Lecture 9: Exam 1 Review

Mark Hasegawa-Johnson

ECE 537: Speech Processing Fundamentals

Outline

Exam 1: Administrative Details

- In class, Wednesday; if you need conflict exam or on-line exam, contact me in advance
- One page handwritten notes, both sides
- No calculator

Content

- Loudness: Intensity, Loudness Level, Masking
- Vocoder: Voiced, Unvoiced, Spectral shape
- Pitch: Autocorrelation, Narrowband signals
- Nasals: Laplace Transform, Plane Waves, Susceptance

Outline

Solutions to the Wave Equation

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

The solution to the 1d wave equation is any combination of a rightward-traveling wave, $r(t)$ and a leftward-traveling wave, $l(t)$:

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

$$v(x,t) = \frac{1}{\rho c}\left(r\left(t - \frac{x}{c}\right) - l\left(t + \frac{x}{c}\right)\right)$$

Acoustic Intensity of a Pure Tone

Suppose that $p(t)$ is a pure tone, with a root-mean-squared (RMS) amplitude of P Pascals, and a frequency of f Hertz.

$$p(t) = \sqrt{2}P\cos(2\pi ft)$$
$$v(t) = \frac{\sqrt{2}P}{\rho c}\cos(2\pi ft)$$

The intensity of this wave is:

$$J = \langle pv \rangle = f\int_0^{1/f} p(t)v(t)dt$$
$$= f\int_0^{1/f} \frac{2P^2}{\rho c}\cos^2(2\pi ft)\,dt \qquad = \frac{P^2}{\rho c}$$

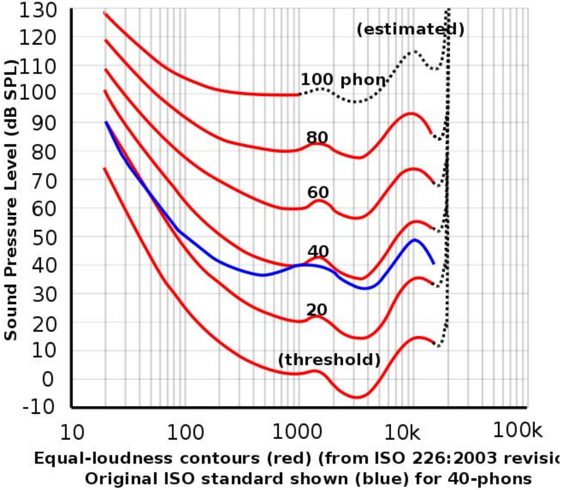
Sound Pressure Level

The intensity level of a sound can be measured with respect to a standard reference level. The standard reference level is $J_r = 10^{-12}$ Watts per square meter.
The level of a sound, measured w.r.t. 10^{-12} W/m^2, is called its "sound pressure level" (SPL). So

$$\beta = 10 \log_{10}\left(\frac{J}{J_r}\right)$$

... has units of "dB SPL."

Loudness Level



Loudness

$$G(L) = \sum_k b_k G(L_k)$$

$G(L_k)$ is a nonlinear function of the loudness level, L_k. The exam will give you a table of these values.

If you want to find the loudness level, L, of the whole sound, you can use

$$L = G^{-1}\left(\sum_k b_k G(L_k)\right)$$

Masking

- If $\Delta f = |f_2 - f_1| < B$, then just add the intensities of the two tones, and calculate loudness from that.
  ($B \in \{100, 200, 400, 800\}$, depending on f_2).
- If $\Delta f \geq B$, then

$$b_2 = \left[\frac{250 + \Delta f}{1000}\right] Q(L_2)$$

  where $Q(L_2)$ is a nonlinear function of L_2. The exam will give you a table of its values.

Outline

Voiced Source: Impulse Train

$$x[n] = \sum_{m=-\infty}^{\infty} \delta[n - mN]$$
$$= \frac{1}{N} \sum_{k=0}^{N-1} e^{j\frac{2\pi kn}{N}}$$

Spectrum of a Bandpass-Filtered Impulse Train

Suppose $x[n]$ is periodic:

$$x[n] = \sum_{k=0}^{N-1} X_k e^{j\frac{2\pi kn}{N}},$$

and we bandpass filter it with a filter $h[n]$:

$$y[n] = h[n] * x[n],$$

then $y[n]$ is periodic with Fourier series coefficients given by:

$$Y_k = H\left(\frac{2\pi k}{N}\right) X_k$$

Unvoiced Source: White Noise

The autocorrelation of a wide-sense stationary signal is:

$$R_{xx}[m] = E\left[x[n]x[n+m]\right]$$

Its power spectrum is:

$$R_{xx}(\omega) = E\left[\frac{1}{N}\left|X(\omega)\right|^2\right] = \mathcal{F}\left\{R_{xx}[m]\right\}$$

A unit variance white noise signal has

$$R_{xx}[m] = \delta[m]$$
$$R_{xx}(\omega) = 1$$

Spectrum of a Bandpass-Filtered Noise

$$y[n] = h[n] * x[n]$$

$$R_{yy}[n] = h[n] * h^*[-n] * R_{xx}[n]$$

$$R_{yy}(\omega) = |H(\omega)|^2 R_{xx}(\omega)$$

Outline

Correlogram

1. Pass the signal through a bank of bandpass filters:

$$x_f[n] = h_f[n] * x[n],$$

   where f denotes the center frequency, in Hertz, and we assume that the bandwidth is one auditory critical band.

