Chapter 2
Speech Production Model

Abstract The continuous speech signal (air) that comes out of the mouth and the nose is converted into the electrical signal using the microphone. The electrical speech signal thus obtained is sampled to obtain the discrete signals and are stored in the digital system for further processing. This is digital speech processing. The speech signal model is broadly classified as the source-filter model and the probabilistic model. Source-filter model assumes the physical phenomenon for the production of speech signal. Probabilistic model like Hidden Markov Model (HMM), Gaussian Mixture Model (GMM) are the mathematical model that does not care about the physical phenomenon. Speech model is used to extract the feature vectors from the speech signal for isolated speech recognition and the speaker recognition. It is used to compress the speech signal for storage like in Code exited linear prediction (CELP). It is useful for converting text into speech, known as speech synthesis. It is also used for continuous speech recognition. This chapter deals with the source-filter model of speech production.

2.1 Introduction

The air that comes out of the lungs passes through the vocal tract and comes out of the mouth and the nose to obtain the continuous speech signal. The air coming out of lungs are either sent directly to the vocal tractor or altered using the vocal chord vibrations before sending to the vocal tract. The speech signals with vocal chord vibrations are known as voiced speech signals. The speech signals without the vocal chord vibrations are known as unvoiced speech signals. The velum is used to close the nose path,so that the speech signal is coming out only through the mouth. The vocal tract path is adjusted using tongue and velum to produce different speech signal. Thus lung, vocal chord, vocal tract, tongue, velum, mouth and nose are the integral part that produces the speech signal (refer Appendix F).

E. S. Gopi, *Digital Speech Processing Using Matlab*,
Signals and Communication Technology, DOI: 10.1007/978-81-322-1677-3_2,

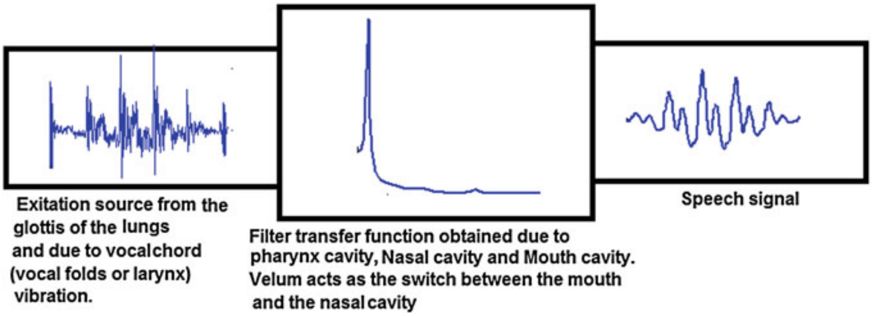

Fig. 2.1 Source-filter model of the speech production

2.2 1-D Sound Waves

The sound waves are longitudinal waves. It produces the disturbance along the direction of the flow (refer Fig. 2.1). The disturbance is in the form of compression and rarefaction. In source-filter model, the source is either the noise (air from the lungs) or the impulse stream (vocal chord vibration with the particular frequency) and the filter is the vocal-tract. The filter is assumed as the cascade connections of the tubes with different cross-sectional area. The length of the tube is usually less than the wavelength of the produced sound wave. Hence speech–sound waves are assumed to travel in one-dimensional direction. This model is known as 1-D sound wave.

2.2.1 Physics on Sound Wave Travelling Through the Tube with Uniform Cross-Sectional Area A

Consider the small segment of the tube (shaded region). When the sound wave crosses the small segment, the change in the physical entities (refer Table 2.1) like force (F), pressure (P), volume flow in terms of volume/s (S), velocity (V) are described below. Let the tube is kept along the direction of X-axis. Let the points I and J (refer Fig. 2.2) are at the distances x and $x + \Delta x$ from the origin. The pressure as point I is represented as P. Hence the pressure at J is computed as follows

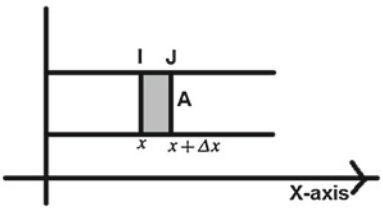

Fig. 2.2 1-D sound wave travelling through the tube with uniform cross-sectional area A

$$P + \frac{\partial P}{\partial x}\Delta x \tag{2.1}$$

The velocity is computed as the volume flow per unit area across the tube ($V = \frac{S}{A}$). Let the velocity at I is given as V. Hence the velocity at J is computed as follows

$$V + \frac{\partial S}{A\partial x}\Delta x \tag{2.2}$$

The volume of the air in the element is computed as $L = A\Delta x$. The rate of change of volume is computed as

$$\frac{\partial L}{\partial t} = \frac{A\partial x}{\partial t} = A\frac{\partial x}{\partial t} \tag{2.3}$$

Note that $\frac{\partial x}{\partial t}$ is the change in the velocity in the element. From (2.2) and (2.3), we get the following

$$\frac{\partial L}{\partial t} = A\frac{\partial S}{A\partial x}\Delta x = \frac{\partial S}{A\partial x}L \tag{2.4}$$

Net force (N_F) in the cross section is obtained as the difference between the force at I and at J.
Force at I is computed as pressure at A $\times$ cross sectional area $= PA$. Force at J is computed as pressure at J $\times$ cross sectional area $= (P + \frac{\partial P}{\partial x}\Delta x)A$. Thus the netforce in the cross section in the x-direction is given as follows.

$$N_F = PA - (P + \frac{\partial P}{\partial x}\Delta x)A = -\frac{\partial P}{\partial x}A\Delta x \tag{2.5}$$

Net force in the cross section is also computed as mass $\times$ accelaration. Also density of the air ρ $\times$ volume of the cross section gives the mass of the air inside the cross section. Recall $V = \frac{S}{A}$ and also note that acceleration is the rate of change of velocity and hence it is computed as $\frac{\partial V}{\partial t} = \frac{\partial S}{A\partial t}$. Hence the netforce is computed as follows

$$N_F = \rho L\frac{\partial S}{A\partial t} \tag{2.6}$$

Equating (2.5), (2.6) and $L = A\Delta x$, we get the following.

$$-\frac{\partial P}{\partial x}A\Delta x = \rho L\frac{\partial S}{A\partial t} \tag{2.7}$$

$$\Rightarrow -\frac{\partial P}{\partial x}A = \rho L\frac{\partial S}{A\Delta x\partial t} \tag{2.8}$$

$$\Rightarrow -\frac{\partial P}{\partial x}A = \rho\frac{\partial S}{\partial t} \tag{2.9}$$

