Fundamentals of digital audio processing

SMC chapt. 1
Outline

1.2 Discrete-time signals and systems
1.4 Spectral analysis of discrete-time signals
1.5 Short-time spectral analysis
1.6 Digital filters
Discrete time signals

chapt. 1.2.1
Digital signal processing

- analog
- sampling
- processing
- reconstruction

Sampling period \( T_s \) \[ \text{sampling frequency } F_s = \frac{1}{T_s} \]
Characterization and Classification of Signals

\[ x(t) : t \in D \rightarrow x(t) \in C \]

- \( D = R, C = R \) \( \rightarrow \) “analog” signal
- \( D = R, C = I \) \( \rightarrow \) “quantized analog” signal
- \( D = I, C = R \) \( \rightarrow \) “sampled” or “discrete-time” signal
- \( D = I, C = I \) \( \rightarrow \) “numerical” or “digital” signal

\( I \) is a countable set \( \{t_0, t_1, t_2, \ldots, t_n, \ldots\} \)

\[ x_d[n] = x_c(nT_s) \]
Characterization and Classification of Signals

- **analog**

\[ I = \{ \ldots, -0.1, 0, 0.1, 0.2, \ldots \}[\text{sec}] \]

- **discrete time**

- **quantized**

- **digital**

Digital signals: time representations

- 8000 samples
- 100 samples
- line with dots

- vertical quantization
  - integer
    e.g. -32768 .. 32767
  - normalized
    e.g. -1 .. (1-Q)
    Q = quantization step

Esempio1.m
Scale formats for digital audio signals
Discrete-Time signals: Sequences

- Discrete-Time signals are represented mathematically as sequences of numbers.
- A sequence of numbers $x$, in which the $n$-th number in the sequence is denoted $x[n]$, is formally written as:

$$x = \{x[n]\}, -\infty < n < \infty \quad n \text{ is integer}$$

$x$ is not defined for non-integer values of $n$
In a practical setting, such sequences can arise from periodic sampling of an analog signal.

In this case, the numeric value of the \( n \)-th number in the sequence is equal to the value of the analog signal \( x_a(t) \) at time \( nT \); i.e.,

\[
x[n] = x_a(nT), \quad -\infty < n < \infty
\]

The quantity \( T \) is called sampling period and its reciprocal \( F_s = 1/T \) is the sampling frequency.
Basic operations

- Sequences are manipulated through various basic operations
  - product:  $y[n] = x_1[n] \times x_2[n]$
  - sum: $y[n] = x_1[n] + x_2[n]$
  - multiplication by a constant $y[n] = a \times x[n]$
  - time shifting or translation $y[n] = x[n - n_0]$
    - for $n_0 \in \mathbb{Z}$
      - $n_0 > 0$ delaying operation
      - $n_0 < 0$ advancing operation
Basic Sequences

- **Unit sample sequence** or impulse

\[
\delta[n] = \begin{cases} 
1, & n = 0, \\
0, & n \neq 0.
\end{cases}
\]

- **Unit step sequence**

\[
u[n] = \begin{cases} 
1, & n \geq 0, \\
0, & n < 0.
\end{cases}
\]

- Impulse as first backward difference of the unit step

\[
\delta[n] = u[n] - u[n - 1].
\]
Sinusoidal and exponential sequences

- **Real sinusoidal sequence**

\[ x[n] = A \cos(\omega_0 n + \phi), \quad -\infty < n < \infty, \]

- **Exponential sequence** if \( \alpha \) is complex

\[ \alpha = |\alpha|e^{j\omega_0} \]

\[ x[n] = A\alpha^n \]

\[ = |A||\alpha|^n \cdot (\cos(\omega_0 n + \phi) + j \sin(\omega_0 n + \phi)). \]
Measures of discrete-time signals

- **Energy**
  \[ E_x = \sum_{n=-\infty}^{\infty} |x[n]|^2. \]

- **Power**
  \[ P_x = \frac{E_x}{N} = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|^2. \]

- **Sound signal**: \( x[n] \) represents acoustic pressure
- **RMS (root mean square) level** \( \sqrt{P_x} \)