2. Compute the autocorrelation in each channel:

$$\phi(f, m) = E\left[x_f[n]x_f[n+m]\right]$$

Autocorrelation of a Sinusoid

Suppose $x[n]$ is periodic, and the critical band contains only one harmonic:

$$x_f[n] = A\cos\left(\frac{2\pi kF_0}{F_s}n + \theta\right)$$

where kF_0 is within the passband of the filter centered at f. Suppose we treat the timing, n, as a random variable. Then

$$\begin{aligned}\phi(f,m) &= E_n\left[x_f[n]x_f[n+m]\right] \\ &= \frac{A^2}{2}\cos\left(\frac{2\pi kF_0}{F_s}m\right)\end{aligned}$$

...which is periodic with a period of $\frac{1}{kF_0}$, and at every multiple thereof, including the pitch period.

Autocorrelation of two sinusoids

Suppose $x[n]$ is periodic, and the critical band contains only two harmonics:

$$x_f[n] = A_k \cos\left(\frac{2\pi k F_0}{F_s} n + \theta_k\right) + A_{k+1} \cos\left(\frac{2\pi (k+1) F_0}{F_s} n + \theta_{k+1}\right)$$

where kF_0 and $(k+1)F_0$ are within the passband of the filter centered at f.

Suppose we treat the timing, n, as a random variable. Then

$$\begin{aligned} \phi(f,m) &= E_n\left[x_f[n] x_f[n+m]\right] \\ &= \frac{A_k^2}{2} \cos\left(\frac{2\pi k F_0}{F_s} m\right) + \frac{A_{k+1}^2}{2} \cos\left(\frac{2\pi (k+1) F_0}{F_s} m\right) \end{aligned}$$

. . . which is periodic at the pitch period $\frac{1}{F_0}$.

Autocorrelation of Narrowband Noise

Suppose $x[n]$ is unit-variance white noise. Then

$$\phi(f,m) = E_n\left[x_f[n]x_f[n+m]\right] = h_f[m] * h_f^*[-m] * R_{xx}[m]$$

But what is that? It turns out to be easier to solve in the frequency domain:

$$\phi(f,\omega) = \left|H_f(\omega)\right|^2 R_{xx}(\omega) = \left|H_f(\omega)\right|^2$$
$$= \begin{cases} 1 & \frac{2\pi(f-B/2)}{F_s} \le |\omega| \le \frac{2\pi(f+B/2)}{F_s} \\ 0 & \text{otherwise} \end{cases}$$

where B is the auditory filter bandwidth. This has the inverse transform of

$$\phi(f,m) = \left(\frac{B}{F_s}\right) \text{sinc}\left(\frac{\pi B}{F_s}m\right) \cos\left(\frac{2\pi f}{F_s}m\right)$$

. . . which is periodic with a period of $\frac{1}{f}$, which varies from filter to filter, and has no relationship to any overall pitch period.

Outline

The Two-Sided Laplace Transform

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

Example:

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

$$X(s) = \frac{1}{s-a}, \quad \text{for } \Re(s) > \Re(a)$$

Delay Property

If

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

then

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

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:

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

Volume Velocity

The relationship between pressure and volume velocity is:

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

The Zero-Pressure Constraint at the Lips

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

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:

$$B(x,s) = \frac{U(x,s)}{P(x,s)}$$
$$= -\frac{A(x)}{\rho c} \coth\left(s(x - d_l)/c\right)$$

The Zero-Velocity Constraint at the Lips

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

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

$$B(x,s) = \frac{U(x,s)}{P(x,s)}$$
$$= -\frac{A(x)}{\rho c}\tanh\left(s(x - d_l)/c\right)$$

Resonances of a Nasal Consonant

Fujimura proposed computing the resonances of a nasal consonant by finding the zeros of the total susceptance,

$$B(s) = B_p(s) + B_n(s) + B_m(s)$$

For the consonant /ŋ/, Fujimura assumed that the mouth cavity has zero volume, thus $B_m(s) = 0$, so resonances of /ŋ/ are the zeros of $B_i(s) = B_n(s) + B_p(s)$.
The resonances of /m/ and /n/ are then modeled by the equation

$$B_i(s) = -B_m(s)$$

Anti-resonance

The anti-resonance (the zeros of the transfer function) are the frequencies at which

$$B_m(s) = \infty$$

Outline

Conclusion: Topics for Exam

- Loudness: Intensity, Loudness Level, Masking
- Vocoder: Voiced, Unvoiced, Spectral shape
  - Not covered: Brownian motion, relaxation oscillator
- Pitch: Autocorrelation, Narrowband signals
  - Not covered: Gammatone filters
- Nasals: Laplace Transform, Plane Waves, Susceptance