Power-balanced modelling of the vocal tract: a recast of the classical lumped-parameter model

Victor Wetzel, Thomas Hélie, Fabrice Silva

POWER-BALANCED MODELLING OF THE VOCAL TRACT: A RECAST OF THE CLASSICAL LUMPED-PARAMETER MODEL

Victor WETZEL
STMS (IRCAM-SU-CNRS)
wetzel@ircam.fr

Thomas HELIE
STMS (IRCAM-CNRS-SU)
thomas.helie@ircam.fr

Fabrice SILVA
LMA, AMU - CNRS - ECM
silva@lma.cnrs-mrs.fr

ABSTRACT

Lumped parameter models are a popular approach for the modelling of the vocal tract. They are well suited to synthesize vowels, their main advantages being their relative simplicity and low computation cost. However, difficulties arise in simulation when trying to reproduce consonants, as their corresponding configurations do not always comply with the quasi-static approximation used to derive these models.

In this communication we introduce a simple fluid mechanical model that consistently describes the power-balanced interaction between the moving walls of the vocal tract and the fluid dynamics, expressed in the port-Hamiltonian systems framework. The capacities of the model are demonstrated through simulations using a numerical scheme that preserves the power balance at discrete time. We then compare this model to the well known lumped parameter model of [Maeda, 1982] using equivalent electrical circuits which proves to be useful to understand the underlying structure that governs the exchange of power between components.

At last we discuss how we intend to extend the aforementioned model to include more complex phenomena while preserving the structure and passivity provided by the port-Hamiltonian framework.

1. PROBLEM STATEMENT AND APPROACH

There are two main types of physical models of the vocal tract: high dimensional models, relying on spatial discretization methods like the Finite Element Method [1, 2]; and low dimensional models based on Digital Waveguides [3, 4] or lumped parameter models [5–7].

These models rely on the hypotheses of linear acoustics and, in accordance, wall movements are restricted to small velocities (vibrating walls or slow movements). In practice, these models are derived from static descriptions but used for/with quasi-static configurations to produce articulations.

Our approach : We intend to connect this vocal tract to a self oscillating model of the larynx to reproduce known behaviours (e.g. fry, phonation). We want the full vocal apparatus model to satisfy the power balance (passivity of internal physics). Later on we wish to investigate the conditions of phonation on this model for healthy and pathological configuration. The preservation of this property is crucial for this work.

A minimal power-balanced self-oscillating model of the larynx already exists [8]. However such a model does not exist for a moving vocal tract: the use of linear acoustics hypothesis does not allow one to model accurately the movement of the wall for an arbitrary displacement and account for matter convection. In [9] we proposed a new simple lumped parameter model that uses a Fluid Structure Interaction (FSI) description to consistently model the interaction between the moving wall and the fluid inside the tract.

In this paper we bring the point of view of the model closer to linear acoustics while preserving its capacity to account for matter convection. We assert its capacities through numerical experiments. Eventually we compare this model to the popular one of Maeda [6]. For modelling we use the port-Hamiltonian framework which guarantees that the system satisfies the power balance. We shortly recall this tool.

Port-Hamiltonian systems : It is a framework established by Maschke and Van der Schaft [10, 11]. It extends the Hamiltonian mechanics to include dissipation phenomena and external interactions.

An open physical systems satisfies the power balance

$$\frac{\mathrm{d}E}{\mathrm{d}t} + P_{diss} - P_{ext} = 0 \tag{1}$$

where E is the energy stored by the system, P_{diss} the dissipated power (which is positive or zero) and P_{ext} the power supplied by the exterior. The full system can be interpreted as a set of components associated with each one of these phenomena.

The energy storing components are described by a state vector $\mathbf{X}$ and the Hamiltonian $\mathrm{H}(\mathbf{X})$ (energy). The dissipating components are described by the vector of dissipating variables $\mathbf{w}$ and the associated dissipation law vector $\mathbf{z}(\mathbf{w})$. The inputs of the system are gathered in the $\mathbf{u}$ vector and the outputs in the $\mathbf{y}$ vector. The product of each pair of quantity gives the associated power: $\frac{\mathrm{d}E}{\mathrm{d}t} = \nabla_{\mathbf{X}}\mathrm{H}(\mathbf{X})^{\intercal}\dot{\mathbf{X}}$ and so on.

The components are connected through a conservative interconnection, here represented by the skew-symmetric matrix $\mathbf{S}(\mathbf{X}) = -\mathbf{S}^{\intercal}$. The full system is written in its algebraic-differential formulation [12]:

$$\underbrace{\begin{pmatrix}\dot{\mathbf{X}}\\ \mathbf{w}\\ -\mathbf{y}\end{pmatrix}}_{\mathbf{f}} = \mathbf{S}(\mathbf{X}) \underbrace{\begin{pmatrix}\nabla_{\mathbf{X}}\mathrm{H}(\mathbf{X})\\ \mathbf{z}(\mathbf{w})\\ \mathbf{u}\end{pmatrix}}_{\mathbf{e}} \tag{2}$$

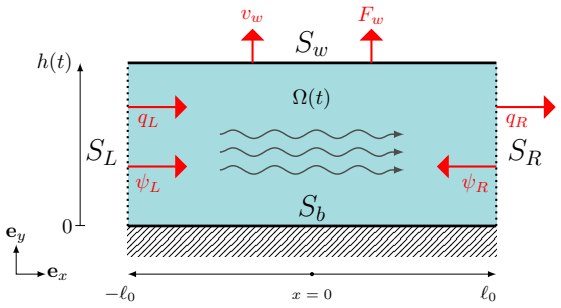

Figure 1: 2D view of the domain. External ports variables are highlighted in red

The power balance is guaranteed by the skew-symmetry of $\mathbf{S}(\mathbf{X})$. We rewrite Eq. (1) in terms of the vectors $\mathbf{f}$ and $\mathbf{e}$ and the interconnection matrix: $\mathbf{f}^\intercal\mathbf{e} = \mathbf{e}^\intercal\mathbf{S}\mathbf{e} = 0$.

