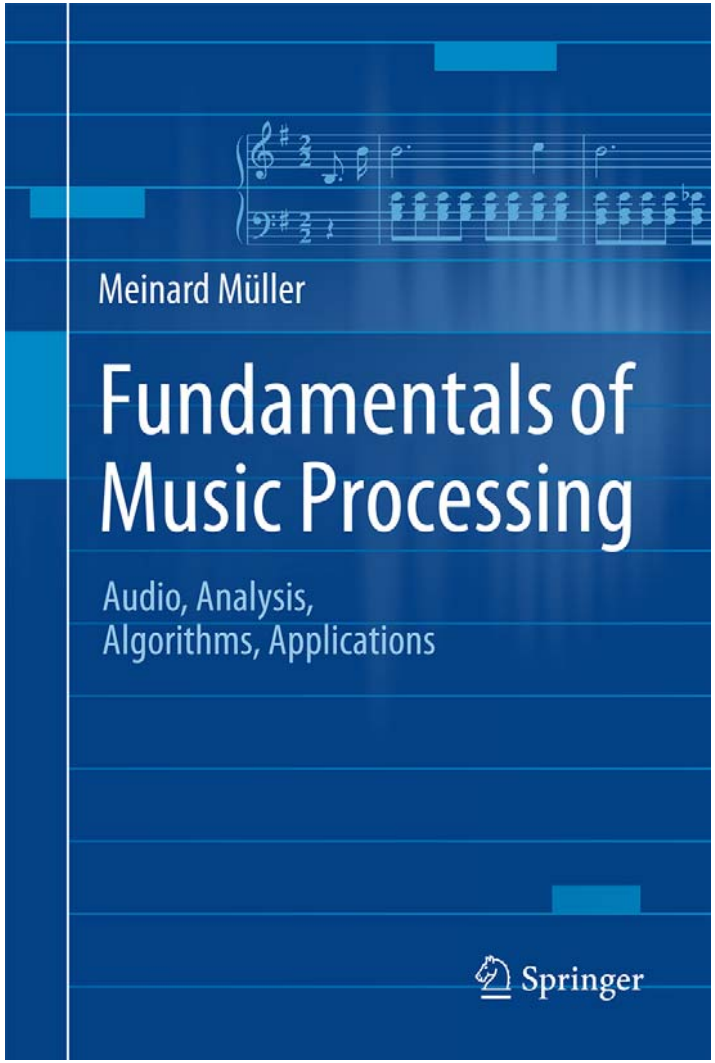
Book: Fundamentals of Music Processing



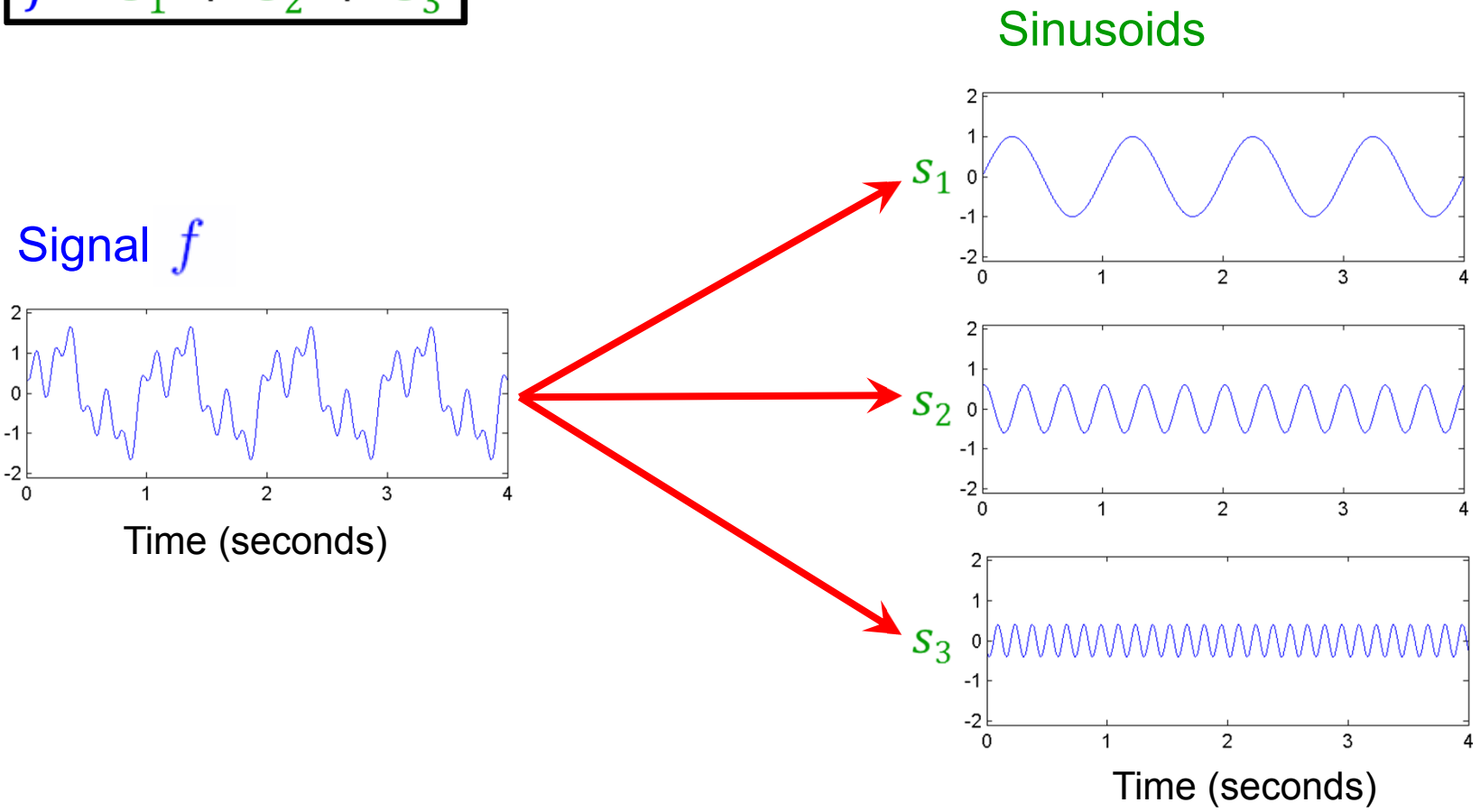
Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

Fourier Transform

Idea: Decompose a given signal into a superposition of sinusoids (elementary signals).

$$f = s_1 + s_2 + s_3$$

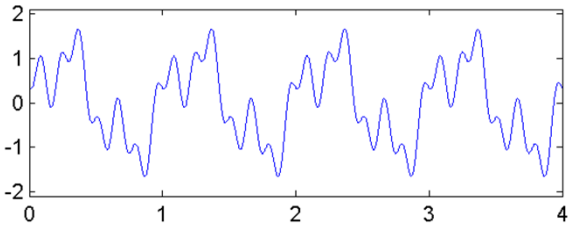


Fourier Transform

Each sinusoid has a physical meaning
and can be described by three parameters:

