OBJECT-BASED SOUND SYNTHESIS

Augusto Sarti, Stefano Tubaro

Dipartimento di Elettronica e Informazione, Politecnico di Milano
Piazza L. Da Vinci 32, 20133 Milano, Italy
E-mail: sarti/tubaro@elet.polimi.it

ABSTRACT

The physical modeling of complex sound generators would be feasible if approached by individually synthesizing and discretizing the objects that contribute to the generation of sounds. This raises the problem of show to correctly implement the interaction between these objects. In this paper we show how to synthesize sounds in an object-based fashion, i.e. by building objects individually synthesized and making them interact with each other through the modeling of a potential interaction topology. We will also show how this interaction topology can be made dynamical and time varying.

1. INTRODUCTION

The physical synthesis of sounds [1, 2] consists of modeling the vibrational phenomena that occur in a complex resonating structure, which can be made of a number of simpler resonators connected together. The vibrational phenomena are normally caused and, possibly, sustained by the interaction with other structures. This way of looking at physical model synthesis suggests an object-based approach to the modeling of sounds, which requires a strategy that allows us to manage all possible interactions between individually synthesized objects, by planning and implementing the interaction topology and solving all possible computability and stability problems beforehand.

One major difficulty in this approach, however arises when we need to connect together two discrete-time models, each of which exhibits an instantaneous connection between input and output. In fact, the direct interconnection of the two systems would give rise to a delay-free loop (an implicit equation) in their implementation algorithm. This problem usually occurs when we try to connect together two individually discretized systems without taking into account any global interconnection constraint. Inserting a delay element in the non-computable loops (i.e. deciding an artificial ordering in the involved operations) or solving the relative implicit equation involves a certain cost or risk in the final digital implementation, especially when discontinuous nonlinearities are present in the model. In fact, too simple a solution will tend to modify the system's behavior and, often time, to cause severe instability. Conversely, a more sophisticated iterative solution will dramatically increase the computational cost, as an implicit equation will have to be solved at each time instance. As a matter of fact, it would be highly desirable for a block-based synthesis strategy to be able to preserve the stability properties of the analog reference system. In fact, this would allow us to select a sampling frequency that is only related to the involved signal bandwidths, rather than to the adopted discretization strategy. In other words, we would like to keep the oversampling factor (of the temporal discretization) as low as possible, without giving up the physicality or the behavioral plausibility of the system. Unlike what it may seem, this problem is, in fact, quite critical when highly nonlinear elements are involved in the model implementation, which is our case not just because systems may be intrinsically nonlinear, but because contact conditions are modeled by step functions.

2. WAVE DIGITAL STRUCTURES

A physical structure (mechanical or fluidodynamical) can be described by an electrical equivalent circuit made of lumped or distributed elements. The equivalence can be established in a rather arbitrary fashion as a physical model is always characterized by a pair of extensive-intensive variables (e.g. voltage-current, force-velocity, pressure-flow, etc.), and reciprocity principles can always be invoked. For example, if we wanted to model the hammer-string interaction in a piano we could first select a simplified model of the actual piano mechanism, and then adopt an electrical equivalent of it, as shown in Fig. 1. In this case the equivalence is established by having forces and velocities correspond to voltages and currents, respectively.

Using the electrical equivalent of the sound-production mechanism provides us with a standard representation of physical models. However, this representation cannot be digitally implemented using a local approach, as a direct interconnection of individually discretized elements would give rise to problems of computability. This is to be attributed to the fact that, when using extensive-intensive (vol-
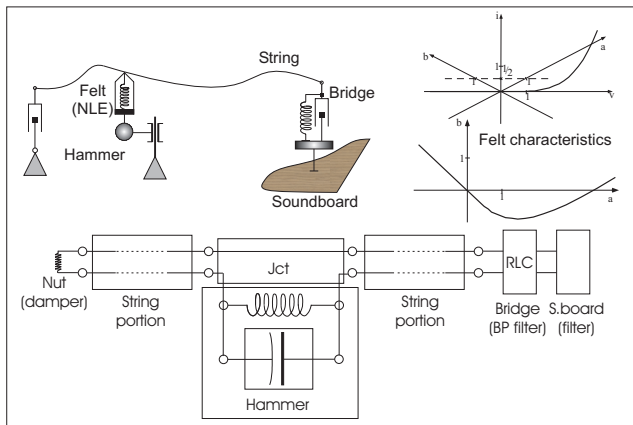
Fig. 1. Construction of the electrical equivalent of a piano model. When the hammer is in contact with the string, the velocities of hammer and string are the same at the contact point, therefore the contact junction is a series junction (current corresponds to velocity, voltage corresponds to force).

