University of California
Berkeley

College of Engineering
Department of Electrical Engineering
and Computer Sciences

Professors : N.Morgan / B.Gold

EE225D

Spring,1999

Medium & High Rate Coding

Lecture 26

Quality, Robustness, Speaker I.D., Possible Vocoding of Other Sounds.

Summary for Chapter 33 - Medium & High Rate Systems

* Voice Excitation & Spectral Flattering
* Voice - Excited Channel Vocoder
* VELP & RELP
* Wave form Coding using predictive methods.(ADPCM)-(CVSD)
* APC
* Subband Coding

Emphasis in an getting a good efficient repesentation of the excitation.

<u>Accent on Error Signal.</u>

Representation of good Excitation is still the basic issue.

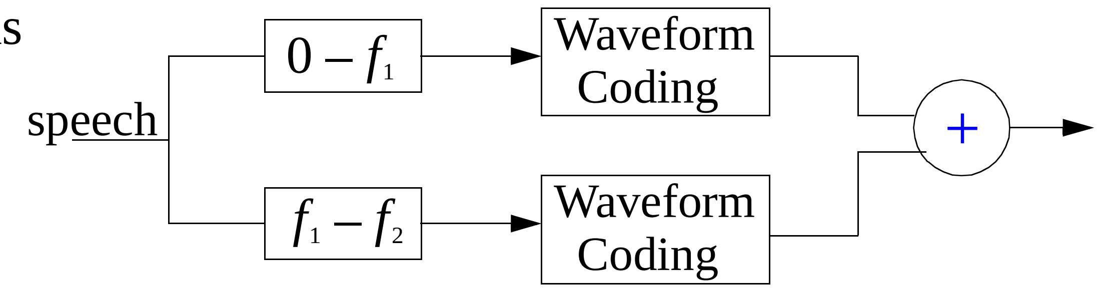
- *Multi pulse LPC
- * CELP ——————Celp is a variation on APC.
- * Reducing Codebook Search Time in CELP.
- * Back Ward Filtering
- * Multi Resolution Codebook Search
- * Partial Sequence Elimination
- * Tree Structured Delta Codebooks
- * Adaptive Codebooks
- * Linear Conbination Codebooks
- * Vector Sum Excited Linear Preduction
- * Adaptive Transform Coding

Different Structures for Wide Band & Medium Band Digital System

* Pure Waveform Coding — Subband Coding Pure Waveform Coding on a channel by channel basis.

* Pure Modelling

* Hybrid Systems



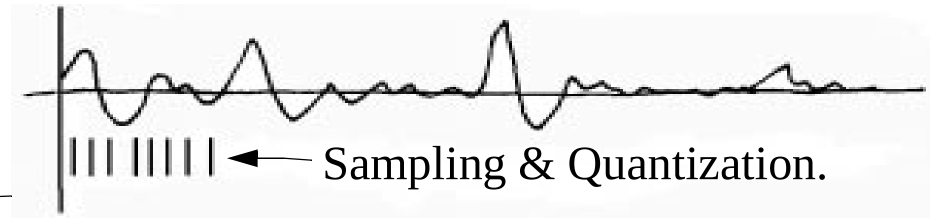
* Voice - Excited System — Excitation derived from same band limited function of the speech.

* Processing of the Error Signal to Produce a Band Limited Excitation Function.

APC
CELP — Many Variation
Multi-Pulse

Pure Waveform Coding

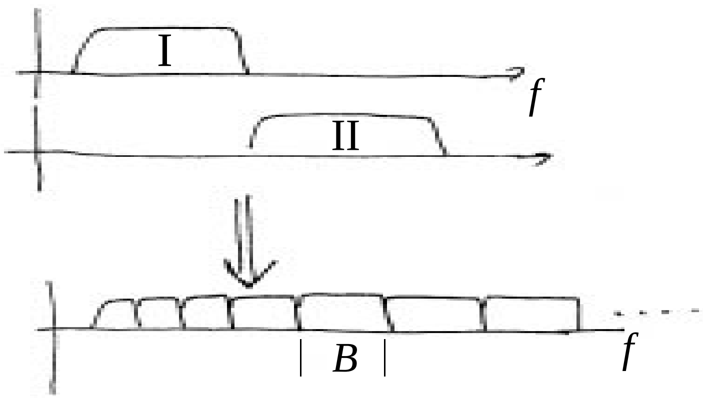
(Very Robust)



Play Molly’s tape.

Possible Viewgraphs :
Fig.33.8, 33.9, 33.10.

Local max construct ADPCM CUSD

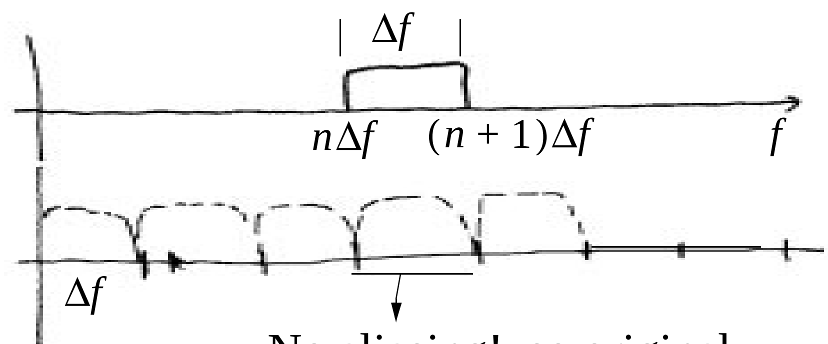


Let’s stick to pure PCM.
from a perceptual viewpoint, fewer bits are needed to encode II, compared to I.

Each band is sampled and quantized.

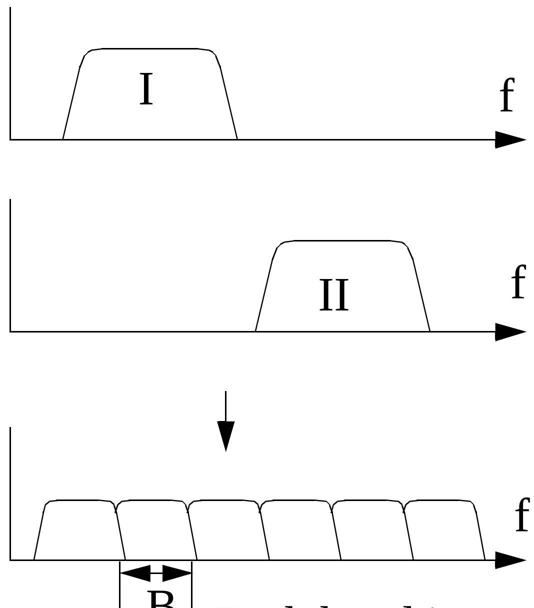
Important Point : if bandwidths are <u>judiciously clean</u>, sampling can be done at Nyquist rate.

$$\boxed{f_s = 2B}$$



No aliasing!, so original can be recovered.

Subband Coding



Let's stick to pure PCM.

From a perceptual Viewpoint, fewer bits are needed to encode II, compared to I.

Important Point: If bandwidths are judicialy clean, sampling can be done at Nyquist rate.

Each band is sampled and quantized.

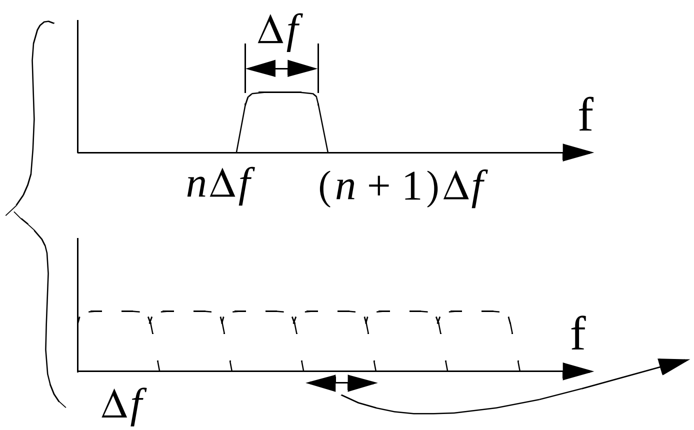


Fig.33.14

$$f_s = 2B$$

No Aliasing! So original can be recovered.

