IDIAP RESEARCH REPORT

IDIAP

Martigny - Valais - Suisse

Investigation of a possible process identity between DRM and Linear Filtering

Sacha KRSTULOVIĆ [a]

IDIAP–RR 97-19

May 1997

Dalle Molle Institute for Perceptual Artificial Intelligence • P.O.Box 592 • Martigny • Valais • Switzerland

phone +41 − 27 − 721 77 11
fax +41 − 27 − 721 77 12
e-mail secretariat@idiap.ch
internet http://www.idiap.ch

[a] IDIAP

IDIAP Research Report 97-19

Investigation of a possible process identity between DRM and Linear Filtering

Sacha KRSTULOVIĆ

May 1997

Abstract. The classical analogy between linear filtering and acoustical filtering by tubes is applied in the non-classical case where the tubes are made of unequal-length sections (such as the DRM case). It is shown that the filtering process identity is substantially more complicated than in the case of equal-length sections. In particular, it prevents the use of the Levinson algorithm for inverting the filtering process and recovering the tube characteristics from sound alone.

Contents

1 Introduction

It is traditionally recognized that the Linear Prediction Coding (or LPC) modelling method has a relationship with the process of acoustical filtering occurring in a set of connected cylindrical pipes. The purpose of the present study is to disclose this relationship in the case of the Distinctive Regions and Modes articulatory model (DRM), which precisely consists in a pile of connected pipes. Such a relationship can then be exploited for the design of an acoustic-articulatory inversion sytem, to determine the parameters of the tube by means of inverse linear filtering.

The tube-LPC relation is rather obvious in the case of a pile of equally lengthy tubes, as we show in section 3. But we also show that the expected process identity is a lot more complicated in the case of the DRM, which is made of unequally lengthy tubes. In particular, the DRM does not appear to be compatible with a lattice structure for the corresponding inverse filter.

We try to use two different approaches to adress the problem :

- starting from the Levinson-Durbin equations, we try to recover the acoustical filtering process equations
- starting from acoustical phenomenons, we try to recover a recursive algorithm that could allow the implementation of a lattice inverse filter

Both of these approaches make the object of a section in the following.

2 Effect of $\mu_i = 0$ in a step of the Levinson recursion

A natural way of trying to solve our process identification problem is by starting with the Levinson recursion and disturbing it by adding the constraints inherited from the special structure of our unequal-lengths tube model.

As a matter of fact, we can consider that the DRM is made of a pile of equal-length tubes, some of them being fastened together in order to form a set of longer and unequally lengthy sections (see figure 1). This amounts to setting some reflection coefficients to zero in the course of our Auto-Regressive predictor design.

The classical form of the Levinson-Durbin algorithm for AR filter design is described in [RJ93] by :

$$E^{(0)} = r_0 \tag{1}$$

$$\mu_{m+1} = \left[r_{m+1} - \sum_{i=1}^{m} a_i^{(m)} r_{m+1-i} \right] / E^{(m)} \tag{2}$$

$$\begin{cases} a_{m+1}^{(m+1)} = \mu_{m+1} \\ a_i^{(m+1)} = a_i^{(m)} + \mu_{m+1} a_{m+1-i}^{(m)} \qquad i = 1, \cdots, m \end{cases} \tag{3}$$

$$E^{(m+1)} = \left(1 - \mu_{m+1}^2\right) E^{(m)} \tag{4}$$

with :

- a_m : LPC coefficients
- μ_m : partial correlation or reflection coefficients
- $r_m = \sum_{n=0}^{N-1-m} x(n)x(n+m)$: values of the (estimated) autocorrelation function.

What does $\mu_{m+1} = 0$ brings about the correlation and the LPC coefficients ?

- From equation (3), it simply means that the predictor has not changed between step m and step $m+1$ of the algorithm.
- From equation (4), it means that the energy of the prediction error stays the same.
- From equation (2), setting $\mu_{m+1} = 0$ induces :

$$r_{m+1} = \sum_{i=1}^{m} a_i^{(m)} r_{m+1-i} \tag{5}$$

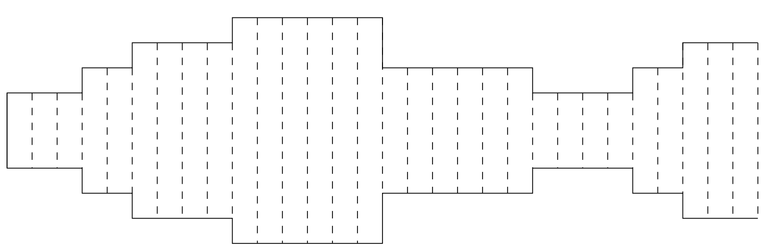

Figure 1: The DRM tube as a concatenation of 30 equal-length sections.

Adopting a matrix notation, we have :

$$r_{m+1} = \left[a_1^{(m)}, a_2^{(m)}, \cdots a_m^{(m)}\right] \begin{bmatrix} r_m \\ r_{m-1} \\ \vdots \\ r_1 \end{bmatrix} \quad (6)$$

This is a way of constraining the autocorrelation matrix.
Given that :

- the autocorrelation matrix is supposed to be estimated from the original signal
- the LPC coefficients at step (m) are determined and fixed at the previous step

then this constraint amounts to imposing some values in the autocorrelation matrix. We do not just try here to neglect some terms in the matrix. This is therefore equivalent to constraining the original sound signal itself, which we just wish to analyse.

Trying to incorporate the DRM constraints in the Levinson recursion leads to the above contradiction. This shows that the original Levinson recursive algorithm cannot be used to solve our problem.