tage-current) pairs of variables, a direct interconnection of the blocks will not account for global constraints such as Kirchhoff laws. One way to overcome this difficulty is to describe the system by means of scattering parameters. This allows us to exploit the concept of adaptation in order to avoid computability problems. A well-known "local" method for designing filters after linear circuits, which is based on this approach, is that of Wave Digital Filters (WDF's) [3]. The method consists of adopting a different pair of "wave" variables $a = v + Ri$ and $b = v - Ri$ for each element of the circuit, R being a free parameter called "reference resistance". This corresponds to a linear change of reference frame, from a (v, i) pair to an (a, b) pair, performed with a linear transformation with one degree of freedom (reference resistance R). The global constraints (Kirchhoff laws) are modeled in the interconnection phase, using multi-port series and parallel adaptors, which also account for all the changes in the reference frames from point to point. The degree of freedom in the specification of the reference frame can be exploited to satisfy an adaptation condition on one of the ports of each adaptor. An adapted port, in fact, will not exhibit a local instantaneous wave reflection, thus guaranteeing that no computability problems will take place.

One key aspect of WDF's is the fact that they preserve many properties of the analog filters that are used as a reference, such as passivity and losslessness [3]. Because of that, in the past few years we witnessed renewed interest in WDF's as the research in musical acoustics started to turn toward synthesis through *physical modeling*. This interest in WDF's is also due to the popularity gained in the past few years by *Digital WaveGuides* (DWG's) [8], which are

close relatives of WDF's. Such structures, in fact, are suitable for modeling resonating structures in a rather versatile and simple fashion.

The similarity between DWG's and WDF's is not incidental, as the former represent the *distributed*-parameter counterpart of WDF's. In fact, they both use (incident and reflected) waves and scattering junctions. Thanks to such similarities, WDF's and DWG's turn out to be fully compatible with each other. However, while DWG's waves are defined with reference to a *physical* choice of wave parameters such as propagation velocity and characteristic impedance, the reference parameters for WDF's waves represents a degree of freedom to be used to avoid computability problems.

It is quite clear that hybrid WDF/DWG structures seem to offer a flexible solution to the problem of sound synthesis through physical modeling. One should keep in mind, however, that both the classical WDF theory and the DWG theory are inherently linear, which raises the problem of how to incorporate nonlinearities into a generic Wave Digital (WD) structure, as they are predominant in musical acoustics.

Nonlinear elements can be quite easily incorporated in WDF structures by exploiting that one degree of freedom that WDF structures have in the combination of reference resistances. In fact, this allows us to adapt the port where the nonlinear element needs to be connected to. Since the wave variables are either voltage or current waves, nonlinear elements that can be incorporated in classical WDF structures are resistors, and their wave nonlinearity (a $b - a$ curve) can be obtained from the Kirchhoff characteristic (a $v - i$ curve) using the transformation that defines wave pairs (a, b) in terms of Kirchhoff pairs (v, i). Nonlinear resistors, however, are only a subset of the nonlinearities encountered in musical acoustics. Among the simplest ones are those nonlinearities are that have a nonlinear capacitors or a nonlinear inductors as their electrical equivalent.

