Linear Shift Invariant Systems

- A system maps one or more input signals onto one or more output signals. For a single input / single output system, we have

$$y(k) = T\{x(k)\}$$

where $T\{.\}$ defines the mapping (or a transfer function)

- Linear System: A system is linear if and only if for any two signals $x_1(k)$ and $x_2(k)$ and any $a, b \in \mathbb{C}$ we have

$$T\{a x_1(k) + b x_2(k)\} = a T\{x_1(k)\} + b T\{x_2(k)\}.$$

Q: what does it mean? Explain by some physical examples.

- Memoryless system: A system is memoryless when $y(k=k_0)$ depends only on $x(k=k_0)$ for any k_0.

- Shift invariant system: A system is shift invariant if for any k_0 $\quad y(k-k_0) = T\{x(k-k_0)\}$

Q: what does it mean? Explain.

- For a linear and shift invariant (LSI) system, the output is governed by a convolution relation.

$$y(k) = x(k) * h(k)$$
$$= \sum_{l=-\alpha}^{+\alpha} x(l)\, h(k-l) = \sum_{l=-\alpha}^{+\alpha} h(l)\, x(k-l) = h(k) * x(k)$$

where h(k) is the Impulse Response of the LSI system.

- Useful system (needs memory) – current output depends on the current and past inputs as well as past outputs.

$$y(k) = b_0\, x(k) + b_1 x(k-1) + \cdots + b_N\, x(k-N)$$
$$+ (-a_1)\, y(k-1) + (-a_2)\, y(k-2) + \cdots + (-a_M)\, y(k-M)$$

Compactly written: $\displaystyle\sum_{m=0}^{M} a_0\, y(k-n) = \sum_{n=0}^{N} b_0\, x(k-n)$

where $a_0 = 1$.

N and M are the order of system.

Frequency Response of LSI systems

Let $x(k) = e^{j(\Omega_0 k + \phi_0)}$. Then $x(k)$ is an eigenfunction of an LSI system.

$$y(k) = x(k) * h(k) = \sum_{l=-\alpha}^{+\alpha} h(l)\, x(k-l)$$
$$= \sum_{l=-\alpha}^{+\alpha} h(l)\, e^{j(\Omega_0(k-l)+\phi_0)}$$
$$= e^{j(\Omega_0 k+\phi_0)} \sum_{l=-\alpha}^{+\alpha} h(l)\, e^{-j\Omega_0 l}$$
$$= e^{j(\Omega_0 k+\phi_0)}\, H(e^{j\Omega_0})$$

$H(e^{j\Omega})$ is the frequency response and

$$H(e^{j\Omega_0}) = H(e^{j\Omega})\Big|_{\Omega=\Omega_0} = \sum_{k=-\alpha}^{+\alpha} h(k)\, e^{-j\Omega k}\Big|_{\Omega=\Omega_0}.$$

Z- transform

Complex variable $Z = e^{sT} = e^{(\alpha + j\omega)T}$ $\quad [s = \alpha + j\omega]$

$$= e^{\alpha T} \cdot e^{j\omega T}$$

$$= r e^{j\Omega} \quad \begin{bmatrix} r = e^{\alpha T} \\ \omega T = \Omega \end{bmatrix}$$

$$Z\{x(k)\} = X(z) = \sum_{k=-\alpha}^{+\alpha} x(k) z^{-k}.$$

ROC (Region of convergence): Z- transform always comes with ROC.

Z-transform converges when $\sum_{k=-\alpha}^{+\alpha} |x(k) r^{-k}| < \infty$.

Explain: what is the role of 'r' and its bearing on ROC.

$$X(k) = \sum_{k=-\alpha}^{+\alpha} x(k) z^{-k} = \sum_{k=-\alpha}^{+\alpha} x(k) (r e^{j\Omega})^{-k}$$

$$= \sum_{k=-\alpha}^{+\alpha} x(k) r^{-k} e^{-j\Omega k}$$

Note $e^{-j\Omega k}$ is always bounded. But r^{-k} can be unbounded or very high depending on r and k. So, for convergence we must need

$$\sum_{k=-\alpha}^{+\alpha} |x(k) r^{-k}| < \alpha.$$

Properties of Z- transform

Property	Time-domain	Z-domain	ROC
Linearity	$a_1 x_1(k) + b x_2(k)$	$a X_1(z) + b X_2(z)$	
Conjugation	$x^*(k)$	$X^*(z^*)$	
Time reversal	$x(-k)$	$X(z^{-1})$	
convolution	$x_1(k) * x_2(k)$	$X_1(z) X_2(z)$	
Time shift	$x(k+k_0)$	$z^{k_0} X(z)$	