<u>Voice -Excited Channel Vocoder</u> (late 1950’s to early 1960’s)

Motivations :

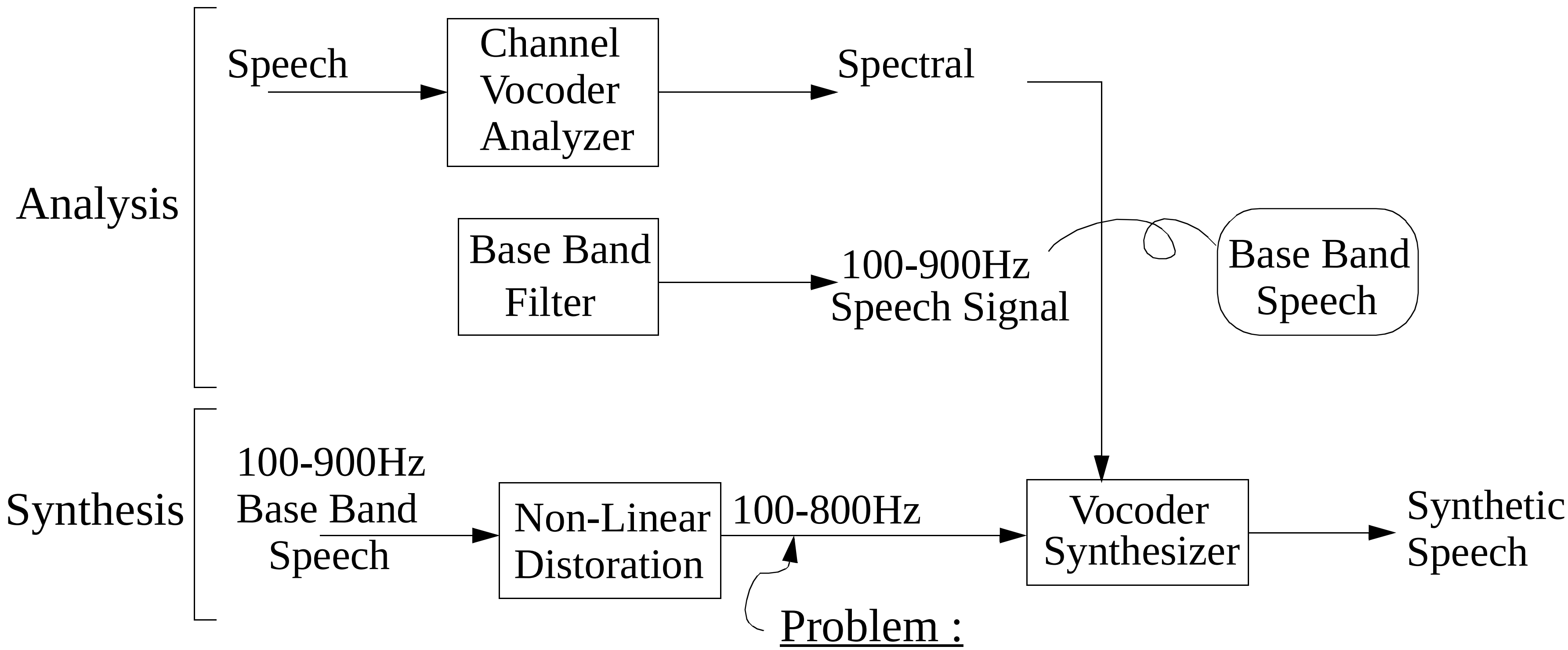
(1) General Faling - Excitation Signal in Channel Vocoder was NOT Robust.

(2) Possible Transmission of a Wide Band (Analog) Speech Signal (0-8kHz) throughan excisting Telephone Channel [300-3000Hz].

* Intuitive grasp of the fact that baseband [0-900Hz], generally carried all the necessary excitation information for vocal channel vibrations.

* A white noise same use a suitable excitatin for voiceless sounds.

Original Idea



Problem :

Distortion of the excitation spectrum.

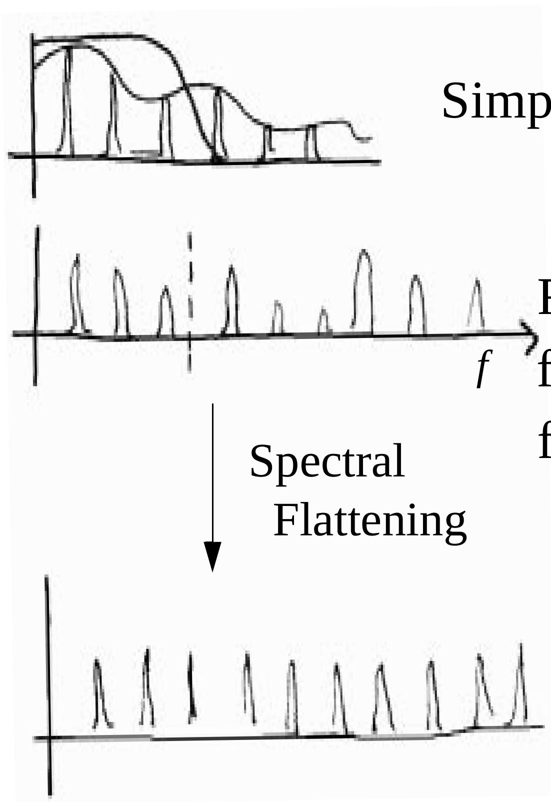
* Intuitive grasp of the fact that baseband [0-900Hz], generally carried all the necessary excitation information for vocal channel vibrations.

* A white noise same use a suitable excitatin for voiceless sounds.

In early 1960's, Schroeder, David et al at BTL, introduced spectral flattering

-got rid of spectral distortion [more or less].

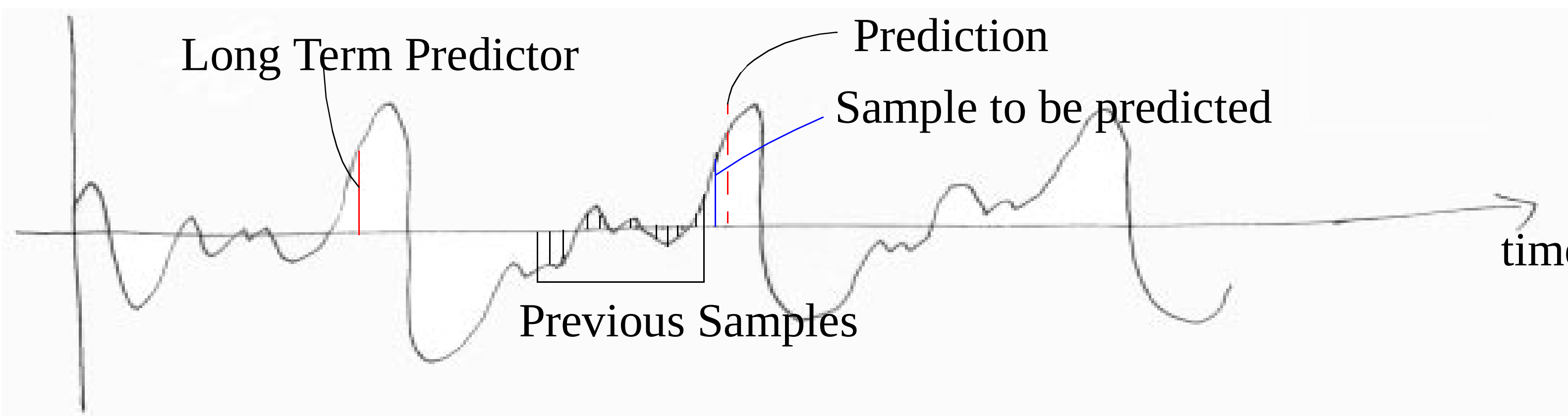
<u>Base Band Signal</u>



Simple non-linearity (rectifier)

Result of non-linearnity distortion products frequencies are harmonics of the fundamental frequency.

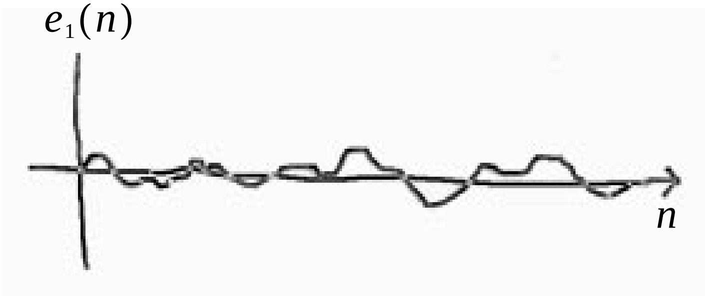
In the late 1960's, schroeder and Atal introduced APC
-Adaptive Predictive Coding.



Two types of prediction.

Predicted Value $\hat{y}(n) = \alpha y(n-T) = y(n) + e(n)$

so $e_1(n) = -y(n) + \alpha y(n-M)$



Perform LPC Analysis on $e_1(n)$

$$e_1(n) = a_1 e_1(n-1) + a_2 e_1(n-2) + \ldots a_k e_1(n-k) + e_2$$

By transmitting $a_1, a_2, \ldots a_k$ and α, M

Major Assumption - $e_2(n)$ is so small

that it can be quantized to 1 bit. BUT SENT at the sampling rate.

Even a ONE bit error signal results in a large bit rate.

If sampling rate is 8kHz, then transmission of error signal costs 8kbs.

Addition of transmitting $\underbrace{d, M}$, a_k's sould be another 2kbs.

Pitch Detection

If pitch is wrong, first error signal $e_1(n)$ is big.

Early APC Systems Operated at 9600bps.

Major Research Efforts : Reduction of Error Signal Bit Rate.

Key to Error Signal Reduction.

* LPC - Error Signal is eliminated and replaced by standard Excitation Signals.

(like Channel Vocoder)

* RELP - Residual Excited linear Prediction.

Low Pass Filter of error Signal - reduced sampling rate.

still one bit quantization.

* VELP - Voice-Excited LP.

So error signal rate can be reduced [hopefully].

They aimed for 4800 bps.

As computer speeds increased, New sptions became available for

real-time Coding of the Error Signal.

Basic Philosophy - Analysis by Synthesis

* Transmitting system has both analyzer and synthesizer available.

* So Synthetic Speech can be generated at the transmitter.

* Using same criteria, the synthetic speech is compared sequentially with the actual speech [perhaps every frame, or every n frames] and synthesizer parameters obtained by analysis VARIED to obtain a GOOD fit between ACTUAL SPEECH vs. SYNTHESIZED SPEECH.

