Virtual Musical Instruments : Contribution to physical modeling and control of self-sustained instruments

Thomas HÉLIE*, Christophe VERGEZ†, Xavier RODET*,
Ircam Centre Georges Pompidou, 1 pl. Igor Stravinsky, Paris, 75004, France

*** Analysis-Synthesis Team, † Instrumental Acoustics Team.**

Keywords: Virtual instrument, Music, Physical Modeling, Behaviour, Control, Inversion.

1 INTRODUCTION

The use of synthesizers and computers sometimes deprives musicians of the experience of playing. One of the goals of physical models[1] is to provide musicians with virtual instruments. These new instruments should react to the musician's control, either by a slight or a dramatic change, as natural instruments would; phrasing, articulation qualities and expressivity should also be reproduced automatically, as a response to musician's gestures.

Potential applications include computer assistance for instrument makers, creation of imaginary instruments, extraction of skilled musician gestures deduced from a sound and pedagogical use, etc... The achievement of virtual instruments requires multiple fields of research including mechanics, numerical simulation, man-machine interaction, analysis of dynamics and oscillating patterns and inversion of dynamical systems.

In the first part of this article, fundamental aspects commonly found in physical modeling of self-sustained instruments are considered: from experimental observations (§2.1) and building of models (§2.2) to their behavioral (§2.3) and gestural (§3) analysis, each proceeding is illustrated through several examples (bowed-string instruments, wind instruments, and more particularly trumpet-like instruments [2]).

Then, it is shown that modeling, behavioural and gestural analysis (in regards to the experimental validations) as a global problem, have direct consequences on the types of expressions used for modeling. Then, we discuss, for excitors (§4.1) and for resonators (§4.2), such expressions which are looked for so that they are well-adapted to the manipulations required by the numerical simulation of the model and by the gestural analysis process.

[1]The laws of mechanics and acoustics are used to describe the natural phenomena accountable for the instrument sound production.

[2]mainly taken from a previous study made by the authors on trumpet-like instruments [1].

2 MODELS OF MUSICAL INSTRUMENTS: STATE OF THE ART

2.1 Experiment

The functioning of a given instrument may be studied experimentally,

- by characterizing the instrument behaviors, i.e. the different *oscillating patterns* when control parameters [3] **G**, representing the musician's *gestures* in the broad sense, are varied (transitions between two patterns correspond to bifurcations of the instrument),
- then more precisely, by analyzing the temporal waveform of dynamic variables[4] for a given *oscillating pattern*, these variables representing the *internal state* **S** of the instrument.

For such experiments and studies, artificial and controlled playing conditions must be re-created: bowing machine for bowed-strings instruments [2], artificial mouth for simple-reed instruments [3], artificial lips for brass instruments (see Fig. (1)) [4], [5].

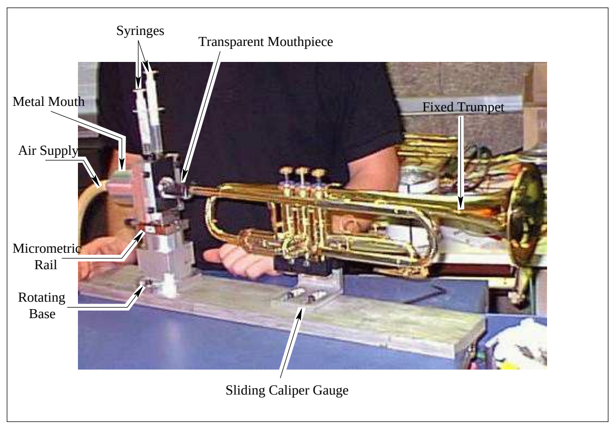

Figure 1: Photograph of an artificial mouth device with water-filled latex lips (from [5]).

[3]bow velocity and pressure, blowing pressure, lips stiffness, ...

[4]positions, speeds, pressures, ...

tions Let $\mathcal{P} \subset \mathbb{R}^{dim\,\mathbf{G}}$ be the control parameter space. The different *oscillating patterns* (periodic, n-periodic, quasi-periodic, chaotic, intermittent, etc...) are located[5] in the regions $\mathcal{R}_\xi$. The set $\mathcal{C}$ of these regions $\mathcal{R}_\xi$ is a partition of $\mathcal{P}$ the *cartography*. From the studies of such cartographies, experimentally established for various instruments, it can be concluded that:

(a) any instrument offers a great diversity and richness of behaviors;

(b) for instruments of a given class, their cartographies $\mathcal{C}$ all display the same topographic structure. These structures, characteristic of each class have been studied in [6] (bowed strings), [7] (clarinet) and [8] (brass).

