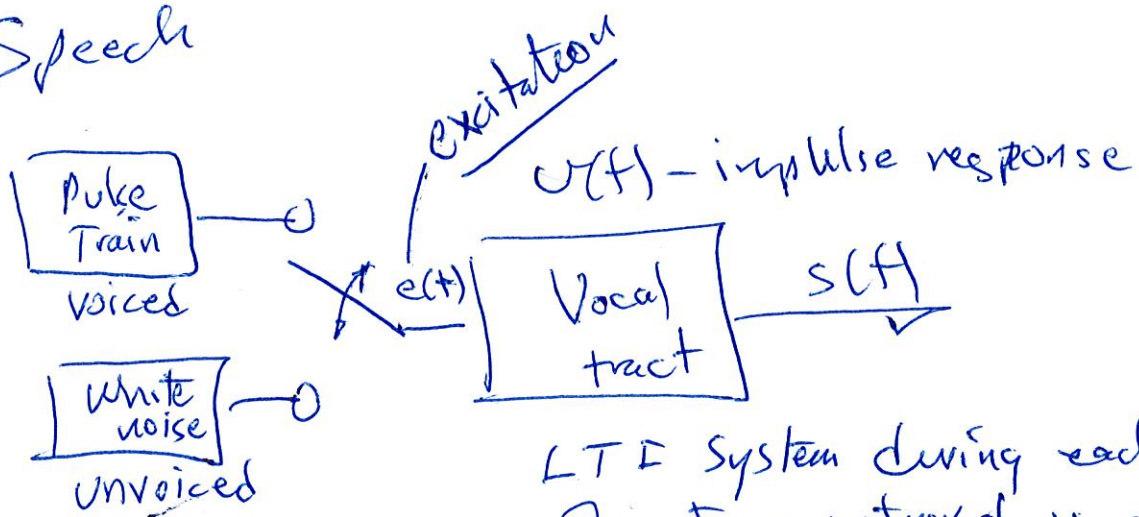
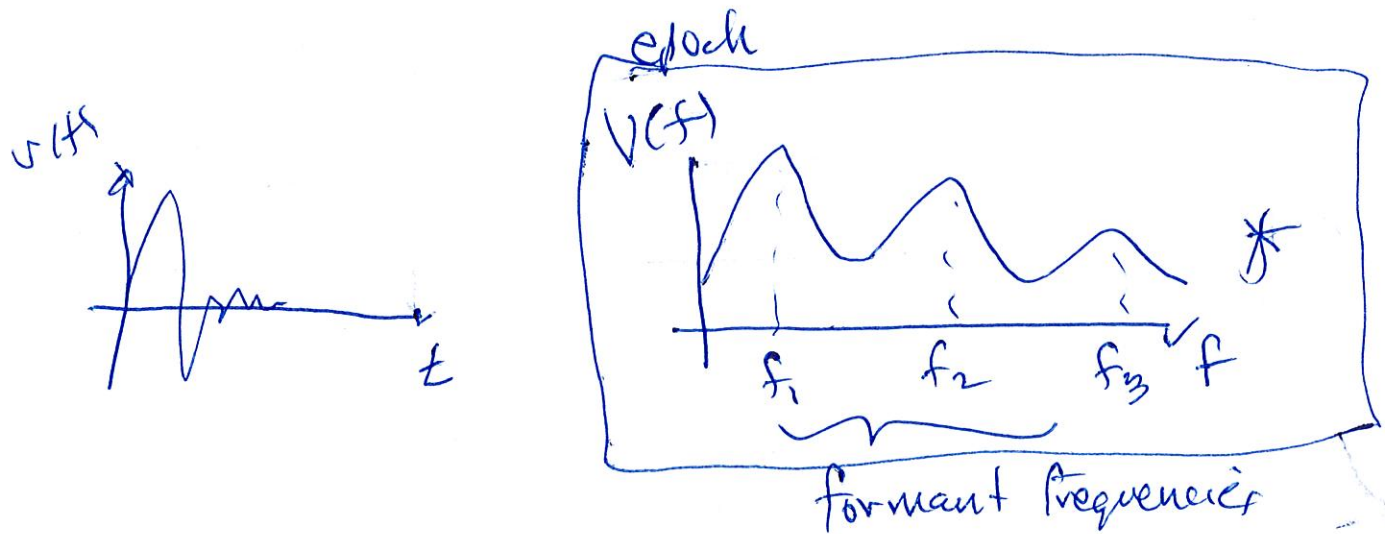