We have therefore to address the problem by its other end, i.e. start from the fluid dynamics of the tube, deduce the general form of the tube's transfer function, and finally find an estimator for the transfer function's parameters.

3 Filtering process of an acoustic tube

3.1 Fluid dynamics roots of the problem

The following section takes up the mathematical development exposed by Wakita in [Wak73], where the emergence of AR filtering equations from the acoustical filtering process is clearly shown in the case of equal-length tubes. The original work is extended to the case of unequal-length tube portions.

Basic system :

The vocal tract is considered to be an acoustic tube divided in M sections of any (time-independent) length.

Assumptions :

- sound waves are plane fluid waves (see [Fla72], pp.24-25, or [MI68], p.467)
- the tube is rigid (no wall impedance)
- losses due to viscosity and heat conduction are neglected.

Equation set :

Posing the problem in terms of fluid dynamics, we can consider that the volume velocity $u_m(t,d)$ and the pressure $p_m(t,d)$ in section m derive from a potential $\Phi_m(t,d)$:

$$u_m(t,d) = -S_m \frac{\partial \Phi_m(t,d)}{\partial d} \tag{7}$$

$$p_m(t,d) = \rho \frac{\partial \Phi_m(t,d)}{\partial t} \tag{8}$$

with :

- t : time variable
- d : distance variable
- S_m : surface of m^{th} section
- ρ : density of air

The evolution of the fluid state is thereafter described by Webster's equation :

$$\frac{\partial^2 \Phi_m(t,d)}{\partial d^2} - \frac{1}{c^2} \frac{\partial^2 \Phi_m(t,d)}{\partial t^2} = 0 \tag{9}$$

where c denotes the sound velocity.

Equation solving :

If we assume that the excitation source (the "glottis" of the tube) delivers a sinusoidal signal, then the solution of this equation is of the classical form :

$$\Phi_m(t,d) = A \exp^{j\omega(t-d/c)} + B \exp^{j\omega(t+d/c)} \tag{10}$$

where A and B are constants[1]. Remarking that $u_m(t,d)$ can be decomposed into a forward-travelling wave $u_m^+(t,d)$ and a backward-travelling wave $u_m^-(t,d)$, the above solution can be decomposed in the

[1] If the excitation signal is made of a linear combination of sine waves, which can be obtained from any signal when applying the Fourier transform, the corresponding solution is a linear combination of the solutions for any individual sinusoidal component. Therefore, the relations developped hereafter do not loose their generality (in the limits of the assumptions made at the beginning) when a non-sinusoidal excitation such as the wave coming out of the vocal cords is applied.

following way :

$$\begin{cases} u_m(t,d) = u_m^+(t,d) - u_m^-(t,d) \\ p_m(t,d) = \frac{\rho c}{S_m}\left\{u_m^+(t,d) + u_m^-(t,d)\right\} \end{cases} \tag{11}$$

with

$$\begin{cases} u_m^+(t,d) = \frac{j\omega S_m A}{c}\exp^{j\omega(t-d/c)} \\ u_m^-(t,d) = \frac{j\omega S_m B}{c}\exp^{j\omega(t+d/c)} \end{cases} \tag{12}$$

At the connection between section m and section $m+1$, the volume velocity and pressure must be continuous. We therefore have the additional relations :

$$\begin{cases} u_{m+1}(t,d_m) = u_m(t,d_m) \\ p_{m+1}(t,d_m) = p_m(t,d_m) \end{cases} \tag{13}$$

with d_m being the distance between the glottis and the connection between sections m and $m+1$ (see figure 2). Since the speed of sound is constant, the distance variable can be related to the time variable and can thus be eliminated. Since there is no loss in a particular section, we also have, *inside* the limits of a section :

$$\begin{cases} u_m^+(t,d) = u_m^+(t,d-\Delta l_m) = u_m^+(t-\frac{\Delta l_m}{c}) \\ u_m^-(t,d) = u_m^-(t,d-\Delta l_m) = u_m^-(t+\frac{\Delta l_m}{c}) \end{cases} \tag{14}$$

with Δl_m being the length of the considered piece of tube. Wakita explains that point very clearly in [Wak73] :

> “ Since no loss is assumed, the volume velocity component $u_{m+1}^+(t,d_m)$ is equal to that component of the volume velocity that started at d_{m+1} at time $\Delta l/c$ [or $\Delta l_m/c$ in our case] earlier, and the volume velocity component $u_{m+1}^-(t,d_m)$ is equal to that component of the volume velocity that will arrive at d_{m+1} at time $\Delta l/c$ [$\Delta l_m/c$] later. Thus the solution of the continuous problem can be obtained by knowing only the values at each junction.”

This step is very important, as dropping the distance variable allows us to express our problem in terms of time series analysis. Furthermore, the fact that the problem can be solved considering only the junctions will allow us to work in a discrete world.