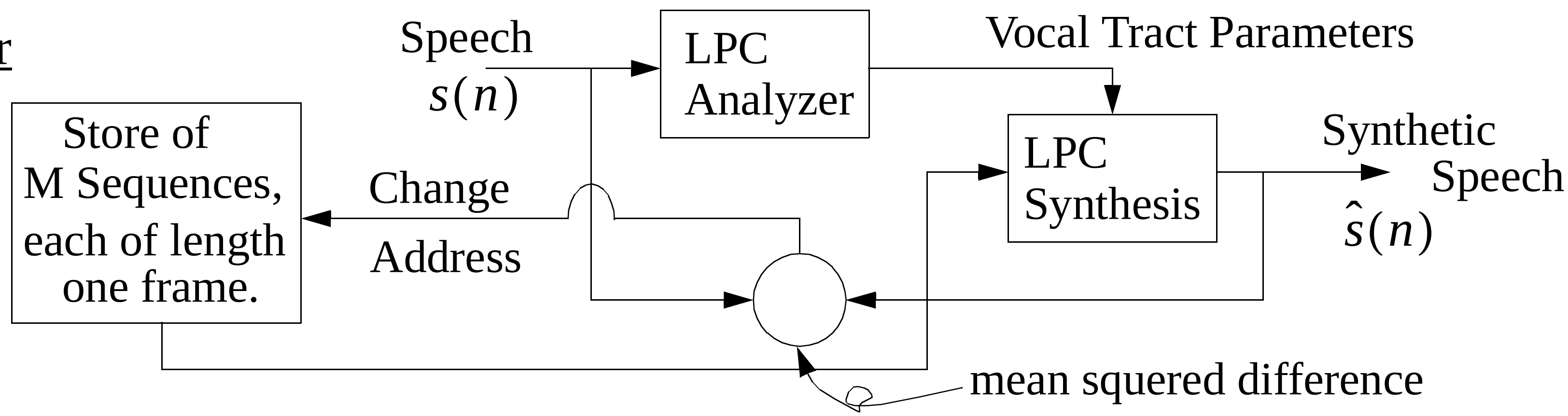
* So the best parameters are sent.

we encountered this idea
in the Setevens Halle concept
in chapter 17.

Basic Idea

- Replace the one-bit error signal of APC with a vector quantized error signal.

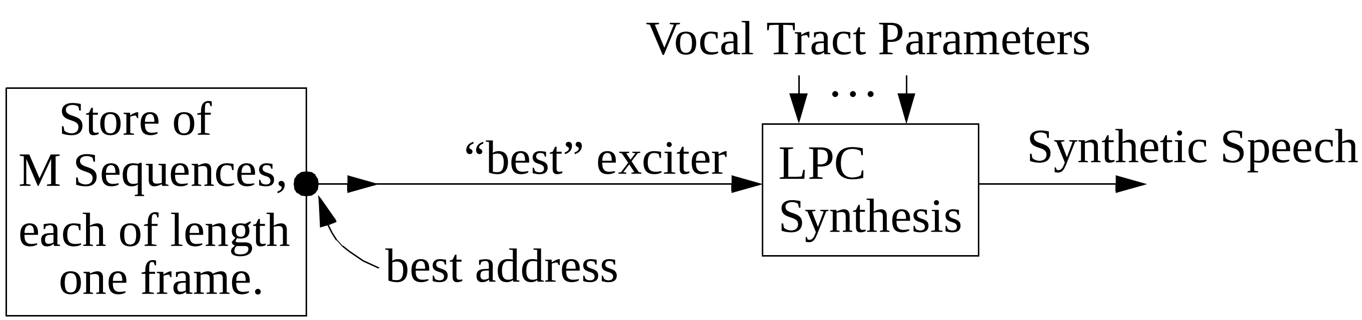
Analyzer



Transmit address of best of the M sequences transmit the Vocal Tract Parameters.

Receiver

This is an example of VQ.



Problems discussed in Chapter 33

* How to create the stone?

* Perceptual weighting filter.

* Delay

* Reduction of Codebook Search. [In 20ms, the analyzer must performe complete analysis-synthesis many times in a single frame.] So computer speed must be such that MANY systems can opearate in real time SIMULTANEOUSLY.

* Adaptive Coding.[Storage is adaptive to a different speaker]

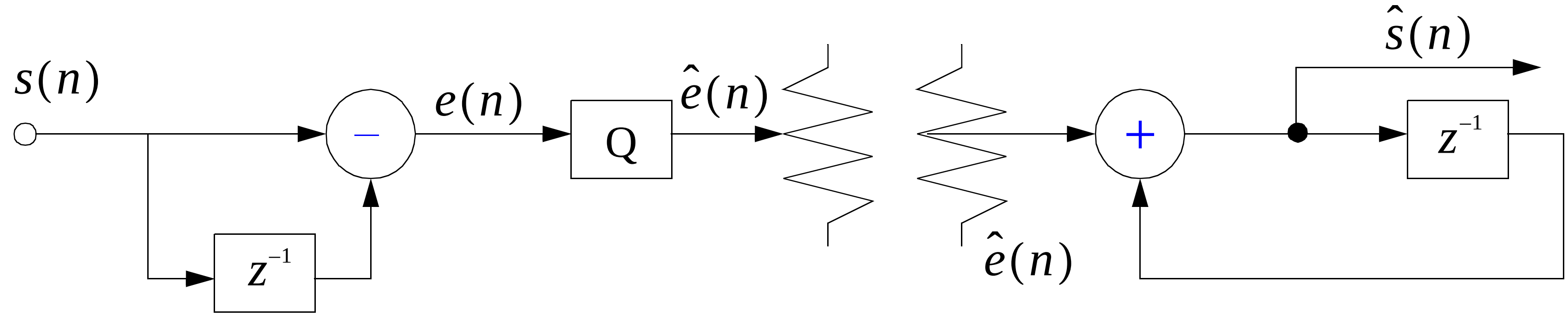


Figure 33.7 : Differential Pulse Code Modulation.

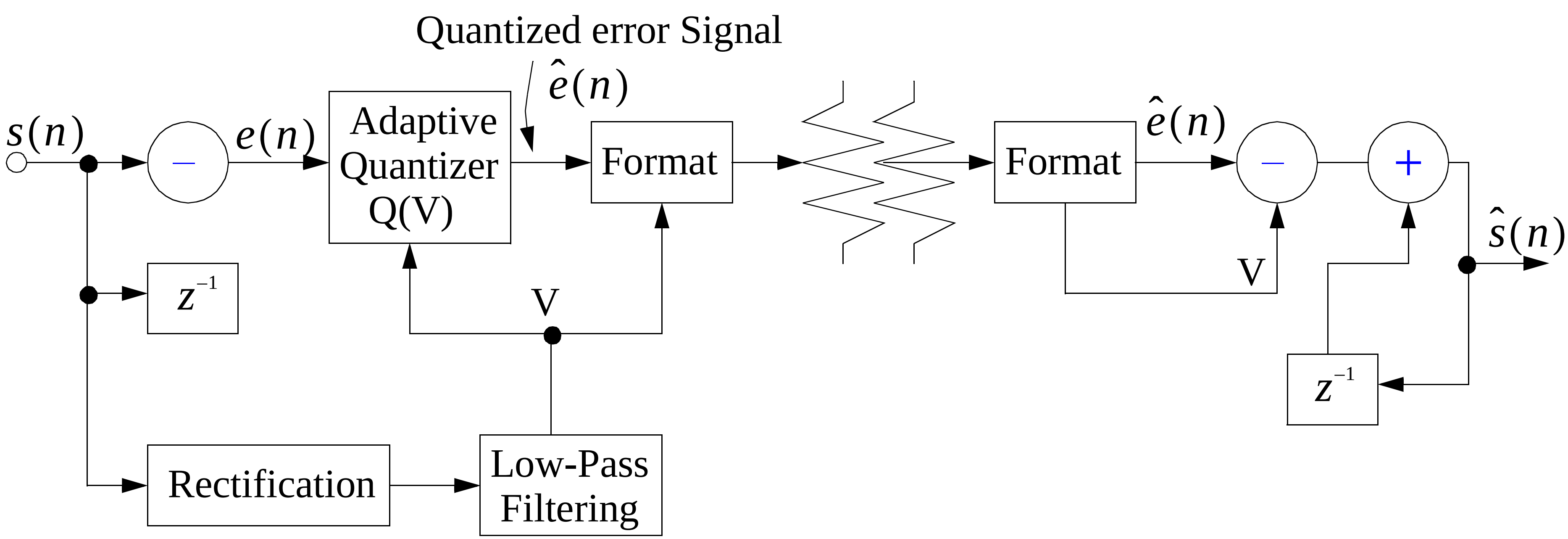


Figure 33.8 : Adaptive Differential Pulse Code Modulation.