Timbre - Ease of play - Instrument quality However, the observed invariance (b) does not mean that

(c) the waveforms of the state $\mathbf{S}$ in a given region $\mathcal{R}$,

(d) or that the localizations (form, size, position) of the region $\mathcal{R}$ in a given cartography $\mathcal{C}$,

are unchanged by instrument-making adjustments[6] or conditioning of the instrument[7] .

In additition, it is well-known from instrument makers and musicians that such modifications directly affect the quality of the instrument:

- the timbre and the potential expressivity depend on the diversity of the waveforms of $\mathbf{S}$ (for timbre) and and of its amplitudes (for nuances), found in a given region $\mathcal{R}$;
- the ease of play depends on the size of the regions of $\mathcal{C}$, but also on the agreement of the ranges of $\mathbf{G}$ with the player's physiological capacities .

Some apparantly weak changes may even induce dramatic effects. For instance, the inner surface of certain african flutes has to be covered with a thin layer of water so that the flute can play [10]. Another example concerns good brass players who can feel the ageing of their instrument: this could be due to the damage of the polish inside the instrument, which alters visco-thermal interactions between the air column and the inner surface [11].

To sum up, the sound quality of any instrument can depend on the most subtle physical factors, even if behavioural analyses show that within a class of instrument, the cartographies $\mathcal{C}$ for different instruments have all the same topographic structure.

[5]These properties are studied for constant parameters (in time).

[6]such as material or shape choices (ex: dimensions of the mouth and the chamfers for the recorder [9], etc...).

[7]hygroscopy for the reeds, type and quantity of rosin for bows, etc...

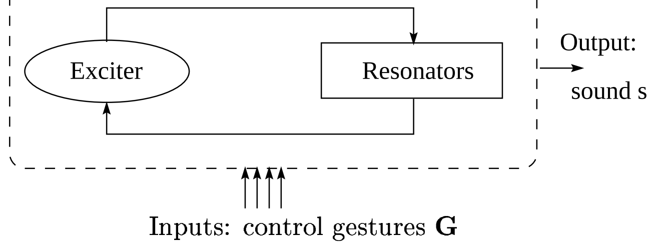

Figure 2: Common functional structure for self-sustained musical instruments.

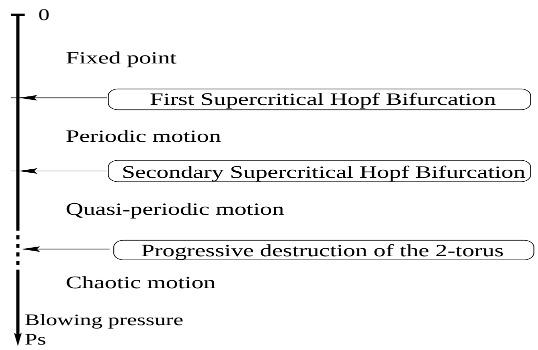

Figure 4: Example of a sequence of bifurcations of a brass model [14] when the blowing pressure is progressively increased from zero and other parameters are fixed.

2.2 Physical modeling

Since Helmholtz [12], it is well established that sustained instruments functioning relies on the general structure displayed in figure 2: a feedback loop system made of an *exciter* and *resonators*. The input is made of the control parameters $\mathbf{G}$ and the only output is the sound s.

The *resonators* are usually modeled at first approximation by linear systems which introduce delays due to wave propagation. The *exciter* is usually a mechanical[8] system which is coupled to resonators. The relations modeling the exciter and the coupling may be written as *non-linear differential equations*[9] since a stable auto-oscillation can only be generated by a dynamical system which is non-linear.

2.3 Behavioural study of a model

Recent studies have shown that physical models present oscillating patterns comparable (see Fig. (3)) with the experimentally observed ones (§2.1): a period doubling scenario for a clarinet model [7], periodic and non periodic oscillating patterns for both a violin model [13] and a trumpet model [14] (see Fig. (4)).

Moreover, most of the dynamical behaviours

[8]fluid or solid.

