State space models for wind-instrument synthesis

Denis Matignon, Philippe Depalle and Xavier Rodet
IRCAM, Analysis/Synthesis Department
31, rue Saint Merri, 75004 Paris, France

ABSTRACT

We present a new approach for sound synthesis, as an attempt to unify the two main categories of sound synthesis techniques: signal-modelling and physical-modelling. This approach uses state space models which allow a precise physical description of the instruments, and the estimation of some unknown parameters by Kalman filtering techniques. The paper focuses on the automatic construction of the state space equations of a class of connected acoustic tubes.

1. INTRODUCTION

This paper deals with the use of state space models for sound synthesis, as an attempt to unify the two main categories of sound synthesis techniques developed until now (Smith 1991): signal-modelling and physical-modelling.

The first one makes use of fairly simple models (e.g. additive synthesis (Risset 1971), or source-filter model (Rodet, Depalle and Pourot 1980)), but allows for automatic extraction of parameters from sampled signals (e.g. FFT, or Linear Prediction). There are two major drawbacks with this signal approach: from a synthesis point of view, the simplicity of the model cannot account for all natural perceptual effects, since the internal structure of the instrument is not specified; and from an analysis point of view, automatic extraction of the parameters of the models often fails when they vary with time.

The second one is based on more complicated models, offering an internal description of the system under study, since it starts from physical principles. The sounds delivered by physical models are of high quality. But one of the drawbacks of the physical approach lies in the difficulty to measure certain control parameters, and thus to simulate the behaviour of the instrument with accurate parameters.

Consequently, we want to keep positive aspects of both previous approaches while avoiding their major drawbacks. To achieve our goal, we use the wide range of techniques provided by control system theory. State space models offer interesting features: on the one hand the internal structure of the system can be precisely described, on the other hand some known parameters can be estimated, using Kalman filtering techniques. And, for example, one can imagine a feedback loop to automatically control the stability of a note played by a "state space instrument".

In the present paper, we are particularly concerned with the automatic construction of the state space equations of a class of connected acoustic tubes.

2. MODEL CONSTRUCTION

The model is based on electroacoustical analogies with transmission lines, and uses the classical lattice network representation. The systems are built from lossless or lossy tubes and boundary conditions; for more details on the subject, we refer to a former publication (Depalle, Rodet, Matignon and Poulliaout 1992).

2.1 Lossless tube

Considering an elementary cylindrical tube (whose length is directly determined by the sampling rate) with its right hand side discontinuity, we can represent it by a single delay and a reflection coefficient r, and then use the classical lattice formulation that leads to much simpler state space equations. The ongoing waves are the controls, and the outgoing waves are the observations.

The concatenation of N elementary cylindrical tubes with different section areas, made through N-1 junctions that we name $J_2$, gives a tube of length $N$; we have obtained the formal calculation of the matrices of the N-input, N-output, N-state variable subsystem.

Then, we have developed a recurrent procedure to concatenate these N subsystems through a $J_3$ junction (fig.0), in order to get the state space equations of a lossless system; notice that the second subsystem, a fingerhole, must be connected once its output boundary condition has been applied. The figure below represents a step of the recurrent construction:
2.2 Lossy tube

Using a more elaborate electroacoustical analogy, we can describe our elementary cylindrical tube by a transmission line including viscothermal losses through resistance \( R \) and condutance \( G \); we have found a new lattice formulation: it becomes non symmetrical, and the \( k^2 \) reflection coefficients are now complex valued (but fortunately not conjugate). A first order approximation, which is justified in the case of small losses, leads to:

\[
k^2 = (\gamma + \delta) e^{2\pi i} \Rightarrow \gamma = \delta = 2\pi i.
\]

The first term is the lossless reflection coefficient \( r \), the two others condense in the junction the influence of viscothermal losses on the acoustic impedance \( \delta \) and the propagation within the tube \( \epsilon \). Both \( \delta \) and \( \epsilon \) are complex valued and frequency dependents; but, in order to get a true space formulation of the system, we use rational approximations of order 1 or 2 (Gustafsson 1991) instead of the frequency dependent factor \( (\epsilon)^{\frac{1}{2}} \).

As in the previous case, it is possible to perform the formal calculation of the matrices of a concatenation of \( N \) elementary cylindrical tubes with different section areas through \( N-1 \) lossy junctions.

