University of California
Berkeley

College of Engineering
Department of Electrical Engineering
and Computer Sciences

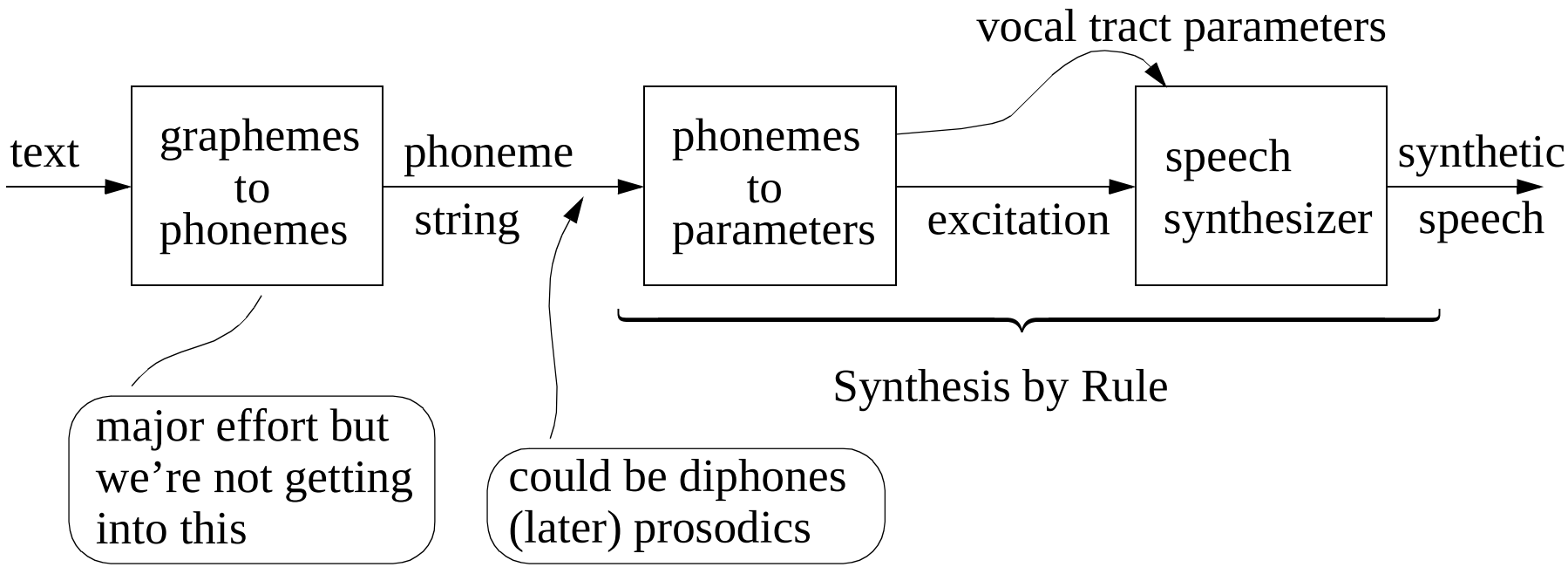
Professors : N.Morgan / B.Gold

EE225D

Spring,1999

Speech Synthesis

Lecture 23



We'll be listening to

* Early synthesizer
* Synthesis by rule ——— researcher enters phoneme
* Complete text-to-speech system

Evolution Desche??has early synthesis was done (page4?)

-Different configuration ⟶ Connected with Vocoders

	Tape Number	Figure
	1. Voder	29.7
	2. Pattern Playback	29.1
	3. PAT	No Fig.
	4. OVE	No Fig.
	5. PAT2	29.8
	6. OVE II	29.2
	7. OVE II (holmes)	
	8. Holmes II Synthesis	29.3
	9. Klatt (Male Fem.)	29.4
	10. Dectalk	
	11. Davo	29.9
	12. Flanagan	29.10&29.11
	13. Speach & Spell	29.5
	14. Multi pulse LPC	
Synthesis by Rule	15. Pattern Playback	29.1
Synthesis by Rule	16. Kelly Gerstman	get Fig from previous 11.3

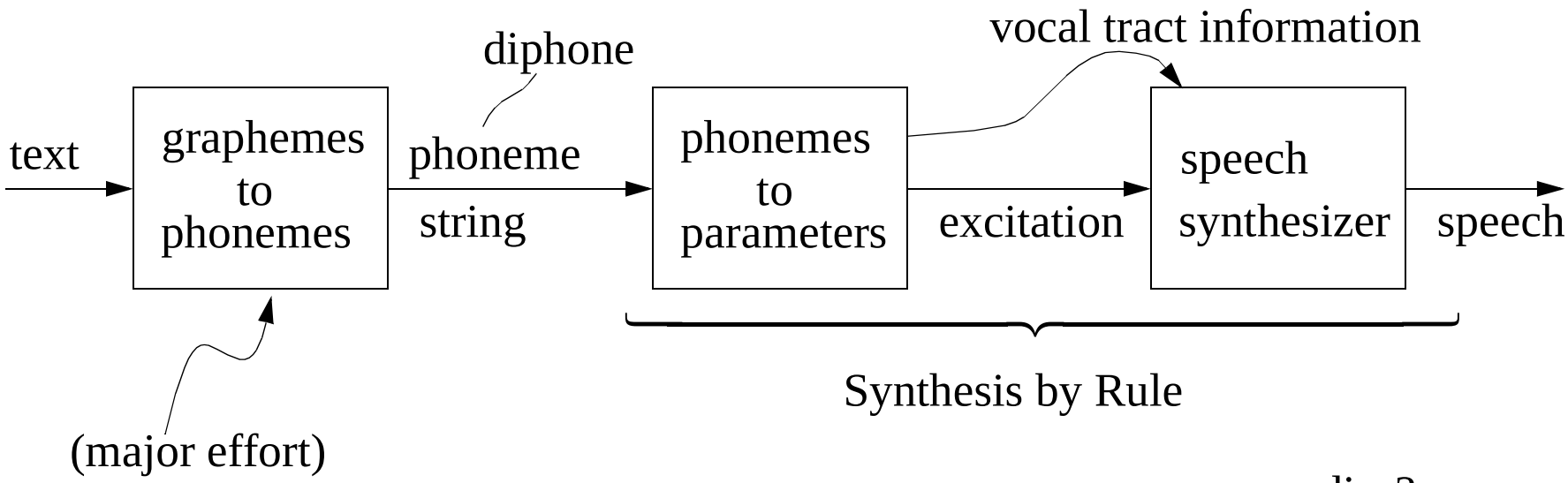
Synthesizers can be

- Channel vocoder, LPC or homomorphic
- Serial formants [each formant is a two-pole network]
- Parallel formants –
- Articulatay models
- Oddball arrangement pattern playback

Evolution

* Researcher piches an utterance, creates a spectrogram.
* Researcher has a synthesizer model at his/her disposal.
* Researcher enter secuence of parmeter values into model.
* Synthesizer “Speaks” and researcher adjusts sounds so utterance searches better, before this. We had the Voder where the instrument was “played” in real time by a skilled performer.

Speech Synthesis



prosodics?

Early Synthesizers
Synthesis by Rule
Complete text-to speech.