Help Session TODAY 4:00p - 6:00p ECE 222

Speech



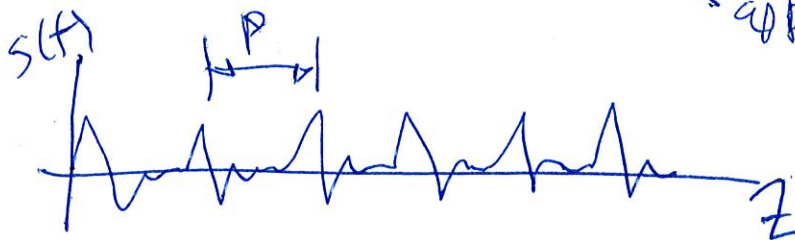
LTE System during each phoneme.
This time interval is called an epoch



$$e(t) = \sum_k s(t - kP)$$

P - pitch period sec.

approximate model



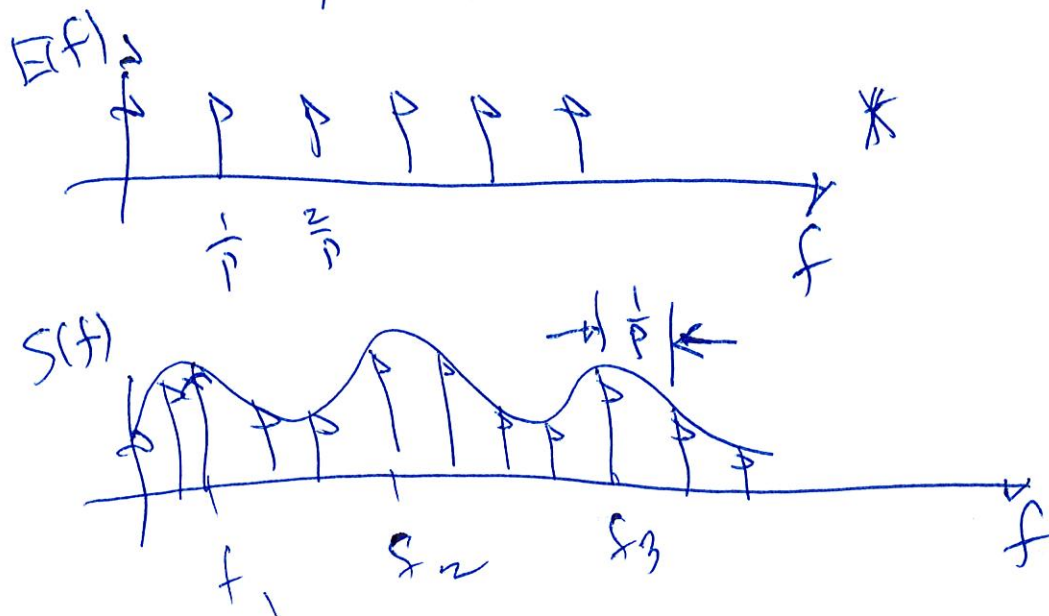
$$s(t) = v(t) * e(t)$$

$$S(f) = V(f) E(f) *$$

$$e(t) = \text{comb}_P[1]$$

(2)