$$f = s_1 + s_2 + s_3$$

Signal f

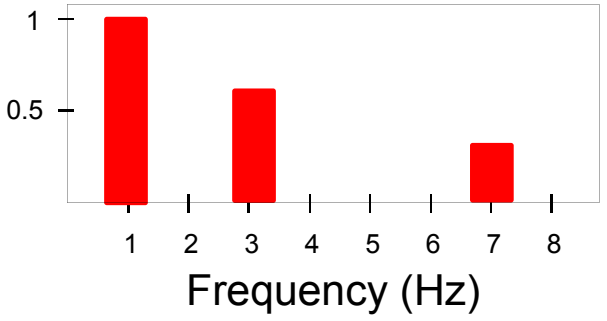


$A_1 = 1$
$\omega_1 = 1$
$\varphi_1 = 0$

$A_2 = 0.6$
$\omega_2 = 3$
$\varphi_2 = -0.2$

$A_3 = 0.4$
$\omega_3 = 7$
$\varphi_3 = 0.4$

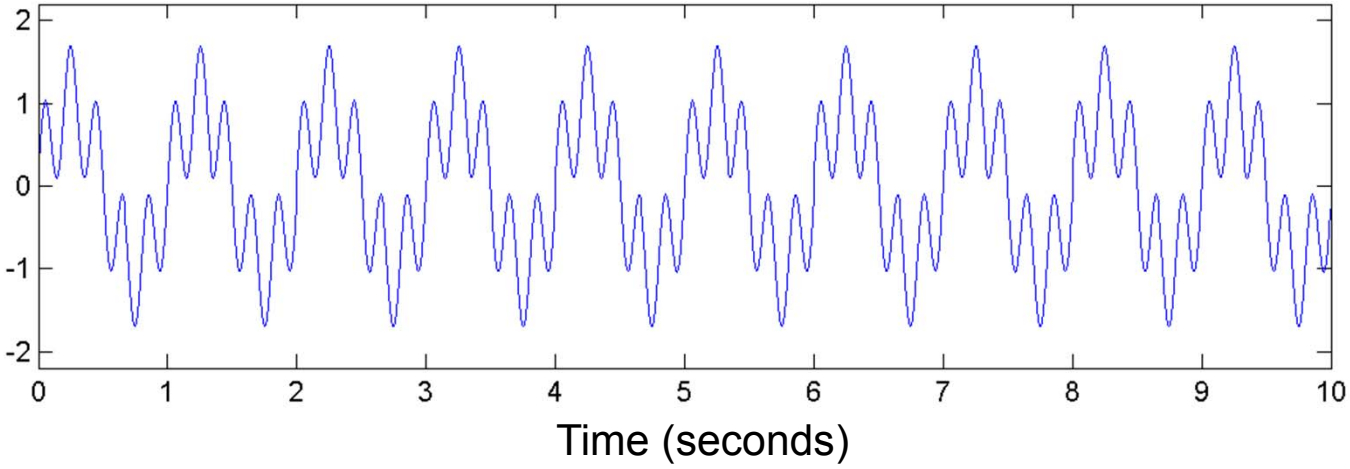
Fourier transform $|\hat{f}|$



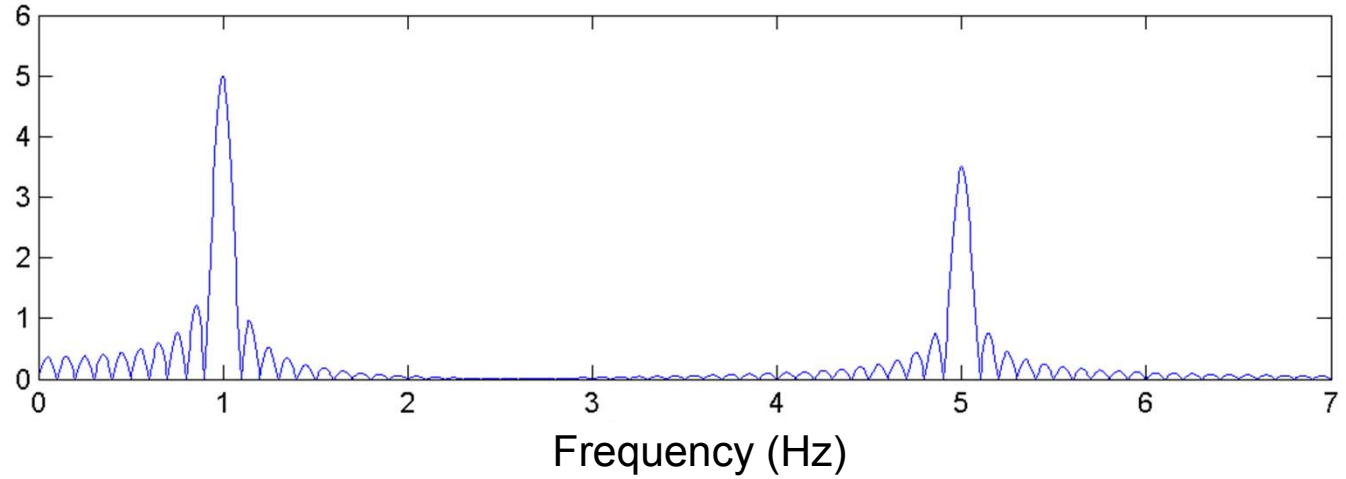
Fourier Transform

Example: Superposition of two sinusoids

f



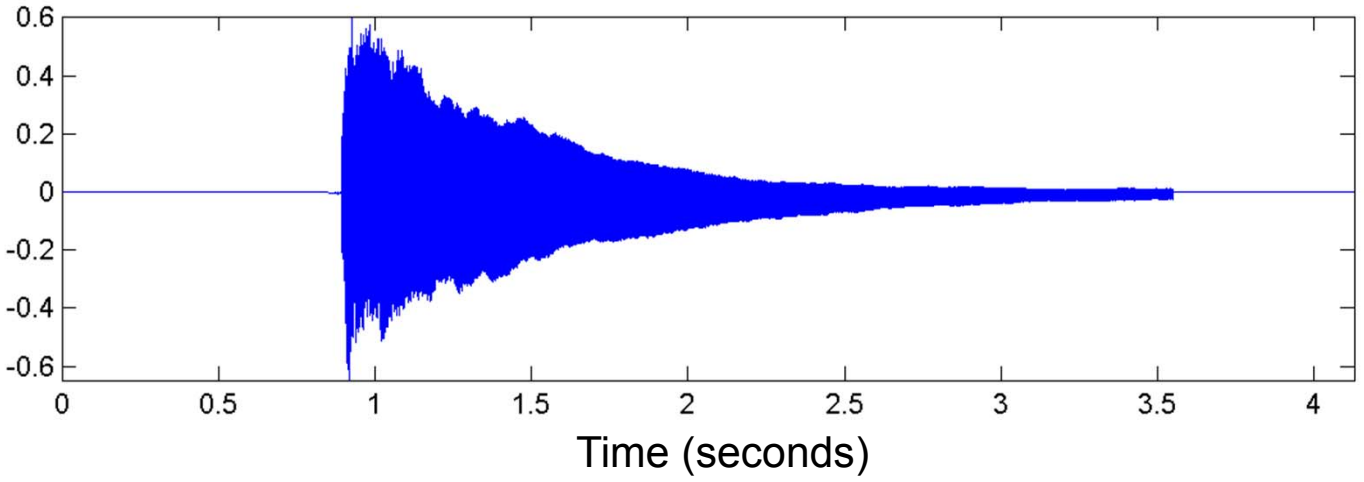
$|\hat{f}|$



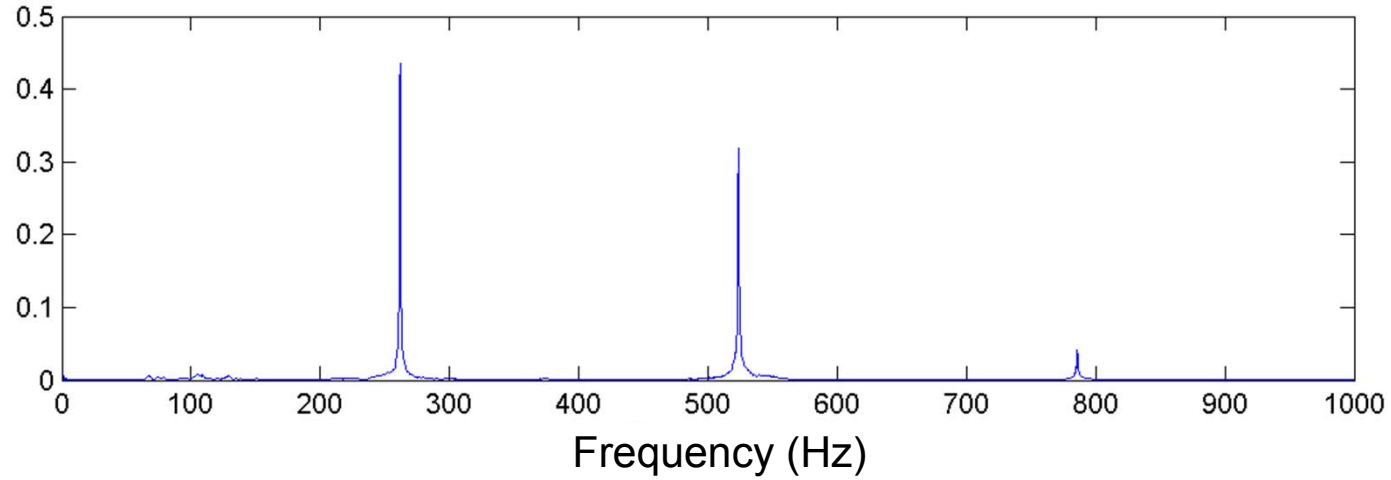
Fourier Transform

Example: C4 played by piano

f



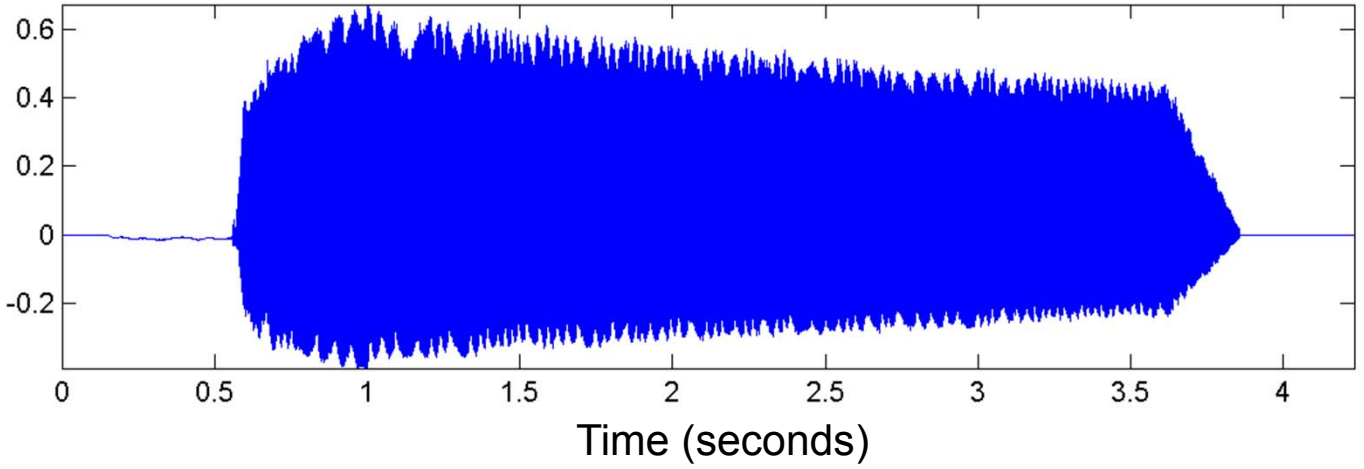
$|\hat{f}|$



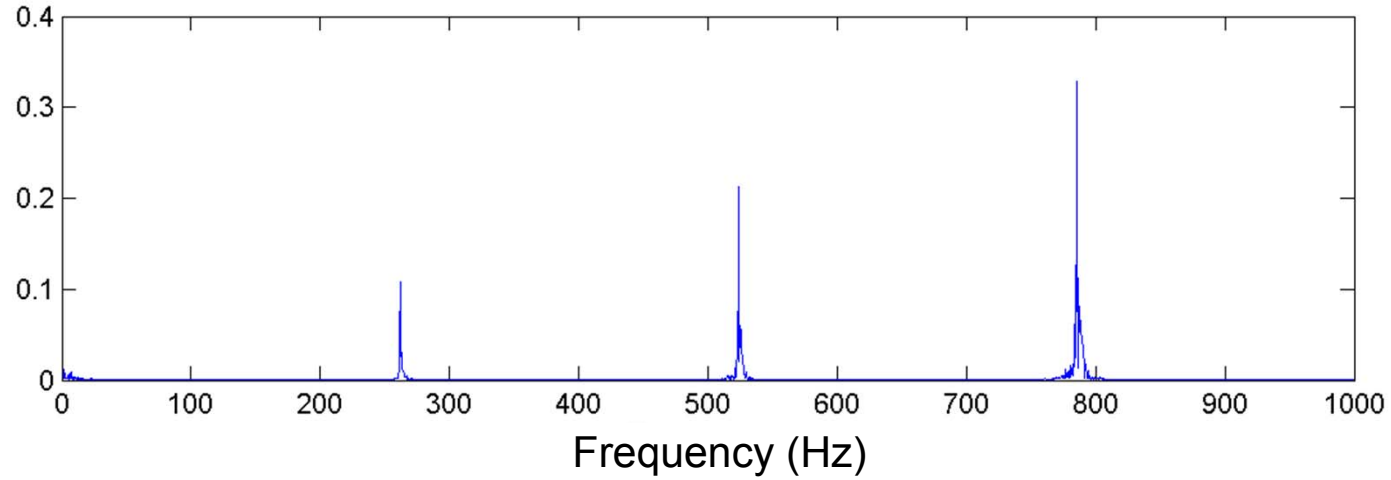
Fourier Transform

Example: C4 played by trumpet

f



$|\hat{f}|$



Fourier Transform

Signal $f: \mathbb{R} \to \mathbb{R}$

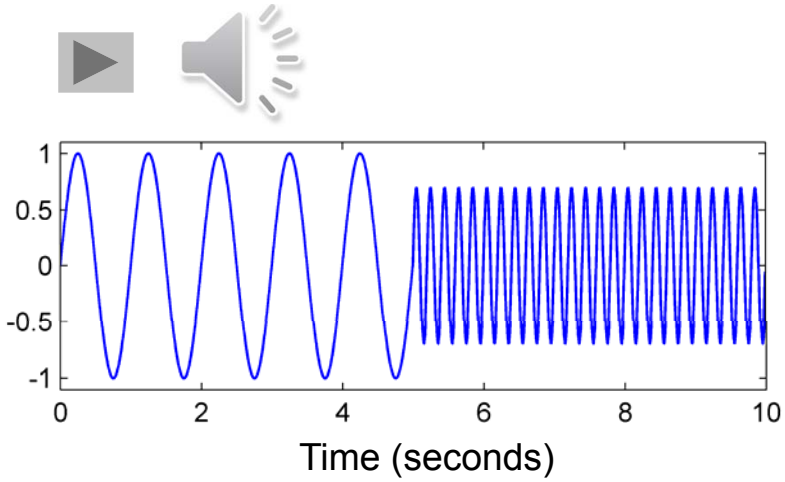
Fourier representation $f(t) = \int_{\omega \in \mathbb{R}} c_\omega \exp(2\pi i \omega t) d\omega$

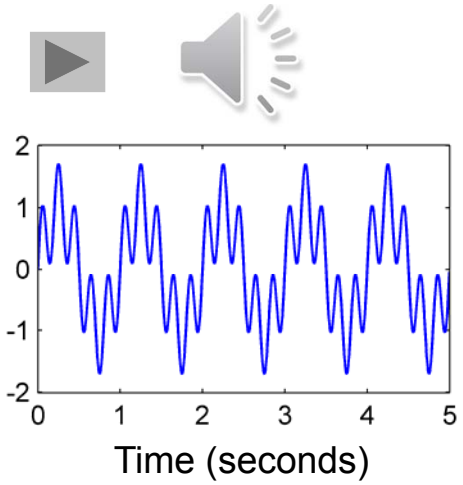
Fourier transform $c_\omega = \hat{f}(\omega) = \int_{t \in \mathbb{R}} f(t) \exp(-2\pi i \omega t) dt$