3.2 From fluids to signals

Starting from fluid dynamics, we end up with the following relations between the forward and backward travelling waves at each junction :

$$\begin{cases} u_{m+1}^+(t-\Delta_m t) - u_{m+1}^-(t+\Delta_{m+1} t) = u_m^+(t) - u_m^-(t) \\ \frac{\rho c}{S_{m+1}}\left\{u_{m+1}^+(t-\Delta_m t) + u_{m+1}^-(t+\Delta_{m+1} t)\right\} = \frac{\rho c}{S_{m+1}}\left\{u_m^+(t) + u_m^-(t)\right\} \end{cases} \tag{15}$$

where :

$$\Delta_m t = \frac{\Delta l_m}{c}$$

Now defining the coefficient :

$$\mu_m = \frac{S_m - S_{m+1}}{S_m + S_{m+1}}$$

and applying to the above equations, we obtain :

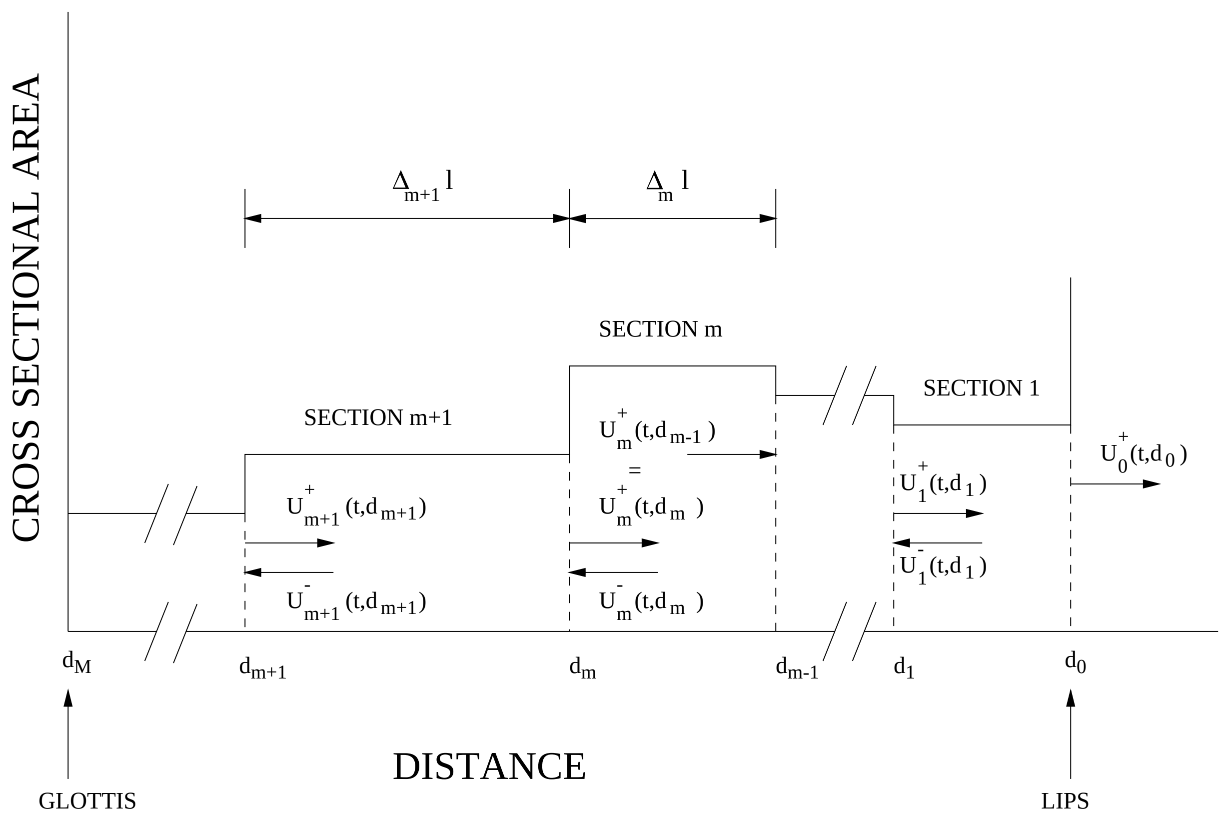

Figure 2: Non-uniform acoustic tube model of the vocal tract. (Inspired from [Wak73].)

$$\left\{\begin{array}{l} u^+_{m+1}(t-\Delta_m t) = \frac{1}{1+\mu_m}\left\{u^+_m(t) - \mu_m u^-_m(t)\right\} \\ \\ u^-_{m+1}(t+\Delta_m t) = \frac{1}{1+\mu_m}\left\{-\mu_m u^+_m(t) + u^-_m(t)\right\} \end{array}\right. \tag{16}$$

Defining a unit length Δl_{unit} as the greatest common divisor of the lenghts Δl_m, we can apply the Z-transform with z defined as $z = e^{j\omega 2\Delta l_{unit}/c} = e^{j\omega 2\Delta_{unit} t}$, and we obtain:

$$\left\{\begin{array}{l} z^{\frac{-n_m}{2}} U^+_{m+1}(z) = \frac{1}{1+\mu_m}\left[U^+_m(z) - \mu_m U^-_m(z)\right] \\ \\ z^{\frac{n_m}{2}} U^-_{m+1}(z) = \frac{1}{1+\mu_m}\left[-\mu_m U^+_m(z) + U^-_m(z)\right] \end{array}\right. \tag{17}$$

i.e.

$$\left\{\begin{array}{l} U^+_{m+1}(z) = \frac{z^{\frac{n_m}{2}}}{1+\mu_m}\left[U^+_m(z) - \mu_m U^-_m(z)\right] \\ \\ U^-_{m+1}(z) = \frac{z^{\frac{-n_m}{2}}}{1+\mu_m}\left[-\mu_m U^+_m(z) + U^-_m(z)\right] \end{array}\right. \tag{18}$$

and, in matrix notation:

$$\left[\begin{array}{c} U^+_{m+1}(z) \\ U^-_{m+1}(z) \end{array}\right] = \frac{z^{\frac{n_m}{2}}}{1+\mu_m}\left[\begin{array}{cc} 1 & -\mu_m \\ -\mu_m z^{-n_m} & z^{-n_m} \end{array}\right]\left[\begin{array}{c} U^+_m(z) \\ U^-_m(z) \end{array}\right] \tag{19}$$