See Table 3.3

Relation to DTFT

for $r = 1$, we have unit circle

$$X(z)\Big|_{r=1} = \sum_{k=-\infty}^{+\infty} x(k)\, z^{-k}$$

$$= \sum_{k=-\infty}^{+\infty} x(k)\, r^{-k}\, e^{-j\Omega k}$$

$$= \sum_{k=-\infty}^{+\infty} x(k)\, e^{-j\Omega k}$$

$$= X(e^{j\Omega})$$

One-sided Z transform

A causal signal $x(k)$ can not have $k < 0$. Basically $x(k) = 0$ if $k < 0$.

$$X(z) = \sum_{k=0}^{\infty} x(k)\, z^{-k} .$$

Z-transform and LSI systems

$$x(k) \longrightarrow \boxed{h(k)} \longrightarrow y(k) = x(k) * h(k) .$$

$$X(z) \longrightarrow \boxed{H(z)} \longrightarrow Y(z) = X(z)\, H(z)$$

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

$Y(z)$ and $X(z)$ are polynomial in z.

$$Y(z) = \sum_{n=0}^{N} b_n z^{-n} \;;\quad X(z) = \sum_{m=0}^{M} a_m z^{-m}$$

$$= \prod_{n=1}^{N} (1 - \alpha_n z^{-1}) \quad\Big|\quad = \sum_{m=0}^{M} (1 - \beta_m z^{-1})$$

This provides poles and zeros.

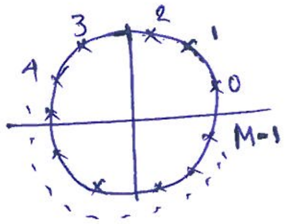
Discrete Fourier Transform (DFT)

- DTFT is continuous. Not suitable for a digital computer.
- We need a sampled version.
- DFT is the samples of DTFT.

Let $x(k)$ is an M-length signal $\quad x(k) = 0$ if $k < 0$ and $k \geq M$.

$$\text{DTFT} \rightarrow X(e^{j\Omega}) = \sum_{k=0}^{M-1} x(k)\, e^{-j\Omega k}$$

$$\Omega = \frac{2\pi m}{M}$$



$$\text{DFT} \rightarrow X_m = X(e^{j\Omega})\Big|_{\Omega = \frac{2\pi m}{M}}$$

$$= \sum_{k=0}^{M-1} x(k)\, e^{-j\frac{2\pi m}{M} \cdot k} \quad , \quad m = 0, 1, \cdots M-1$$

$$\text{IDFT} \rightarrow x(k) = \frac{1}{M} \sum_{m=0}^{M-1} X_m\, e^{j\frac{2\pi m}{M} k} \quad , \quad k = 0, 1, \cdots, M-1$$

- Can we give an MATLAB example. A sinusoidal signal.

DFT of Windowed Sequence

- Let a window $W(K)$ is multiplied with an input signal $x(k)$.

$$y(k) = w(k)\, x(k)$$

- $$Y(e^{j\Omega}) = W(e^{j\Omega}) * X(e^{j\Omega}) \longrightarrow \text{leads to spread.}$$

- DFT $Y_m = Y(e^{j\Omega}) \Big|_{\Omega = \frac{2\pi m}{M}}$

- So, we get sequence of sampled points from a spreaded frequency response.

Spectral Resolution and Zero-padding

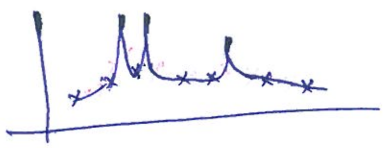
- $x(k)$ is a finite M-length sequence.

- we can create a new sequence $x'(k)$ by zero-padding $L-M$ zeros where $L > M$.

- $$x'(k) = \underbrace{[\overbrace{x(k)}^{M}\ 0\ 0 \cdots 0]}_{L}$$

- Now if we take an L-length DFT, so we sample more on the continuous DTFT.

- This may give a better spectral resolution where two close spaced sinusoidal signals may be seperately viewed.