[9]Note that even if the wave propagation inside a resonator is non-linear (case of brass instruments for strong pressures), this effect is not responsible for the oscillation, but mainly for the brightness of the sound due to the transfer of energy from low to higher harmonics.

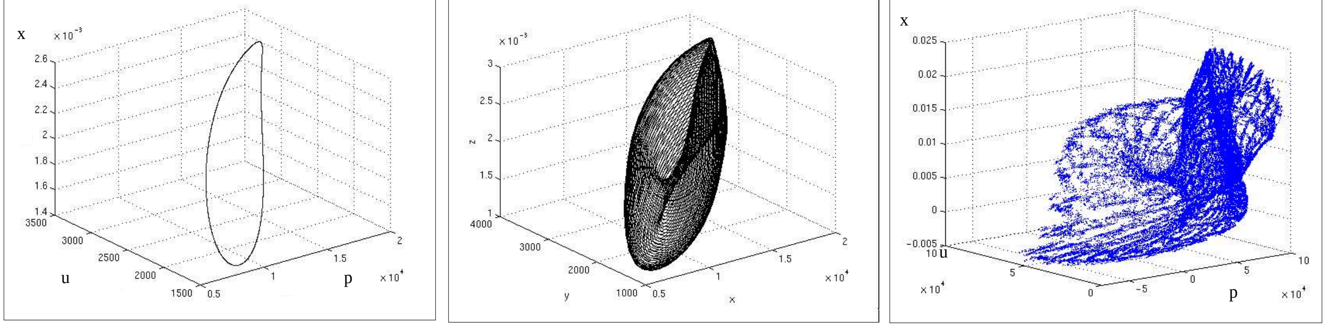

Figure 3: Various simulated oscillating patterns (from left to right: periodic, quasi-periodic and chaotic) for a brass model when the blowing pressure parameter is increased (from [14]). x, y and z axis are respectively air pressure and volume flow between the lips, and lips opening.

have proved to be perceptively relevant ([15], [16], [17], [18]): they are immediately associated to the corresponding instrument class by a listener.

If simplified models are used[10], cartographies $\mathcal{C}$ still keep similar topographic structures and the represented instruments are still recognized despite notable sonic differences. This strenghtens the argument that these cartographies and their "invariant complexity" are representative of the instrument classes. But, at the same time, this complexity makes the control of a virtual instrument as difficult as the one of a real instrument.

3 INVERSION OF A PHYSICAL MODEL

In this section, we study the identification process of the control trajectories which make the model produce a target sound.

3.1 Definitions and technical approach

The inversion of the dynamical system modeling the behavior of an instrument answers the question: "how should I control my instrument to obtain a target sound?". Then, writting the relations of the model[11]

$$\mathcal{M}(\mathbf{G}, s, t) = \mathbf{0} \tag{1}$$

the inversion consists in determining the inputs $\mathbf{G}(t)$ of the system when its output $s(t)$ (the *sound*) is known.

There are two main approaches to deal with this problem: to determine the reciprocal application of the model starting from the equations of this model (this step is called inversion), or to use the so-called "black box" techniques such as model training approaches.

In addition to the choice of the sound features used for the training, the reliability of this last method requires browsing the various model behaviors as widely as possible. On the contrary, the first approach exploits, in an *a priori* and exhaustive way, all the knowledge brought by the model (including behaviours, transients, multiphonics, etc...). This is why, the first approach is detailed in the next section.

3.2 Inversion of physical models of self-sustained instruments: an ill-posed problem

It has been shown in a physical models of brass instruments [20] that for each sound a non countable set of solutions $\mathbf{G}$ exists. As an example (Fig. (5)), the damping factor $d(t)$ and the contraction factor $c(t)$ of the model of the lips evolve on a surface which is a function of the sound $s(t)$.

In fact, this property is common to all the self-sustained models: if the vector of the gestural parameters $G(t)$ has a dimension n, the inversion leads to an *under-defined* system which simply imposes $\forall t,\ G(t) \in \mathcal{M}_t$ where $\mathcal{M}_t$ is usually a hypersurface of dimension $n-1$. This means that changing a gestural parameter does not change the sound produced if the other parameters are consequently modified.

