BOUNDARY LAYER RESISTANCE IN TIME DOMAIN SIMULATIONS OF THE VOCAL TRACT

Peter Birkholz and Dietmar Jackèl

Institute for Computer Sciences, University of Rostock
Albert-Einstein-Str. 21, 18055 Rostock, Germany (Europe)
phone: +49 (381) 498-3428, fax: +49 (381) 498-3426, email: {piet,dj}@informatik.uni-rostock.de

ABSTRACT

This paper describes how the boundary layer resistance can be considered in time domain simulations of the vocal tract system. The magnitude of the resistance is proportional to the square root of the frequency and is therefore difficult to include in traditional one-dimensional time domain simulations. We propose a method how the resistance can be included in such models based on a discrete representation of the velocity profile in the boundary layer. The boundary layer resistance can be considered as an important factor for high quality articulatory speech synthesis.

1. INTRODUCTION

A widely used model of the vocal tract for articulatory speech synthesis is the inhomogeneous transmission line with lumped elements [6, 8, 4]. In this model, the vocal tract is approximated by a finite number of abutting tube sections (Fig. 1 (a)). Each of them is represented by a lumped element of an inhomogeneous transmission line (Fig. 1 (b)).

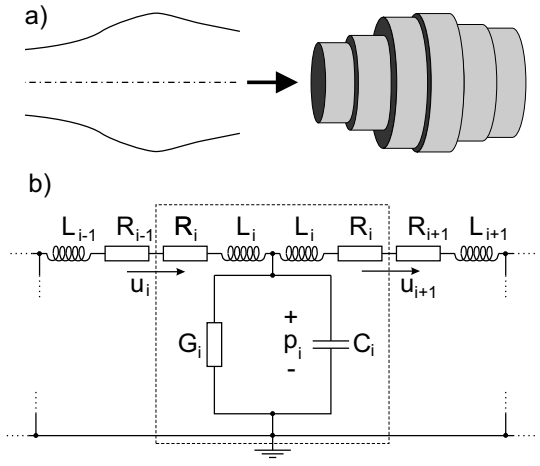

Figure 1: a) Spatial discretization of a part of the vocal tract. b) Representation of a tube section as a lumped element (dashed box) of an inhomogeneous transmission line.

In this analogy, electric current corresponds to the volume velocity u and voltage corresponds to the acoustic pressure p. The inductivities L_i and the capacities C_i stand respectively for the inertance and the compliance of the air in the tube section i. The resistances R_i and the conductances G_i stand for energy losses due to viscous friction and heat conduction.

This research was funded by the DFG (German Research Foundation).

This representation of the vocal tract allows to set up the relations between the pressures and volume velocities in the individual tube sections in a straightforward manner. Depending on the requirements, this can be done in the time or frequency domain.

For the purpose of speech synthesis, these relations are most frequently set up in the time domain by means of differential equations. The discretization of these equations is the foundation for the simulation of acoustic propagation in the vocal tract. The advantage of time domain simulations is that many dynamic aspects of speech production (acoustic interaction between glottis and vocal tract, changes of the area function over time, etc.) are automatically considered. However, a problem for the synthesis in the time domain is the correct treatment of losses due to viscous friction and heat conduction, because both R and G are frequency-dependent. Stevens investigates in [7, S. 162] the contribution of different loss mechanisms to the perceptively relevant damping of the formants and accordingly to their bandwidths. Compared to other losses, heat conduction can generally be neglected. However, the loss due to viscous friction contributes second most to the bandwidth of the formants, right after the radiation loss, for middle and high frequencies. It is therefore desirable to correctly consider this loss mechanism in time domain simulations of the vocal tract.

According to [3, 8], the boundary layer resistance is given by

$$R = \frac{lS}{2A^2}\sqrt{\frac{\rho\omega\mu}{2}}, \qquad (1)$$

where l is the length, S is the circumference and A is the area of the tube section under consideration. ρ is the ambient density of air, μ the viscosity and ω the radian frequency.

In traditional time domain simulations of the vocal tract, R was mostly either evaluated at a constant frequency or instead of Eq. (1), the frequency-independent formula for the *steady* flow resistance (Hagen-Poiseuille resistance) was used [6]. In [4], R is evaluated with the vibration frequency of the vocal folds.