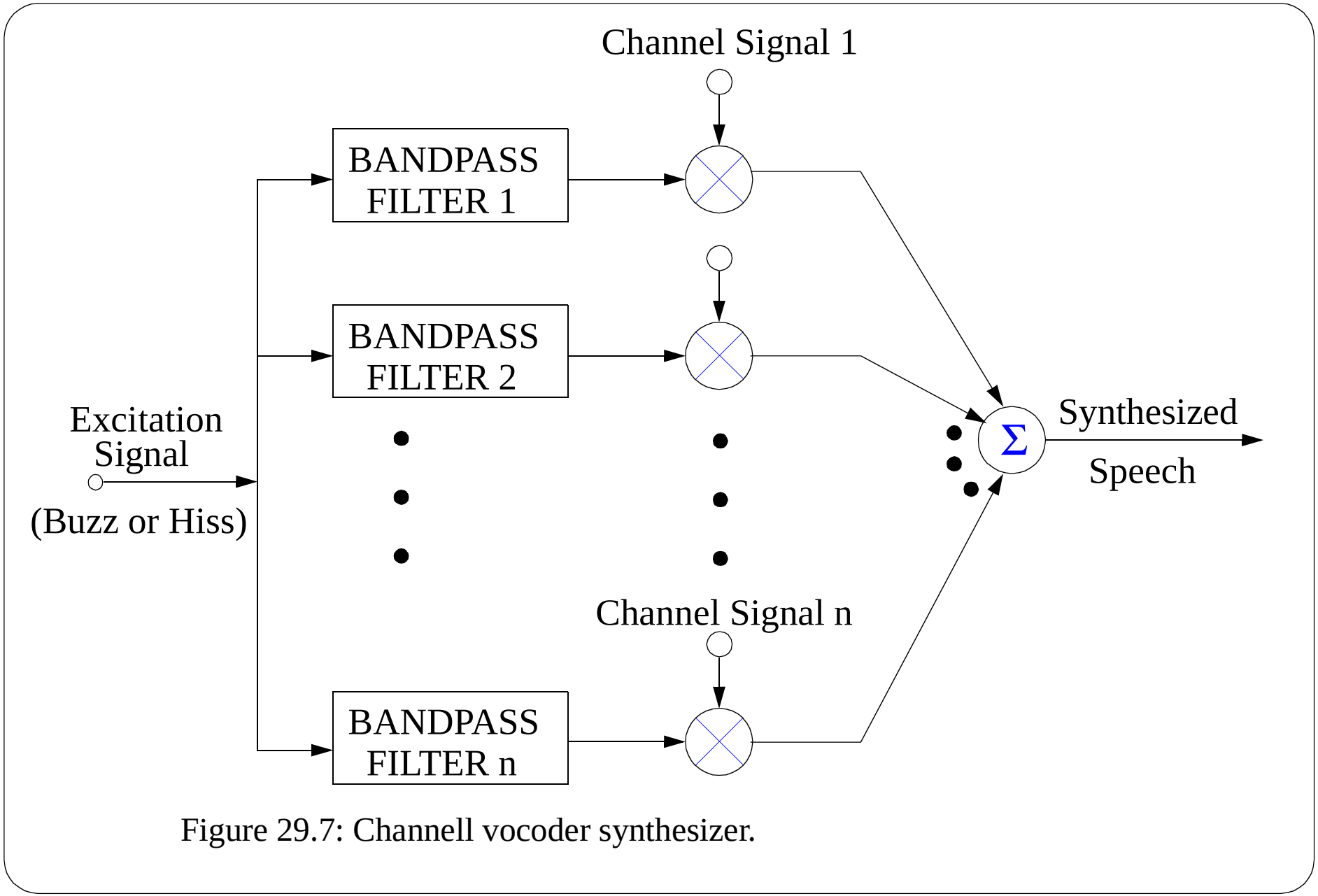


Figure 29.7: Channell vocoder synthesizer.

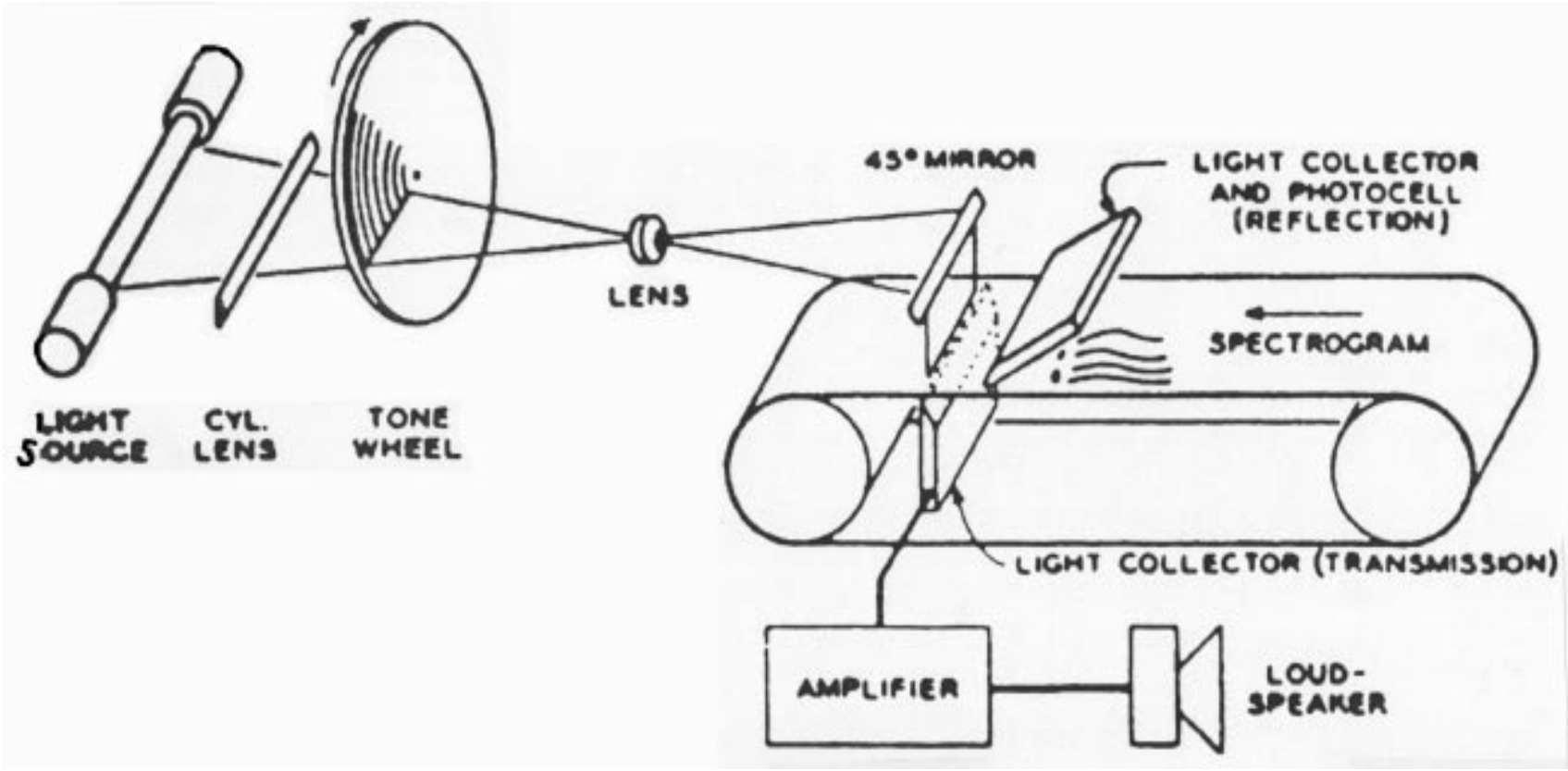


Figure 29.8 : Light Collector, mirror, Tone wheel, Spectrogram etc.