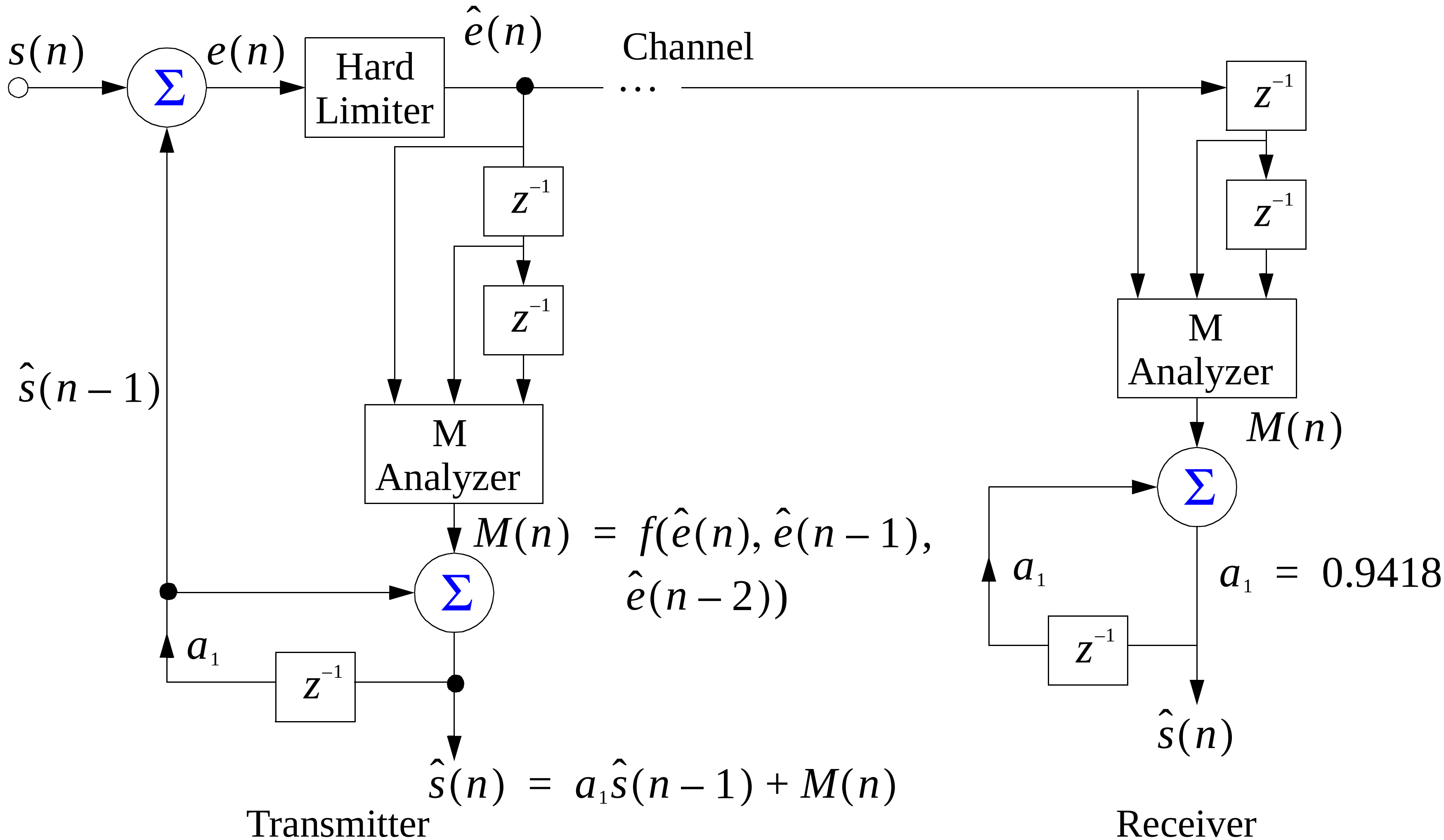


Figure 33.9 : Continuously Variable Slope Delta Modulation.

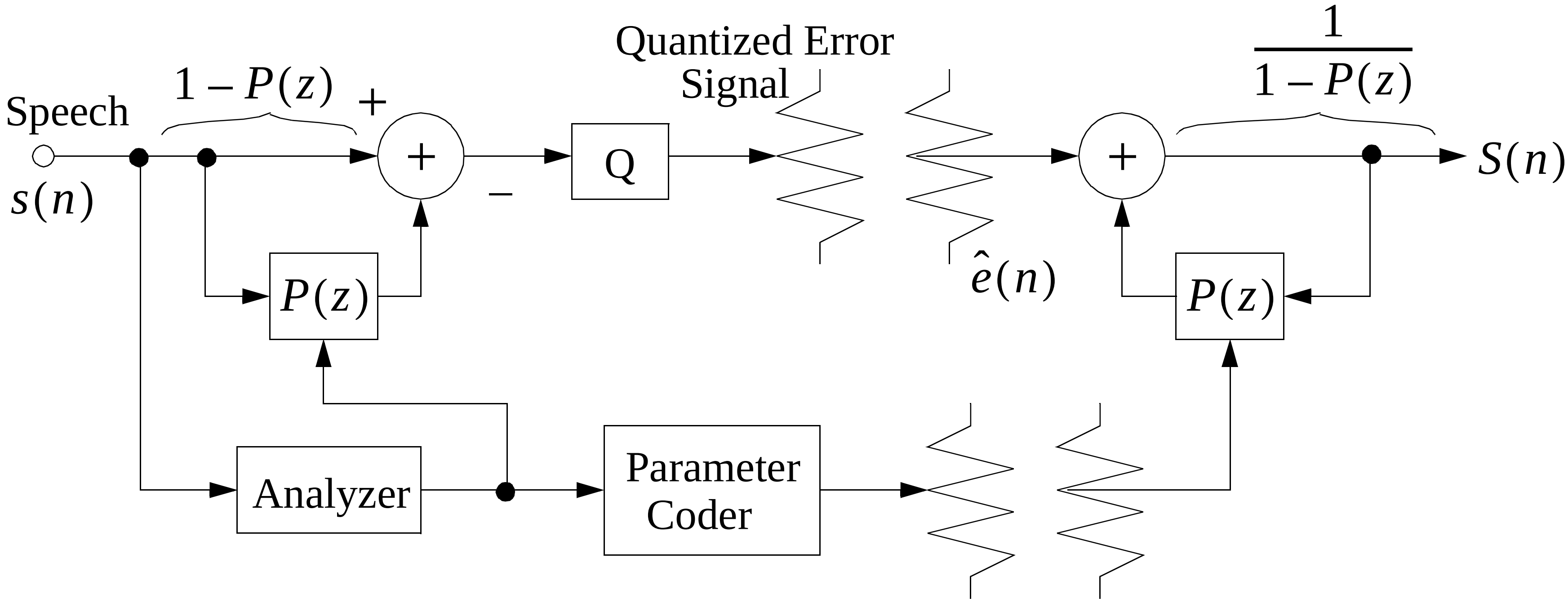


Figure 33.10 : Rudimentary Linear Prediction Concept

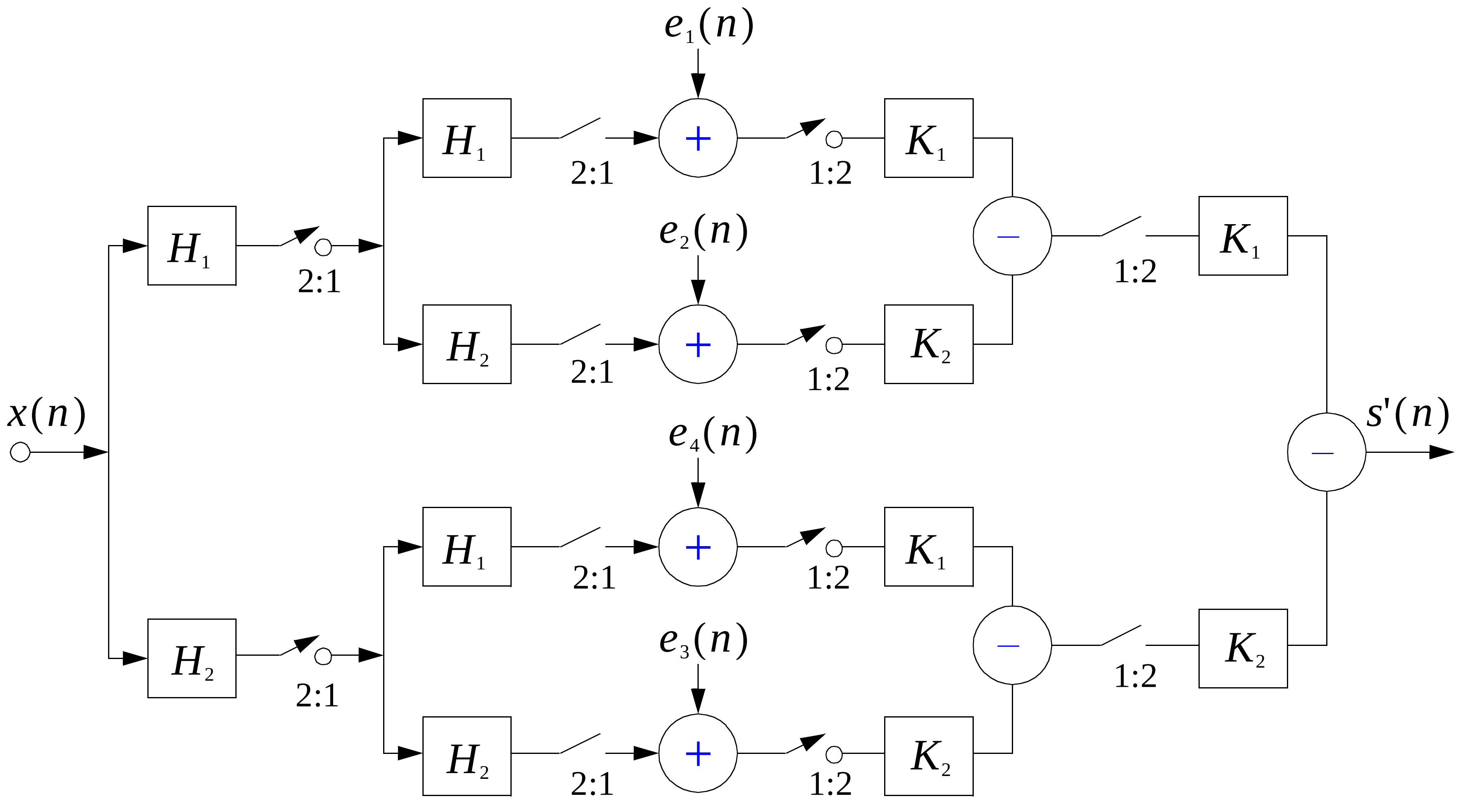


Figure 33.14 : Four Channel Subband Coder with Quadrature Mirror Filters.