Other approaches (e.g. [5], although for the particular case of a reflection-type line analog of the vocal tract) use digital filters to simulate the frequency dependence of the resistance. This strategy is more accurate than those mentioned above, but very time-consuming. Apart from that, the ideal filter with the frequency response $H(\omega) = sqrt(\omega)$ would for our purpose have to be a causal zero-phase filter, which is not realizable. Therefore, filters that are realizable have either a linear phase and a long group delay (FIR-filter) or they are an IIR-filter approximation with nonlinear phase.

For this reasons, we propose a different approach to consider the loss due to viscous friction by simulating its physical

cause, namely the velocity profile in the boundary layer. In Sec. 2 we recall the physical foundations and describe the simulation method. The results will be presented and discussed in Sec. 3.

2. METHOD FOR SIMULATING THE FREQUENCY DEPENDENT RESISTANCE

2.1 Conservation of momentum

Viscous friction is mathematically captured in the equations of motion which describe the principle of the conservation of momentum. First, we consider the case of the unsteady, one-dimensional flow through a tube, where the pressure p and the *mean* velocity $\overline{v}$ depend only on the time t and the position coordinate x. If we assume only small perturbations from the thermodynamic equilibrium and neglect perturbation terms of second order we get

$$\frac{\partial \overline{v}}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial x} - \frac{\tau}{\rho}\frac{S}{A}. \tag{2}$$

In this equation, the flow resistance of the tube is expressed by means of the shear stress τ at the wall

$$\tau = \mu \left.\frac{\partial v}{\partial y}\right|_{wall},$$

where the y-axis is oriented perpendicular to the wall in the direction of the fluid. Obviously, τ depends on the velocity profile at the wall. Hence, it is not sufficient to know only the pressure p and the *mean* velocity $\overline{v}$ (or alternatively the volume velocity $u = A\overline{v}$) during the simulation as in the one-dimensional transmission line models, but we also need to know the gradient of the flow velocity at the tube wall.
In traditional time domain simulations [4, 6], the last term in Eq. (2) is usually approximated as $R\overline{v}$ or Ru, where R depends only on the tube geometry. With these approaches, the frictional force is proportional to the volume velocity, but generally not to τ!
Since frequency dependent losses can not be included in this way, we calculate besides p and u also the concrete velocity profile in the immediate proximity of the tube wall, where the boundary layer evolves. This means that we need a separate velocity profile representation for every place in the discretized vocal tract, where the mean velocity $\overline{v}$ differs. This is the case at the entrance and at the exit of each single tube section. For a vocal tract consisting of M tube sections we thus have to simulate $2M$ profiles in addition to the pressures and volume velocities.
For the motion of the fluid in the boundary layer we assume a plane parallel shear flow in one direction. The governing equation can be derived from the Navier-Stokes equations as

$$\frac{\partial v}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{\mu}{\rho}\frac{\partial^2 v}{\partial y^2}, \tag{3}$$

which has only one velocity component $v = v(x,y)$ oriented in x-direction. This equation is apparently similar to Eq. (2). However, while the last term in Eq. (3) expresses the friction among neighbouring fluid layers, it describes the friction of the fluid 'as a whole' at the tube wall in Eq. (2).
When the width of the simulated boundary layer is δ, the following boundary conditions must hold true for any point in time: $v(x=0,y,t) = 0$ and $v(x=\delta,y,t) = \overline{v}$ [1]. Because of these two conditions Eq. (2) and (3) are interlinked.

2.2 Discretization

In order to solve the equations (2) and (3) in the simulation process, they must be discretized in spatial and time domain. First, we focus our attention to the spatial discretization of the velocity profile in the boundary layer according to Fig. 2.

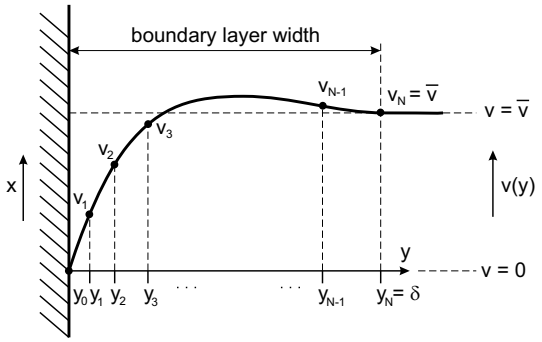

Figure 2: Discrete boundary-layer representation.