Modeling such nonlinearities with classical WDF principles is known to give rise to problems of computability, since closed loops without delays cannot be avoided in the resulting WD structure. In order to avoid such problems, a solution for a wave implementation that includes reactive nonlinearities was proposed in [6]. In this solution, new waves were defined in order to be suitable for the direct modeling of algebraic nonlinearities such as capacitors and inductors. In fact, with respect to the new waves, the description of the nonlinear element became equivalent to that of a resistor. In order to adopt such new waves, a special two-port element that performs the change of variables is defined and implemented in a computable fashion. The reactive nonlinear element is thus modeled in a new WD domain, where its description becomes memoryless. Roughly speaking, with respect to the new wave variables, the behavior of the nonlinear bipole becomes resistor-like, therefore the two-port junction that performs the change of wave vari-

ables plays the role of a device that transform the reactance into a resistor.

A further extension of these ideas can be found in [7], where a more general family of digital waves is defined, which allow us to model a wider class of nonlinearities. This generalization of WDF principles include dynamic multiport junctions and adaptors, which synergically combine ideas of nonlinear circuit theory (mutators) and WDF theory (adaptors). This generalization provides us with a certain degree of freedom in the design of WD structures. In fact, not only can we design a dynamic adaptor in such a way to incorporate the whole dynamics of a nonlinear element into it, but we can also design a dynamic adaptor that will incorporate an arbitrarily large portion of a linear structure. It can be easily proven [7] that, under mild conditions on their parameters, such multiport adaptors are nonenergetic, therefore the global stability of the reference circuit is preserved by the wave digital implementation. For this reason, such multiport junctions can be referred to as dynamic *adaptors*.

The class of digital waves that we use for modeling a "port" in the WD domain is basically of the form

$$A(z) = V(z) + R(z)I(z) \,, \ B(z) = V(z) - R(z)I(z) \,,$$

where $R(z)$ is a "reference transfer function" (RTF). With this choice, the class of nonlinearities that can be modeled in the WD domain is, in fact, that of all algebraic bipoles of the form

$$p = g(q) \,, \quad P(z) = H_v(z)V(z) \,, \quad Q(z) = H_i(z)I(z) \quad ,$$

where p and q are related to v and i, respectively, through a finite difference equation, while $R(z) = H_v(z)/H_i(z)$. The above choice of digital waves allows us to model a wide class of nonlinear dynamical elements, such as nonlinear reactances (e.g. nonlinear springs) or, more generally, linear circuits containing a *lumped* nonlinearity. The *memory* of the nonlinear element is, in fact, incorporated in the dynamic adaptor or in the mutator that the nonlinearity is connected to. As a consequence, our adaptors cannot be memoryless, as they are characterized by reflection filters instead of reflection coefficients.

With this more general definition of the digital waves, we can define the adaptation conditions for any linear bipole by selecting the reference transfer function in such a way as to eliminate the instantaneous input/output connection in its WD implementation (instantaneous adaptation). An "adapted" bipole will thus be modeled in the WD domain as $B(z) = z^{-1}K(z)A(z)$, where the delayed reflection filter $K(z)$ can also be identically zero.

The interconnection between WD elements is implemented through a network of elementary (series or parallel) dynamic adaptors, as shown in Fig. 2. These adaptors take care of the necessary transformation (with memory) between variables, as each wave pair is referred to a different RTF. This network of elementary adaptors constitutes a dynamic macro-adaptor that can be proven to be non-energetic [7]. This is an important feature of such elements as it allows us to guarantee that the passivity properties of the individual elements of the reference analog circuit be preserved by their WD counterpart. In fact, we have already verified that parallel and series multiport junctions are intrinsically nonenergetic provided that the port RTF's be stable. A computable interconnection through nonenergetic junction of elements having the same passivity properties as the reference ones will preserve the stability properties of the reference analog circuit.
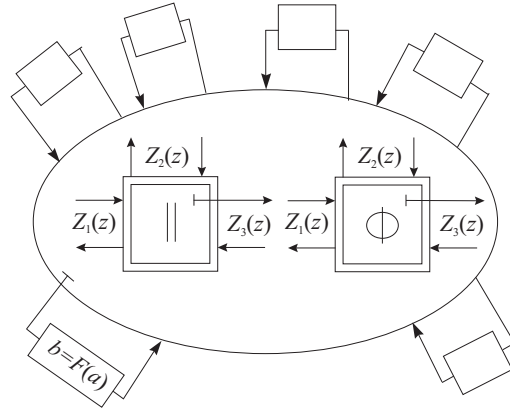

Fig. 2. Macro-adaptors in extended WDF structures are obtained by arbitrarily interconnecting together a number of dynamic adaptors. Such macro-adaptors model the local topology of "istantaneously decoupled" subsystems.