Port Hamiltonian systems (abbreviated pHs hereafter) are stable by interconnection: the passive connection of two pHs is a pHs.

Organization of the paper - In Sec. 2 we recall the FSI model of [9] and modify it to take its description closer to linear acoustics. In Sec. 3, we perform two simple numerical experiments to demonstrate the capacities of the model. In Sec. 4, we establish equivalent electrical circuits and compare it with the model of [6]. Finally Sec. 5 discusses the capacities and limitations of our model.

2. MODELLING

The model we established in ref. [9] consistently includes the contribution of the moving wall in the power balance. However its description is closer to fluid mechanics than acoustics. In this section we: (i) simplify the internal energy of the gas (also called "compression energy") that models its thermodynamic behaviour ; (ii) include the rest state within the model.

In subsection 2.1, we briefly recall the model we established in [9]. Subsection 2.2 simplifies the thermodynamic part of the system under the hypothesis of small volumetric mass variations. In subsection 2.3, we rewrite the model so that it takes into account the fact that it assumes an equilibrium point at atmospheric pressure. Finally we write the full system.

2.1 Description of the original model

Here we recall the principle ingredients that constitute the model we proposed in ref. [9]: the hypothesis; the energy functions; and its port-Hamiltonian formulation. For the sake of brevity, we do not detail the computations. Note that in this article, we make and motivate a different choice of constant for the thermodynamic energy.

Geometry - The vocal tract is approximated by the concatenation of several rectangular shaped tract, as the one depicted on Fig. 1. Each tract is delimited by four surfaces: two motionless open surfaces S_L and S_R [1] ; one motionless impermeable surface S_b (bottom wall); one impermeable mobile surface S_w (top wall).

Physics (fluid mechanical point of view) - The flow is: **(H1)** irrotational; **(H2)** the volumetric mass ρ is homogeneous within the domain $\Omega(t)$ (see Fig. 1). The gas: **(H3)** verifies the perfect gas assumption; **(H4)** undergoes an isotropic and adiabatic process; **(H5)** the calorific capacity at constant volume C_V is assumed to be constant within the considered range of temperature fluctuations.

Inertial components - In [9] we supplied a velocity field that verifies the boundary conditions. Under hypothesis **(H1-2)**, the inertial energy E_i associated with the chosen velocity field is:

$$E_i(t) = \frac{1}{2}\iiint_\Omega \rho\mathbf{v}\cdot\mathbf{v}\,\mathrm{d}V = \frac{1}{2}\rho V\left(v_{mx}^2 + \frac{v_c^2}{3} + \frac{v_w^2}{3}\right) \tag{3}$$

where: v_{mx}, v_c are the degrees of freedom associated with the axial flow (x direction); v_w is the transverse velocity on the top wall (y direction); ρ is the volumetric mass density; V the volume of the tract. Hereafter this energy will be written in terms of the momenta $\Pi_x = \rho V v_{mx}$, $\Pi_y = \rho V v_w/3$ and $\Pi_c = \rho V v_c/3$.

Compression components - Under hypothesis **(H3-5)**, the internal specific energy of the gas e_c is:

$$\mathrm{e}_c(\rho) = \frac{P_0}{\rho_0(\gamma-1)}\left[\left(\frac{\rho}{\rho_0}\right)^{\gamma-1} - \gamma\right] \tag{4}$$

where: P_0 is the atmospheric pressure; ρ_0 is the volumetric mass density of air at P_0; γ is the ratio of heat capacities at constant pressure and constant volume. Computations are detailed in [9]. Here we choose to set $\mathrm{e}_c(\rho_0) = {}^{-P_0}/_{\rho_0}$ instead of zero in order to get an expression of the thermodynamics energy as close as possible to the classical linear acoustics one. (see Eq. (9d) later in this paper).

The volume-averaged pressure can be retrieved from the compression energy $\mathrm{H}_c(\rho, V) = \iiint_\Omega \mathrm{e}_c\,\mathrm{d}\Omega = \rho V \mathrm{e}_c(\rho)$ as $\langle p\rangle_\Omega = \rho^2\partial_\rho \mathrm{e}_c(\rho) = \rho/V\,\partial_\rho \mathrm{H}_c - \partial_V \mathrm{H}_c$.

Port Hamiltonian Formulation - We choose the formulation [9, eq. 12-14]. Because of the choice of the constant $\mathrm{e}_c(\rho_0)$, the compression energy slightly differs. The state vector $\mathbf{X}_1$ and the Hamiltonian $\mathrm{H}_1(\mathbf{X}_1)$ and its gradient $\nabla_{\mathbf{X}_1}\mathrm{H}_1$ are given by:

$$\mathbf{X}_1 = \begin{bmatrix}\nu_L & \nu_R & \Pi_y & V & \rho\end{bmatrix}^\intercal \tag{5a}$$

$$\mathrm{H}_1(\mathbf{X}_1) = \mathrm{H}_i(\mathbf{X}_1) + \mathrm{H}_c(\rho, V) \tag{5b}$$

$$\mathrm{H}_i(\mathbf{X}_1) = \frac{\rho V}{2\ell_0^2}\left(\nu_L^2 + \nu_R^2 - \nu_L\nu_R\right) + \frac{3\Pi_y^2}{2\rho V} \tag{5c}$$

$$\mathrm{H}_c(\rho, V) = \frac{VP_0}{\gamma-1}\left[\left(\frac{\rho}{\rho_0}\right)^{\gamma} - \gamma\frac{\rho}{\rho_0}\right] \tag{5d}$$

$$\nabla_{\mathbf{X}_1}\mathrm{H}_1 = \begin{bmatrix}q_L, & q_R, & v_w, \\ & \dfrac{-\mathrm{H}_i + \mathrm{H}_c}{V}, & \dfrac{-\mathrm{H}_i + \mathrm{H}_c}{\rho} + \rho V\partial_\rho \mathrm{e}_c\end{bmatrix}^\intercal \tag{5e}$$

where $\nu_L = \ell_0\left(v_{mx} + {}^{v_c}/_3\right)$ and $\nu_R = \ell_0\left(v_{mx} - {}^{v_c}/_3\right)$ are introduced so that the corresponding Hamiltonian gradient

[1] L for the leftmost surface, R for the rightmost surface

components are the mass flows q_L and q_R through the open surface $S_i = \{S_L, S_R\}$.