- Tells which frequencies occur, but does not tell when the frequencies occur.
- Frequency information is averaged over the entire time interval.
- Time information is hidden in the phase

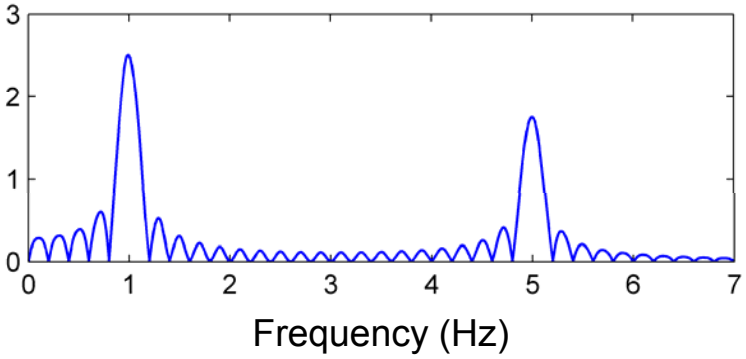
Fourier Transform

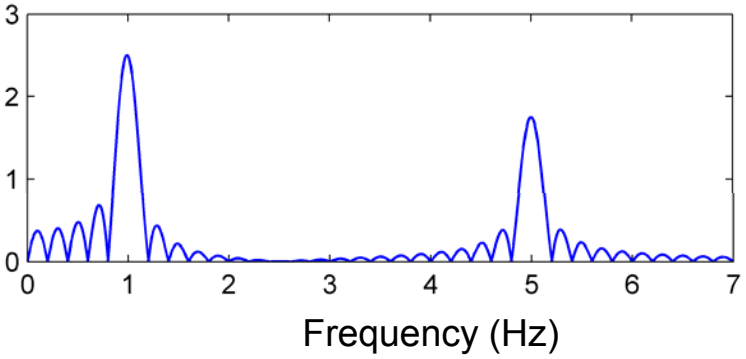
f





$|\hat{f}|$





Short Time Fourier Transform

Idea (Dennis Gabor, 1946):

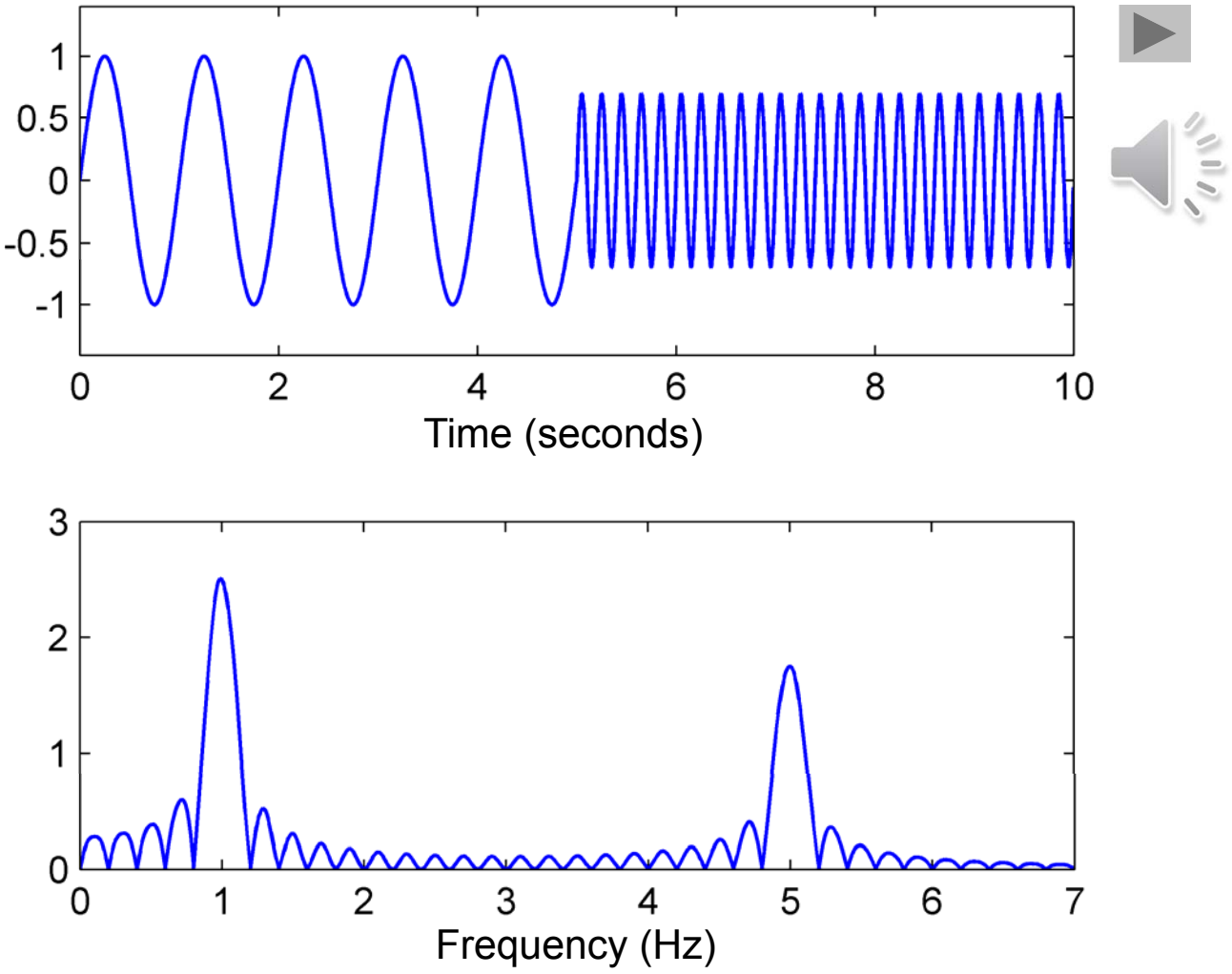
- Consider only a small section of the signal for the spectral analysis

  → recovery of time information

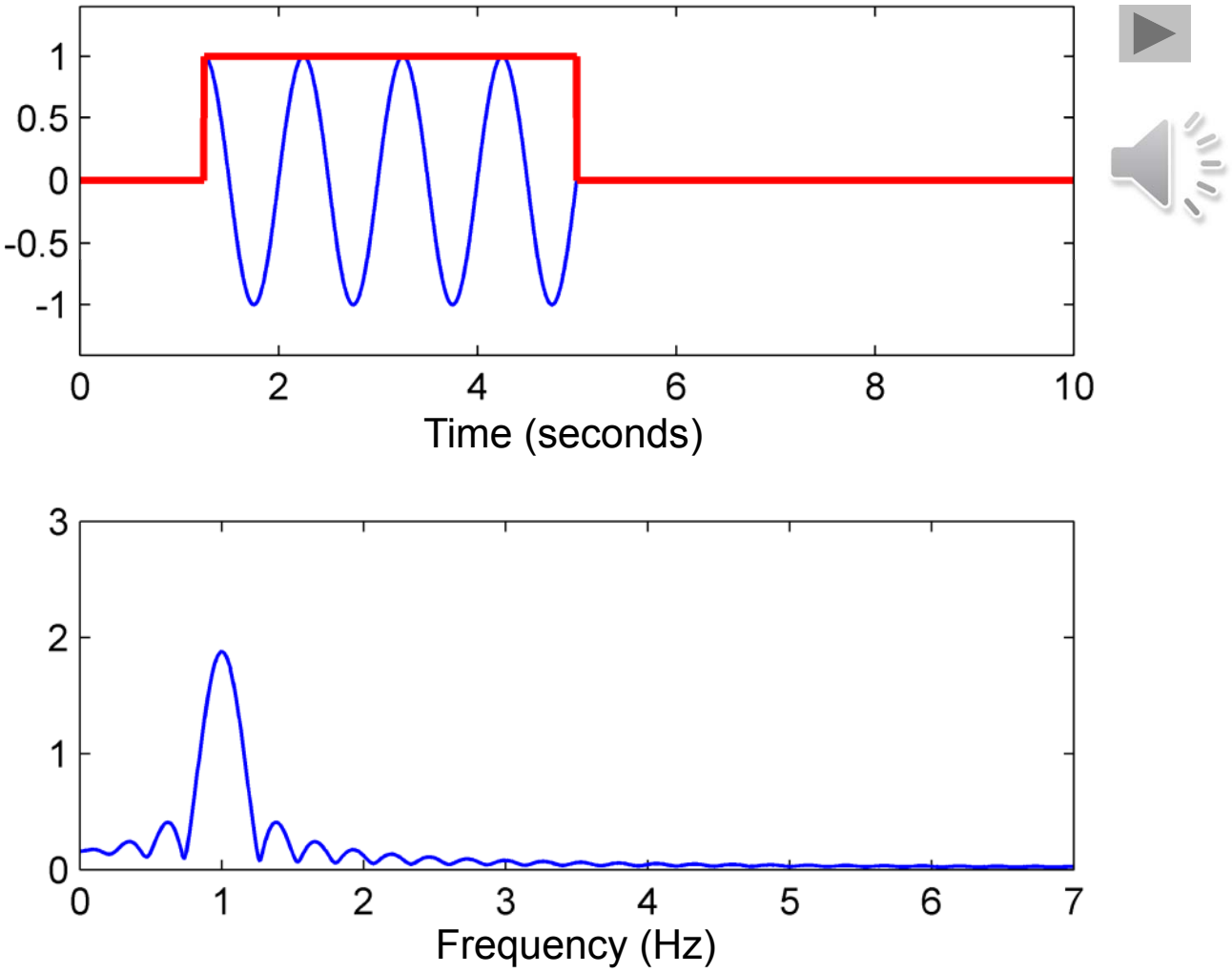
- Short Time Fourier Transform (STFT)

- Section is determined by pointwise multiplication of the signal with a localizing window function