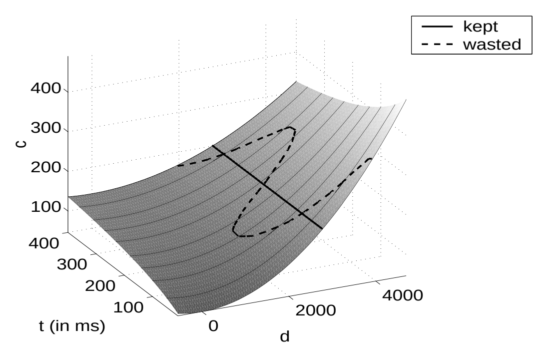

Figure 5: Example of a solution set of gestures (d, c) for a brass model

But practically, these solutions are not all physically meaningful. Indeed, most of them have unrealistic magnitude orders, are oscillating too fast, or sim-

[10] provided the "shape of the non-linearities" is similar.

[11] Note that these relations are quite general ($\mathcal{M}$ may contain differential operators, delays, etc...) but have a strong particularity: the internal state $\mathbf{S}$ does not appear! Indeed, this form, which can be usually obtained for models of self-sustained instrument models by eliminating the states variables (see for ex. [19]), is very useful in the inversion process.

musician's ability has not yet been taken into account in the inversion process !

3.3 Selection of the musician's gestures

As defining and using a physical model of musician is a complicated task, the prefered choice is to make a hypothesis on the "right gestures" to be selected

- by looking for the best gesture, i.e. the one which minimizes a given criterion $\mathcal{F}(\mathbf{G}, t)$ during all the playing time,
- or by imposing a parametric shape of gestures $\mathbf{G}^{\mathbf{p}}(t)$ on consecutive time intervals of small duration δ.

Then, the resolution of the inverse problem can be respectively obtained by:

- minimizing $\mathcal{F}(\mathbf{G}, t)$ under the constraint that $\mathbf{G}(t)$ makes the model produce the target sound
$$\mathbf{G}^*(t) = \underset{\mathbf{G}}{\arg} \min_{\mathbf{G}, \mathbf{\Lambda}} \int_{t_1}^{t_2} \Big[\mathcal{F}(\mathbf{G}, t) + \mathbf{\Lambda}(t)\mathcal{M}(\mathbf{G}, s, t)\Big] \, dt \quad (2)$$
where $\mathbf{\Lambda}(t)$ is the Lagrangian multiplicator. An interesting *cost function* is $\mathcal{F}_W(\mathbf{G}, t) = \|W\mathbf{G}'(t)\|^2$ where W is a weighting matrix, and for which the optimal gestures are the *lower ones.*

- looking for the best parameters p^* such that[12]
$$\mathbf{p}^* = \underset{\mathbf{p}}{\arg} \min \int_{\tau}^{\tau+\delta} \left\|\mathcal{M}(\mathbf{G}^{\mathbf{p}}, s, t)\right\|^2 \, dt \quad (3)$$
where simple examples of gestural evolution are
$$\text{the constant gestures}: \quad \mathbf{G}^{\mathbf{a}}(t) = \mathbf{a} \quad (4)$$
$$\text{the affine gestures}: \quad \mathbf{G}^{\mathbf{a},\mathbf{b}}(t) = \mathbf{a}t + \mathbf{b} \quad (5)$$

A first study concerning a model for the lips of a trumpet player (with gestural evolution Eq. (4)) showed that a single solution was obtained. Moreover, for tests on synthetic sounds, the method is efficient even for transients and multiphonic sounds [20].

3.4 Conclusion

In addition to the obvious applications of gestural control, benefits of inversion are also important for modeling improvements: usually, the validation of a model is done by comparing the proximity of the outputs $s_{\text{exp}}(t)$ and $s_{\mathcal{M}}(t)$ produced by calibrated controls $\mathbf{G}(t)$, for respectively, the real instrument and the modelled/simulated one. However, the closeness of two non-linear systems may be difficult to measure since for a given $\mathbf{G}(t)$, outputs may be very different[13].

[12] Note that $\mathbf{G}^{\mathbf{p}}(t)$ does not exactly generate $s(t)$, unlike the first method.

[13] such a difference could correspond for example to a slight shift between basins of attraction of the model and of the instrument.

for a given sound $s(t)$, the experimental and the reconstructed gesture.

Therefore, *inversion* plays a role in the virtual instrument research as an important as *modeling*, *behaviour analysis* in regards to *experimental validations.*

4 CONSEQUENCES ON MODELING

Now, we focus on the investigation of the best means for modeling the *exciter* and *resonators*: these models should be as concise and physically realistic as possible, and above all, well-adapted to the manipulations required by the model simulation and the inversion process. Moreover, strategies to cope with unavoidable properties which are ill-adapted for these manipulations are proposed.

