Analyse et synthèse de sons de piano par modèles physiques et de signaux

Julien Bensa

▶ **To cite this version:**

Julien Bensa. Analyse et synthèse de sons de piano par modèles physiques et de signaux. Acoustics [physics.class-ph]. Université de la Méditerranée - Aix-Marseille II, 2003. English. NNT : . tel-00003588

HAL Id: tel-00003588

https://theses.hal.science/tel-00003588

Submitted on 17 Oct 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE L'UNIVERSITE DE LA MEDITERRANEE

AIX-MARSEILLE II

FACULTE DES SCIENCES DE LUMINY

Spécialité : ACOUSTIQUE, TRAITEMENT DE SIGNAL, INFORMATIQUE APPLIQUES A LA MUSIQUE

présentée par

BENSA Julien

pour obtenir le grade de DOCTEUR D'UNIVERSITE

Analyse et synthèse de sons de piano par modèles physiques et de signaux

---

Thèse soutenue le 23 mai 2003

devant le jury composé de :

Pr.	Antoine CHAIGNE	rapporteur
Pr.	Julius O. SMITH III	rapporteur
Pr.	Bruno COCHELIN	examinateur
Dr.	Olivier GIPOULOUX	examinateur
Dr.	Richard KRONLAND-MARTINET	directeur de thèse
M.	Patrick BOUSSARD	membre invité

Exemples sonores disponibles à l'adresse http://www.lma.cnrs-mrs.fr/~bensa

Remerciements

Je voudrais tout d'abord remercier mon directeur de thèse, Richard Kronland-Martinet, avec lequel j'ai eu beaucoup de plaisir à travailler. En plus de me soutenir avec une grande gentillesse et ouverture d'esprit pendant ces trois années, de m'offrir la possibilité de rencontrer un grand nombre de chercheurs — y compris à l'étranger —, il a su me communiquer son énergie et son enthousiasme pour la recherche.

Je remercie également Antoine Chaigne et Julius Smith, pour avoir accepté de rapporter mon travail ainsi que les autres membres de mon jury, Patrick Boussard, Bruno Cochelin et Olivier Gipouloux.

Merci à tous les membres de l'équipe S2M pour sa chaleureuse ambiance de travail et parmi eux Mitsuko, pour m'avoir accueilli dans son bureau après mon retour à Marseille, Thierry pour ses conseils avisés, sa disponibilité et ses talents de programmeur temps réel, Solvi pour sa bonne humeur permanente et ses glegs écoeurants, Olivier pour ses prouesses linuxiennes, Philippe pour ses remarques pertinentes, Christophe, Fred, Greg, Irina, Jean, Marie-France, Snorre, Vincent.

Un grand merci à Anaïs pour m'avoir toujours soutenu pendant mes années de thèse ainsi qu'à Guillaume, Joe, Fred, Ol et autres skyhawks pour toutes les bonnes tranches. Merci à Tamara pour m'avoir prêté sa voiture rouge et pour avoir traduit une bonne partie de cette thèse et à tous les anglo-francophones qui m'ont aidé, Stefan, Stefania et Carr. Merci à mes parents et mes soeurs pour leur soutien, leur patience, leur gentillesse, et pour ne m'avoir jamais trop questionné sur la suite des évènements. Merci à Kristoffer, Attilio ainsi qu'à mes oncles et tantes, les proven caux particulierement, qui m'ont logé, nourri, blanchi quand j'étais SDF. Merci à tous ceux qui m'ont demandé si j'étais " enfin passé à la seconde corde".

Contents

List of Figures

Abstract

Nous avons conçu un modèle de synthèse de sons de piano qui utilise à la fois des méthodes de type modèle physique et des méthodes de type modèle de signal. Il reproduit fidèlement le timbre de l'instrument tout en simulant les principaux phénomènes physiques qui apparaissent en condition de jeu.

Nous avons réalisé un ensemble de mesures sur une expérimentation et sur un véritable piano. Pour chaque note du piano et plusieurs dynamiques, nous avons mesuré la vitesse du marteau et l'accélération du chevalet. Divers modèles de synthèse de vibration de cordes de piano et d'interaction marteau-corde ont été comparés afin de reproduire les signaux mesurés. Tout d'abord, nous avons étudié les modèles physiques sous la forme d'équations aux dérivées partielles (EDP). La synthèse sonore réalisée avec de tels modèles manque de réalisme. Pour améliorer cette synthèse, nous avons utilisé des techniques d'optimisation, calculant les paramètres du modèle afin que la synthèse soit aussi proche que possible d'un point de vue perceptif du signal mesuré. Une amélioration nette du son produit a été obtenue, sans qu'il soit totalement satisfaisant.

Nous avons ensuite étudié les modèles de type "guides d'onde numériques". Ces modèles simulent avec une grande précision la propagation des ondes dans une structure et sont peu coûteux en tant de calcul. Nous avons montré que les paramètres de ce type de modèle s'expriment en fonction des paramètres d'un modèle physique EDP. En couplant ces guides d'onde, nous sommes parvenus à modéliser les transferts d'énergie entre les cordes, responsables de phénomènes importants perceptivement comme les battements et la double décroissance des modes. Nous avons de plus montré qu'il est possible de traiter le problème inverse, c'est à dire d'obtenir les relations analytiques permettant de déduire les paramètres du modèle à partir des paramètres modaux des signaux mesurés. Cette calibration a permis de resynthétiser avec une grande fidélité les signaux originaux. La source du modèle guide d'onde, dont le comportement est directement lié à l'interaction marteau-corde, a été simulée par un modèle de signal, la synthèse soustractive. Le modèle de synthèse a finalement été implémenté en temps réel. L'instrument réalisé constitue un piano numérique dont on peut modifier le timbre.

Chapter 1

Introduction

1.1 Motivations, aims and method

1.1.1 Piano synthesis

A digital piano that accurately emulates the sound and behaviour of its acoustic counterpart has both scientific and commercial merit. It has advantages because of the many conveniences it offers to the user: it can be more easily transported, it can be played quietly by adjusting a volume control or played in complete silence by wearing headphones, and it can be easily recorded via a direct audio input to the sound card of a computer.

Sampling, a synthesis technique that uses an on-board bank of pre-recorded audio samples is the most common method used to develop commercial digital pianos. There are two major problems with this technique. The memory available for sample storage is limited and this makes it impossible to sample the wide range of nuances available to each individual note on a full size, 88 key, acoustic piano. It therefore becomes necessary to use transposition and/or interpolation of existing samples to obtain the missing information. The second disadvantage of sampling occurs when playing the instrument. The timbre of a note on the piano depends on whether the note was played alone or whether it was played along with other notes. This is because of the interactions between certain physical elements of the piano, and in particular, the energy that is transfered between the strings whose keys are suppressed, causing a modification in their vibration. Since the notes on a sampled piano are recorded in isolation, the sampler cannot take into account this important phenomenon that occurs only at the time of play (one cannot know ahead of time how many and which notes will be played at any given time). A model is needed, therefore, to simulate any change in timbre caused by a change in the behaviour of the system.

The piano is a struck-string instrument. Its range covers 88 notes with the lowest being

27.5 Hz and the highest being 4186 Hz. The piano consists of a large number of components, each one having a role that is more or less important to the sound production mechanism. Each note can use either one, two, or three strings with physical properties that differ from one note to another, making it necessary therefore to model each string independently. The resulting timbre depends on the interaction between the hammer and the string, the coupling between the string, and the way in which the sound radiates from the soundboard.

The piano also has certain advantages for real-time interactive sound synthesis. Its keyboard interface is very well suited to an input device for controlling a synthesis model. A MIDI keyboard (reproducing the piano touch) is all that is required to obtain input control parameters—this perhaps being because of the unsustained nature of the piano. Unlike wind or bowed instruments, the piano does not require continuous input control by the player in order to change the sound over the duration of a note. Rather, the evolution of the piano's sound is based entirely on the player's attack and release of a key (providing the pedal is not held down).

There are many different types of sound synthesis models for the piano. Signal models are generally computationally efficient enough to run in real time and can be very accurate in reproducing the sound of an existing piano. However, these types of models fall short when it comes to incorporating the player into the control-instrument loop. Since they make no direct link between the player's actions and the physics of the instrument, important playing conditions have no effect on the produced sound.

Physical models on the other hand, have the benefit of being able to simulate the interaction between player and instrument—of course this comes at a computational cost (though this cost is becoming less of an issue as computers become increasingly powerful). The parameters of physical models are difficult to accurately estimate from measured signals, and often their sound quality is inferior.

The quality of the synthesis depends on how many details of the acoustic system are taken into account in the physical model. Though algorithms are becoming more efficient and computer computation ever more powerful, there still remains a delicate balance between quality of sound and algorithmic complexity that is yet to be fully achieved in a quality digital piano.

1.1.2 Objectives

In this research, we propose a model that both achieves certain quality of sound and is efficient enough computationally to run interactively in real time. Our hearing is the best judge of whether the former criteria are met, and since human perception and experience are highly subjective, the sound quality of a synthesized musical instrument is very difficult

to ascertain. How does one judge sound quality without a reference? How does the listener know if the synthesized sound is emulating a particular piano or a generic piano-like sound? And if the listener hears a quality piano sound, how does one determine whether it's really because the model is good and not because of that particular listener's experience listening to pianos? One way to eliminate some of the subjectivity involved in this task is to synthesize a particular piano and have that piano available to the listener as a reference. We have therefore, in the first part of this study, tried to identically reproduce the sound of a Yamaha Disklavier grand piano (belonging to our Laboratory). All measurements required by the piano model were made on this instrument. Our task therefore falls under the category of analysis/synthesis, one well suited to the LMA/S2M group which has solid experience in this domain.

The second objective involves the control of the synthesis model. Since the algorithms used in the model must be efficient enough to work in real-time, the sound quality of the model will ultimately depend on the computational power of the computer.

The control parameters of the model should be similar to those controlled by the pianist on a real piano, i.e., the beginning and the end of the sonic event, the note that is played and with what velocity. In order to give the user the chance to modify the timbre of the virtual piano, s/he is also given access to certain physical parameters such as the length and tension of the strings, their inharmonicity, the strength of the the hammer, etc. This model could therefore also be useful in designing and building new acoustic pianos. It can also allow the user to hear alternate ways of playing the real acoustic instrument.

Though the user will likely own only one digital piano, s/he will be able to play on many different instruments simply by selecting the appropriate parameters ahead of time. By selecting, modifying, or combining piano parameters, the user is effectively building his/her piano of choice. And when certain parameter values surpass scientific reality, the user produces a virtual piano that cannot be built in the real world and sounds are likely to be produced that unlike any acoustic piano, creating artistic opportunity for the musician willing to experiment.

1.1.3 Methods

The steps taken to complete this research are outlined by the structure of this document. We begin by classifying physical systems (as described in Section 1.2) from a perceptual point of view. We then describe existing synthesis methods and their implementation as applied to the piano in Section 1.3. As the primary aim of this study is not to exhaustively classify and test all possible synthesis methods but rather to develop one sole algorithmic model, those methods that do not meet our criteria are abandoned at this point. Though there are certainly multiple ways to achieve our goals, we chose the methods that were most

suited to our experience and that best met our criteria. We then applied these methods to the synthesis of the piano while also proposing new models and new ways to, in particular, improve parameter estimation procedures in order to achieve a resynthesis of the piano sound that is as close to perfect as possible.

We begin by modeling the strings (chapter 2) and then the interaction of the strings with the hammer (chapter 3). This presupposes that the structure of the model will be, as we have previously specified, directly related to the physics of the instrument. We then propose in the chapter 4 a model that is practically complete save certain aspects involving the sound board—a topic more thoroughly developed in the thesis of Mitsuko Aramaki [**?**]. Finally, in chapter 5, we show how the model is implemented in real-time by simplifying certain aspects of the original model.

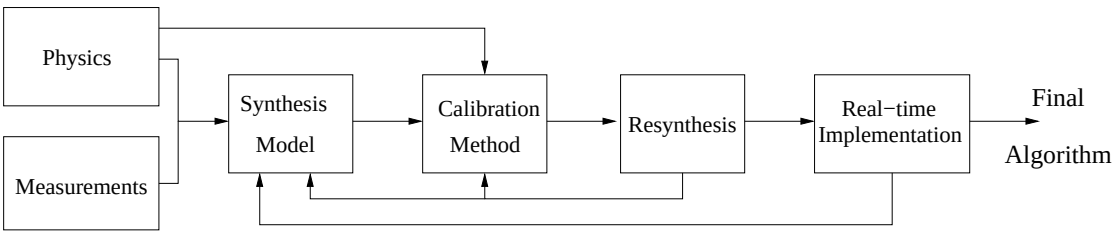

Figure 1.1: *Principle stages of the modeling.*

The principle stages of this procedure are outlined in figure 1.1.

- The synthesis model is developed based on the knowledge of physical systems and observations as made through laboratory measurements.

- The model's parameters are estimated using either signal measurements from a real piano or experimentation. Two methods were used to estimate the parameters. Whenever possible the model was inverted and the model's parameters were expressed as a function of the measured signal's parameter's such as amplitude, frequency and the rate of decay in the partials. When the model was impossible to invert, global optimization methods was used.

- Once the synthesis model was complete, we compared its produced sound with that of the original. If need be, we refined the calibration method or tweaked the model by ear until it reproduced the original signal to our satisfaction.

1.2 Acoustical properties of the piano

Before trying to model the instrument and doing its sound synthesis, we need to understand the main phenomena involved in the sound production. Many studies have described the behavior of the varied elements of the piano. The strings and the hammer have been the subject of specifics experimentations and publications [60] [117] [45]; the study of the radiation of the soundboard [47] [31] is more recent and has made a lot of progress the last years thanks to the appearance of accurate sensors (laser vibrometer, accelerometers, etc.). In this section, we mainly describe phenomena which have an acoustic function in the transformation of the gesture of the pianist in a perceptible sound. Some of the phenomena described have been taken into account in the design of the piano model and can be directly measured from experimentations. Other phenomena, which are less important from a perceptual point of view, have not been modelled since they make the synthesis models too complicated and thus incompatible with real-time simulations.

1.2.1 An intricate chain

The first piano was built by Bartolomeo Cristofori in 1709 in Florenze. Starting from a harpsichord, he substituted the plectra for a basic hammer (a wooden core with a thin leather). The oldest existing piano, made by its inventor in 1720, is on display at the Metropolitan Museum of Art in New York. Since its first appearance, the piano had to go to many changes to reach its modern form. The hammers, the strings, the frame have changed, the string tension has consequently increased. H.A. Conklin [30] has studied the difference between the Cristofori piano and a modern piano. He reports that if the principle is still the same, the influence of design factors has evolved and transformed this instrument, leading to a piano adapted for an immense concert room. In this study, we concentrate mainly on the grand piano. But most of the results given and models proposed are also valid for an upright piano.

Figure 1.2 shows a schematic view of a grand piano [5]. The string is suspended between the agraffe and the hitch pin which is attached to an iron frame. Between those two fixed supports, the string runs across the bridge. The bridge is a thin wooden bar transmitting the vibration of the string to the soundboard (a wooden plate which can be found under the frame). The distance between the agraffe and the bridge determines the vibrating part of the string. The end of the string which is closer to the keyboard is connected to the tuning pins on the bin block. The string is struck by a hammer, transforming the kinetic energy given by the artist to vibrational energy in the string.

The piano sound is actually constituted by two parts produced thanks to two different

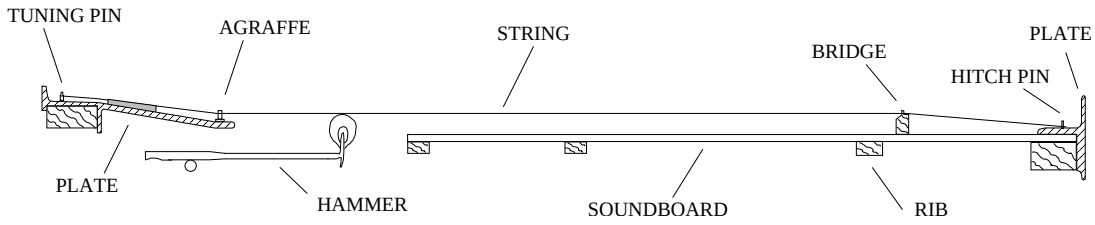

Figure 1.2: *Schematic view of a grand piano (from [5]).*

vibrating ways [21]. The first way produces a percussive sound, result of the impact of the key on the frame. The corresponding vibration precedes slightly the main sound and depends on the way the pianist strikes the key. The second way starts with the hammer, which strikes the strings. The interaction between the hammer and the strings leads to two complex transverse waves which propagates back and forth [8]. Each time these waves are reflected at the bridge level, a part of their vibrational energy is transfered to the soundbard and radiated. Besides the impact noise and the transverse waves, the piano sound is the result of many other vibrational phenomena. One can also hear slightly the longitudinal waves, and the sound due to the vibration of strings in the high register. Moreover, coupling phenomena between strings lead to important perceptual features like beating and two-stage decays [115] which characterize the piano sound.

1.2.2 The strings

The piano strings have a great function in the behavior of the instrument. They are too thin to radiate any sound but they have strong resonances. They select the energy received from the hammer and given to the soundboard around several frequencies. Therefore, the design of a string scale has to be done with a particular care.
The strings are made of steel wire. In order to obtain good elastic properties and efficiency, the string tension has to be high, around 700 N. But it is usual in designing piano scales to limit string pull to no more than 65% of the specified tensile strength of the wire [32]. There are 88 notes on the piano keyboard (from A0 to C8) but there are more than 88 strings: in the middle and the treble range, the hammer strikes two or three strings to reach higher acoustic output. In order to have an homogeneous tone quality, the length of the strings should change exactly in inverse proportion to the fundamental frequency. But if it were so, the lowest strings would be too long to build a realistic piano and the scale would be larger than acceptable. This is avoided by also changing the mass of the strings especially in the bass range.

Inharmonicity

The "harmonics" of the piano sound are not exact integral multiples of the fundamental frequency. The whole spectrum is stretched and its components are called partials (see figure 1.3). The frequency difference between partials is higher than the fundamental frequency and typically, the 10th partial could have the frequency of the 11th if there were no inharmonicity. We could think that this inharmonicity is an artefact that has to be minimized. But it also contributes to the piano tone and is mainly responsible for its specificity [32]. Fletcher, Blackham, and Stratton [38] pointed out that a slightly inharmonic spectrum added certain "warmth" into the sound. Simulations have shown that a piano without inharmonicity would sound unnatural [85]. Inharmonicity has then to be absolutely taken into account in our piano model.

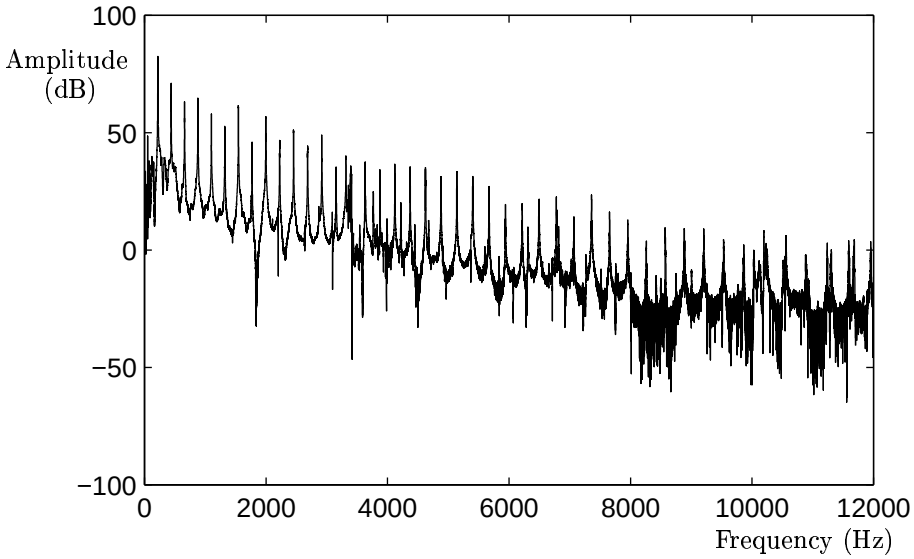

Figure 1.3: *Spectrum of the note A3 of a grand piano.*

The stretched tuning of piano strings can be almost entirely attributed to their inherent stiffness [117] [84] which produces a restoring force in the string (the finite bridge impedance can also slightly shift the mode frequencies of the string). This leads to a dispersion of the waves during the propagation. The phase velocity of the high frequency waves is greater than the one of the low frequency waves. A small "precursor", corresponding to the highest frequencies which reach the bridge before the low frequencies, can then be measured at the beginning of the sound [112]. Consider the wave equation of a stiff and lossless string. It is a combination of the equation for ideal string and ideal bar:

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2} - K^2 \frac{\partial^4 y}{\partial t^4} \tag{1.1}$$

$$c = \sqrt{\frac{T}{\rho}} \qquad K = \sqrt{\frac{EI}{\rho}} \tag{1.2}$$

where y is the transversal string displacement, x the position along the string, t the time, c the wave speed, K the stiffness coefficient ρ the linear mass and T the tension. E is the young modulus of the string and I its moment of inertia. Solving Eq. (1.1) for small stiffness, the modal frequencies of the string will be [77, 38, 40]:

$$f_n = n f_0 \sqrt{1 + Bn^2} \quad \text{for} \quad n > 0 \tag{1.3}$$

with

$$f_0 = \frac{1}{2L}\sqrt{\frac{T}{\rho}} \qquad B = \frac{EI\pi^2}{TL^2} \qquad I = \frac{\pi d_0^4}{64} \tag{1.4}$$

where f_n is the modal frequency, f_0 the fundamental frequency, n the partial index, B the inharmonicity factor and d_0 the diameter of the string.

This inharmonicity factor is proportional to the diameter of the string (raised to the 4th power). The larger the diameter, the more the string behaves as a beam. Therefore, in order to increase the string mass in the bass range without increasing the inharmonicity, the strings are wrapped. Wrapped strings are made by winding a helical coil of covering wire, usually copper or copper-coated steel, onto a tensioned steel core wire [32]. Since this inharmonicity increases in inverse proportion to the length of the strings squared, grand pianos are less inharmonic than upright pianos. We have estimated the inharmonicity coefficient B thanks to formula 1.3 for each note by making measurements on a grand piano. Figure 1.4 shows the evolution of B over the piano range.

Except for the first notes, the inharmonicity increases with the range. This is in accordance with other measurements found in the literature [32]. But, when listening to piano sound, it seems to be more inharmonic in the lower register than in the higher one. One of the reasons is that the bass strings countain more partials than the treble ones, allowing an easier perception of the inharmonicity. Another reason is that the threshold of audibility is close to the actual inharmonicity of the piano for the treble range [65] [75]. Inharmonicity affects also piano tuning. Since the tune is made by using the beating between two different notes, the stretching of partials will cause a stretching of the fundamental frequencies as well. In a correctly tuned piano, the upper and treble fundamental tones are higher in frequency and the bass fundamental tones are lower in frequency than would be the case without inharmonicity. This is a cause of troubles when the piano is played with other instruments.