- **Sound intensity**: \( I = \) power per unit area of the wavefront
- **Sound pressure level SPL (dB)**
  \[ SPL = 10 \log_{10}(I/I_0) \quad (dB) \]
  \[ SPL_2 - SPL_1 = 10 \log_{10}(p_2^2/p_1^2) = 20 \log_{10}(p_2/p_1) \]
Discrete time systems

Chapt. 1.2.2
Discrete-time systems

- **Generic system**
  \[ y[n] = T\{x\}[n]. \]

- **Example: ideal delay**
  \[ y[n] = T_{n_0}\{x\}[n] = x[n - n_0], \]

- **Example: moving average**
  \[ y[n] = T_{MA}\{x\}[n] = \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} x[n - k]. \]
Classes of discrete-time systems

- **Memory**: “current” value $y[n]$ computed using...
  - with memory: past/future values of $x$
  - memoryless: only current value $x[n]$, e.g. $y[n] = \sin(x[n])$

- **Linearity**: principle of superposition
  - $\mathcal{T}\{a_1x_1 + a_2x_2\}[n] = a_1\mathcal{T}\{x_1\}[n] + a_2\mathcal{T}\{x_2\}[n]$

- **Time-invariance**: time shift of input seq. $\Rightarrow$ time shift of output seq.
  - $\mathcal{T}\{\mathcal{T}_{n_0}\{x\}\}[n] = y[n - n_0] \quad \forall n_0,$

- **Causality**:
  - Current value $y[n_0]$ depends only on present and past values of $x$

- **Stability**: every limited input $x$ produces limited output $y$
  (Bounded-Input Bounded-Output, BIBO stability)
  - $|x[n]| \leq B_x \quad \forall n.$
Linear Time-Invariant (LTI) systems

- **Linear Time-Invariant** (LTI) system: is both linear and time invariant
- **Definition:** *unit sample sequence (or impulse)*
  \[ \delta[n] = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \]
  then \[ x[n] = \sum_{k=-\infty}^{\infty} x[k] \delta[n - k]. \]

- If the system is LTI:
  \[ y[n] = T \left\{ \sum_{k=-\infty}^{+\infty} x[k] \delta[n - k] \right\} = \sum_{k=-\infty}^{+\infty} x[k] T \{ \delta[n - k] \} = \sum_{k=-\infty}^{+\infty} x[k] h[n - k] \]

- **Convolution** operation:
  \[ y[n] = T \{ x \}[n] = (x * h)[n] \]

- **h[n]** is the system **impulse response**
  - **FIR** (finite impulse response) systems: \( h[n] \) has a finite number of non-null samples
  - **IIR** (infinite...) systems: \( h[n] \) has an infinite number of non-null samples
LTI systems: properties

- Convolution is linear, associative, commutative, distributive over sum

System properties translate into $h[n]$ properties

- Causal LTI systems:
  \[ h[n] = 0 \quad \forall n < 0 \]

- BIBO-stable LTI systems:
  \[ \sum_{k=-\infty}^{+\infty} |h[k]| < \infty \]
Examples

- A room is a LTI system
  - Sound source (impulse) at point $x_0$, microphone at point $x$
  - Room impulse response (RIR)

- A still listener is a LTI system (more or less...)
  - Sound source at angle $(\theta, \varphi)$, microphone inside ear canal
  - Head-Related Impulse Response (HRIR)
Impulse Response

- Time of flight
- Direct sound
- Early reflections
- Reverberant tail

Unit pulse $\delta$

System under test

System's impulse response
A simple linear system

Real-world system (one input, one output)

CD player → Amplifier → Loudspeaker → Microphone → Analyzer

“SYSTEM”

Block diagram

\[ x(\tau) \rightarrow h(\tau) \rightarrow y(\tau) \]

Input signal \rightarrow System’s Impulse Response (Transfer function) \rightarrow Output signal

see [http://www.openairlib.net](http://www.openairlib.net)
Discrete time spectral analysis

chapt. 1.4
The Family of Fourier Transform

Four classes of signals
- continuous vs. discrete
- periodic vs. aperiodic

<table>
<thead>
<tr>
<th>Time Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>periodic</td>
</tr>
<tr>
<td>discr. time</td>
</tr>
</tbody>
</table>