4.1 Exciter and coupling

Even if some simplifications encountered in the literature that work for certain instruments [14] lead to non-differential instantaneous relations, these types of expressions are exceptional. Brass instruments for which inertia of the lips cannot be neglected, is an obvious counter-example. This justifies why, as seen in §2.2, the simplest formulation used for modeling an excitor and the coupling is a non-linear ordinary *time-differential* equation (or a system of such equations).

However, for many instruments, some relations included in Eq. (1) may be degenerated on small time intervals. Moreover, differential systems are not the best suited formulation for modeling certain phenomena such as "noisy" sounds (e.g. breath for the wind instruments or friction for the violin). These both difficulties are adressed below.

Local degeneration of some relations. For many instruments, some of the n relations given by Eq. (1) can become degenerated (*ie* $0 = 0$) on small time intervals $\mathcal{I}$. This occurs, for example, when the lips (brass) or the reed and the mouthpiece (clarinet) are in a closed configuration. In such cases, a well-adapted method for realizing the inversion consists of locally removing the p degenerated relations and using the $n-p$ remaining relations. In the constrained criterion methods (Eq. (3)), the langrangian multiplicator dimension $\dim \Lambda(t)$ is then $n-p$ ($\forall t \in \mathcal{I}$).

Practically, the time intervals $\mathcal{I}$ may be identified from a wrong conditionning calculus in the computation of the inversion process. Then, the resolution is done for the $n-p$ relations modeling the instrument on $\mathcal{I}$.

"Noisy" sound components. Even if "noisy" components may be produced in real intruments by deterministic mechanisms (turbulences, chaos), to use a stochastic processes $\mathbf{N}(t)$ in physical

[14] ex: neglecting the mass of the reed for the clarinet and saxophones, or inertia of the bow for the violin, etc...

models can be preferable. In fact, real mechanisms usually lead to complex models whereas a statistical approach can lead to simpler phenomenological models. Moreover, to use a *stochastic non-linear time differential system* does not complicate simulation. However, the criteria used in Eq. (3) and Eq. (3) are not adapted any more. A solution consists in respectively modifying them in the following way:

$$\mathbf{G}^*(t) = \underset{\mathbf{G}}{\arg} \min_{\mathbf{G},\mathbf{N},\mathbf{\Lambda}} \int_{t_1}^{t_2} \Big[\mathcal{F}(\mathbf{G},t) + \mathcal{P}_{\mathbf{G}}(\mathbf{N},t) + \mathbf{\Lambda}(t)\mathcal{M}\big(\mathbf{G},s,t\big)\Big]\, \mathrm{d}t \qquad (6)$$

$$\mathbf{p}^* = \underset{\mathbf{p}}{\arg} \min_{\mathbf{p},\mathbf{N},\mathbf{\Lambda}} \int_{\tau}^{\tau+\delta} \Big[\mathcal{P}_{\mathbf{G}}(\mathbf{N},t) + \mathbf{\Lambda}(t)\mathcal{M}\big(\mathbf{G}^{\mathbf{p}},s,t\big)\Big]\, \mathrm{d}t \qquad (7)$$

where the *cost function* $\mathcal{P}_{\mathbf{G}}(\mathbf{N},t)$ which penalizes $\mathcal{F}$ may be chosen as the *entropic* disorder measure of the noise $\mathbf{N}(t)$ defined by $-\ln f_{\mathbf{G}}\big(\mathbf{N}(t)\big)$, where $f_{\mathbf{G}}$ the probability density.

4.2 Resonator and radiation

Radiation and resonators are the place of propagation phenomena. Their precise modeling is complex since they require the resolution of *partial differential equations* (*P.D.E.*) for arbitrary geometries and boundary conditions, which include non-linearities for large amplitude waves.

Because of this complexity, various approaches are commonly found in the literature:

i acoustic measurements on the resonators (impedance [21] & [22], reflectometry [23], measurements in an anechoic room) allows one to compute numerical acoustical transfer functions (e.g. impulse response or reflection function in the time domain);

ii appropriate simplifications (simple geometries, non dissipative medium, linearisation, etc...), allow one to solve the resulting *P.D.E.* :

- **ii**$_1$ analytically, leading to a continuous *delayed system* (*D.S.*),
- **ii**$_2$ numerically, leading to a discrete distributed *D.S.* ,

(ii$_1$) and (ii$_2$) giving parametrical models.