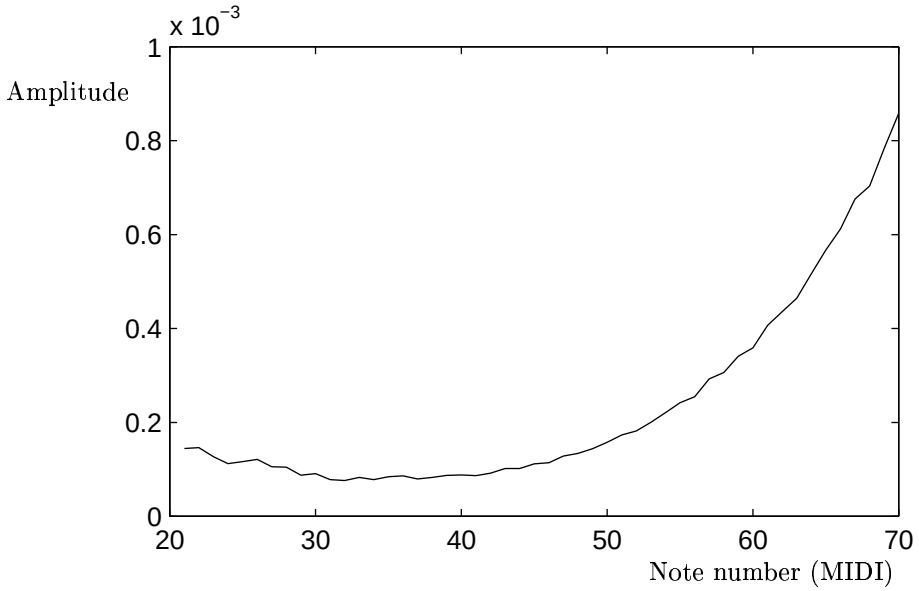

Figure 1.4: *Measured inharmonicity factor B.*

Losses

There are two different losses mechanism in the piano string: losses in the string itself, and those due to the vibration of its extremeties. The first can be divided into three types, according to the frequency [21]. Above 1 kHz, the greatest losses are due to viscous friction in the air and increase with the decreasing of the frequency. After 1 kHz, the visco-elastic losses in the steel are greater and increase with the frequency. The losses by heat conduction are weak but greater around 1 kHz. For more details, see [112]. The second loss mechanism allows the radiation and perception of the sound by transferring energy to the soundboard and to other strings (see next paragraph). It depends on the string tension and will be different from one note to another, according to its position on the soundboard [5].

Coupling phenomena

In the middle and the treble range of the piano, there are two or three strings for each note in order to increase the efficiency of the energy transmission towards the bridge. The vibration produced by this coupled system is not the superposition of the vibrations produced by each string. Having several strings for each note not only facilitates the energy transfer but also influences the behaviour of the sound. Considering a temporal signal of a piano sound (see Figure 1.5), we can notice beating and double decay on the amplitude.

Similarly, we can notice in the spectrum (Figure 1.6) that partials are actually constituted

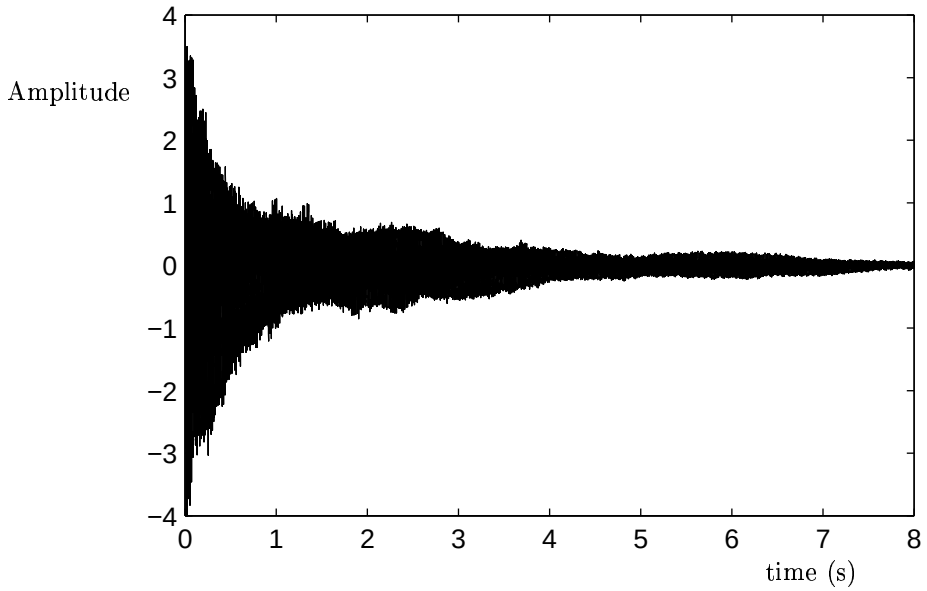

Figure 1.5: *Recorded waveform for a note C2 of a grand piano.*

by two close components. If we assume that the strings are not tuned to the same frequency but neglect the effect of coupling, the result will be beating between partials as well. But the strings are coupled to the bridge. There is energy transfer between the strings giving rise to new modal frequencies.

G. Weinreich [115] has studied in detail this phenomenon in the case of two strings coupled at the bridge level and vibrating in the vertical plan. But its conclusion can be also extended to the case of one string vibrating in two orthogonal planes (those two modes of vibration are also coupled at the bridge). He shows that frequencies of the two eigenmodes of vibration depend on the type of coupling (resistive, reactive) and of the frequency difference between the two systems without coupling. In the case of a purely reactive bridge admittance, the eigenfrequencies are purely real (there is no damping in the system) and always different (even if they were equal without coupling). This leads to beating in the sound. If the bridge impedance is purely resistive and if the frequency difference between the two uncoupled modes is not too large, the eigenfrequencies will have the same real part (there is no beating) but different imaginary parts and, consequently, there will be two different decay rates. If the admittance has both real and imaginary parts, as it is the case for a real bridge, there will be beating and double decay.

Practically, when the hammer strikes two strings, they will first vibrate in phase. The

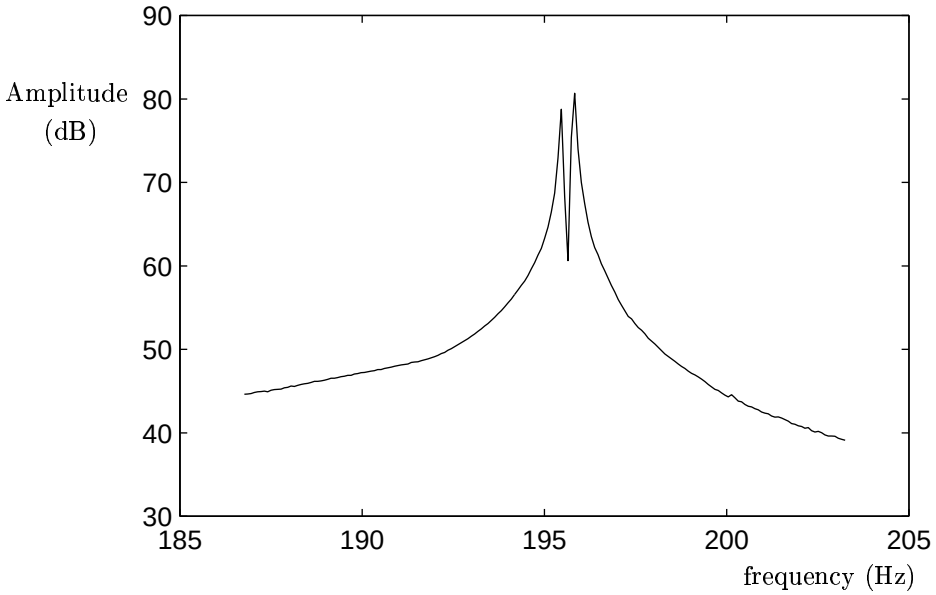

Figure 1.6: *One component of the spectrum of a note C2.*

vibrations are consequently large and the energy is quickly transmitted to the soundboard. This leads to the "immediate" sound which has the greatest decay rate. Then the two phases of each string will be different and the strings will exchange energy via the bridge. The losses are then weaker as well as the decay rate: this is the "remanent" sound. There is the same kind of phenomena in the case of one bi-polarized string. The hammer excites first the vertical mode, which is more efficiently coupled to the bridge than the horizontal mode. The energy transmission is more effective in this direction as well. The sound produced by the vertical polarization will be dominant and the decay will be greater. Then, the string will vibrate in the horizontal polarization, leading to a smaller transmission of its energy and to a smaller decay rate.

The coupling phenomena provides a characteristic dynamic to the piano sound. Beats are used by the professional tuner to precisely mistune the doublets or triplets of strings since the beat periodicity is correlated to the frequency difference between the strings modes. This way, they can adjust the duration and the tone of the sound.

Longitudinal modes

A dynamic consequence of the longitudinal elasticity of the string is the existence of longitudinal waves. According to Giordano and Korty [48], they are produced by a non linear

way, which is the stretching of the string due to the finite amplitude of the transverse displacements. Those longitudinal modes lead to perfectly audible sound in the piano case and are especially perceptible in the bass range. As Conklin claims, those "longitudinal string modes should be taken into account in the design of piano scales and wrapped strings in order to obtain the best tone" [32]. They have to be tuned to the transversal modes for a more pleasing tone. Longitudinal modes have a higher wave speed propagation than transverse modes. They create an harmonic series of fundamental frequency $\sqrt{E/\rho}/(2L)$ [112], much higher than those of transverse waves. According to Conklin, the frequency ratio between the first longitudinal and transversal modes is typically from 4200 to 5200 cents. Therefore, they are mainly present at the attack of the sound. The high frequency precursor vibrations we have discussed before has then to be attributed to the vibrations of the piano frame/key bed but also to longitudinal string vibrations.

Non linearities

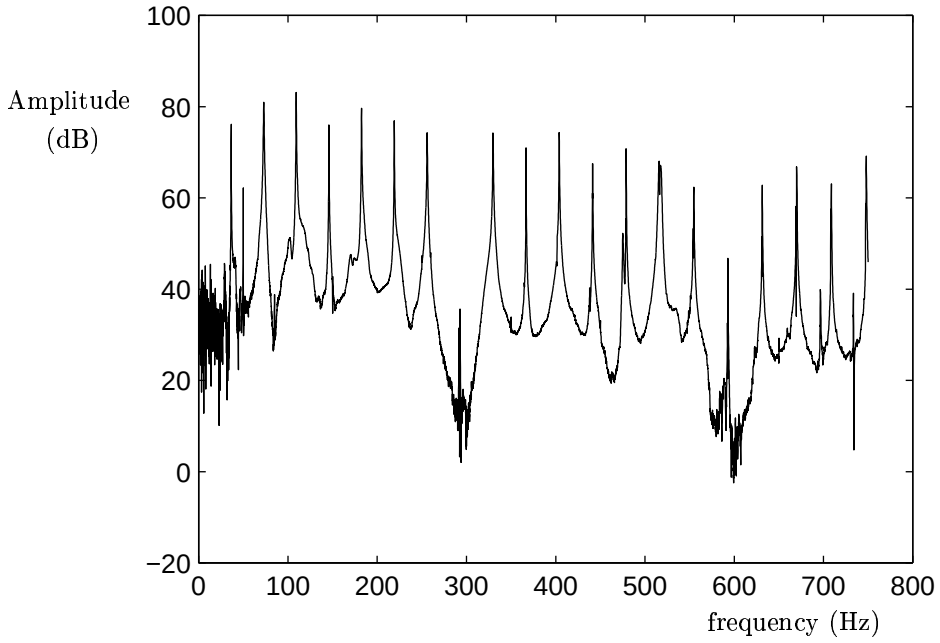

Figure 1.7: *Spectrum from a note D1 of a grand piano.*

Let's have a look to the spectrum of Figure 1.7. The corresponding signal was measured on an experimental setup described in Section 2.6.2. The string was struck with a hammer located at 1/8 of its length and observed using a laser vibrometer at a distance 1/8 from the bridge. The string has then been excited at a node of vibration which means that if

we only consider the linear phenomena, the corresponding partial (here the eighth partial) should not be visible, which is not the case. Consider now the Figure 1.8 showing a part of the spectrum of the measured signal and of the corresponding synthesized signal using the linear model described in Section 2.2.2. The original signal should contain only a series of inharmonic partials whose frequencies are given by the relation Eq. (1.3). We can see on the measured signal that there are many other partials (marked with a '*' on Figure 1.8) on the spectrum. Those phenomena are not foreseen by the linear theory. There are several nonlinear phenomena intervening in piano string vibration. Some of them have a perceptible role and have been studied in the literature. As we will see, the given explanations are sometimes contradictory, which means that it calls for further research.

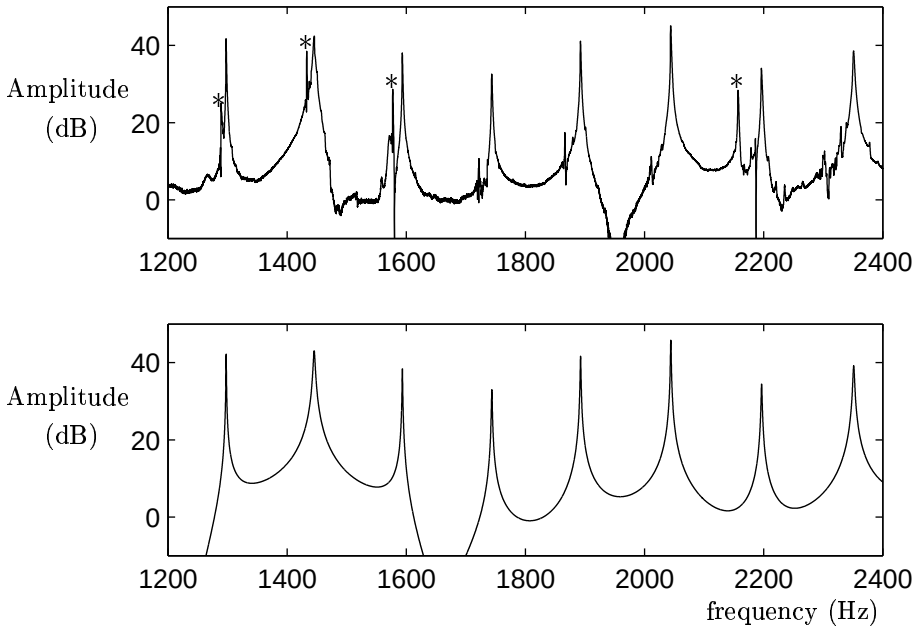

Figure 1.8: *Measured (top) and modelled (bottom) piano signal using a linear model. The '*' mark partials produced by non linear phenomena.*

One of these non linear phenomena and perhaps the most obvious is the tension variation in the string during the movement (see i.e. [110]). Indeed, due to transverse waves, the shape of the string changes during the motion and its length changes as well. The tension will then be modulated at a double frequency of the fundamental according to the amplitude of the motion. It is possible to measure a slight decrease of the frequency of the partials at the beginning of the sound [73]. In the piano case (contrary to harp or guitar) this phenomenon is not perceptible. It is too weak and probably masked by other phenomena. According to

Legge and Fletcher [73], tension variation is the starting point of mode coupling. Authors have investigated both theoretically and experimentally energy transfer among partials due to such non linear phenomena. For their study, they took the model of a string tightened between two supports. One rigid and another one which can have a mechanical finite impedance in a direction perpendicular or parallel to the string. They show that the energy transfer between modes is due to the tension modulation. In the case of completely rigid terminations, the tension modulation will only affect original modes. If the bridge has infinite impedance in a direction parallel to the string and a finite impedance in a perpendicular direction, modes of frequencies ω_n and ω_m can transfer energy to mode frequencies $2\omega_n \pm \omega_m$. If the bridge also has a finite impedance in the string direction, the non linear mechanism gives driving terms of angular frequencies $2\omega_n$ and $2\omega_m$.

In [58], Hall and Clark claim that there are at least two reasons showing that even if the hammer strikes the string at a node, the corresponding modes will not be completely missing. They show that, first, the small residual strengths in the nearly missing modes can be well accounted for by the finite bridge admittance and secondly the finite width of the region of contact between real piano hammers and their strings will avoid completely missing modes.

Conklin attributes in [33] the so called "phantom partials" observed in piano tones to nonlinear mixing. This phenomenon can occur because the tension varies during transverse vibration and produces longitudinal string forces at the bridge level. If the reason seems to be close to the one given by Legge and Fletcher, one can notice that Conklin does not make any assumption concerning the bridge impedance. By imposing a displacement to a string (tightened between two fixed supports) at the center, he shows experimentally that the force measured on the bridge can be expressed by $f_t = 0.0972d^2$ where f_t is the tension end force in Newtons and d is the center displacement in mm. When two frequencies ω_n and ω_m are present in the string, this force will generate components of frequencies $2\omega_n$, $2\omega_m$ and $\omega_n \pm \omega_m$. This non linear mixing could occur even if the bridge has an infinite impedance and then generates other partials than the original ones, which is in contradiction with what Legge and Fletcher claimed. The discussion is not closed yet and further investigation has to be done.

Finally, we would like to give another element which has to be taken into account in the discussion. For purposes of modeling, we have made a lot of experiments and measured the vibrating behavior of piano strings in several experimental conditions. We would like to emphasize the fact that what is called "phantom partial" in the literature can also be produced by coupling phenomena between the two polarizations of one string or, in the piano case, by coupling between several strings. We have experimentally noticed that those couplings

produce doublet (or more) components which can sometimes lead to totally independent components in the spectra. Moreover, the coupling changes the frequency of the regular inharmonic components. This may complicate the experimental detection of "real phantom partials" produced by non linear phenomena.

1.2.3 The hammer

The action mechanism

The action of the piano hammer is a complex mechanism which was first developed by Cristofori. Its action was remarkably ingenious and has been later improved by Erard in 1820, with the "double escapement". Minor improvements have been made since this date.

The piano action consists of 4 main parts: the key, the leverbody, the hammer and the damper. The motion of the key is transferred to the hammer via the leverbody. Due to the levering in the action (ratio 1:5), the hammer velocities are much higher than the key velocities. The force transmission from key to hammer is disconnected shortly before the hammer-string contact. This is called the "escapement". The hammer strikes then the string freely. After the hammer-string contact, the hammer is caught. The damper is lifted off the string by the key before hammer-string contact and let down when the key is released. The action is aimed at allowing the best control of the hammer velocity, especially in the low dynamic range, and the fastest possible repetition of a single note. In this way, the grand piano has a more precise and sensible action for the control of dynamic and repeated notes. When playing pianissimo, the impact of the hammer on the string is not strong enough to throw back the hammer. It is done thanks to the gravity for a grand piano and to a spring for an upright piano. But in the case of an upright piano, a security margin has to be taken, which means that the escapement velocity of the hammer has to be a little higher, leading to a weaker control in this dynamic range. Moreover, the double escapement, which usually equipped grand piano allows a sooner reload of the system leading to a faster repetition of the same note.

A. Askenfelt and E.V. Jansson have studied the timing of the action of a grand piano [6] and noticed that the delay introduced by the action depends on the dynamic level. The freetime for the hammer before the string contact became shorter, the key bottom contact was reached earlier compared to the string contact; the hammer-string interaction duration decreased. The pianist has then to compensate those differences. In a second article [7], A. Askenfelt and E.V. Jansson have tried to solve a contradicting standpoint between pianists and researchers. Pianists often claims that they can change the tone of the sound by using different "touch" at a same dynamic level, whereas researchers said that since the escapement

interrupt the transmission force before the hammer-string contact, the only control parameter is the hammer velocity at the escapement time. One first explanation is that the shock of the finger on the key and the shock of key on the keypad depends on the touch and are part of the piano sound. The key motion can be different according to the dynamics: at loud dynamics, strong oscillations superimposed on the key motion. The second explanation is that the hammer exhibits various resonances and their level differs according to the type of touch (legato, staccato...). The story of the hammer motion during acceleration is influenced by those modes and consequently by the hammer-string contact. Those two results show that depending on the touch, the piano action would offer a possibility for the pianist to influence the tone.

Hayashi et *al* have proposed a model of the hammer action [59]. All the mechanism is modeled as a spring-mass oscillator. Their model seems to perfectly reproduce the movement of the hammer before the escapement. Authors show that depending on the displacement of the finger on the key (constant speed, constant acceleration), modes of the hammer can be excited. Moreover, for a constant low speed, the hammer can break contact with the action parts before the escapement point and strikes the string. After it strikes, it is pushed up by the parts and it can strikes the string again. This phenomenon is referred to as "double-striking".

The interaction with the strings

The piano tone as many of musical instruments, depends on the dynamics. The greater the velocity of the hammer, the brighter the resulting sound. Figure 1.9 shows several piano spectra for different dynamics. It is obvious that there is a global increase of the amplitude but also a modification of the spectral shape. Figure 1.10 shows the behavior of the centroid of the spectrum modulus [13]. The centroid is correlated to the perception of brightness. It is defined as the sum of the amplitudes of each component times its frequency, over the sum of the amplitudes of the components. The centroid is clearly an increasing function of the hammer velocity which is in accordance with the perceptual statement of an increase in brightness. The audition is actually more sensible to the variation of the tone rather than the variation of the intensity and the perception of dynamic is then strongly correlated to the tone. This progressive increase of the high partials of the spectrum denotes a non linear behavior due to the interaction between the string and the hammer.

The hammers are constituted by a hardwood core with one or two layers of wool felt that increase in thickness from treble to bass [30]. In contemporary piano, the hammers have also hardwood cores of graduated size. Hammer mass typically decreases over the 88-note compass by a factor between two and three. Moreover, the ratio of hammer mass to string mass increases more widely. The hammer/string mass ratio is of great importance since it

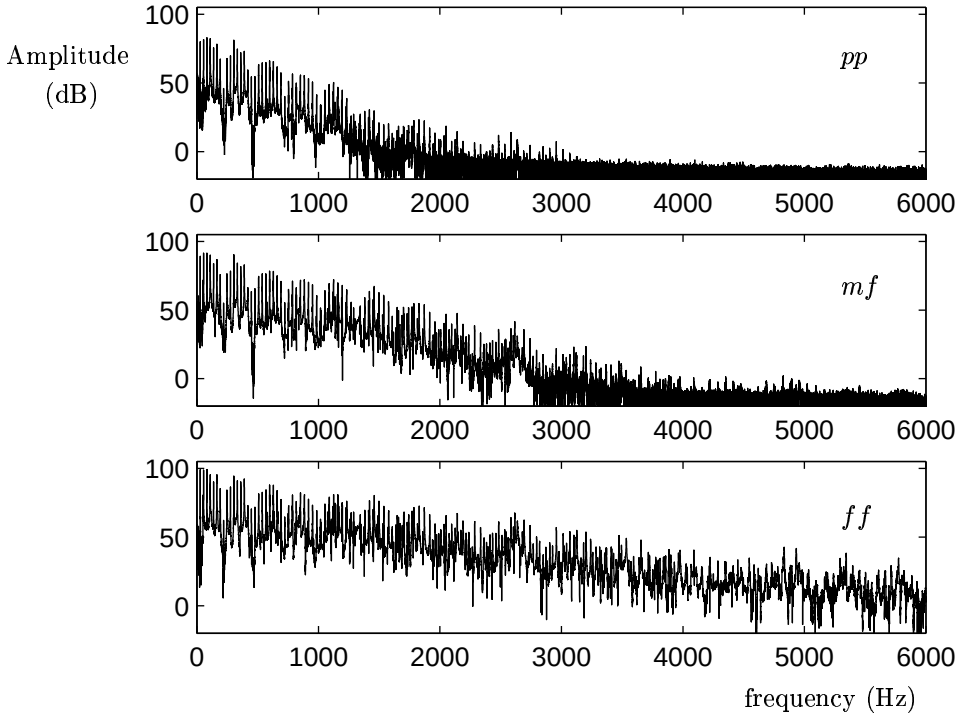

Figure 1.9: *Three spectra for three different dynamics (pp, mf, ff) from a note A0.*

will affect the contact duration between the hammer and the string and therefore influences the tone (see next paragraph). The hardness of piano hammer felt also directly affects the quality of the tone. Harder felt results in stronger high partials, i.e., a brighter tone. The felt of the hammer is not homogenous. Conklin believes that the existence of a hardness gradient is chiefly responsible for the spectral changes that occur with changes of dynamic level. The brightness of the tone can be adjusted by "voicing". Voicing is done after a piano has been completely assembled and tuned to produce a "good" tone, maintain an homogeneous timbre from one key to another and giving a "personality" to the instrument. The felt can be made softer by needling and harder by "filling" away with sandpaper some of the softer outer felt or by using a hardening agent.