From ideal gas law inside the cross section, we get, $PL = nRT$, where P is the pressure, L is the volume inside the cross section, $n = \frac{\text{mass}}{\text{molecular weight of the air}} = \frac{\text{volume} \times \text{density}}{\text{molecular weight of the air}} = \frac{L\rho}{M}$ is the number of moles, R is the gas constant, T is the temperature in kelvin. From the above discussion, we get the following.

$$PL = \frac{L\rho}{M}RT \Rightarrow P = \frac{\rho}{M}RT \tag{2.10}$$

The square of the speed of the sound depends only on temperature and is given as $c^2 = \gamma \frac{RT}{M}$. Hence,

$$P\gamma = \frac{\rho\gamma}{M}RT = \rho c^2 \tag{2.11}$$

The transfer of sound energy inside the cross section is faster so that we can assume that there is no transfer of heat energy and hence we assume it as the adiabatic process inside the cross section. Hence it obeys $PL^{\gamma} =$ constant. This implies the following

$$PL^{\gamma} = \text{constant} \tag{2.12}$$

$$\Rightarrow \frac{\partial (PL^{\gamma})}{\partial t} = 0 \tag{2.13}$$

$$\Rightarrow P\gamma L^{\gamma-1}\frac{\partial L}{\partial t} + L^{\gamma}\frac{\partial P}{\partial t} \tag{2.14}$$

$$\Rightarrow \frac{\partial L}{\partial t}\frac{P\gamma}{L} + \frac{\partial P}{\partial t} = 0 \tag{2.15}$$

From (2.4), (2.11) and (2.15), we get the following

$$\frac{\partial S}{A\partial x}L\frac{P\gamma}{L} + \frac{\partial P}{\partial t} = 0 \tag{2.16}$$

$$\Rightarrow \frac{\partial S}{A\partial x}P\gamma + \frac{\partial P}{\partial t} = 0 \tag{2.17}$$

$$\Rightarrow \rho c^2\frac{\partial S}{\partial x} = -A\frac{\partial P}{\partial t} \tag{2.18}$$

The sound flow in the segment is described by (2.9) and (2.18).

2.2.2 *Solution to (2.9) and (2.18)*

Differentiating (2.9) with respect to t and (2.18) with respect to x, we get the following

$$-\frac{\partial P}{\partial x}A = \rho\frac{\partial S}{\partial t} \tag{2.19}$$

$$\Rightarrow -\frac{\partial^2 P}{\partial x \partial t} A = \rho \frac{\partial^2 S}{\partial t^2} \tag{2.20}$$

$$\rho c^2 \frac{\partial S}{\partial x} = -A \frac{\partial P}{\partial t} \tag{2.21}$$

$$\Rightarrow \rho c^2 \frac{\partial^2 S}{\partial x^2} = -A \frac{\partial^2 P}{\partial t \partial x} \tag{2.22}$$

Using (2.20) and (2.22), we get the following

$$\rho c^2 \frac{\partial^2 S}{\partial x^2} = \rho \frac{\partial^2 S}{\partial t^2} \tag{2.23}$$

$$\Rightarrow \frac{\partial^2 S}{\partial t^2} = c^2 \frac{\partial^2 S}{\partial x^2} \tag{2.24}$$

Let $u = x + ct$ and $v = x - ct$. $\frac{\partial^2 S}{\partial x^2}$ is computed in terms of u and v as follows

$$\frac{\partial S}{\partial t} = \frac{\partial S}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial S}{\partial v}\frac{\partial v}{\partial t} \tag{2.25}$$

$$\Rightarrow \frac{\partial S}{\partial t} = \frac{\partial S}{\partial u} c + \frac{\partial S}{\partial v}(-c) \tag{2.26}$$

$$\Rightarrow \frac{\partial^2 S}{\partial t^2} = c(\frac{\partial^2 S}{\partial u^2}(c) + \frac{\partial^2 S}{\partial u \partial v}(-c) - \frac{\partial^2 S}{\partial v \partial u}(c) - \frac{\partial^2 S}{\partial v^2}(-c)) \tag{2.27}$$

$$\Rightarrow \frac{\partial^2 S}{\partial t^2} = c^2(\frac{\partial^2 S}{\partial u^2} + \frac{\partial^2 S}{\partial v^2} - 2\frac{\partial^2 S}{\partial u \partial v}) \tag{2.28}$$

Similarly $\frac{\partial^2 S}{\partial x^2}$ is computed as follows

$$\frac{\partial^2 S}{\partial x^2} = (\frac{\partial^2 S}{\partial u^2} + \frac{\partial^2 S}{\partial v^2} + 2\frac{\partial^2 S}{\partial u \partial v}) \tag{2.29}$$

Using (2.28) and (2.29), (2.24) is rewritten as follows

$$c^2(\frac{\partial^2 S}{\partial u^2} + \frac{\partial^2 S}{\partial v^2} + 2\frac{\partial^2 S}{\partial u \partial v}) = c^2(\frac{\partial^2 S}{\partial u^2} + \frac{\partial^2 S}{\partial v^2} - 2\frac{\partial^2 S}{\partial u \partial v}) \tag{2.30}$$

$$\Rightarrow \frac{\partial^2 S}{\partial u \partial v} = 0 \Rightarrow \frac{\partial S}{\partial u} = f(v) \tag{2.31}$$

$$\Rightarrow S(x,t) = g(v) + h(u) = g(x - ct) + h(x + ct) \tag{2.32}$$

Note that f, g and h are arbitrary functions. Represent $g(x - ct)$ as $S^+(t - \frac{x}{c})$ and $g(x + ct)$ as $-S^-(t + \frac{x}{c})$, we get the following

$$S(x,t) = S^+(t - \frac{x}{c}) - S^-(t + \frac{x}{c}) \tag{2.33}$$

Using (2.18) we get the pressure equation as follows

$$\frac{\partial S}{\partial x} = (-\frac{1}{c})(S'^{+}(t - \frac{x}{c}) + S'^{-}(t + \frac{x}{c})) \tag{2.34}$$

$$\Rightarrow \frac{\partial P}{\partial t} = c\rho(S'^{+}(t - \frac{x}{c}) + S'^{-}(t + \frac{x}{c})) \tag{2.35}$$

$$\Rightarrow P(x,t) = \frac{\rho c}{A}(S^{+}(t - \frac{x}{c}) + S^{-}(t + \frac{x}{c})) \tag{2.36}$$

Note that $S^{+}(t-\frac{x}{c})$ is the volume flow in the positive direction of x-axis and $S^{-}(t-\frac{x}{c})$ is the volume flow in the negative direction of x-axis. Thus the netflow in the positive direction is given as (2.33). Also the pressure at (t, x) is the constant times absolute sum of volume flow in both the directions as given in (2.36).