- periodic
- aperiodic
The Family of Fourier Transform

- angular frequency: $\omega = 2\pi f$
- normalized angular frequency: $\omega_d = 2\pi f / F_s$ range $[-2\pi, +2\pi]$
- sampling DTFT at $\omega_k = 2\pi k / N$ => $f_k = k F_s / N$

<table>
<thead>
<tr>
<th></th>
<th>Time</th>
<th>Duration</th>
<th>Finite</th>
<th>Infinite</th>
<th>discr.</th>
<th>time</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete FT (DFT)</td>
<td></td>
<td></td>
<td>$X[k] = \sum_{n=-0}^{N-1} x[n] e^{-j\omega_k}$</td>
<td>$X(\omega_d) = \sum_{n=-\infty}^{+\infty} x[n] e^{-j\omega_d n}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$k = 0, 1, \ldots, N - 1$</td>
<td>$\omega_d \in [-\pi, +\pi]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourier Series (FS)</td>
<td></td>
<td></td>
<td>$X(k) = \frac{1}{T} \int_{0}^{P} x(t) e^{-j\omega_k t} dt$</td>
<td>$X(\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$k = -\infty, \ldots, +\infty$</td>
<td>$\omega \in (-\infty, +\infty)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>discrete freq. $k$</td>
<td>continuous freq. $\omega$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discrete-Time Fourier Transform (DTFT)

- Continuous time signals

\[ \mathcal{F}\{x\}(\omega) = X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft} dt = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t} dt \]

- Discrete-time signals: Discrete-Time Fourier Transform (DTFT)

\[ \mathcal{F}\{x\}(\omega_d) = X(\omega_d) = \sum_{n=-\infty}^{+\infty} x(nT_s)e^{-j2\pi f \frac{n}{F_s}} = \sum_{n=-\infty}^{+\infty} x[n]e^{-j\omega_d n} \]

- Inverse DTFT (IDTFT)

\[ \mathcal{F}^{-1}\{X\}[n] = x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega_d)e^{j\omega_d n} d\omega_d \]

\[ \omega_d = \frac{2\pi f}{F_s} \] is the normalized (digital) angular frequency
## Properties

<table>
<thead>
<tr>
<th>Property</th>
<th>Time-domain sequences</th>
<th>Frequency-domain DTFTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity</td>
<td>$ax[n] + by[n]$</td>
<td>$aX(\omega_d) + bY(\omega_d)$</td>
</tr>
<tr>
<td>Time-shifting</td>
<td>$x[n - n_0]$</td>
<td>$e^{-j\omega_d n_0} X(\omega_d)$</td>
</tr>
<tr>
<td>Frequency-shifting</td>
<td>$e^{j\omega_0 n} x[n]$</td>
<td>$X(\omega_d - \omega_0)$</td>
</tr>
<tr>
<td>Frequency differentiation</td>
<td>$nx[n]$</td>
<td>$j \frac{dX}{d\omega_d}(\omega_d)$</td>
</tr>
<tr>
<td>Convolution</td>
<td>$(x * y)[n]$</td>
<td>$X(\omega_d) \cdot Y(\omega_d)$</td>
</tr>
<tr>
<td>Multiplication</td>
<td>$x[n] \cdot y[n]$</td>
<td>$\frac{1}{2\pi} \int_{-\pi}^{\pi} X(\theta)Y(\omega_d - \theta) d\theta$</td>
</tr>
<tr>
<td>Parseval relation</td>
<td>$\sum_{n=-\infty}^{+\infty} x[n]y^<em>[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega_d)Y^</em>(\omega_d) d\omega_d$</td>
<td></td>
</tr>
</tbody>
</table>
Frequency aliasing