$$E(f) = \frac{1}{P} \text{rep} \frac{1}{P} [S(f)]$$



STDTFT

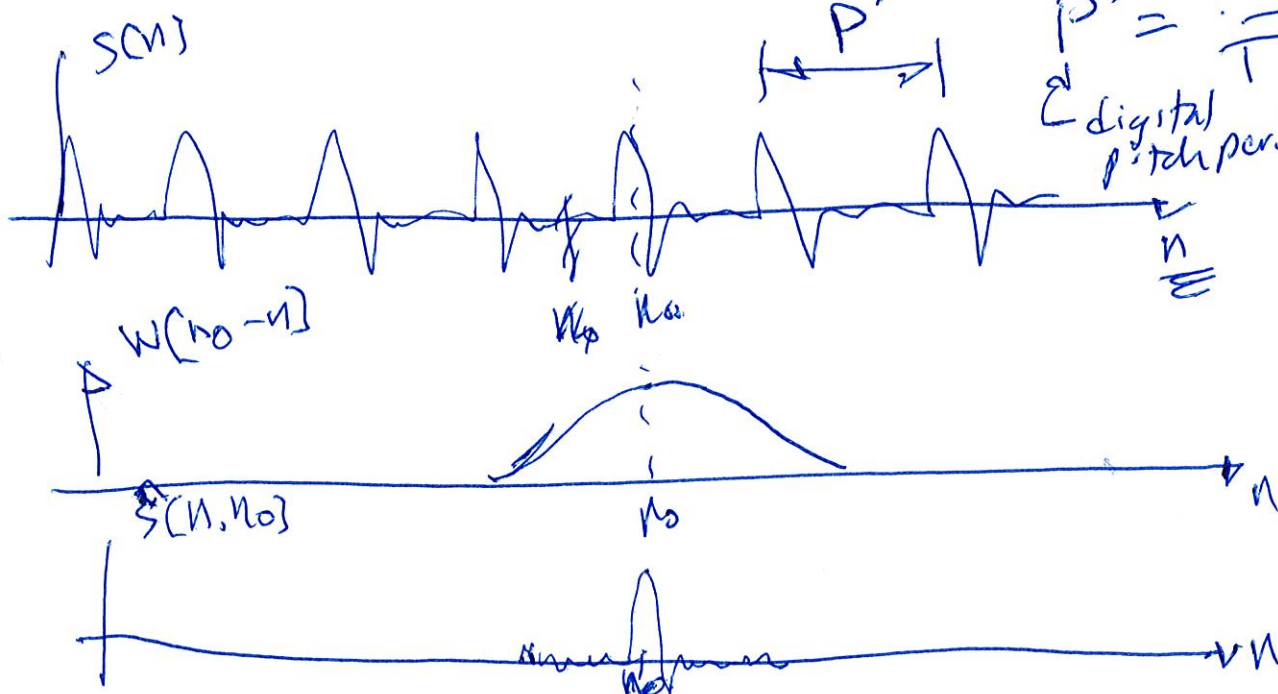
$$\hat{S}(\omega, n_0) = \sum_n \hat{S}[n, n_0] e^{-j\omega n}$$

$$\begin{aligned} \hat{S}[n, n_0] &= s[n] w[n_0 - (n - n_0)] \\ &= s[n] w[n_0 - n] \end{aligned}$$

analog pitch period

$$P' = \frac{P \text{ sec.}}{T \text{ sec. samples}}$$

digital pitch period

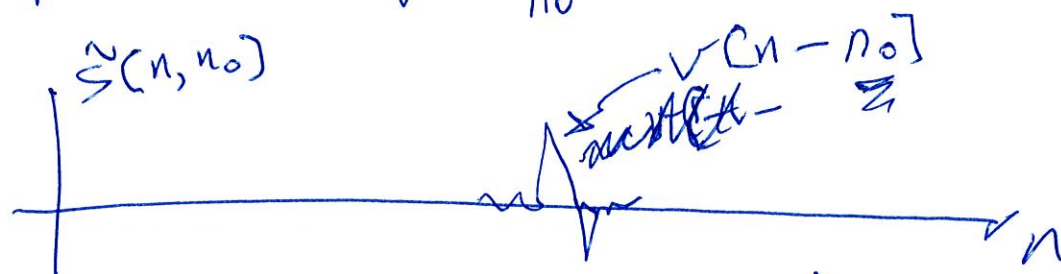
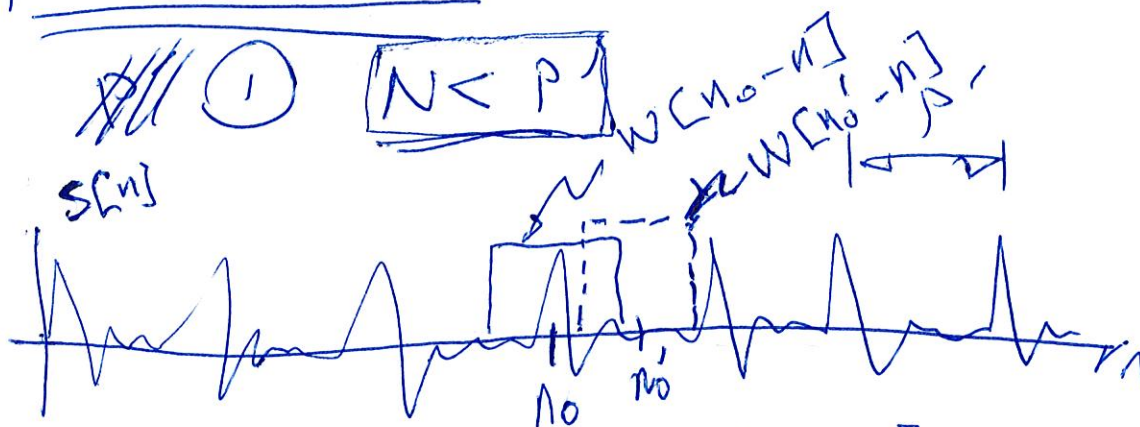