Many studies have tried to describe and model the hammer/string interaction. One of the most complete study can be found in [8]. When the hammer strikes the string, it generates a pair of waves traveling outward from the striking point (see Figure 1.11).

In the piano, the string is struck close to one of its terminations, the agraffe. During the striking process, the string is temporarily divided in two parts by the hammer. The wave propagating toward the agraffe will be reflected shortly after the hammer impact and follow closely behind the original wave traveling toward the bridge. The returning wave will then

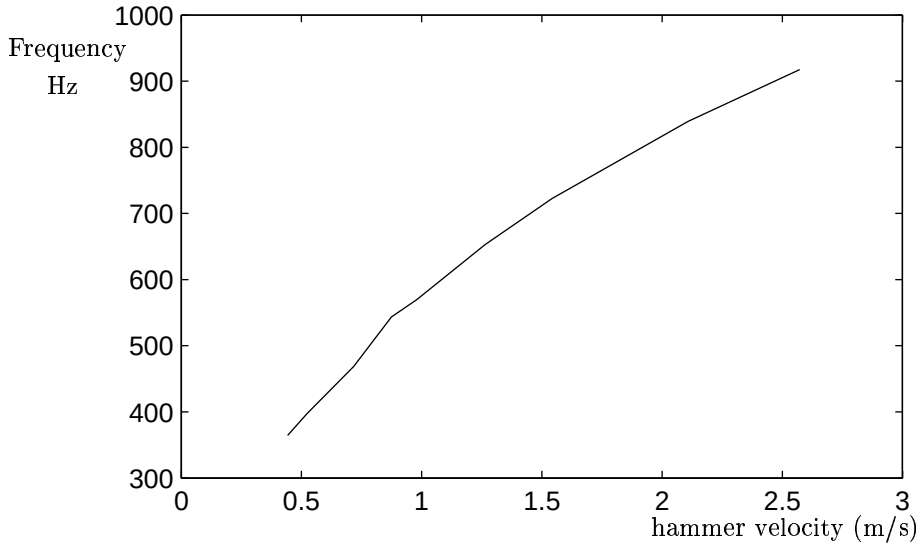

Figure 1.10: *Evolution of the centroid of the piano spectrum as a function of the hammer velocity for the note A0.*

meet the hammer, a non rigid support. There will be multiple reflections in the agraffe side of the string. The hammer will partly transmit and reflect the wave. Figure 1.12 shows the velocity of two points measured using a laser vibrometer, one on the agraffe side and the second on the bridge side. The reflected waves can be clearly seen in the first milliseconds of the signal measured on the agraffe side. Transmitted waves appears as well on the bridge side. When the hammer acceleration is too weak to cope with the gravity and the force exerted by the reflected waves, the hammer string contact ends. One can then understand why the hammer/string contact duration affects the tone. The spectra of the resulted sound depend also on the striking point. This results in a comb filtering effect since modes of the string which have a node at the striking point will be attenuated.

To model the hammer-string contact as a simple mass striking a string (modeled, for example, using Eq. (1.1)) allows to give an analytical solution of the problem, but is totally unrealistic. The next step consist in adding a linear spring to the mass to take into account the felt compliance. Hall has reviewed in [54] [55] [56] articles dealing with such a model. He proposed a solution, half analytical, half numerical, and shows that this approach is still too far from the real phenomena. To reproduce the modification of the tone according to the dynamic, one must take into account the non linear behavior of the felt. We have used such a model and we will then give more details in Section 3.2.

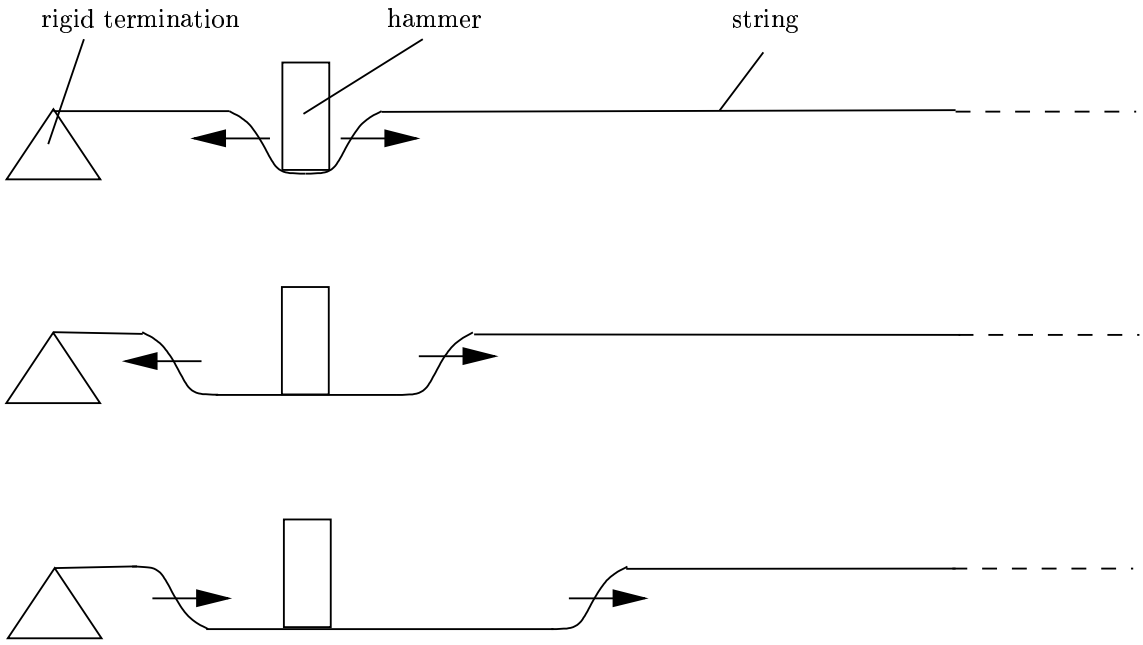

Figure 1.11: *Evolution of the string shape during the hammer/string contact.*

1.2.4 The soundboard

The soundboard radiates the vibration received from the string via the bridge, according to its dimension and geometry. It consists in a thin wooden (epicea) plate [31]. Ribs are glued behind the board with regular intervals in order to increase its rigidity in a direction perpendicular to the fibers of the wood. The soundboard allows the transform of the mechanical energy from the vibrating strings in to acoustical energy radiated in the surrounding medium. This energy transfer depends on the impedance ratio between the strings and the soundboard in the vertical direction. The strings are distributed over the entire length of the bridge and then deliver their vibration energy at different points on the soundboard, encounting different matching conditions. The soundboard has its own resonance modes. To be well radiated, a partial frequency of the string has first to be close to a mode resonance of the soundboard and secondly, has to be superior to the "critical frequency" (see [5]). Measurements from Wogram [5] have shown that the soundboard is incapable of radiating sound below 100 Hz. Giordano in [47] shows that the impedance of the soundboard decreases as a function of the note. Therefore, the damping is weaker in the bass range than in the treble. As we said previously, the soundboard has not been specifically studied and modelled here.

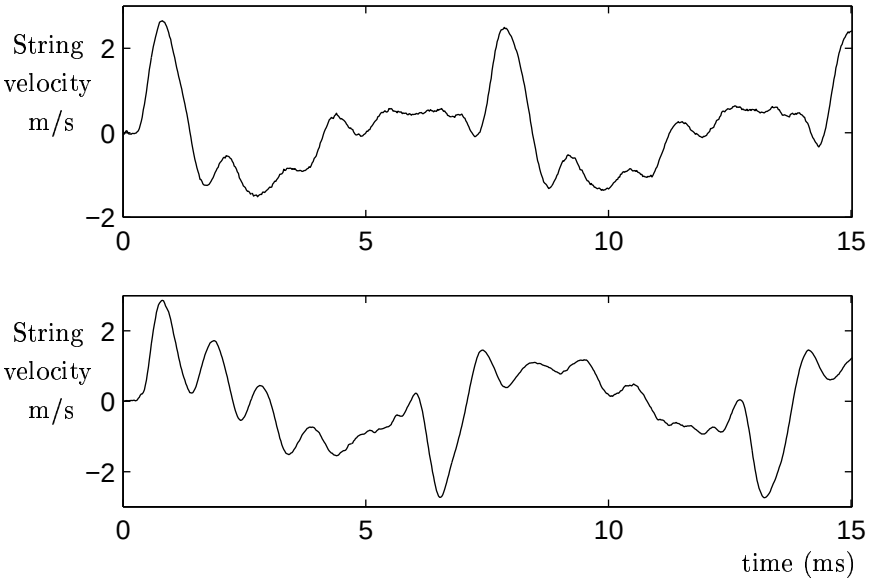

Figure 1.12: *String velocity at the bridge side (top) and at the agraffe side (bottom) measured using a laser vibrometer on the experimental setup described in 2.6.2. These correspond to the curves given in [8].*

1.2.5 The dampers and the pedals

The extinction of the sound after the released of the key is done thanks to the damper. The pianist can control the fall of the damper and consequently act on the extinction. There are few publications dealing with this phenomenon. Measurements and explanation can be found in [8, 6]. The damper has a limited mass and cannot terminate the vibration instantaneously, but the decay time can be made relatively short. Authors studied the effect of the dampers by simultaneous measurements of string and damper motion. The damping depends on the ratio string mass/damper mass. The damping is fairly inefficient in the bass. It takes approximately 250 ms for the string vibration to decay whereas this decay is 100 ms in the midrange and 50 ms in the treble.

Pianos have either two or three pedals: the sustain pedal, the una corda pedal (or in upright pianos, the soft pedal) and in some cases the susteno pedal. The sustain pedals lifts the dampers of all strings. The string can then vibrate after the release of the key. As all the strings can vibrate freely, the sympathetic coupling affects the resulting tone. The una corda pedal shifts the action mechanism so that the hammer will strikes only one or two strings. This causes a reduction of the sound pressure level (1 or 2 dB) but also changes the timbre.

The string which is not directly excited vibrates thanks to energy transfer via the bridge. In upright piano, there is no una corda pedal but a soft pedal which attenuates the sound thanks to a piece of felt lowered between the strings and the hammer. The susteno pedal (used only on grand piano) sustains the notes that have been struck before depressing the pedal.

1.3 Synthesis methods for piano tones: an overview

In this section we describe the main techniques used in sound synthesis for piano tones. It is not the purpose of this overview to exhaustively describe different existing methods, but to consider those which, among the most commonly used, could be functional for this study. Those methods have advantages and drawbacks, but some of them offer a better matching of the conditions we have described in Section 1.1.2: doing a synthesis as realistic as possible from a perceptual point of view while keeping a physical control of the parameters and allowing a real-time implementation. Details of these methods will not be given in this Section. Methods that have been used will be widely described in the different chapters of this document while the others are referenced for those who might be interested in a more detailed description. The classification given is based on the works of [68], [111] et [97]. We will distinguish two main class of synthesis algorithms: synthesis using signal models and synthesis using physical models. Signal models aim at reconstructing a perceptual effect without being concerned with the specific source that made the sound whereas physical models aim at simulating the behavior of existing or virtual sound sources.

1.3.1 Signal models

Signal models use a mathematical description of sounds. Their advantages are simplicity and their easy implementation. We will distinguish three types of signal models: the processed recordings, the global synthesis methods, and the linear methods.

Processed recordings

The sampling method is generally used in commercial digital pianos. It consists of playing a sound which has been previously recorded. This method is very accurate in reproducing a specific sound but is not able to reproduce the changes in playing conditions. Commercial digital pianos store separate tones of the instrument in a memory and play it back when a key is depressed [86] [111]. Several recordings are necessary for the same note to take into account the modification of the tone with the dynamics. This method requires a large amount of data, therefore only a few seconds are recorded. After the attack, the waveform

is reconstruct by looping, which means, continuously repeating the steady state part of the tone. Then, the amplitude and timbre evolution of the tone has to be simulated by a time-varying amplitude envelope and a time-varying filter [12]. In less expansive samplers, there might not be enough memory to store every tone. Typically, every third or fourth semitone are stored. Other tones are obtained by pitch shifting of the reference note, using simple time-domain methods that affect the length of the signal [86]. The drawback of this method is that the transposed noise of the attack sounds usually unrealistic. In many samplers, the memory requirements are tackled by data reduction techniques which degrades the perceived audio quality. Nevertheless, the highest quality synthesized piano sounds are coming from devices using sampling technique. The concept is straightforward and the implementation of the algorithm is relatively simple. But this method synthesizes the note separately, which means that physical phenomena, like energy transfer between the strings, restrike of the same string are not taken into account. This method has been successfully implemented for piano sound, since the musician can not act on the sound after the note was played. But it is very difficult to use such an algorithm for instruments like violin where the musician has more influence on the sound.

Multiple wavetable synthesis is close to sampling synthesis. The musical tones are produced using several wavetables, i.e. sound signals stored in a computer memory. The sound signals are multiplied with an amplitude envelope and then summed together to produce the synthetic sound [111]. In wavetable cross-fading, only portions of the signal are synthesized and then cross-faded to the next. This procedure is repeated over the course of the event. In wavetable stacking, every whole sound signal is shaped and they are then summed to produce the synthetic sound. It is in a way similar to additive synthesis (see further).

Global synthesis methods

Global synthesis methods aim at shaping a simple signal (a sinusoid for example) using a function. Those methods are nonlinear since the operations on the signal are not simple additions and amplifications. The synthesis uses relatively simple algorithm and exhibits a small number of parameters, but the analysis procedures are complicated. It is then generally difficult to control the shaping of a sound by those methods, since the timbre is related to the synthesis parameters in a non linear way.

The most well known example of global synthesis is audio frequency modulation (FM) originally introduced by J.Chowning in 1973 [29]. This synthesis method has been adapted

from the theory of FM for radio frequencies to audio synthesis purposes and revolutionized commercial synthesizers. Chowning observed that by connecting two sinusoidal oscillators, namely the carrier and the modulator, complex audio spectra can be easily achieved. The two oscillators are connected in such a way that the frequency of the carrier is modulated with the modulating wave form. If both the carrier and modulating waves are sinusoids, the amplitudes of the resulting harmonics can be calculated by the Bessel functions (see [29] for more details).

Another method, based on the nonlinear distortion of an input signal was developed by Le Brun [22]. We have tried this method to take into account the non-linear behavior of the hammer/string interaction and we will then give more details in section 3.8.4. An advantage is that the amplitude variation of the signal results in a large alteration of the output spectrum. If the input signal is sinusoidal, the harmonics ratios could be accurately controlled by using Chebyshev polynomials as distorsion functions. The signal can be then bandlimited and the aliasing can be avoided.

Linear models

Linear models like additive synthesis or substractive synthesis aim at constructing a signal using its time-frequency representation. In this way, they are using the process of human perception, which analyze audio signals according to their spectral contents. A wide range of sounds can be produced using those techniques. But generally, a large number of parameters is needed for the description of a given instrument.

Additive synthesis is one of the simplest and most intuitive of those spectral techniques. It consists in constructing a complex tone by summing elementary sounds, generally sinusoidal signals modulated in amplitude and frequency. It can be interpreted as a method to model the time-varying spectra of a tone by a set of discrete lines in the frequency domain [97]. Three control functions are needed for every sinusoidal oscillator: the amplitude, frequency and phase of each component. In many cases, the phase is left out and the output signal can then be represented as [111]:

$$y(n) = \sum_{k=0}^{M-1} A_k(n) sin[2\pi F_k(n)] \tag{1.5}$$

where y is the output signal, n the time index, M is the number of the sinusoidal oscillator, $A_k(n)$ is the time varying amplitude of the k^{th} partial and $F_k(n)$ is the frequency of the k^{th} partial. For periodic or quasi-periodic sounds, these components have frequencies that are multiples of one fundamental frequency. Analysis method can be applied to map

a natural sound into a series of sinusoidal oscillator. Amplitudes and frequencies of the components are mapped peaks in the frequency domain, and can be detected by using the *Short Time Fourier Transform* (STFT) of the original signal. We have synthesized piano tones by applying this method. The spectral components are well reproduced as well as their time varying behaviour as we can see on figure 1.13.

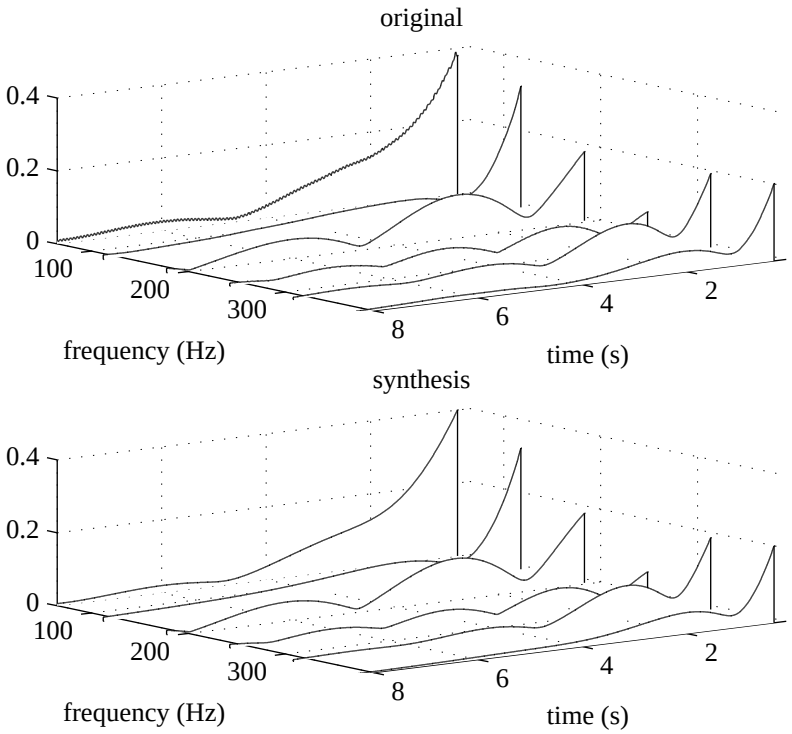

Figure 1.13: *Amplitude of the first 6 components of an original piano tone measured on a grand piano (top) and its resynthesis using additive method (bottom)*

.

The resynthesized sound is then very close to the original one. An advantage of this method is its flexibility and the potential for intimate and dynamic modifications of the sound. But its drawback is the huge number of control parameters. Group additive synthesis has been used to reduce this number of control parameters. Partials are group together and have common frequency and amplitude enveloppes. These grouped partials are then combined to form wavetables. According to K.Lee and A.Horner [72] who have used this method to model piano tones, it reduces the amount of data by about 50% compared to additive synthesis, while preserving the perceptual identity of the sound.

Granular synthesis is a set of techniques based on composing the synthetic sounds from short signal elements in the time domain called "sound atoms" or grains. Grains can have a duration ranging from one millisecond to more than a hundred milliseconds. The granular synthesis can be classified according to how the grains are obtained. In asynchronous granular synthesis (AGS) [86], grains are scattered over a region in the time-frequency domain called cloud. The grains can have similar waveforms or different waveforms. The waveform can be a windowed sinusoid, a sampled signal or obtained from a physics-based model [24]. In pitch synchronous granular synthesis (PSGS) [79], grains are derived from the *Short Time Fourier Transform* (STFT) of the original signal. The length of the rectangular windows used in STFT is the period of the final synthetic sound. Then each grain corresponds to a period of the signal.

The Spectral Modeling Synthesis (SMS) [91] is a method which aims at synthesizing a sound by adding separately the deterministic components and the stochastic components. The deterministic part is obtained from a given musical sound by first calculate the STFT of each windowed portion of the signal [111]. From the complex spectra produces by the STFT, the prominent peaks and their trajectories are obtained using a peak continuation algorithm. The stochastic components are then calculated by substracting the deterministic part from the original signal in the frequency domain.

In subtractive synthesis, the process is opposite to the one used in additive synthesis: the sound is constructed by removing undesired components from an initial, complex sound such as a noise. This method is then closely linked to the theory of digital filtering. This approach allows the separation between a source signal and a filter. This technique has been used especially to produce synthetic speech [76].

The source-resonator synthesis is in a way similar to substrative synthesis. A wide-band excitation signal is filtered using resonant filters. But this approach also corresponds to what happens in many physical system, assuming that there is no feedback from the resonator to the excitator. There is sometimes a confusion between the terms “source” and “excitation”. In this document, we will call “source” the concept and “excitation” the corresponding signal. The sound transformations related to these methods often use this property to make hybrid sounds or crossed synthesis of two different sounds by combining the excitation source of a sound and the resonant system of another [83]. A wellknown example of cross-synthesis is the sound of a talking cello obtained by associating an excitation of a cello string and a resonance system corresponding to the time-varying formants of the vocal track. The source-resonator technique was used by Laroche and Meiller [71] for the synthesis of piano sound. To

model percussive sounds using this approach, the authors made two assumptions: first, the freely vibrating body generates a sound that is composed of exponentially decaying sinusoids (which is the case for piano sounds, under the hypothesis of small perturbations and linear elasticity); secondly, the independence between the source and the values of frequencies and damping factors. From a physical point of view, the two parts of the model can be interprated as follows [71] : the strings, which correspond to the vibrating structures are represented by the resonant filter, and the hammer, which is the physical excitator is represented by a short duration excitation signal. This analogy is not totally exact since in the piano case, the hammer and the string interact during several milliseconds. The excitation which has to be injected in the resonator is then the result of this interaction. Laroche and Meiller show that realistic piano tones can be produced using this method. Moreover, they used one excitation to generate several sounds, typically five piano tones belonging to the same octave. We have also used this method to produce piano sound. More details are given in Sections 3.8 and 4.

1.3.2 Physical Models

Unlike signal models which aim at modeling the resulting signal, the main feature of physical modeling approaches is that they describe the sound generation system with respect to its physical behaviour. Such systems can be constructed either from knowledge of the physical laws that govern the vibrating parts of the system and expressing them mainly as mathematical equations, or from the behaviour of the solution of these equations [68]. There are two reasons for developing physics-based model. One is the understanding of the physical phenomena involved in sound production and the other is sound synthesis. But if the purpose is to model a real instrument and to reproduce its tone as close as possible, those two motivations are inexorably related and the accuracy of the synthesized sound will validate the design of the physical model. Physical models give a realistic response to the interactions of the musicians, since the input parameters are directly related to the physical features of the instrument and consequently to the gesture of the musicians. They are generally complicated and necessitate high computational costs.

Numerical Solving of a wave equation

The first method used in physical modeling was based on finite difference equations. The basic principle is to obtain the mathematical equation describing the vibrating motion and solve it in a finite set of points. The finite difference equation obtained simulates the propagation of waves in the system. Hiller and Ruiz [62, 87] proposed one of the first finite difference method for sound synthesis purposes. They applied this principle to plucked strings and proposed a PDE taking into account the frequency dependency loss and the stiffness of the string.

