Lecture 7: Analysis of Nasal Consonants

Mark Hasegawa-Johnson

ECE 537: Speech Processing Fundamentals

Outline

The Two-Sided Laplace Transform

$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt$$

Why You've Never Seen a Two-Sided Laplace Transform Before

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

Warning: here are a few signals for which $X(s)$ fails to converge for any value of s:

$$x(t) = 1$$
$$x(t) = \cos(200\pi t)$$
$$x(t) = e^{-0.1t}$$

Why It's Useful Anyway

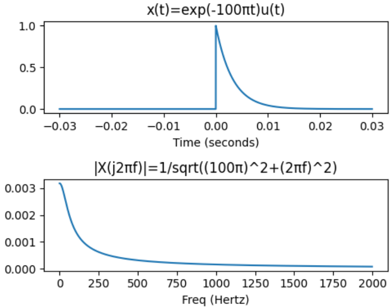
- For practical purposes, the two-sided Laplace transform only converges for two types of signals:
    - Exponentially-damped causal signals
    - Finite-duration, finite-amplitude signals
- ... but if it converges, it is way, way, way easier to work with than any other type of transform.

Example: Causal Exponential

$$x(t) = e^{at}u(t)$$

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$
$$= \frac{1}{s-a}$$

The above transform is valid for $\Re(s) > \Re(a)$.



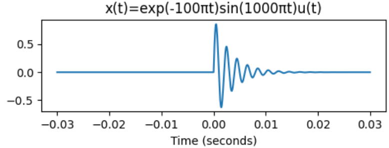
Example: Damped Sinusoid

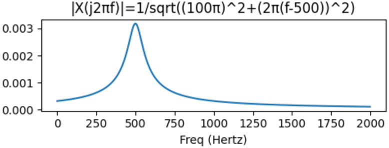
$$x(t) = e^{(a+j\omega)t}u(t)$$

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$

$$= \frac{1}{s-(a+j\omega)}$$

The above transform is valid for $\Re(s) > \Re(a)$.



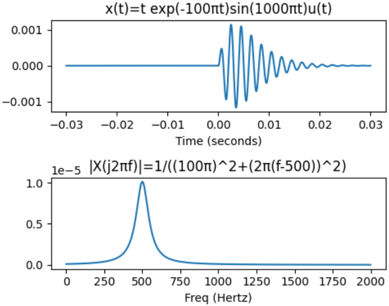


Example: Gammatone

$$x(t) = te^{(a+j\omega)t}u(t)$$

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt$$
$$= \left(\frac{1}{s-(a+j\omega)}\right)^2$$

The above transform is valid for $\Re(s) > \Re(a)$.

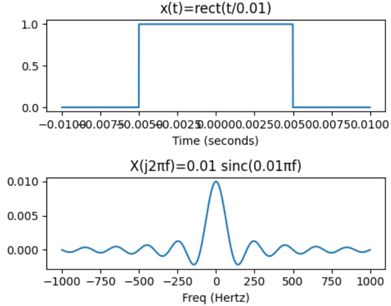


Example: Rectangle

$$x(t) = \text{rect}\left(\frac{t}{T}\right) = \begin{cases} 1 & |t| \le \frac{T}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$X(s) = \frac{1}{s}\left(e^{sT/2} - e^{-sT/2}\right)$$
$$= \frac{2}{s}\sinh\left(\frac{sT}{2}\right),$$

Like the tranforms of every other finite-duration, finite-amplitude signal, the above transform is valid for all finite values of s.