3. OBJECT INTERACTION

Let us consider an object that could potentially interact with a number of other objects in a sound environment. For example, we could think of a mallet that could collide, at different times, with a number of drum-like resonators. Indeed, this situation cannot be implemented with a fixed interaction topology such as the one of Fig. 4. In order to make this dynamic topology possible, we need to be able to connect or disconnect objects while the system is running. This can be achieved by exploiting the fact that a connection between systems is *irrelevant* when their contact condition is not satisfied.

As a simple example, let us consider the case of hammer-string interaction in the piano mechanism. The WD structure that corresponds to the equivalent circuit of Fig. 1 is shown in Fig. 3, where the macro-block M corresponds to the contact point between hammer and string. The nonlinear element (NLE) connected to the $R - C$ mutator [6, 5, 7]

(the double-boxed two-port junction of Fig. 1, whose aim is to "transform" the nonlinear capacitor into a nonlinear resistor) corresponds to the nonlinear spring that models the felt deformation and, at the same time, the contact condition. It can be easily shown that, when the contact condition is not satisfied, the series adaptor that connects the hammer to the two portions of the string becomes "transparent" for the two portions of waveguides that model the string. This fact suggests us that removing the whole connection by replacing that series adaptor with a direct connection between the two waveguide portions would not modify the behavior of the resonator.
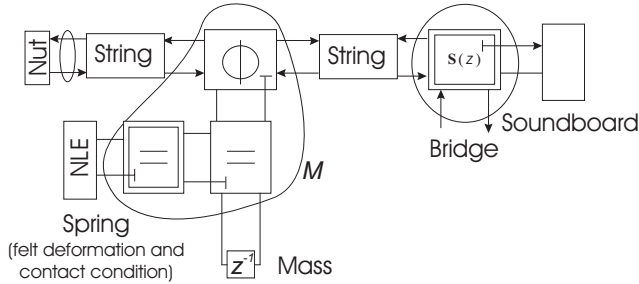

Fig. 3. WD structure for the modeling of piano sounds with fixed interaction topology. The contact condition is incorporated in the nonlinear element that is connected to the macro-adaptor M. When the contact condition is not satisfied, the macro-adaptor M becomes irrelevant and the string keeps evolving as if the macro-adaptor was not there.

The above reasoning can be extended to more complex resonators and has a significant impact onto our implementation scheme. In fact, there are two important direct consequences that are worth mentioning:

- systems that are not "close" to contact can be disconnected and may evolve independently;

- if the topology of the DWG network that implements the resonator is fixed, then a measure of "proximity" can be used for deciding whether and where to insert a transparent junction on the delay lines, in order to "preset" the contact point.

In general, while for a bipole the condition of adaptation corresponds to the possibility of "extracting" a delay element from it, for a multi-port element this is no longer true. In fact, the port adaptation only implies that no *local* instantaneous reflections can occur, while nothing can be said about instantaneous reflections through outer paths. If it is true that a delay can actually be extracted from a port, then we talk about *instantaneous decoupling*, which is a stronger condition than adaptation. The concept of *instantaneous decoupling* is important as it allows us to split the synthesis and the initialization of large WD structures into

that of smaller substructures [5, 4]. If N portions of a WD structure that are connected together through a decoupling N-port block ($N \geq 2$), which is a multi-port element that exhibits at least $N - 1$ decoupling ports, then such portions are said to be instantaneously decoupled, as they do not instantaneously interact with each other. One other reason why this decoupling condition is important is that it allows us to model WD structures that contain more than one nonlinearity. We know, in fact, that only one of all the ports of a macro-adaptor (oval block of Fig. 2) can be adapted, therefore only one nonlinearity can be connected to it. Through a decoupling N-port block, however, we can connect together N macro-adaptors, each of which is allowed one nonlinear element.