The algebraic-differential formulation is displayed in Eq. (6) (on next page). It accounts for the following inputs: ψ_L and ψ_R are the total specific enthalpy applying on the open surfaces; and F_w the net force exerted by the wall on the fluid. The corresponding outputs are the mass flow q_L and q_R, and v_{ext} the mean normal velocity on the wall.

2.2 Linearizing the thermodynamic behaviour

We will consider in this model fluctuations $\tilde{\rho}$ of the density ρ in the vicinity of the rest value ρ_0. In order to avoid numerical issues (related to the noninteger exponent), we simplify the specific internal energy defined in Eq. (4) by its first-order Taylor series approximation which, without any other approximation, lead to the compression energy:

$$\mathrm{H}_c(\tilde{\rho}, \tilde{V}) = \rho V \tilde{\mathrm{e}}_c = P_0(V_0 + \tilde{V}) \left(\frac{\gamma}{2} (\frac{\tilde{\rho}}{\rho_0})^2 - 1 \right) \quad (7)$$

which still takes into account the variation of volume, which is not the case of classical linear acoustics theory.

2.3 Including the equilibrium state within the model

The original model defines the state vector in terms of absolute quantities (see Eq. (5a)). We now wish to express it in terms of fluctuation around its equilibrium state: resting gas (no flow) at atmospheric pressure.

The functioning point $\mathbf{X}^*$ of the system under constant inputs $\mathbf{u}^* = \left[{P_0\gamma}/{(\gamma-1)\rho_0}, {P_0\gamma}/{(\gamma-1)\rho_0}, P_0 S_w \right]$ verifies $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{X}^* = 0$. This choice of $\mathbf{u}^*$ correspond to the action of the atmosphere (no flow, atmospheric pressure) on the system.

Then we have $\mathbf{X}^* = \begin{bmatrix} 0 & 0 & 0 & \rho_0 & V_0 \end{bmatrix}^\intercal$ where V_0 may be arbitrarily set. We choose to set it to its value a the initial time: $V_0 = V(t=0)$. We define a shifted state vector $\tilde{\mathbf{X}} = \begin{bmatrix} \tilde{\nu}_R & \tilde{\nu}_R & \tilde{\Pi}_y & \tilde{\rho} & \tilde{V} \end{bmatrix}^\intercal$ such as $\mathbf{X}_1 = \mathbf{X}^* + \tilde{\mathbf{X}}$.

The shifted Hamiltonian $\tilde{\mathrm{H}}$ is given by using the following relation:

$$\tilde{\mathrm{H}}(\tilde{\mathbf{X}}, \mathbf{X}^*) = \mathrm{H}(\mathbf{X}^* + \tilde{\mathbf{X}}) - \mathrm{H}(\mathbf{X}^*) - \nabla_{\mathbf{X}} \mathrm{H}(\mathbf{X}^*)^\intercal \tilde{\mathbf{X}} \quad (8)$$

with $\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{X}^*$.

Remark 1: This procedure is not a simplification or an approximation. It is simply a mere change of variable, shifting the "viewpoint" as we are now looking at the fluctuations around the state of equilibrium $\mathbf{X}^*$. The physics described by the "shifted" system is strictly equivalent to the original system with the simplified compression energy.

2.4 Port-Hamiltonian formulation of the shifted model

$$\tilde{\mathbf{X}} = \begin{bmatrix} \tilde{\nu}_L & \tilde{\nu}_R & \tilde{\Pi}_y & \tilde{\rho} & \tilde{V} \end{bmatrix}^\intercal \quad (9a)$$

$$\tilde{\mathrm{H}}(\tilde{\mathbf{X}}) = \tilde{\mathrm{H}}_i(\tilde{\mathbf{X}}) + \tilde{\mathrm{H}}_c(\tilde{\rho}, \tilde{V}) \quad (9b)$$

$$\tilde{\mathrm{H}}_i(\tilde{\mathbf{X}}) = \frac{m(\tilde{\rho}, \tilde{V})}{2\ell_0^2}(\tilde{\nu}_L^2 + \tilde{\nu}_R^2 - \tilde{\nu}_R\tilde{\nu}_L) + \frac{3\tilde{\Pi}_y^2}{2m(\tilde{\rho}, \tilde{V})} \quad (9c)$$

$$\tilde{\mathrm{H}}_c(\tilde{\rho}, \tilde{V}) = \frac{1}{2}\gamma P_0 (V_0 + \tilde{V})(\frac{\tilde{\rho}}{\rho_0})^2 \quad (9d)$$

$$\nabla_{\tilde{\mathbf{X}}} \tilde{\mathrm{H}} = \begin{bmatrix} \tilde{q_L} & \tilde{q_R} & \tilde{v_w}, \\ \frac{-\mathrm{H}_i + \mathrm{H}_c}{V}, & \frac{-\mathrm{H}_i + \mathrm{H}_c}{\rho} + m(\tilde{\rho}, \tilde{V})\partial_\rho \tilde{\mathrm{e}} \end{bmatrix}^\intercal \quad (9e)$$

where $m(\tilde{\rho}, \tilde{V}) = \rho V = (\tilde{\rho} + \rho_0)(\tilde{V} + V_0)$.

The interconnection matrix is simply obtained by substituting $\mathbf{X}$ with $\mathbf{X}^* + \tilde{\mathbf{X}}$ in the interconnection matrix $\mathbf{S}_1$ (see Eq. (6) where $\mathbf{S}(\mathbf{X} = \mathbf{X}^* + \tilde{\mathbf{X}})$).