For convenience, we assume

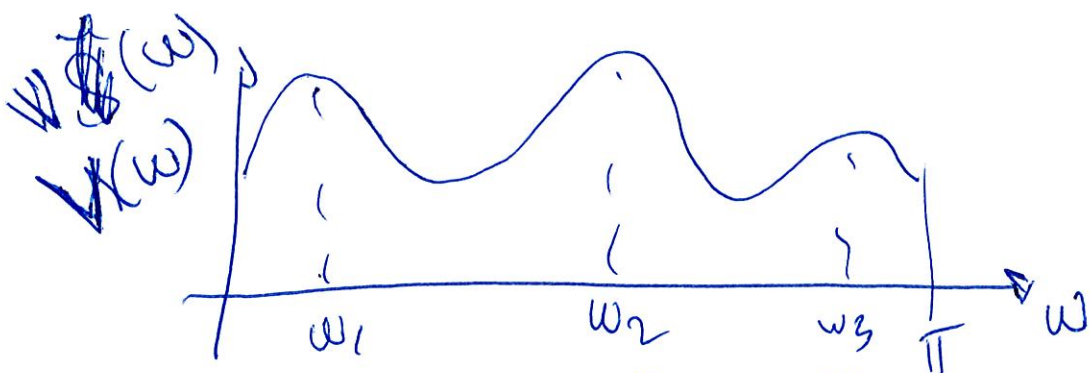
$$W[n] = \begin{cases} 1, & |n| < \frac{N-1}{2} \\ 0, & \text{else} \end{cases}$$

(3) $\frac{N-1}{2}$ window length

TWO CASES:



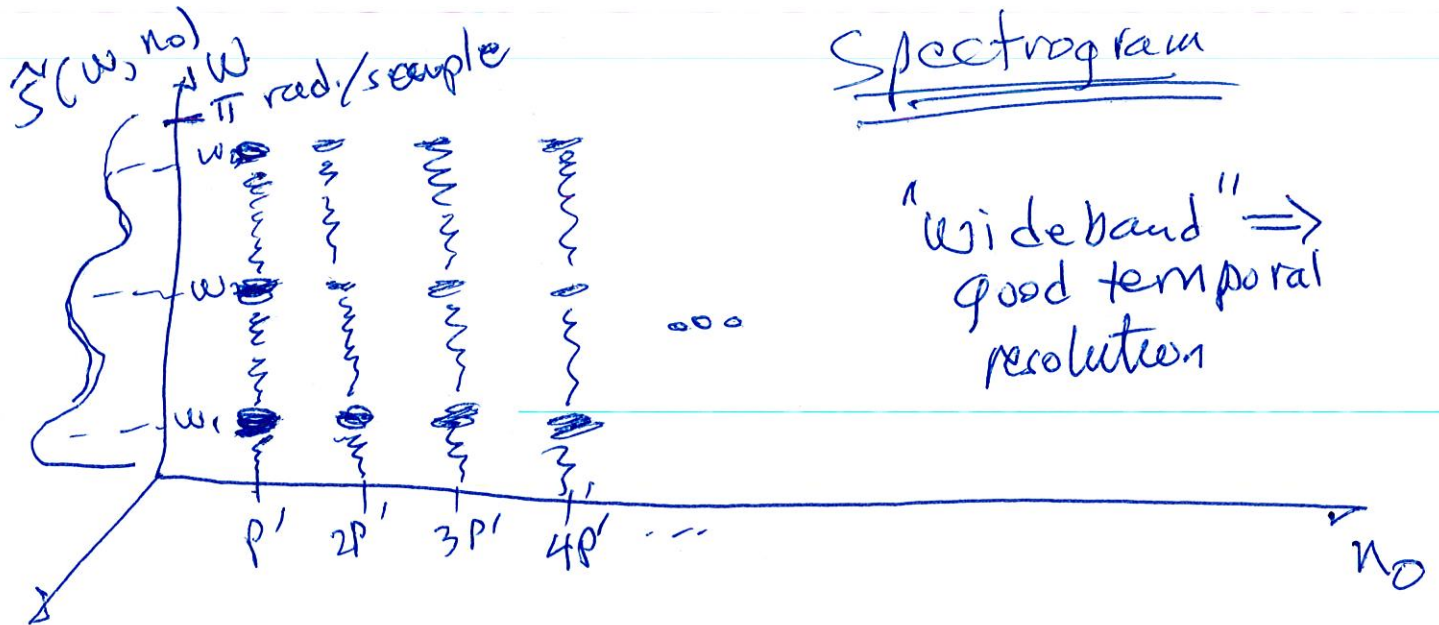
$$\tilde{S}(\omega, n_0) = \sum_n \tilde{s}[n, n_0] e^{-j\omega n}$$



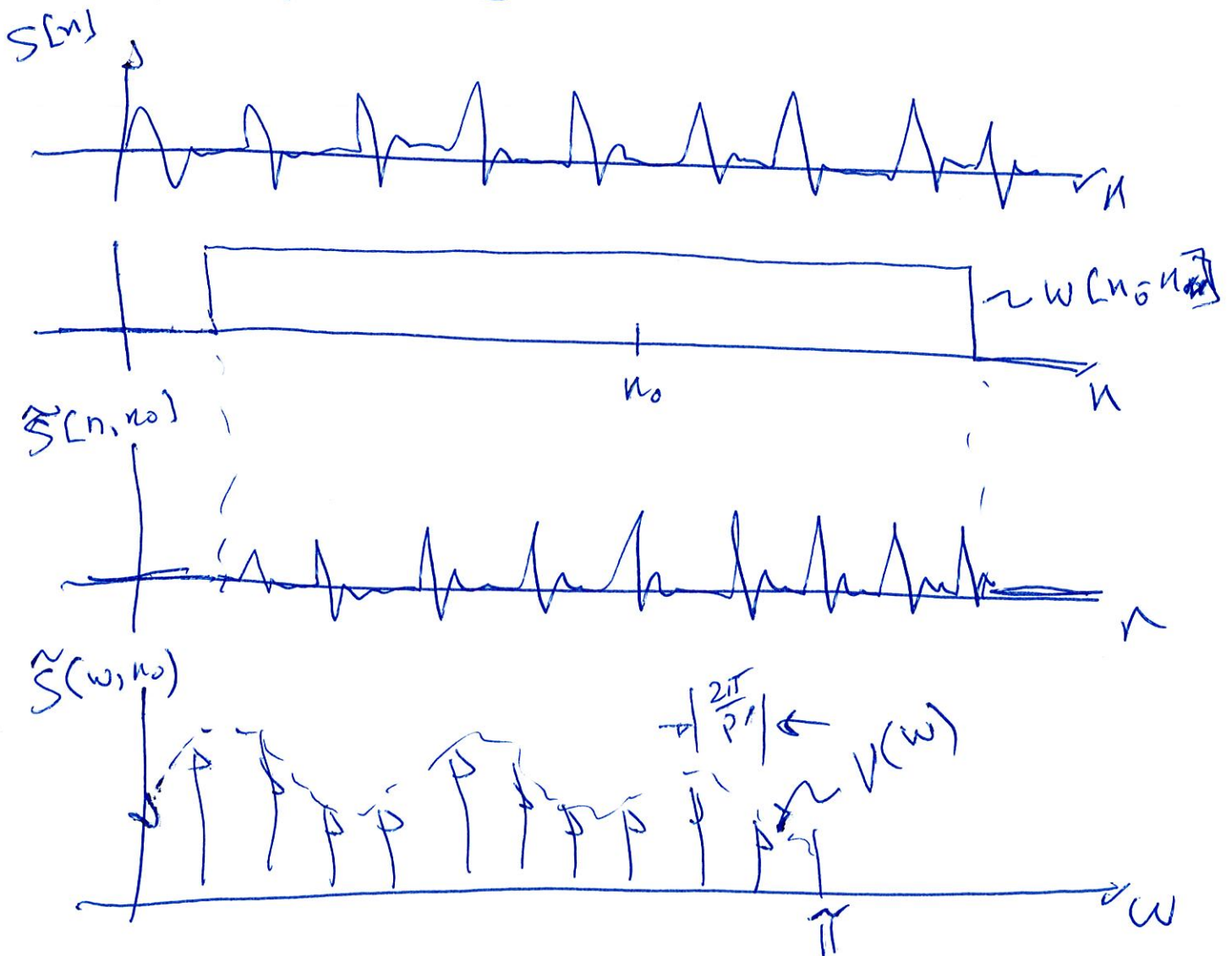
Time delay of $x(t)$ causes multiplication by $e^{-j\omega n_0}$ in frequency domain. We will ignore this.