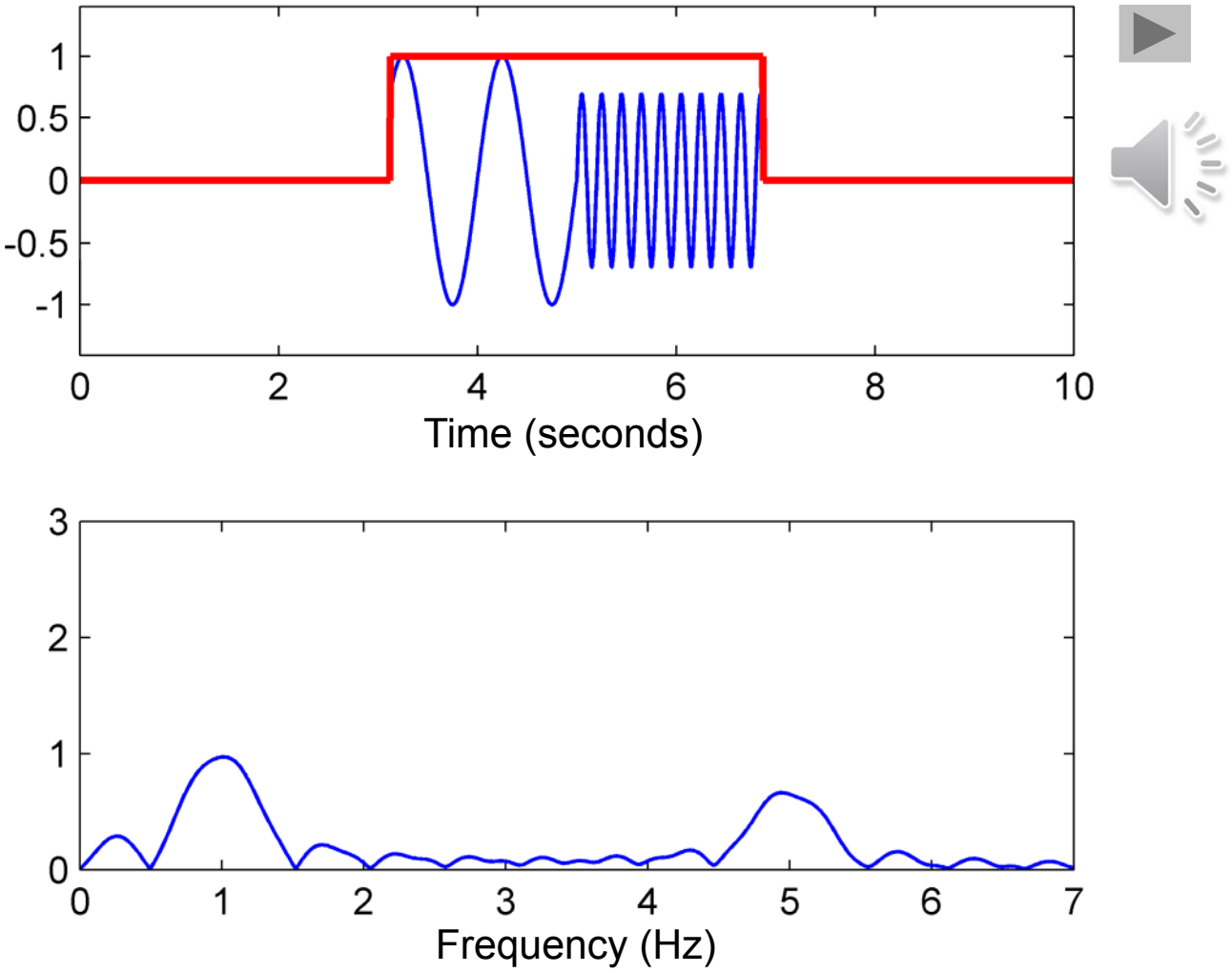
Short Time Fourier Transform



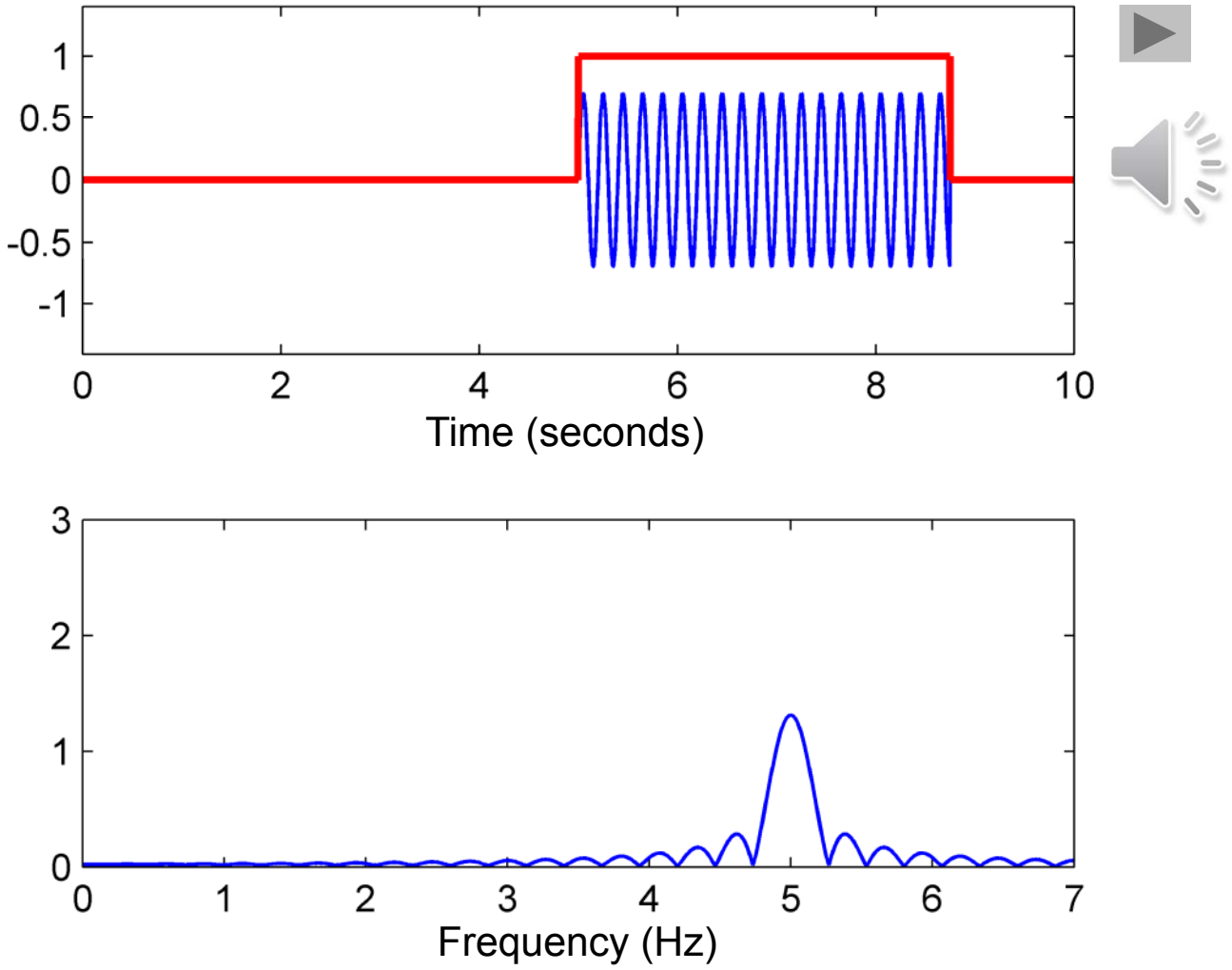
Short Time Fourier Transform



Short Time Fourier Transform

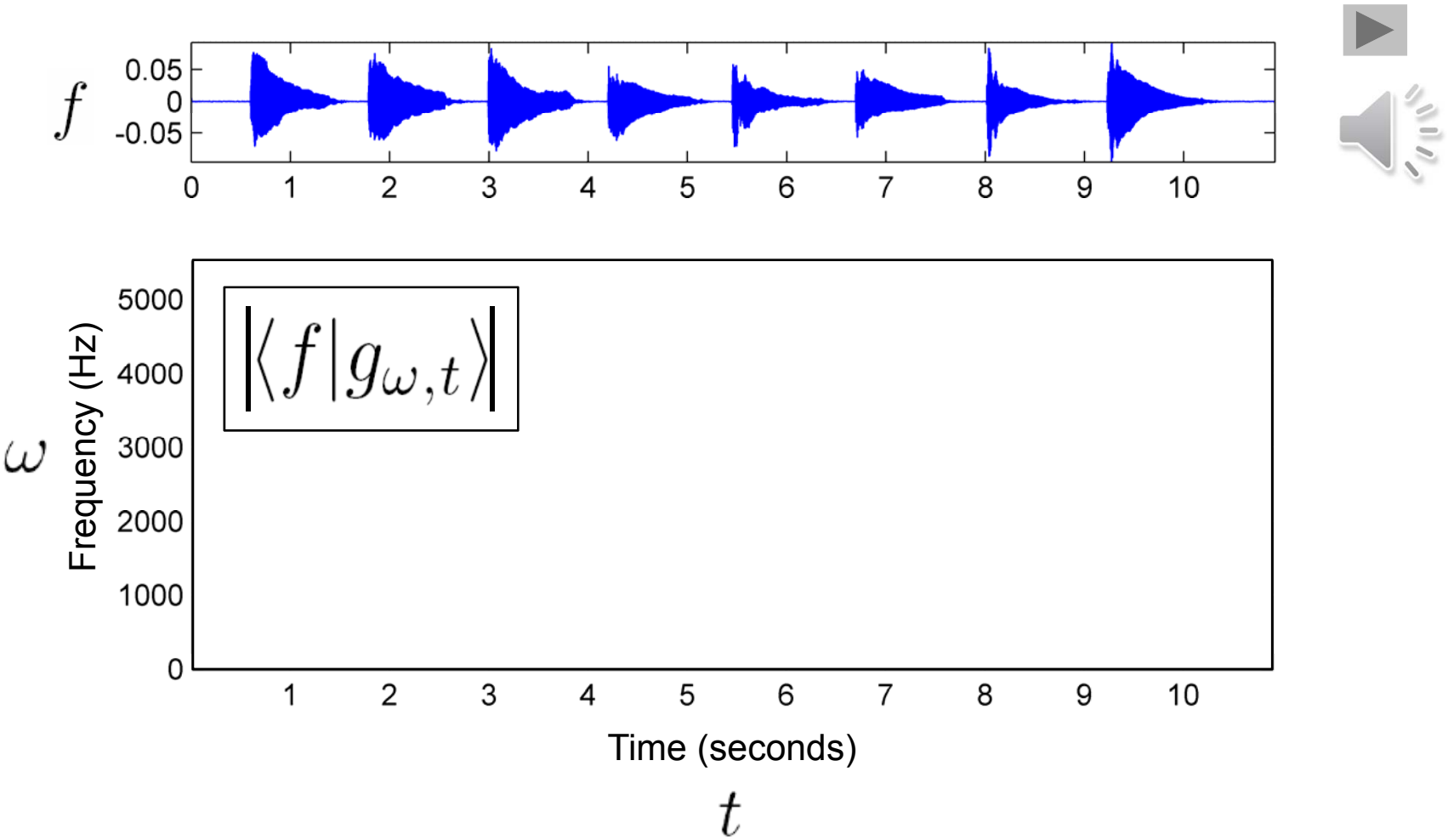


Short Time Fourier Transform



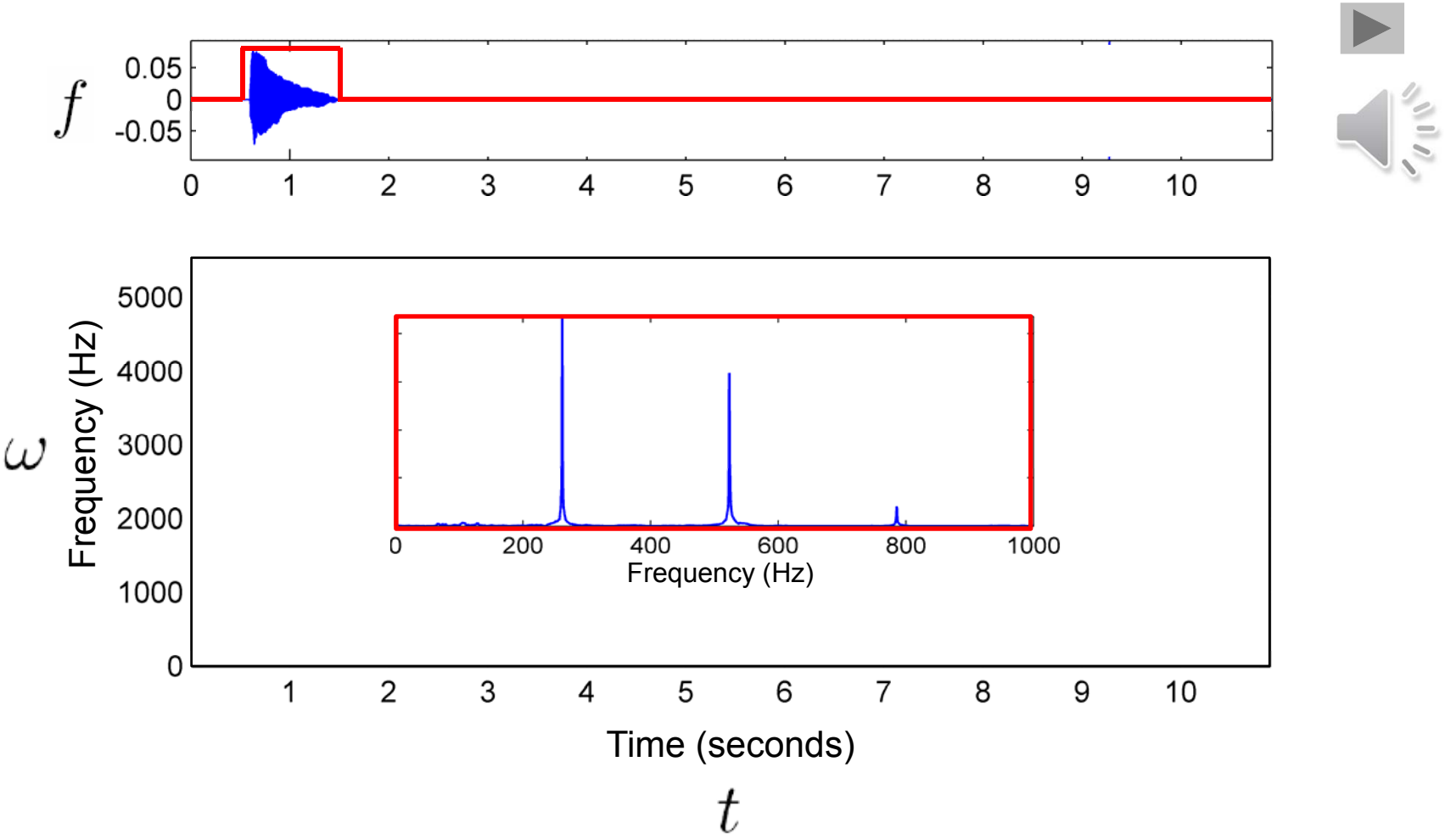
Time-Frequency Representation

Spectrogram



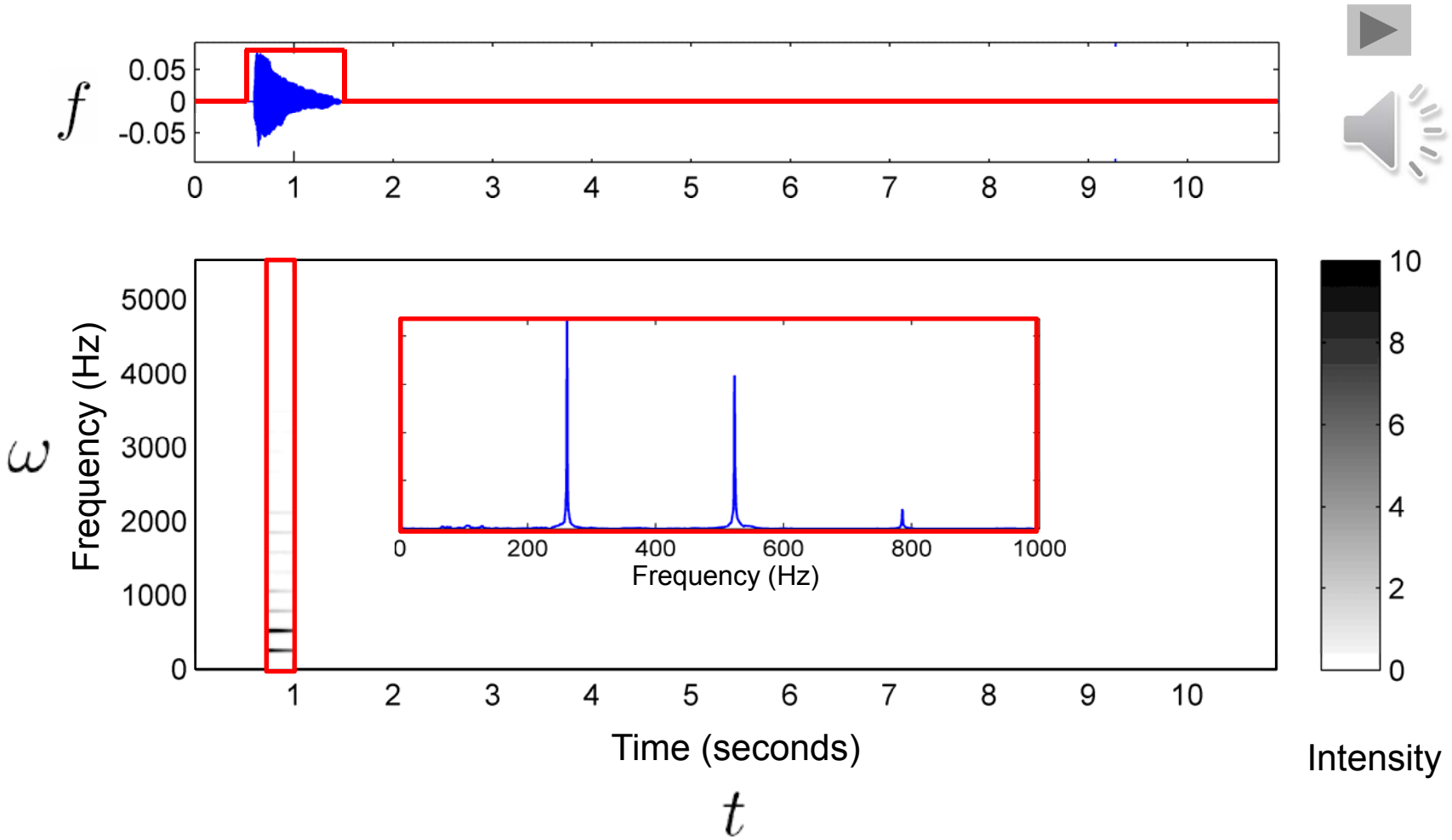