2.5 Assembly of the full vocal tract

The assembly of two tracts is performed taking into account the constraints at the junction between tracts, i.e., the continuity of the mass rates and of the total enthalpy. We solved this problem using the equivalent component method, adapted for pHs by Najnudel et al. [13]. For the sake of conciseness, we refer to ref. [9, eq. 15] for the interconnection of the two-segments vocal tract.

We will now demonstrate the capacities of the model through numerical experiments with this two tract model.

3. NUMERICAL EXPERIMENTS

3.1 Numerical method: Corrected Discrete Gradient

We recall the passive guaranteed numerical scheme proposed by Lopes in [14]. It jointly approximates $\dot{X}$ and $\nabla \mathrm{H}$ with $\delta\mathbf{X}\, F_s$ and $\overline{\nabla}^d \mathrm{H}$ such as:

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}t} \approx \frac{\mathbf{X}_{n+1} - \mathbf{X}_n}{h} = \delta\mathbf{X}\, F_s$$

$$\nabla_{\mathbf{X}} \mathrm{H}(\mathbf{X}) \approx \overline{\nabla}^d \mathrm{H}(\mathbf{X}, \delta\mathbf{X}) \triangleq \begin{cases} \frac{\mathrm{H}(x+\delta x) - \mathrm{H}(x)}{\delta \mathbf{X}} & \text{if } |\delta x| > \epsilon_{DG} \\ \frac{\partial \mathrm{H}}{\partial x}(x) & \text{otherwise} \end{cases}$$

$$F_s \delta\mathbf{X} = S(\mathbf{X}, \delta\mathbf{X}) \overline{\nabla}^d \mathrm{H}$$

where $\overline{\nabla}^d \mathrm{H}$ is piecewisely defined in order to avoid numerical cancellation errors for small values of $\delta\mathbf{X}$. Here: $F_s^{-1} = h$ is the time step; $\delta\mathbf{X}$ the increment of the states; ϵ_{DG} is a small constant that is about the machine's precision.

This numerical scheme is usually quite robust even for non linear systems. However, we experienced divergence in our simulations even for very simple configurations. It was found to have two causes. First, in this model, we are considering de facto small fluctuations $shift\rho$ within an iterative incremental method, which leads to potential cancellation errors in the finite difference expression. Then,

$$
\begin{pmatrix} \dot{\nu}_L \\ \dot{\nu}_R \\ \hline \dot{\Pi}_y \\ \hline \dot{V} \\ \dot{\rho} \\ \hline\hline -q_L \\ q_R \\ \hline -v_w \end{pmatrix} = \underbrace{\left(\begin{array}{cc|ccc|cc|c} & & -\Pi_y/\rho V & 0 & -1/V & 1 & 0 & \\ & & \Pi_y/\rho V & 0 & 1/V & 0 & -1 & \\ \hline \Pi_y/\rho V & -\Pi_y/\rho V & & -S_w & \rho S_w/V & & & 1 \\ \hline 0 & 0 & S_w & & & & & \\ 1/V & -1/V & -\rho S_w/V & & & & & \\ \hline\hline -1 & 0 & & & & & & \\ 0 & 1 & & & & & & \\ \hline & & -1 & & & & & \end{array}\right)}_{\mathbf{S}_1(\mathbf{X}_1=[\nu_L,\nu_R,\Pi_y,\rho,V]^\intercal} \begin{pmatrix} \partial_{\nu_L}\mathrm{H}_1 \\ \partial_{\nu_R}\mathrm{H}_1 \\ \hline \partial_{\Pi_y}\mathrm{H}_1 \\ \hline \partial_V \mathrm{H}_1 \\ \partial_\rho \mathrm{H}_1 \\ \hline\hline \psi_L \\ \psi_R \\ \hline F_w \end{pmatrix} \begin{array}{l} = q_L \\ = q_R \\ = v_w \\ \\ \left.\begin{array}{l}\\\\\end{array}\right\} \frac{\rho}{V}\partial_\rho \mathrm{H} - \partial_V \mathrm{H} =< p >_\Omega \\ \\ \text{Total specific enthalpy left} \\ \text{Total specific enthalpy right} \\ \text{Net force of wall on fluid} \end{array} \tag{6}
$$

Parameter	Value
Sampling frequency F_s	$1 \cdot 10^5$
Machine's precision	$2 \cdot 10^{-16}$
ϵ_{DG}	$5 \cdot 10^{-7}$

Table 1: Parameters for the simulation

this piecewise definition of the discrete gradient has a discontinuity proportional to $\frac{\partial^2 \mathrm{H}}{\partial x^2}|\delta x|$ at switch. The small increments of the small fluctuations triggers the switch more than in other previous studies, with potential loss of accuracy.

However the fix is quite easy and twofold. First, we use a custom value of ϵ_{DG} set to the square root of the machine's precision, in order to limit the effect of cancellation error to the order of magnitude of round-off error. Secondly we extend the order of the development for the case where $|\delta x| \geqslant \epsilon_{DG}$ such as:

$$
\overline{\nabla}^d \mathrm{H} \triangleq \begin{cases} \frac{H(x+\delta x)-\mathrm{H}(x)}{\delta x} & \text{if} |\delta x| > \epsilon_{DG} \\ \frac{\partial \mathrm{H}}{\partial x} + \frac{1}{2}\frac{\partial^2 \mathrm{H}}{\partial x^2} \cdot \delta x & \text{otherwise} \end{cases} \tag{10}
$$

This corrected version of the discrete gradient has been implemented in the Python library PyPHS [15].

3.2 Configurations and value of the parameters

The full system is depicted on Fig. 2. We consider a simple case: two tracts connected by the constraints $q_{R1} = q_{L2}$, which simply states that the amount of matter exiting the left interface of the second tract enters the first tract. Soft tissues surrounding the vocal tract are modelled as a spring-damper system (one for each moving wall, without coupling) which allows a velocity-based control on the walls. The parameters' value are taken from Ishizaka et al. [16, table I] (relaxed cheek configuration, see Table 2).

