Automatic Speech Recognition
Lecture Note 2: Digital Signal Processing

1. **The z Transform:** The z transform of a discrete-time signal $x[n]$ is defined as

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n},$$

where z is a complex variable. Note that z transform is a linear operation. Moreover, the z transform of $x[n-m]$ is equal to the z transform of $x[n]$ multiplied by z^{-m}.

2. **The Inverse z Transform:**

$$x[n] = \frac{1}{2\pi j} \oint X(z)z^{n-1}dz.$$

3. **Convolution:** The convolution of two discrete-time signals $x[n], h[n]$ is defined by

$$y[n] = \sum_m x[m]h[n-m].$$

Theorem: $Y(z) = X(z)H(z)$.

proof: Direct application of the z transform to $y[n]$ yields

$$\begin{aligned} Y(z) &= \sum_n y[n]z^{-n} = \sum_n \sum_m x[m]h[n-m]z^{-n} \\ &= \sum_n \sum_m x[m]h[n-m]z^{-(n-m)}z^{-m} \\ &= \sum_m x[m]z^{-m} \sum_n h[n-m]z^{-(n-m)} \\ &= X(z)H(z). \end{aligned}$$

4. **Linear Time-Invariant Systems:** A system is linear if linear combination of input signals produces linear combination of corresponding output signals. A system is time-invariant if delayed input signal produces delayed corresponding output signal. A system is linear time-invariant (LTI) if it is linear and time-invariant.

Note: The output of an LTI system is the convolution of input sequence and the impulse response.

5. **The Fourier Transform:** The special case of the z transform when z lies on the unit circle is the *Fourier transform*,

$$X(e^{j\omega}) = \sum_n x[n] e^{-j\omega n}.$$

The *inverse Fourier transform* is

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega.$$

The sequence of $x[n]$ is represented as a sum of sinusoids. The contribution of frequency ω is $X(e^{j\omega})$.

Note that the convolution theorem for the z transform holds similarly for Fourier transforms.

6. **Sampling:** A discrete-time sequence $x[n]$ can be produced from a continuous signal $x(t)$ by sampling. If the sampling is periodic with period T, then the resultant spectrum is periodic with period $\frac{1}{T}$. In fact it is the sum of equally spaced replica of the spectrum of $x(t)$.

Nyquist Sampling Theorem: A continuous signal can be reconstructed by its periodic samples if the sampling frequency is greater than twice the highest frequency component in the signal.

7. **Linear Difference Equations:** A first-order linear difference equation is defined by

$$y[n] = Ky[n-1] + x[n].$$

Taking the z transform, one has

$$Y(z) = Kz^{-1}Y(z) + X(z),$$

which leads to the transfer function of

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - Kz^{-1}}.$$

Thus there is a zero at $z = 0$ and a pole at $z = K$. A pole (zero) is a point in the z plane where the transfer function is singular (zero).

A *resonance frequency* is the angular frequency ω where the magnitude of the transfer function $|H(e^{j\omega})|$ peaks. In the above example, if K is positive (negative), the resonance frequency is $\omega = 0$ (π).

A second-order difference equation is defined by

$$y[n] = Ay[n-1] + By[n-2] + x[n],$$

and the transfer function is

$$\frac{1}{1 - Az^{-1} - Bz^{-2}}.$$

This system has two poles as conjugate complex numbers if $A^2 \leq -4B$. In this case, let $B = -r^2, A = 2r\cos\theta$, then

$$z_{1,2} = re^{\pm j\theta}.$$

Resonance can also be obtained by cascading two linear systems, with

$$H_1(z) = 1 - z^{-M}$$

and

$$H_2(z) = \frac{1}{1 - 2\cos\theta z^{-1} + z^{-2}}.$$

Here $H_1(z)$ represents M equally spaced zeros on the unit circle, and $H_2(z)$ represents 2 poles also on the unit circle. These poles can be used to cancel the zeros on the unit circle, resulting in resonances.