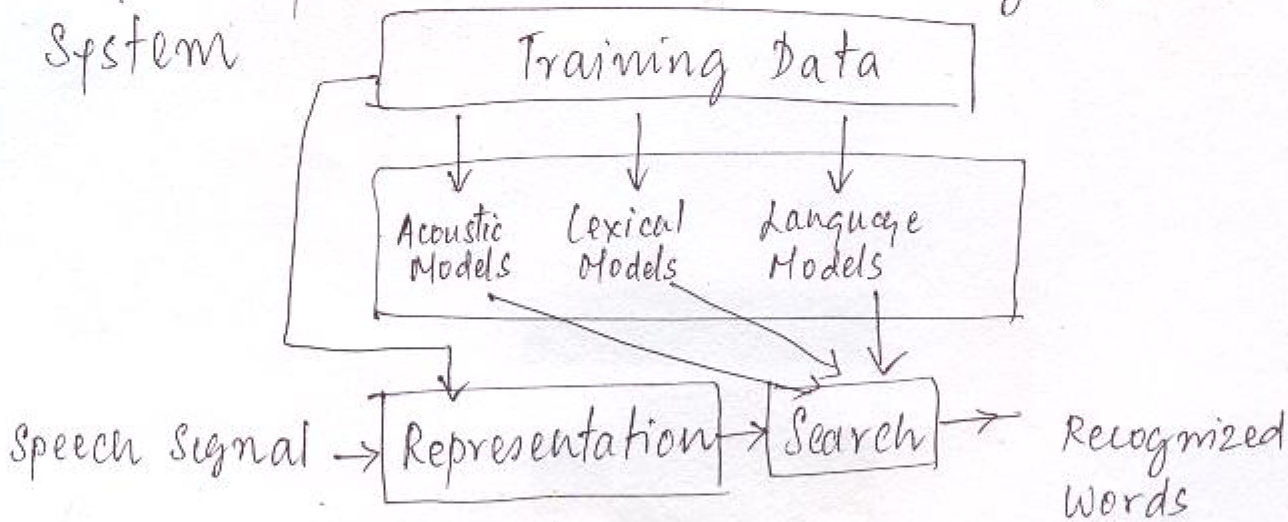
Major Components in a Speech Recognition System – J. Glass



1. Acoustic theory of Speech Production
2. Properties of Speech Sounds
3. Representation
4. Pattern Recognition
5. VQ / GMM
6. HMM

---

* Use a Microphone

* Record 'Mack' " "

* label in matlab, sfs.

I. Signal Processing Components in Speech Processing.

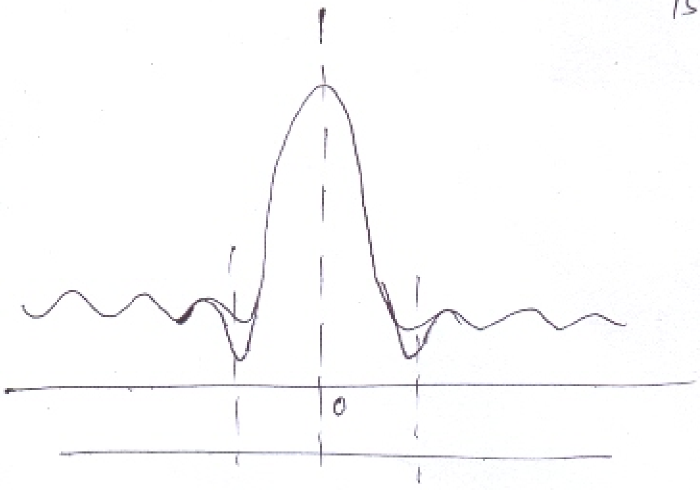
Speech
↓
Spectral Shaping
↓
Spectral Analysis
↓
Spectral Modeling
↓
Parametric transform
↓
Observation vectors.