As for the radiation, we add a purely resistive radiation impedance Z_r at the right interface of the second tract such as $\psi_{R2} = Z_r q_{R2}$. This would correspond to the connection of a semi infinite tube (of greater section S_r and with an anechoic end) to the exit of the vocal tract.

Static configuration - Total tract length is 17.5cm, A_1 is 0.4cm^2 and A_2 is 3.6cm^2 ($h_1 = L_1 = 2\,$cm, and $h_2 = L_2 = 6\,$cm). Walls are motionless and right open surface is a pure resistance ($Z_r = \rho_0 c/S_r$ with $S_r = 64\text{cm}^2$). We perform a 0-blocking impulse response at the input ψ_L. All other inputs are set to zero (motionless walls).

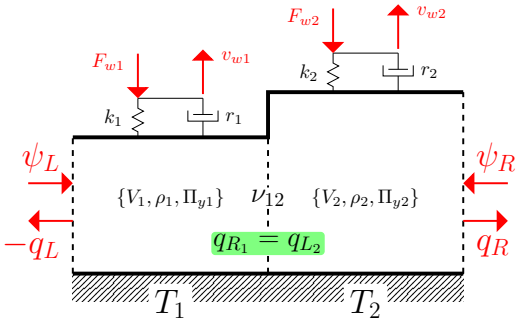

Figure 2: Configuration for the simulation: the assembly constraint is highlighted in green; input (inward arrows) and outputs (outward arrow) in red; q denotes the mass rate; ψ the total specific enthalpy ; we append a purely resistive component at the rightmost surface S_R.

Dynamic configuration - In this experiment, only the wall of the rightmost tract moves (see Fig. 3. The total duration of the simulation is 3s with three steps: (i) $t \in [0; 0.7\text{s}[$ the vocal tract is set with the parameters of the static configuration and remains motionless; (ii) $t \in [0.7; 0.75\text{s}[$ the second tract closes, going from height $h = 6\text{cm}$ to $h = 3\text{cm}$; (iii) the system remains motionless until the end at $t = 3\text{s}$. Linear fades in and out (duration 5 ms) are added to the corresponding velocity to smooth out these transitions.

We perform a 0-blocking impulse response at the beginning of step (i) and step (iii), precisely at $t = 0s$ and $t = 2.3\text{s}$.

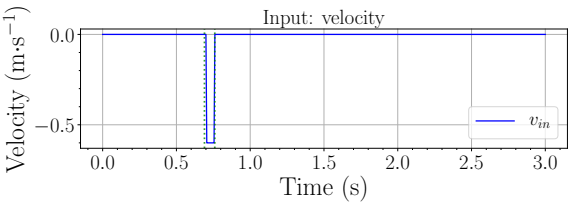

Figure 3: Inputs and outputs at the walls

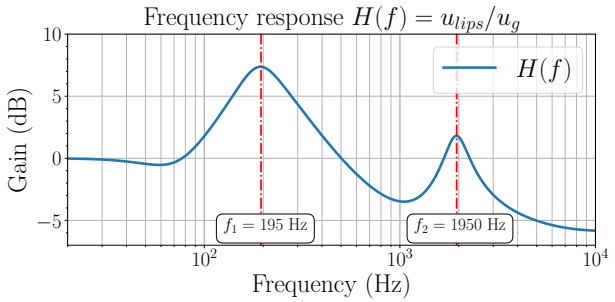

Figure 4: **Static configuration -** Frequency representation: the system exhibits two acoustical resonances at 196Hz and 1950Hz

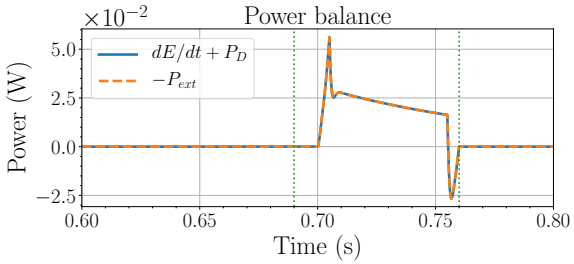

Figure 5: **Dynamic configuration - Focus on step (ii):** $t \in [0.6; 0.8]$ **-** Power balance for the dynamic configuration: green dots shows the limits between each step; power balance is satisfied with a maximum local error of $2.5 \cdot 10^{-14}$W (at the end of step (ii))

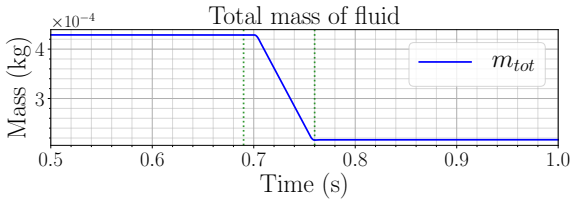

Figure 6: **Dynamic configuration -** Evolution of the mass of fluid contained in the system: here we show the time section $t \in [0.5; 1.0]$; the total mass diminishes as the tract closes.

3.3 Results

As the system is non linear we cannot establish transfer function for a given configuration. Therefore we propose to plot $H(f)$ the ratio of the Fourier transforms of the 0-blocking impulse response for the flow rates at the glottis $u_g(f)$ and at the lips $u_{lips}(f)$ (respectively the leftmost and rightmost open surfaces). For convenience we will call it "frequency representation" as it allows us to identify the resonances of the system.

Static configuration - The frequency representation of the 0-blocking impulse response is shown on Fig. 4.

We identify two resonances with frequencies $f_1 = 196$Hz and $f_2 = 1950$Hz, and the first resonance being higher in amplitude (about 5dB). The low number of tracts causes frequency warping (2 tracts to approximate a 17.5cm vocal tract), which is similar to the one affecting spatial finite difference schemes.

Dynamic configuration - Here we may enhance three aspects. First, the simulation is passive as visible on Fig. 5: the power balance is preserved throughout the movement of the wall, the sum of the power stored and the power dissipated equating with the power flowing through external ports. Moreover this model accounts for matter transport even after shifting around the rest state. In fact air is being ejected out of the system as the volume of the second tract and the total amount of mass within the vocal tract expectedly decrease as shown on Fig. 6.