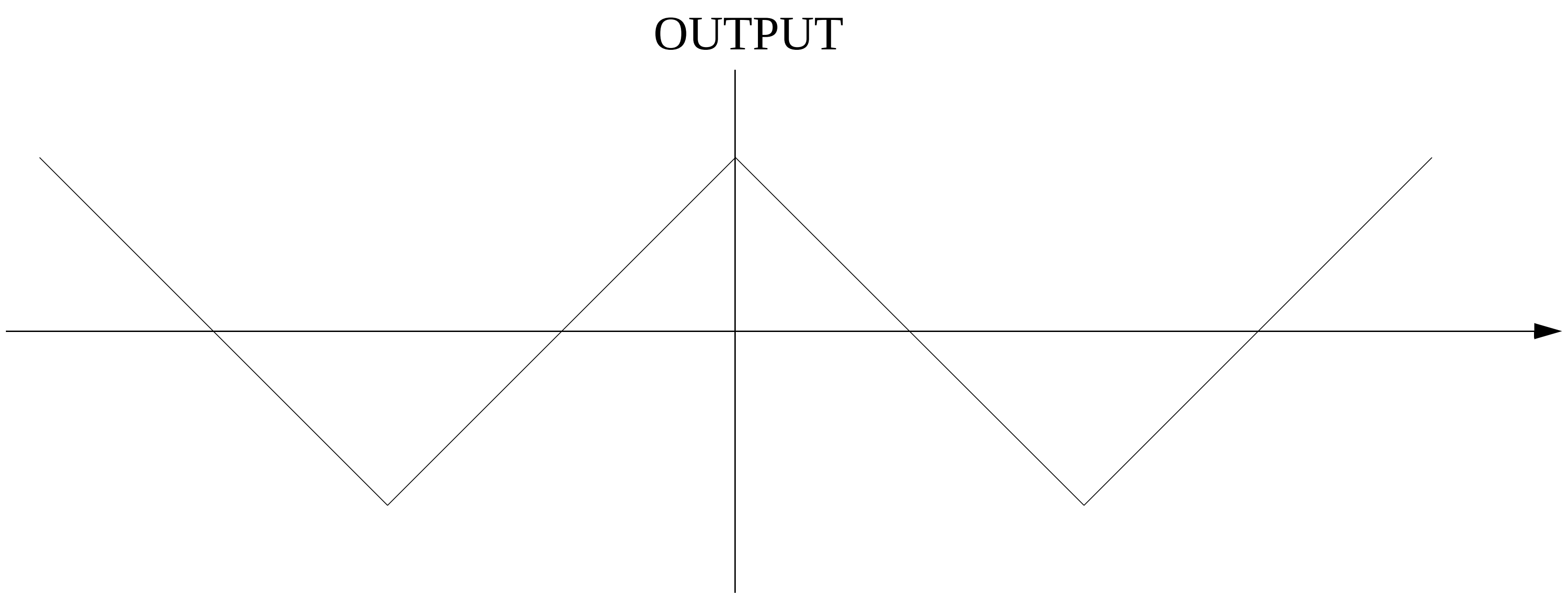
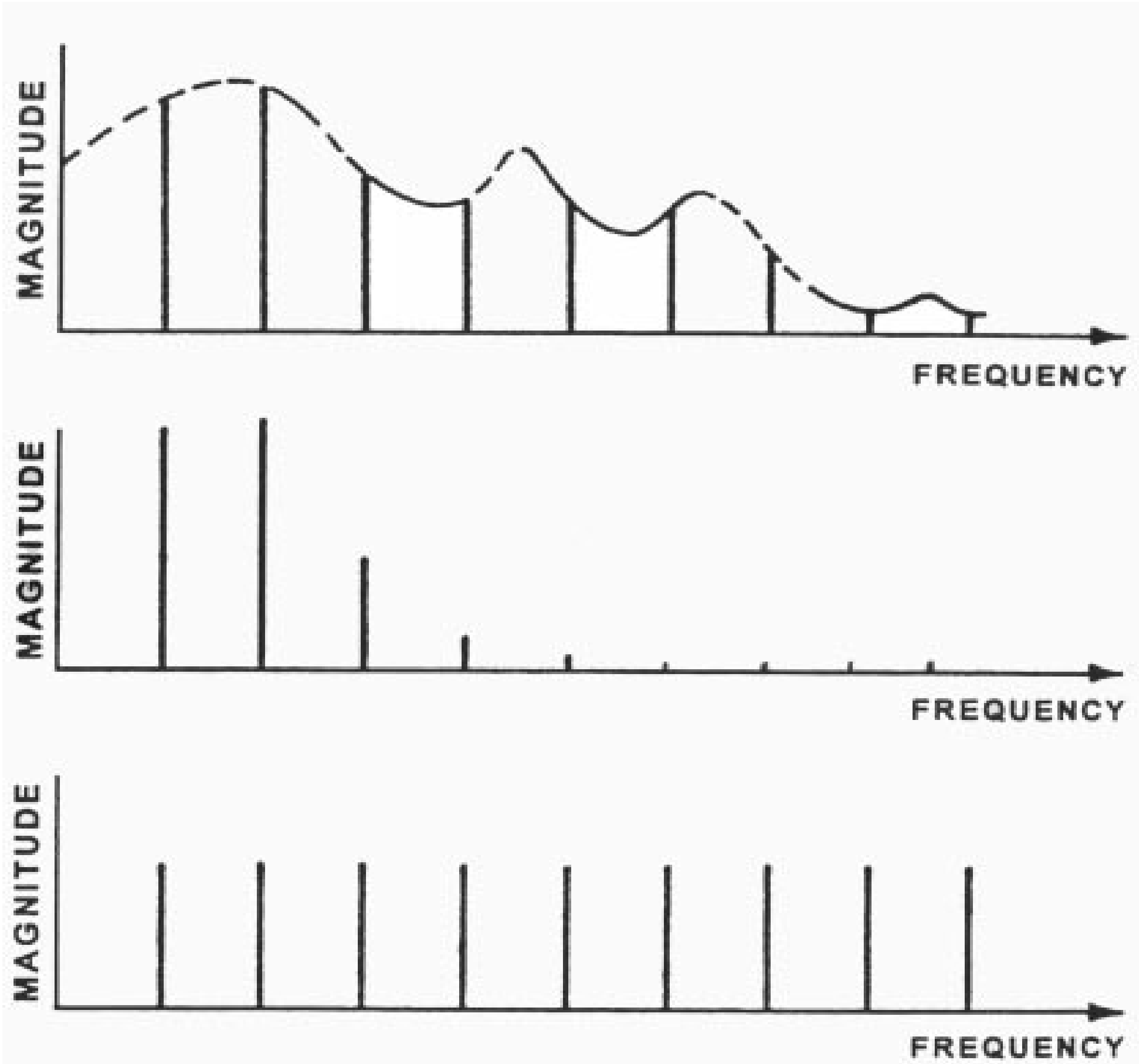