2.3 Vocal Tract Model as the Cascade Connections of Identical Length Tubes with Different Cross-Sections

Consider that the vocal tract is modelled as the cascade of three identical length (L) tubes with the cross sectional areas as A_1, A_2 and A_3 respectively (refer Fig. 2.3). The inlet volume flow in the forward direction of the ith tube is represented as P_i. The outlet volume flow in the forward direction of the ith tube is represented as R_i. Similarly the inlet and outlet volume flow in the reverse direction of the ith tube is represented as S_i and Q_i respectively. The relationship between P_i, Q_i, R_i, S_i are given as follows

$$P_i(t) = R_i(t + \tau) \tag{2.37}$$

$$\Rightarrow R_i(t) = P_i(t - \tau) S_i(t) = Q_i(t + \tau) \tag{2.38}$$

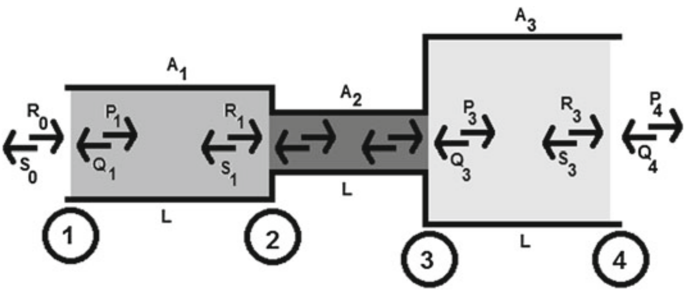

Fig. 2.3 Cascade of three tubes with different cross-sectional areas as the model of the vocal tract

where τ is the delay. The delay is the time required for the sound wave to travel through the single tube of length L, which is computed as $\tau = L/c$. If the signal is sampled with sampling time T_s, we get the following

$$R_i(nT_s) = P_i(nT_s - \tau) \tag{2.39}$$

$$S_i(nT_s) = Q_i(nT_s + \tau) \tag{2.40}$$

If $\tau = \frac{T_s}{2}$ and representing the (2.39) and (2.40) in discrete form, we get the following

$$R_i(n) = P_i(n - \frac{1}{2}) \tag{2.41}$$

$$S_i(n) = Q_i(n + \frac{1}{2}) \tag{2.42}$$

Representing in z-domain, we get the following

$$R_i(Z) = P_i(Z)Z^{\frac{1}{2}} \tag{2.43}$$

$$S_i(Z) = Q_i(Z)Z^{-\frac{1}{2}} \tag{2.44}$$

Let the input vector of the ith segment is represented as $I_{Si} = [P_i \ Q_{Si}]^T$ and the output vector of the segment is represented as $O_{Si} = [R_i \ S_i]^T$. They are related with the matrix $[M_{Si}]$ as $[O_{Si}] = [M_{Si}][I_{Si}]$, where M_{Si} is given as follows. Representing in the matrix form in z-domain, we get the following $\begin{bmatrix} Z^{\frac{1}{2}} & 0 \\ 0 & Z^{-\frac{1}{2}} \end{bmatrix}$. Let the input vector of the ith junction is represented as $I_{Ji} = [R_{i-1} \ S_{i-1}]^T$ and the output vector of the segment is represented as $O_{Ji} = [P_i \ Q_i]^T$ (refer Fig. 2.3). They are related using the matrix M_{Ji} as $[O_{Ji}] = [M_{Ji}][I_{Ji}]$. M_{Ji} is computed as follows. At the ith junction, there is the continuation in the pressure and the volume flow as mentioned below.

- Volume flow continuity

$$R_{i-1} - S_{i-1} = P_i - Q_i \tag{2.45}$$

- Pressure continuity

$$\frac{\rho c}{A_{i-1}}(R_{i-1} + S_{i-1}) = \frac{\rho c}{A_i}(P_i + Q_i) \tag{2.46}$$

$$\Rightarrow \frac{(R_{i-1} + S_{i-1})}{A_{i-1}} = \frac{(P_i + Q_i)}{A_i} \tag{2.47}$$

$$\Rightarrow (R_{i-1} + S_{i-1})(A_i) = (P_i + Q_i)(A_{i-1}) \tag{2.48}$$

Multiplying (2.45) with A_i we get the following

$$(R_{i-1} - S_{i-1})A_i = (P_i - Q_i)A_i \tag{2.49}$$

Adding (2.48) and (2.49), we get the following

$$2R_{i-1}A_i = P_i(A_i + A_{i-1}) + Q_i(A_i - A_{i-1}) \tag{2.50}$$

$$\Rightarrow R_{i-1} = P_i\frac{(A_i + A_{i-1})}{2A_i} + Q_i\frac{(A_i - A_{i-1})}{2A_i} \tag{2.51}$$

Subtracting (2.48) and (2.49), we get the following

$$2S_{i-1}A_i = P_i(A_{i-1} - A_i) + Q_i(A_{i-1} + A_i) \tag{2.52}$$

$$\Rightarrow S_{i-1} = P_i\frac{(A_{i-1} - A_i)}{2A_i} + Q_i\frac{(A_{i-1} + A_i)}{2A_i} \tag{2.53}$$

Let $r_i = \frac{A_i - A_{i-1}}{A_i + A_{i+1}}$ and hence

$$\frac{A_{i-1}}{A_i} = \frac{1 - r_i}{1 + r_i} \tag{2.54}$$

Using (2.54), we get the following

$$\frac{(A_i + A_{i-1})}{2A_i} = \frac{1}{2}(1 + \frac{A_{i-1}}{A_i}) = \frac{1}{1 + r_i} \tag{2.55}$$

$$\frac{(A_{i-1} - A_i)}{2A_i} = \frac{1}{2}(\frac{A_{i-1}}{A_i} - 1) = \frac{-r_i}{1 + r_i} \tag{2.56}$$

Thus R_{i-1} and S_{i-1} are expressed interms of r as follows

$$R_{i-1} = P_i\frac{1}{1 + r_i} - Q_i\frac{-r_i}{1 + r_i} \tag{2.57}$$

$$S_{i-1} = P_i\frac{-r_i}{1 + r_i} + Q_i\frac{1}{1 + r_i} \tag{2.58}$$

Thus the matrix M_{Ji} is given as $\frac{1}{1+r_i}\begin{bmatrix} 1 & -r_i \\ -r_i & 1 \end{bmatrix}$. It is noted that the matrix M_{Ji} is identical in z-domain also. The transfer function of the system is given as $\frac{P_4(Z)}{R_0(Z)}$. This is computed using the relationship between the vector $I_0(Z)$ and $O_4(Z)$ as $O_4(Z) = M_{J1}(Z)M_{S1}(Z)M_{J2}(Z)M_{S2}(Z)M_{J3}(Z)M_{S3}M_{J4}(Z)I_0(Z)$, which is computed as follows.