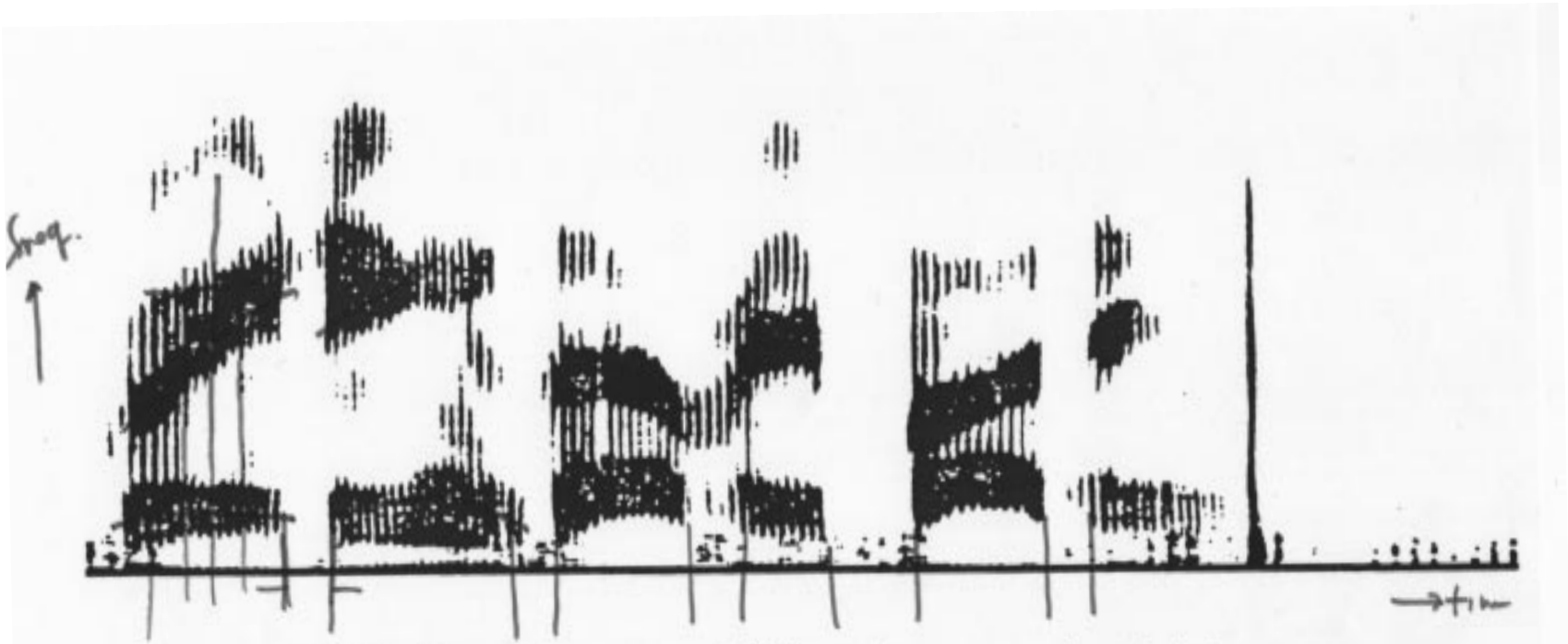


Figure 2.12: Spectrogram of "Greetings everybody" by announcer

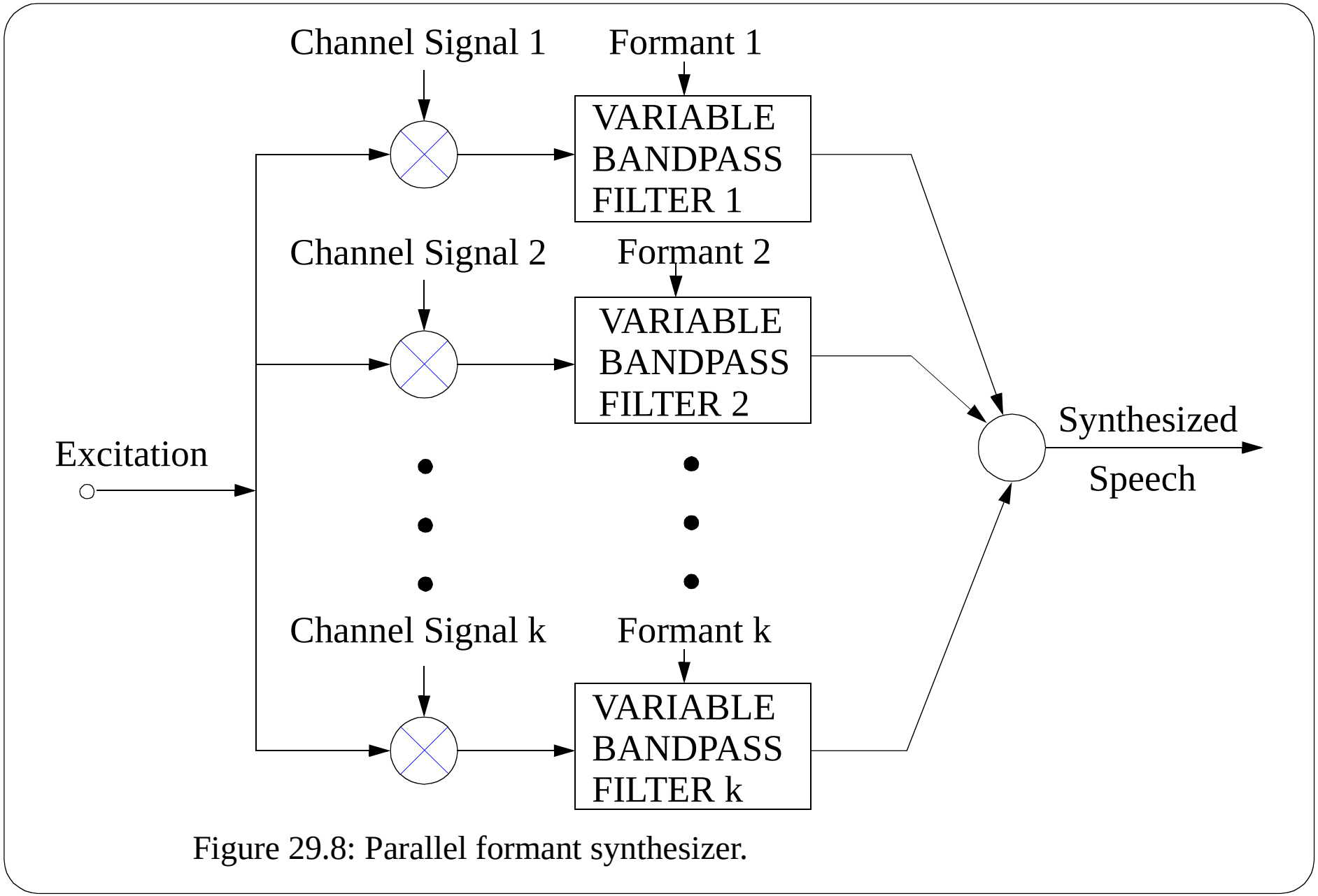


Figure 29.8: Parallel formant synthesizer.