Figure 33.2 : Zig-Zag Network



Frequency

Figure 33.1 : Spectral Flattening of the Base-Band Signal

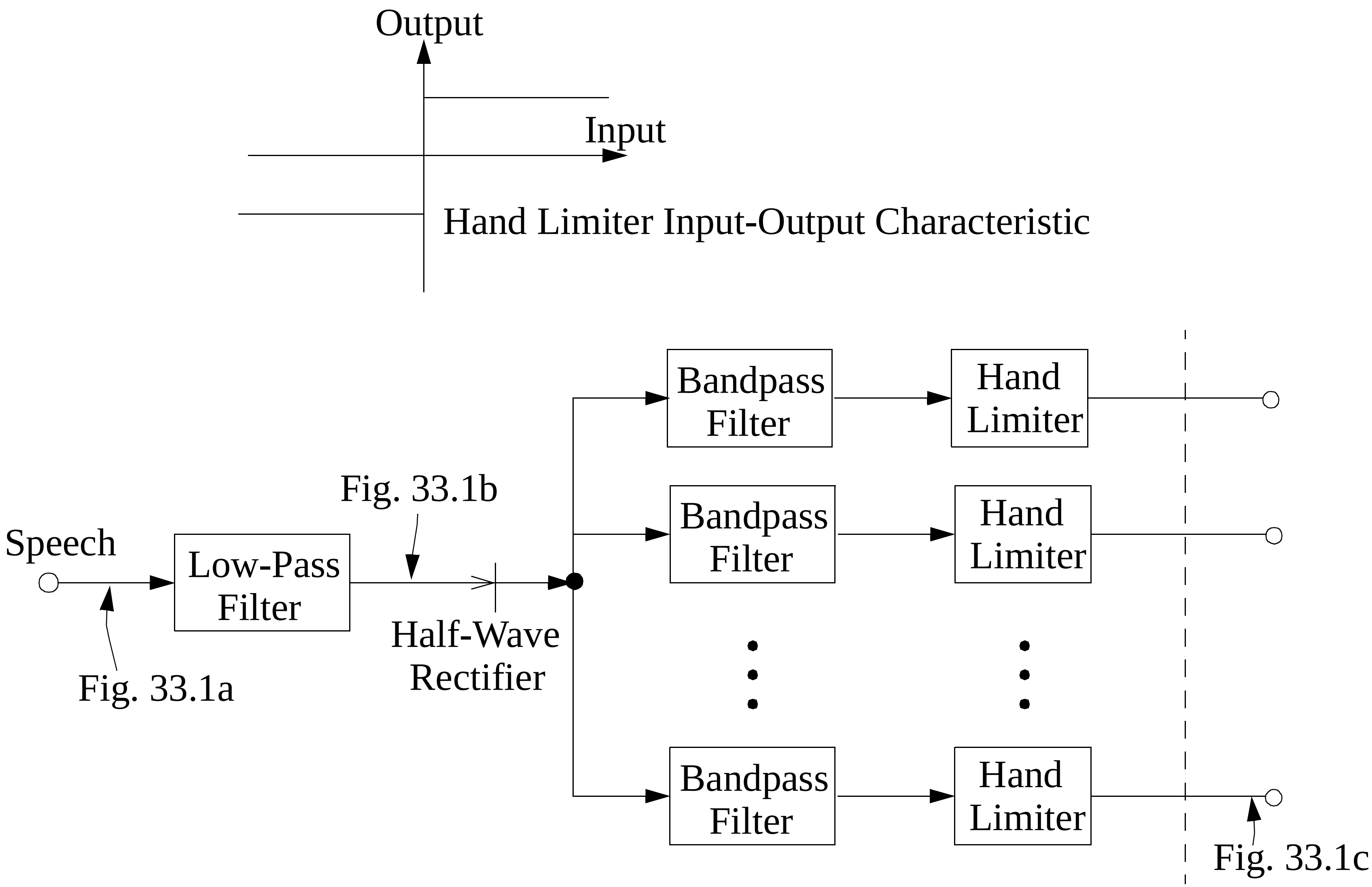
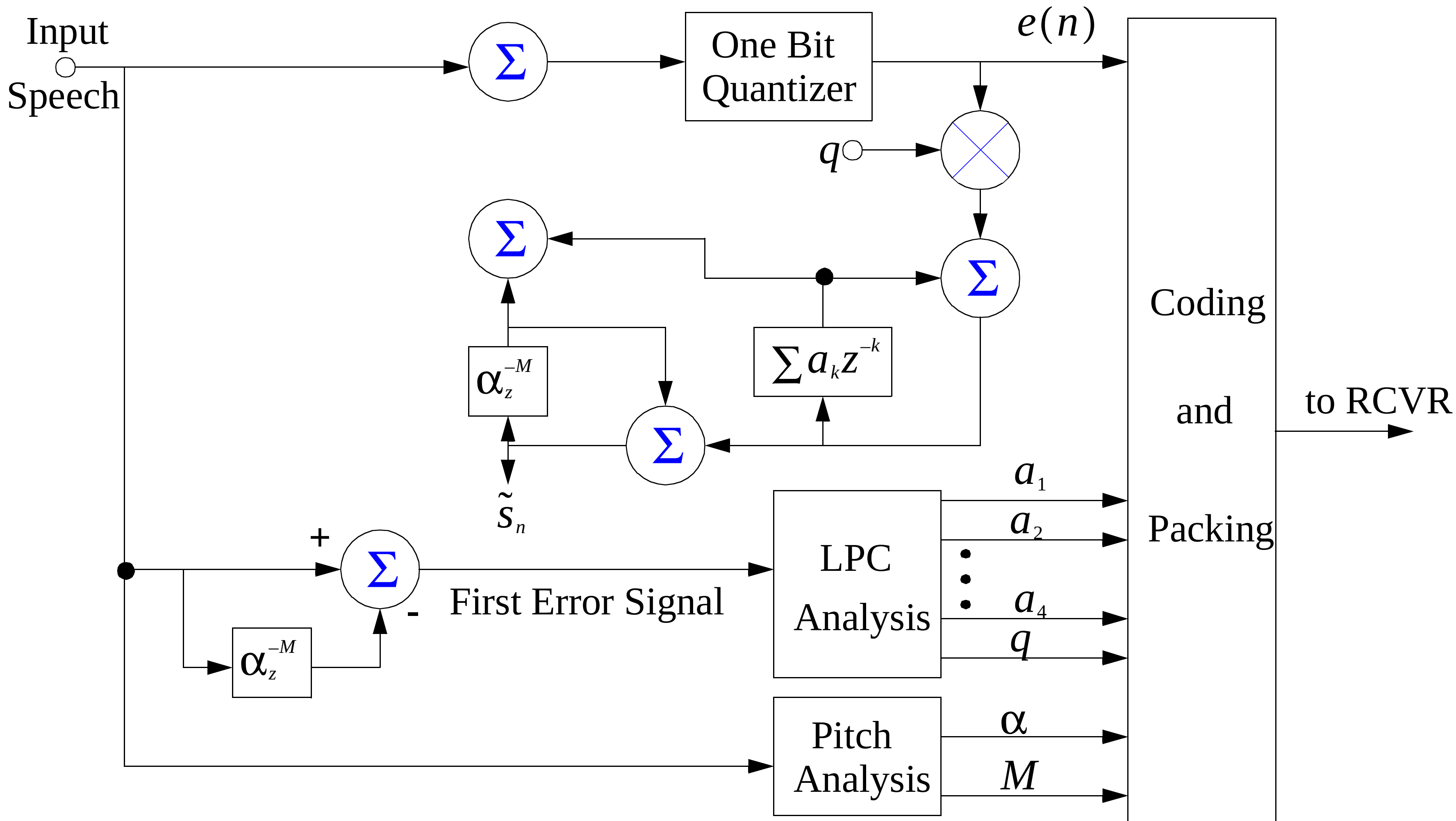
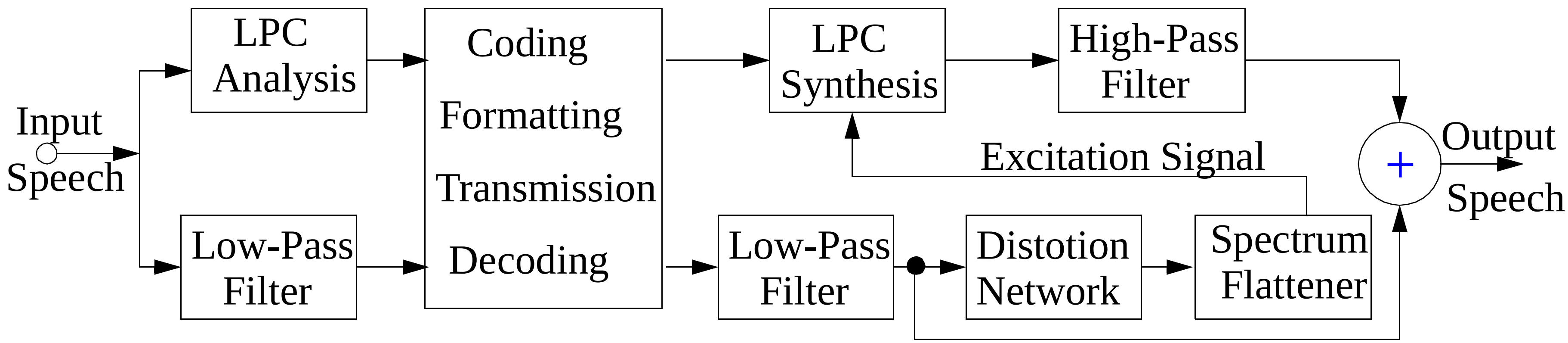


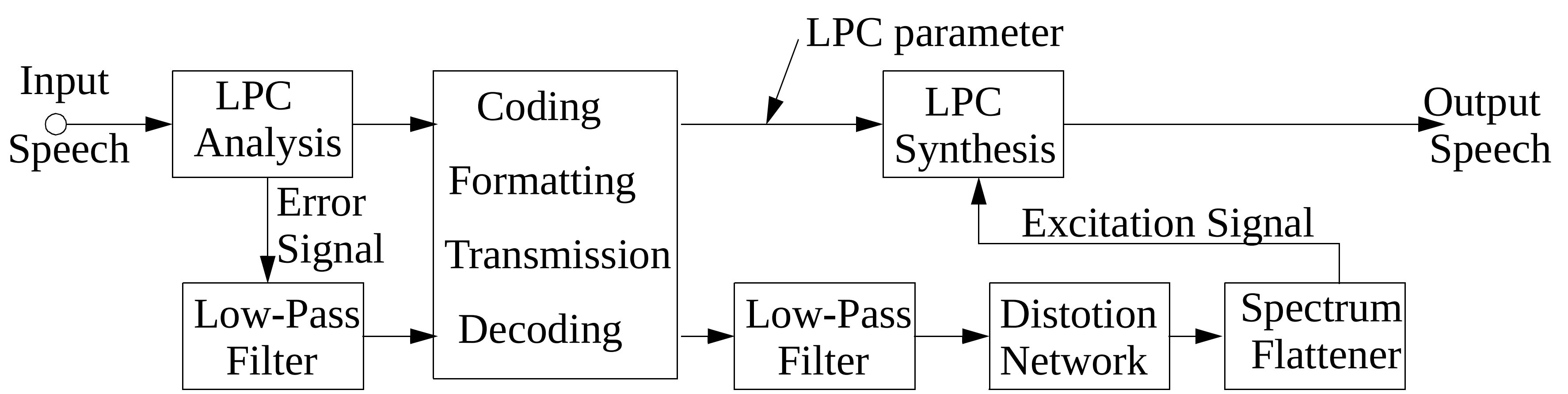
Figure 33.3 : Implementation of Spectral Flattening.