Finally, Fig. 7 depicts the frequency response before and after the movement (step (i) and (iii)). The frequencies of the acoustical resonances evolve as the geometry changes from $f_1 = 196$Hz and $f_2 = 1950$Hz to $f_1 = 140$Hz and $f_2 = 3800$Hz.

Remark 2: The model has been shifted around the equilibrium point associated with atmospheric conditions. However it still accounts for matter convection, which, in terms of amplitude, may be bigger than the acoustical values (typically about the same magnitude as the wall's velocity v_w).

Using a passive guaranteed numerical scheme, we achieved simulations for a simple model. We have shown that the displacement of the wall is accurately reproduced, and that the system exhibits a number of acoustical resonance related to the number of tracts. Increasing the latter should allow a better control of the resonance frequency and thus of the formants of the synthesized signals.

4. EQUIVALENT ELECTRICAL CIRCUITS

We now focus on an electrical circuit representation in order to better understand the coupling of the various physical phenomenon included in the model. For the sake of brevity, we establish the circuit of the fully non linear and non shifted system (see Eqs. (5) and (6)) as their interconnection structure are strictly equivalent.

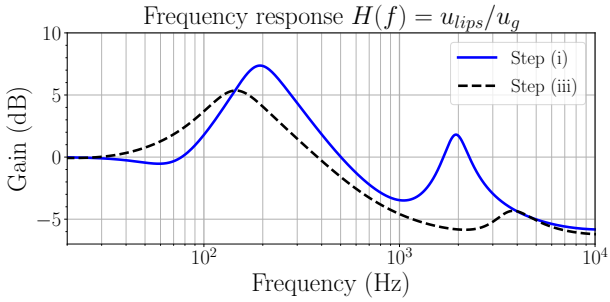

Figure 7: **Dynamic configuration -** Frequency response before (solid line) and after (dashed line) the motion of the right tract's wall (respectively step (i) and (iii)).

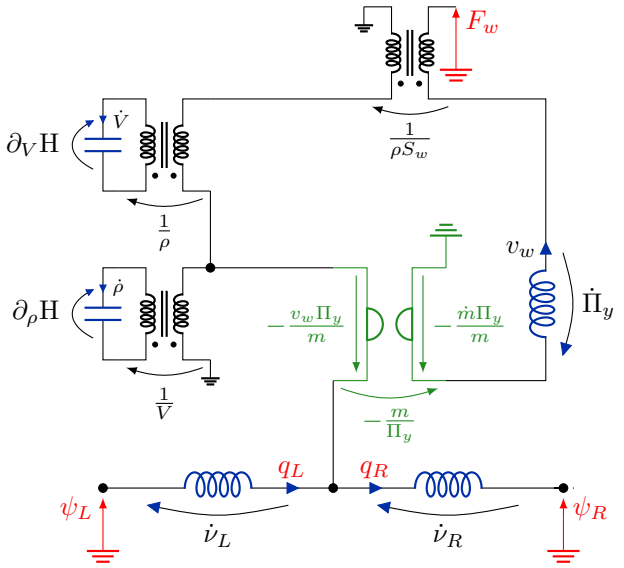

Figure 8: Equivalent electrical circuit of the first model (see subsection 2.1): potential is analogous to total specific enthalpy; current is analogous to mass flow rate; $\mathbf{X} = [\nu_L\, \nu_R\, \Pi_y\, V\, \rho]^\intercal$; in red, external ports; in blue energy-storing components; in green, conservative/dissipative components.

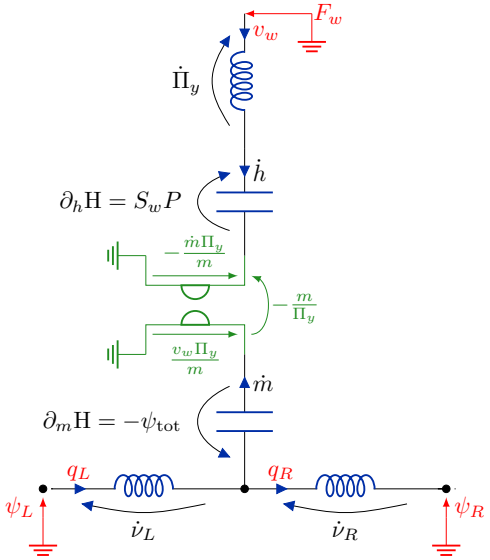

Figure 9: (Equivalent and to be compared with circuit shown on figure 8) - Equivalent electrical circuit: potential is analogous to total specific enthalpy; current is analogous to mass flow rate; $\mathbf{X} = [\nu_L\, \nu_R\, \Pi_y\, h\, m]^\intercal$; color scheme remains identical as in Fig. 8.

4.1 Application to our model

We choose the following analogy: the total specific enthalpies ψ and mass rates q are represented as voltages and currents, respectively. With this convention, inertial components are represented by inductors and compliant components by capacitors.

The structure of the axial flow (x direction) is quite simple with only two inductors associated with the inertia ν_L and ν_R. The inductor Π_y (in the upper branch) accounts for inertia of the transverse flow (y direction). The gyrator and the two capacitors (and their associated transformers) forms a capacitive quadripole which can be interpreted as a kind of "2D thermodynamic spring" that couples both axial and transverse flows.

4.2 Simplifying the structure

The structure of this quadripole seems pretty complicated. This appears to result from the choice of variables (ρ, V) to account for compression. This is due to the intensive nature of the volumetric mass density.

In classical thermodynamics, a given system can be described by a set of extensive variables, in this case: the number of moles N; the volume occupied by the gas V. The pressure (with a minus sign) is thus defined as the partial derivative of the internal energy w.r.t. V.

In our case we wanted to remain close to classical acoustical variables and chose to use the intensive variable $\rho = {}^m\!/_V$ which encodes both information about mass and geometry. This information redundancy between ρ and V is responsible for the rather complicated structure of the circuit 8.