Time-Frequency Representation

Spectrogram



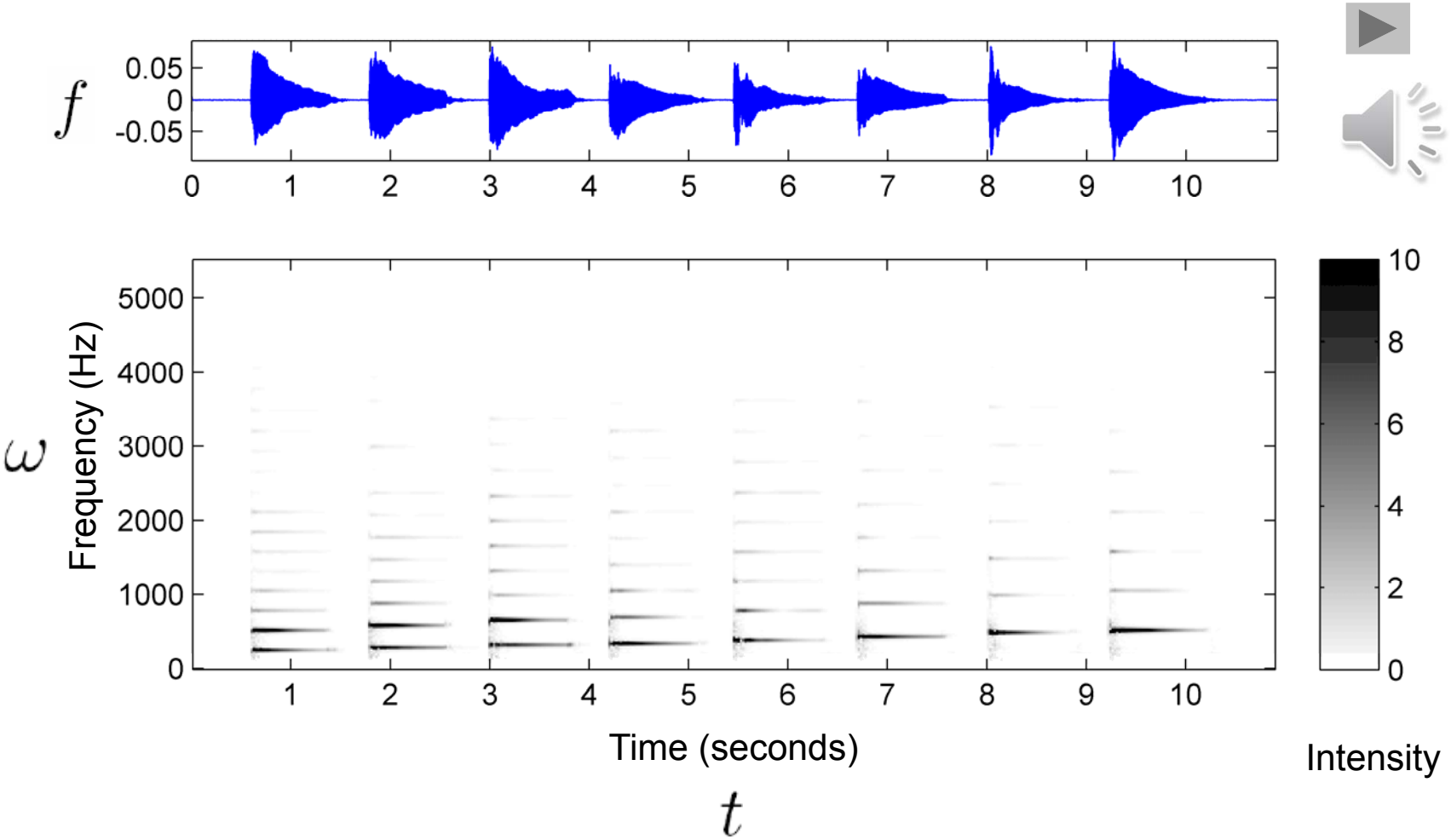
Time-Frequency Representation

Spectrogram



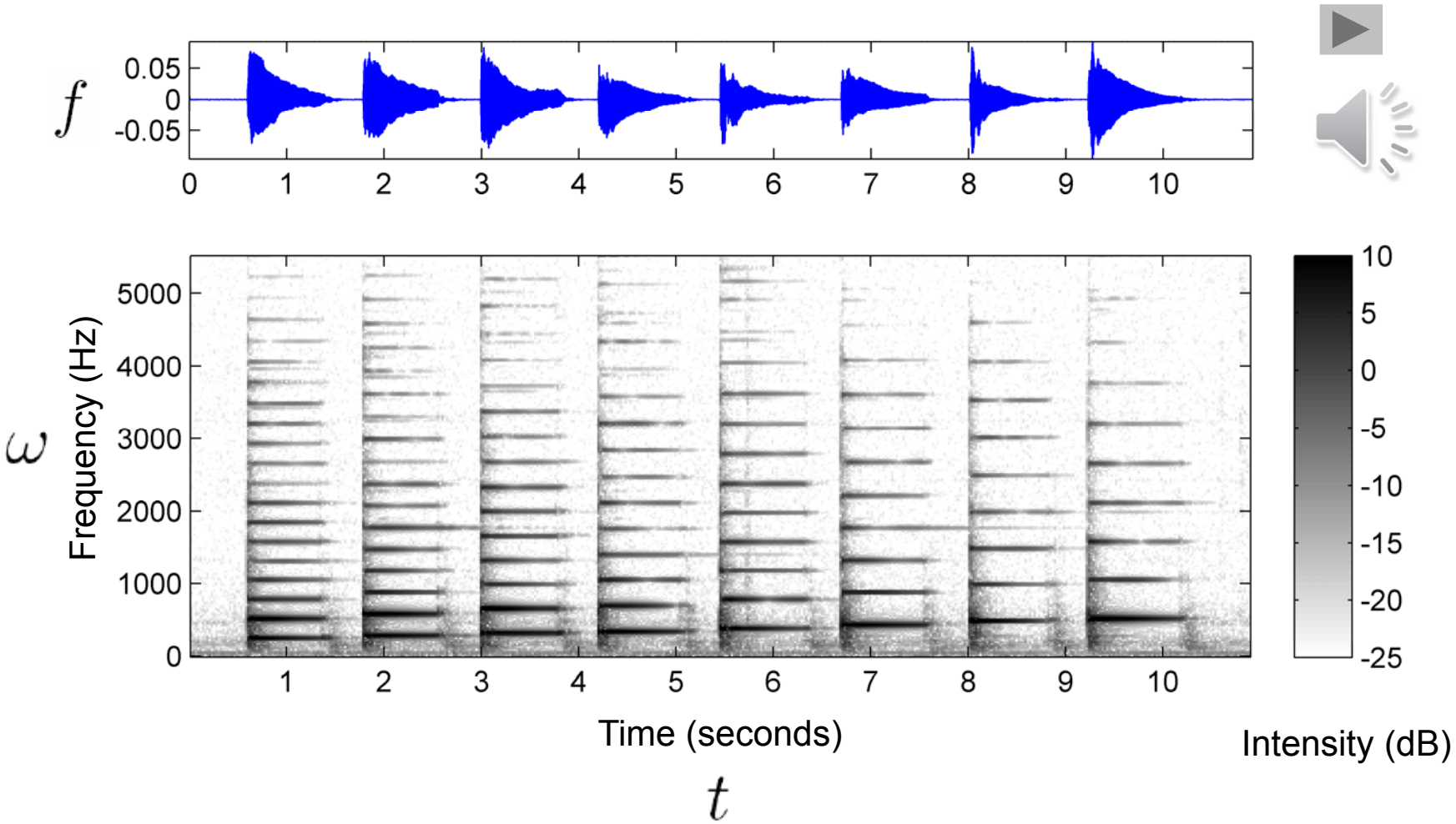
Time-Frequency Representation

Spectrogram



Time-Frequency Representation

Spectrogram



Time-Frequency Representation

Time-Frequency Localization

- Size of window constitutes a trade-off between time resolution and frequency resolution:

