Introduction to **Audio Content Analysis**

module 3.3.1: time-frequency representations — Fourier transform

alexander lerch

Georgia Tech Center for Music Technology
College of Design

introduction
overview

corresponding textbook section

section 3.3.1
appendix B

- **lecture content**
  - FT of continuous signals
  - FT properties
  - FT of sampled signals
  - Short Time FT (STFT)
  - DFT

- **learning objectives**
  - name and explain definition and properties of the FT

introduction
overview

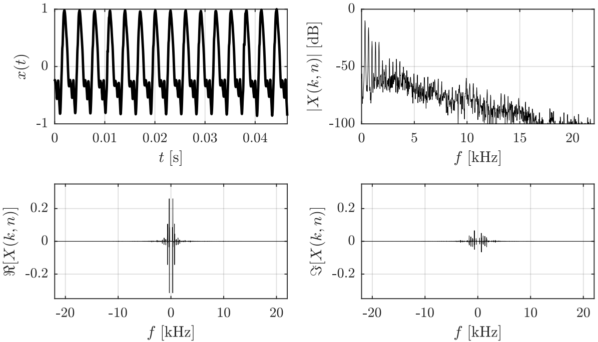
corresponding textbook section

section 3.3.1
appendix B

- **lecture content**
  - FT of continuous signals
  - FT properties
  - FT of sampled signals
  - Short Time FT (STFT)
  - DFT

- **learning objectives**
  - name and explain definition and properties of the FT

fourier transform
introduction



top	time domain signal	magnitude spectrum in dB
bottom	real spectrum	imaginary spectrum

matlab source: plotFourierTransform.m

fourier transform

definition (continuous)

$$X(\mathrm{j}\omega) = \mathfrak{F}[x(t)] = \int\limits_{-\infty}^{\infty} x(t) e^{-\mathrm{j}\omega t}\, dt$$

sidenote: Fourier series coefficients

$$a_k = \frac{1}{T_0} \int\limits_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}\omega_0 k t}\, dt$$

- $T_0 \rightarrow \infty$ to allow the analysis of aperiodic functions

$\Rightarrow$ $k\omega_0 \rightarrow \omega$

fourier transform
definition (continuous)

$$X(\mathrm{j}\omega) = \mathfrak{F}[x(t)] = \int\limits_{-\infty}^{\infty} x(t) e^{-\mathrm{j}\omega t}\, dt$$

sidenote: Fourier series coefficients

$$a_k = \frac{1}{T_0} \int\limits_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}\omega_0 k t}\, dt$$

- $T_0 \to \infty$ to allow the analysis of aperiodic functions

$\Rightarrow$ $k\omega_0 \to \omega$

fourier transform

representations

$$
\begin{aligned}
X(\mathrm{j}\omega) &= \Re[X(\mathrm{j}\omega)] + \Im[X(\mathrm{j}\omega)] \\
&= \underbrace{|X(\mathrm{j}\omega)|}_{\text{magnitude}} \cdot \underbrace{\Phi_{\mathrm{X}}(\omega)}_{\text{phase}}
\end{aligned}
$$

$$
\begin{aligned}
|X(\mathrm{j}\omega)| &= \sqrt{\Re[X(\mathrm{j}\omega)]^2 + \Im[X(\mathrm{j}\omega)]^2} \\
\Phi_{\mathrm{X}}(\omega) &= \operatorname{atan2}\left(\frac{\Im[X(\mathrm{j}\omega)]}{\Re[X(\mathrm{j}\omega)]}\right)
\end{aligned}
$$

- complex spectrum either represented as magnitude & phase or as real & imaginary
- magnitude spectrum has no negative values

fourier transform

representations

$$
\begin{aligned}
X(\mathrm{j}\omega) &= \Re[X(\mathrm{j}\omega)] + \Im[X(\mathrm{j}\omega)] \\
&= \underbrace{|X(\mathrm{j}\omega)|}_{\text{magnitude}} \cdot \underbrace{\Phi_{\mathrm{X}}(\omega)}_{\text{phase}}
\end{aligned}
$$

$$
\begin{aligned}
|X(\mathrm{j}\omega)| &= \sqrt{\Re[X(\mathrm{j}\omega)]^2 + \Im[X(\mathrm{j}\omega)]^2} \\
\Phi_{\mathrm{X}}(\omega) &= \operatorname{atan2}\left(\frac{\Im[X(\mathrm{j}\omega)]}{\Re[X(\mathrm{j}\omega)]}\right)
\end{aligned}
$$

note:

- complex spectrum either represented as magnitude & phase or as real & imaginary
- magnitude spectrum has no negative values

fourier transform

property 1: invertibility

$$\begin{aligned} x(t) &= \mathfrak{F}^{-1}[X(\mathrm{j}\omega)] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\mathrm{j}\omega)e^{\mathrm{j}\omega t}\,d\omega \end{aligned}$$

