

Model of Speech Production

What is a source filter model

A source filter model is a mathematical model that represent a speech signal combining sound source with a linear acoustic filter

It's usual modeled with discrete time lti system.

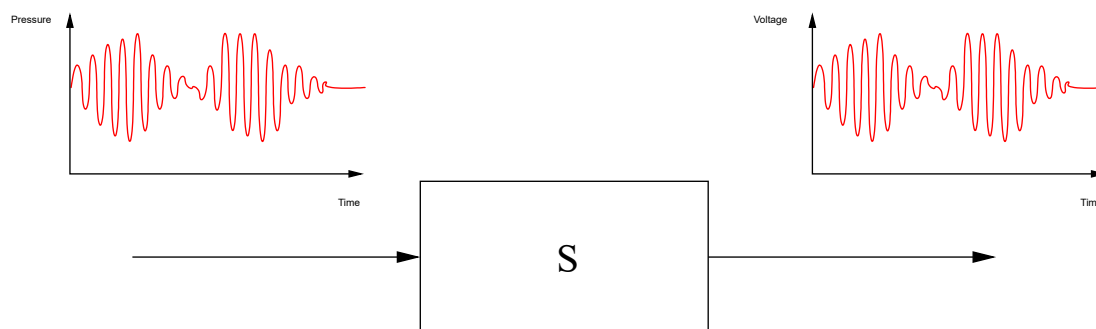
It aims to to model tghe way the vocal tract shapes the speech sound.

What is an acoustic filter

An acoustic filter is a mathematical model that represent the vocal tract.

What is a microphone?

A microphone is a system that convert an acoustic signal into an electrical signal.



What is a system?

A system in signal processing is model that takes in input a signal and produce an output signal. We can see this as a transfer function. that relates the input signal to the output signal.

We can see them as functions that maps the function space into another function space. For example:

$$F_{space} \rightarrow F_{space}$$

$$S : [Time] \rightarrow [Pressure] \rightarrow [Time \rightarrow Pressure]$$

For example, for dicrsrete time signals the function space is:

$$X = [\mathbb{Z} \rightarrow \mathbb{R}]$$

where \mathbb{Z} is the set of integers and \mathbb{R} is the set of real numbers.

What is a causal system?

A causal system is a particular type of system in which the output doesn't depend on future values of the input signal, but only on the previous and current values of it.

An example of causal system

The *unit delay* system is a causal system. It's a system that takes the input signal and produce the same signal but delayed by one sample.

What is a linear system

A linear system is a system that must satisfy 2 properties:

- Additivity: $S(x_1 + x_2) = S(x_1) + S(x_2)$
- Homogeneity: $S(ax) = aS(x)$

The first property means that the system is additive, that is, the output of the sum of two signals is the sum of the output of the two signals.

The second property means that the system is homogeneous, that is, the output of a signal multiplied by a constant is the same as the output of the signal multiplied by the constant.

Important: If we know the response of a system to some signal we can compute all the linear combinations of the signals.

What is a time-invariant system

A time invariant system is a system that doesn't change over time. That is, **if we shift the input signal the output signal will be shifted by the same amount.**

$$S(X(n)) = y(n) \xrightarrow{\text{time invariance}} S(X(n - k)) = y(n - k)$$

We can see that the output is shifted by the same amount as the input.

What is a Linear Time Invariant LTI system?

Is a system that is both linear and time invariant. So it must satisfy the properties of:

- Additivity
- Homogeneity
- Time invariance

What is a discrete time unit impulse?

Is a discrete time signal in which all the values are zero except for one value that is 1 at time 0.

$$\delta(n) = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{otherwise} \end{cases}$$

Why any discrete time signal can be expressed as sum of scaled time shift and unit impulses?

Any discrete time signal can be expressed as a sum of scaled time shift and unit impulses because of the linearity and time invariance properties of the LTI systems.

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k)$$

Example: Let's say we have a signal $x(n)$ that is a sum of two impulses:

$$x(n) = 2\delta(n-1) + 3\delta(n-2)$$

We can express this signal as a sum of scaled time shift and unit impulses:

$$x(n) = 2\delta(n-1) + 3\delta(n-2) = 2\delta(n) + 3\delta(n-1)$$

What is a discrete time unit step signal?

Is a particular type of signal that is zero for all negative values and 1 for all positive values of time n .

$$u(n) = \begin{cases} 1 & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases}$$

This is related to the *discrete time unit impulse*

What is the impulse response of a system?

Is the output of a discrete time system when the input is a **unit impulse**.

$$h(n) = S(\delta(n))$$

where h is the impulse response of the system S .

What is the convolutional sum?

The convolutional sum is a mathematical operation that takes two signals and produce a third signal. It's defined as:

$$y(n) = x(n) * h(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k)$$

where $x(n)$ is the input signal, $h(n)$ is the impulse response of the system and $y(n)$ is the output signal. The operator $*$ is the convolutional sum.

What is a eigenfunction of a system?

An eigenfunction of a system is a function that when passed through the system is scaled by a constant.

We say that a function $p(n)$ is an eigenfunction of a system S if:

$$T(p(n)) = \lambda p(n)$$

where λ is the eigenvalue of the system.

What a difference equation is?

It's a mathematical equation that describes the evolution of a sequence of values over time.

Where can we use difference equations?

It's good because we can describe the *discrete time LTI system* with difference equations.

$$y(n) = \sum_{k=0}^N a_k x(n-k) = \sum_{k=0}^M b_k y(n-k)$$

where the first part is the input signal and the second part is the output signal.