$$\begin{bmatrix} R_0(Z) \\ S_0(Z) \end{bmatrix} = \frac{1}{1+r_1}\begin{bmatrix} 1 & -r_1 \\ -r_1 & 1 \end{bmatrix}\begin{bmatrix} Z^{\frac{1}{2}} & 0 \\ 0 & Z^{-\frac{1}{2}} \end{bmatrix}\frac{1}{1+r_2}\begin{bmatrix} 1 & -r_2 \\ -r_2 & 1 \end{bmatrix}\begin{bmatrix} Z^{\frac{1}{2}} & 0 \\ 0 & Z^{-\frac{1}{2}} \end{bmatrix}$$
$$\frac{1}{1+r_3}\begin{bmatrix} 1 & -r_3 \\ -r_3 & 1 \end{bmatrix}\begin{bmatrix} Z^{\frac{1}{2}} & 0 \\ 0 & Z^{-\frac{1}{2}} \end{bmatrix}\frac{1}{1+r_4}\begin{bmatrix} 1 & -r_4 \\ -r_4 & 1 \end{bmatrix}\begin{bmatrix} P_4(Z) \\ Q_4(Z) \end{bmatrix}$$

Note that R_0 is coming from the lung openings and P_4 is coming out of mouth and nose. As there is no feedback in the mouth opening during speech, Q_4 is equated to zero and solving the $\frac{P_4(Z)}{R_0(Z)}$ gives the transfer function. Note that r_i is defined as the reflection co-efficient of ith segment. The values for r_i ranges from -1 to 1. It is also noted that A_0 is the area of the opening from the lungs to the vocal chord (glottis opening) and A_4 is assumed to be large finite value. On simplification, we get the following

$$\begin{bmatrix} R_0(Z) \\ S_0(Z) \end{bmatrix} = \frac{1}{(1+r_1)(1+r_2)(1+r_3)(1+r_4)} \begin{bmatrix} Z^{\frac{1}{2}} & -r_1 Z^{-\frac{1}{2}} \\ -r_1 Z^{\frac{1}{2}} & Z^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} Z^{\frac{1}{2}} & -r_2 Z^{-\frac{1}{2}} \\ -r_2 Z^{\frac{1}{2}} & Z^{-\frac{1}{2}} \end{bmatrix}$$
$$\begin{bmatrix} Z^{\frac{1}{2}} & -r_3 Z^{-\frac{1}{2}} \\ -r_3 Z^{\frac{1}{2}} & Z^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} Z^{\frac{1}{2}} & -r_4 Z^{-\frac{1}{2}} \\ -r_4 Z^{\frac{1}{2}} & Z^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} Z^{-\frac{1}{2}} & 0 \\ 0 & Z^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} P_4(Z) \\ 0 \end{bmatrix}$$

On further simplification, we get the following

$$\begin{bmatrix} R_0(Z) \\ S_0(Z) \end{bmatrix} = \frac{1}{(1+r_1)(1+r_2)(1+r_3)(1+r_4)} \begin{bmatrix} Z + r_1 r_2 & -r_2 - r_1 Z^{-1} \\ -r_1 Z - r_2 & r_1 r_2 + Z^{-1} \end{bmatrix}$$
$$\begin{bmatrix} Z + r_3 r_4 & -r_3 - r_4 Z^{-1} \\ -r_3 Z - r_4 & r_3 r_4 + Z^{-1} \end{bmatrix} \begin{bmatrix} Z^{-\frac{1}{2}} & 0 \\ 0 & Z^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} P_4(Z) \\ 0 \end{bmatrix}$$

Thus the transfer function of the vocal tract is given as follows

$$\frac{P_4(Z)}{R_0(Z)} = \frac{Z^{-\frac{1}{2}}}{(Z + r_1 r_2)(Z + r_3 r_4) + (r_2 + r_1 Z^{-1})(r_4 + r_3 Z)} \tag{2.59}$$

$$= \frac{Z^{-\frac{1}{2}}}{Z^2 + (r_1 r_2 + r_3 r_4 + r_2 r_3)Z + r_1 r_2 r_3 r_4 + r_2 r_4 + r_1 r_3 + r_1 r_4 Z^{-1}} \tag{2.60}$$

$$= \frac{Z^{-\frac{3}{2}}}{1 + (r_1 r_2 + r_3 r_4 + r_2 r_3)Z^{-1} + +r_1 r_4 Z^{-2} + r_1 r_2 r_3 r_4 + r_2 r_4 + r_1 r_3} \tag{2.61}$$

Note that the factor $Z^{-\frac{3}{2}}$ is due to the delay introduced by the three segments. The transfer function of the vocal tract is identified as the ALL POLE third order filter. In general if vocal chord is assumed to have r segments, we get the transfer function becomes the rth order all pole filter with the delay factor of $Z^{-r/2}$. Thus the generalized transfer function of the rth order vocal tract filter is given as follows

$$V(Z) = \frac{Z^{-\frac{r}{2}}}{1 - \sum_{k=1}^{k=r} a_k Z^{-k}} \tag{2.62}$$

The length of the vocal tract is approximately 15 cm for adults. If the sampling frequency is $F_s = 8000\,Hz$, the length of the each segment is given as $\frac{c}{2F_s} = \frac{340}{16000} = 0.02125\,\mathrm{m}$. Hence number of segments are usually assumed as $\frac{0.15}{0.02125} \cong 7$. Hence the order of the filter is assumed around 7 for the sampling frequency of 8000 Hz.

2.4 Modelling the Vocal Tract from the Speech Signal

The sound wave that comes out from the lungs is the noise, which passes through the vocal tract filter to produce the particular speech signal. This type of speech signal is known as unvoiced speech signal. The sound wave that is produced by the vocal chord gets mixed with the wave that comes out of the lungs is passed through the vocal tract filter to produce the particular speech signal. This type of speech signal is known as unvoiced speech signal. In both the cases, the speech signal is modelled as the convolution of the sound source with the vocal tract filter. In Z-domain, speech signal is the product of the Z-transformation of the source signal with the Z-Transformation of the vocal tract. Let the source signal is represented as $I(Z)$ and output speech signal $S(Z)$ and are related as $\frac{S(Z)}{I(Z)} = V(Z) = \frac{Z^{-\frac{n}{2}}}{1-\sum_{k=1}^{k=n} a_k Z^{-k}}$. Rewriting the expression without delay we get the following.

$$\frac{S(Z)}{I(Z)} = \frac{1}{1-\sum_{k=1}^{k=n} a_k Z^{-k}} \Rightarrow S(n) = I(n) + \sum_{k=1}^{k=r} a_k S(n-k) \tag{2.63}$$

As the amplitude of the input signal is negligible, we can approximate $S(n)$ as $S(n) \cong \sum_{k=1}^{k=r} a_k S(n-k)$. This equation is known as prediction equation because nth sample of the speech signal is predicted using the past r samples of the identical speech signal. The co-efficiets $a_k \forall k = 1 \cdots n$ are known as Linear Predictive Co-efficients (LPC). The LPC completely describes the vocal tract filter. These are obtained from the speech signal using the following techniques.