Decoupling multi-port blocks are quite frequently encountered in musical acoustics, especially when using DWG to implement reverberating structures. An example of block-based sound synthesis structure where the decoupling condition allows us to model a large number of nonlinear elements is that of the acoustic piano. In this case, in fact, a number of wave digital hammers are connected, each through a DWG model of a string, to the same (decoupling) resonating structure (soundboard).

In conclusion, the global structure of a WD implementation of a physical model can be seen as a number of decoupled interconnection blocks such as those of Fig. 4, whose aim is to connect together either linear macro-blocks or instantaneous nonlinear blocks. The presence of decoupling ports, allows us to approach the synthesis/initialization problem in a block-wise fashion. For example, if an interconnection block is connected to a set of adapted macro-blocks of the form $B(z) = z^{-1}K(z)A(z)$, then we can separate the synthesis/initialization of the macro-blocks of the form $K(z)$ from that of the interconnection block [5, 4]. A similar reasoning holds for two decoupled portions of the global WD structure.

The contact conditions allow us to *unplug* and isolate subsystems, while decoupling blocks allow us to approach the synthesis and the initialization of WD structures in a block-wise fashion.

Macro-adaptors – An N-port macro-adaptor can be automatically built through a tableau-based approach, specifically designed for WD structures [5, 4]. Its description, in fact, can be given in the form $\mathbf{S}(z)\mathbf{C}(z) = \mathbf{0}^T$, where $\mathbf{S}(z)$ is a $2N \times N$ Tableau matrix, $\mathbf{0}$ is a vector with N zero elements and $\mathbf{C}(z) = [A_1 \cdots A_N B_1 \cdots B_N]^T$ is the vector of digital waves. A generic macro-adaptor can be thought of as a network of elementary (parallel or series) three-port adaptors with memory that belong to a pre-defined collection. This allows us to construct $\mathbf{S}(z)$ by "pasting" a number of pre-defined 6×3 matrices into a larger sparse matrix. This matrix equation can be quite easily rearranged and inverted in order to obtain a state-update equation, or else it can be
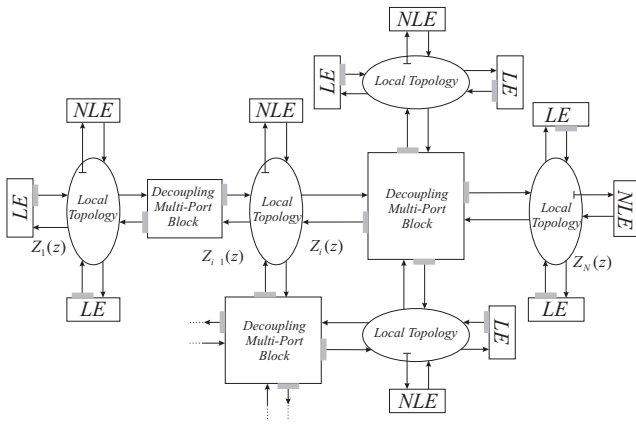
Fig. 4. Structure of a nonlinear block-based WD system with fixed interaction topology. The gray boxes at the ports of decoupling multi-port block denote the presence of a delay element, which guarantees that neither instantaneous local reflections nor instantaneous reflections through outer loops will occur.

solved iteratively using some efficient numerical method for sparse matrix equations.

As our macro-adaptors are not memoryless, they need to be properly initialized, which is a critical operation for WD models of mechanical systems as it usually affects the mutual position and contact conditions of mechanical elements. The determination of the state update equation can be seen as a direct form of the synthesis problem, as output signals are computed from input signals and memory content. Initialization, on the other hand, can be seen as an inverse problem, as memory content must be derived from output and input signals. As the nonlinearity is "lumped", this operation can be quite easily performed through nonlinearity inversion and matrix inversion.