Then, the former recurrent construction holds, but with lossy \( J_j \) junctions.

2.3 Input boundary condition

This represents the loading of the source on the system. A simplified model uses a constant coefficient, named \( R_0 \) since it accounts for the interaction with the glottis in a model of the vocal tract; a dependence upon frequency would lead to better approximation.

But in fact, one of the interests of the state space formulation lies in the possibility of coupling physical systems together: it is therefore interesting to connect the input of our tube system with a non-linear oscillator (Roden 1992), such as a reed.

2.4 Output boundary condition

This represents the loading of the outer space on the system, which is well-known as the radiation effect. A simplified model can be designed with a constant coefficient, named \( R_f \) because it accounts for the radiation of the lips in a model of the vocal tract.

But, by definition, \( R_f \) is a rational fraction in terms of the acoustic impedance of the last elementary tube and the radiation load impedance: taking the results from theoretical acoustics (Brunetti 1983), we can express \( R_f \) as a function of frequency and then approximate it by an ARMA model; two configurations have been modelled: the radiation of the last tube in a semi-infinite space, and an infinite space.

3. EXPERIMENTAL RESULTS AND APPLICATIONS

In the preceding section, we saw what the elementary modules are and how they can be connected one with another. From the state space representation, an external ARMA representation can be formally obtained. Thus, we can test...
our model on the subclass of tubes of variable section area with no fingerholes and with boundary conditions, by examining the transfer function between the single input and the single output (SISO).

3.1 Models of the vocal tract
In the leakless model, the transfer functions of the larynx filter found in the cases $R_L = \pm 1$ exactly corresponds to the well-known Levinson polynomials; care must be taken that such a system exhibits infinite resonances (i.e., poles located on the unit circle); therefore, it is preferable to implement $|R_L| < 1$
In the leaky model, we can express the transfer functions of a system of connected leaky tubes loaded by a radiation impedance; we obtain new polynomials: extended Levinson polynomials, the structure of which is rather complicated. But we can calculate them using formal calculus.

3.2 A worked out numerical example
Let us consider an example of application of a two tube system of lengths $2L$ and $L$, where $L = 96.5$ mm; diameters $d_1 = 10$ mm, $d_2 = 20$ mm. We use a sampling rate of 32 kHz (the first tube has $p = 36$ elementary tubes, and the second $q = 18$). The input boundary condition is $R_{G1} = 1$, the output boundary condition is $R_L = 0.99$ or radiation loud; we choose the half-space radiation loud: the output of the tube is through a large piece of wood, for example. In the most general case, the extended Levinson polynomial of our system is:

$$X(z) = \frac{1}{2} \left[ R_{G1} e^{2 \pi i L / 2} z - z^{-1} R_L e^{2 \pi i L / 2} z + \frac{1}{2} \right]$$

We compare the effects of leaky and radiation on the logarithm of the magnitude of the frequency response of the tube.

On fig. 1, we show the influence of radiation on leaky models: we observe a model of three modes that regroups and is damped as high frequencies: in fact, the physical system can be seen as three tubes of equal length L. The solid line represents a constant coefficient, and the dotted line represents the half-space radiation: notice the gradual shift and damping of the second mode.

On fig. 2 and 3, we show partial views at low and high frequencies; curves $n=1$ are the unit model with $R_L = 0.99$; the resonances are very sharp and show no damping at high frequencies. Curves $n=2$, in plain line, are the leaky model with $R_L = 0.99$; the resonances are all the more damped and shifted as frequency increases. Curves $n=3$ are the leaky model with half-space radiation; the second mode has disappeared at high frequencies.

4. CONCLUSION
State space models seem to be very promising, keeping positive aspects of both signal-modelling and physical-modelling.
Our model allows a wide range of applications, extending from signal processing (generation of a frequency response to filter a source) to physical modelling (digital simulation of a woodwind instrument).
In the very near future, further developments will consist in the realization of a formal calculus program for building synthesizers and the connection of sources.

5. ACKNOWLEDGEMENTS
The DREF of the French DGA is gratefully acknowledged for financing the thesis of Denis Manigotton.

6. REFERENCES
fig. 1: Influence of radiation

fig. 2: Partial view at Low Frequencies

fig. 3: Partial view at High Frequencies