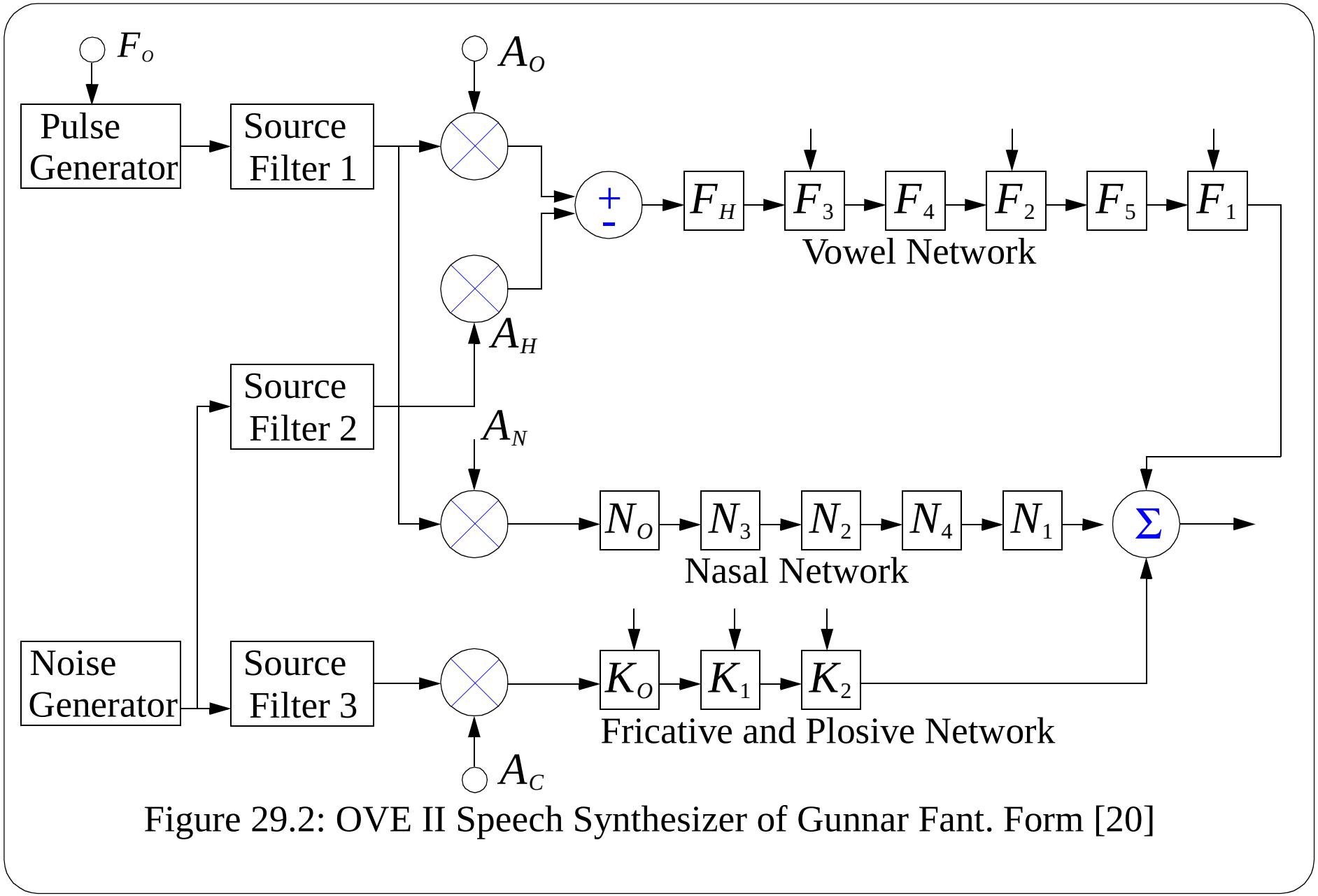
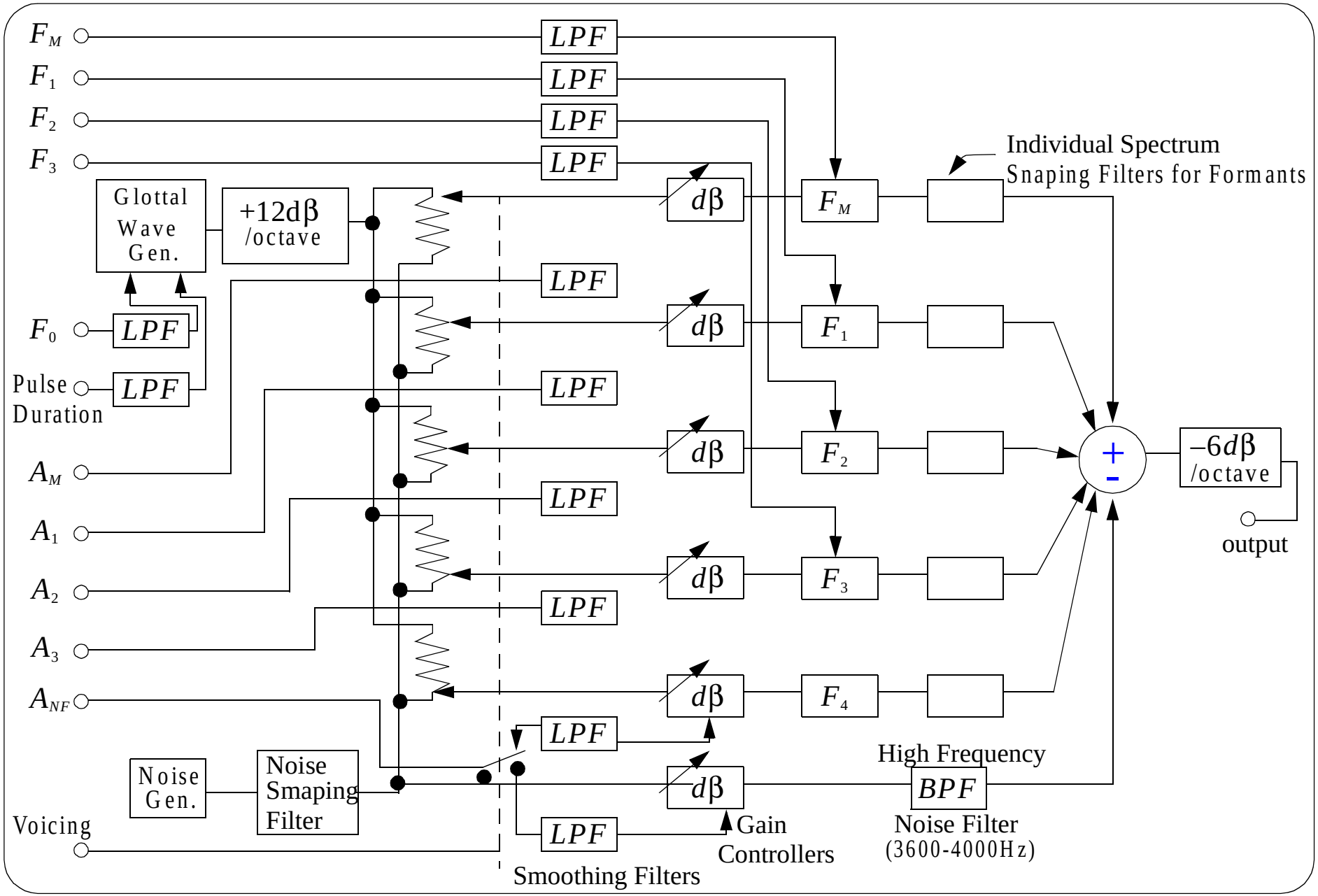


Figure 29.2: OVE II Speech Synthesizer of Gunnar Fant. Form [20]


F_M
F_1
F_2
F_3
F_0
Pulse Duration
A_M
A_1
A_2
A_3
A_{NF}
Voicing
LPF
Glottal Wave Gen.
+12dβ /octave
dβ
Individual Spectrum Snaping Filters for Formants
F_M
F_1
F_2
F_3
F_4
+
-
−6dβ /octave
output
Noise Gen.
Noise Smaping Filter
BPF
High Frequency Noise Filter (3600-4000Hz)
Gain Controllers
Smoothing Filters

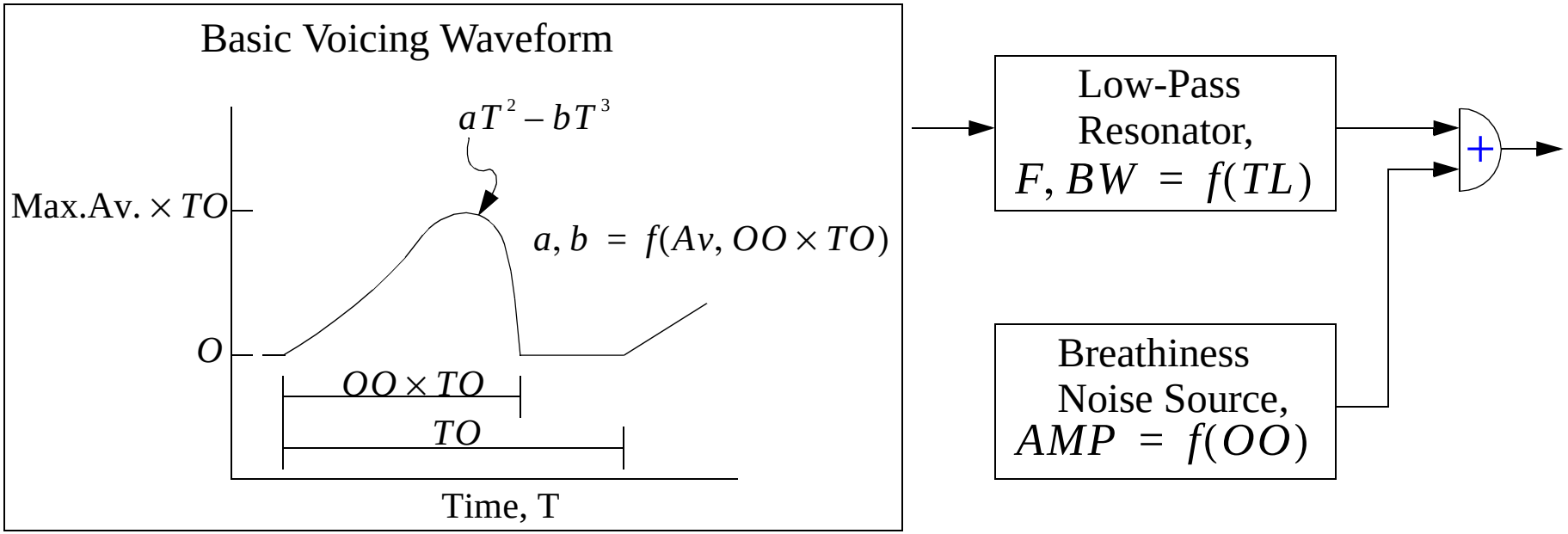


Figure 29.4: The Klatt Synthesizer. From [35]. (cont.)