Hyperbolic Functions

The three main hyperbolic functions are:

$$\cosh(x) = \frac{1}{2}\left(e^{x} + e^{-x}\right)$$
$$\sinh(x) = \frac{1}{2}\left(e^{x} - e^{-x}\right)$$
$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}$$

Notice that hyperbolic functions are just like trigonometric functions, but easier:

$$\cos(\theta) = \frac{1}{2}\left(e^{j\theta} + e^{-j\theta}\right)$$
$$\sin(\theta) = \frac{1}{2j}\left(e^{j\theta} - e^{-j\theta}\right)$$
$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$$

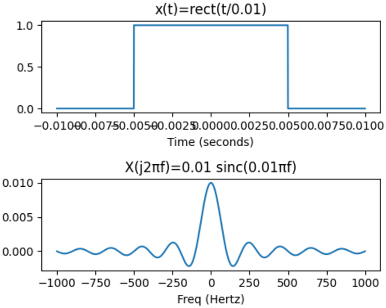
Example: Rectangle

$$x(t) = \text{rect}\left(\frac{t}{T}\right)$$

$$X(s) = \frac{2}{s}\sinh\left(\frac{sT}{2}\right)$$

In particular, if we plug in $s = j\omega$, we get that

$$X(j\omega) = T\text{sinc}\left(\frac{\omega T}{2}\right)$$



Outline

Delay Property

If

$$y(t) = x(t-d),$$

then

$$\begin{aligned} Y(s) &= \int_{-\infty}^{\infty} y(t) e^{-st} dt \\ &= X(s) e^{-sd} \end{aligned}$$

Outline

The One-Dimensional Wave Equation

No matter what type of wave you're talking about (water, radio, light, sound, Slinky), they are all written in exactly this way:

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2},$$

where t =time, x =position, c =speed of the wave, and $p(x,t)$ is the quantity that moves in waves.

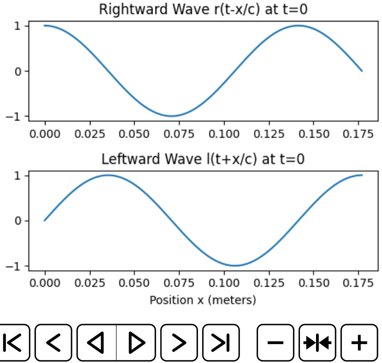
Solution to the Wave Equation

$$\frac{\partial^2 p(x,t)}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p(x,t)}{\partial t^2},$$

The general form of the solution is:

$$p(x,t) = r\left(t - \frac{x}{c}\right) + l\left(t + \frac{x}{c}\right)$$

where $r(t)$ is a rightward wave, and $l(t)$ is a leftward wave.



Laplace Transform of the Solution to the Wave Equation

$$p(x,t) = r\left(t - \frac{x}{c}\right) + l\left(t + \frac{x}{c}\right)$$

Let's take the Laplace transform of that:

$$\begin{aligned} P(x,s) &= \int_{-\infty}^{\infty} p(x,t)e^{-st}dt \\ &= R(s)e^{-xs/c} + L(s)e^{xs/c} \end{aligned}$$

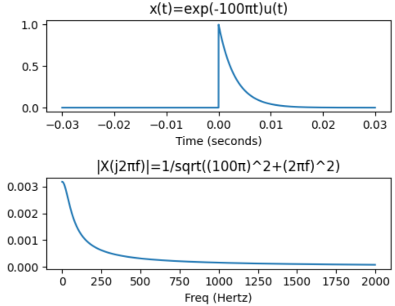
Example: Causal Exponential

For example, suppose:

$$p(x,t) = e^{a\left(t-\frac{x}{c}\right)} u\left(t-\frac{x}{c}\right)$$

Its transform is just

$$P(x,s) = \frac{1}{s-a} e^{-sx/c}$$



Volume Velocity

- A wave equation is always caused by the trading of energy between two quantities, e.g., between electricity and magnetism, or between kinetic and potential energy in a Slinky or guitar string.
- In acoustics, pressure = potential energy, velocity = kinetic energy.
- In a tube like the vocal tract, instead of using average velocity, it makes more sense to use volume velocity.

Definition: Volume Velocity is average air particle velocity, multiplied by the cross-sectional area of the tube.

$$\left[\frac{m}{s}\right] \times \left[m^2\right] = \left[\frac{m^3}{s}\right]$$

- Why this makes sense: if you blow 1 liter/second into the small end of a tube, you expect 1 liter/second to come out the big end.