Chaigne and Askenfelt [26] [27] proposed one of the most realistic model of a struck piano string by combining the PDE of Hiller and Ruiz with a non-linear hammer model. They have then compared simulations and experiments. Waveforms were in good accordance with the signal measured on the real piano. We will describe their work in details in Section 2.2.1 and Section 3.2. A more complete piano model (taking into account the soundboard) using finite differences was presented by Hikichi and Osaka in [61]. Finite differences can be used to simulate the vibrational behaviour of many other structures such as bars, membranes, tubes, etc.

Modal synthesis and mass-spring networks

Modal synthesis describes sound-producing objects as a linear combination of vibrating substructures which are defined by modal data [1]. Substructures are then connected and can also respond to external excitations. Each substructures is modeled using its modal data, i.e. the frequencies and damping coefficients of the resonant modes and the shapes of each mode. Modes are supposed to be uncoupled, which means that there is no energy transfer. The modes are excited by an external force applied at a given point on the structure. Analysis techniques for the determination of the modal data have been developed. Modal data can be obtained analytically for simple structures, but for complex structures, an analysis based on measuring experiments is necessary. Modal synthesis method is general and can be applied to structure of arbitrary complexity. The modal synthesis has been mainly developed at IRCAM in Paris. A commercial software application called Modalys has been produced.

Another method was proposed by Cadoz et al. [23] consisting on using a set of masses, dampers and springs to model mechanical structures. The CORDIS system used this method and was the first system capable to produce sound based on a physical model in real time.

The Karplus-Strong algorithm and Digital Wave Guides

Karplus and Strong [67] have proposed a very efficient algorithm to simulate plucked string and drum timbres. This algorithm is an extension of the wavetable synthesis technique (see Section 1.3.1). As we have seen previously, wavetable synthesis consists in the periodic reading of data in a computer memory. In the Karplus-Strong algorithm (KS), the content of the wavetable evolves with time. The wavetable changes each time a sample is read. A block diagram of the simple wavetable algorithm and of the KS algorithm are shown on Figure 1.14 [111]. In the KS algorithm, the lowpass filter will account for the decay of the tone. The transfer function of the modifier filter is

$$H(z) = 1/2(1 + z^{-1}) \tag{1.6}$$

This filter can then be related to the losses of the physical system, allowing a physical interpretation of this signal model.

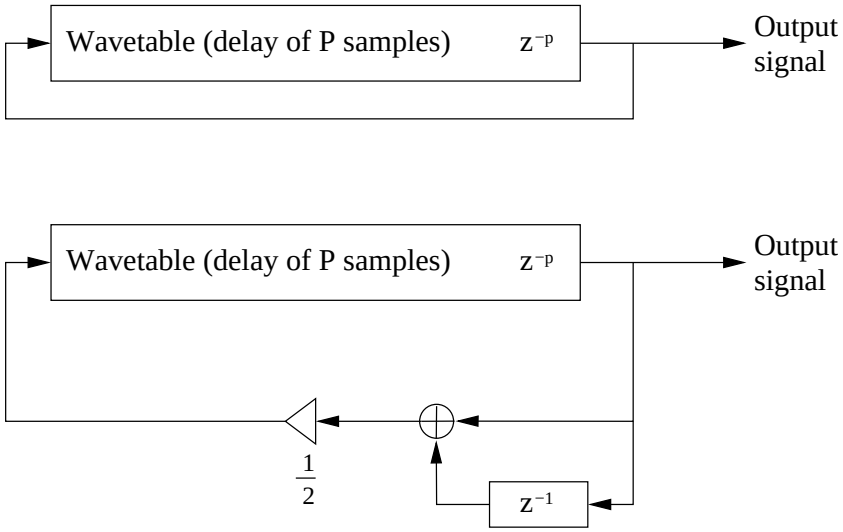

Figure 1.14: *The simple wavetable synthesis algorithm (top) and the Karplus-Strong model for plucked string tones (bottom) (from [111]).*

Smith [92] [64] has extended this algorithm and developed the concept of Digital Waveguide synthesis. The digital waveguide approach provides efficient computational models for musical instruments primarily in the string, wind, and brass families [101]. Digital waveguides and finite differences are closely related in that they are both based on the discretization of the wave equation. The efficiency of the digital waveguide approach lies in lumping the losses and dispersion of the structure to one point (assuming that the system is Linear and Time Invariant - LTI). Digital waveguides have been developed specifically for piano synthesis [104, 114, 101, 43]. These approaches are close to the one presented in this document. We will described in details digital waveguide algorithm in Section 2.3 and those approaches in Sections 2, 3 and 4.

1.4 Conclusion

Several physical phenomena are involved in the sound-production mechanism in the piano. Some are taken into account in the model proposed here, though others, of presumably lesser perceptual significance, are ignored. String inharmonicity and frequency-dependent losses are dealt with. Inter-string couplings, between string pairs or triples, giving rise to two-

stage decay, are also covered; however, only string vibration in direction perpendicular to the soundboard is modeled (vibration in the parallel plane does not transmit energy efficiently to the soundboard). Longitudinal waves, also present in a vibrating string, are also not modeled, and neither are nonlinear effects; their perceptual effects are of much less importance than those mentioned above. As far as the hammer is concerned, the felt nonlinearity on the hammer/string interaction is a necessary feature to be modeled, as is the modulation of the spectrum as a function of the hammer striking point.

Because we would like to transfer the physics directly to the synthesis method, we will not make any use of prerecorded sound (samples) or of non-physically motivated techniques such as FM synthesis. In addition, the parameter estimation methods that we will make use of will not give rise to a perfect reproduction of the analysis sound. Linear models give good results, from the perceptual point of view, but the relationship with physical parameters is neither simple, nor easy to control. We have thus looked mainly at physical and quasi-physical models: synthesis via numerical integration (of the equations of motion of the hammer-string system, see Sections 2.2 and 3.4)), synthesis by source/filter methods, and through the use of digital waveguides.

Chapter 2

String modeling

2.1 Introductory remarks

We have seen that there are several methods, by which synthesis of piano string vibration is possible. Ideally, we have focused on those which allow a satisfactory resynthesis from a perceptual point of view, have little cost in calculation time and have its parameters directly tied to the physical phenomena. We will show that the digital waveguide satisfies these first two requirements in the case of Linear Time Invariant (LTI) systems. On the other hand, the synthesis by digital solving (finite difference scheme) of a partial differential equation (PDE) enables the use of physical synthesis parameters, thereby satisfying the third of these requirements. In forming a link between these two methods, we express the parameters of the digital waveguide (DWG) as a function of those of the PDE. We show, then, that the DWG equally satisfies this third requirement and will be then, finally, used as a synthesis model.

Section 2.2 describes two different PDE models that allow the simulation of piano string vibration. Section 2.3 gives the basic principles of synthesis by digital waveguide model and Section 2.4 compares the signals produced by each of these methods, for a set of physical parameters given. This work (Sections 2.2, 2.3 and 2.4) was the result of collaboration with researchers of Stanford University and will be published in [14]. The second part of this chapter is concerned particularly with the modeling of coupled strings by digital waveguide.

2.2 PDE string model

We will present in this Section two different PDE string models. The first have been proposed by Hiller and Ruiz [87, 62] and is one of the most realistic string models. It takes into account stiffness and frequency dependent loss phenomena but does not allow the identification with

the digital waveguide model (see Sections 1.3.2 and 2.3). In order to compare both approaches, we will then consider another PDE model allowing this identification.

2.2.1 A PDE piano string model

Several models of transverse wave propagation on a stiff string, of varying degrees of complexity, have appeared in the literature [87, 62, 11, 26, 27]. These models, intended for the synthesis of musical tones, are always framed in terms of a partial differential equation (PDE), or system of PDEs; usually, the simplified starting point for such a model is the *one-dimensional wave equation* (see Section 1.2.2 and [39]), and the most realistic features, such as dispersion, frequency-dependent loss and nonlinear hammer excitation (in the case of the piano string), are incorporated through several extra terms. Hiller and Ruiz [87, 62] have proposed one of the most advanced such model and used it as the basis for a synthesis technique, through the use of finite differences. Chaigne and Askenfelt [26] [27] have used this model to show that the time waveform on a struck piano string is simulated in this way to a remarkable degree of fidelity. Frequency-dependent loss, is modeled through the use of a third order time derivative perturbation to the dispersive wave equation; while a physical justification for the use of such a term is tenuous, it does give rise to perceptually important variations in damping rates. This model is described by the following equation:

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2} - \kappa^2 \frac{\partial^4 y}{\partial x^4} - 2b_1 \frac{\partial y}{\partial t} + 2b_3 \frac{\partial^3 y}{\partial t^3} \tag{2.1}$$

The first term on the right-hand side of the equation, in the absence of the others, gives rise to wave like motion, with speed c. The second "ideal bar" term [51] introduces dispersion, or frequency-dependent wave velocity, and is parameterized by a stiffness coefficient κ. The third and fourth terms allow for loss (b_1 and b_3 are the loss parameters), and if $b_3 \neq 0$, decay rates will be frequency-dependent. A complete model, for a piano string, is obtained by including hammer excitation term $f(x,t)$, possibly accounting for nonlinear effects, on the right-hand side, and by restricting the spatial domain to a finite interval and supplying a realistic set of boundary conditions (see Section 3.4).

In order to compare the solution given by the finite difference method to the one obtained using waveguide models, we have to calculate explicit formulae for dispersion and loss curves. By inserting a solution of the form

$$y(x,t) = e^{st+j\beta x} \tag{2.2}$$

into Eq. (2.1), we obtain a dispersion relation,

$$-2b_3 s^3 + s^2 + 2b_1 s + c^2\beta^2 + \kappa^2\beta^4 = 0 \tag{2.3}$$

This is a cubic in the variable s, which means that Eq. (2.1) could admit three solutions. The third solution corresponds to a fast exponential growth [87] which can, under certain condition, be unstable (see [14]). This third solution was not numerically of importance for the simulation proposed by Chaigne in [26, 27]. But, as we will see in Section 2.3, the digital waveguide model we choose, simulates the propagation of only two transverse waves traveling in both directions. It will then be formally impossible to relate the PDE model of Hiller and Ruiz with the digital waveguide we would like to use. We will propose in the next Section another PDE model which avoids a third solution and allows the identification to the waveguide model.

2.2.2 A second order model

A second-order model of one-dimensional wave propagation

The model of string vibration we will describe here has been proposed by Stefan Bilbao. It is of second order in time differentiation; frequency-dependent loss is introduced via mixed time-space derivative terms [28]. It is described by the equation

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2} - \kappa^2 \frac{\partial^4 y}{\partial x^4} - 2b_1 \frac{\partial y}{\partial t} + 2b_2 \frac{\partial^3 y}{\partial x^2 \partial t} \tag{2.4}$$

where b_1 and b_2 are the loss parameters.

This model differs from that of Hiller and Ruiz [62] only by the replacement of the term $2b_3 \frac{\partial^3 y}{\partial t^3}$ by $2b_2 \frac{\partial^3 y}{\partial x^2 \partial t}$. It also allows for frequency-dependent loss, but the system itself is of a quite different character, due to the decreased degree of the equation with respect to the time variable. The characteristic equation is now

$$s^2 + 2q(\beta)s + r(\beta) = 0 \tag{2.5}$$

with

$$q(\beta) = b_1 + b_2\beta^2 \qquad r(\beta) = c^2\beta^2 + \kappa^2\beta^4$$

and the roots are

$$s_\pm = -q \pm \sqrt{q^2 - r}$$

The condition that the initial value problem corresponding to Eq. (2.4) be *well posed* is that these roots have real parts which are bounded from above as a function of β. A more restrictive (and physically relevant) condition is that these roots have non-positive real part for all β, so that all exponential solutions are non-increasing. It is simple to show that this will be true for

$$q(\beta), r(\beta) \geq 0 \tag{2.6}$$

For q and r satisfying Eq. (2.6), the imaginary parts of these roots correspond to oscillation frequencies, and the real parts to loss. Clearly, for real wavenumbers β such that $q^2 \leq r$, the real parts of $s_\pm$ are simply $-q$. This case corresponds to normal damped wave propagation; notice in particular that if q depends on β, then damping rates will be wavenumber (and thus frequency) dependent. If $q^2 > r$, then both roots are purely real and non-positive, yielding damped non-traveling solutions. For $b_1, b_2 \geq 0$, condition Eq. (2.6) is satisfied and this PDE obviously possesses exponentially decaying solutions, and what is more, loss increases as a function of wavenumber. The PDE of Eq. (2.4) possesses traveling wave solutions when $q^2 \leq r$, which, for realistic values of the defining parameters for a piano string, includes the overwhelming part of the audio spectrum.

In order to relate this PDE model with a digital waveguide numerical simulation method, it is useful to write the expressions for dispersion and loss directly. Taking

$$s_\pm = \alpha \pm j\omega \tag{2.7}$$

over the range of β for which traveling wave solutions exist, we obtain

$$\alpha(\beta) = -b_1 - b_2\beta^2 \tag{2.8}$$

$$\omega(\beta) = \sqrt{-(b_1 + b_2\beta^2)^2 + c^2\beta^2 + \kappa^2\beta^4} \tag{2.9}$$

Boundary Conditions

Let us now restrict the spatial domain for the problem defined by Eq. (2.4) to $x \in [0, L]$. As Eq. (2.4) is of fourth order in the spatial derivatives, we need to supply two boundary conditions at either end, i.e., at $x = 0$ and $x = L$. Following Chaigne [26], we apply fixed boundary conditions,

$$y|_{x=0} = y|_{x=L} = \left.\frac{\partial^2 y}{\partial x^2}\right|_{x=0} = \left.\frac{\partial^2 y}{\partial x^2}\right|_{x=L} = 0 \tag{2.10}$$

For a solution of the form $y(x,t) = e^{st+j\beta x}$, from dispersion relation Eq. (2.5), there are thus four solutions for β in terms of s,

$$\beta_+^2(s) = \frac{-\gamma + \sqrt{\gamma^2 - 4\kappa^2(s^2 + 2b_1 s)}}{2\kappa^2}$$

$$\beta_-^2(s) = \frac{-\gamma - \sqrt{\gamma^2 - 4\kappa^2(s^2 + 2b_1 s)}}{2\kappa^2}$$

with

$$\gamma = c^2 + 2b_2 s$$

At frequency s, thus, any linear combination

$$y(x,t) = e^{st}\Big(\quad a_{+,+}e^{j\beta_+ x} + a_{+,-}e^{-j\beta_+ x} \\ + \quad a_{-,+}e^{j\beta_- x} + a_{-,-}e^{-j\beta_- x}\Big)$$

is a solution to Eq. (2.4). Applying the boundary conditions Eq. (2.10) to this solution gives the matrix equation

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ e^{j\beta_+ L} & e^{-j\beta_+ L} & e^{j\beta_- L} & e^{-j\beta_- L} \\ -\beta_+^2 & -\beta_+^2 & -\beta_-^2 & -\beta_-^2 \\ -\beta_+^2 e^{j\beta_+ L} & -\beta_+^2 e^{-j\beta_+ L} & -\beta_-^2 e^{j\beta_- L} & -\beta_-^2 e^{-j\beta_- L} \end{bmatrix}}_{\triangleq A(s)} \begin{bmatrix} a_{+,+} \\ a_{+,-} \\ a_{-,+} \\ a_{-,-} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Non-trivial solutions can occur only when $\det(A) = 0$, giving the relation

$$\left(\beta_+^2 - \beta_-^2\right)^2 \sin(\beta_+ L)\sin(\beta_- L) = 0 \tag{2.12}$$

Discounting the case $\beta_+^2 = \beta_-^2$ (which yields an identically zero solution $y(x,t)$), then obviously, solutions are of the form $\beta_+ = n\pi/L$, for integer $n \neq 0$ (similarly for β_-), and the modal frequencies s_n are, from the solutions Eq. (2.8) and Eq. (2.9) of the dispersion relation Eq. (2.5),

$$s_n = \alpha(n\pi/L) + j\omega(n\pi/L)$$

over wavenumbers for which a traveling solution exists (for small b_1 and b_2, this will be true for all n).

Finite difference schemes

In order to solve Eq. (2.4) numerically, we may approximate its solution over a grid with spacing X, and with time step T. To obtain an "explicit scheme", Eq. (2.4) can be written as

$$\begin{aligned} \delta_t^2 y &= c^2\delta_x^2 y - \kappa^2\delta_x^2\delta_x^2 y \\ &\quad -2b_1\delta_{t,0}y + 2b_2\delta_x^2\delta_{t,-}y + O(T, X^2) \end{aligned} \tag{2.13}$$

where the various difference operators are defined by

$$\begin{aligned} \delta_x^2 y(x,t) &= \frac{1}{X^2}\left(y(x+X,t) - 2y(x,t) + y(x-X,t)\right) \\ \delta_t^2 y(x,t) &= \frac{1}{T^2}\left(y(x,t+T) - 2y(x,t) + y(x,t-T)\right) \\ \delta_{t,0} y(x,t) &= \frac{1}{2T}\left(y(x,t+T) - y(x,t-T)\right) \\ \delta_{t,-} y(x,t) &= \frac{1}{T}\left(y(x,t) - y(x,t-T)\right) \end{aligned}$$

All these operators are "centered" about the point (x,t), except for the backward difference operator $\delta_{t,-}$, which is used in order to obtain an explicit algorithm. The approximation is first-order accurate in the time step T, and second-order accurate in the space step X. We may now rewrite Eq. (2.13) as a difference scheme, operating on the grid function y_m^n, indexed by integer m and n, which will serve as an approximation to $y(x,t)$ at the location $x = mX$, $t = nT$:

$$y_m^{n+1} = F(y_{m+2}^n, y_{m+1}^n, y_m^n, y_{m-1}^n, y_{m+1}^{n-1}, y_m^{n-1}, y_{m-1}^{n-1}) \tag{2.14}$$

with F a linear function.

To obtain an "implicit scheme", we have to rewrite Eq. (2.13)

$$\begin{aligned} \delta_t^2 y &= c^2\delta_x^2 y - \kappa^2\delta_x^2\delta_x^2 y \\ &\quad -2b_1\delta_{t,0}y + 2b_2\delta_x^2\delta_{t,+}y + O(T,X^2) \end{aligned} \tag{2.15}$$

where the various difference operators are now defined by

$$\begin{aligned} \delta_x^2 y(x,t) &= \frac{1}{X^2}\left(y(x+X,t+T) - 2y(x,t+T) + y(x-X,t+T)\right) \\ \delta_t^2 y(x,t) &= \frac{1}{T^2}\left(y(x,t+T) - 2y(x,t) + y(x,t-T)\right) \\ \delta_{t,0} y(x,t) &= \frac{1}{2T}\left(y(x,t+T) - y(x,t-T)\right) \\ \delta_{t,+} y(x,t) &= \frac{1}{T}\left(y(x,t+T) - y(x,t)\right) \end{aligned}$$

which leads to the implicit scheme:

$$G(y_{m+2}^{n+1}, y_{m+1}^{n+1}, y_m^{n+1}, y_{m-1}^{n+1}, y_{m-2}^{n+1}) = H(y_m^n, y_{m-1}^n, y_{m+1}^{n-1}, y_m^{n-1}, y_{m-1}^{n-1}) \tag{2.16}$$

with G and H linear functions.

We have here given one formulation to obtain an explicit or implicit scheme. They have both advantages and drawbacks. The explicit scheme is less expensive from a computational point of view (the new iterate is directly obtained from the old ones) but a stability condition is needed to ensure convergence: the space step and the time step are related and cannot be chosen independently (as small is the space step, as high is the sampling frequency). On the other hand, the implicit scheme is unconditionally stable (in the linear case, time and space steps may be chosen independently), is less dispersive, but needs more calculation time since it necessitates the inversion of a linear system at each iteration. Depending on the kind of simulation we are looking for, we may use an explicit or implicit scheme. We have used an explicit scheme for numerically integrating the PDE Eq. (2.4). For a detailed description of those numerical methods, see [25].

Synthesis

We have numerically calculated the solution of Eq. (2.4) by finite differences. Since we have obtained the analytical expression of the solution, the numerical solving is not necessary for the time being, but it will be when we couple the PDE model with an equation describing the interaction with the hammer. This interaction is not being taken into account in this section; we have simply imposed an initial and specific displacement to the string. The sound synthesis results obtained are then those of a plucked string and are significantly different from those of a piano.

2.3 Digital Waveguide model

Digital waveguide modeling is one of the most efficient methods for physics-based real-time modeling of musical instruments. As we have seen previously (Section 1.3.2), this method has been developed by Smith and is particularly usefull for piano string modeling [98].

2.3.1 Principle

As mentioned in Section 2.2.1, to arrive at a PDE modeling the piano string, it is fruitful to start with the ideal wave equation and add perturbation terms to give more realistic frequency-dependent dispersion and losses. The perturbed PDE is then numerically integrated via a finite-difference scheme. The digital waveguide approach interchanges the order of these operations: the ideal wave equation is integrated first using a trivial finite-difference scheme, and the resulting solutions are perturbed using *digital filters* to add frequency-dependent loss and dispersion.

It has been known since d'Alembert [34] that the ideal one-dimensional wave equation is solved exactly by arbitrary (sufficiently smooth) wave shapes propagating in both directions. The lossless wave equation for a vibrating string can be expressed as

$$\frac{\partial^2 y}{\partial x^2} = c^2 \frac{\partial^2 y}{\partial t^2} \tag{2.17}$$

where y is the displacement of the string, x the position along the string and c the wave speed. This differential equation has an explicit solution and can be solved by an arbitrary function of the form

$$y(x,t) = y_l(x - ct) + y_r(x + ct) \tag{2.18}$$

The functions $y_l(x - ct)$ and $y_r(x + ct)$ can be interpreted as respectively leftgoing and rightgoing traveling waves at the speed propagation c. The general solution of the wave equation

is a linear combination of those two traveling waves. The functions y_l and y_r have to be twice derivative. This means physically that the slope of the vibrating string can only have values much lower than one.

The digital waveguide formulation works directly with these traveling wave components. An isolated traveling wave is trivially simulated in practice using a *delay line*. An ideal vibrating string is then modeled as a pair of delay lines, one for each direction of travel. For digital implementation, the traveling waves are *sampled* at intervals of T seconds in time and X meters in space by changing the variables

$$\begin{cases} x \longrightarrow x_m = mX \\ t \longrightarrow t_n = nT \end{cases} \tag{2.19}$$

where m and n are the new integral-valued time and space variables. These variables are related by

$$c = \frac{X}{T} \tag{2.20}$$

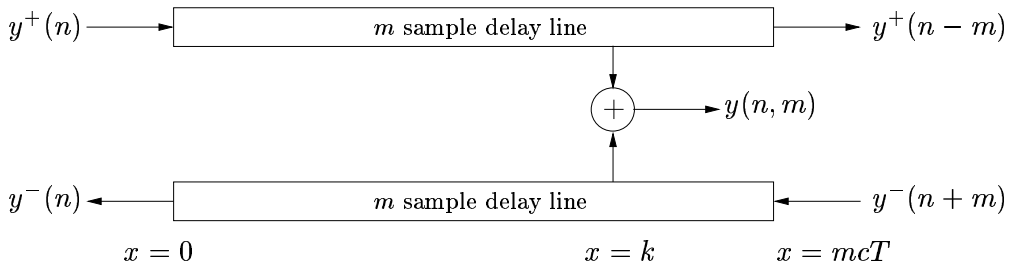

Figure 2.1: *The one-dimensional digital waveguide [98].*

By Shannon's sampling theorem [102], the solution remains exact, in principle, at all frequencies up to half the sampling rate. To avoid aliasing, all initial conditions and ongoing excitations must be *band-limited* to less than half the sampling rate $f_s = 1/T$. Substituting Eq. (2.19) into the traveling-wave solution Eq. (2.18) of the wave equations gives

$$\begin{aligned} y(t_n, x_m) &= y_r(t_n - x_m/c) + y_l(t_n + x_m/c) \\ &= y_r(nT - \frac{mX}{c}) + y_l(nT + \frac{mX}{c}) \\ &= y_r(T(n-m)) + y_l(T(n+m)) \end{aligned}$$

Since T multiplies all arguments, we can suppress it by defining [98]

$$y^+(n) = y_r(nT) \qquad y^-(n) = y_l(nT)$$

The + superscript denotes the traveling wave component going to the right and the − superscript denotes the traveling wave going to the left. The two discrete functions $y^+(n-m)$ and $y^-(n+m)$ can be interpreted as the output of an m-sample delay line whose input are respectively $y^+(n)$ and $y^-(n)$ [98]. The output from the waveguide at point k is obtained by summing the delay line variables at that point. A simulation diagram is shown in Figure 2.1. Any ideal one-dimensional waveguide can be simulated in this way. It is important to notice that contrary to finite differences, the solution provided by the waveguide is exact at the discrete points in the lossless case as long as the waveforms are originally bandlimited to one half of the sampling rate. On the other hand, using digital waveguide, we efficiently calculate the string vibration only at preselected point. Time and space are related using Eq (2.20) and we do not get the exact solution for the entire length of the string, loosing spatial information. This is not of importance for sound synthesis purposes, because only the behavior of the string at the bridge is of interest in most stringed instruments.