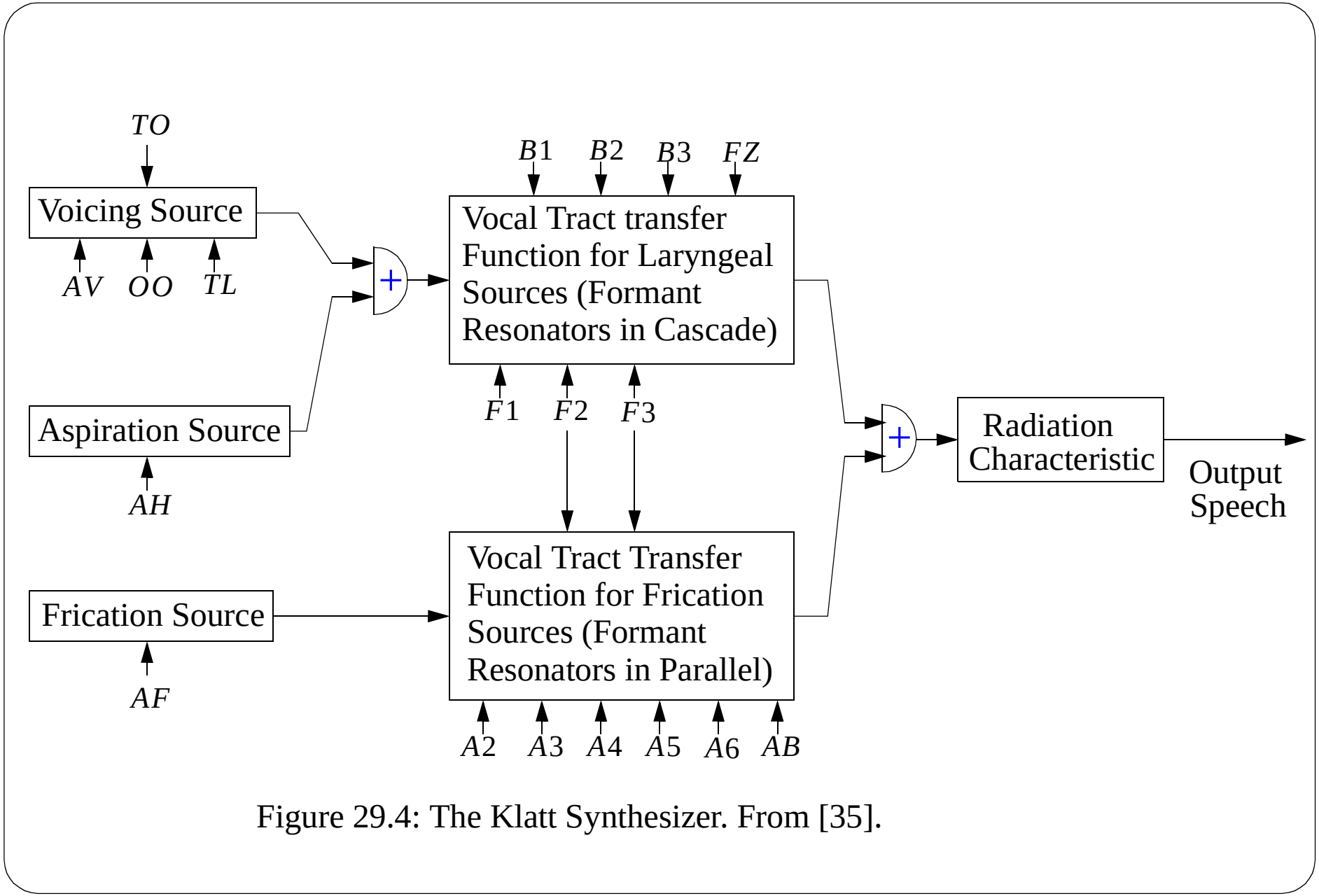


Figure 29.4: The Klatt Synthesizer. From [35].

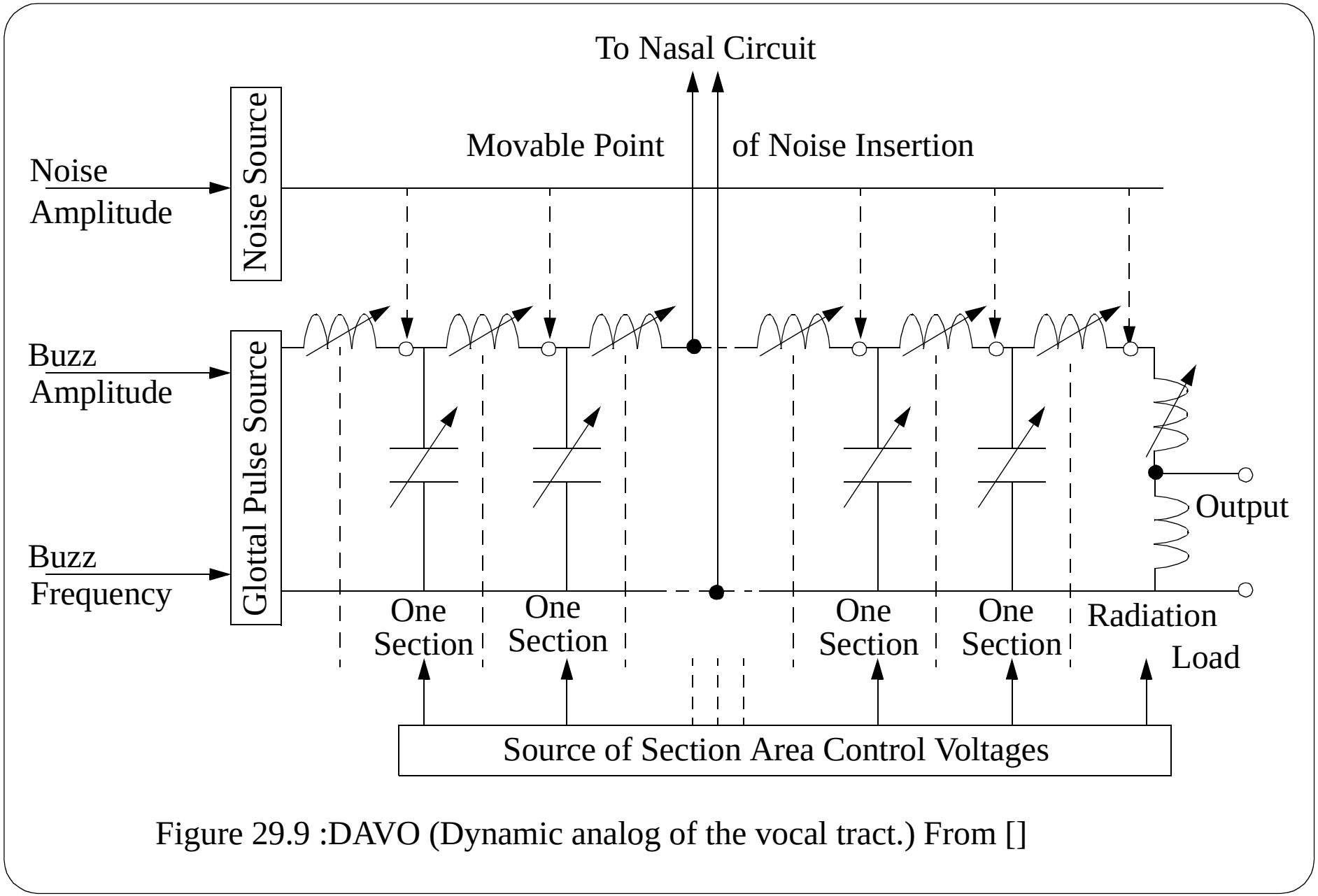
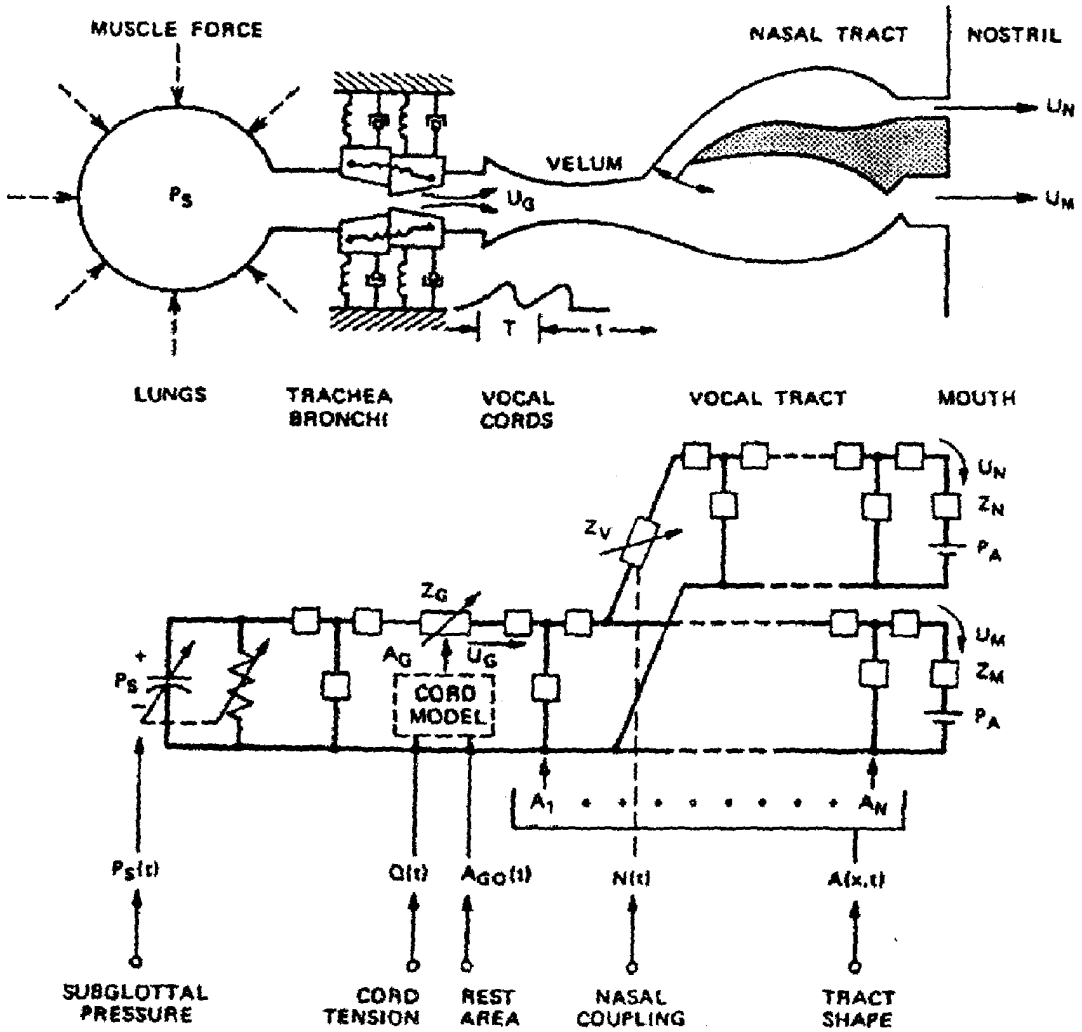