For the global problem (simulation, inversion and model validation), a parametric model (ii) is more appropriate than a non parametric one (i). Indeed, it may easily encompass all the configurations of a resonator (e.g. fingerings and intermediate states during a change of fingerings, sophisticated playing techniques using for example half pressed keys or valves) whereas (i) would require a large table of sampled functions. Moreover, the unavoidably truncated functions may under-estimate or ignore low frequency and very long memory phenomena[15]. This may cause robustness problems for inversion since the compensation of these under-represented phenomena may drive the process to wrong trajectories, and then perceptibly disturb it.

Hence, using a parametric model increases the quality of both the model simulation and the inversion process if long memory effects and transitions are modelled. Moreover, this allows us to envisage an *adaptative inversion process* which would adjust the instrument making parameter values[16]. The simplest hoped *D.S.* modeling propagation is a pure delayed plus a linear ordinary time-differential equation. This formulation which becomes more complex as soon as realistic geometries for resonators are considered keep usually its linear property.

However, nonlinear propagation phenomena and degeneretion of relations may occur and cause some difficulties.

Nonlinearity: as seen in section 2.2, this property naturally arises for several instruments at high sound levels. It does not hinder the use of Eq. (3) or Eq. (3) but makes their numerical resolution more complicated. In particular, the nonlinear character of the resonator is likely to increase the number of possible solutions. Therefore global nonlinear optimisation techniques [24] should be used to select the final solution.

Structural degeneration: structural degeneration in models of resonators are observed when an equation disappears (or a new equation appears) in the system of equations modeling the resonator. In practice, this corresponds to the closure/opening of a hole, the pressure of a finger down on a string, a change of valve position ... Two differents strategies are proposed:

- Considering in parallel as many models as needed. The selection of which model to invert is done according to model breakdown criteria. This solution is quite simple but computationally expensive.

- Using a refined physical model of the resonator which encompasses all the configurations (as suggested above). Then, the pursuit of the transition parameters (e.g. positions of fingers) allows one to follow these configurations. Numerical resolution becomes again more complicated but transients during different notes can be inverted as well as the above-mentioned sophisticated playing techniques.

---

[15] Working in the Fourier's transform space does not generally suppress these problems which are converted into a "bin resolution" one.

[16] A method consists for example in including these parameters in $\mathbf{G}$ and in increasing their associated weighting in the matrix W (Eq. (3)).

5 CONCLUSION

After a first description of the state of the art for physical modeling of self-sustained musical instruments, the inversion process naturally appeared as a useful tool for good control of such models. In addititionto this benefit, this tool also gives new possibilities for checking the validy of the models and improving them (see Fig. (6)).

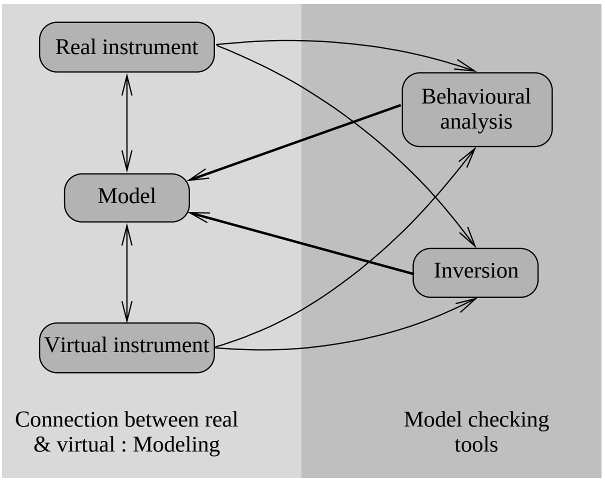

Figure 6: Links between modeling, behavioural analysis and inversion.

Then, *modeling, behavioural analysis and inversion (in regards to the experimental validations)* may be gathered in a global problem. With this in mind, the study of virtual musical instruments leads us to prefer parametric expressions to numerical descriptions: practically, it is better to match experimental measures on a model and to use this model than, for example, to numerically calculate transfer functions from these measures and to use them as the model.

REFERENCES

[1] T. Hélie, C. Vergez, J. Lévine, and X. Rodet. Inversion of a physical model of trumpet. In *Proceedings of ICMC99*, Pékin, China, October 1999.