At n_0' , $\tilde{S}(\omega, n_0') \approx 0$

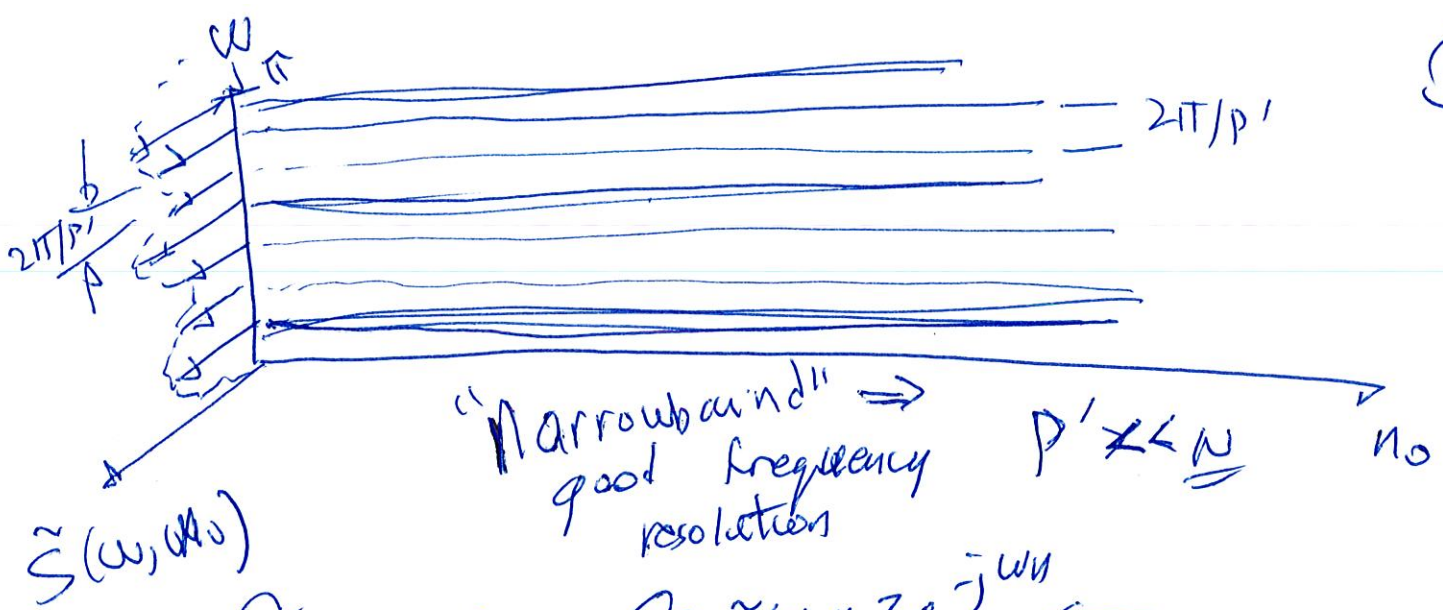
(4)