If we assume that the lips end is connected to a tube of infinite section, it amounts to the following boundary condition at front end (or lips end) of our model:

$$S_{-1} = \infty \Rightarrow \mu_0 = 1$$

Applying this condition, we can write:

$$\left[\begin{array}{c} U^+_{m+1}(z) \\ U^-_{m+1}(z) \end{array}\right] = z^{\frac{1}{2}\sum_{k=0}^{m} n_k} K_m \left[\begin{array}{c} D^+_m(z) \\ D^-_m(z) \end{array}\right]\left\{U^+_0(z) - U^-_0(z)\right\} \tag{20}$$

with

$$\left[\begin{array}{c} D^+_m(z) \\ D^-_m(z) \end{array}\right] = \left[\begin{array}{cc} 1 & -\mu_m \\ -\mu_m z^{-n_m} & z^{-n_m} \end{array}\right]\left[\begin{array}{cc} 1 & -\mu_{m-1} \\ -\mu_{m-1} z^{-n_{m-1}} & z^{-n_{m-1}} \end{array}\right]\cdots\left[\begin{array}{cc} 1 & -\mu_1 \\ -\mu_1 z^{-n_1} & z^{-n_1} \end{array}\right]\left[\begin{array}{c} 1 \\ -z^{-n_0} \end{array}\right] \tag{21}$$

and

$$K_m = \prod_{i=0}^{m} \frac{1}{1+\mu_i} \tag{22}$$

Neglecting the overall delay $z^{\frac{1}{2}\sum_{k=0}^{m} n_k}$ and the gain K_m, **the true transfer function for the forward travelling volume velocity is there denoted by $D^+_m(z)$, and can be built recursively by applying:**

$$\left[\begin{array}{c} D^+_{m+1}(z) \\ D^-_{m+1}(z) \end{array}\right] = \left[\begin{array}{cc} 1 & -\mu_{m+1} \\ -\mu_{m+1} z^{-n_{m+1}} & z^{-n_{m+1}} \end{array}\right]\left[\begin{array}{c} D^+_m(z) \\ D^-_m(z) \end{array}\right] \tag{23}$$

We can also show by mathematical induction from equation (21) that we have:

$$D^-_m(z) = -z^{-\sum_{k=0}^{m} n_k} D^+_m(1/z) \tag{24}$$

Developing (23) and applying (24), we obtain:

$$\begin{cases} D^+_{m+1}(z) = D^+_m(z) + \mu_{m+1} z^{-\sum_{k=0}^{m} n_k} D^+_m(1/z) \\ D^+_{m+1}(1/z) = \mu_{m+1} z^{\sum_{k=0}^{m} n_k} D^+_m(z) + D^+_m(1/z) \end{cases} \quad (25)$$

We can remark that if we change the variable z to $1/z$ in the first of the above formulae, we obtain the formula in the second line. Both formulae are equivalent with regard to the relationship they imply between $D^+_{m+1}(z)$ and $D^+_m(z)$.

We now develop this relationship in order to study more precisely the form and the growth of the transfer function. This developpment is made in the case of equal length sections and then in the case of unequal-length sections such as in the DRM.

3.3 Case 1 : the length of the sections is uniform

In this case, the transmission delay induced in every piece of tube is the same. Therefore, we can set $n_k = 1\ \forall k$, i.e. $z^{-n_k} = z^{-1}\ \forall k$ in all the above equations. z is then defined as $z = e^{j\omega 2\Delta l/c}$, Δl being the length of every piece of tube.

We know from equation (21) that $D_m^+(z)$ is of the form :

$$D_m^+(z) = \sum_{i=0}^{m} a_i^{(m)} z^{-i} \tag{26}$$

and we have also :

$$D_{m+1}^+(z) = \sum_{i=0}^{m+1} a_i^{(m+1)} z^{-i} \tag{27}$$

The relation (25) gives :

$$\sum_{i=0}^{m+1} a_i^{(m+1)} z^{-i} = \sum_{i=0}^{m} a_i^{(m)} z^{-i} + \mu_{m+1} z^{-(m+1)} \sum_{i=0}^{m} a_i^{(m)} z^{i} \tag{28}$$

i.e.

$$\sum_{i=0}^{m+1} a_i^{(m+1)} z^{-i} = \sum_{i=0}^{m} a_i^{(m)} z^{-i} + \mu_{m+1} \sum_{i=0}^{m} a_i^{(m)} z^{i-(m+1)} \tag{29}$$

or, changing the mute index i to $(m+1-i)$ in the second sum :

$$\sum_{i=0}^{m+1} a_i^{(m+1)} z^{-i} = \sum_{i=0}^{m} a_i^{(m)} z^{-i} + \mu_{m+1} \sum_{i=1}^{m+1} a_{m+1-i}^{(m)} z^{-i} \tag{30}$$

Identifying the coefficients of the polynoms in z^{-i} on each side of the equal sign, we obtain :

$$\begin{cases} a_0^{(m+1)} = a_0^{(m)} \\ a_1^{(m+1)} = a_1^{(m)} + \mu_{m+1} a_m^{(m)} \\ \vdots \\ a_m^{(m+1)} = a_m^{(m)} + \mu_{m+1} a_1^{(m)} \\ a_{m+1}^{(m+1)} = \mu_{m+1} a_0^{(m)} \end{cases} \tag{31}$$

which can be formalized in one line as :

$$a_i^{(m+1)} = a_i^{(m)} + \mu_{m+1} a_{m+1-i}^{(m)} \tag{32}$$

These equations are similar to those linking partial correlation coefficients μ_i to prediction coefficients a_i in the Levinson-Durbin algorithm for LPC modelling. If we analyse speech with a sampling frequency of $F_s = \frac{c}{2\Delta l}$, and if we estimate the reflection coefficients in a proper way (for instance using Itakura's covariance method, see [Ita71]), the equivalence between an LPC model of order M and the filtering process of the tube needs no more assumptions to hold.