One can prove that

\[ x_d[n] = x(nT_s) \]

\[ X_d(\omega_d) = F_s \sum_{m=-\infty}^{+\infty} X(\omega_dF_s + 2m\pi F_s) \]

\[ \Rightarrow X_d(\omega_d) \text{ is a periodization of } X(\omega) \]
Duality of sampling and periodicity

- Sampling in time corresponds to replication in the frequency domain
Duality of sampling and periodicity

- Sampling in frequency corresponds to replication in the time domain
Time Function

Boxcar
\[ G(t) = \begin{cases} 1, & |t| < \tau/2 \\ 0, & |t| > \tau/2 \end{cases} \]

Triangle
\[ G(t) = \begin{cases} 1-|t|/\tau, & |t| < \tau \\ 0, & |t| > \tau \end{cases} \]

Gaussian
\[ G(t) = e^{-t^2/2\tau^2} \]

Impulse
\[ G(t) = \delta(t) \]
\[ = 0, \quad t \neq 0 \]

Sinusoid
\[ G(t) = \cos \omega_0 t \]

Comb.
\[ G(t) = \text{comb}(t) \]
\[ = \sum_{-\infty}^{\infty} \delta(t-n\tau) \]

Frequency Function

Sinc
\[ S(f) = \tau \text{sinc}(f\tau) \]
\[ = (1/\pi f) \sin(\pi f\tau) \]

Sinc²
\[ S(f) = \tau \text{sinc}^2(f\tau) \]
\[ = (1/\pi^2 f^2\tau) \sin^2(\pi f\tau) \]

Gaussian
\[ S(f) = \tau (2\pi)^{1/2} e^{-(\pi f\tau)^2} \]

DC Shift
\[ S(f) = 1 \]

Single Freq.
\[ S(f) = \frac{1}{2} (\delta(f+f_0) + \delta(f-f_0)) \]

Comb.
\[ S(f) = \sum_{-\infty}^{\infty} \delta(f-n\tau) \]
Sampling theorem

- Signal $x(t)$ can be reconstructed from its sampled version if and only if its spectrum is 0 outside the so-called baseband
  - for all $|\omega| \geq \pi/F_s$
- The upper limit of the baseband is called Nyquist frequency
  - $f_{Ny} = F_s/2$ Hz
- audio signals: $F_s > 40$ kHz
Spectrum: analog vs. digital signal

- Sampling leads to a replication of the analog signal spectrum
- Reconstruction of the analog signal:
  - low pass filtering the digital signal
Short-time Fourier analysis

chapt. 1.5
Why is Fourier analysis so important?

- Interesting signals are time-varying quantities
- A time-varying value which only increases over time is not only a physical impossibility but a recipe for disaster
- Oscillations are nature’s and man’s way of keeping things in motion without trespassing all physical bounds
- Sinusoidal oscillations are the purest form of such a constrained motion
- Fourier: one could express any given phenomenon as the combined output of a number of sinusoidal “generators”

- Any linear time-invariant system transforms a sinusoid in a sinusoid at the same frequency
  - tool for the analysis and design of signal processing structures
- Ear as spectral analyser
Motivations

 Sounds are time-varying signals
  therefore, it is important to develop analysis techniques that allow to inspect some of their time-varying features.

 The “Short-Time Fourier Transform” (STFT) allows joint analysis of the temporal and frequency features of a sound signal
  in other words it allows to follow the temporal evolution of the spectral parameters of a sound.

 The main building block of the STFT is the Discrete Fourier Transform (DFT)
Discrete Fourier Transform - DFT

- Special case of DTFT applied to finite-length sequences
- Sampling of the DTFT

\[ X[k] = X(\omega_d)|_{\omega_d = 2\pi k/N} = \sum_{n=0}^{N-1} x[n]e^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N-1} x[n]W_N^{kn}, \quad 0 \leq k \leq N - 1 \]

(N-th root of one

- Inverse DFT

\[ x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k]W_N^{-kn}, \quad 0 \leq n \leq N - 1. \]