To simplify the structure, we performed a variable change between from the set of compression variable (ρ, V) to (m, h) (full port Hamiltonian formulation in Appendix 9.2). The corresponding equivalent circuit is shown in Fig. 9,

Overall, the structure is much simpler[2] as we eliminated all transformers and loops within the circuit. It also allows for a better physical interpretation. We cut the circuit in half as if the gyrator was the ground. The components below the gyrator model the behaviour of a compressible flow through a rigid pipe (constant volume). The ones above the gyrator model the behaviour of a gas in a closed piston (constant mass). Both phenomenon are coupled with a gyrator so that the full system accounts for mass and volume variation.

4.3 Comparison with the lumped parameter model

In this section we compare our model with the lumped parameter model of Maeda [6]. More precisely we compare the underlying structure of both model using their associated equivalent electrical circuits.

The model of Maeda is build from the linear acoustics hypothesis. The quasi-staticity hypothesis is assumed so that the wall can move freely. Therefore the analogy is different from our fluid mechanical model: current denotes the flow rate; voltage denotes the acoustical pressure. Its equivalent electrical circuit is depicted on Fig. 10.

We now compare this model to ours. For the both of them, the axial flow is represented by two inductors. The rest of the circuit is connected between these two inertial components. Both circuits exhibit the same "T-shaped"

[2] which is also visible in the interconnection matrix, now only including 0 and $+/-1$, see Eq. (11).

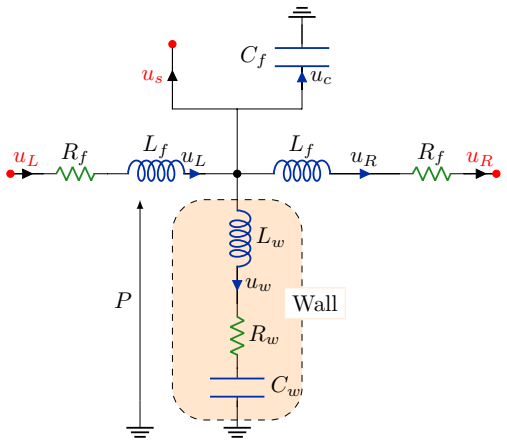

Figure 10: Equivalent electrical circuit of Maeda [6]: current is flow rate; voltage is the acoustical pressure; to be compared with the circuit shown on Fig. 9

structure.

These are the only similarities. Our model has an inductor associated with the transverse flow. The fluid's compliance is described by a set of components that accounts for both mass and volume change. Moreover the contribution of the moving wall, the fluid's transverse mass and the fluid's compliance are in series. These caracteristics are absent in the model of Maeda.

In his model, Maeda adds a simple loss model that accounts for viscous losses (resistance of a laminar flow in a cylindrical duct) represented by the two resistances R_f in the axial branch. It allows the wall to get very close to the bottom one without the fluid accelerating in a non-physical manner. We wish to add a similar dissipation mechanism in our model as it will extend its capacities and realism.

5. DISCUSSION

5.1 Capacities of the model

Our model consistently describes the balanced exchange of power at the moving wall. Despite its fluid mechanics point of view, it uses a description that is close to acoustics. It is a relatively simple non linear model with only five components and a gyrator. Its pH formulation ensures the power balance for any movement of the wall.

5.2 Future improvements

As we mentioned earlier, our model lacks a flow dissipation component. It is not necessary for vowel configurations, but remains critical when a tract is almost closed, which is the case for stop consonants. Without such a dissipation, the fluid accelerates in a non-physical manner. To address this we wish to add the same kind of resistance model as the one of Maeda (see end of subsection 4.3).

A more ambitious improvement is to model the scenario where the tract is completely closed ($h = 0$, $m = 0$). This can be achieved with switching pHs (see ref. [10]) of constraints. For a certain value of h, the system will switch between two different states: open tract and closed tract. This changes the tract's model but also those of the contiguous tract. This could be critical for the modelling of stop consonants.

6. SUMMARY

In this paper we derived a FSI model from [9] that accounts for linear fluctuations of the volumetric mass density. Therefore it describes both the FSI behaviour (e.g. matter convection) as well as the acoustical one. With numerical experiments and a tailored passive numerical scheme, we showed that the model indeed is able to reproduce these behaviours. Finally we established equivalent electrical circuits which helped better understand the structure of the model and compare it with the one of Maeda [6]. We then discussed the future possible improvement to further extend the capacities of the model.

7. ACKNOWLEDGMENTS

T. Hélie acknowledges ANR-DFG project INFIDHEM (ANR 16 CE92 0028).

8. REFERENCES

[1] T. Vampola *et al.*, "FE Modeling of Human Vocal Tract Acoustics. Part I: Production of Czech Vowels," *Acta Acust united Ac*, vol. 94, no. 3, pp. 433–447, 2008.

[2] M. Arnela and O. Guasch, "Two-dimensional vocal tracts with three-dimensional behavior in the numerical generation of vowels," *J. Acous. Soc. Am.*, vol. 135, no. 1, pp. 369–379, 2014.

[3] J. L. Kelly and C. C. Lochbaum, "Speech synthesis," in *Stockholm Speech Commun Seminar, RIT, Stockholm*, 1962.

[4] V. Välimäki and M. Karjalainen, "Improving the kelly-lochbaum vocal tract model using conical tube sections and fractional delay filtering techniques," in *3rd Int. Conf. Spoken Language Processing*, 1994.

[5] J. L. Flanagan *et al.*, "Synthesis of Speech From a Dynamic Model of the Vocal Cords and Vocal Tract," *Bell System Tech. J.*, vol. 54, no. 3, pp. 485–506, 1975.

[6] S. Maeda, "A digital simulation method of the vocal-tract system," *Speech Comm.*, vol. 1, no. 3-4, pp. 199–229, 1982.

[7] P. Birkholz, "Articulatory Synthesis of Singing," *Proc. Interspeech*, pp. 4001–4004, 2007.