3.4 Case 2: the length of the sections is not uniform

If the length of the sections is not uniform, we must deal with the irregular delays and the z^{-n_k} not being equal to z^{-1}. To formalize the growth of the transfer function in a readable way, we will borrow the notation of the summation indexes to the set theory. Let Ω_m be the set of all possible indexes k for the discrete delays n_k, and let Γ be a set containing one of the possible index combinations[2].

We know from equation (21) that $D_m^+(z)$ is a polynomial in z and we can now express its form as:

$$D_m^+(z) = \sum_{\Gamma \subset \{0,1,\cdots m\}} a_\Gamma^{(m)} z^{-\sum_{k\in\Gamma} n_k} \tag{33}$$

or:

$$D_{m+1}^+(z) = \sum_{\Gamma \subset \{0,1,\cdots m+1\}} a_\Gamma^{(m+1)} z^{-\sum_{k\in\Gamma} n_k} \tag{34}$$

The relation (25) now gives:

$$\sum_{\Gamma\subset\Omega_{m+1}} a_\Gamma^{(m+1)} z^{-\sum_{k\in\Gamma} n_k} = \sum_{\Gamma\subset\Omega_m} a_\Gamma^{(m)} z^{-\sum_{k\in\Gamma} n_k} + \mu_{m+1} z^{-\sum_{k\in\Omega_{m+1}} n_k} \sum_{\Gamma\subset\Omega_m} a_\Gamma^{(m)} z^{\sum_{k\in\Gamma} n_k} \tag{35}$$

i.e.

$$\sum_{\Gamma\subset\Omega_{m+1}} a_\Gamma^{(m+1)} z^{-\sum_{k\in\Gamma} n_k} = \sum_{\Gamma\subset\Omega_m} a_\Gamma^{(m)} z^{-\sum_{k\in\Gamma} n_k} + \mu_{m+1} \sum_{\Gamma\subset\Omega_m} a_\Gamma^{(m)} z^{\sum_{k\in\Gamma} n_k - \sum_{k\in\Omega_{m+1}} n_k} \tag{36}$$

For a particular subset Γ of our index set Ω_m, we can show the following:

$$\begin{aligned} \textstyle\sum_{k\in\Gamma} n_k - \sum_{k\in\Omega_{m+1}} n_k &= \underbrace{\sum_{k\in\Gamma} n_k - \sum_{k\in\Omega_m} n_k} \quad -n_{m+1} \\ &= \quad -\textstyle\sum_{k\in\overline{\Gamma}} n_k \quad\quad -n_{m+1} \end{aligned} \tag{37}$$

$\overline{\Gamma}$ being the complementary set of Γ so that $\Gamma \cup \overline{\Gamma} = \Omega_m$.
Equation (36) then becomes:

$$\sum_{\Gamma\subset\Omega_{m+1}} a_\Gamma^{(m+1)} z^{-\sum_{k\in\Gamma} n_k} = \sum_{\Gamma\subset\Omega_m} a_\Gamma^{(m)} z^{-\sum_{k\in\Gamma} n_k} + \mu_{m+1} \sum_{\Gamma\subset\Omega_m} a_\Gamma^{(m)} z^{-\sum_{k\in\overline{\Gamma}} n_k + n_{m+1}} \tag{38}$$

In this case, the analytical identification of the polynomial coefficients has to be performed on a case-by-case basis.

For instance, let us express it in the case of the DRM. In this case, we have 8 sections of unequal length with $\Delta l_{unit} = L/30$ (L being the total length of the full tube). The lengths of the sections are distributed as follows from lips to glottis: $\Delta l_0 = 3\Delta l_{unit}$, $\Delta l_1 = 2\Delta l_{unit}$, $\Delta l_2 = 4\Delta l_{unit}$, $\Delta l_3 = 6\Delta l_{unit}$, $\Delta l_4 = 6\Delta l_{unit}$, $\Delta l_5 = 4\Delta l_{unit}$, $\Delta l_6 = 2\Delta l_{unit}$, $\Delta l_7 = 3\Delta l_{unit}$.
The recursion allowing to compute the DRM transfer function is then of the form:

[2] We have $\Omega_m = \{0, 1, \cdots m\}$ and $\Gamma \subset \Omega_m$. This means that Γ belongs to the set of all subsets of Ω_m.
We can remark that this later set defines a σ-algebra on the set of delays n_k. A measure on this set could be defined as $\sum_{k\in\Gamma} n_k$. We don't know if such measure theory notions have already been used in the framework of polynomial transfer functions analysis, but an expert in measure theory might find here a lead to an alternate way of investigating our problem.

$$\begin{bmatrix} D^+(z) \\ D^-(z) \end{bmatrix} = \begin{bmatrix} 1 & -\mu_7 \\ -\mu_7 z^{-3} & z^{-3} \end{bmatrix} \begin{bmatrix} 1 & -\mu_6 \\ -\mu_6 z^{-2} & z^{-2} \end{bmatrix} \cdots \begin{bmatrix} 1 & -\mu_2 \\ -\mu_2 z^{-4} & z^{-4} \end{bmatrix} \begin{bmatrix} 1 & -\mu_1 \\ -\mu_1 z^{-2} & z^{-2} \end{bmatrix} \begin{bmatrix} 1 \\ -z^{-3} \end{bmatrix} \tag{39}$$