Each profile is described by means of $N+1$ interpolation points (y_i, v_i) that are distributed between the wall and the fixed maximal width of the boundary layer $\delta = y_N$. Since the gradient of the velocity profile generally decreases with increasing distance to the wall, the point density decreases towards the center of the tube, too. The boundary conditions require that for any point in time $v_0 = 0$ and $v_N = \overline{v} = u/A$.
Equation (3) describes the temporal change of the velocities at the interpolation points. Both terms on the right hand side of the equation represent different causes for a velocity change. The contribution of the first term is proportional to the pressure gradient and implies the same velocity change Δv for all interpolation points (apart from the first point at the wall). However, the velocity change due to the second term representing viscosity differs at each point. In our simulation, we evaluate both terms successively. Thereby, we accept a small error, but on the other hand we facilitate a very efficient calculation.
The velocity changes at the interpolation points due to the first term we ascribe to the change in volume velocity. The reason for this strategy becomes clear as soon as Eq. (2) is discretized. By comparison of the equations (2) and (3) we can write

$$\left(\frac{\partial v}{\partial t}\right)_1 = -\frac{1}{\rho}\frac{\partial p}{\partial x} \approx \frac{1}{A}\frac{\partial u}{\partial t}$$

and therewith $(\partial v)_1 \approx (1/A)\partial u$. This means in the time-discrete case, that the interim values $v_i^*[n]$ of the velocities at the interpolation points i at the time sample n can be calculated as

$$v_i^*[n] = v_i[n-1] + \frac{u[n]}{A} - v_N[n-1], \quad i = 1 \ldots N. \tag{4}$$

The separate evaluation of the viscosity term in Eq. (3) leads to the diffusion equation

[1] The influence of the velocity profile in the thin boundary layer on $\overline{v}$ is neglected.

$$\left(\frac{\partial v}{\partial t}\right)_2 = \frac{\mu}{\rho}\frac{\partial^2 v}{\partial y^2}. \tag{5}$$

For the discretization of this equation we implement an implicit finite-difference method which has the advantage to be unconditionally stable for any time step Δt. When the sampling index n is dropped for all velocities and furthermore $h_i = y_{i+1} - y_i$, we obtain

$$\frac{v_i - v_i^*}{\Delta t} = \frac{\mu}{\rho}\frac{(v_{i+1} - v_i)/h_i - (v_i - v_{i-1})/h_{i-1}}{\frac{1}{2}(h_{i-1} + h_i)}, \quad i = 1 \ldots N-1. \tag{6}$$

The solution for the unknown quantities $v_i[n]$ must be performed simultaneously. A fast solution method for this problem is the *Thomas algorithm* [2]. Because of the boundary condition at $y = \delta$ we have furthermore: $v_N[n] = v_N^*[n]$.

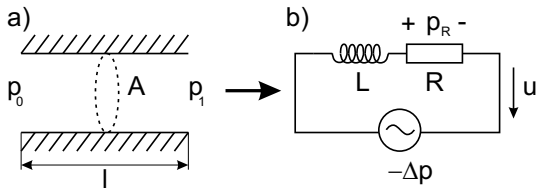

Figure 3: Simple acoustic system described by the equation of motion.

What still remains to be done is the discretization of Eq. (2) in order to determine the volume velocity $u[n]$ at each sampling point n. With the substitution of $\partial p/\partial x$ by $\Delta p/l$ and the relation $\overline{v} = u/A$ we obtain

$$\frac{\rho l}{A}\frac{du}{dt} = -\Delta p - \tau\frac{lS}{A}. \tag{7}$$

This equation directly describes the simple acoustic system in Fig. 3 (a). It consists of a tube section with the length l and the area A, which has the sound pressure $\Delta p = p_1 - p_0$ between its two open ends. The dimensions of the tube section are assumed to be small enough that the compression of the air can be neglected. With the relations $L = \rho l/A$ and $p_R = \tau Sl/A$ Eq. (7) can be rewritten as

$$-\Delta p = L\dot{u} + p_R. \tag{8}$$

This illustrates the equivalence of the problem to the electric circuit in Fig. 3 (b). When we discretize $\dot{u}$ with the trapezoid rule

$$\dot{u}[n] \approx \frac{2}{\Delta t}(u[n] - u[n-1]) - \dot{u}[n-1] \tag{9}$$

and approximate the wall shear stress τ by means of the velocity profile as

$$\tau = \mu\left.\frac{\partial v}{\partial y}\right|_{wall} \approx \mu\frac{v_1^*[n]}{h_0}, \tag{10}$$

we obtain an equation, which can be solved for the unknown volume velocity sample $u[n]$, when the pressure difference Δp is given.