What is a frequency response?

A frequency response is a function that describes how the system responds to different frequencies.

It's a mathematical representation of the system behavior respect to the frequency domain.

It's a particular case of **TRANSFER FUNCTION**

What is a z-transform?

It's a mathematical operation that takes a discrete time signal and produce a complex function of a complex variable z .

The transfer function is the z-transform of the impulse response.

How is related the z-transform with the Fourier Transform?

The z-transform is a generalization of the Fourier Transform. The Fourier Transform is a particular case of the z-transform when $z = e^{j\omega}$.

A correlation is that the DTT is the z-transform over the unit circle, that is, when $z = e^{j\omega}$.

How does the shifting works on z-transform?

If a signal is shifted by $x(n)$ by k samples, the z-transform of the shifted signal is:

$$X(z) = z^{-k} X(z)$$

where $X(z)$ is the z-transform of the original signal.

What is a unit delay

A unit delay is a system that takes the input signal and produce the same signal but delayed by one sample.

Usually, with the z-transform, we can say that the unit delay is represented as a shift register in the time domain or as a multiplication by z^{-1} in the z-domain.

What is a transfer function

A transfer function in signal processing is a mathematical function that represents how an input signal is transformed into an output signal using a linear and time invariant system (LTI)

Remember: Discrete time complex exponentials signal is a *eigen function* of a linear time invariant system. (LTI)

The transfer function it's a complex number.

Rational transfer function

Using time shifting property and convolutional property we can see that the transfer function of an LTI system is the z-transform of the output signal divided by the z-transform of the input signal.

$$H(z) = \frac{Y(z)}{X(z)}$$

So, the transfer function of an LTI system is rational.

$$H(z) = \frac{P(z)}{Q(z)}$$

where $P(z)$ and $Q(z)$ are polynomials in z .

Remember, if $P(z) = 0$, the system goes to 0. If $Q(z) = 0$, the system goes to infinity.

- Roots of P are called **zeros**
- Roots of Q are called **poles**

Convolutional property of the z-transform

If a DTS (Discrete Time Signal) is the impulse response of a discrete time LTI system, the output of the system can be computed using the convolutional sum.

Finite impulse response system (FIR)

It's a system of a discrete time LTI system where the output depends on a finite set of values of the input signal. It's known even as FIR filter.

The difference equation of a FIR system depends only on the current and past values of the input signal.

The coefficient b_k are the impulse response of the system, where the impulse response is the output of the system when the input is a unit impulse.

Where the FIR filters are used?

In audio and image processing, communication and biomedical signal analysis.

NOTE: Discrete time LTI system have an **INFINITE** impulse response system

What is an ideal low-pass filter?

It's a filter that passes all the frequencies components of the input signal higher than a threshold frequency and blocks everything below that frequency. The frequency is called **cutoff frequency** ω_c .

Why the ideal low-pass filter is not realizable?

The impulse response have values from $-\infty$ to ∞ . We can calculate an approximation using a window function and transforming it into a finite impulse response system filter.

What is an infinite impulse response system (IIR)

It's a type of discrete time LTI system where the output depends both on a finite number of input and output samples. It's known as a IIR Filter.

`scipy.signal.lfilter` is an example of an IIR filter.

Can be described by a difference equation of the form:

$$y(n) = \frac{1}{a_0} \left(\sum_{k=0}^M b_k x(n-k) - \sum_{k=1}^N a_k y(n-k) \right)$$

where a_0 is often set to 1.

This system is an infinite impulse response system if there is at least another non-zero a_k parameter.

What is a window function?

A window function are signals concentrated in time. They are finite duration discrete time signal that are applied to another signal by multiplying them.

Some examples are:

- Rectangular window:
- Hamming window
- Hanning window
- Blackman window

What are formants?

Formant is the resonance in the vocal tract that represent a peak of energy in the speech signal spectrum. They are important for speech recognition and synthesis.

Formant are numbered from lower to higher frequencies.

The frequency of the formants uniquely identify the vowel.

What is a cascade combination of systems?

A cascade combination of two systems is a system that takes the output of the system 1 and gives it as input to the system 2.

$$h(t) = h_1(t) * h_2(t)$$

Maybe we get multiple tubes as a result of the cascade combination.

What is a continuous time resonator?

A system that models a resonance with two poles and no zeroes in the transfer function.

Can be used to model a single resonance of the vocal tract, called **formant**. It's defined by formant frequency and formant bandwidth.

F = 500Hz and BW = 100Hz

What is a second order All Pole IIR system

It's a discrete time system that has two poles and no zeroes in the transfer function.

What does the linear combination do?

It tries to predict a signal sample using the linear combination of the previous samples.

The prediction error is called **residue**

What is the inverse of the linear prediction?

Is an All-Pole filter that can be seen as a vocal tract transfer function.

What the goal of the linear prediction optimization?

The goal is to minimize the energy of the prediction error. So the goal is finding the coefficients a_k that minimize the energy of the prediction error.

The name of the equations that create this problem is called **Yule-Walker equations**

How does the autocorrelation method work?

It assumes that the input signal is a finite duration discrete time signal that's the output from a window process.

The output for the autocorrelation method returns us the coefficients of the linear prediction.

What is a feature of the Toeplitz matrix?

The Toeplitz matrix is a matrix in which all the elements in the diagonal are the same. It's used in the autocorrelation method.