- time domain signal can be **perfectly reconstructed** — no information loss
- FT and IFT are very similar, largely equivalent

fourier transform

property 1: invertibility

Georgia Tech | Center for Music Technology College of Design

$$
\begin{aligned}
x(t) &= \mathfrak{F}^{-1}[X(\mathrm{j}\omega)] \\
&= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} X(\mathrm{j}\omega)e^{\mathrm{j}\omega t}\,d\omega
\end{aligned}
$$

note:

- time domain signal can be **perfectly reconstructed** — no information loss
- FT and IFT are very similar, largely equivalent

fourier transform

property 2: superposition

$$
\begin{aligned}
y(t) &= c_1 \cdot x_1(t) + c_2 \cdot x_2(t) \\
\mapsto & \\
Y(\mathrm{j}\omega) &= c_1 \cdot X_1(\mathrm{j}\omega) + c_2 \cdot X_2(\mathrm{j}\omega)
\end{aligned}
$$

- FT is a *linear* transform

fourier transform

property 2: superposition

$$\begin{aligned} y(t) &= c_1 \cdot x_1(t) + c_2 \cdot x_2(t) \\ \mapsto & \\ Y(\mathrm{j}\omega) &= c_1 \cdot X_1(\mathrm{j}\omega) + c_2 \cdot X_2(\mathrm{j}\omega) \end{aligned}$$

note:

- FT is a *linear* transform

fourier transform

property 3: convolution and multiplication

$$
\begin{aligned}
y(t) &= \int_{-\infty}^{\infty} h(\tau) \cdot x(t-\tau)\, d\tau \\
\mapsto & \\
Y(\mathrm{j}\omega) &= H(\mathrm{j}\omega) \cdot X(\mathrm{j}\omega)
\end{aligned}
$$

- convolution in time domain means multiplication in frequency domain
- convolution in frequency domain means multiplication in time domain

fourier transform

property 3: convolution and multiplication

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau) \cdot x(t-\tau)\, d\tau \\ \mapsto & \\ Y(\mathrm{j}\omega) &= H(\mathrm{j}\omega) \cdot X(\mathrm{j}\omega) \end{aligned}$$

note:

- convolution in time domain means multiplication in frequency domain
- convolution in frequency domain means multiplication in time domain

fourier transform

property 4: Parseval's theorem

$$\int\limits_{-\infty}^{\infty} x^2(t)\, dt = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \left|X(\mathrm{j}\omega)\right|^2 \, d\omega$$

- energy of the signal is preserved when switching between time and frequency domains

fourier transform

property 4: Parseval's theorem

$$\int\limits_{-\infty}^{\infty} x^2(t)\,dt = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \left|X(\mathrm{j}\omega)\right|^2\,d\omega$$

note:

- energy of the signal is preserved when switching between time and frequency domains

fourier transform

property 5: time & frequency shift

- **time shift**

$$x(t - t_0) \mapsto X(\mathrm{j}\omega)e^{-\mathrm{j}\omega t_0}$$

- **frequency shift**

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} X(\mathrm{j}\omega - \omega_0)e^{\mathrm{j}\omega t}\, d\omega = e^{\mathrm{j}\omega_0 t} \cdot x(t)$$

- time shift results in phase shift
- frequency shift results in modulation of time signal

fourier transform

property 5: time & frequency shift

- **time shift**

$$x(t - t_0) \mapsto X(\mathrm{j}\omega)e^{-\mathrm{j}\omega t_0}$$

- **frequency shift**

$$\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} X(\mathrm{j}\omega - \omega_0)e^{\mathrm{j}\omega t}\, d\omega = e^{\mathrm{j}\omega_0 t} \cdot x(t)$$

note:

- time shift results in phase shift
- frequency shift results in modulation of time signal

fourier transform

property 6: symmetry

$$
\begin{aligned}
|X(\mathrm{j}\omega)| &= |X(-\mathrm{j}\omega)| \\
\Phi_{\mathrm{X}}(\omega) &= -\Phi_{\mathrm{X}}(-\omega)
\end{aligned}
$$

- spectrum of (real) signal is conjugate complex
  - magnitude spectrum is symmetric to ordinate
  - phase spectrum is symmetric to origin
- even signals have no imaginary spectrum
- odd signals have no real spectrum

fourier transform

property 6: symmetry

$$
\begin{aligned}
|X(\mathrm{j}\omega)| &= |X(-\mathrm{j}\omega)| \\
\Phi_\mathrm{X}(\omega) &= -\Phi_\mathrm{X}(-\omega)
\end{aligned}
$$

note:

- spectrum of (real) signal is conjugate complex
  - magnitude spectrum is symmetric to ordinate
  - phase spectrum is symmetric to origin
- even signals have no imaginary spectrum
- odd signals have no real spectrum

fourier transform

property 7: time & frequency scaling

$$
\begin{aligned}
y(t) &= x(c \cdot t) \\
\mapsto \\
Y(\mathrm{j}\omega) &= \frac{1}{|c|} X\left(j\frac{\omega}{c}\right)
\end{aligned}
$$

- scaling of abscissa in one domain leads to inverse scaling in the other domain

fourier transform

property 7: time & frequency scaling

$$
\begin{aligned}
y(t) &= x(c \cdot t) \\
\mapsto & \\
Y(\mathrm{j}\omega) &= \frac{1}{|c|} X\left(j \frac{\omega}{c}\right)
\end{aligned}
$$

note:

- scaling of abscissa in one domain leads to inverse scaling in the other domain

fourier transform

sampled time signals 1/2

- sampled time signal can be modeled as multiplication of original signal with delta pulse $\delta_\mathrm{T}(t)$
- multiplication in time domain $\mapsto$ convolution in frequency domain

$$\begin{aligned} \mathfrak{F}[x(i)] &= \mathfrak{F}[x(t) \cdot \delta_\mathrm{T}(t)] \\ &= \mathfrak{F}[x(t)] * \mathfrak{F}[\delta_\mathrm{T}(t)] \\ &= X(\mathrm{j}\omega) * \Delta_\mathrm{T}(\mathrm{j}\omega) \end{aligned}$$

note

- even if time domain signal is discrete, its Fourier transform is *still continuous*
- spectrum is *repeated periodically*

fourier transform

sampled time signals 1/2

- sampled time signal can be modeled as multiplication of original signal with delta pulse $\delta_\mathrm{T}(t)$
- multiplication in time domain $\mapsto$ convolution in frequency domain

$$\begin{aligned}\mathfrak{F}[x(i)] &= \mathfrak{F}[x(t)\cdot\delta_\mathrm{T}(t)] \\ &= \mathfrak{F}[x(t)] * \mathfrak{F}[\delta_\mathrm{T}(t)] \\ &= X(\mathrm{j}\omega) * \Delta_\mathrm{T}(\mathrm{j}\omega)\end{aligned}$$

note

- even if time domain signal is discrete, its Fourier transform is *still continuous*
- spectrum is *repeated periodically*

fourier transform

sampled time signals 1/2

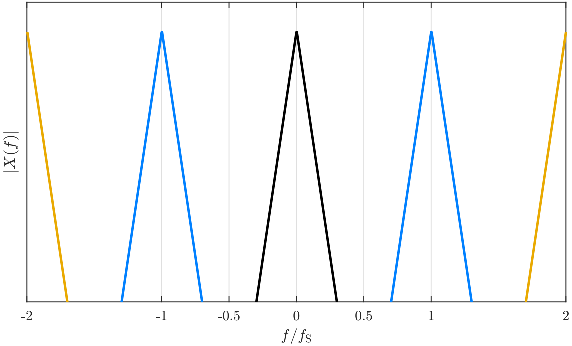
- sampled time signal can be modeled as multiplication of original signal with delta pulse $\delta_\mathrm{T}(t)$
- multiplication in time domain $\mapsto$ convolution in frequency domain

$$
\begin{aligned}
\mathfrak{F}[x(i)] &= \mathfrak{F}[x(t)\cdot\delta_\mathrm{T}(t)] \\
&= \mathfrak{F}[x(t)] * \mathfrak{F}[\delta_\mathrm{T}(t)] \\
&= X(\mathrm{j}\omega) * \Delta_\mathrm{T}(\mathrm{j}\omega)
\end{aligned}
$$

note

- even if time domain signal is discrete, its Fourier transform is *still continuous*
- spectrum is *repeated periodically*

fourier transform
sampled time signals 2/2



matlab source: matlab/animateSampling.m

fourier transform

STFT 1/2

short time Fourier transform (STFT):
Fourier transform of a short time segment

- **reasons**:
  - remember block based processing
  - segment can be seen as quasi-periodic or stationary
- **implementation**:
  - pretend signal is 0 outside of the segment

  $\Rightarrow$ multiplication of signal and *window function*

fourier transform

STFT 1/2

short time Fourier transform (STFT):
Fourier transform of a short time segment

- **reasons**:
  - remember block based processing
  - segment can be seen as quasi-periodic or stationary
- **implementation**:
  - pretend signal is 0 outside of the segment