Adaptive Predictive Coding of Speech

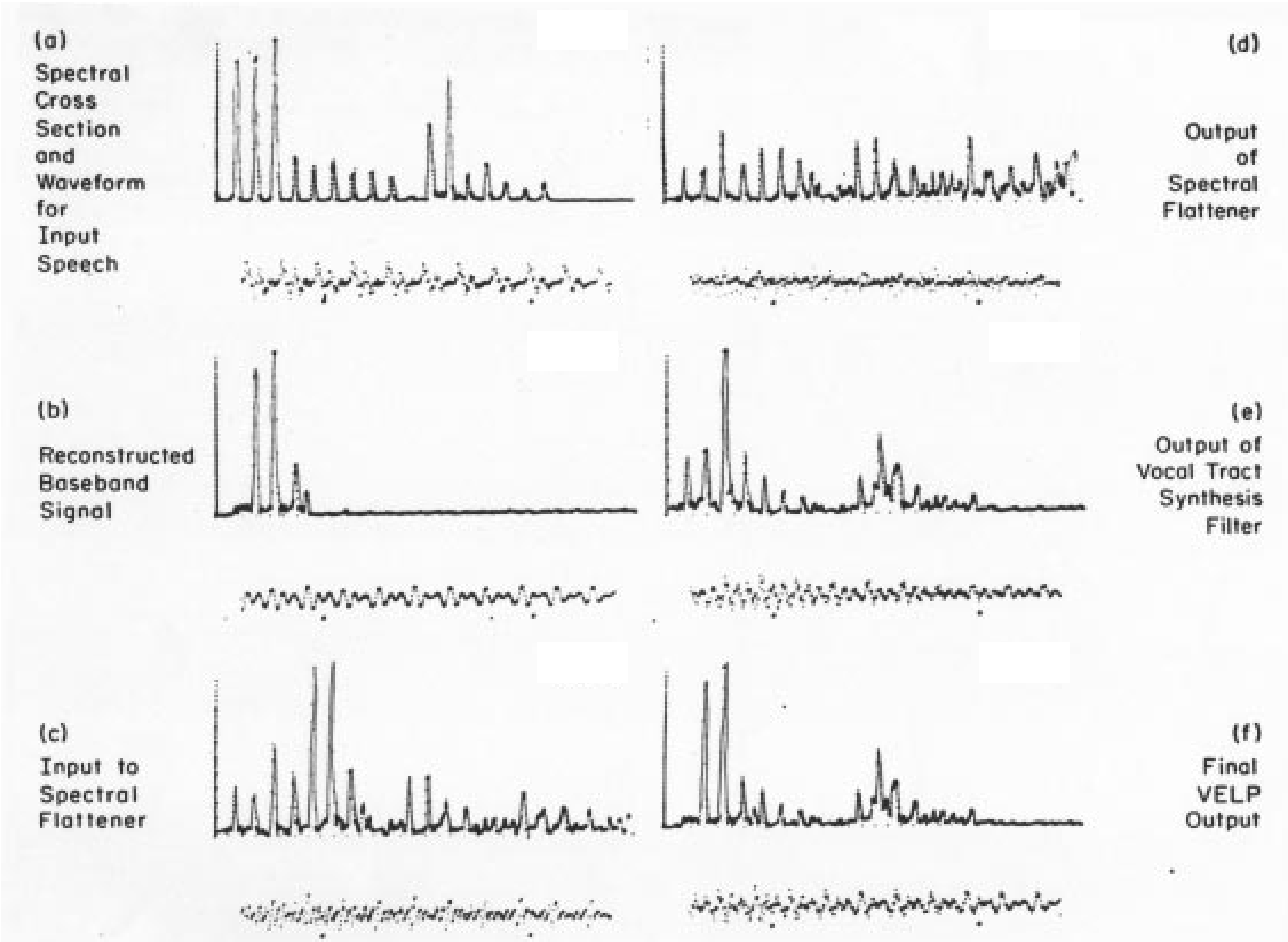


(a) VELP (Voice-Excited Linear Prediction)



(b) RELP (Residual-Excited Linear Prediction)

Block Diagrams of VELP and RELP.



Spectra at different nodes in the VELP algorithm.

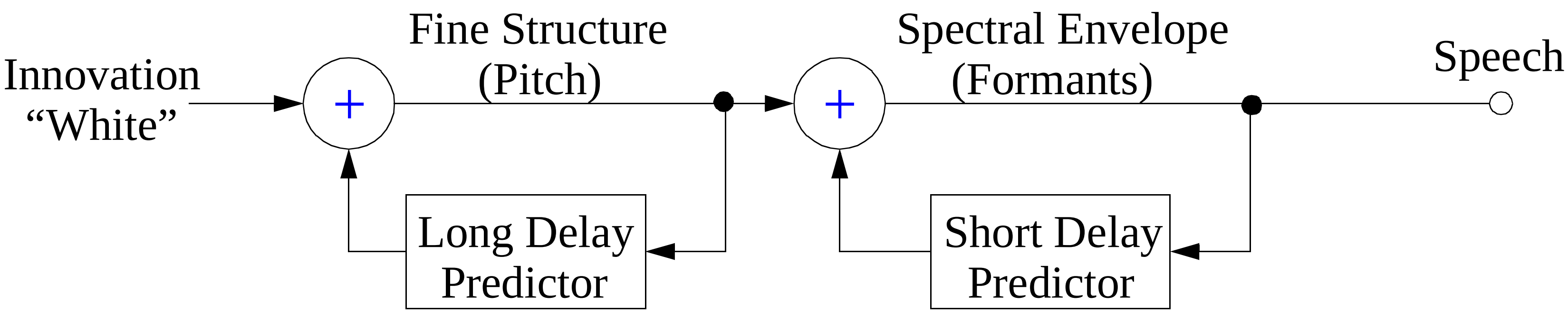


Figure 1 : Speech Synthesis Model with Short and Long Delay Predictors.

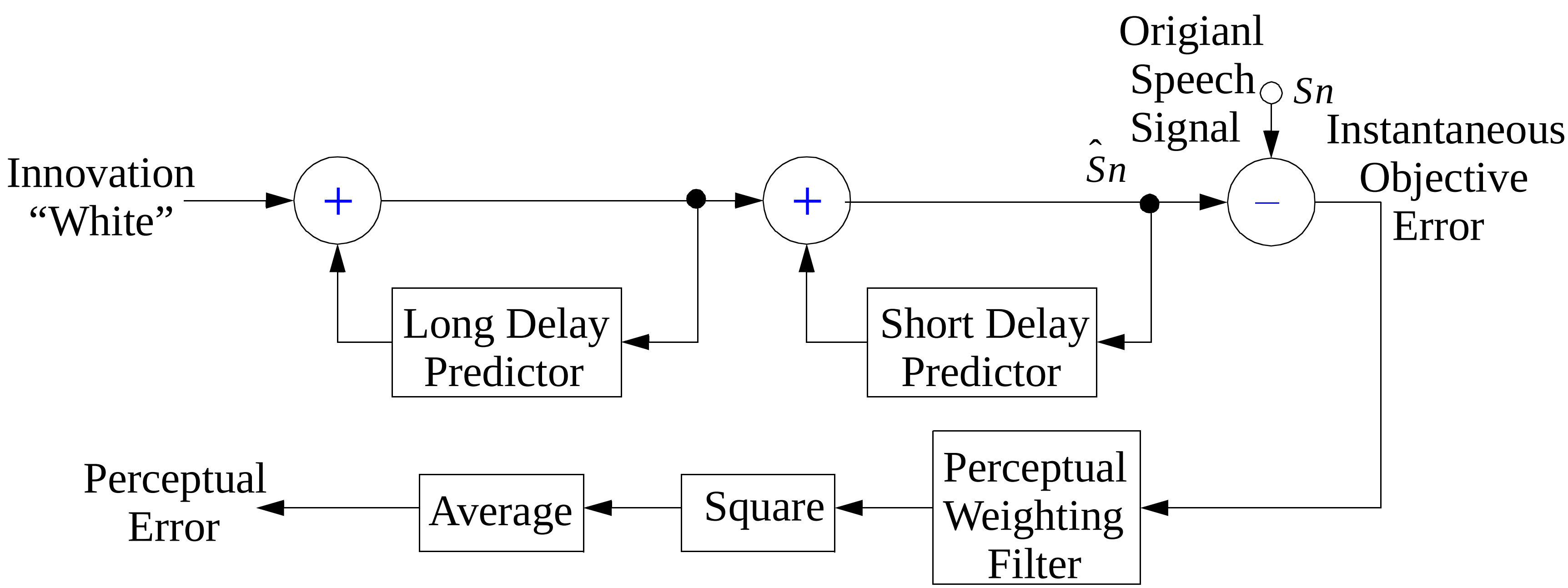


Figure 2 : Block Diagram Illustrating the Procedure for Selecting the Optimum innovation Sequence.

Analysis by Synthesis

How to create the store?

Reduction of delay.