- Similar properties to DTFT
  - In particular: periodic convolution

\[ z[n] = (x \ast y)[n] \triangleq \sum_{m=0}^{N-1} x[n]y[n - m] \quad \Rightarrow \quad Z[k] = (X \cdot Y)[k], \]
Discrete Fourier Transform

\[ X(k) = \text{DFT}[x(n)] = \sum_{n=0}^{N-1} x(n)e^{-j2\pi nk/N} \quad k = 0, 1, \ldots, N - 1. \]

\[ X = \text{abs(fft(x,N))}/N; \]

- **Magnitude**
  \[ |X(k)| = \sqrt{X_R^2(k) + X_I^2(k)} \quad k = 0, 1, \ldots, N - 1 \]

- **Phase**
  \[ \varphi(k) = \arctan \frac{X_I(k)}{X_R(k)} \quad k = 0, 1, \ldots, N - 1 \]
Discrete Fourier Transform (example)

- FFT with 16 points
  - cosine (16 points)
  - magnitude (16 points)
    - normalization: 0 dB for sinusoid ±1
  - magnitude (frequency points)
    - $k f_s / N$
    - step $f_s / N$

- magnitude dB vs. Hz
  $$20 \log_{10} \left( \frac{X(k)}{N/2} \right)$$
DFT and FFT (Fast-Fourier Transform)

- DFT is the transformation:
  \[
  X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn}, \quad 0 \leq k \leq N - 1
  \]

- FFT is a fast algorithm for DFT computation
  - \( N \) power of 2: \( N = 2^s \)
    \[
    X[k] = \sum_{n=0}^{N/2-1} x[2n] W_N^{2kn} + \sum_{n=0}^{N/2-1} x[2n + 1] W_N^{k(2n+1)}
    \]
    \[
    = \sum_{n=0}^{N/2-1} x[2n] W_N^{2kn} + W_N^k \sum_{n=0}^{N/2-1} x[2n + 1] W_N^{2kn}
    \]
    \[
    = \sum_{n=0}^{N/2-1} x[2n] W_N^{kn} W_{N/2} + W_N^k \sum_{n=0}^{N/2-1} x[2n + 1] W_N^{kn}
    \]
  - A \( N \)-long DFT is a combination of two \( N/2 \)-long DFTs… then \( O(N \log N) \) implementations (recursive or iterative)
Resolution and leakage

- Complex exponential sequence
  \[ x[n] = e^{j\omega_0 n} \quad 0 \leq n < N \]

- This sequence is periodic only if \( \omega_0 N = 2\pi k \) for \( k = k_0 \)
  - There are only \( N \) such sequences that are periodic on \( N \) samples
  - For the periodic sequences the DFT is:
    \[ X[k] = \delta(k - k_0) \]
  - For the non-periodic ones there is leakage

\[ N = 64, k_0 = 20 \]

\[ N = 64, k_0 = 20.5, \]
Leakage and zero-padding

Zero-padding procedure

$$x[n] = \begin{cases} e^{j2\pi k_0 n/N}, & 0 \leq n < \frac{N}{2}, \\ 0, & \frac{N}{2} \leq n < N, \end{cases}$$

- Interpretation: it's the product of a sequence twice long with a rectangular window

$$x[n] = e^{j2\pi k_0 n/N} \cdot w_{N/2}[n] \quad w_{N/2}[n] = \begin{cases} 1, & 0 \leq n < \frac{N}{2}, \\ 0, & \frac{N}{2} \leq n < N \end{cases}$$

- then the DFT of $x[n]$ is the DFT of the rectangular window, shifted by $k_0$ elements
Zero-padding and resolution

Zero-padding a sequence can be exploited to increase the resolution of the spectral analysis.

\[ x[n] = e^{j2\pi k_0 n/N} \]
\[ N = 64, \quad k_0 = 20.5, \]

\[ x[n] = e^{j2\pi k_0 n/N} \cdot w_{N/2}[n] \]
\[ N = 128, \quad k_0 = 41. \]
Short-time Fourier transform (STFT)

**Definition:**

\[ X_n(\omega_d) = \sum_{m=-\infty}^{+\infty} w[n-m]x[m]e^{-j\omega_dm} \]

- \( w[n] \) is a window of length \( N \) that shifts in time.
- The STFT is the sum of the DFTs of \( x[n] \) windowed by time-shifted windows.
- Can be used to perform **time-frequency analysis**.