Anatomy of Hearing

(7)

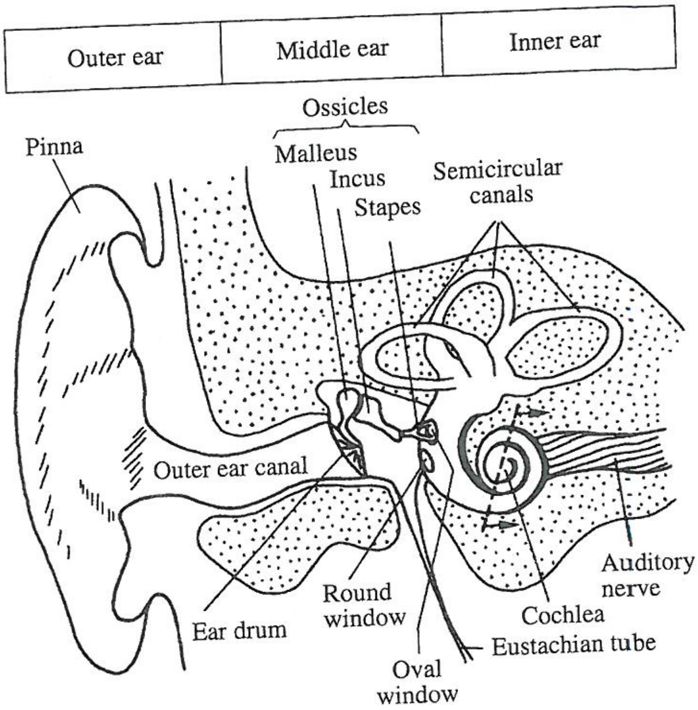


Figure 2.17: Schematic drawing of the ear ([Zwicker, Fastl 1999])

The principal hearing organ is divided into three sections

- outer ear
- middle ear, and
- inner ear

outer ear canal is nearly 3 cm long. It has a ~~...~~ resonance frequency between 3 to 4 KHz (within the frequency of range). That is why the ear is more sensitive at this frequency range.

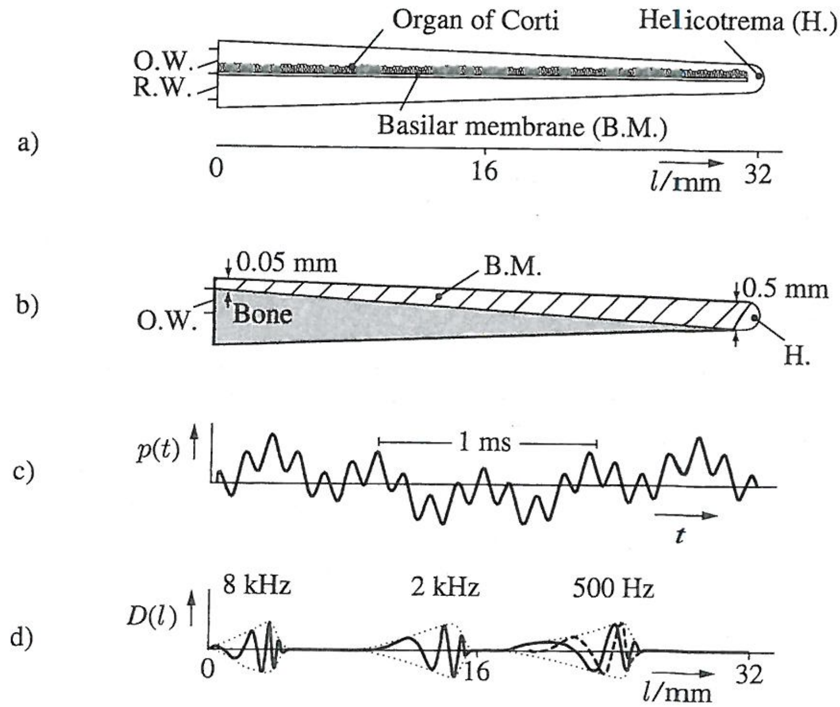


Figure 2.18: Frequency-to-place transformation along the basilar membrane (adapted from [Zwicker, Fastl 1999]); O.W. = Oval Window; R.W. = Round Window
a) Top-down view into the unwound cochlea
b) Side view
c) Tripple-tone audio signal (sound pressure, 500 Hz, 2000 Hz, 8000 Hz)
d) Displacement $D(l)$ by excitation (traveling waves and envelopes)