Figure 29.9 :DAVO (Dynamic analog of the vocal tract.) From []

Figure 29.10 : Schematic of the vocal cord-vocal tract system.



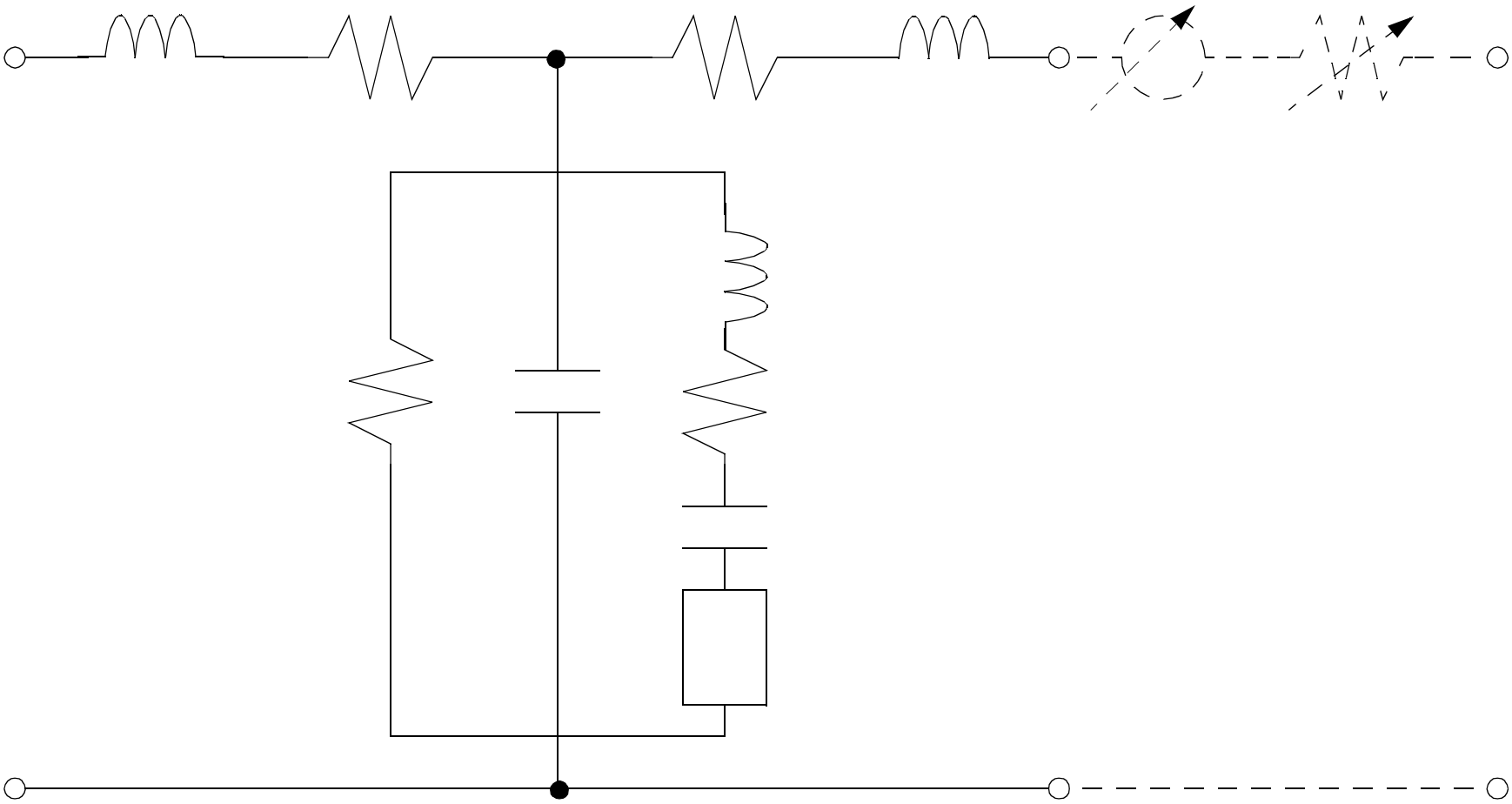
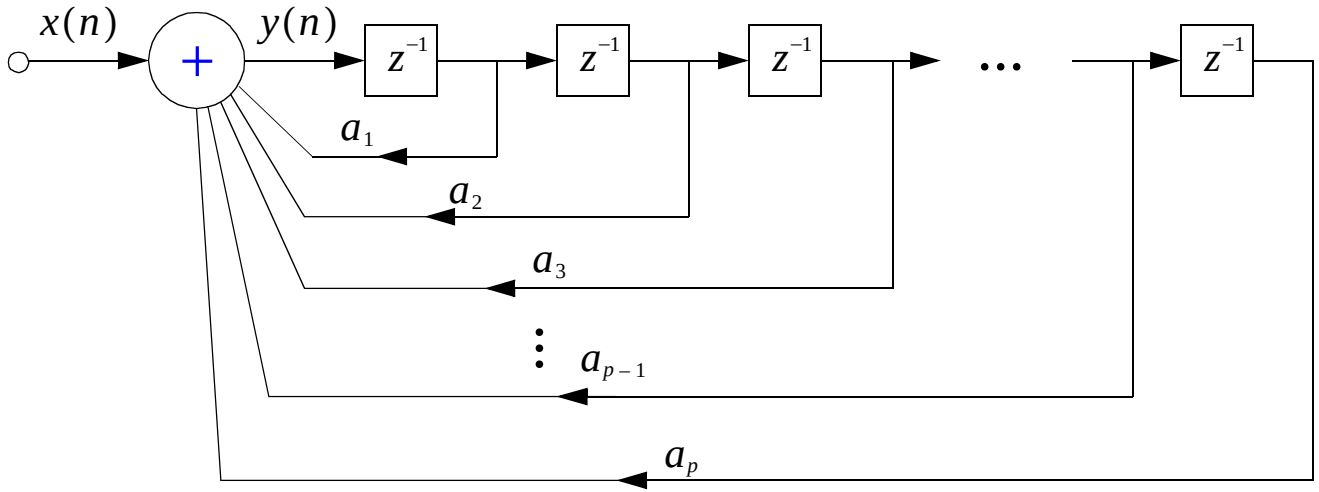
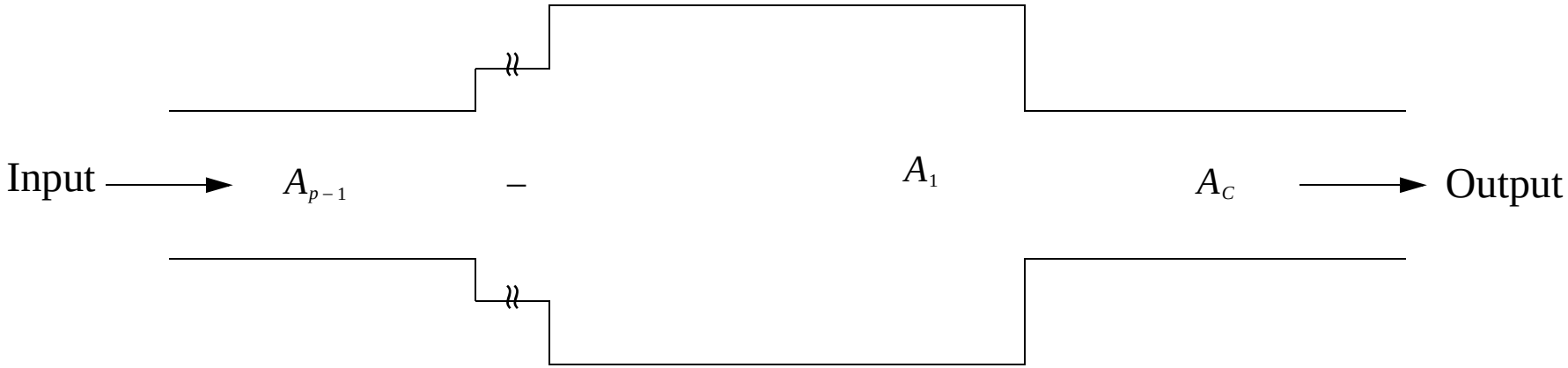