II. If Highest freq. component in an analog signal $x_a(t)$ is $f_{max} = B$, then $F_s > 2B$ - only then can it be recovered using

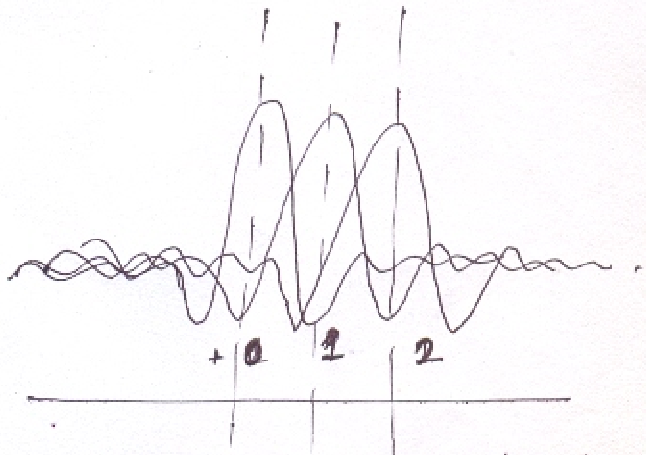
$$x_a(t) = \sum_{\forall n} x_a(nT) \frac{\sin(\pi/T)(t-nT)}{(\pi/T)(t-nT)}$$

At $t = nT$; $x_a(t) = \sum_{\forall n} x_a(nT)$ which is original analog signal.

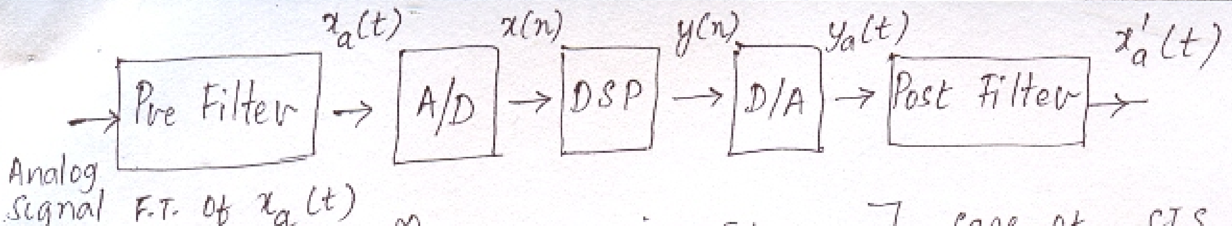
otherwise; $x_a(t) = \sum_{\forall n} x_a(nT) \, \mathrm{sinc}((\pi/T)(t-nT))$ which is weighted sum of sinc functions



Sinc function



Set of shifted sinc functions.



F.T. of $x_a(t)$

CTS

$$X_a(F) = \int_{-\infty}^{\infty} x_a(t)\, e^{-j2\pi Ft}\, dt$$

case of CTS

$x_a(t)$ can be recovered as

$$x_a(t) = \int_{-\infty}^{\infty} X_a(F)\, e^{j2\pi Ft}\, dF \quad \text{—} (1)$$

DTS

$$X_a(\omega) = \sum_{\forall n} x(n)\, e^{-j\omega n} \quad - (2)$$

or

$$X(f) = \sum_{\forall n} x(n)\, e^{-j2\pi fn} \quad - (3)$$

$x(n)$ can be recovered as

$$x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)\, e^{j\omega n}\, d\omega \quad - (4)$$

→ changing variable ω to f we get

$$x(n) = \int_{-1/2}^{1/2} X(f)\, e^{j2\pi fn}\, df \quad - (5)$$

Consider (1) and substitute $t = nT = \frac{n}{F_s}$

$$\therefore\ x_a(nT) \equiv x(n) = \int_{-\infty}^{\infty} X_a(F)\, e^{j2\pi n(F/F_s)}\, dF \quad - (6)$$

But (5) = (6);