2.3 Simulation

In order to evaluate the proposed method, we have examined how well the frequency dependent resistance can be simulated with the system in Fig. 3. The system was excited by sinusoidal pressure differences Δp with different frequencies. In the system flows exactly one volume velocity u. Hence, because the area of the tube is constant over the entire length, only one velocity profile needs to be simulated.
The simulation is executed as follows. For each sampling point n we first calculate the actual volume velocity $u[n]$ starting from Eq. (8). Right after that the velocities at the interpolation points are adjusted according to the equations (4) and (6).
The simulation of acoustic propagation in a system with multiple tube sections must be performed in a manner analogous to this. The difference to the simple system is, that a different velocity profile must be simulated at each place in the tube, where the mean velocity $\overline{v}$ differs.
For the simulation of the simple system in Fig. 3 we have used the following default values. The tube dimensions are $l = 5$ cm, $r = 1$ cm, $A = \pi r^2$, $S = 2\pi r$. For the boundary layer we set $\delta = 1$ mm, $N = 16$, $y_i = \delta(i/N)^2$. The sampling rate was $f_s = 88.2$ kHz and the time step accordingly $\Delta t = 1/f_s$. For each excitation frequency, the temporal evolution of the volume velocity and the pressure loss p_R due to viscous friction at the wall according to Eq. (8) was displayed graphically.

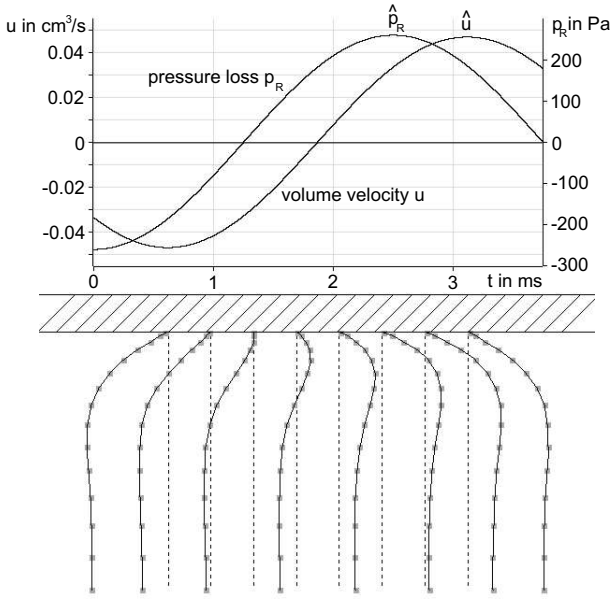

Figure 4: Top: Graphs for the volume velocity and the pressure loss due to viscous friction. Bottom: Velocity profiles in the boundary layer at different points in time.

Fig. 4 shows a piece of the graphs for u and p_R for an excitation frequency of 200 Hz. It can be clearly recognized, that the volume velocity lags the pressure loss by a small phase angle. This is also predicted theoretically for the case of sinusoidal signals [3]. The lower part of the picture shows the simulated velocity profiles in the boundary layer for different points in time according to the time-axis in the upper picture. From the quotient of the maximal amplitudes of p_R and u we have calculated the resistance due to the boundary layer:

$R_{sim} = \widehat{p}_R/\widehat{u}$. R_{sim} can directly be compared to the resistance $R(\omega)$ determined analytically by Eq. (1).

3. RESULTS AND DISCUSSION

Fig. 5 shows the results of the simulation in Sec. 2.3. The solid curve (reference line) shows the boundary layer resistance as a function of the frequency according to Eq. (1), whereas the single points represent the resistance values obtained in the time domain simulation.

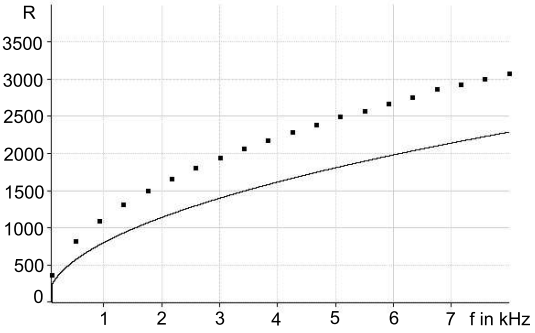

Figure 5: Exact (solid curve) and simulated (dots) resistance values.