When observing the growth of the transfer function between step 3 and step 4 for instance (see equations developed in figure 4), and replacing the Γ indexes by integer indexes corresponding to the place of the increasing negative powers of z, we obtain the following set of equations :

$$\left\{ \begin{array}{lllll} a_0^{(4)} & = & a_0^{(3)} & = & 1 \\ a_1^{(4)} & = & a_1^{(3)} & & \\ a_2^{(4)} & = & a_2^{(3)} & & \\ a_3^{(4)} & = & a_3^{(3)} & & \\ a_4^{(4)} & = & a_4^{(3)} & & \\ a_5^{(4)} & = & a_5^{(3)} & + & \mu_4 a_7^{(3)} \\ a_6^{(4)} & = & a_6^{(3)} & & \\ a_7^{(4)} & = & & & \mu_4 a_6^{(3)} \\ a_8^{(4)} & = & a_7^{(3)} & + & \mu_4 a_5^{(3)} \\ a_9^{(4)} & = & & & \mu_4 a_4^{(3)} \\ a_{10}^{(4)} & = & & & \mu_4 a_3^{(3)} \\ a_{11}^{(4)} & = & & & \mu_4 a_2^{(3)} \\ a_{12}^{(4)} & = & & & \mu_4 a_1^{(3)} \\ a_{13}^{(4)} & = & & & \mu_4 \end{array} \right. \tag{40}$$

In the general case, we see that if we try to operate a polynomial coefficients identity starting from equation (25), we cannot meet the Levinson-Durbin relation tying prediction coefficients a_i and reflection coefficients μ_i. As the basic idea of the Levinson algorithm is to find a relation between an m^{th} order predictor and its $(m+1)^{th}$ order successor, we try in the following section to come up with something similar in the irregular lengths case.

$$\begin{bmatrix} 1 \\ z^{-1} \end{bmatrix}$$

$$\begin{bmatrix} 1-\frac{\mu_1}{z} \\ -\frac{\mu_1}{z}+z^{-2} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_2\,\mu_1-\mu_1}{z}-\frac{\mu_2}{z^2} \\ -\frac{\mu_2}{z}+\frac{\mu_2\,\mu_1-\mu_1}{z^2}+z^{-3} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_2\,\mu_1-\mu_1+\mu_3\,\mu_2}{z}+\frac{-\mu_2+\mu_3\,\mu_1-\mu_3\,\mu_2\,\mu_1}{z^2}-\frac{\mu_3}{z^3} \\ -\frac{\mu_3}{z}+\frac{-\mu_2+\mu_3\,\mu_1-\mu_3\,\mu_2\,\mu_1}{z^2}+\frac{\mu_2\,\mu_1-\mu_1+\mu_3\,\mu_2}{z^3}+z^{-4} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_2\,\mu_1-\mu_1+\mu_4\,\mu_3+\mu_3\,\mu_2}{z}+\frac{-\mu_2+\mu_4\,\mu_3\,\mu_2\,\mu_1-\mu_4\,\mu_3\,\mu_1+\mu_3\,\mu_1-\mu_3\,\mu_2\,\mu_1+\mu_4\,\mu_2}{z^2}+\frac{-\mu_4\,\mu_3\,\mu_2-\mu_3+\mu_4\,\mu_1-\mu_4\,\mu_1\,\mu_2}{z^3}-\frac{\mu_4}{z^4} \\ -\frac{\mu_4}{z}+\frac{-\mu_4\,\mu_3\,\mu_2-\mu_3+\mu_4\,\mu_1-\mu_4\,\mu_1\,\mu_2}{z^2}+\frac{-\mu_2+\mu_4\,\mu_3\,\mu_2\,\mu_1-\mu_4\,\mu_3\,\mu_1+\mu_3\,\mu_1-\mu_3\,\mu_2\,\mu_1+\mu_4\,\mu_2}{z^3}+\frac{\mu_2\,\mu_1-\mu_1+\mu_4\,\mu_3+\mu_3\,\mu_2}{z^4}+z^{-5} \end{bmatrix}$$

$$\vdots$$

Figure 3: **Regular tube transfer function growth.** Note the regular increase in the polynomial degrees. (Equations computed with the help of a symbolic computation program.)

$$\begin{bmatrix} 1 \\ -z^{-3} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_1}{z^3} \\ -\frac{\mu_1}{z^2}-z^{-5} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_2\,\mu_1}{z^2}+\frac{\mu_1}{z^3}+\frac{\mu_2}{z^5} \\ -\frac{\mu_2}{z^4}-\frac{\mu_1}{z^6}-\frac{\mu_2\,\mu_1}{z^7}-z^{-9} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_2\,\mu_1}{z^2}+\frac{\mu_1}{z^3}+\frac{\mu_3\,\mu_2}{z^4}+\frac{\mu_2}{z^5}+\frac{\mu_3\,\mu_1}{z^6}+\frac{\mu_3\,\mu_2\,\mu_1}{z^7}+\frac{\mu_3}{z^9} \\ -\frac{\mu_3}{z^6}-\frac{\mu_3\,\mu_2\,\mu_1}{z^8}-\frac{\mu_3\,\mu_1}{z^9}-\frac{\mu_2}{z^{10}}-\frac{\mu_3\,\mu_2}{z^{11}}-\frac{\mu_1}{z^{12}}-\frac{\mu_2\,\mu_1}{z^{13}}-z^{-15} \end{bmatrix}$$