Time-varying structures – Changing any model parameters in a WD structure usually affects all the other parameters as they are bound to satisfy global adaptation conditions. Temporal variations of the nonlinearities are easily implemented by employing special WD two-port elements that are able to perform a variety of transformations on the nonlinear characteristics (non-homogeneous scaling, rotation, etc.). Temporal variations of RTFs, on the other hand, are implemented through a global re-computation of all model parameters on the behalf of a process that works in parallel with the simulator [5, 4]. This operation requires the re-mapping of the nonlinearities as well. This parameter update, however, is not computationally intensive as it is performed at a rate that is normally only a fraction of the signal rate (e.g. 100 times slower). It is important to remember, however, that abrupt parameter changes must be carefully dealt with in order not to affect the global energy in an uncontrollable fashion.

Automatic synthesis – Some methods are already available for synthesizing linear macro-blocks, therefore the automatic synthesis procedure is based on the assumption that such elements are already available in the form of a collection of pre-synthesized structures. In its current state, the system that we developed is able to automatically compile the source code that implements a WD structure based on standard WDF adaptors and new dynamic adaptors chosen from a reasonably wide collection [5, 4]. The information that the system starts from is a semantic description of the network of interactions between all such elements.

Currently, the family of blocks includes WD mutators [6] and other types of adaptors developed for modeling typical nonlinear elements of the classical nonlinear circuit theory (both resistive and reactive). The available linear macro-blocks belong to the family of the DWG's [8], while the nonlinear maps are currently point-wise described in the Kirchhoff domain and then automatically converted in a piecewise linear WD map. Typical lumped WDF blocks are masses, springs, dampers, nonlinearities, ideal generators and filters (especially allpass filters, for the fine tuning of strings or acoustic tubes, or to account for the dispersive propagation in some enharmonic elastic structures such as bells, low piano strings, etc.). Typical distributed-parameter blocks are simple DWG implementation of strings and acoustic tubes, generalized DWG that account for rigidity and losses in a distributed fashion, reverberators based on Toeplitz matrices, green functions, DWG models of 2D and 3D structures such as membranes and bells.

The parameters can be modified "on the fly" in order to make the structure time-varying. A parallel process deals with the problem of re-computation of all WD parameters, depending on their changes expressed in the Kirchhoff domain.

4. AN EXAMPLE OF APPLICATION

Our approach to object-based sound synthesis has been tested on a variety of applications of musical acoustics. Starting from an appropriate semantic descriptions of the building blocks and their topology of interconnection, we used our authoring tool to automatically generate C++ source code for the implementation of a number of typical acoustic musical instruments. The timbral classes implemented with this method are hammered strings (piano, electric piano), plucked strings (guitar), bowed strings (violin), reed instruments (clarinet, oboe), jet-flow acoustic tubes (flute, organ pipes), percussions, etc.

One of these examples, namely the grand piano, has been developed with a twofold goal in mind: to test our solution to the problem of the mechanical modeling of a non-trivial acoustic instrument; and to test our approach to the construction of a dynamic topology of interconnection.

The basic mechanism of hammer-string interaction is shown in Fig. 1, which corresponds to the block-based WD model of Fig. 3. As we can see in Fig. 5, the trajectories of the hammer and of the string at contact point and the temporal evolution of the force that the hammer exerts on the string are very "physical" and realistic. In fact, the hammer tends to bounce back a bit more every time a wave is reflected by the nut or the bridge and returns at the contact point, causing the ripples in the force's profile. This behavior turns out to have a very realistic impact on the resulting sound. The plotted output corresponds to the acoustic signal at the bridge.

The global implementation of the piano model has been entirely built using a rather extended network of WDF and DWG elements. The DWG model of each string includes stiffness and losses. The bridge is modeled as a bandpass filter (the WD-equivalent of an RLC filter) and is connected to a soundboard model based on a DWG network. The string's fine tuning is performed using high-order all-pass filters. A limited number of hammers are used dynamically to hit a full-scale resonator such as the one described above, with a dynamical management of the contact conditions.