Reduction of Codebook Search

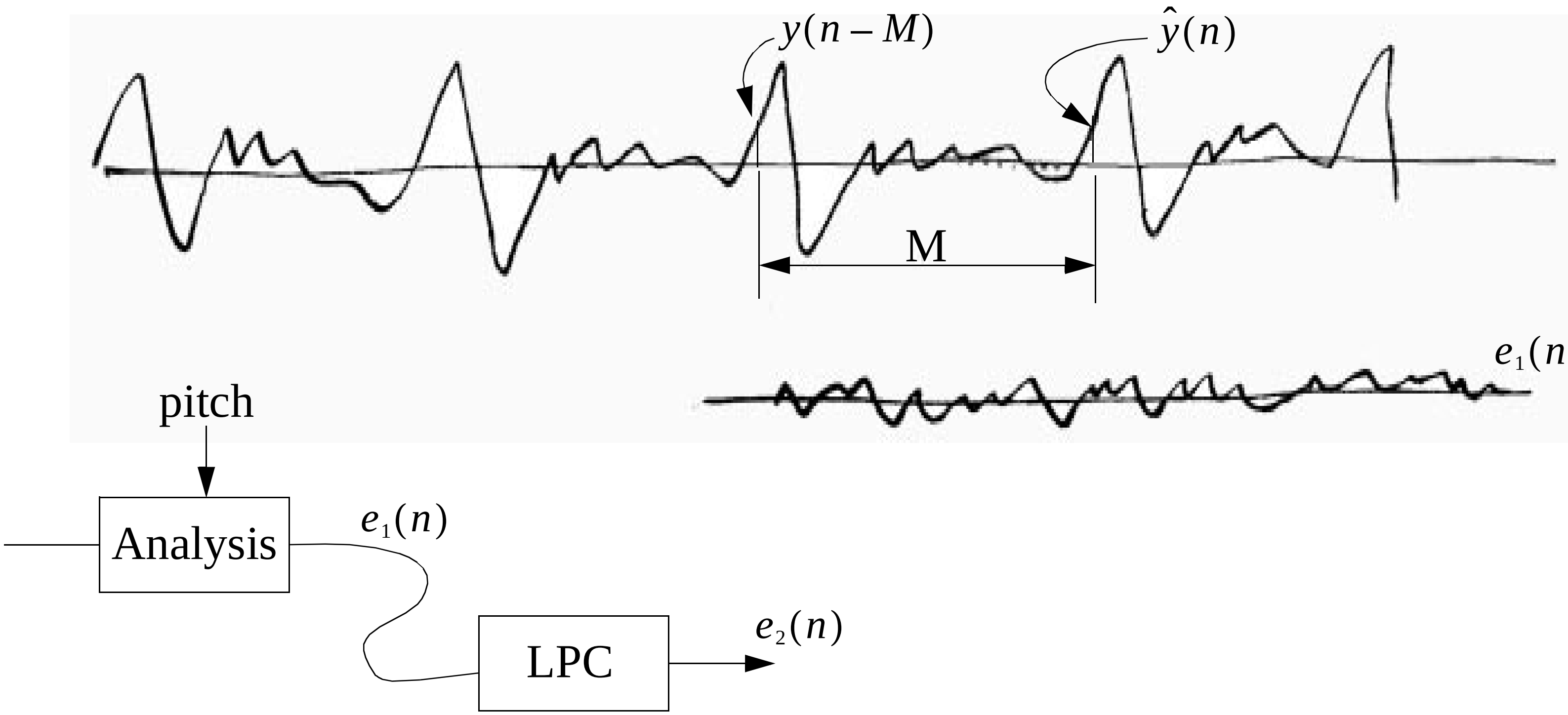
Adaptive Codebook.

APC 10,000bps.

Bessie Smith

Louis Armstong.

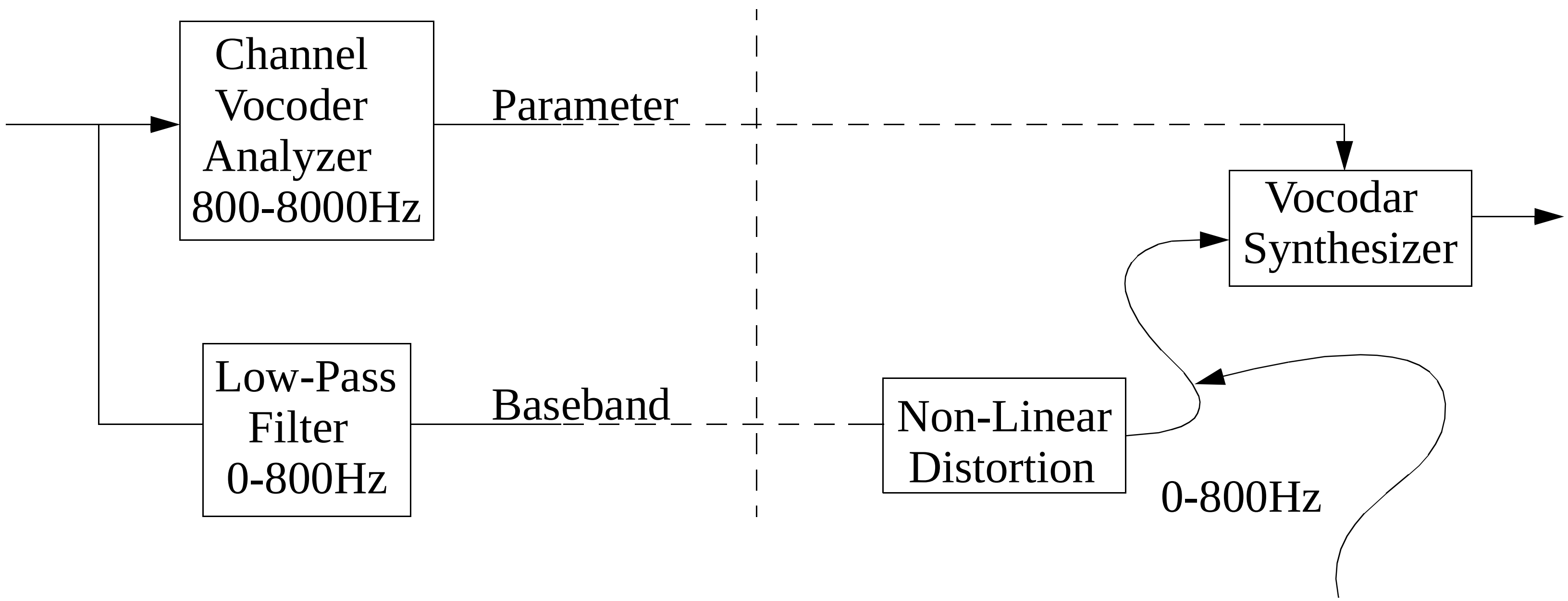
APC - Adaptive Preductor Coder Late 1960's



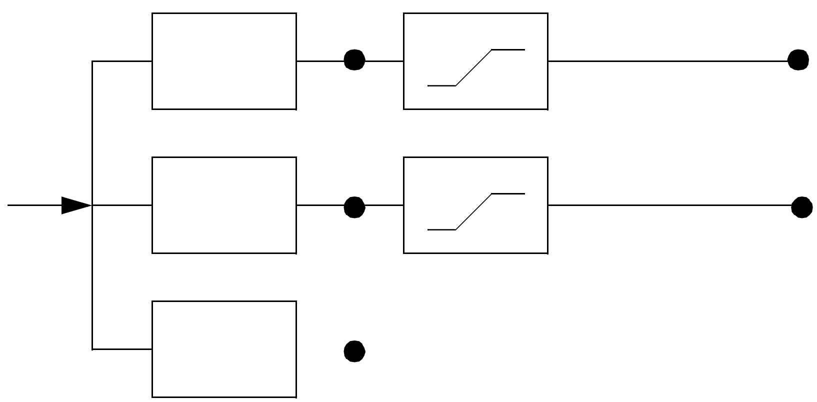
$$\hat{y}(n) = \alpha y(n-M) = y(n) + e_1(n)$$

$$e_1(n) = -y(n) + \alpha y(n-M)$$

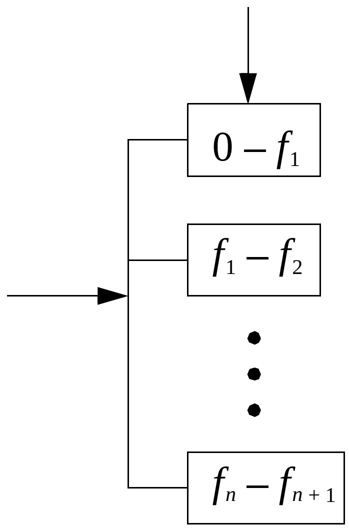
$$\longrightarrow e_1(n) = a_1 e_1(n-1) + a_2 e_1(n-2) + \ldots + \underbrace{e_2(n)}_{\text{i bit}}$$



Impossible to get Flat Spectrum.
10,000 bits per second.

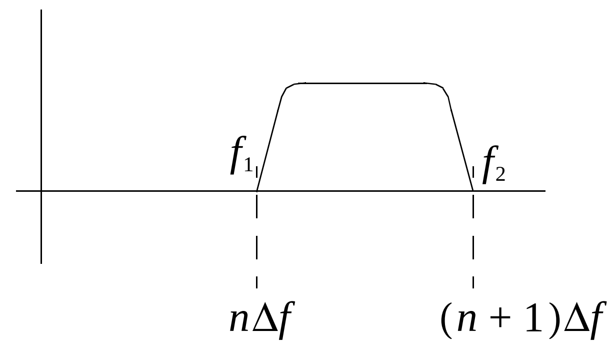


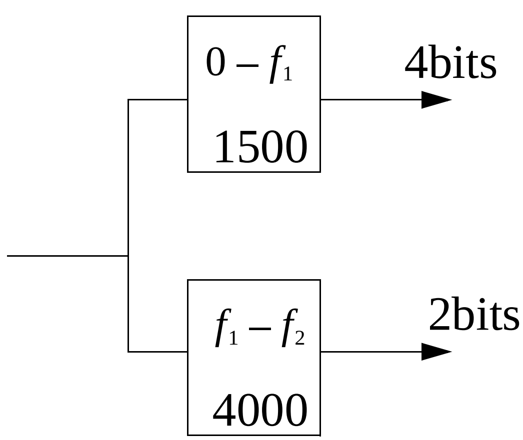
Subband Coding



64k bits

20-30k bits





Medium & High Rate Systems

Quality Possible Vocoding of Sounds other than Speech

Robustness Speaker I.D.

Waveform - PCM - 8kHz, 8bits $\longrightarrow$ 64kbits.
(Subband Coding)

DPCM ADPCM CVSD

Basic Issue

Efficient Representation of the Excitation.

Voice-Excited Systems — LPC
Channel Vocoders

Prediction - APC $\longleftarrow$
CELP
Multi-Pulse