**Module of the STFT:** spectrogram

- Example: chirp
  \[ x[n] = A \cos \left( \frac{\omega_0}{2}(nT_s)^2 \right). \]

\[ \omega_0 = 2\pi \cdot 800 \text{ rad/s}^2 \]

\[ f = \omega \cdot n \cdot T_s \]
STFT: parameters

- **Length** $N$ of the window
  - By increasing $N$, frequency resolution is improved and temporal resolution is degraded, and vice versa
  - **Uncertainty principle**

- **Overlap** $N-m$ of the DFTs
  - If $m=N$ there is no overlap between windowed portions of the signal

- **Window shape**
  - can affect frequency leakage
    - Because the DFT of a windowed sinusoid is the shifted DFT of the window
  - There are better windows than the rectangular window
STFT: Windows

- To reduce leakage:
  - weight audio samples by a window

(Hann) \[ w[n] = \frac{1}{2} \left[ 1 + \cos \left( \frac{2\pi n}{2M + 1} \right) \right] \]

(Hamming) \[ w[n] = 0.54 + 0.46 \cos \left( \frac{2\pi n}{2M + 1} \right) \]

(Blackman) \[ w[n] = 0.42 + 0.5 \cos \left( \frac{2\pi n}{2M + 1} \right) + 0.08 \cos \left( \frac{4\pi n}{2M + 1} \right) \]
Spectrograms

\[ x(n) \]

DFT

\[ N = 8 \]

Amplitude

Frequency

Time signal

Spectrogram

Time
Waterfall representation

Waterfall Representation of Short-time FFTs

$x(n)$

Signal $x(n)$

Magnitude in dB

$f$ in Hz

$n$
STFT: example

- Parameters affect information displayed in spectrogram
  - file: diner.wav (from “Tom's Diner” by Suzanne Vega)
  - First: N=256, window=hanning, overlap=N/2
  - Second: N=2048, window=hamming, overlap=N/4
STFT example: piano solo

\[ F_s = 44.1 \text{ kHz}, \quad H=441, \quad N = 4096 \text{ (left) and } N = 1024 \text{ (right)} \]
STFT example: rock

\[ F_s = 44.1 \text{ kHz}, \quad H=441, \quad N = 4096 \text{ (left) and } N = 1024 \text{ (right)} \]
STFT example: symphony

\[ F_s = 44.1 \text{ kHz}, \quad H=441, \quad N = 4096 \text{ (left) and } N = 1024 \text{ (right)} \]
Filters: transfer function and frequency response

chapt. 1.6
Z-transform

\[ Z\{x\}(z) = X(z) = \sum_{n=-\infty}^{+\infty} x[n]z^{-n} \quad z \in \mathbb{C} \]

- Defined (i.e., the series converges) only on a certain subset of \( \mathbb{C} \): region of convergence (ROC)
- Becomes the DTFT when \( z \) is on the unit circle: \( z = e^{j\omega} \)

<table>
<thead>
<tr>
<th>Property</th>
<th>Sequences</th>
<th>( z )-Transforms</th>
<th>ROCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity</td>
<td>( ax[n] + by[n] )</td>
<td>( aX(z) + bY(z) )</td>
<td>contains ( R_x \cap R_y )</td>
</tr>
<tr>
<td>Time-shifting</td>
<td>( x[n - n_0] )</td>
<td>( z^{-n_0}X(z) )</td>
<td>( R_x )</td>
</tr>
<tr>
<td>( z )-scaling</td>
<td>( z_0^n x[n] )</td>
<td>( X(z/z_0) )</td>
<td>(</td>
</tr>
<tr>
<td>( z )-differentiation</td>
<td>( nx[n] )</td>
<td>( -z\frac{dX(z)}{dz} )</td>
<td>( R_x )</td>
</tr>
<tr>
<td>Conjugation</td>
<td>( x^*[n] )</td>
<td>( X^<em>(z^</em>) )</td>
<td>( R_x )</td>
</tr>
<tr>
<td>Time-reversal</td>
<td>( x^*[-n] )</td>
<td>( X^<em>(1/z^</em>) )</td>
<td>( 1/R_x )</td>
</tr>
<tr>
<td>Convolution</td>
<td>( (x * y)[n] )</td>
<td>( X(z) \cdot Y(z) )</td>
<td>contains ( R_x \cap R_y )</td>
</tr>
</tbody>
</table>