Volume Velocity

The relationship between pressure and volume velocity is:

$$p(x,t) = r\left(t - \frac{x}{c}\right) + l\left(t + \frac{x}{c}\right),$$

$$u(x,t) = \frac{A(x)}{\rho c}\left(r\left(t - \frac{x}{c}\right) - l\left(t + \frac{x}{c}\right)\right),$$

where $A(x)$ is cross-sectional area, and ρ is the density of air.

Laplace Transform of Volume Velocity

The relationship between pressure and volume velocity is:

$$
\begin{aligned}
P(x,s) &= R(s)e^{-sx/c} + L(s)e^{sx/c},\\
U(x,s) &= \frac{A(x)}{\rho c}\left(R(s)e^{-sx/c} - L(s)e^{sx/c}\right)
\end{aligned}
$$

Outline

Radiation from the Lips

The cross-sectional area of the room is much larger than the cross-sectional area of your lips. No matter how much air you blow out of your lips, you are not going to appreciably change the air pressure in the room.
We can express this constraint by saying that, at the lips ($x = d_l$),

$$p(d_l, t) \approx 0$$
$$P(d_l, s) \approx 0$$

Public Domain,
https://commons.wikimedia.org/wiki/File:

The Zero-Pressure Constraint at the Lips

Recall that

$$p(x,t) = r\left(t - \frac{x}{c}\right) + l\left(t + \frac{x}{c}\right),$$
$$P(x,s) = R(s)e^{-sx/c} + L(s)e^{sx/c}.$$

If we apply the condition that $p(d_l, t) = 0$, we learn that $l(t)$ is a reflection of $r(t)$, delayed by $2d_l/c$ and multiplied by -1:

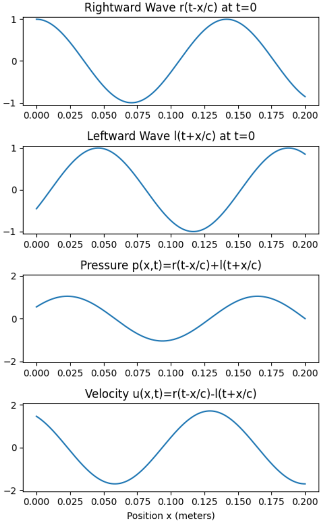
$$l(t) = -r\left(t - \frac{2d_l}{c}\right)$$
$$L(s) = -R(s)e^{-2sd_l/c}$$

Reflections from the Lips

The equation $p(d_l, t) = 0$ is satisfied if $l(t)$ is a reflection of $r(t)$, multiplied by -1:

$$l(t) = -r\left(t - \frac{x - 2d_l}{c}\right)$$

$$L(s) = -R(s)e^{-2sd_l/c}$$



Dividing $R(s)$ from both Numerator and Denominator

Fujimura had the following brilliant insight. (1) Both pressure and velocity are proportional to $R(s)$. (2) $R(s)$ is totally arbitrary, it can be anything. (3) The math will be easier if we just get rid of $R(s)$.

He did this by defining a quantity called "susceptance" ($B(x,s)$), the ratio of volume velocity over pressure:

$$\begin{aligned}
B(x,s) &= \frac{U(x,s)}{P(x,s)} \\
&= \frac{A(x)}{\rho c}\left(\frac{R(s)e^{-sx/c} - L(s)e^{sx/c}}{R(s)e^{-sx/c} + L(s)e^{sx/c}}\right) \\
&= \frac{A(x)}{\rho c}\left(\frac{e^{-s(x-d_l)/c} + e^{s(x-d_l)/c}}{e^{-s(x-d_l)/c} - e^{s(x-d_l)/c}}\right),
\end{aligned}$$

where the last line takes advantage of the constraint at the lips:

$$L(s) = -R(s)e^{-2sd_l/c}$$

Outline

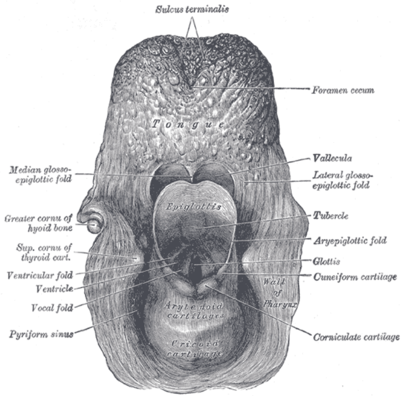
Closure at the Glottis

The cross-sectional area of the glottis is much smaller than the cross-sectional area of your vocal tract.