2.3.2 Dispersion and frequency-dependent losses in digital waveguide model

Consider a wave-like solution propagating from a point M_1 to a point M_2 along a string (see Figure 2.2, top). The distance M_1M_2 will be arbitrarily called l and the propagation time d at the *minimal* phase velocity. At the observation point M_2, the wave will have arrived after having undergone the effects of loss and dispersion. In terms of digital waveguides, the wave will undergo a pure delay (in the frequency domain, a multiplicative phase factor $\exp(-i\omega d)$), times a multiplicative factor $F(\omega)$ representing the loss and the dispersion experienced by the wave during this interval (see Figure 2.2, bottom). Since loss and dispersion are, for this system, linear time-invariant (LTI) phenomena, even when frequency dependent, the perturbations needed for added realism in the digital waveguide string model are LTI *digital filters*. Since LTI filters *commute*, we may *lump all* of the filtering associated with propagation in one direction into a *single* LTI filter.

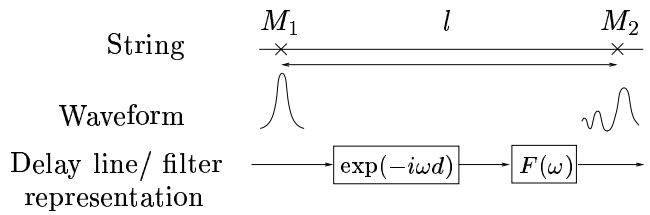

Figure 2.2: *The physical system and its coresponding delay line/filter.*

Figure 2.3 shows the simulation diagram for a digital waveguide model of a rigidly ter-

minated string. The string is excited by the signal E at a given location and its vibration is observed via the signal S at another location. The string is then divided into three parts. Sampled traveling velocity waves propagate to the right along the upper rail, and to the left along the lower rail. The rigid terminations cause inverting reflections (the two -1 scale factors). The delay lines are denoted D_i, $i = 1, 2, 3$, and the F_i blocks are digital filters taking into account loss and dispersion phenomena on each segment of the string.

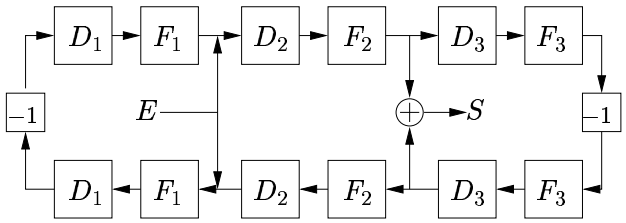

Figure 2.3: *Digital waveguide model of a rigidly terminated string.*

For purposes of computing the output signal S from the input signal E, Figure 2.3 may be further simplified:

- The two filters labeled F_1 can be replaced by a single filter F_1^2 (by commuting one of them with the inverting two delay lines D_1 and -1 gain). A similar simplification is possible for F_3.
- In the same way, the two delay lines labeled D_1 can be replaced by a single delay line D_1^2 (having twice the length of D_1), and the two D_3 blocks can be replaced by one D_3^2 block.

In general, any uniform section of a linear vibrating string which is excited and observed only at its endpoints can be accurately modeled (in one vibrational plane) by a pair of digital delay lines, each in series with a digital filter.

Since losses and dispersion are relatively *weak* in vibrating strings and acoustic bores, a *low-order* filter can approximate the distributed filtering associated with a particular stretch of string or bore. In practice, the desired loss and dispersion filters are normally derived from measurements such as the decay time of overtones in the freely vibrating string (see, e.g., [93, 71, 113]). In the next section, the filter will be derived from the stiff-string model of Eq. (2.4). Interestingly, the filter(s) so designed can be mapped back to an equivalent PDE, including many higher-order terms (which may or may not have a physical interpretation). Lumping of traveling-wave filtering in this way can yield computational savings by orders of magnitude relative to more typical finite difference schemes [99, 66].

2.4 From the wave equation to digital waveguide

2.4.1 Relating digital waveguide to the physical model

We address here the problem of relating digital waveguide filter parameters to the loss and dispersion curves from the physical model discussed in Section 2.2.2. For that, we consider the continuous frequency representation of the loop filter and show its relation with the physical parameters. The digital waveguide parameters can then be obtained by discretization. We will address the problem of the time domain implementation of the digital waveguide in Section 5.

According to Eqs. (2.2) and (2.7), the transformations of the wave due to propagation along the string segment of Figure 2.2 can be represented in terms of a digital waveguide filter by a multiplicative phase factor $\exp(sd + j\beta l)$. Ideally, the modulus and phase of this expression are related to the filter F by

$$|F(\omega)| = |e^{sd+j\beta l}| = e^{\alpha d} = e^{-(b_1+b_2\beta^2)d} \tag{2.21}$$

$$\arg(F(\omega)) = \arg(e^{sd+j\beta l}) = \omega d + \beta l \tag{2.22}$$

In order to write the expressions of the modulus and the phase of the loop filter in terms of the frequency ω, it is necessary to express the wavenumber β in terms of ω. From Eq. (2.9), solving for β, one gets

$$\beta^2(\omega) = \frac{-\eta_1 \pm \sqrt{\eta_1^2 + 4\eta_2(b_1^2 + \omega^2)}}{2\eta_2}$$

with

$$\eta_1 = c^2 - 2b_1b_2 \qquad \eta_2 = \kappa^2 - b_2^2$$

Since β is real (see Section 2.2.2), we keep only the solution for which the term inside the root is positive. Then,

$$\beta(\omega) = \pm\sqrt{\frac{-\eta_1 + \sqrt{\eta_1^2 + 4\eta_2(b_1^2 + \omega^2)}}{2\eta_2}}$$

Given that, for realistic piano string modeling, $b_1 \simeq 1$, $b_2 \simeq 10^{-4}$, $c \simeq 200$, $\kappa \simeq 1$ and $\omega \simeq 400$, we make the simplifying assumptions

$$b_1b_2 \ll c^2 \qquad b_2^2 \ll \kappa^2 \qquad b_1^2 \ll \omega^2 \tag{2.23}$$

which permits the following approximation of β:

$$\beta(\omega) \simeq \pm\frac{c}{\kappa\sqrt{2}}\sqrt{-1 + \sqrt{1 + 4\frac{\kappa^2}{c^4}\omega^2}}$$

In practice, it is helpful to work with more perceptually significant parameters. For that purpose, we will now suppose that the string is of length L, with perfect reflections at the extremities. The delay D, which corresponds to the propagation time of the minimum phase-velocity wave along the length L can be expressed as:

$$D = L/c = \pi/\omega_0. \tag{2.24}$$

where ω_0 is the fundamental frequency (rad/sec) of the ideal (stiffless) string and c is the minimum phase-velocity given by

$$c = \omega_0 L/\pi.$$

Thus, D is the propagation time over distance L for a sinusoidal traveling wave tuned to the first resonant mode of the ideal string. This choice of nominal propagation-time D simplifies the frequency-domain computations to follow; however, since phase velocity increases with frequency, the associated propagation filter F will be non-causal in the time domain. This poses no difficulty for frequency-domain implementation.

Next, we may express β as

$$\beta(\omega) \simeq \pm\frac{\pi}{L\sqrt{2B}}\sqrt{\xi} \tag{2.25}$$

with

$$\xi = -1 + \sqrt{1 + 4B\omega^2/\omega_0^2}$$

in terms of the inharmonicity factor B [38] given by

$$B = \kappa^2\omega_0^2/c^4$$

We now have to choose the sign of β in the expression for the phase. Since we want the output signal to be *delayed* with respect to the input signal, the loop filter/pure delay combination has to be causal. This means that the phase of the whole transfer function must be negative, i.e.,

$$-\omega D + \arg(F(\omega)) < 0. \tag{2.26}$$

This indicates that we choose the negative solution for β in the phase expression. Finally, using Eq. (2.25), we arrive at approximate expressions for the modulus and phase of the filter F as a function of frequency ω:

$$|F(\omega)| \simeq \exp\left(-D\left[b_1 + \frac{b_2\pi^2\xi}{2BL^2}\right]\right) \tag{2.27a}$$

$$\arg(F(\omega)) \simeq \omega D - \pi\sqrt{\frac{\xi}{2B}} \tag{2.27b}$$

These expressions serve as the link the PDE model of Eq. (2.4) to the lumped filters of the digital waveguide, under approximations Eq. (2.23). For the sake of simplicity, one can choose an "elementary filter" $\delta F = F^{1/D}$ such that

$$\begin{cases} |\delta F(\omega)| \simeq \exp\left(-\left[b_1 + \frac{b_2 \pi^2 \xi}{2BL^2}\right]\right) \\ \arg(\delta F(\omega)) \simeq \omega - \omega_0 \sqrt{\frac{\xi}{2B}} \end{cases} \tag{2.28}$$

The filters F of the digital waveguide, which correspond to propagation over a time duration D, can be then easily expressed in terms of δF by

$$F = \delta F^D$$

The stability of the digital waveguide model is ensured if the modulus of the filter δF is less than one. This condition is here always fulfilled, since the expression in the exponential of Eq. (2.27a) is negative.

2.4.2 Numerical simulations

We now address the validity of the analytical approach of the previous section to determine digital waveguide loop filters. The modulus and phase of the filters corresponding to the digital waveguide model have been directly linked to the parameters of the physical model.

Table I: Typical physical model parameters for piano tones C2, C4 and C7, after [27].

	C2	C4	C7	units
L	1.23	0.63	0.10	m
c	160.9	329.6	418.6	$\mathrm{m{\cdot}s^{-1}}$
κ	0.58	1.25	1.24	$\mathrm{m^2{\cdot}s^{-1}}$
b_1	0.25	1.1	9.17	$\mathrm{s^{-1}}$
b_2	$7.5{\times}10^{-5}$	2.7×10^{-4}	2.1×10^{-3}	$\mathrm{m^2{\cdot}s^{-1}}$
F_s	16 000	32 000	96 000	$\mathrm{s^{-1}}$

For a given set of PDE parameters, thus, we can design a complete digital waveguide model simulating the signal of the vibrating string at a given location, for a predetermined excitation location. For purposes of comparison, we have generated output signals using both a finite difference scheme (for more details see [14]), and the digital waveguide model for the same set of model parameters.

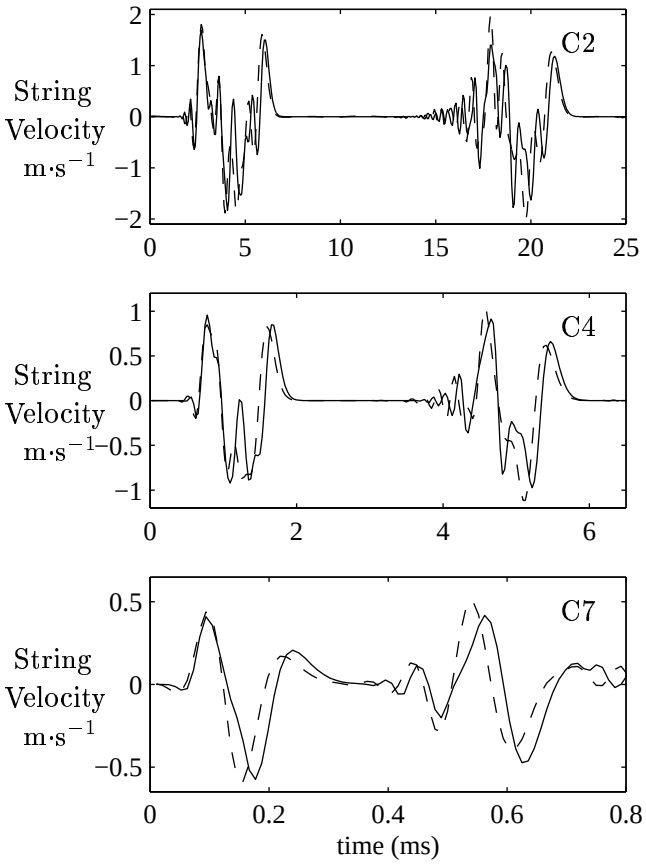

Figure 2.4: *Velocity signals obtained from a finite difference scheme (solid line) and a waveguide model (dotted line), for three different notes and for two periods of sound. The string is excited at distance $L/8$ from one endpoint, and the output signal is measured at distance $9L/10$.*

The digital waveguide has been computed in the frequency domain, allowing the use of Eq. (2.27a). The excitation, a Gaussian function simulating the initial velocity of the string at a distance $L/8$ from one end is the same in both cases. The signal is observed at the location $9L/10$. We have performed simulations for the notes C2, C4 and C7 using the parameters in Table I. Figure 2.4 shows the two first periods of the waveforms generated by both approaches. The amplitudes are similar. Nevertheless, there is a slight discrepancy between the two signals, due to the numerical dispersion introduced by the finite difference scheme. This discrepancy can be better seen by comparing the phase velocity of the two signals on Figure 2.5 (the phase velocity of the signal produced by the waveguide is similar to theoretical phase velocity of the model. We have then plotted only one curve).

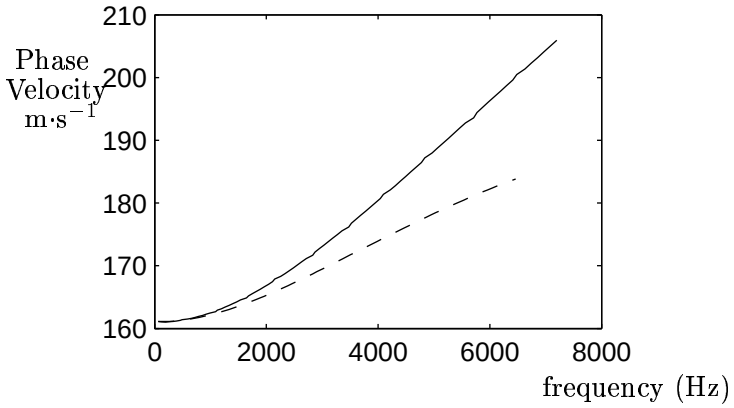

Figure 2.5: *Phase velocity for the waveguide model (solid line) and for difference scheme as a function of the frequency for the note C2.*

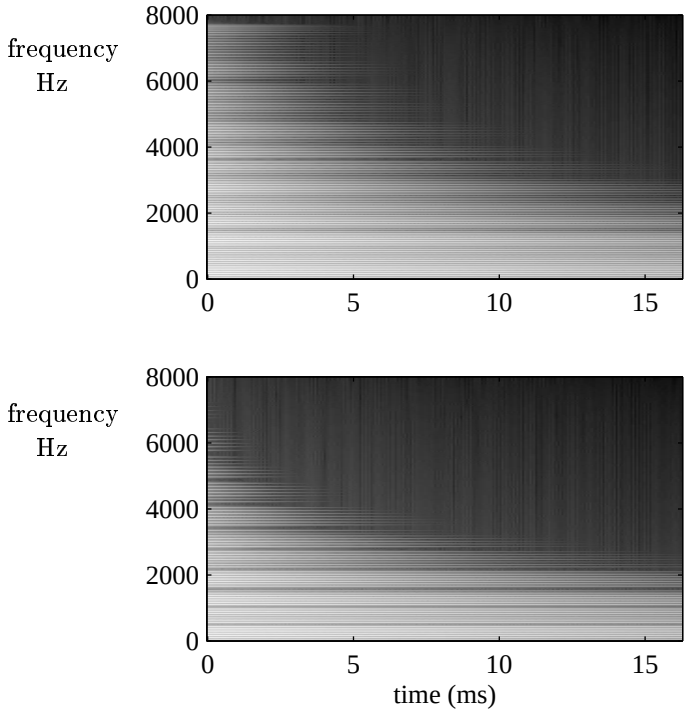

Figure 2.6: *Spectrograms of the output of the finite difference scheme (at top) and the waveguide model (at bottom) for the note C2.*

The long-time behavior of the generated signal is also very similar. In Figures 2.6, 2.7 and 2.8 spectrograms obtained over the whole length of the sound are shown. It is clear

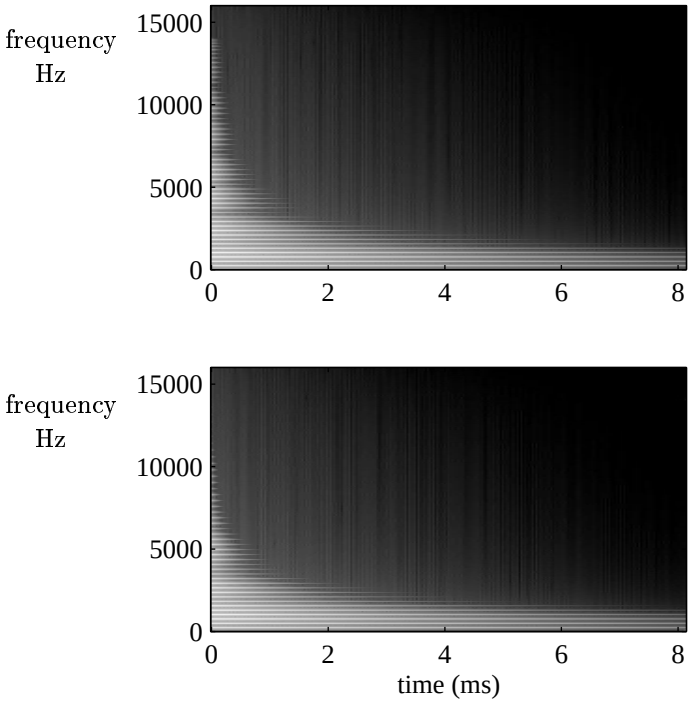

Figure 2.7: *Spectrograms of the output of the finite difference scheme (at top) and the waveguide model (at bottom) for the note C4.*

that the global damping behavior is similar in the low-frequency range. However, at high frequencies, the finite-difference model suffers from an artificially high propagation gain. The fundamental frequencies are essentially equal, but the wave dispersion due to string stiffness is different due to the *numerical* dispersion introduced by the difference scheme. In summary, these figures illustrate the extent to which the waveguide model provides a more accurate digital simulation of stiff, lossy strings with respect to both attenuation and dispersion of wave propagation.

2.5 Simple and coupled digital waveguide model

The digital waveguide model is actually a source-resonator model as we defined it in Section 1.3.2 : a wide-band excitation signal filtered using a resonant filter. It efficiently simulates string vibrations. But, as we will see, it can be further simplified. Moreover, we show that by coupling several waveguides, it is possible to accurately reproduce coupling phenomena between strings. Finally, we give the resolution of the inverse problem, allowing to calibrate the coupled digital waveguide model thanks to experimental data.

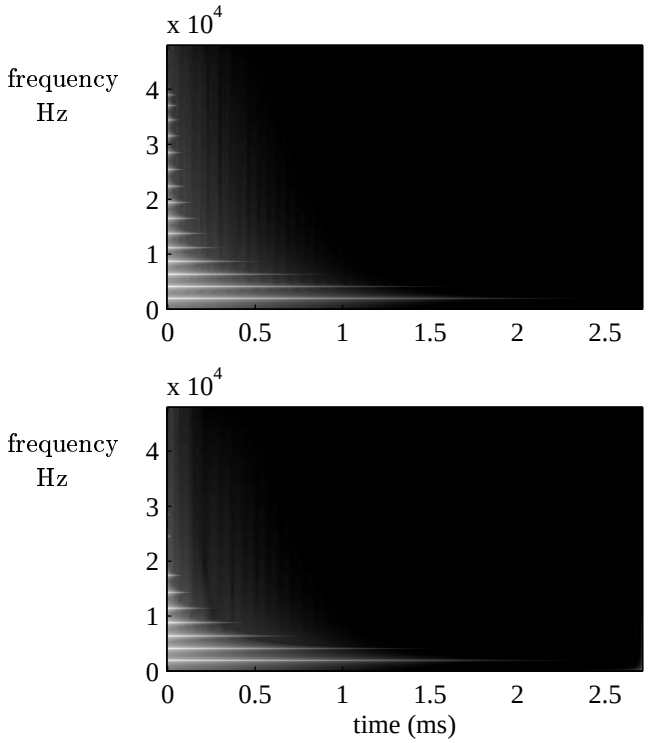

Figure 2.8: *Spectrograms of the output of the finite difference scheme (at top) and the waveguide model (at bottom) for the note C7.*

The construction of such an equivalent model using PDE is more complicated and expensive from a computational point of view. The calibration would be more difficult and does not a priori allow a perfect resynthesis of a given sound.

2.5.1 Source-resonator approach for modeling piano tones

The use of a source-resonator model requires that the source and the resonator are uncoupled. This is not the case in the piano since the hammer interacts with the strings during 2 to 5 milliseconds. A significant part of the piano sound characteristics is due to this interaction. Even though this observation is true from a physical point of view, this short interaction period is not in itself of great importance from a perceptual point of view. As we have seen in Section 1.2.1 the attack is actually constituted by two parts produced thanks to two different vibrating ways [21]: one percussive, a result of the impact of the key on the frame, and another which starts when the hammer strikes the strings. Actually, Schaeffer [89] shows that cutting the first milliseconds of a piano sound (for a bass note, for which the impact of the key on the frame is less perceptible) doesn't alter the perception of the sound. We

have carried out such an experiment using data recorded on an experimental setup and have concluded that, from a perceptual point of view, assuming that there is no impact noise due to the impact of the key on the frame, the hammer-string interaction is not audible in itself even though it undoubtedly plays an important role as an initial condition for the string motion. This is a substantial point justifying the dissociation of the string model and the source model in the design of our synthesis model. Thus, the resulting model consists of a source-resonator system. The resonator is modeled using a physically related model: a digital waveguide. The source model will be presented in Sections 3 and 4. Such an approach has already be studied in the piano case in the literature (S.A. Van Duyne and J.O. Smith [114][104], F. Avanzini, B. Bank [12], G. Borin, G. De Poli and D. Rocchesso [9]). As we have seen in the introduction, we have here worked in an analysis-synthesis framework, which means that contrary to what have been done before, our point was to accurately reconstruct the signal measured on single and coupled strings.