2.4.1 *Autocorrelation Method*

The LPC's are obtained such that

$$E((S(n) - \sum_{k=1}^{k=r} a_k S(n-k))^2) \tag{2.64}$$

is minimized. In this E is the expectation operator and $(S(n) - \sum_{k=1}^{k=r} a_k S(n-k))^2$ is the squared error obtained in predicting the nth sample of the speech signal using the past r samples. minimizing (2.64) is achieved by partial differentiating the (2.64) with respect to unknown variables $a_j \forall j = 1 \cdots r$ and equate to zeros as mentioned below

$$\frac{\partial E((S(n) - \sum_{k=1}^{k=r} a_k S(n-k))^2)}{\partial a_j} \tag{2.65}$$

$$\Rightarrow E(2(S(n) - \sum_{k=1}^{k=r} a_k S(n-k))S(n-j)) = 0 \tag{2.66}$$

$$\Rightarrow R_S(j) = \sum_{k=1}^{k=r} a_k R_S(j-k) \tag{2.67}$$

$R_S(j)$ is the autocorrelation of the speech signal. The speech signal under consideration for the particular duration is assumed to be Wide Sense Stationary (W.S.S) and hence the autocorrelation depends on the difference of the index. As speech signal is the real signal and W.S.S, autocorrelation is symmetric function. This technique is known as autocorrelation method.

2.4.1.1 Solving (2.67) Using Levinson–Durbin Algorithm

The auto correlation $R_S(j)$ is computed as $E(S(n)S(n-j))$. Consider $S(n)$ is the random variable obtained by sampling across the random process S at the time instant n and $S(n-j)$ is the random variable obtained by sampling across the random process S at the time instant $n-j$. To compute $E(S(n)S(n-j))$, we need the joint probability density function $f_{S(n)S(n-j)}(\alpha, \beta)$. This is not available in practice. Hence the $R_S(j)$ is estimated from the sample speech signal itself. The estimation is done along the process assuming that the speech signal is ergodic in autocorrelation as follows

$$R_S(j) = \sum_{n=-\infty}^{n=\infty} S(n)S(n-j) \tag{2.68}$$

In practice, the computation is done for the longer duration of above 20 ms. The (2.68) is written in the matrix form for $r = 5$ as follows

$$\begin{bmatrix} R_S(1) \\ R_S(2) \\ R_S(3) \\ R_S(4) \\ R_S(5) \end{bmatrix} = \begin{bmatrix} R_S(0) & R_S(-1) & R_S(-2) & R_S(-3) & R_S(-4) \\ R_S(1) & R_S(0) & R_S(-1) & R_S(-2) & R_S(-3) \\ R_S(2) & R_S(1) & R_S(0) & R_S(-1) & R_S(-2) \\ R_S(3) & R_S(2) & R_S(1) & R_S(0) & R_S(-1) \\ R_S(4) & R_S(3) & R_S(2) & R_S(1) & R_S(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}$$

Due to symmetric nature of autocorrelation function, the equation in the matrix form is rewritten with $R_s(-n) = R_s(n)$.

Fig. 2.4 Matrix highlighting the toeplitz structure

$$\begin{bmatrix} R_S(0) & R_S(1) & R_S(2) & R_S(3) & R_S(4) \\ R_S(1) & R_S(0) & R_S(1) & R_S(2) & R_S(3) \\ R_S(2) & R_S(1) & R_S(0) & R_S(1) & R_S(2) \\ R_S(3) & R_S(2) & R_S(1) & R_S(0) & R_S(1) \\ R_S(4) & R_S(3) & R_S(2) & R_S(1) & R_S(0) \end{bmatrix}$$

The autocorrelation matrix thus obtained is the toeplitz matrix because, it is the symmetric matrix with the identical diagonal elements. (refer Fig. 2.4).

2.4.1.2 Levinson–Durbin Algorithm

Consider the vocal tract with 5 Linear predictive co-efficients (5th order LPC) (represented as $\overline{x_5^4} = [x_5(0)\ x_5(1)\ x_5(2)\ x_5(3)\ x_5(4)]^T$) are obtained by solving the equation mentioned in Fig. 2.5. If the vocal tract is modelled with 4 LPC (4th order LPC), the co-efficients $\overline{x_4^3} = [x_4(0)\ x_4(1)\ x_4(2)\ x_4(3)]$ are obtained by solving the equation mentioned in Fig. 2.6. The key idea in Levinson–Durbin algorithm is to obtain the 5th order LPC from the 4th order LPC. They are related as follows. Let $[c_5(0)c_5(1)c_5(2)c_5(3)c_5(4)]^T$ is the correction vector. Note that $x_5(4) = c_4(4)$.

$$\begin{bmatrix} x_5(0) \\ x_5(1) \\ x_5(2) \\ x_5(3) \\ x_5(4) \end{bmatrix} = \begin{bmatrix} x_4(0) \\ x_4(1) \\ x_4(2) \\ x_4(3) \\ 0 \end{bmatrix} + \begin{bmatrix} c_4(0) \\ c_4(1) \\ c_4(2) \\ c_4(3) \\ c_4(4) \end{bmatrix} \tag{2.69}$$

Representing the equation in Fig. 2.5 using (2.69), we get the following

Fig. 2.5 Equation for obtaining 5 LPC

$$\left[\begin{array}{cccc|c} a_0 & a_1 & a_2 & a_3 & a_4 \\ a_1 & a_0 & a_1 & a_2 & a_3 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ a_3 & a_2 & a_1 & a_0 & a_1 \\ \hline a_4 & a_3 & a_2 & a_1 & a_0 \end{array}\right] \left[\begin{array}{c} x_5(0) \\ x_5(1) \\ x_5(2) \\ x_5(3) \\ \hline x_5(4) \end{array}\right] = \left[\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \hline a_5 \end{array}\right]$$

$$\begin{bmatrix} A_3 & \overline{a_4^r} \\ \overline{a_4^r}^T & a_0 \end{bmatrix} \begin{bmatrix} \overline{x_5^3} \\ x_5(4) \end{bmatrix} = \begin{bmatrix} \overline{a_4} \\ a_5 \end{bmatrix}$$