Figure 29.11 : Circuit of an individual T-Section.

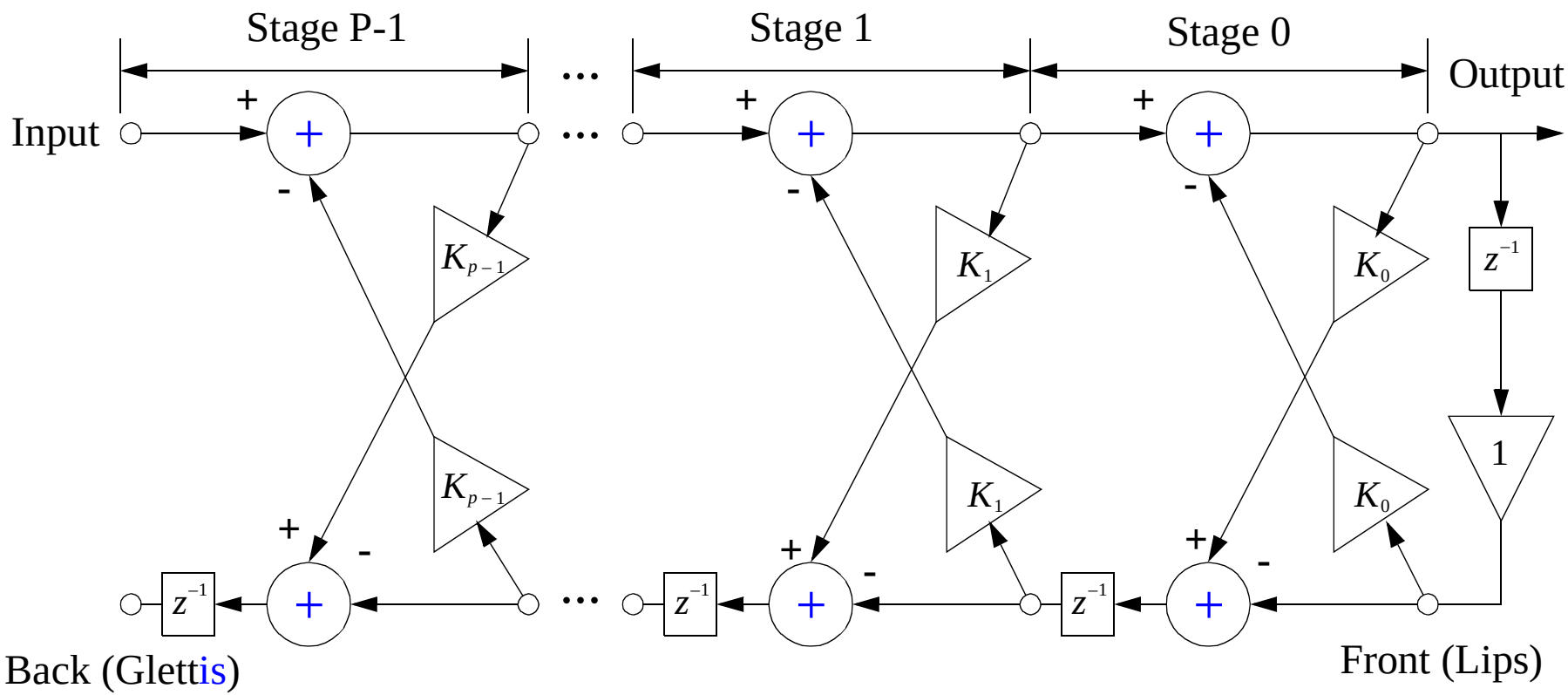


(a) Direct-Form Digital Filter with Variable "a" Coefficients



(b) Acoustic Tube with Variable Area Functions

Figure 29.5 : Two configurations for all pole synthesizers based on LPC analysis.(cont.)



(c) All-Pole lattice Network with Variable “k” Parameters

Figure 29.5 : Two configurations for all pole synthesizers based on LPC analysis.

a) shows a direct form inplementation of the difference equation giving a synthsyzer output as a weighted sum of its past values plus the excitation input. b) shows a model of the acoustic tube with variable cross-sectional area that could give rise to such a characteristic. c) shows an intepretation of this model that suggests a lattice form for the filter.

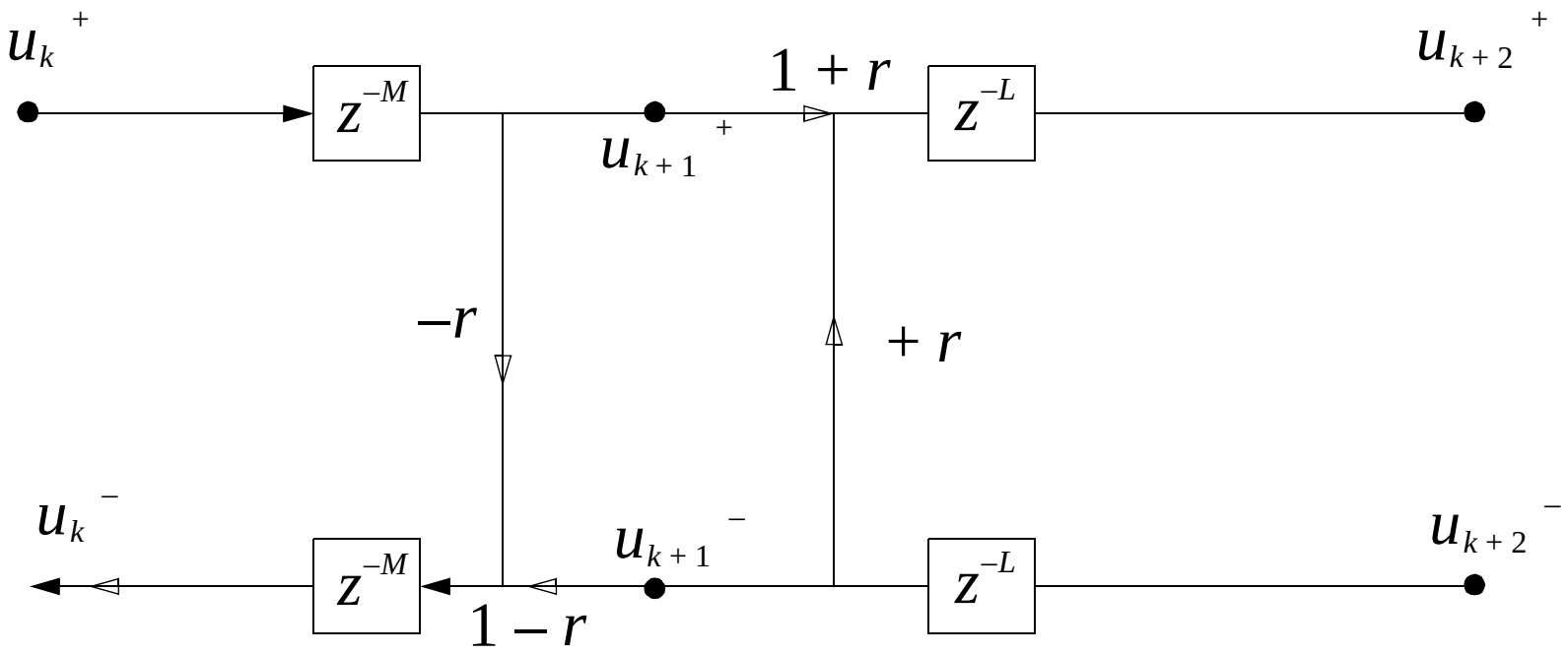


Figure 11.2 : Two section digital wave guide.