Initial value theorem: If \( x[n] \) causal (i.e. \( x[n] = 0 \ \forall \ n < 0 \)), then \( \lim_{z \to \infty} X(z) = x[0] \).
LTI systems and Z-transform

- \[ y[n] = T\{x\}[n] = (x * h)[n] \]
  \[ Y(z) = H(z)X(z) \]

- \( H(z) \) is the system **transfer function**

- \[ H(e^{j\omega_d}) \] is the system **frequency response**
  - it is the DTFT of the impulse response \( h[n] \),
  - it exists only if the ROC includes the unit circle
  \[ Y(e^{j\omega_d}) = H(e^{j\omega_d})X(e^{j\omega_d}) \]

- \[ |H(e^{j\omega_d})| \] is the system **magnitude response**

- \[ \text{arg}\left[H(e^{j\omega_d})\right] \] is the system **phase response**
Frequency filtering

- Effect of a LTI system in the frequency domain
  - For a sinusoidal input: \( x[n] = A \cos(\omega_0 n) \)
    \[
    y[n] = (h \ast x)[n] = A \left| H(e^{j\omega_0}) \right| \cdot \cos(\omega_0 n + \arg[H(e^{j\omega_0})])
    \]
    - magnitude response at \( \omega_0 \) defines the \textit{gain},
    - phase response at \( \omega_0 \) defines the \textit{phase delay}
  
- For a generic input:
  \[
  \left| Y(e^{j\omega_d}) \right| = \left| H(e^{j\omega_d}) \right| \cdot \left| X(e^{j\omega_d}) \right|, \\
  \arg[Y(e^{j\omega_d})] = \arg[H(e^{j\omega_d})] + \arg[X(e^{j\omega_d})]
  \]
  - attenuation or boost of certain frequency bands (e.g., equalization)
  - frequency-dependent phase delay (phase distortion)

- LTI system as a \textit{frequency selective filter}, or simply a \textit{filter}. 
Also, there are all-pass filters with flat magnitude response (they only act on signal phase)
About phase response

- Linear phase response is often desirable
  - Means same phase delay at all frequencies
- Measure of linearity of the phase response:
  - **group delay**

\[
\tau(\omega_d) = -\frac{d}{d\omega_d} \left\{ \text{arg} \left[ H (e^{j\omega_d}) \right] \right\}
\]

\[
x[n] = a[n]e^{\omega_0 n}
\]
Filters with rational transfer function

- **LTI system defined by difference equation:**

\[
\sum_{k=0}^{N} a_k y[n - k] = \sum_{k=0}^{M} b_k x[n - k] \iff \sum_{k=0}^{N} a_k z^{-k} Y(z) = \sum_{k=0}^{M} b_k z^{-k} X(z),
\]

- N.B. one also needs initial conditions
- or causal system with \(x[n]=0\) for \(n<0\). Then

\[
y[n] = -\sum_{k=1}^{N} \frac{a_k}{a_0} y[n - k] + \sum_{m=0}^{M} \frac{b_m}{a_0} x[n - m]
\]

- **Rational transfer function (ratio of polynomials):**

\[
H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}}
\]

- **FIR systems:** \(a_k = 0\) for all \(k>0\)
  - always stable, no poles
- **IIR systems:** \(a_k \neq 0\) for some \(k>0\)
  - stability: all poles must be inside unit circle
Block diagrams

- **Main elements**
  - Delay
  - Multiplier

- **Computational structure of a generic filter with rational transfer function**
### Acoustic Interpretation

<table>
<thead>
<tr>
<th>Spectral Features</th>
<th>Filter Type</th>
<th>Acoustic Analog</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recursive (poles)</td>
<td>Stored Energy Resonance</td>
</tr>
<tr>
<td></td>
<td>Non-Recursive (zeros)</td>
<td>Cancelled Energy Anti-Resonance</td>
</tr>
</tbody>
</table>

- **Induced Pitch**
Pass-band and stop-band diagrams with cutoff frequency $f_c$ (half power point or -3 dB point) and bandwidth.