2.5.2 One string digital waveguide model

To model the wave propagation in a piano string, we can simplify the waveguide model of Figure 2.3 to design an "elementary" waveguide model (represented on Figure 2.9). Consider a wave traveling back and forth along a piano string. Since the system is linear and time invariant, all the dissipation and dispersion phenomena over one period can be lumped into only one digital filter and only one delay is necessary to take into account the propagation time. Thus, our model consists of a single loop system including:

- a delay line (a pure delay filter denoted D) simulating the duration the waves take to travel back and forth in the medium,
- a filter (denoted F) taking into account the dissipation and dispersion phenomena, together with the boundary conditions. The modulus of F is then related to the damping of the partials and the phase to the inharmonicity in the string.

The output S which represents the vibrating signal measured at one end of the string (at the bridge) and the input E corresponds to the frequency-dependent energy transferred to the string. The excitation location on the string will also be taken into account by the input: the source spectrum will be modulated with respect to the hammer position.

The transfer function T of this elementary digital waveguide is given by:

$$T(\omega) = \frac{S(\omega)}{E(\omega)} = \frac{F(\omega)e^{-i\omega D}}{1 - F(\omega)e^{-i\omega D}} \tag{2.29}$$

with ω the radial frequency.

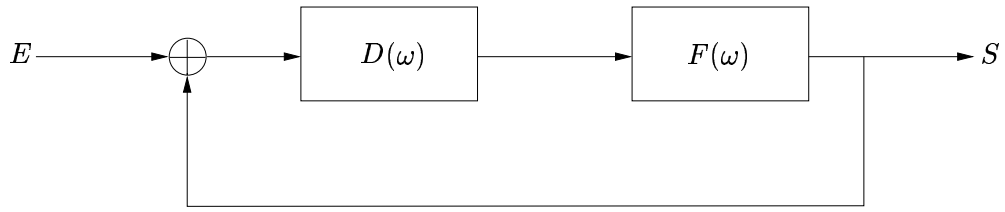

Figure 2.9: *Elementary digital waveguide representing a single string.*

2.5.3 Two and three coupled digital waveguides model

As we have seen in Section 1.2.2, coupling phenomena between strings are very important from a perceptual point of view. We have then designed a coupled waveguide, for two and three coupled strings allowing to reproduce those phenomena.

Smith [12] proposed a coupling model with two elementary waveguides. He assumed that the two strings were coupled to the same termination, and that the losses were lumped to the bridge impedance. This technique leads to a simple model only necessitating one loss filter. Nevertheless, the decay times and the coupling of the modes are not independent. Another approach proposed by Karjalainen [13] consists in coupling two digital waveguides through real gain amplifiers. In that case, the coupling is the same for each partial and the time behavior of the partials is similar. We have designed two models which are an extension of the Karjalainen's approach. They consist in separating the time behavior of the components by using complex-valued and frequency-dependent amplifiers (linear filters) to couple the waveguides (see Figures 2.10 and 2.11). Those models accurately simulate the energy transfer between the strings.

Each string is modeled using an elementary waveguide (named G_1, G_2, G_3). The coupled model is then obtained by connecting the output of each elementary waveguide to the input of the others through coupling filters. The coupling filters simulate the wave propagation along the bridge and are thus correlated to the distance between the strings. In the case of a doublet of strings, the two coupling filters (named C) are identical. In the case of a triplet of strings, the coupling filters of adjacent strings (named C_a) are equal but differ from the coupling filters to the extreme strings (named C_e). The excitation signal is assumed to be the same for each elementary waveguide since we suppose the hammer to struck the strings in a similar way.

The transfer functions of the coupled model are given by (all the filters are frequency

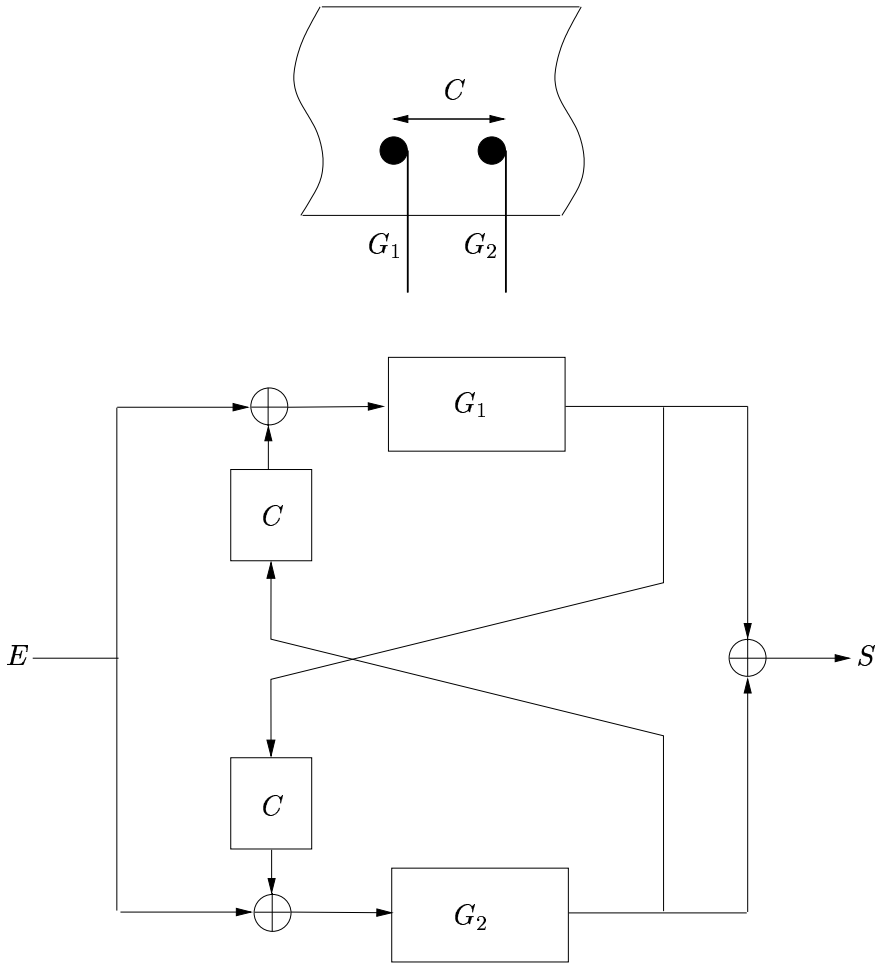

Figure 2.10: *The two-coupled digital waveguide model.*

dependent but for sake of simplicity, we note G_i for $G_i(\omega)$ $(i = 1, 2, 3)$ and C_j for $C_j(\omega)$) $(j = e, a)$

$$T(\omega) = \frac{G_1 + G_2 + 2CG_1G_2}{1 - C^2G_1G_2} \tag{2.30}$$

for the two strings model and

$$T(\omega) = \frac{G_1 + G_2 + G_3 + 2C_a(G_1G_2 + G_2G_3) + 2C_eG_1G2 + (4C_eC_a - C_e^2)G_1G_2G_3}{1 - C_a^2(G_1G_2 + G_2G_3) - C_e^2G_1G_3 - 2C_a^2C_eG_1G_2G_3} \tag{2.31}$$

for the three strings model.

The strings have the same length and are tuned to a very close pitch. Without loss of generality, one can assume that $D_1 = D_2 = D_3 = D$, since the difference in delays can be

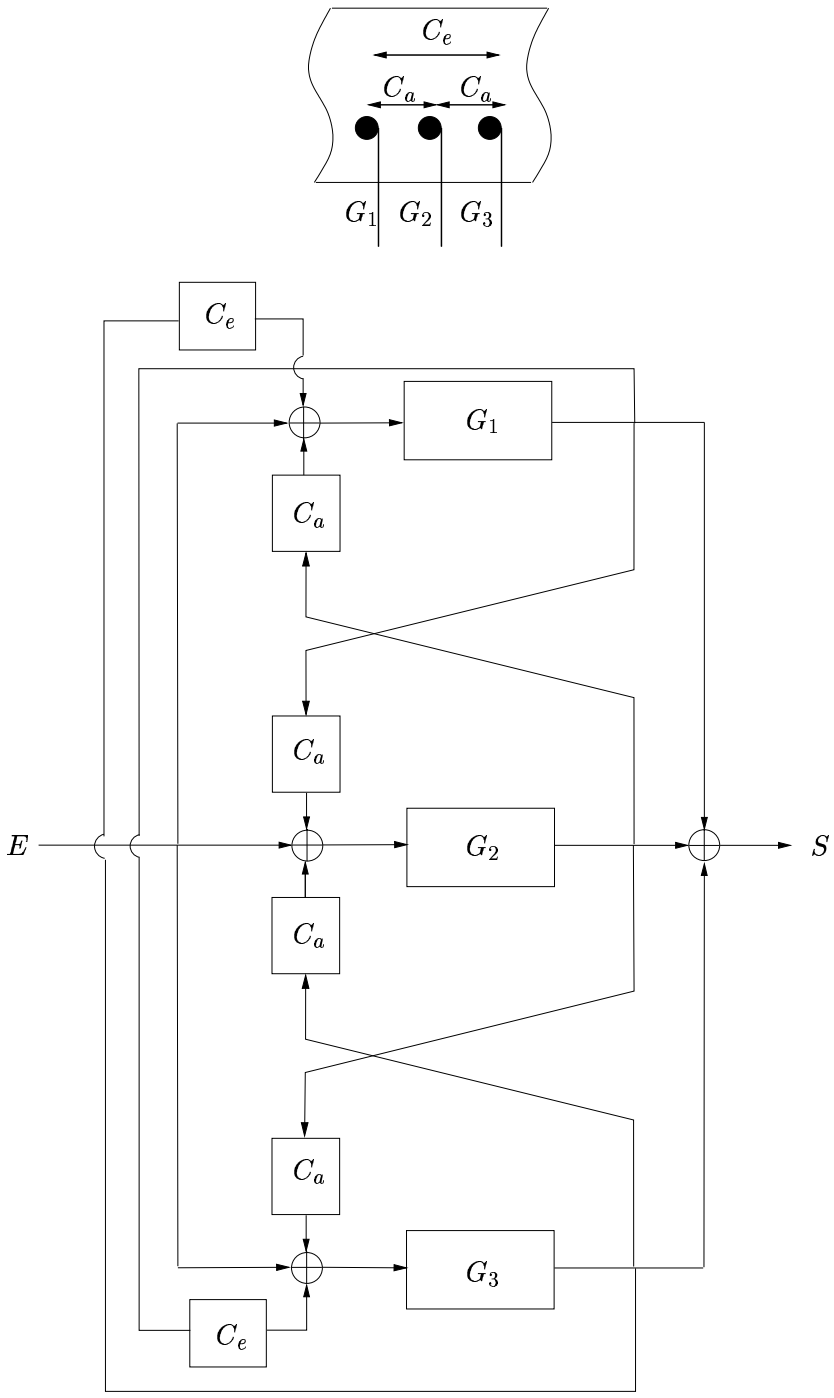

Figure 2.11: *The three-coupled digital waveguide model.*

taken into account in the phase of the filter F_i. By replacing the elementary waveguides by their transfer function (Eq. (2.29)), we obtain

$$T(\omega) = \frac{F_1 e^{-i\omega D} + F_2 e^{-i\omega D} + 2F_1F_2(C-1)e^{-i2\omega D}}{1 - (F_1 + F_2)e^{-i\omega D} + F_1F_2(1-C^2)e^{-i2\omega D}} \tag{2.32}$$

for the two strings model and

$$T(\omega) = \frac{N_1}{N_2} \tag{2.33}$$

with

$$\begin{aligned} N_1 = F_1 + F_2 + F_3 + 2[(C_a - 1)(F_1F_2 + F_2F_3) + (C_e - 1)F_1F_3] \\ +F_1F_2F_3[3 + 4C_eC_a - 4C_a - 2C_e - C_e^2] \\ N_2 = 1 - (F_1 + F_2 + F_3) + (F_1F_2 + F_2F_3 + F_1F_3)(1 - C_a^2) \\ +F_1F_2F_3(2C_a^2 + C_e^2 - 2C_a^2C_e - 1 \end{aligned}$$

for the three strings model.

Figures 2.12 shows the waveform and a partial of a signal produced using the two-coupled waveguide model. We can see beats and double decays on the amplitude of the signal. The partials are actually constituted by two or more components, depending on the number of waveguides we coupled.

2.6 Analysis and resynthesis using digital waveguides

In order to resynthesize a given piano sound, we have to accurately calibrate the model parameters. For that, we deal with the inverse problem, which consits in finding the expression of the model parameters as a function of parameters estimated on the measured signal (i.e, the modal parameters: *frequencies* of the modes, *amplitudes* and *damping coefficients*).

2.6.1 Inverse problem

single string

In the case of only one string (modeled using an elementary digital waveguide), the problem of relating the parameters of the waveguide to the parameters of the signal measured is quite easy. We have seen in Section 2.2 that a function of the form

$$y(x,t) = y_0 e^{st + j\beta x}$$

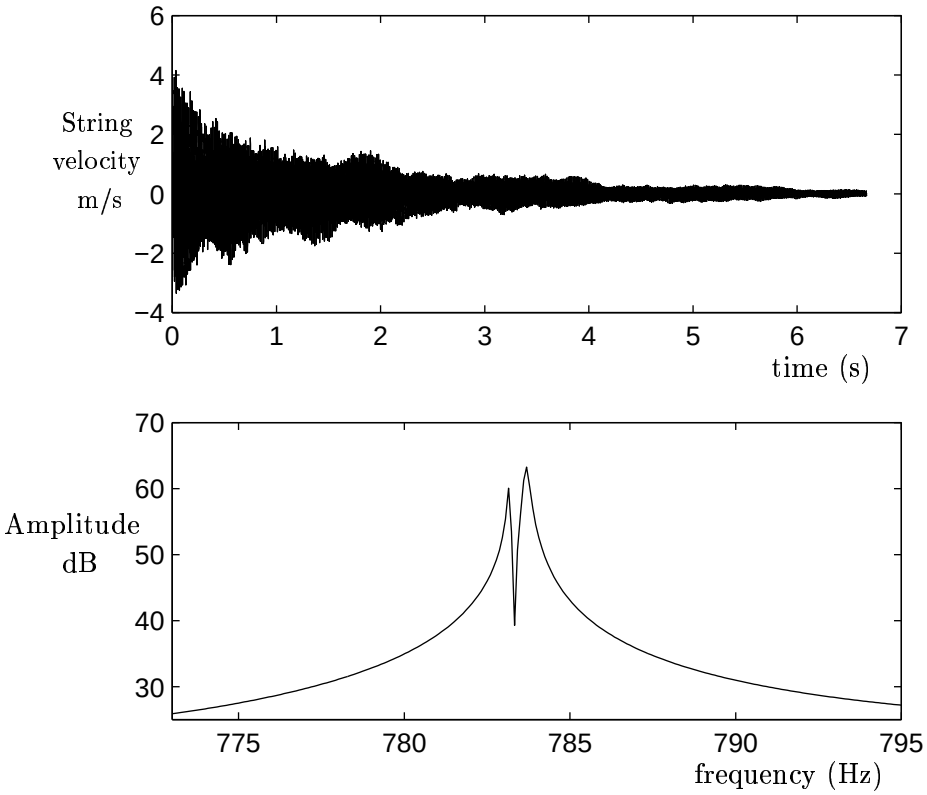

Figure 2.12: *Signal produced by the two coupled waveguide model. Waveform (top) and a doublet of the corresponding spectrum (bottom)*

is solution of the wave equation, assuming here that β is negative By applying boundary conditions, we concluded that the wave number β exists for a discrete set of values. We write then $\beta_k = -k\pi L$ where k is the partial number. The signal measured at a given location on the string can then be expressed as a sum of exponentially decaying sinusoids

$$s(t) = \sum_{k=1}^{\infty} a_k e^{-\alpha_k t} e^{i\omega_k t} \tag{2.34}$$

with a_k the initial amplitude of the k^{th} partial, depending on the excitation position on the string and on the position of the observation point, α_k and ω_k the damping coefficient and the frequency of the k^{th} partial.

The modulus and phase of F are related to α_k and ω_k by the expressions Eqs. (2.21) and

(2.22):

$$|F(\omega_k)| = e^{-\alpha_k D} \tag{2.35}$$

$$\arg(F(\omega_k)) = \omega_k D + \beta_k L = \omega_k D - 2k\pi \tag{2.36}$$

The calculation of the modulus and phase of the filter F necessitates the knowledge of α_k and ω_k for each partial. This can be easily done by using the analytical signal (see Section 2.6.3).

In practice, resynthesis also requires estimating the excitation signal. This excitation can be explicitly obtained by deconvolution. To avoid deconvolution difficulties such as division by values close to zero, we deconvoluate at the modal frequencies. We denote by $S(\omega)$, the Fourier transform of the output signal $s(t)$. In the frequency domain we have $S(\omega) = E(\omega)T(\omega)$ which can be written

$$E(\omega) = S(\omega)T^{-1}(\omega) \tag{2.37}$$

as we suppose that T (Eq. (2.29)) can be inverted. $S(\omega)$ is the Fourier transform of Eq. (2.34)

$$S(\omega) = \sum_{k=1}^{\infty} \frac{a_k}{\alpha_k + i(\omega - \omega_k)} \tag{2.38}$$

Then at the modal frequencies, we have

$$E(\omega_k) = \frac{a_k(1 - F(\omega_k)e^{-i\omega_k D})}{\alpha_k F(\omega_k)e^{-i\omega_k D}}$$

In order to simplify this expression, we will use a limited expansion of $e^{-\alpha_k}$, assuming that for a realistic piano string $\alpha_k \ll 1$.

$$F(\omega_k) = e^{-\alpha_k D}e^{i\omega_k D - 2k\pi} \simeq \frac{e^{i\omega_k D - 2k\pi}}{\alpha_k D + 1}$$

Finally, we get

$$E(\omega_k) \simeq a_k D \tag{2.39}$$

This expression gives a good estimation of the excitation signal, leading to a good resynthesis. Actually, it takes into account both the amplitude and the relative phase of the component (a_k is complex valued here).

Coupled strings

We will first show in this paragraph how the parameters of a two-coupled digital waveguide model can be expressed as function of the modal parameters. The case of three coupled

waveguides is quite similar.
The signal measured at the bridge level is the result of the vibration of two coupled strings. As we have seen, each partial is actually constituted by two components, having frequencies which are slightly different from the frequencies of each individual string (see Section 1.2.2). Similarly to Eq. (2.34) we write the measured signal as a sum of exponentially damped sinusoids

$$s(t) = \sum_{k=1}^{\infty} a_{1k} e^{-\alpha_{1k} t} e^{i\omega_{1k} t} + a_{2k} e^{-\alpha_{2k} t} e^{i\omega_{2k} t} \tag{2.40}$$

with a_{1k} and a_{2k} the initial amplitudes, α_{1k}, α_{2k} and ω_{1k}, ω_{2k} the damping coefficients and the frequencies of the components of the k^{th} partial. The Fourier transform of $s(t)$ is

$$S(\omega) = \sum_{k=1}^{\infty} \frac{a_{1k}}{\alpha_{1k} + i(\omega - \omega_1 k)} + \frac{a_{2k}}{\alpha_{2k} + i(\omega - \omega_2 k)} \tag{2.41}$$

We will now identify this expression locally in frequency with the output of the two-coupled waveguide model. For this purpose, since $T(\omega)$ is a rational fraction of second order polynomia in $e^{-i\omega D}$ (see Eq. (2.32)), it can be decomposed into a sum of two rational fractions of the first-order polynomia in $e^{-i\omega D}$

$$T(\omega) = \frac{P(\omega) e^{-i\omega D}}{1 - X(\omega) e^{-i\omega D}} + \frac{Q(\omega) e^{-i\omega D}}{1 - Y(\omega) e^{-i\omega D}} \tag{2.42}$$

The vibrations generated by the model are assimilated to a superposition of two series of partials whose frequencies and decay times are governed by the quantities $X(\omega)$ and $Y(\omega)$. By identification between Eqs. (2.32) and (2.42), we can determine the following system of 4 equations:

$$P + Q = F_1 + F_2 \tag{2.43}$$

$$PY + QX = 2F_1F_2(1 - C) \tag{2.44}$$

$$X + Y = F_1 + F_2 \tag{2.45}$$

$$XY = F_1F_2(1 - C^2) \tag{2.46}$$

We now identify Eq. (2.41) with the excitation signal times the transfer function T (Eq. (2.42))

$$S(\omega) = E(\omega)T(\omega) \tag{2.47}$$

Assuming that two successive modes do not overlap (these assumptions are verified for the piano sound) and by writing $X(\omega) = |X(\omega)|e^{i\Phi_X(\omega)}$ and $Y(\omega) = |Y(\omega)|e^{i\Phi_Y(\omega)}$, we express Eq. (2.47) near each double resonance as:

$$\frac{a_{1k}}{\alpha_{1k} + i(\omega - \omega_1 k)} + \frac{a_{2k}}{\alpha_{2k} + i(\omega - \omega_2 k)} \simeq \frac{E(\omega)P(\omega)e^{-i\omega D}}{1 - |X(\omega)|e^{-i(\omega D - \Phi_X(\omega))}} + \frac{E(\omega)Q(\omega)e^{-i\omega D}}{1 - |Y(\omega)|e^{-i(\omega D - \Phi_Y(\omega))}}$$

We will identify term by term the members of this equation. We take, for example

$$\frac{a_{1k}}{\alpha_{1k} + i(\omega - \omega_1 k)} \simeq \frac{E(\omega)P(\omega)e^{-i\omega D}}{1 - |X(\omega)|e^{-i(\omega D - \Phi_X(\omega))}}$$

The resonance frequencies of each doublet ω_{1k} and ω_{2k} correspond to the minimum of the two denominators $1 - |X(\omega)|e^{-i(\omega D - \Phi_X(\omega))}$ and $1 - |Y(\omega)|e^{-i(\omega D - \Phi_Y(\omega))}$. If we assume that modulus $|X(\omega)|$ and $|Y(\omega)|$ are close to one (this assumption is realistic because the propagation is weakly damped), we can determine the values of ω_{1k} and ω_{2k} :

$$\omega_{1k} = \frac{\Phi_X(\omega_{1k}) + 2k\pi}{D} \qquad \omega_{2k} = \frac{\Phi_Y(\omega_{2k}) + 2k\pi}{D} \tag{2.48}$$

Taking $\omega = \omega_{1k} + \epsilon$ with ϵ arbitrary small:

$$\frac{a_{1k}}{\alpha_{1k} + i\epsilon} \simeq \frac{E(\omega_{1k} + \epsilon)P(\omega_{1k} + \epsilon)e^{-i\Phi_X(\omega_{1k}+\epsilon)}e^{-i\epsilon D}}{1 - |X(\omega_{1k} + \epsilon)|e^{-i\epsilon D}}$$

A limited expansion of $e^{-i\epsilon D} \simeq 1 - i\epsilon D + \theta(\epsilon^2)$ around $\epsilon = 0$ (at the $zero^{th}$ order for the numerator and at the first order for the denominator) gives

$$\begin{aligned} E(\omega_{1k} + \epsilon)P(\omega_{1k} + \epsilon)e^{-i\Phi_X(\omega_{1k}+\epsilon)}e^{-i\epsilon D} &\simeq E(\omega_{1k})P(\omega_{1k})e^{-i\Phi_X(\omega_{1k})} \\ 1 - |X(\omega_{1k} + \epsilon)|e^{-i\epsilon D} &\simeq 1 - |X(\omega_{1k})|(1 - i\epsilon D) \end{aligned}$$

Assuming that $P(\omega)$ and $|X(\omega)|$ are locally constant (in the frequency domain), we identify term by term (the two members are considered as functions of the variable ϵ). We deduce the expressions of $|X(\omega)|$ and $|Y(\omega)|$ as a function of the amplitudes and decay times coefficients for each mode.

$$|X(\omega_{1k})| = \frac{1}{\alpha_{1k}D + 1} \qquad |Y(\omega_{2k})| = \frac{1}{\alpha_{2k}D + 1} \tag{2.49}$$

We also get the relations:

$$E(\omega_{1k})P(\omega_{1k}) = a_{1k}DX(\omega_{1k}) \tag{2.50}$$

$$E(\omega_{2k})Q(\omega_{2k}) = a_{2k}DY(\omega_{2k}) \tag{2.51}$$

From the measured signal, we will estimate the modal parameters a_{1k}, a_{2k}, α_{1k}, α_{2k}, ω_{1k} and ω_{2k}. Thanks to Eqs. (2.48) and (2.49) we can calculate X and Y. We still have six unknown variables P, Q, E, C, F_1 and F_2. But we also have a system of six equations (Eqs. (2.43) - (2.46), (2.50), (2.51)). Assuming that the two resonance frequencies are close and that the variables P, Q, E, C, F_1, F_2, X, Y have a locally smooth behavior, we can then express the waveguide parameters as function of the temporal parameters. For sake of simplicity, we note $P_k = P(\omega_{1k}) = P(\omega_{2k})$

Using Eqs. (2.43) and (2.45), we obtain $P_k + Q_k = X_k + Y_k$. Thanks to Eqs. (2.50) and (2.51) we finally get the expression of the excitation signal at the resonance frequencies

$$E_k = \frac{D(a_{1k}X_k + a_{2k}Y_k)}{X_k + Y_k} \tag{2.52}$$

We deduce the expressions of P_k and Q_k

$$P_k = a_{1k}X_k \frac{X_k + Y_k}{a_{1k}X_k + a_{2k}Y_k}$$
$$Q_k = a_{2k}Y_k \frac{X_k + Y_k}{a_{1k}X_k + a_{2k}Y_k}$$

Using Eqs. (2.44) and (2.46) we get the expression of the coupling filter

$$C_k = \frac{(a_{1k} - a_{2k})(X_k - Y_k)}{(a_{1k} + a_{2k})(X_k + Y_k)} \tag{2.53}$$

Using Eqs. (2.44), (2.45) and (2.53) we obtain after some calculation the expression of the waveguide filters

$$F_{1k} = \frac{X_k + Y_k}{2} + \frac{(X_k^2 - Y_k^2)\sqrt{a_{1k}b_{2k}}}{2\sqrt{(a_{1k}X_k + a_{2k}Y_k)(a_{1k}Y_k + a_{2k}X_k)}} \tag{2.54}$$

$$F_{2k} = \frac{X_k + Y_k}{2} - \frac{(X_k^2 - Y_k^2)\sqrt{a_{1k}b_{2k}}}{2\sqrt{(a_{1k}X_k + a_{2k}Y_k)(a_{1k}Y_k + a_{2k}X_k)}} \tag{2.55}$$

The same approach can be used to get the expression of the three-coupled waveguide parameters as a function of the modal parameters. The signal s will be, for each partial, a sum of three exponentially damped sinusoids. We obtain a system of nine equations which does not admit any analytical solution. We solve it numerically and obtain the values of the parameters E, F_1, F_2, F_3, C_a, C_e for a given set of modal parameters.