Fig. 2.6 Equation for obtaining 4 LPC

$$\left[\begin{array}{ccc|c} a_0 & a_1 & a_2 & a_3 \\ a_1 & a_0 & a_1 & a_2 \\ a_2 & a_1 & a_0 & a_1 \\ \hline a_3 & a_2 & a_1 & a_0 \end{array}\right] \left[\begin{array}{c} x_4(0) \\ x_4(1) \\ x_4(2) \\ \hline x_4(3) \end{array}\right] = \left[\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ \hline a_4 \end{array}\right]$$

$$\begin{bmatrix} A_2 & \overline{a_3^r} \\ \overline{a_3^r}^T & a_0 \end{bmatrix} \begin{bmatrix} \overline{x_4^2} \\ x_4(3) \end{bmatrix} = \begin{bmatrix} \overline{a_3} \\ a_4 \end{bmatrix}$$

$$\begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 \\ a_1 & a_0 & a_1 & a_2 & a_3 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ a_3 & a_2 & a_1 & a_0 & a_1 \\ a_4 & a_3 & a_2 & a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_4(0) + c_4(0) \\ x_4(1) + c_4(1) \\ x_4(2) + c_4(2) \\ x_4(3) + c_4(3) \\ 0 + c_4(4) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \tag{2.70}$$

Using the notations used in Fig. 2.5, (2.70) is represented as the following. Also let $c_4(4) = k_4$.

$$A_3\overline{x_4^3} + A_3\overline{c_4^3} + \overline{a_4^r}c_4(4) = \overline{a_4} \tag{2.71}$$

It is noted $A_3\overline{x_4^3} = \overline{a_4}$ and hence

$$A_3\overline{c_4^3} = -\overline{a_4^r}k_4 \tag{2.72}$$

$$\Rightarrow \overline{c_4^3} = -A_3^{-1}\overline{a_4^r}k_4 \tag{2.73}$$

It is also noted the following from the Fig. 2.5.

$$\overline{a_4^r}^T \overline{x_4^3} + \overline{a_4^r}^T \overline{c_4^3} + a_0k_4 = a_5 \tag{2.74}$$

$$\Rightarrow \overline{a_4^r}^T \overline{x_4^3} + \overline{x_4^{3r}}^T A_3^T \overline{c_4^3} + a_0k_4 = a_5 \tag{2.75}$$

$$\Rightarrow \overline{a_4^r}^T \overline{x_4^3} + \overline{x_4^{3r}}^T A_3 c_4^3 + a_0k_4 = a_5 \tag{2.76}$$

Table 2.1 List of notations

Symbol	Notations
A	Area of cross section (m^2)
P	Pressure (Kg/ms^2)
S	Volume flow (m^3/s)
V	Velocity of the sound wave (m/s)
L	Volume (m^3)
ρ	Density of the air
m	Mass of the air
M	Molecular mass of the air
R	Gas constant
T	Temperature
c	Speed of the air
γ	Adiabatic constant
x	Distance from the origin on the x-axis (m)
t	Time (s)

Using (2.73), we get the following

$$\Rightarrow \overline{a_4^r}^T \overline{x_4^3} - \overline{x_4^{3r}}^T \overline{a_4^r} k_4 + a_0 k_4 = a_5 \tag{2.77}$$

$$k_4 = \frac{a_5 - \overline{a_4^r}^T \overline{x_4^3}}{a_0 - \overline{x_4^{3r}}^T \overline{a_4^r}} \tag{2.78}$$

Thus the Levinson–Durbin algorithm is described by (2.70), (2.73) and (2.78) and steps involved are summarized as follows

1. $a_0 x_1(0) = a_1, \Rightarrow x_1(0) = \overline{x_1^0} = \frac{a_1}{a_0}$
2. $\begin{bmatrix} a_0 & a_1 \\ a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_2(0) \\ x_2(1) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$
   $\begin{bmatrix} x_2(0) \\ x_2(1) \end{bmatrix} = \begin{bmatrix} x_1(0) \\ 0 \end{bmatrix} + \begin{bmatrix} c_1(0) \\ c_1(1) \end{bmatrix}$
3. Compute $k_1 = c_1(1) = \frac{a_2 - \overline{a_1^r}^T \overline{x_1^0}}{a_0 - \overline{x_1^{0r}}^T \overline{a_1^r}^T} = \frac{a_2 - a_1 x_1(0)}{a_0 - x_1(0) a_1}$
4. Compute $\overline{c_1^0} = -A_0^{-1} \overline{a_1^r} k_1 = -a_0^{-1} a_1 k_1$
5. Compute $\overline{x_2^1} = \begin{bmatrix} x_2(0) \\ x_2(1) \end{bmatrix} = \begin{bmatrix} x_1(0) \\ 0 \end{bmatrix} + \begin{bmatrix} c_1(0) \\ c_1(1) \end{bmatrix}$
6. $\begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & a_0 & a_1 \\ a_2 & a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_2(0) \\ x_2(1) \\ x_2(2) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$
   $\begin{bmatrix} x_3(0) \\ x_3(1) \\ x_3(2) \end{bmatrix} = \begin{bmatrix} x_2(0) \\ x_2(1) \\ 0 \end{bmatrix} + \begin{bmatrix} c_2(0) \\ c_2(1) \\ c_2(2) \end{bmatrix}$
7. Compute $k_2 = c_2(2) = \frac{a_3 - \overline{a_2^r}^T \overline{x_2^1}}{a_0 - \overline{x_2^{1r}}^T \overline{a_2^r}^T}$
8. Compute $\overline{c_2^1} = -A_1^{-1} \overline{a_2^r} k_2$

9 Compute $\overline{x_3^2}$

10 In general, compute $k_i = c_i(i) = \frac{a_{i+1} - \overline{a_i^r}^T \overline{x_i^{i-1}}}{a_0 - \overline{x_i^{(i-1)r}}^T \overline{a_i^r}^T}$

11. Compute, $\overline{c_i^{i-1}} = -A_i^{-1} \overline{a_i^r} k_i$
12. Compute $\overline{x_i^{i-1}}$
13. Repeat the steps 10, 11, 12 for $i = 1 \cdots n$ to obtain the nth order LPC $\overline{x_n^{n-1}}$.