- Pass-band: Frequency range where the gain is above 0 dB.
- Stop-band: Frequency range where the gain is below 0 dB.
- Cutoff frequency $f_c$: Point where the gain drops to half of its maximum value (3 dB down).
- Bandwidth: Frequency range between the two points where the gain is 0.7071 times the maximum gain.
Example: 1st order filters

FIR:

\[
H_{lp}(z) = \frac{1}{2} \left( 1 + z^{-1} \right), \quad H_{hp}(z) = \frac{1}{2} \left( 1 - z^{-1} \right).
\]

Intuitively

- \(H_{lp}\) outputs the average of the last two samples: zero at \(z = -1\)
- \(H_{hp}\) outputs the half-difference of the last two samples: zero at \(z = 1\)

IIR:

\[
H_{lp}(z) = \frac{1 - \alpha}{2} \frac{1 + z^{-1}}{1 - \alpha z^{-1}}, \quad H_{hp}(z) = \frac{1 + \alpha}{2} \frac{1 - z^{-1}}{1 - \alpha z^{-1}}.
\]

- Real pole in \(z = \alpha\)
- Stability if and only if \(|\alpha| < 1\)
- Changing \(\alpha\) changes filter selectivity
The simplest IIR filter

single pole

\[ y(n) = \frac{1}{2} y(n - 1) + \frac{1}{2} x(n) \]

\[ H(z) = \frac{1/2}{1 - \frac{1}{2} z^{-1}} \]

\[ H(\omega) = \frac{1/2}{1 - \frac{1}{2} e^{-j\omega}} \]
Fir example

- computation of frequency response:
  - impulse response
  - magnitude resp.
  - pole/zero plot
  - phase resp.

- h=[-.1 .15 .3 .15 -.1]

- Fir example computation of frequency response:
  h=[-.1 .15 .3 .15 -.1]
  - impulse response
  - magnitude resp.
  - pole/zero plot
  - phase resp.
Summary

Take-home concepts

- **Sampling** (sampling theorem, sampling/Nyquist frequencies)
- **Spectrogram** (STFT, DFT, parameters, uncertainty principle)
- **Impulse response** (LTI systems, impulse, convolution)
- **Frequency response** (magnitude/phase, transfer function, ...)
- **Filter classification** and **block diagrams** (\( z^{-k} \) and delays, ...)
Matlab/Octave & Puredata examples

- **Matlab**
  - Spectrograms: `chirp_spectrograms.m`, `diner_spectrogram.m`
  - Envelope generator: `envgen.m`
  - Amplitude/frequency controlled sin. oscillator: `sinosc.m`
  - Sinusoids with varying ampl/freq: `costsin.m`, `modsin.m`
  - First-order low- and high-pass filters: `filter_examples.m`

- **Puredata**
  - Amplitude/frequency controlled sin. oscillator: `sinosc.pd`
  - White/pink noise generators: `white_pink_noise.pd`
  - Examples from Pd Help Browser
    - **D04**: ADSR envelope for amplitude modulation
    - **D06**: frequency modulation and portamento
    - **D14**: frequency modulation and vibrato
    - **H01-02-03**: high/low/band-pass filters
    - **H05**: “voltage-controlled” filters (VCF) and “On the Run”
    - **H16**: voltage controlled filters and “analog” synthesizer
  - High order and band-pass filters: `highorderfilters.pd`, `filtercolors.pd`
Matlab/Octave functions/scripts

- chirp
- wavread (deprecated)
- wavwrite (deprecated)
- audioread
- audiowrite
- sound
- soundsc
- fft
- specgram (deprecated)
- spectrogram
- hamming
- hanning
- blackman
- filter
- freqz
- grpdelay