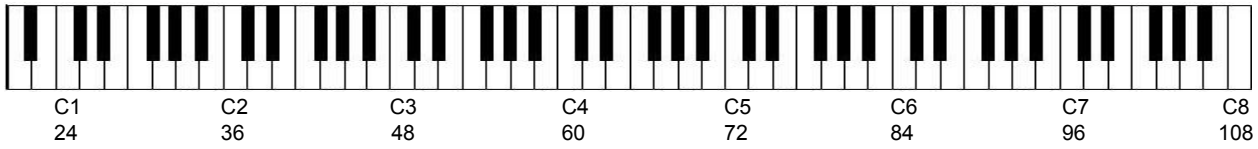
  Large window : poor time resolution
  good frequency resolution

  Small window : good time resolution
  poor frequency resolution

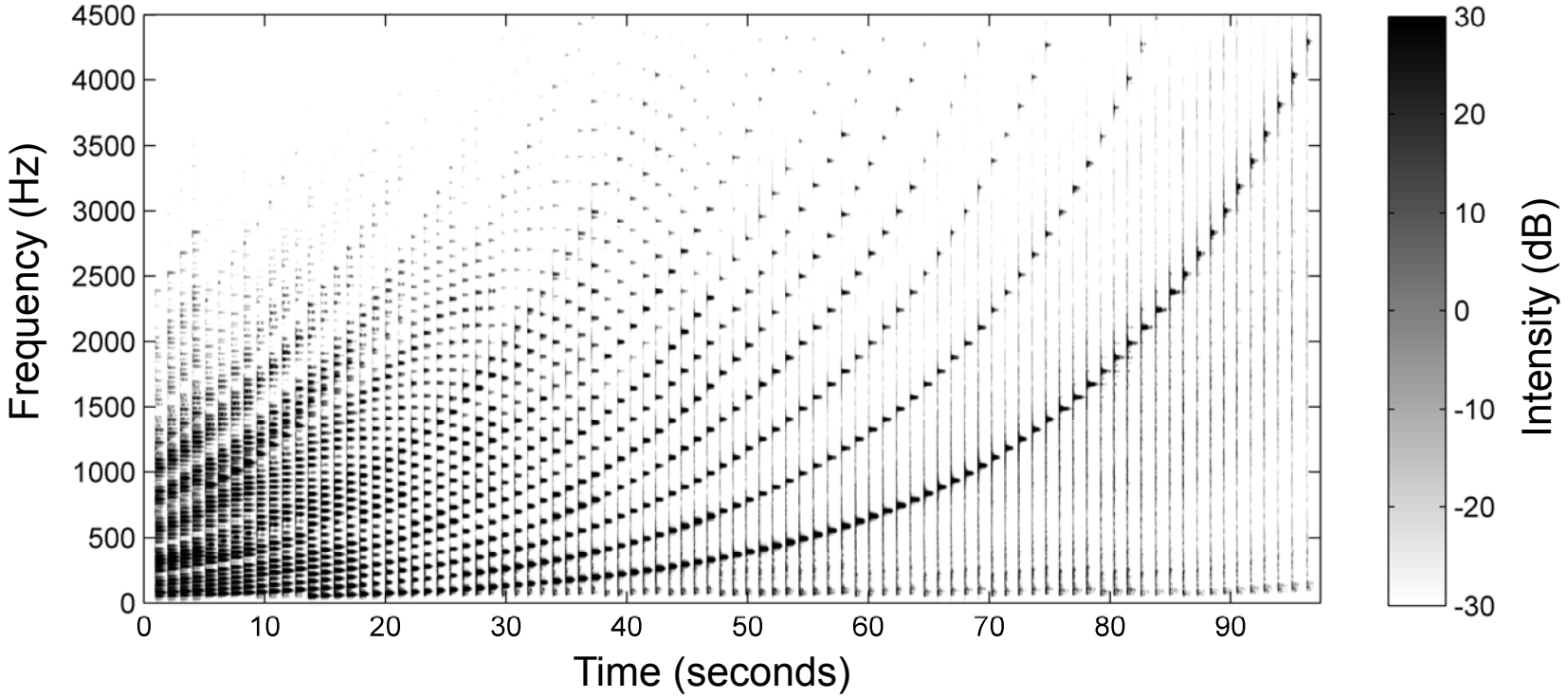
- Heisenberg Uncertainty Principle: there is no window function that localizes in time and frequency with arbitrary precision.