(2) $N \gg p'$



5



STDTFT $\hat{S}(w, n_0) = \sum_n \underline{\hat{S}[n, n_0]} e^{-jwn}$ $\leftarrow$

A different interpretation of the STDTFT

recall $\hat{S}[n, n_0] = W[n_0 - n] s[n]$

$\Rightarrow \hat{S}(w, n_0) = \sum_{n=-\infty}^{\infty} W[n_0 - n] s[n] e^{-jwn}$

convolution !!!

Use commutative of convolution

$\hat{S}(w, n_0) = \sum_n s[n_0 - n] e^{-jw(n_0 - n)} W[n]$

6

1/10

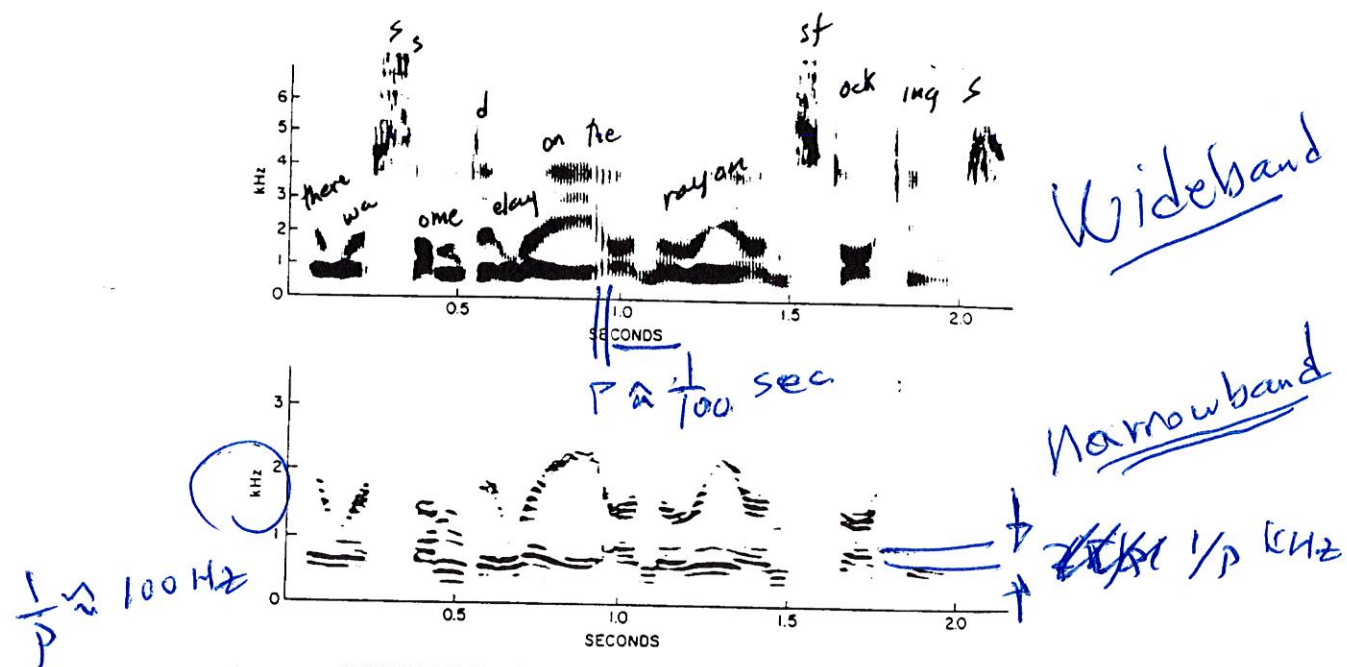


FIGURE 3-4. Speech spectrograms of the utterance "there was some delay on the rayon stockings"; (a) wideband spectrogram; (b) narrowband spectrogram.