[2] G. Weinreich and R. Caussé. Elementary stability considerations for bowed-string motion. *J. Acoust. Soc. Amer.*, 89(2):887–895, 1991.

[3] J. Gilbert. *Etude des instruments de musique à anche simple: extension de la méthode d'équilibrage harmonique, rôle de l'inharmonicité des résonances, mesure des grandeurs d'entrée*. PhD thesis, Université du Maine, Novembre 1991.

[4] J. Gilbert and J. F. Petiot. Brass Instruments, Some Theoretical and Experimental Results. In *ISMA Proceedings*, pages 391–400, 1997.

[5] C. Vergez and X. Rodet. Model of the trumpet functioning: real time simulation and experiments with an artificial mouth. In *Proceedings ISMA'97*, Edinburgh, August 1997.

[6] J. C. Schelleng. The bowed string and the player. *J. Acoust. Soc. Amer.*, 53:26–41, 1973.

[7] C. Maganza, R. Caussé, and F. Laloë. Bifurcations, period doublings and chaos in clarinetlike systems. *Europhysic Letters*, 1(6):295–302, March 1986.

[8] J. S. Cullen, J. Gilbert, and D. M. Campbell. Brass instruments: linear stability analysis and experiments with an artificial mouth. *Acta Acustica*, 86(4):704–724, July/August 2000.

[9] Claire Ségoufin. *Production du son par interaction écoulement/résonateur acoustique*. PhD thesis, Université de Paris VI, 2000.

[10] Fabre, C. Ségoufin, R. Caussé, and C. Vergez. Ouldémé flutes. Work in progress.

[11] J. Gilbert. Les Cuivres (tutorial html pages). http://www.loa.espci.fr/sfa/fr/gsam/jp/cuivres /textes/cuivres.htm#_Toc500315537, 2000.

[12] H. L. F. Helmholtz. *"On the sensation of tone"*. Masson, 1877. Reprinted by Dover, from the first edition in 1954.

[13] J. Puaud, R. Caussé, and V. Gibiat. Quasi-périodicité et bifurcations dans la note de loup. *J. Acoustique*, 4:253:259, 1991.

[14] C. Vergez and X. Rodet. Trumpet and trumpet player: a highly nonlinear interaction studied in the framework of nonlinear dynamics. *International Journal of Bifurcation and Chaos*, 2001. (to appear in July).

[15] M. P. Verge. *Aeroacoustics of confined jets, with applications to the physical modeling of recorder-like instruments*. PhD thesis, Technische Universiteit Eindhoven, 1995.

[16] C. Vergez and X. Rodet. Comparison of real trumpet playing, latex model of lips and computer model. In *Procedings ICMC'97*, Thessalonike, September 1997.

[17] J. M. Adrien and J. Morison. MOSAIC: A Modular Program For Synthesis Using Modal Superposition. In *Modèles Physiques: Création Musicale et Ordinateur*, volume 2, pages 371–384. ACROE, 1990. Current name of MOSAIC is MODALYS.

[18] E. Ducasse. Modélisation d'instruments de musique pour la synthèse sonore: application aux instruments à vent. In *Proceedings du* 1^{er} *Congrés Français d'Acoustique*, pages pp C2–837:C2–840, February 1990.

[19] T. Hélie. Étude de méthodes d'estimation des paramètres d'un modèle physique d'instrument de musique - application à la trompette. Master's thesis, DEA ATS - Université de Paris XI, 1998.

[20] T. Hélie, C. Vergez, J. Lévine, and X. Rodet. Inversion of a physical model of a trumpet. In *Proceedings of CDC99*, Phoenix, USA, December 1999.

[21] J. Backus. Input Impedance Curves for the reed woodwind instruments. *J. Acoust. Soc. Amer.*, 56:1266–1279, 1974.

[22] R. Caussé, J. Kergomard, and X. Lurton. Input Impedance of brass musical instruments. Comparison between experiment and numerical models. *J. Acoust. Soc. Amer.*, 75:241–254, 1984.

[23] David B. Sharp. *Acoustic pulse reflectometry for the measurement of musical wind instruments*. PhD thesis, University of Edinburgh, 1996.

[24] E. Walter and L. Pronzalto. *Identification de modèles paramètriques à partir de données expérimentales*. Masson, Paris, 1994.