Audio Features

Example: Chromatic scale

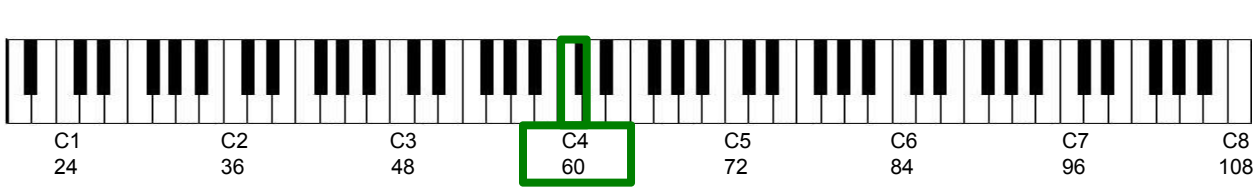


Spectrogram

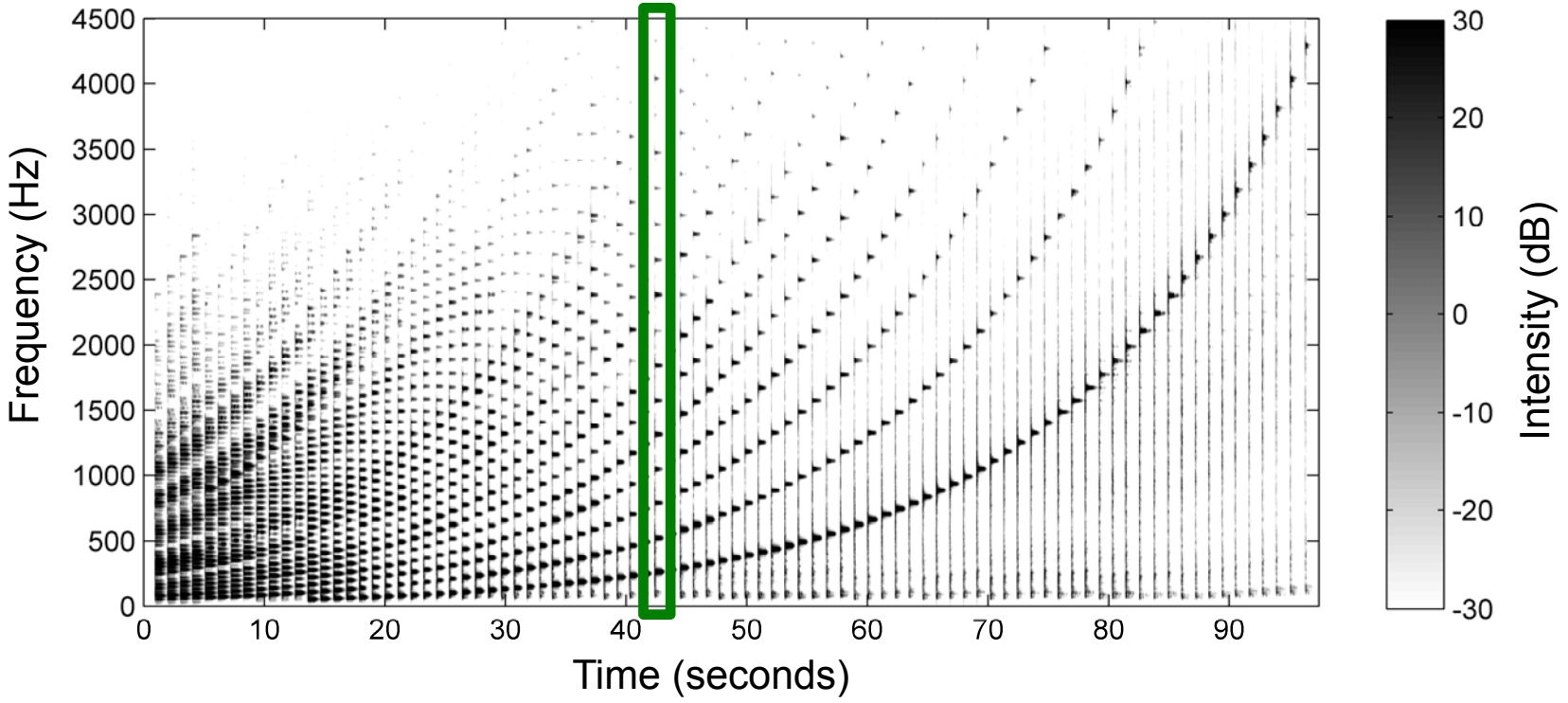


Audio Features

Example: Chromatic scale

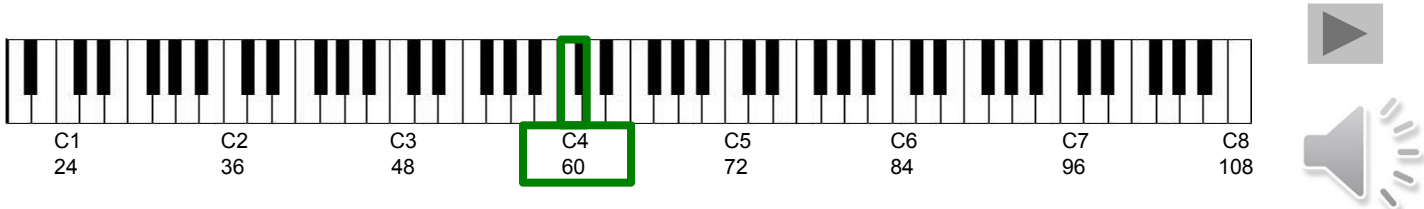


Spectrogram

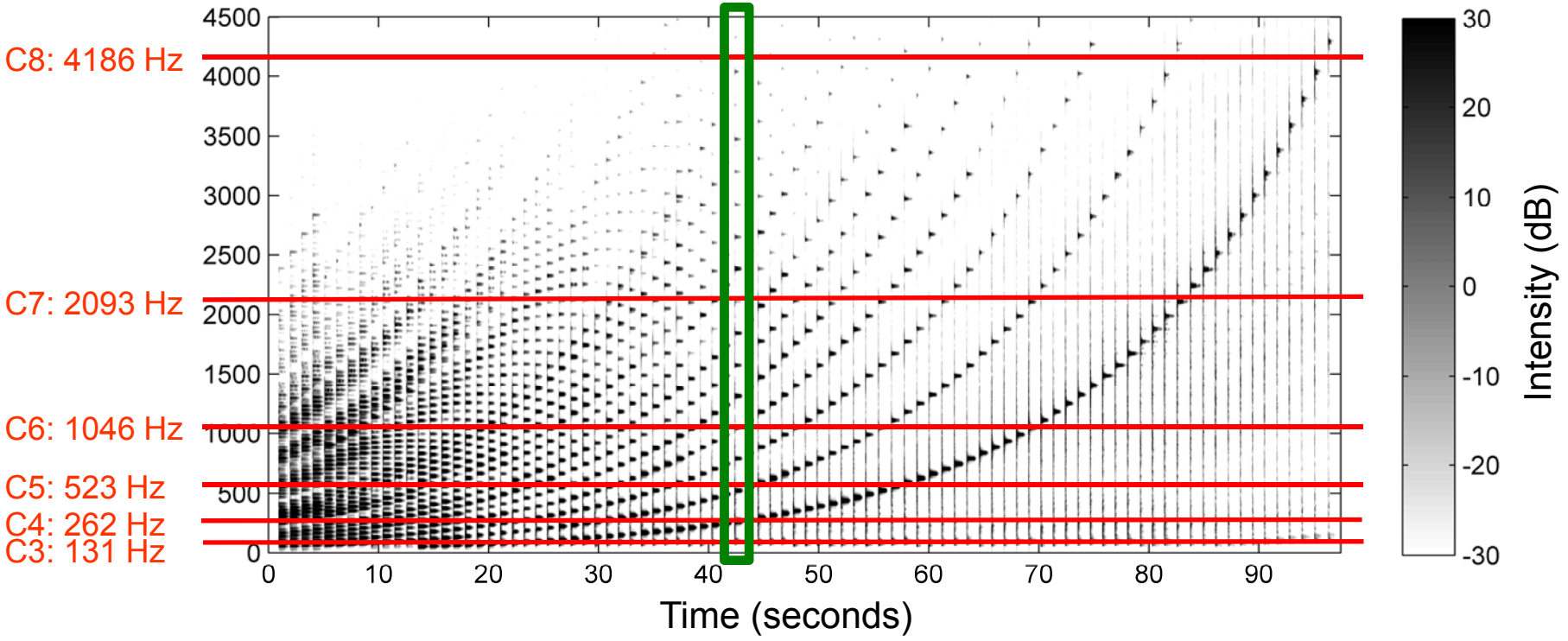


Audio Features

Example: Chromatic scale

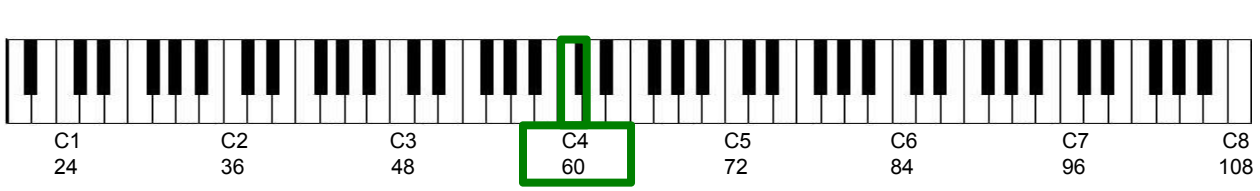


Spectrogram

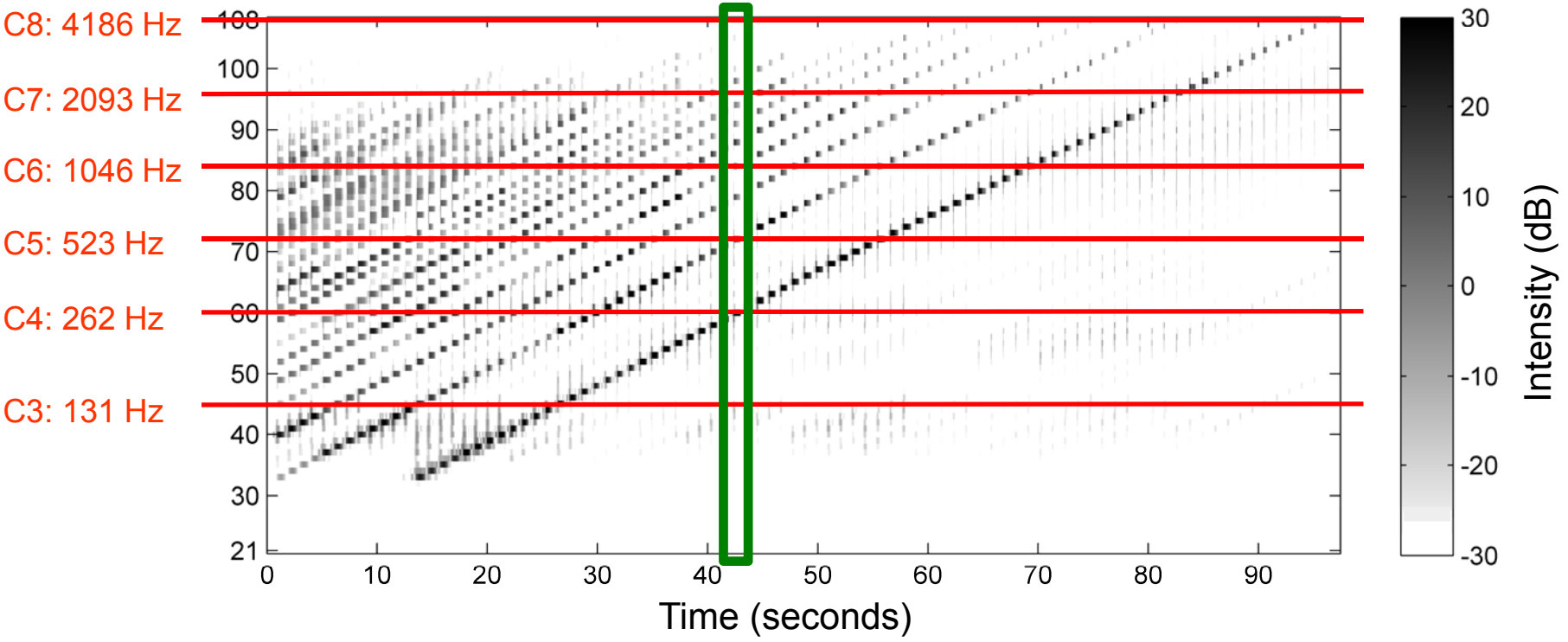


Audio Features

Example: Chromatic scale

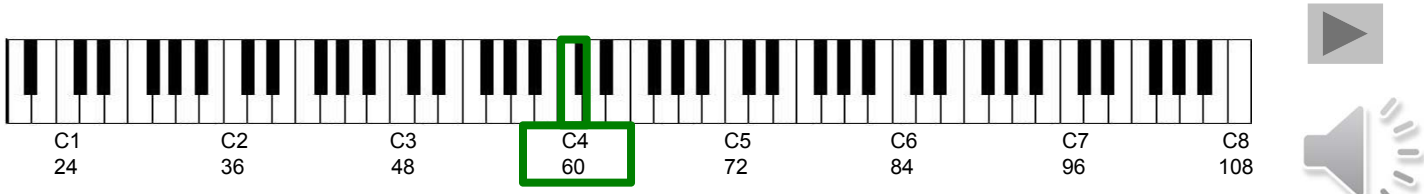


Log-frequency spectrogram

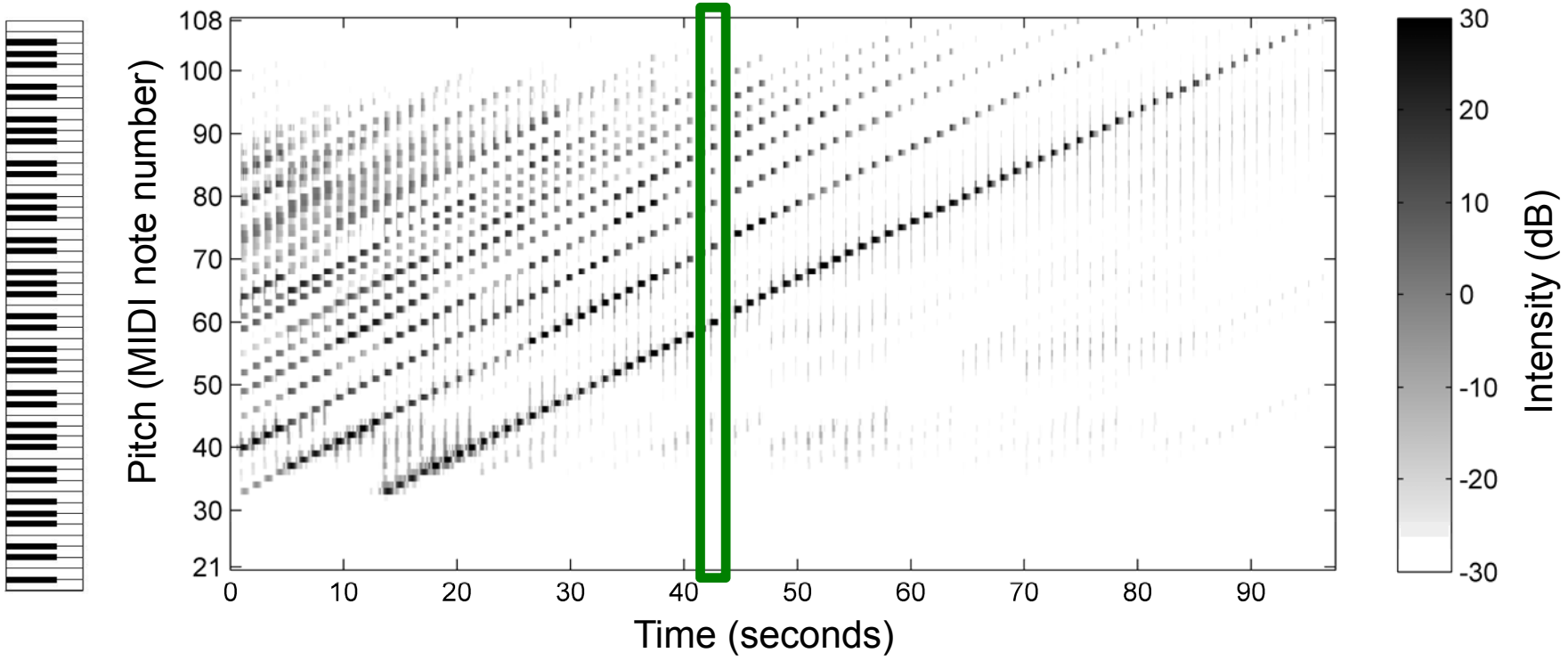


Audio Features

Example: Chromatic scale

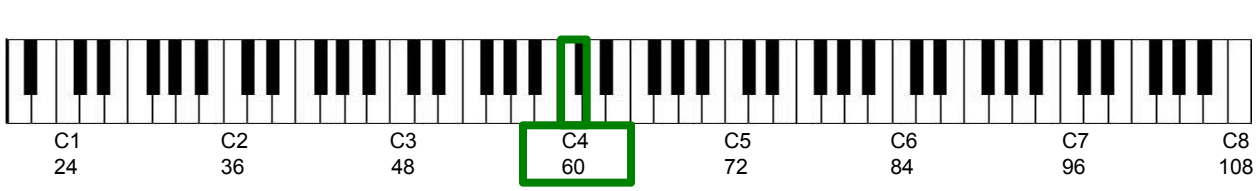


Log-frequency spectrogram

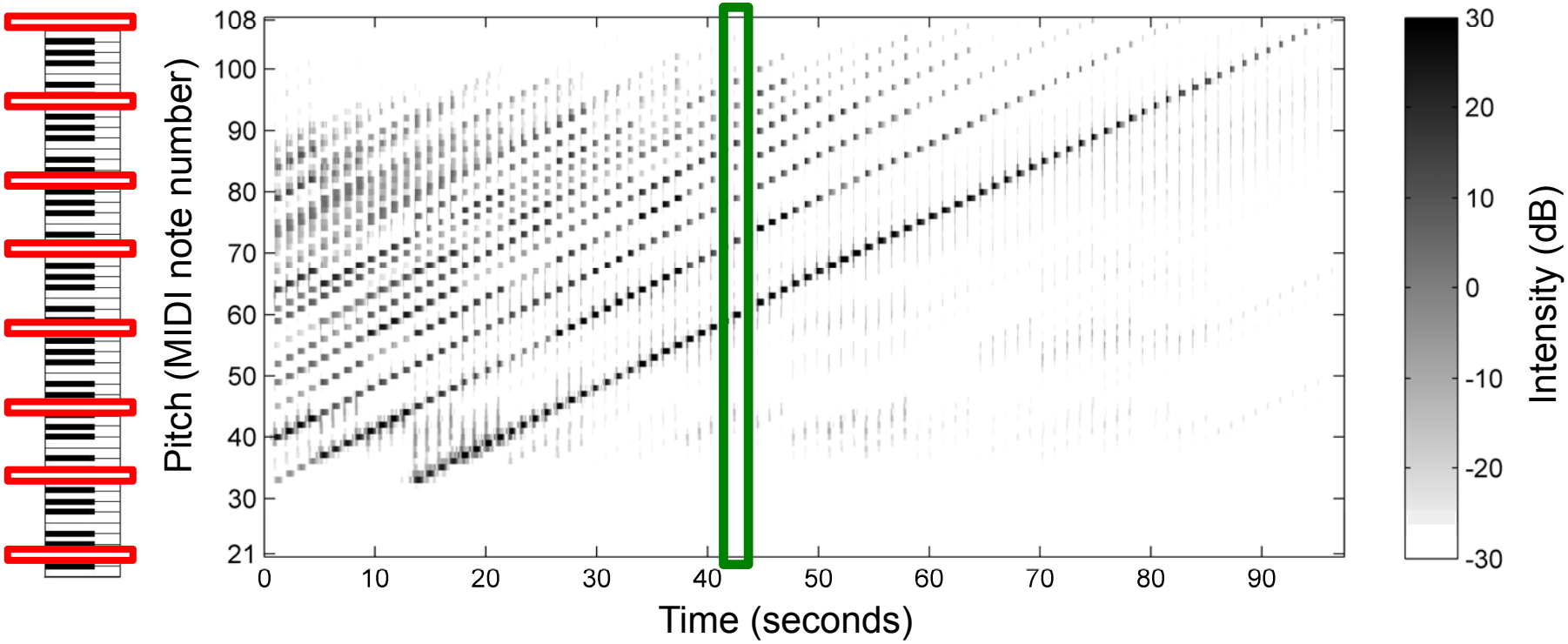