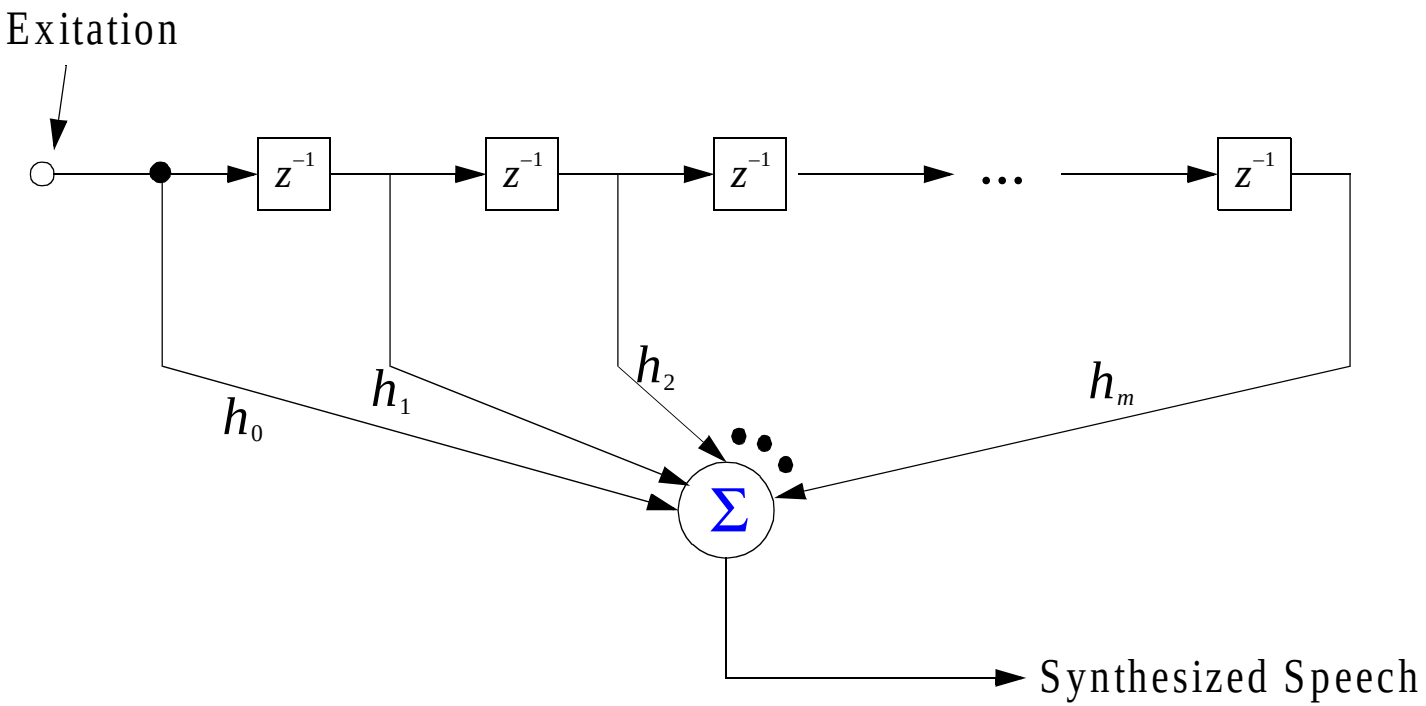


Figure 29.6 : All-Zero synthesizer based on depstral analysis.

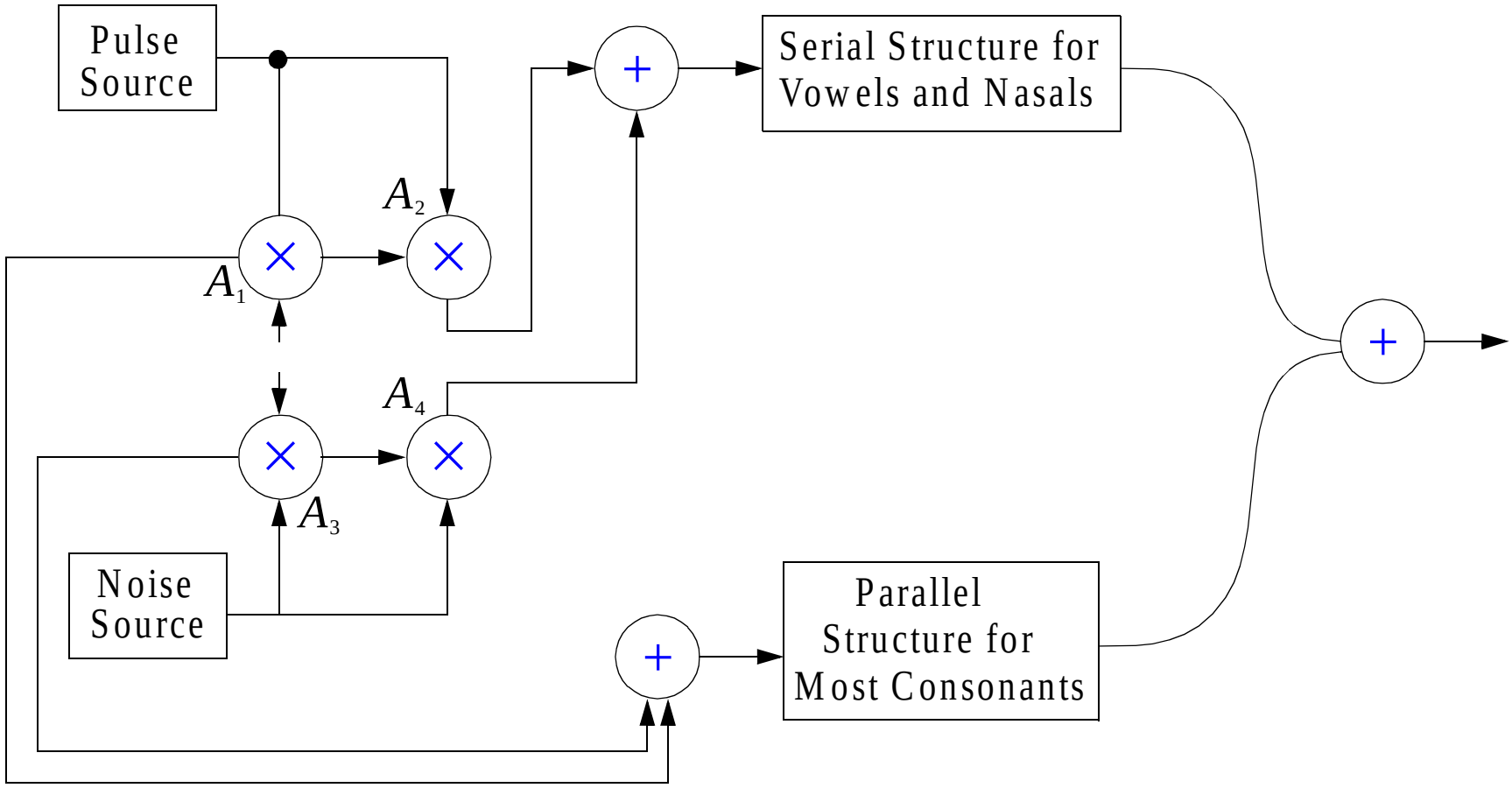


Figure 29.12 : Structure of Klatt Synthesizer.