No matter how high the pressure gets inside the vocal tract, it is not able to push very much air back through the glottis.

We can express this constraint by saying that, at the glottis ($x = d_g$),

$$u(d_g, t) \approx 0$$
$$U(d_g, s) \approx 0$$



Public Domain, https://commons.wikimedia.org/wiki/File:Gray955.png

The Zero-Velocity Constraint at the Glottis

Recall that

$$u(x,t) = \frac{A(x)}{\rho c}\left(r\left(t-\frac{x}{c}\right) - l\left(t+\frac{x}{c}\right)\right),$$

$$U(x,s) = \frac{A(x)}{\rho c}\left(R(s)e^{-sx/c} - L(s)e^{sx/c}\right).$$

If we apply the condition that $u(d_g,t)=0$, we learn that $l(t)$ is a reflection of $r(t)$, delayed by $2d_g/c$ and multiplied by $+1$:

$$l(t) = r\left(t - \frac{x-2d_g}{c}\right)$$

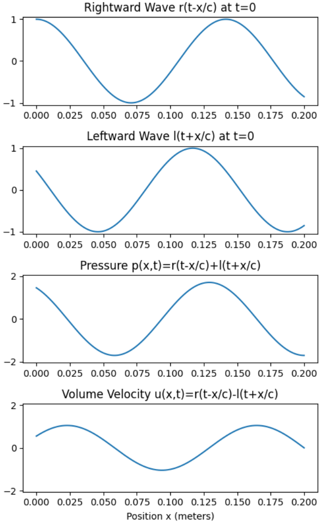
$$L(s) = R(s)e^{-2sd_g/c}$$

Reflections from the Glottis

The equation $u(d_g, t) = 0$ is satisfied if $l(t)$ is a reflection of $r(t)$, multiplied by -1:

$$l(t) = r\left(t - \frac{x - 2d_g}{c}\right)$$

$$L(s) = R(s)e^{-2sd_g/c}$$



Dividing $R(s)$ from both Numerator and Denominator

Fujimura had the following brilliant insight. (1) Both pressure and velocity are proportional to $R(s)$. (2) $R(s)$ is totally arbitrary, it can be anything. (3) The math will be easier if we just get rid of $R(s)$.

He did this by defining a quantity called "susceptance" ($B(x,s)$), the ratio of volume velocity over pressure:

$$\begin{aligned} B(x,s) &= \frac{U(x,s)}{P(x,s)} \\ &= \frac{A(x)}{\rho c}\left(\frac{R(s)e^{-sx/c} - L(s)e^{sx/c}}{R(s)e^{-sx/c} + L(s)e^{sx/c}}\right) \\ &= \frac{A(x)}{\rho c}\left(\frac{e^{-s(x-d_g)/c} - e^{s(x-d_g)/c}}{e^{-s(x-d_g)/c} + e^{s(x-d_g)/c}}\right), \end{aligned}$$

where the last line takes advantage of the constraint at the glottis:

$$L(s) = R(s)e^{-2sd_g/c}$$

Outline

Conclusions

- A Laplace transform is a Fourier transform, but easier:

$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt$$

- Every one-dimensional wave equation has the following solution:

$$P(x,s) = R(s)e^{-sx/c} + L(s)e^{sx/c}$$
$$U(x,s) \propto R(s)e^{-sx/c} - L(s)e^{sx/c}$$

- Air pressure drops to nearly zero at the lips, causing a reflection inside the vocal tract:

$$L(s) = -R(s)e^{-sd_l/c}$$

- Volume velocity drops to nearly zero at the glottis, causing a reflection inside the vocal tract:

$$L(s) = R(s)e^{-sd_g/c}$$