Audio Features

Example: Chromatic scale



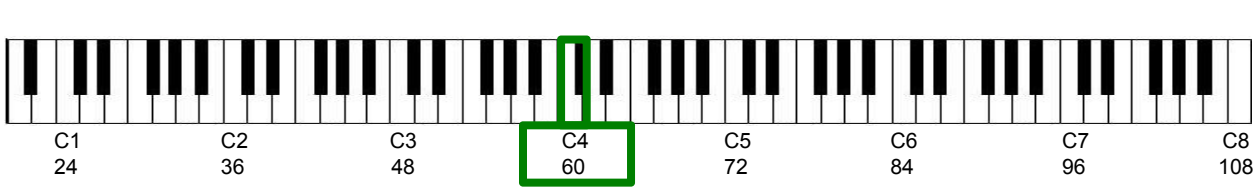
Log-frequency spectrogram



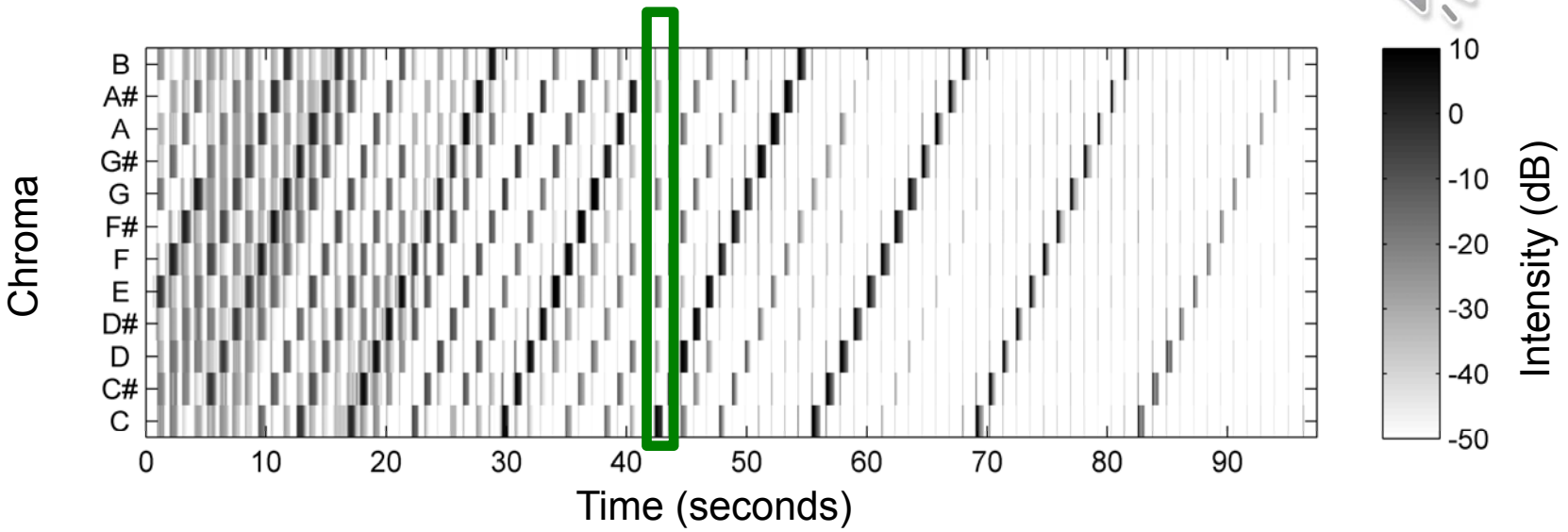
Chroma C

Audio Features

Example: Chromatic scale

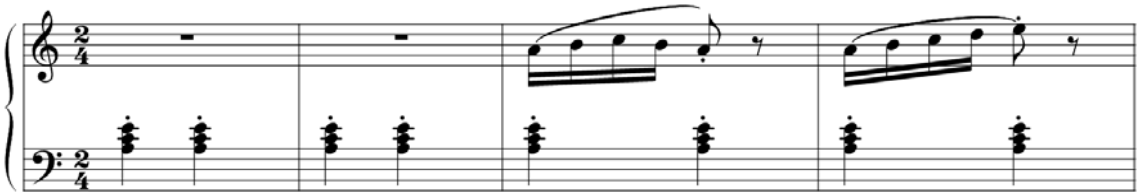


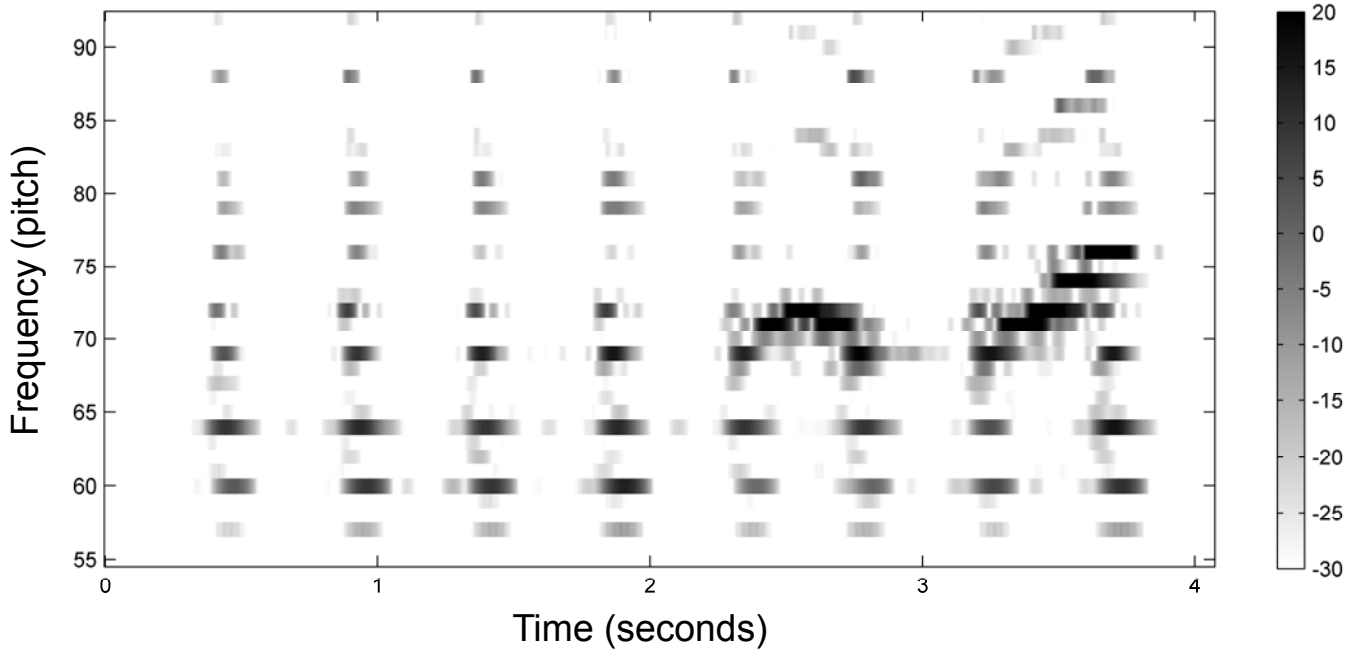
Chromagram



Audio Features

Chroma features





Audio Features

- There are many ways to implement chroma features
- Properties may differ significantly
- Appropriateness depends on respective application
- Chroma Toolbox (MATLAB)
  https://www.audiolabs-erlangen.de/resources/MIR/chromatoolbox
- LibROSA (Python)
  https://librosa.github.io/librosa/
- Feature learning: “Deep Chroma”
  [Korzeniowski/Widmer, ISMIR 2016]