  $\Rightarrow$ multiplication of signal and *window function*

fourier transform

STFT 1/2

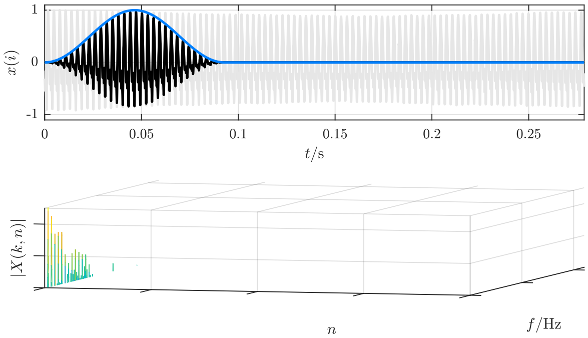
short time Fourier transform (STFT):
Fourier transform of a short time segment

- **reasons**:
    - remember block based processing
    - segment can be seen as quasi-periodic or stationary
- **implementation**:
    - pretend signal is 0 outside of the segment

    $\Rightarrow$ multiplication of signal and *window function*

fourier transform

STFT 2/2



matlab source: matlab/animateStft.m

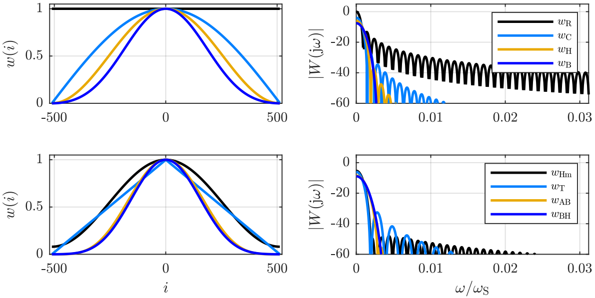
fourier transform

STFT: window functions

- time domain multiplication $\mapsto$ frequency domain convolution
- time domain shape determines frequency domain shape of the window

fourier transform
STFT: window functions

- time domain multiplication $\mapsto$ frequency domain convolution
- time domain shape determines frequency domain shape of the window



matlab source: plotSpectralWindows.m

fourier transform

STFT: window functions

- time domain multiplication $\mapsto$ frequency domain convolution
- time domain shape determines frequency domain shape of the window

spectral leakage characterization

- main lobe width
- side lobe height
- side lobe attenuation

fourier transform
DFT

digital domain: requires discrete frequency values:
$\Rightarrow$ discrete Fourier transform

$$X(k) = \sum_{i=0}^{\mathcal{K}-1} x(i) e^{-\mathrm{j}ki\frac{2\pi}{\mathcal{K}}}$$

with

$$\Delta\Omega = \frac{2\pi}{\mathcal{K}T_\mathrm{S}} = \frac{2\pi f_\mathrm{S}}{\mathcal{K}}$$

2 interpretations:

- sampled continuous Fourier transform
- continuous Fourier transform of periodically extended time domain segment

fourier transform
DFT

digital domain: requires discrete frequency values:
$\Rightarrow$ discrete Fourier transform

$$X(k) = \sum_{i=0}^{\mathcal{K}-1} x(i) e^{-\mathrm{j}ki\frac{2\pi}{\mathcal{K}}}$$

with

$$\Delta\Omega = \frac{2\pi}{\mathcal{K}T_\mathrm{S}} = \frac{2\pi f_\mathrm{S}}{\mathcal{K}}$$

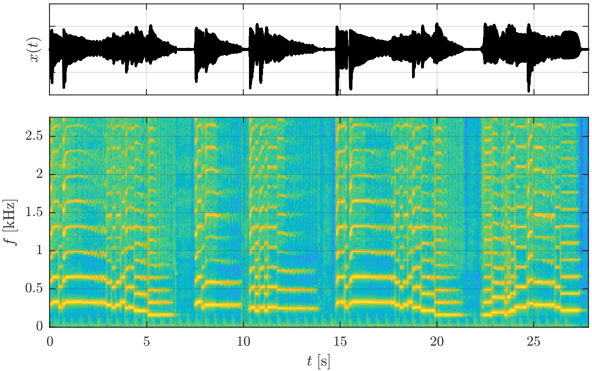
2 interpretations:

- sampled continuous Fourier transform
- continuous Fourier transform of periodically extended time domain segment

fourier transform

spectrogram

- spectrogram allows to visualize temporal changes in the spectrum
- displays the *magnitude spectrum* only



matlab source: plotSpecgram.m

summary
lecture content

- **Fourier Transform (of a real signal)**
  - is conjugate complex
  - often represented as magnitude + phase
  - invertible
  - linear
  - convolution in time domain is multiplication in frequency domain
  - energy preserving
  - time shift result in phase shift, frequency shift results in amplitude modulation
  - symmetric
  - time scaling result in inverse frequency scaling
- **FT of sampled signals**:
  - is periodic with sample rate
- **STFT**
  - window results in spectral leakage (convolution in freq domain)
- **DFT**
  - discrete in both time and freq domain