$$\begin{bmatrix} 1+\frac{\mu_2\,\mu_1}{z^2}+\frac{\mu_1}{z^3}+\frac{\mu_3\,\mu_2}{z^4}+\frac{\mu_2}{z^5}+\frac{\mu_4\,\mu_3+\mu_3\,\mu_1}{z^6}+\frac{\mu_3\,\mu_2\,\mu_1}{z^7}+\frac{\mu_4\,\mu_3\,\mu_2\,\mu_1}{z^8}+\frac{\mu_4\,\mu_3\,\mu_1+\mu_3}{z^9}+\frac{\mu_4\,\mu_2}{z^{10}}+\frac{\mu_4\,\mu_3\,\mu_2}{z^{11}}+\frac{\mu_4\,\mu_1}{z^{12}}+\frac{\mu_4\,\mu_2\,\mu_1}{z^{13}}+\frac{\mu_4}{z^{15}} \\ -\frac{\mu_4}{z^6}-\frac{\mu_4\,\mu_2\,\mu_1}{z^8}-\frac{\mu_4\,\mu_1}{z^9}-\frac{\mu_4\,\mu_3\,\mu_2}{z^{10}}-\frac{\mu_4\,\mu_2}{z^{11}}-\frac{\mu_4\,\mu_3\,\mu_1+\mu_3}{z^{12}}-\frac{\mu_4\,\mu_3\,\mu_2\,\mu_1}{z^{13}}-\frac{\mu_3\,\mu_2\,\mu_1}{z^{14}}-\frac{\mu_4\,\mu_3+\mu_3\,\mu_1}{z^{15}}-\frac{\mu_2}{z^{16}}-\frac{\mu_3\,\mu_2}{z^{17}}-\frac{\mu_1}{z^{18}}-\frac{\mu_2\,\mu_1}{z^{19}}-z^{-21} \end{bmatrix}$$

$$\vdots$$

Figure 4: **DRM tube transfer function growth.** Note the disturbance in the polynomial degrees.

3.5 Relation between polynomial coefficients, reflection coefficients and the Yule-Walker equation system

In the case of equal-length tube sections, the degree of the polynomial tube transfer function increases by -1 at each step of its growth. When trying to estimate the transfer function by solving the Yule-Walker equations in a recursive fashion such as using the Levinson-Durbin algorithm, the problem is the following:

> given a set of m polynomial coefficients $a_i^{(m)}$, resulting from solving a $m \times m$ system of Yule-Walker equations at step m, and given one reflection coefficient depending upon the $(m+1) \times (m+1)$ correlation matrix, what are the (m+1) polynomial coefficients $a_i^{(m+1)}$ of the transfer function at step (m+1) (or what is the solution of the $(m+1) \times (m+1)$ Yule-Walker system at the next step) ?

In the case of the non-equal length tubes, the degree of the polynomial increases by a certain amount p, very often different from 1. If we want to apply the classical RMS criterion for estimating our predictor at a particular step m (see appendix A) , the estimation still corresponds to solving a linear system of the form:

$$\left[1, a_1^{(m)}, \cdots, a_m^{(m)}\right] \mathcal{R}_m = [0, 0, \cdots, 0] \tag{41}$$

But here, due to the application of irregular delays for the computation of the correlation matrix $\mathcal{R}_m$, the matrix looses the Toeplitz structure and in some cases the symmetry. The problem is therefore :

> given a set of m polynomial coefficients resulting from solving Yule-Walker-like, non-Toeplitz equations at step m, and given a *single* reflection coefficient related to new correlation values, what are the (m+p) polynomial coefficients of the transfer function at step $(m+1)$?

This is an ill-posed problem, as we miss $(p-1)$ known parameters to solve our $(m+p) \times (m+p)$ Yule-Walker-like system of equation. Even though we get p new correlation values, they are merged into one reflection coefficient, and we loose $(p-1)$ degrees of liberty.

The problem is therefore uncompatible with a simple inverse filtering scheme using a simple ("monodimensional") lattice structure. Recursive solutions of an other nature may possibly be found in the domain of numerical analysis, but their design and implementation would exceed the scope of the present study.

One could argue that knowing the structure of the transfer function, and given the correlation matrix, we could solve the Yule-Walker-like system at step m and $m+1$ and then deduce the reflection coefficients μ_m from the obtained $a_i^{(m)}$ and $a_i^{(m+1)}$. Experimental attempts to do so have led to numerical errors (probably due to ill-conditioned correlation matrices) making the method untractable. For instance, we haven't been able to verify the relation between the predictor at step 3 and the predictor at step 4 in the DRM case illustrated by equation set (40).

4 Conclusion

As we show in the present study, the DRM articulatory model leads to an ill-posed problem when trying to identify its acoustical filtering action with a simple AR linear filtering process. Although an inverse filter might be found in a numerical analysis framework or in an acoustical theory framework, the difficulty of reaching a solution diminishes the interest of using the DRM model in an acoustic-articulatory inversion system that would be based on an inverse filtering scheme.

References

[Fla72] J.L. Flanagan. *Speech analysis, synthesis and perception.* Springer Verlag, 1972.

[Ita71] F. Itakura. Extraction of feature parameters of speech by statistical method. In *Proceedings of the 8th symposium on Speech Information Processing*, volume II, pages 5–2 – 5–12, February 1971.

[MI68] P.M. Morse and K.U. Ingard. *Theoretical acoustics.* Mc Graw-Hill, 1968.

[RJ93] L. Rabiner and B.H. Juang. *Fundamentals of Speech Recognition.* Prentice Hall, 1993.