2.6.2 Experimental setup

We have designed an experimental setup (see Figure 2.13) allowing the measurement of the vibration of one, two or three strings struck by a hammer for different velocities. On the top

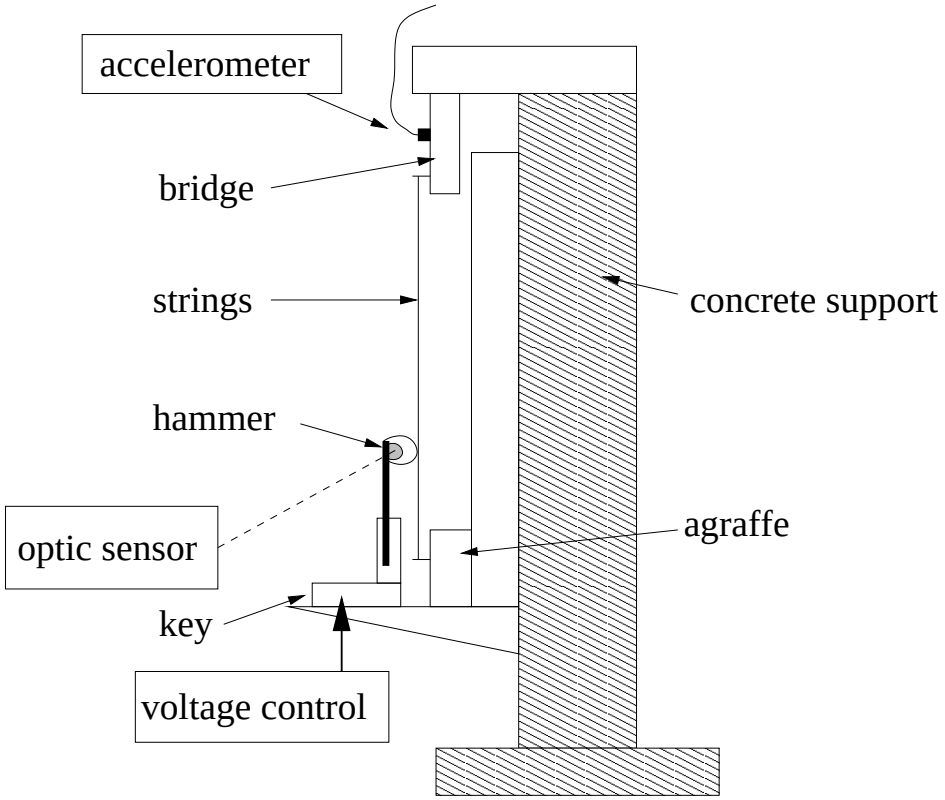

Figure 2.13: *The experimental setup.*

of a massive concrete support, we have attached a piece of a bridge taken from a real piano. On the other extremity of the structure, we have attached an agraffe on a hard wood support. The strings are then tightened between the bridge and the agraffe and tuned manually to the closest sounding pitch. It is clear that the strings are not totally uncoupled to their support. Nevertheless, this experiment has been used to record signals of struck strings, in order to validate the synthesis models, and was entirely satisfying for this purpose. One, two or three strings are struck with a hammer linked to an electronically piloted key. By imposing different voltages to the system, one can control the hammer velocity in a reasonably reproducible way. The precise velocity is measured immediately after escapement using an *optic sensor* (MTI 2000, probe module 2125H) pointing at the side of the head of the hammer. The vibration at the bridge level is measured by an *accelerometer* (B&K 4374). The signals are directly recorded on Digital Audio Tape. We have collected acceleration signals corresponding to hammer velocities varying between 0.8 ms^{-1} and 5.7 ms^{-1}.

The study of the interaction between the hammer and the string is not the purpose

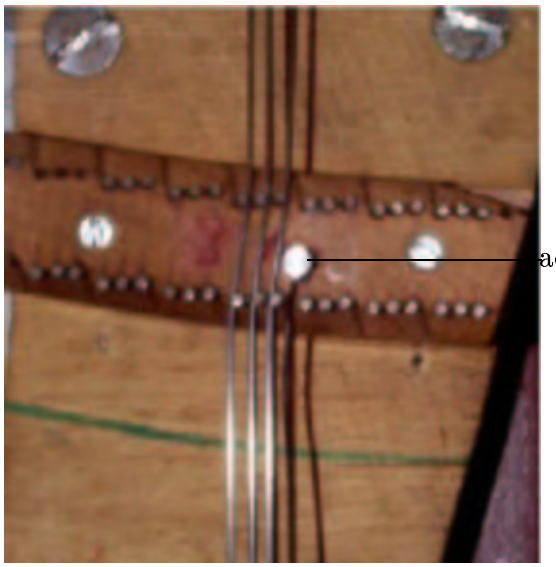

Figure 2.14: *Accelerometer at the piece of bridge.*

of this chapter. The hammer gives its energy to the string and excites a large frequency band allowing us to measure the vibratory behavior of the string. We will only analyze the string vibration during the linear part of the phenomena, i.e., after the interaction with the hammer, which is sufficient to totally characterize the resonator. Exciting the strings for several hammer velocities will however allow to validate the source-resonator approach. Indeed, if both elements of the model are correctly estimated and separated from the analysis, the resonances, which characterize the strings, should be invariant as a function of the hammer velocity.

2.6.3 Parameters estimation

The measured signal contains many partials which can be composed of several components due to the coupling between strings (see Section 1.2.2). To estimate the amplitudes, frequencies and damping factors associated with each component of the measured signal, we use a two-step process. The first step consists in isolating each component by a band pass filtering and then estimate the *average frequency* using the corresponding *analytic signal*. As a second step, a parametric method is used on this analytic signal to model each component and obtain the amplitude, frequency and damping factor. The diagram of Figure 2.17 summarizes the analysis process. This two step process was the most convenient and accurate way to estimate the parameters of the original signal. Indeed, due to the large number of components, the global parametric estimation on the whole signal led to unconvincing results.

Figure 2.15: *The hammer action and the optic sensor.*

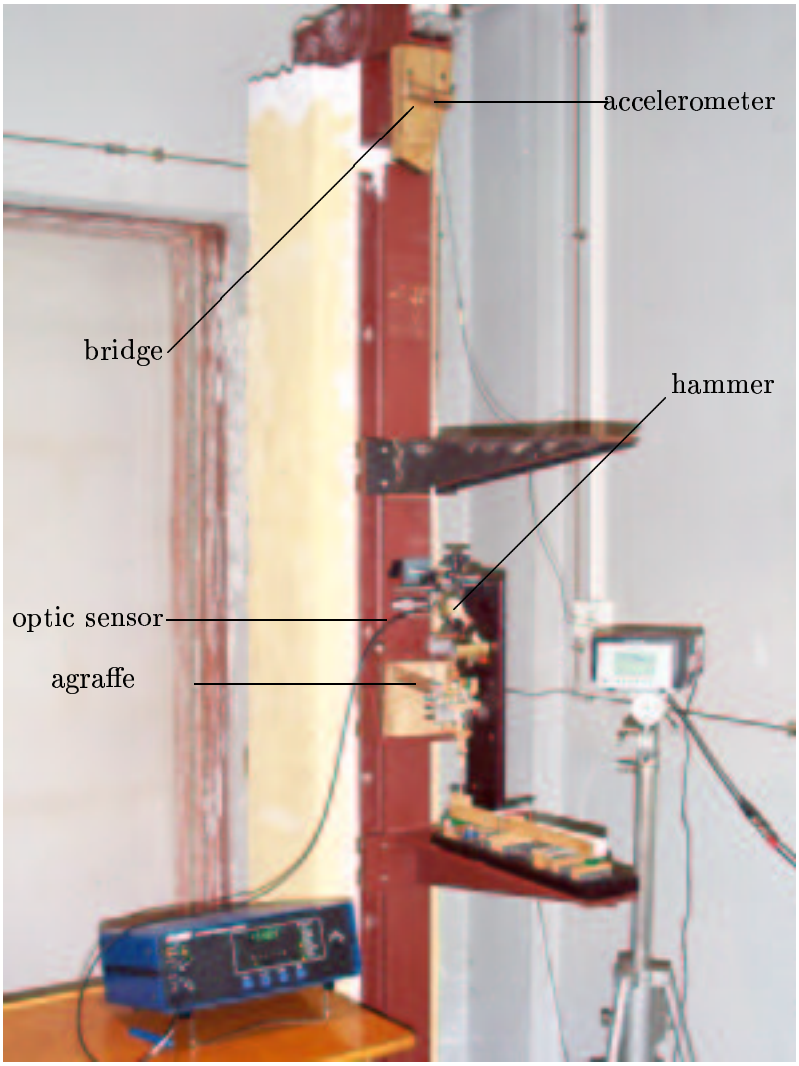

Figure 2.16: *The experimental setup.*

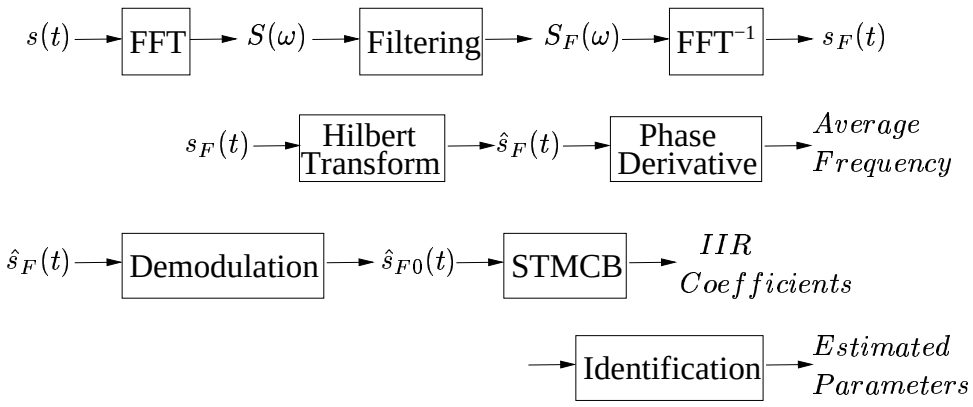

Figure 2.17: *The two steps of the analysis process.*

To isolate each partial in frequency, we used a truncated *Gaussian* window, the frequency bandwidth of which was chosen so as to minimize smoothing effects over the attack duration and to avoid overlapping two successive partials in frequency. The Gaussian was employed since its time-bandwidth product is minimized. As a consequence, it optimizes the exponential damping suport after convolution with a causal single component for a given band pass filtering. Then we use the *Hilbert transform* to get the analytic signal. The analytic signal (first introduced by D. Gabor [41]) is a complex-valued signal associated with a real-valued signal. If $s_F(t)$ is our real-valued filtered signal, the corresponding analytic signal $\hat{s}_F(t)$ is formally obtained using the relation $\hat{s}_F = (1 + iH)s_F$ where H denotes the Hilbert transform operator. An important property of such an analytic signal is that its Fourier spectrum is identical to that of the original signal, except that it vanishes for negative frequencies. Consequently, it allows an unambiguous definition of the frequency. The frequency is obtained using the relation

$$\nu(t) = \frac{1}{2\pi}\frac{\partial}{\partial t}\mathrm{Im}(\log(\hat{s}_F(t))) \tag{2.56}$$

In the case of a partial which contains a doublet or a triplet, this frequency is an average frequency.

As a second step, a parametric modeling of each partial is performed. To work on a small number of samples for better efficiency from a computational point of view, we first demodulate the component by its average frequency, allowing it to be subsampled drastically. Typically, we work on a 512-point sequence, with an original signal one thousand times longer (about 10 seconds). The time origin of the signal is then shifted to remove the biased part around the onset. Moreover, to avoid problems related to the poor signal to noise ratio at

high frequency, the duration on which the estimation is performed is restrained to the significant part of the signal. The resulting signal is then modeled using the *Steiglitz-McBride method* [105], which consists in finding an IIR filter with a prescribed time domain impulse response. For that purpose, the method attempts to minimize the squared error between the impulse response corresponding to the IIR filter and the input signal. If one constrains the method to estimate a one zero and two poles (for example in the case of two coupled strings) IIR filter, then amplitude, frequency, and decay time are directly deduced from the coefficients of the filter. Indeed, by identifying near each component, the expression of the IIR filter and the Z-transform of Eq. (2.40), we write

$$\frac{c_1 + c_2 z^{-1}}{1 + d_2 z^{-1} + d_3 z^{-2}} = \frac{a_{1k}}{1 - e^{-(\alpha_{1k} - i\omega_{1k})} z^{-1}} + \frac{a_{2k}}{1 - e^{-(\alpha_{2k} - i\omega_{2k})} z^{-1}} \tag{2.57}$$

where c_1, c_2, d_2 and d_3 are the IIR coefficients.
After identification, we can deduce the parameters a_{1k}, a_{2k}, α_{1k}, α_{2k}, ω_{1k} and ω_{2k} as a function of the filter coefficients. As the signal has been demoluted, we must add the average frequency to ω_{1k} and ω_{2k} to get the actual frequencies of the components. We have also tested various minimization methods (among them classical Linear Prediction Coding and Prony) and found that the iterative Steiglitz-McBride method was more robust for our problem.

2.6.4 Results

From the signals collected on the experimental setup, a set of data has been extracted. For each hammer velocity, the waveguide filters and the corresponding excitation signals have been calculated using relations Eqs. (2.52)- (2.55).

Resonator stability

Figure 2.18 shows the modulus of the filter response F for the first twenty five partials in the case of tones produced by a single string. Here the hammer velocity varies from 0.7 ms^{-1} to 4 ms^{-1}. One can notice that the modulus of the waveguide filters with respect to the hammer velocity are similar. The resonator represents the strings that are unchanged during the experiment. If the estimated resonator is still the same for different hammer velocities, it means that all the non-linear behavior due to the dynamic as been taken into account in the excitation part. The resonator and the source have been well uncoupled. This result validates our approach based on a source-resonator separation. Nevertheless, one can notice a slight decrease of the filter modulus as a function of the hammer velocity for high frequency partials. This non-linear behavior is not directly linked to the hammer string contact. It is mainly due to non-linear phenomena involved in the wave propagation. The larger the

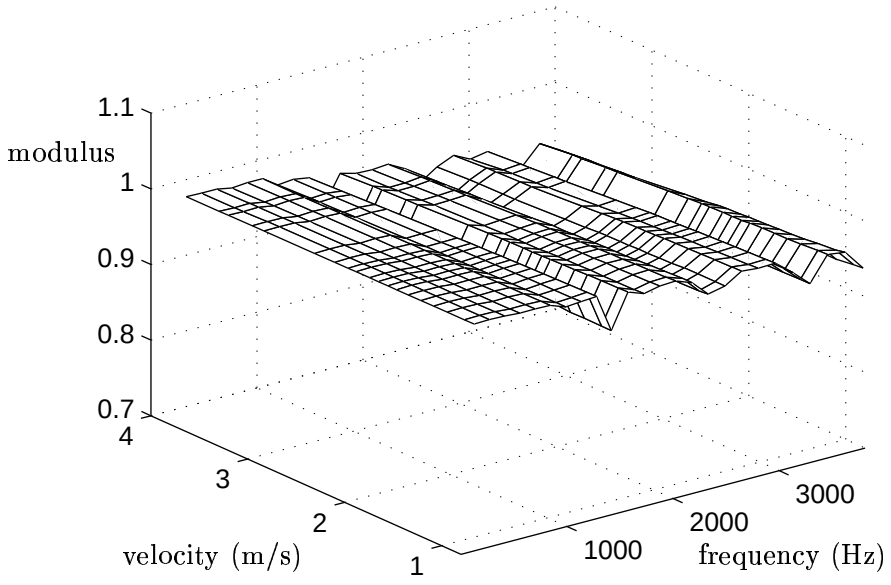

Figure 2.18: *Amplitude of the filter F as a function of the frequency and of the hammer velocity.*

amplitude of the motion, the larger the internal losses. The filter modulus slowly decreases from a value close to 1. Since the higher partials are more damped than the lower ones, the amplitude of the filter decreases while the frequency increases. The value of the filter modulus (close to 1) suggests that the losses are weak. This is true in the piano string case and even more obvious on this experimental setup, since the lack of a soundboard limits the acoustic field radiation. More losses are expected in the real piano.

Let's now consider the multiple strings case. From a physical point of view, the behavior of the filters F_1, F_2 and F_3 (which characterize the intrinsic losses) of the coupled digital waveguides should be similar to the one of the filter F for a single string, since the strings are supposed identical. This is verified except for high frequency partials. One can see this behavior on Figure 2.19 for the filter F_2 of the three-coupled waveguide model. Some artifacts pollute the drawing at high frequencies. These artifacts are mainly due to the poor signal/noise ratio at high frequency (above 2000 Hz) and low velocity. Nevertheless, this does not alter the synthetic sound since the corresponding partials are weak and of short duration.

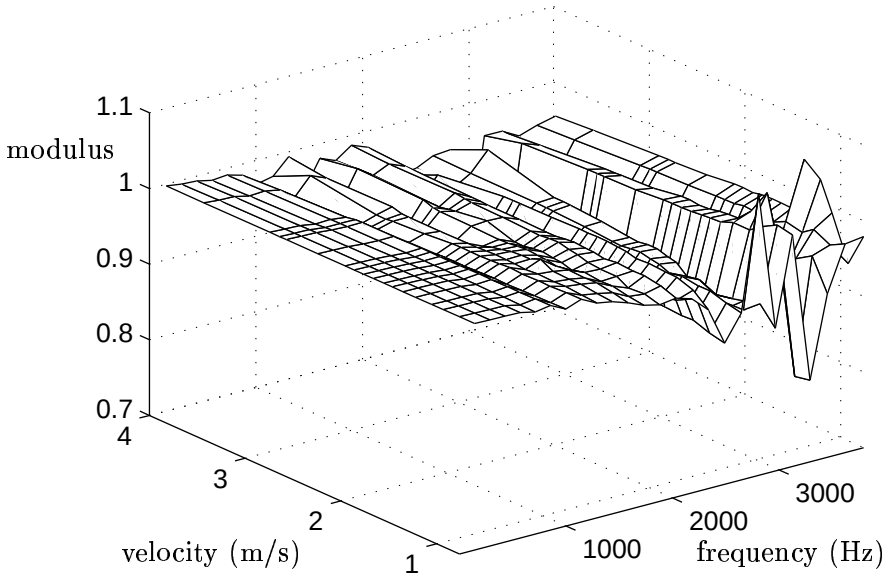

Figure 2.19: *Amplitude of the filter F_2 (three-coupled waveguide model) as a function of the frequency and of the hammer velocity.*

The phase is also of great importance since it is related to the group delay of the signal and consequently directly linked to the frequency of the partials. The phase is then a non-linear function of the frequency (see Eq. (2.27a)). The phase is constant with the hammer velocity (see Figure 2.20), since the frequencies of the partials are always the same (linearity of the wave propagation).

The coupling filters simulate the energy transfer between the strings and are frequency dependent. Figure 2.21 represents one of these coupling filters for different values of the hammer velocity. The amplitude is quite constant with respect to the hammer velocity (up to signal/noise ratio at high frequency and low velocity), showing that the coupling is independent of the amplitude of the vibration. The coupling seems to rise with the frequency. The peaks at frequencies 700 Hz and 1300 Hz correspond to a maximum of the energy transfer and are related to the impedance of the experimental setup termination.