2.4.2 Auto Covariance Method

The LPC obtained using the Autocorrelation method needs long duration of speech signal(>20ms) and hence the vocal tract model is not very accurate. But the computation time to obtain the LPC has been reduced by using Levinson–Durbin algorithm. More accurate vocal tract model is obtained for every 2ms speech signal data. This is obtained using covariance method as described below. The (2.66) is rewritten again for clarity using expectation operator as follows

$$E(2(S(n) - \sum_{k=1}^{k=r} a_k S(n-k))S(n-j)) = 0 \tag{2.79}$$

$$\Rightarrow E(S(n)S(n-j) = \sum_{k=1}^{k=r} a_k E(S(n-k)S(n-j)) \tag{2.80}$$

In autocorrelation method, $E(S(n-k)S(n-j))$ is computed as $\sum_{n=-\infty}^{n=\infty} S(n-k)S(n-j)$ (In practice for the long duration of greater than 20 ms) and hence can be represented as $R_S(j-k)$. But in case of auto covariance method, $E(S(n-k)S(n-j))$ is computed as $\sum_{n=0}^{n=L-1} S(n-k)S(n-j) = C_{kj} = \sum_{n=0}^{n=L-1} S(n-j)S(n-k) = C_{jk}$. Hence the LPC using the co-variance method is computed by solving (2.81).

$$C_{0j} = \sum_{k=1}^{k=r} a_k C_{kj} \tag{2.81}$$

The (2.81) for $r = 5$ is represented as follows

$$\begin{bmatrix} C_{01} \\ C_{02} \\ C_{03} \\ C_{04} \\ C_{05} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \tag{2.82}$$

Note that the matrix in (2.82) is the symmetric matrix, but not the toeplitz matrix. Hence Levison–Durbin cannot be used to solve the (2.81). It is also noted that the diagonal elements are greater than the other elements of the matrix and hence the matrix is positive-semi-definite matrix. Hence this can be solved using diagonalization of the matrix or Gauss-elimination method. Note that positive-semi definite symmetric matrix is always diagonalizable (refer Appendix C) with non-negative eigenvalues as the diagonal elements of the diagonal matrix. The computation time required to solve (2.81) is greater than the time required to solve (2.67).

2.5 Lattice Structure to Obtain Excitation Source for the Typical Speech Signal

The transfer function of the vocal tract is modelled as Nth order all pole filter which is represented as $V(Z)$.

$$V(Z) = \frac{1}{1 + \sum_{k=1}^{k=N} a_k z^{-k}} \tag{2.83}$$

If the $E(Z)$ is the excitation source signal and $S(Z)$ is the output signal, they are related as $S(Z) = E(Z)V(Z)$. In discrete domain they are related as $e(n) = s(n) + \sum_{k=1}^{k=N-1} a_k s(n-k)$. The excitation source $e(n)$ needs the FIR filter coefficients a_k. This can also be realized using lattice structure. The excitation source computed for mth order filter is obtained directly from the excitation source for $(m-1)$th filter. Hence fixing up the order of the model becomes easier in real time in modelling the vocal tract filter. The lattice structure for the $1st$ order filter is as given in Fig. 2.8. Let $f_0(n) = g_0(n) = s(n)$ and $f_1(n) = e^1(n)$. Note that $e^i(n)$ is the nth sample of the excitation source with ith order model. They are related as follows

$$f_1(n) = f_0(n) + k_1 g_0(n-1),\ g_1(n) = k_1 f_0(n) + g_0(n-1) \tag{2.84}$$

The relationship using the first order filter is given as follows

$$e^1(n) = s(n) + a_1^1 s(n-1) \tag{2.85}$$

Comparing (2.84) and (2.85), we get $k_1 = a_1^1$. For the second order filter, we get the following

$$f_2(n) = f_1(n) + k_2 g_1(n-1) \tag{2.86}$$

$$\Rightarrow f_2(n) = f_0(n) + k_1 g_0(n-1) + k_2(g_0(n-2) + k_1 f_0(n-1)) \tag{2.87}$$

$$= s(n) + k_1 s(n-1) + k_2 s(n-2) + k_1 k_2 s(n-1) \tag{2.88}$$

$$g_2(n) = g_1(n-1) + k_2 f_1(n) \tag{2.89}$$

The relationship using the second order filter is given as follows

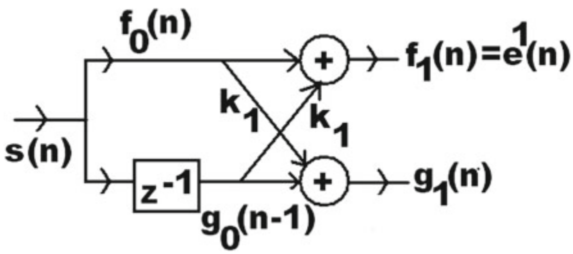

Fig. 2.7 Lattice Structure for the first order filter

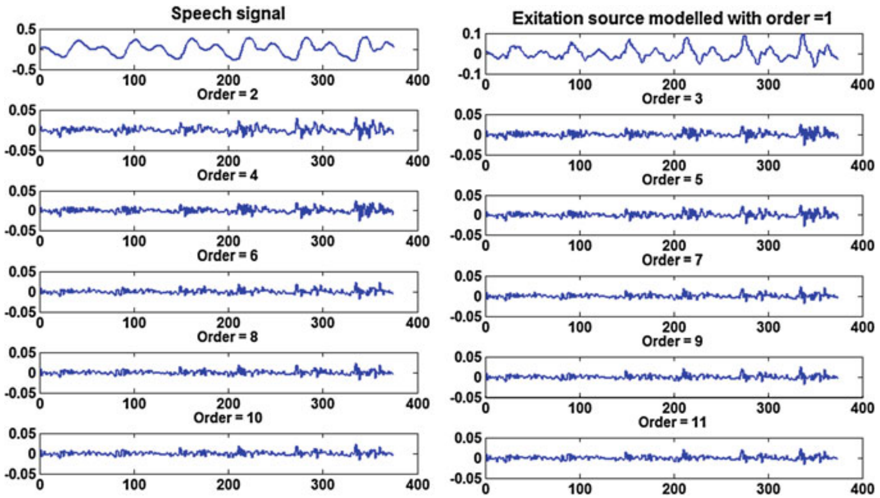

Fig. 2.8 Speech signal and the corresponding excitation source modelled using lpc with different co-efficients

$$f_2(n) = e^2(n) = s(n) + a_1^2 s(n-1) + a_2^2 s(n-2) \tag{2.90}$$

Comparing (2.88) and (2.90), we get the following

$$a_1^2 = k_1 + k_1 k_2 \tag{2.91}$$

$$a_2^2 = k_2 \tag{2.92}$$

In general it is noted that $a_r^r = k_r$. Also it is noted that $g_2(n) = k_1 f_0(n-1) + g_0(n-2) + k_2(f_0(n) + k_1 g_0(n-1))$, which is simplified as follows

$$g_2(n) = k_2 s(n) + (k_1 + k_1 k_2) s(n-1) + s(n-2) \tag{2.93}$$

$$g_2(n) = a_2^2 s(n) + a_1^2 s(n-1) + s(n-2) \tag{2.94}$$

Comparing (2.90) and (2.94), we understand that if the filter co-efficients of $f_2(n)$ are arranged in the reverse order, we get the filter co-efficients for $g_2(n)$. In z-domain $G_2(Z) = Z^{-2} F_2(Z^{-1})$. In general

$$G_N(Z) = Z^{-N} F_N(Z^{-1}) \tag{2.95}$$

2.5.1 Computation of Lattice Co-efficient from LPC Co-efficients

Let the Nth order transfer function $\frac{1}{V(Z)}$ is represented as $A_N(Z)$.