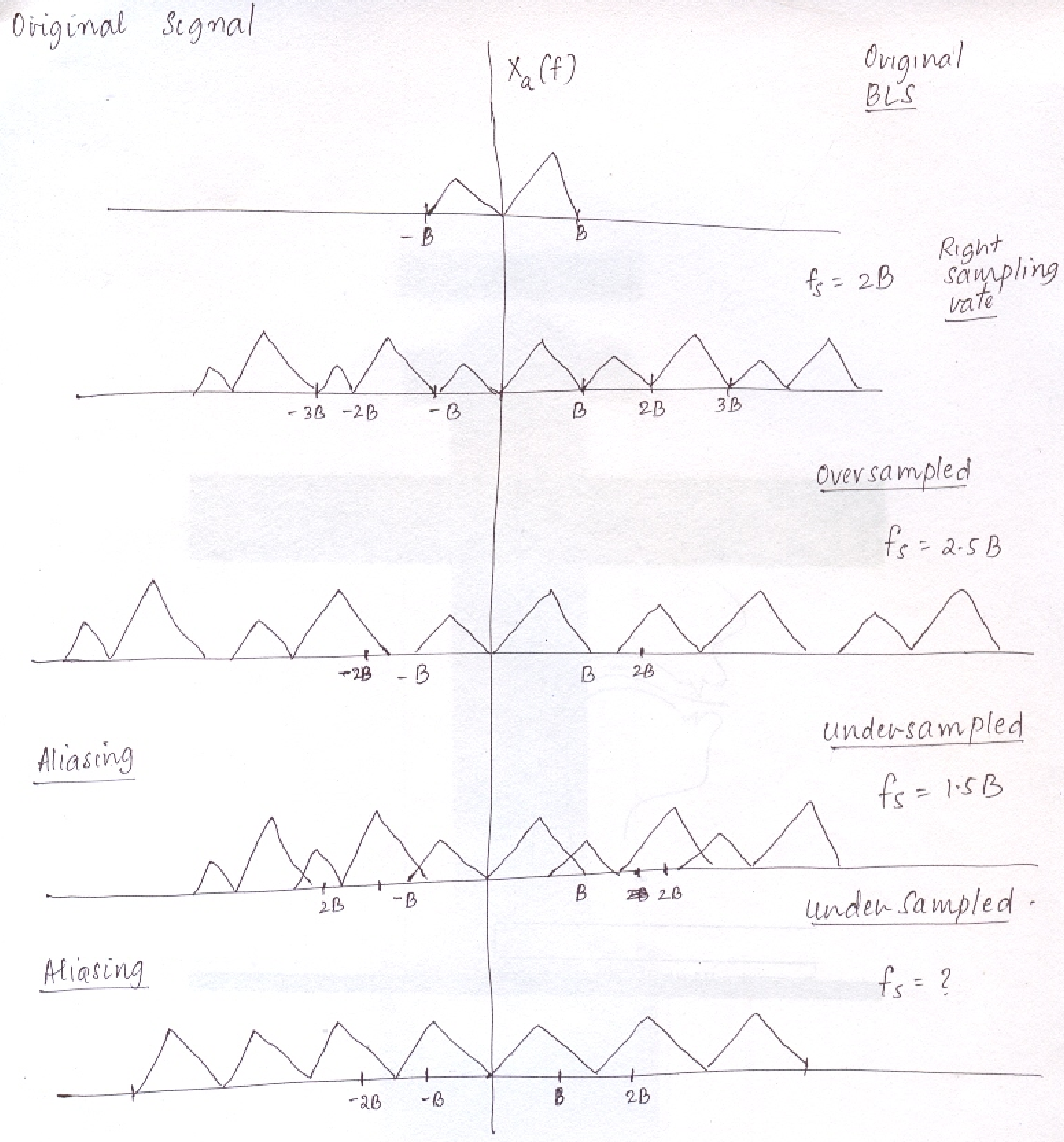
$$\therefore \int_{-1/2}^{1/2} X(f)\, e^{j2\pi fn} = \int_{-\infty}^{\infty} X_a(F)\, e^{j2\pi n(F/F_s)} \quad - (7)$$

But We know $\boxed{f = F/F_s}$

Substitute into LHS and simplify we get

$$\boxed{X(F/F_s) = F_s \sum_{\forall k} X_a(F - kF_s)}$$

⇒ spectrum of the sampled signal "Symmetric"?

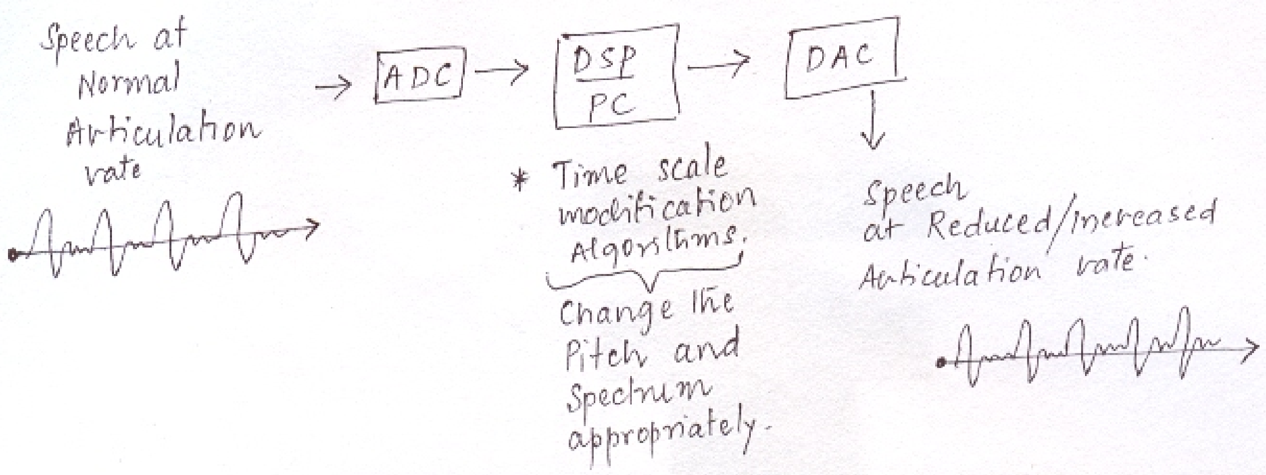


Graphical Illustration of Sampling and Reconstruction.