Indeed, the computational complexity of the resulting algorithm in this case coincides with the complexity of the resonating structure, whose role in the characterization of timbres is predominant. However, some simpler implementations already run real-time on low-cost PC platforms. For example, the WD model of an electro-mechanical piano (e.g. Wurlitzer or Fender-Rhodes) can easily run with full polyphony (61 or 73 keys) on a Pentium III (350MHz).

5. CONCLUSIONS

The proposed approach has proven effective for the automatic and modular synthesis of a wide class of physical structures encountered in musical acoustics. In fact, the wave tableau approach we implemented makes the construction and the implementation of the interaction topology systematic. In its current state, the implementation of the described synthesis system is able to assemble the synthesis structure from a syntactic description of its objects and their interaction topology, providing the user with a first CAD approach to the construction of an interactive sound environment.

6. REFERENCES

[1] G. Borin, G. De Poli, A. Sarti: "Sound Synthesis by Dynamic Systems Interaction", in *Readings in Computer-Generated Music*, edited by Denis Baggi, IEEE Computer Society Press, p. 139-160, 1992.

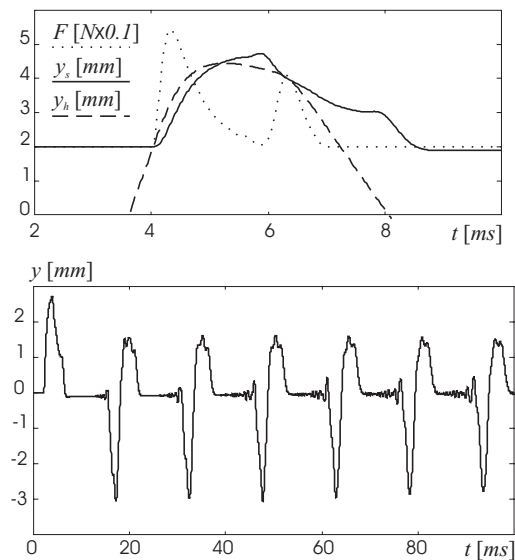[2] G. Borin, G. De Poli, A. Sarti: "Musical Signal Synthesis", in *Musical signal processing*, C. Roads, S. Pope,

Fig. 5. Model simulation of an acoustic piano through a WD model. F is the force that the hammer applies to the string, y_h is the position of the hammer and y_s is the position of the string at the contact point. Finally, y is the position of the string at the soundbridge.

A. Piccialli, G. De Poli editors, Swets and Zeitlinger Pub., Lisse NL, 1996.

[3] A. Fettweis: "Wave digital filters: theory and practice". *Proceedings of the IEEE*, Vol. 74, No. 2, pp. 327–270, Feb. 1986.

[4] F. Pedersini, A. Sarti, S. Tubaro: "Block-wise Physical Model Synthesis for Musical Acoustics". *IEE Electronic Letters*, Vol. 35, No. 17, Aug. 1999, p. 1418-19.

[5] F. Pedersini, A. Sarti, S. Tubaro, R. Zattoni: "Toward the automatic synthesis of nonlinear wave digital models for musical acoustics". *IX European Signal Processing Conference*, September 8-11, 1998, Rhodes, Greece. Vol. IV, pp. 2361-2364.

[6] A. Sarti, G. De Poli: "Generalized Adaptors with Memory for Nonlinear Wave Digital Structures". *VIII European Signal Processing Conference*, 1996, Trieste, Italy, Vol. 3, pp. 1773-6.

[7] A. Sarti, G. De Poli: "Toward Nonlinear Wave Digital Filters". *IEEE Transactions on Signal Processing*. Vol. 47, No. 6, June 1999.

[8] J.O. Smith, "Principles of digital waveguide models of musical instruments", in *Applications of digital signal processing to audio and acoustics*, edited by M. Kahrs and K. Brandenburg, Kluwer, 1998, pp. 417-466.