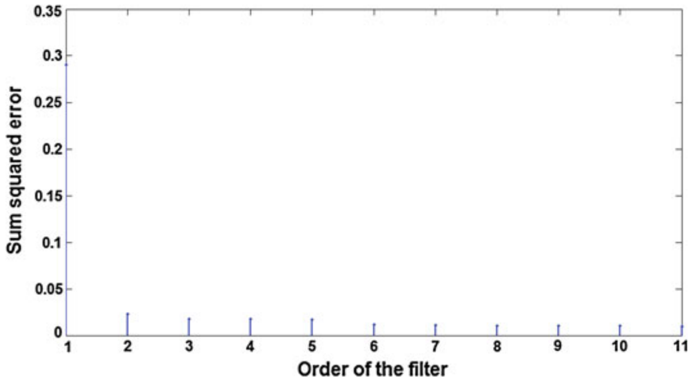

Fig. 2.9 Sum squared error (sum squared value of the samples of the excitation source) obtained using lpc model versus order of the lpc filter (number of lattice co-efficients)

$$A_N(Z) = 1 + a_1^N z^{-1} + a_2^N z^{-2} + a_3^N z^{-3} + \cdots + a_N^N z^{-N} \tag{2.96}$$

Note that a_N^N is the lattice co-efficient k_N. So if the $(N-1)$th order transfer function $A_{N-1}(Z)$ is obtained , the co-efficient of z^{N-1} of $A_{N-1}(Z)$ is obtained as k_N. The relation between $A_N(Z)$ and $A_{N-1}(Z)$ is needed and are obtained as follows. The excitation source signals obtained for various order and the corresponding sum squared values are displayed in Figs. 2.8 and 2.9 respectively.

$$F_N(Z) = F_{N-1}(Z) + k_N G_{N-1}(Z) Z^{-1} \tag{2.97}$$

$$G_N(Z) = k_N F_{n-1}(Z) + G_{N-1}(Z) Z^{-1} \tag{2.98}$$

$$S(Z) A_N(Z) = E_N(Z) = F_N(Z) \tag{2.99}$$

$$\Rightarrow A_N(Z) = A_{N-1}(Z) + k_N \frac{G_{N-1}(Z)}{S(Z)} Z^{-1} \tag{2.100}$$

Let $B_N(Z) = \frac{G_N(Z)}{S(Z)}$ and we get the following.

$$A_N(Z) = A_{N-1}(Z) + k_N B_{N-1}(Z)^{-1} \tag{2.101}$$

$$B_N(Z) = k_N A_{N-1}(Z) + B_{N-1}(Z) Z^{-1} \tag{2.102}$$

$$\Rightarrow A_N(Z) = A_{N-1}(Z) + k_N (B_N(Z) - k_N A_{N-1}(Z)) \tag{2.103}$$

$$\Rightarrow A_{N-1}(Z) = \frac{A_N(Z) - k_N B_N(Z)}{1 - k_N^2} \tag{2.104}$$

Note that $B_0(Z) = \frac{G_0(Z)}{S(Z)} = 1$. Thus the steps involved in computing lattice parameters from Nth order lpc are summarized as follows.

```
function [L,k,E,ERROR]=lattice(S1,FS)
L=lpc(S1,11);
k=[];
for i=1:1:length(L)-1
    k=[k L(length(L))];
    R=L(length(L):-1:1);
temp=(L-k(i)*R)/(1-(k(i)^2))
temp1=temp(1:1:length(temp)-1);
L=temp1;
end
close all
k=k(length(k):-1:1);
ERROR=[];
r=1;
for j=4:1:12
e=[];
for i=1:1:11
    f{i}(1)=0;
    g{i}(1)=0;
end
for n=2:1:length(S1)
f{1}(n)=S1(n);
g{1}(n)=S1(n);
g{1}(n-1)=S1(n-1);
for i=2:1:j
f{i}(n)=f{i-1}(n)+k(i-1)*g{i-1}(n-1) ;
g{i}(n)=g{i-1}(n-1)+k(i-1)*f{i-1}(n);
end
e=[e f{j}(n)];
end
E{r}=e;
ERROR=[ERROR sum(E{r}.^2)];
r=r+1;
end
figure
stem([2:1:11],ERROR)
figure
for i=2:1:10
    subplot(5,2,i)
    plot(E{i})
end
subplot(5,2,1)
plot(S1)
```

1. Obtain $A_N(Z) = 1 + a_1^N z^{-1} + a_2^N z^{-2} + a_3^N z^{-3} + \cdots + a_N^N z^{-N}$ using the lpc
2. $K_N = a_N^N$
3. Compute $B_N(Z) = Z^{-N} A_N(Z^{-1})$. Trick is to arrange the co-efficients of $A_N(Z)$ in the reverse order to obtain $B_N(Z)$.
4. Compute $A_{N-1}(Z) = \frac{A_N(Z) - k_N B_N(Z)}{1 - k_N^2}$. Identify the co-efficient of $(N-1)$ to obtain K_{N-1}
5. Repeat 3 and 4 to obtain the lattice co-efficients.

```
%levinsondurbin.m
function [res]=levinsondurbin(a)
%a is the vector with size 1xn
temp1=a(2:1:length(a));
for j=1:1:length(a)-1
A{j}=toeplitz(a(1:1:j));
end
x{1}=[a(2)/a(1)];
k(1)=(a(3)-a(2)*x{1}(1))/(a(1)-a(2)*x{1}(1));
c{1}=-inv(A{1})*a(2)*k(1) ;
c{1}=[c{1};k(1)];
x{2}=[x{1} ;0];
x{2}=x{2}+c{1};
for r=2:1:length(a)-2

k(r)=(a(r+2)-rev(a,r)*x{r})/(a(1)-[a(2:1:r+1)]*x{r});
c{r}=-1*inv(A{r})*rev(a,r)' *k(r) ;
c{r}=[c{r};k(r)];
x{r+1}=[x{r};0];
x{r+1}=x{r+1}+c{r};
end
res=[1; -1*x{r+1}];
```

http://www.springer.com/978-81-322-1676-6

Digital Speech Processing Using Matlab
Gopi, E.S.
2014, XVI, 182 p. 103 illus., Hardcover
ISBN: 978-81-322-1676-6