[Wak73] H. Wakita. Direct estimation of the vocal-tract shape by inverse filtering of acoustic speech waveforms. *IEEE Transactions on Audio and Electroacoustics*, AU-21:417–427, October 1973.

Acknowledgement: Thanks to E. Mayoraz for allowing some time to the fruitful discussions we have had.

A Application of the Correlation method with the RMS criterion in the case of the DRM

In the case of the DRM, we know that the inverse filter transfer function should have the form :

$$A(z) = \sum_{\Gamma \subset \Omega_m} a_\Gamma z^{-\sum_{k\in\Gamma} n_k} = \frac{Y(z)}{X(z)} \tag{42}$$

with $\Omega_m = \{0, 1, \cdots, m\}$ and given m discrete delays n_k related to the geometry of the tube.
This form corresponds to the following difference equation :

$$y_n = \sum_{\Gamma \subset \Omega_n} a_\Gamma x_{n-\sum_{k\in\Gamma} n_k} \tag{43}$$

The input of the speech AR model is defined as an impulse train δ_n. The inverse filter modelling error can therefore be expressed by substracting the input of the model to the output of the inverse filter. Hence the expression defining the output error :

$$\begin{aligned} \varepsilon_n &= y_n - \delta_n & (44)\\ &= \sum_{\Gamma \subset \Omega_n} a_\Gamma x_{n-\sum_{k\in\Gamma} n_k} - \delta_n & (45) \end{aligned}$$

Since we want our speech model to have the form $X(z) = \frac{\eta}{A(z)}$, the impulse train can equivalently be replaced by an impulse of height η in the above equation :

$$\varepsilon_n = \sum_{\Gamma \subset \Omega_n} a_\Gamma x_{n-\sum_{k\in\Gamma} n_k} - \eta\delta_{n_0} \tag{46}$$

Finding the optimal inverse filter parameters corresponds to minimizing the mean squared error defined as :

$$\begin{aligned} E &= \sum_{n=0}^{N-1-\sum_{k\in\Omega_n} n_k} \varepsilon_n^2 & (47)\\ &= \sum_{n=0}^{N-1-\sum_{k\in\Omega_n} n_k} \left[\sum_{\Gamma \subset \Omega_n} a_\Gamma x_{n-\sum_{k\subset\Gamma} n_k} - \eta\delta_{n_0}\right]^2 & (48) \end{aligned}$$

The differentiation of E with respect to each a_Γ (except $a_\emptyset$, which is set to 1) gives :

$$\frac{\partial E}{\partial a_\Gamma} = 2\sum_{\Lambda \subset \Omega_n} a_\Lambda \left[\sum_{n=0}^{N-1-\sum_{k\in\Omega_n} n_k} x_{n-\sum_{k\in\Lambda} n_k} x_{n-\sum_{k\in\Gamma} n_k}\right] - 2\eta \left[\sum_{n=0}^{N-1-\sum_{k\in\Omega_n} n_k} x_{n-\sum_{k\in\Gamma} n_k}\delta_{n_0}\right] \tag{49}$$

Using the autocorrelation function usually defined as :

$$R_{i-j} = \sum_{n=0}^{N+M-1} x_{n-i}x_{n-j} = \sum_{n=0}^{N-1-|i-j|} x_n x_{|i-j|} \tag{50}$$

but considering only the terms of the form :

$$\begin{aligned} R_{\Gamma,\Lambda} &= \sum_{n=0}^{N-1-\sum_{k\in\Omega_n} n_k} x_{n-\sum_{k\in\Lambda} n_k} \cdot x_{n-\sum_{k\in\Gamma} n_k} & (51)\\ &= \sum_{n=0}^{N-1-\sum_{k\in\Omega_n} n_k} x_n \cdot x_{n+\left|\sum_{k\in\Gamma} n_k - \sum_{k\in\Lambda} n_k\right|} & (52) \end{aligned}$$

and setting equation (49) to zero, we obtain the linear equation system :

$$\sum_{\Lambda \subset \Omega_m} a_\Lambda R_{\Lambda,\Gamma} = 0 \tag{53}$$

with $\Gamma \subset \Omega_m$ and $\Gamma \neq \emptyset$.
Considering that for $\Lambda = \emptyset$ we have $a_\Lambda = 1$, we can finally express this system as :

$$\sum_{\Lambda \subset \Omega_m} a_\Lambda R_{\Lambda,\Gamma} = R_{\Lambda,\emptyset} \tag{54}$$

with $\Gamma \subset \Omega_m$, $\Gamma \neq \emptyset$ and $\Lambda \subset \Omega_m$, $\Lambda \neq \emptyset$.
This system does not have a Toeplitz structure. We can also notice that for some given sets of delays n_k, some of the values of $R_{\Gamma,\Lambda}$ for different Γs and Λs will be the same[3]. This implies that the equation system (54) contains some duplicate lines and columns. To make the system solvable, duplicate lines have to be removed and duplicate columns merged into one by addition. This operation just amounts to reducing the number of unknowns to make it equal to the order of the polynomial transfer function we want to determine. This is where the system looses its former symmetry.

[3] This happens when two different subsets Γ_1 and Γ_2 correspond to the same polynomial degree in $A(z)$, i.e. $\sum_{k \in \Gamma_1} n_k = \sum_{k \in \Gamma_2} n_k$. See for instance the case of the DRM at step 4 : we have $n_k \in \{3, 2, 4, 6\}$; considering $\Gamma_1 = \{2, 4\}$ and $\Gamma_2 = \{6\}$, we have $R_{\Lambda,\Gamma_1} = R_{\Lambda,\Gamma_2}$.