[8] T. Hélie and F. Silva, "Self-oscillations of a vocal apparatus: A port-hamiltonian formulation," in *GSI*, vol. 10589 LNCS, pp. 375–383, 2017.

[9] V. Wetzel *et al.*, "Power balanced time-varying lumped parameter model of a vocal tract: modelling and simulation," in *26th ICSV*, Montréal, Canada, 2019.

[10] A. Van Der Schaft, "Port-hamiltonian systems: An introductory survey," in *Proc. Int. Congr. Math.*, vol. 3, pp. 1339–1365, 2006.

$$
\left(\begin{array}{c}
\dot{\nu}_L \\ \dot{\nu}_R \\ \dot{\Pi}_y \\ \dot{h} \\ \dot{m} \\ \hline\hline v_w \\ \mathrm{d}m \quad (=\dot{m}) \\ \hline\hline -q_L \\ q_R \\ -v_{\text{ext}}
\end{array}\right) =
\left[\begin{array}{cc|c|c|c}
 & & -1 & -1 & +1 \\ \hline
 & & +1 & +1 & -1 \\ \hline
 & & -1 & -1 & +1 \\ \hline
 & +1 & & & \\ \hline
+1 & -1 & & & \\ \hline\hline
 & +1 & & & \\ \hline
+1 & -1 & & & \\ \hline\hline
-1 & & & & \\ \hline
 & +1 & & & \\ \hline
 & -1 & & &
\end{array}\right]
\left(\begin{array}{c}
q_L \\ q_R \\ \hline v_w \\ \hline \langle P\rangle_\Omega\, S_w \\ -\psi_{\text{tot}} \\ \hline\hline F_m = -\frac{\Pi_y}{m}\,\mathrm{d}m \\ \psi_m = v_w\frac{\Pi_y}{m} \\ \hline\hline \psi_L \\ \psi_R \\ F_{\text{ext}}
\end{array}\right) \tag{11}
$$

[11] B. Maschke and A. van der Schaft, "Port-controlled hamiltonian systems: Modelling origins and system theoretic properties," *IFAC Proc. Vol.*, vol. 25, no. 13, pp. 359 – 365, 1992.

[12] A. Falaize and T. Hélie, "Passive Guaranteed Simulation of Analog Audio Circuits: A Port-Hamiltonian Approach," *Appl. Sci.*, vol. 6, no. 10, p. 273, 2016.

[13] J. Najnudel *et al.*, "Analog circuits and Port-Hamiltonian realizability issues: a resolution method for simulations via equivalent components." in *Audio Eng. Soc. Conv.*, New York, United States, 2018.

[14] N. Lopes *et al.*, "Explicit second-order accurate method for the passive guaranteed simulation of port-Hamiltonian systems," *IFAC-PapersOnLine*, vol. 48, no. 13, pp. 223–228, 2015.

[15] A. Falaize, "PyPHS website," (last access 2020-10) https://pyphs.github.io/pyphs/.

[16] K. Ishizaka *et al.*, "Direct determination of vocal tract wall impedance," *IEEE Trans. Acous., Speech, and Signal Proc.*, vol. 23, no. 4, pp. 370–373, 1975.

9. APPENDIX

9.1 Values of the parameters used in the numerical experiments

Symbol	Value	S.I. Unit
P_0	$101.325 \cdot 10^3$	$\mathrm{kg \cdot m^{-1} \cdot s^{-2}}$
γ	1.4	
ρ_0	1.225	$\mathrm{kg \cdot m^{-3}}$
ℓ_1	0.0425	m
ℓ_2	0.0425	m
ℓ_{total}	0.175	m
k_1	845	$\mathrm{N \cdot m^{-1}}$
r_1	0.8	$\mathrm{kg \cdot s^{-2}}$
k_2	845	$\mathrm{N \cdot m^{-1}}$
r_2	0.8	$\mathrm{kg \cdot s^{-2}}$
S_{ray}	0.064	$\mathrm{m^2}$
Z_{ray}	$\rho c_0 / S_{ray}$	$\mathrm{kg \cdot m^{-2} \cdot s^{-1}}$

Table 2: 2 tracts configuration parameters

9.2 Simplifying the structure: computation details

First, we express the gyrator structure as a matrix/vector multiplication.

$$
\begin{aligned}
\mathbf{w} &= \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} v_w \\ q_L - q_R \end{pmatrix} \\
\mathbf{z}(\mathbf{w}) &= \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -\Pi_y/m \\ \Pi_y/m & 0 \end{pmatrix}}_{G_z} \mathbf{w}
\end{aligned} \tag{12}
$$

Because of the skew symmetry of G_z, the power dissipated by the interconnection is zero: $\mathbf{w}^\intercal \mathbf{z}(\mathbf{w}) = P_{\text{diss}} = 0$.

We choose to express the compression part of the system as a function of the mass of fluid m and the height of the tract h such as:

$$
\rho = \frac{m}{S_w h} \quad ; \quad V = S_w h \tag{13}
$$

Under these two changes, the system becomes:

$$
X = \begin{bmatrix} X_L & X_R & \Pi_y & h & m \end{bmatrix}^\intercal \tag{14a}
$$

$$
\begin{aligned}
\mathrm{H}(\mathbf{X}) = {} & \frac{1}{2}\frac{m}{\ell_0^2}\left(\nu_L^2 + \nu_R^2 - \nu_L\nu_R\right) + \frac{3\Pi_y^2}{2m} \\
& + \frac{P_0 S_w h_0}{\gamma - 1}\left[\left(\frac{m}{m_0}\right)^\gamma \left(\frac{h_0}{h}\right)^{\gamma-1} - \gamma\frac{m}{m_0}\right]
\end{aligned} \tag{14b}
$$

$$
\nabla \mathrm{H} = \begin{bmatrix} q_L & q_R & v_w & \langle P\rangle_\Omega\, S_w & \psi \end{bmatrix}^\intercal \tag{14c}
$$

The algebraic-differential formulation is shown in Eq. (11).