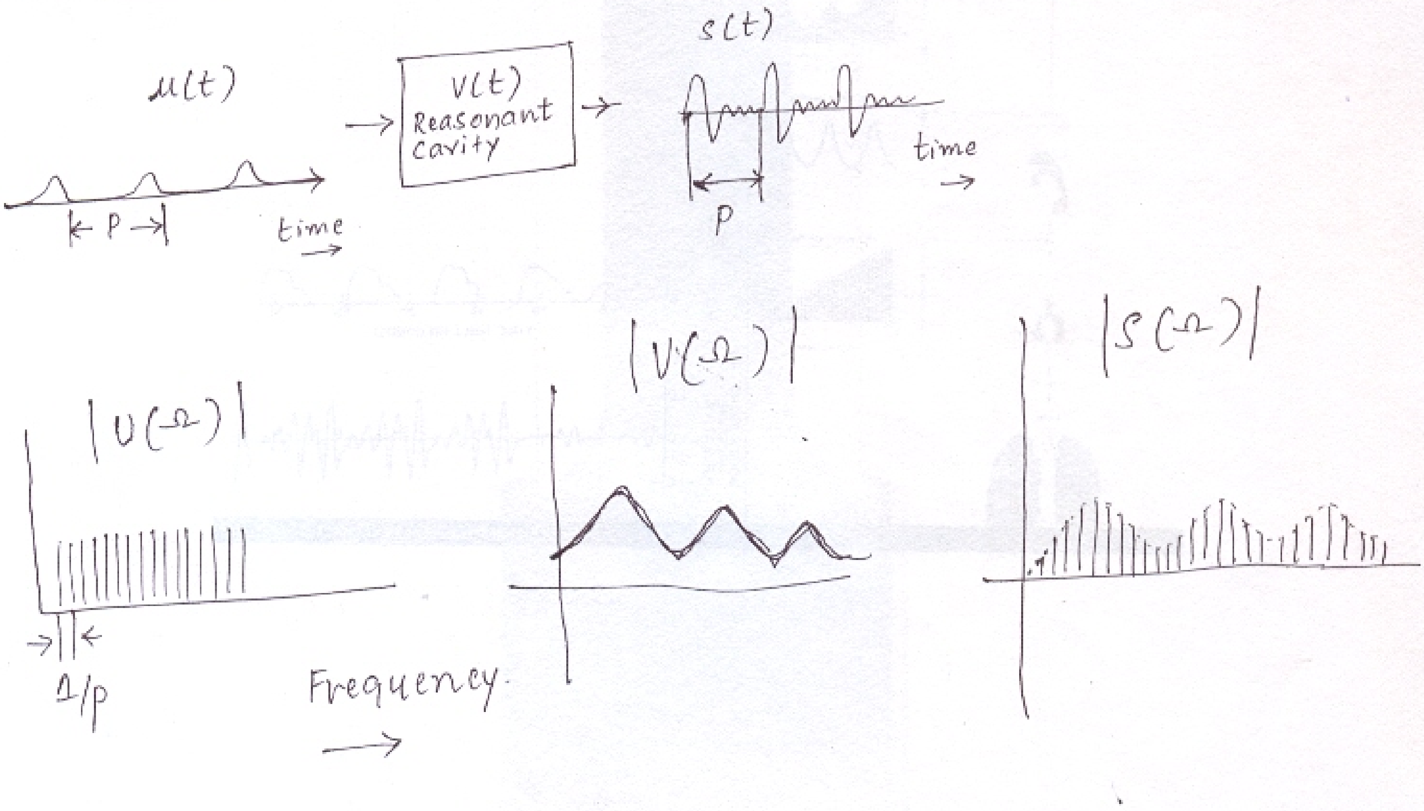
Discrete time speech signal processing

Ch. 1. Tom Quatieri.

eg: Time scale Modification



Vowel Production



Speech Signal Processing Overview :