It can be seen that all simulated resistance values lie above the corresponding reference values. The deviation is about 40% at 100 Hz and 34% at 8000 Hz. For frequencies above 8 kHz the percentage deviation keeps on decreasing.
The reasons for these differences can be found in the discretization schemes as well as in the undertaken approximations. When we tried to discretize the diffusion equation (5) with the explicit finite-difference method instead of the implicit one, all deviations lay clearly below 30%. However, with the explicit scheme only low excitation frequencies could be simulated, because the selected sampling rate f_s and the spatial discretization of the velocity profile violate the stability condition [2, p.57]. A further reason for the deviations is the separated evaluation of the two terms on the right hand side of Eq. (3). From the practical point of view, the differences between the simulated values and the references are not critical, since they can easily be compensated by a correction factor when the equations for the simulation are set up.
Besides the simple system test, described in Sec. 2.3, we also examined the impact of the boundary layer simulation on the bandwidths of the formants of different vowels in a time domain model of the vocal tract. The foundation for the simulation was the method described by Maeda in [6]. The area functions for the vowels were obtained by means of the articulatory model presented in [1].
In the implemented simulation model, we offered the possibility to switch between the presented boundary layer resistance and the frequency-independent Hagen-Poiseuille flow resistance, which was originally used in [6]. In this way, we could easily assess the improvements that can be made with the proposed method. The bandwidths of the formants were measured in the transfer functions of the vocal tract that we calculated from simulated impulse responses by means of the discrete Fourier transform. For comparison we also calculated the exact values of the bandwidths for the same area functions with a simulation of the vocal tract in the *frequency domain* [8].

Vowel	Resistance	F1	F2	F3	F4
/a:/	Hagen-Poiseuille	14.8	20.8	103.5	9.3
	Boundary layer	23.1	36.9	115.5	39.4
	Exact	20.2	34.8	114.7	30.7
/ae:/	Hagen-Poiseuille	7.1	49.3	165.7	56.5
	Boundary layer	13.3	59.2	180.5	76.1
	Exact	14.9	59.9	177.5	72.7
/u:/	Hagen-Poiseuille	4.0	5.3	2.3	1.8
	Boundary layer	6.3	9.9	16.8	24.4
	Exact	15.4	24.2	22.7	22.1

Table 1: Bandwidths of formants for different simulation models.

The results of this investigation are summarized for three vowels in table 1. The bandwidths resulting from the boundary layer simulation are substantially closer to the exact values than in the simulations with the Hagen-Poiseuille resistances. The differences could also be clearly perceived acoustically, when we simulated vowels by means of the time domain model.
In some simulated vowels with the original method from [6] we could also perceive a disturbing high frequency noise component, which probably originated from too low bandwidths of the higher formants or from numerical inaccuracies. The distinctively low bandwidths in the simulations with the Hagen-Poiseuille resistances can clearly be observed in table 1. When we used the boundary layer simulation for the synthesis of the vowels the disturbing noise component was suppressed to a great extend.

REFERENCES

[1] P. Birkholz and D. Jackèl, "A three-dimensional model of the vocal tract for speech synthesis," in *Proc. 15th ICPhS 2003*, Barcelona, Spain, Aug. 2003, pp. 2597–2600.

[2] J.-J. Chattot, *Computational Aerodynamics and Fluid Dynamics*. Springer Berlin Heidelberg, 2002.

[3] J. L. Flanagan, *Speech Analysis Synthesis and Perception*. Academic Press Inc., 1965.

[4] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," *The Bell System Technical Journal*, vol. 51, no. 6, pp. 1233–1268, 1972.

[5] J. Liljencrants, *Speech Synthesis with a Reflection-Type Line Analog*. Dissertation, Royal Institute of Technology, Stockholm., 1985.

[6] S. Maeda, "A digital simulation method of the vocal-tract system," *Speech Communication*, vol. 1, pp. 199–229, 1982.

[7] K. N. Stevens, *Acoustic Phonetics*. The MIT Press, 1998.

[8] H. Wakita and G. Fant, "Toward a better vocal tract model," *STL-QPSR*, vol. 1, pp. 9–29, 1978.