9

Psychoacoustic properties of the Auditory Organ

Hearing and Loudness

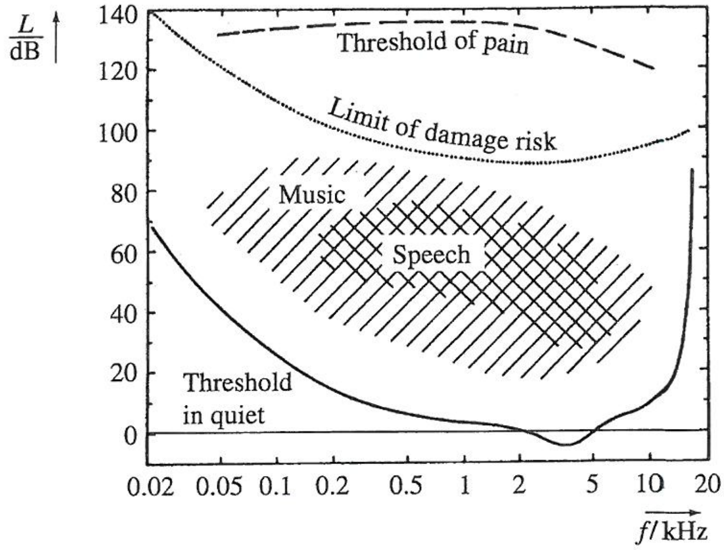


Figure 2.19: The hearing area ([Zwicker, Fastl 1999])

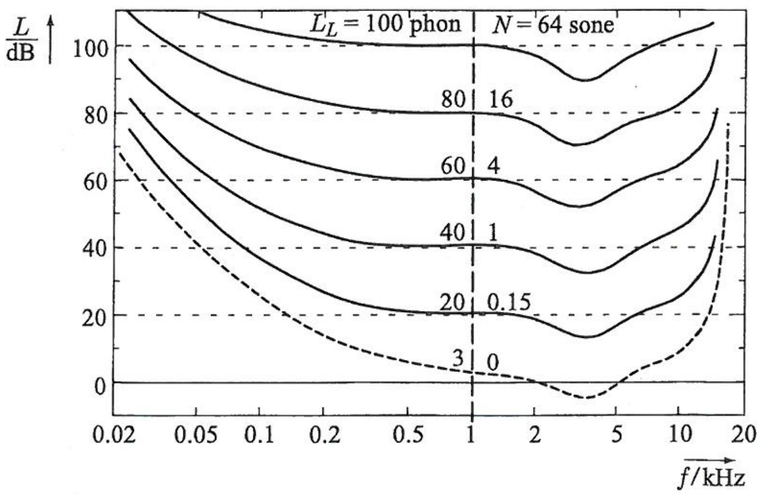


Figure 2.20: Contours of equal loudness level L_L for sinusoidal tones ([Zwicker, Fastl 1999])

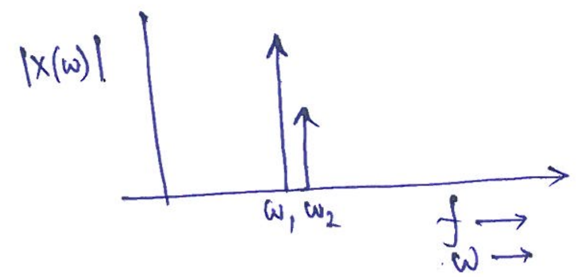
Masking

- mainly used in Audio Coding (MP3 songs)

- Masking occurs in many situations, where a dominant sound renders a weaker sound inaudible. One example is music at "disco level" which might completely mask the weak ringing of a mobile phone. In that case, the high level music signal is masker and the ringing tone is a test tone.

- Example:

$$x(k) = a_1 \sin(\omega_1 k + \phi_1) + b \sin(\omega_2 k + \phi_2)$$



For perfect signal knowledge, we need to code $\{a_1, \omega_1, \phi_1, b, \omega_2, \phi_2\}$

But if the first tone of ω_1 frequency is much higher the it masks the weaker tone of ω_2 frequency.

So, we should only code $\{a_1, \omega_1, \phi_1\}$ for Audio Coding (MP3)