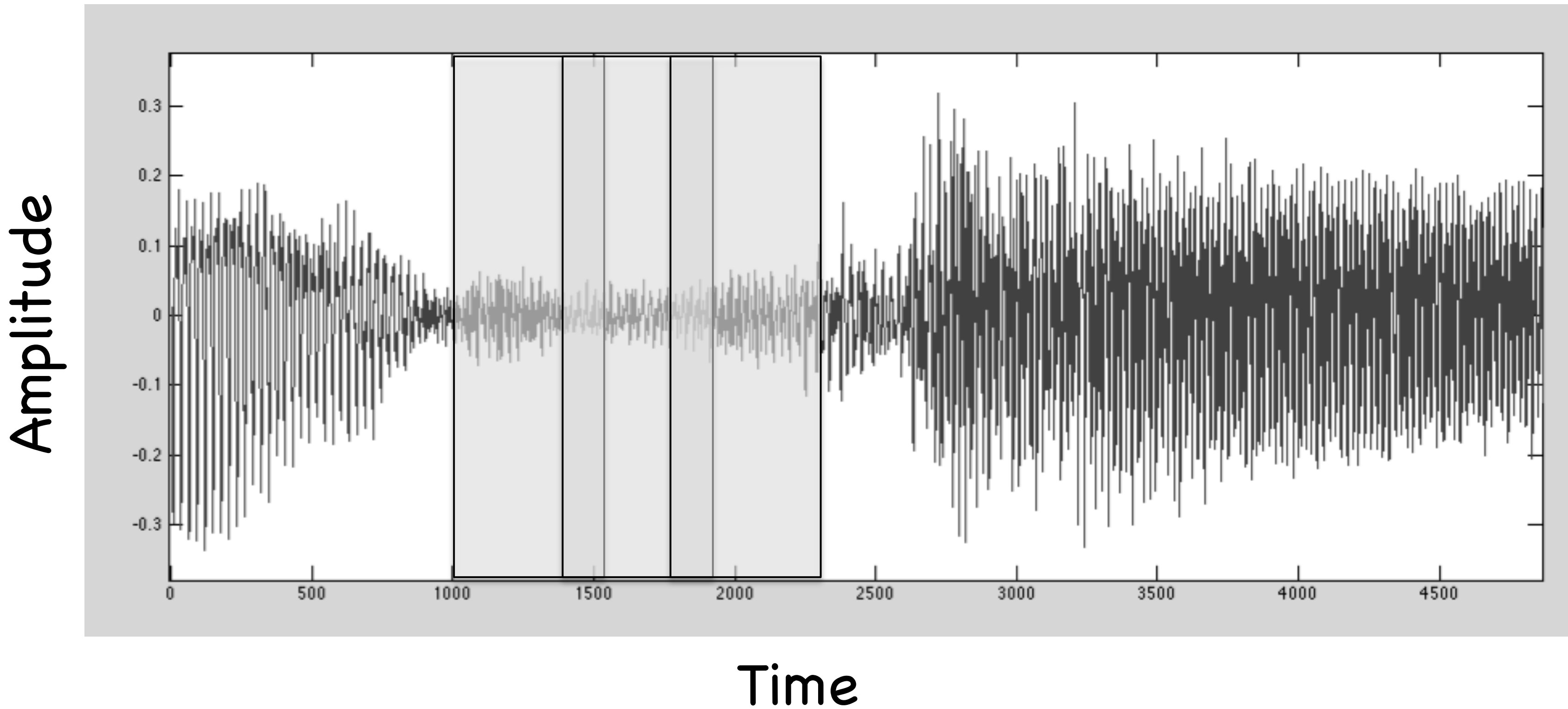
Zoom in on spectrogram …



$\Delta f = R/N \qquad \Delta t = T = N/R$

so … $\Delta f \Delta t = 1$

Using overlap in spectrogram



Gives the illusion of better time resolution…