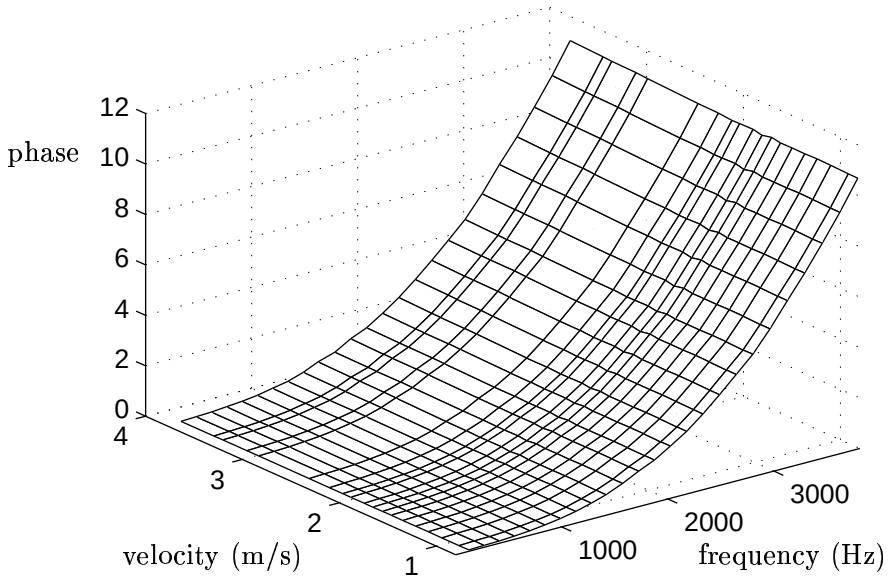

Figure 2.20: *Phase of the filter F as a function of the frequency and of the hammer velocity.*

Resynthesis

At this point, one can resynthesize a given sound by using one or multi-coupled digital waveguide and the parameters extracted from the analysis. For the synthetic sounds to be identical to the original one requires precisely describing the filters. This can be achieved by implementing the model in the frequency domain, as described in 2.5.2 and 2.5.3, allowing to take into account the exact amplitude and the phase of the filters. Nevertheless for real-time synthesis purposes, filters can be approached by IIR low order filter as it is usually done (see for example [113]).

By injecting the excitation signal obtained by deconvolution into the waveguide model, one can reproduce the signal measured on the experimental setup. From a perceptual point of view, the resulting sound is indistinguishable from the original ones. Figure 2.22 shows the original and resynthesized spectra for a one string signal. The perceptual characteristics of the original spectrum are well preversed. Figures 2.23, 2.24 and 2.25 show the amplitude modulation laws of the first six partials of the original and the re-synthesized sound. The variations of the temporal envelope are generally well retained, and for the coupled system,

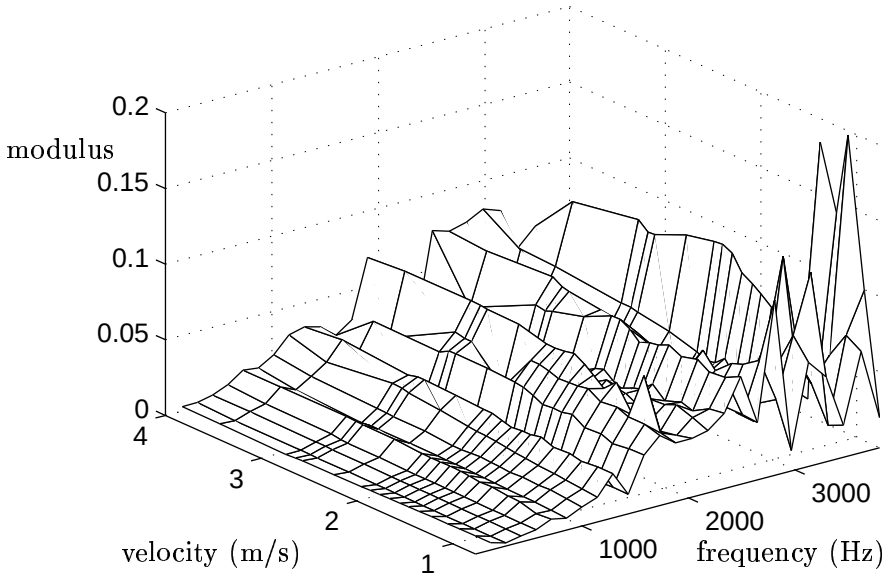

Figure 2.21: *Modulus of the filter C_a as a function of the frequency and of the hammer velocity.*

the beat phenomena are well reproduced. The slight differences that one can observe are not audible. They are due to fine physical phenomena that are not taken into account in our model. For example, let us consider the second and sixth partials of the original sound in Figure 2.23. We can see beats which show coupling phenomena on only one string. Indeed, the horizontal and vertical modes of vibration of the string are coupled through the bridge (see Section 1.2.2). This coupling was not taken into account in this study since the phenomenon doesn't alter the perceptual effect. The accuracy of the re-synthesis validates a posteriori the use of our model and the source-resonator approach.

2.7 Conclusion: PDE model and FDs versus waveguide model

As we have seen, the numerical solution of a PDE model using finite differences and the use of a digital waveguide both allow simulation of the propagation of waves in a piano string, taking into account the main physical features. The most notable distinction is the complete lack of numerical dispersion (which leads to mode mistuning) in the waveguide implementation. On the other hand, the waveguide allows the efficient computation of a solution ("sound")

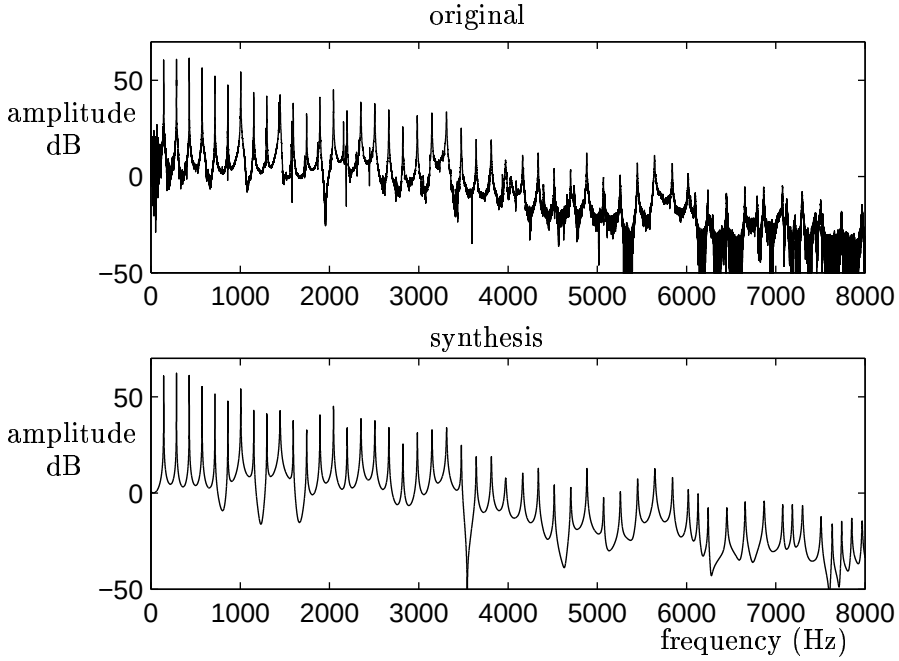

Figure 2.22: *Spectra of the original signal (top) and of the synthesis signal (bottom) for a single string model.*

only at pre-selected points on the string, whereas a finite difference scheme computes the entire string state in (sampled) physical form. This is not a drawback for sound synthesis applications, because only the behavior of the string at the bridge is of interest in most stringed instruments. Moreover, physically accurate outputs from additional points along the string are easily added to a digital-waveguide simulation at the price of one small digital filter each.

The digital waveguide, when computed in the frequency domain, possesses another important advantage. As we have seen, it is possible, through the use of an arbitrary filtering function F, to implement any amount of loss or dispersion at a given frequency. The form of the filter can be of arbitrary complexity, without any effect on the efficiency of the implementation. In the analogy between the PDE model and the digital waveguide model presented in Section 2.4, we were forced to choose a specific form for the modulus and phase of the filter F (Eq. (4.2)) and to make assumptions about values of the parameters for a real piano string. This analogy is then verified only for a realistic piano string. But we have shown in Section 2.6 that it is possible to determine the modulus and phase of this function from measured data, in such a way as to reproduce the piano sound perfectly, at least from the perceptual

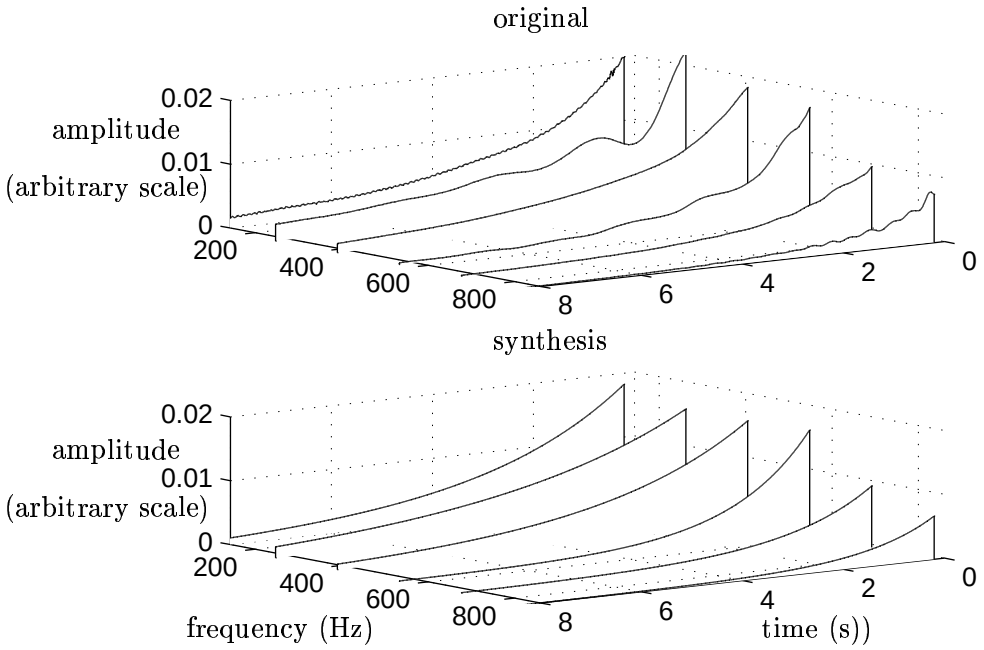

Figure 2.23: *Amplitude modulation laws for the first six partials, one string.*

point of view, even when multiple strings are struck.

For these reasons, we have chosen to make use of the digital waveguide model in order to simulate wave propagation in strings. It is worth mentioning that if the filter F is to be implemented in the time domain, much less freedom of the choice of modulus and phase will be allowed, at least if the implementation is to be efficient. It will become necessary to approximate the measured filter response by that of an FIR or IIR filter with a reasonable number of coefficients. The sounds produced will remain, however, of a good quality.

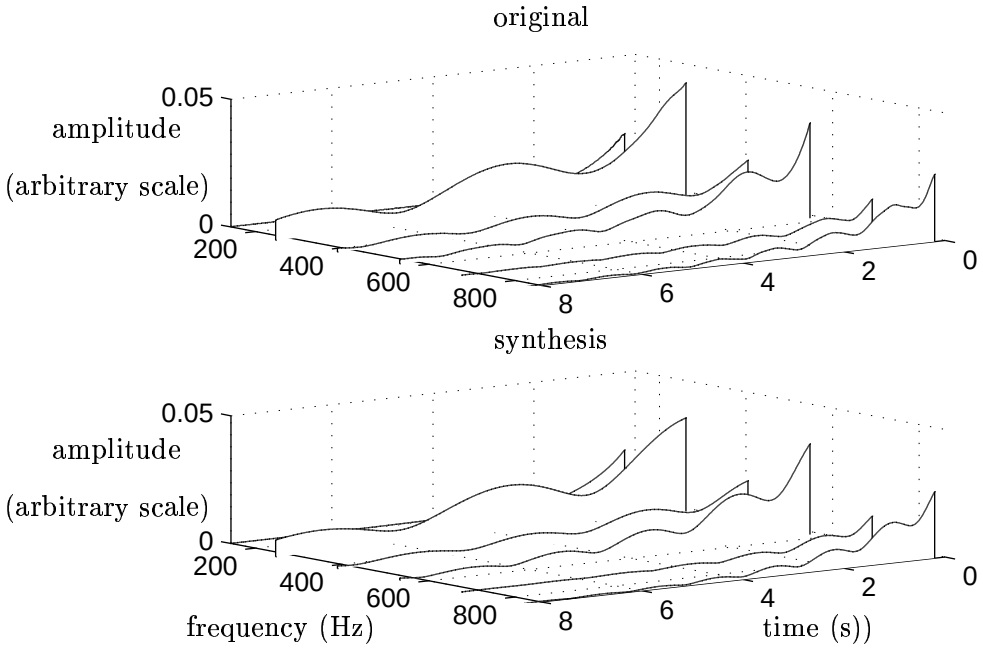

Figure 2.24: *Amplitude modulation laws for the first six partials, two strings.*

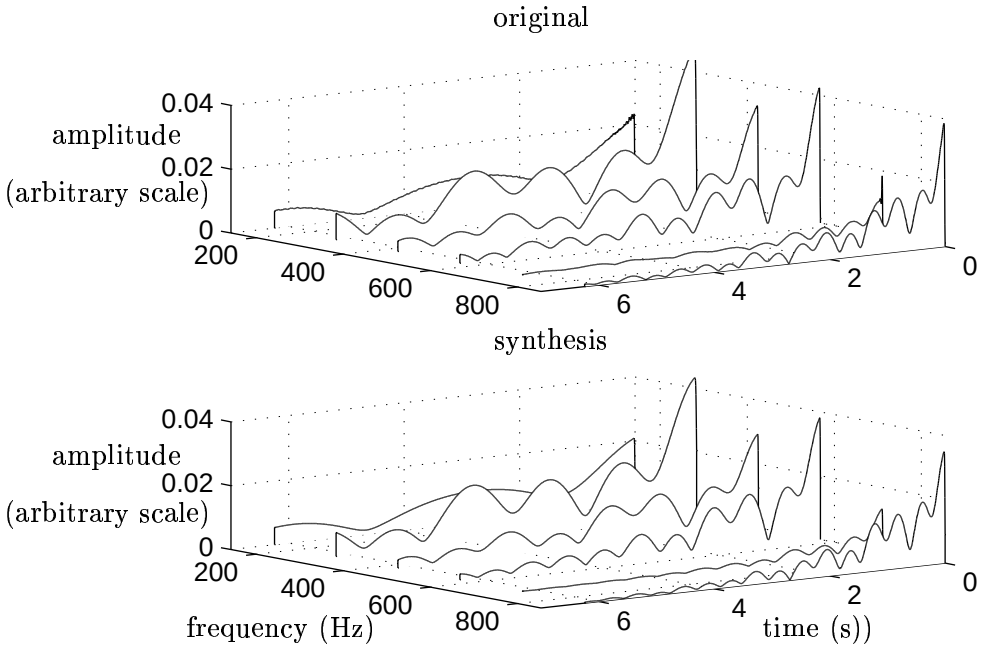

Figure 2.25: *Amplitude modulation laws for the first six partials, three strings.*

Chapter 3

The hammer-string interaction

3.1 Introductory remarks

In this chapter, we present several approaches to the problem of modeling the interaction between the piano string and hammer, as well as its effect on the timbre of the resulting sound. This interaction intervenes in a crucial way in the production of this timbre, giving a struck piano string a sound quite distinct from that of other keyboard instruments such as the harpsichord, and varying considerably even among pianos. Piano tuners place great importance on the hammer, and, as we discussed in Section 1.2.3, manually adjust the mechanical properties of these hammers in order to achieve a balanced sound over the entire range of the keyboard. The nonlinear behavior of the hammer felt is a key attribute, and will be incorporated into the models we will discuss.

First, we present a model which takes the form of a coupled set of differential equations, one describing the motion of the hammer, and the other that of the string. This system is solved numerically via finite differences, and optimization method of calibration of the model parameters is also discussed. We then approach the problem from the point of view of digital waveguides (used to model the string motion) and wave digital filters (to model the hammer), and show advantages of this method compared to the finite differences. Finally, we show that it is possible to simulate the hammer/string interaction (that is, the source signal to the digital waveguide), by subtractive synthesis, a signal-based method. This method, which is more distant from the physically motivated models, allows a resynthesis of measured sounds of a very high quality.

3.2 The hammer as a non linear oscillator

A linear hammer model has been described in Section 1.2.3. We concluded that the only way to take into account the modification of the piano tone with the dynamic is to use a non-linear hammer model. Ghosh [45] was one of the first to propose a non linear spring model of the felt [107], obeying the powerlaw:

$$f_H = K_H x^p \tag{3.1}$$

where x refers to the compression of the felt, K_H is the stiffness coefficient and p the stiffness exponent. Using the Ghosh's model, Hall and Askenfelt [57] measured the values of p for several hammers: p varies between 1.5 to 3.5. Suzuki [108] have approximated the force of the hammer felt using a sum of power laws:

$$f_H = K_1 x^2 + K_2 x^3 + K_3 x^4 \tag{3.2}$$

with K_1, K_2, K_3 the stiffness coefficients. This model was able to reproduce some of the measured phenomena but was not satisfactory since K_1 is negative and the force is consequently negative for small felt compression. Realistic numerical simulations have been obtained using the Ghosh's model and a lossless and stiffless string model by Boutillon [20]. Chaigne and Askenfelt [26] have compared experimental measurements and simulations using a lossy and stiff string. The first milliseconds of the two signals were well reproduced but the resulting synthesized sound corresponding to the displacement of one point of the string) was not perceptually satisfying.

For a more realistic model, some papers suggest an hysteretic law for the hammer felt. The felt is compressed and extended several times during the contact. As the relaxation is not instantaneous, the hardness of the felt increases during the contact. Boutillon [20] tried to model this phenomena using several values of p for the increasing and decreasing part of the spring stiffness characteristic. But using this model, the felt deformation tends to zero with the unloading of the force. Other experiments [116] show that the felt is still deformed after the force is removed. Stulov [107] proposed a mathematical model to take into account the hysteretic feature of the felt. Starting from a simple model of material with memory (obtained by replacing constant elastic parameters by time-dependent operators), he found the following relation governing the hammer force F as a function of the felt compression Δy

$$f_H = K_H x^p - \frac{\epsilon}{\tau} exp(-\frac{t}{\tau}) \int_0^t [x(\xi)]^p exp(\frac{\xi}{\tau}) d\xi \tag{3.3}$$

where ϵ is the hysteresis constant, and $\tau = t/\tau_0$ with τ_0 the relaxation constant. Good agreement with experiments was achieved. Giordano and Miller [49] have shown that this

model well reproduce the force exerted by a hammer on a sensor, in the static case. They have also mounted a very small accelerometer on the string, opposite to the hammer-string contact, and shown that this model is not able to simulate the force characteristic in "dynamic" conditions. Another model which was originally developed by Hunt and Crossley [63] to describe the interaction between two colliding objects has been used by Rochesso and Avanzini in [10] to model the hammer-string contact hysteresis. The losses are empirically introduced by adding a dashpot to the Ghosh's model. Authors show that the force-compression diagrams are closed to the ones of a real hammer. This model has the advantage to conserve energy. The force depends on the felt compression x and of the compression speed $\dot{x}$

$$f_H = K_H x^p (1 + \mu \frac{dx}{dt}) \tag{3.4}$$

where K_H is the hammer stiffness ($K_H > 0$) and μ the felt loss coefficient ($\mu > 0$).

3.3 Coupling a non-linear hammer and a PDE string model

We describe in this section the hammer and string model we have used to simulate the hammer-string interaction. It consists of a set of PDEs governing the string motion and the hammer displacement coupled by a power law describing the compression and losses of the felt as a function of the hammer and string displacement. The force law is proposed by Hunt and Crossley [63]. The string PDE equation corresponds to the one presented in Section 2.2.2.

$$\begin{aligned}
&\frac{\partial^2 y_s}{\partial t^2} = c^2 \frac{\partial^2 y_s}{\partial x^2} - \kappa^2 \frac{\partial^4 y_s}{\partial x^4} - 2b_1 \frac{\partial y_s}{\partial t} + 2b_2 \frac{\partial^3 y_s}{\partial x^2 \partial t} + g(x, x_0) f_H(y_H - y_s) \\
&M_H \frac{d^2 y_H}{dt^2} = -F_H(y_H - y_s) \\
&y_s(x,0) = y_{s0}(x), \ \frac{\partial y_s}{\partial t}(x,0) = v_{s0} \\
&y_s(0,t) = y_s(L,t) = 0, \ \frac{\partial^2 y_s}{\partial x^2}(0,t) = \frac{\partial^2 y_s}{\partial x^2}(L,t) = 0 \\
&y_H(0) = y_s(x_0,0), \ \frac{\partial y_H}{\partial t}(0) = v_{H0}
\end{aligned} \tag{3.6}$$

where

$$\begin{aligned}
F_H(u(t)) &= \begin{cases} K_H u^p (1 + \mu \frac{\partial u}{\partial t}) \text{ if } u > 0 \\ 0 \text{ if } u \leq 0 \end{cases}, \\
f_H(u(t)) &= F_H(u(t)) \left\{ \rho_L \int g(x, x_0) dx \right\}^{-1}
\end{aligned} \tag{3.7}$$

with y_s the string displacement, y_H the hammer displacement, c the wave speed, κ the stiffness coefficient, b_1 and b_2 the loss parameters, ρ_L the linear mass density of string, K_H the hammer stiffness and μ the felt loss coefficient. x_0 is the hammer position on the string, y_{s0} and v_{s0} are the initial string displacement and velocity and v_{H0} is the initial hammer velocity. $g(x, x_0)$ is a dimensionless spatial window. We took here $g(x, x_0)$ equal to one on one cell of the mesh.

As expected, this non-linear system does not admit any analytical solution. We will show in Section 3.4 that this model well reproduce the main phenomena involved in the hammer string interaction, namely the non-linear behavior with respect to the hammer velocity and the hysteresis of the hammer force.

3.4 Numerical solution using time and space finite differences

To solve the system Eq. (3.6), we use an alternative scheme to the explicit and implicit ones describe in 2.2.2. At first we reduce the time derivative order by introducing new variables: the string velocity v_s and the hammer velocity v_H.

$$v_s = \frac{\partial y_s}{\partial t} \qquad v_H = \frac{\partial y_H}{\partial t} \tag{3.8}$$

We obtain the following system of first order in time:

$$\begin{aligned}
&\frac{\partial y_s}{\partial t} = v_s \\
&\frac{\partial y_H}{\partial t} = v_H \\
&\frac{\partial v_s}{\partial t} = c^2 \frac{\partial^2 y_s}{\partial x^2} - \kappa^2 \frac{\partial^4 y_s}{\partial x^4} - 2b_1 v_s + 2b_2 \frac{\partial^2 v_s}{\partial x^2} + g(x, x_0) f_H(y_s - y_H, v_s - v_H) \\
&M_H \frac{dv_H}{dt} = -F_H(y_s - y_H, v_s - v_H) \\
&y_s(x, 0) = y_{s0}(x), \; v_s(x, 0) = v_{s0} \\
&y_s(0, t) = y_s(L, t) = 0, \; \frac{\partial^2 y_s}{\partial x^2}(0, t) = \frac{\partial^2 y_s}{\partial x^2}(L, t) = 0 \\
&y_H(0) = y_s(x_0, 0), \; v_H(0) = v_{H0}
\end{aligned} \tag{3.10}$$

System (3.10) may be written

$$\frac{\partial X}{\partial t} = \mathcal{A}(X) \tag{3.11}$$

where $X = {}^t(y_s, v_s, y_H, v_H)$ and $\mathcal{A}$ the spatial differential operator corresponding to the right-hand side of (3.10). We discretize (3.11) in time using a θ-scheme:

$$\frac{X^{n+1} - X^n}{\Delta t} = (1 - \theta)\mathcal{A}(X^{n+1}) + \theta \mathcal{A}(X^n) \tag{3.12}$$

By choosing $\theta \neq 1$, an implicit scheme is obtained. $\theta = 1$ gives the explicit scheme. We choose $\theta = 1/2$ to obtain the Crank Nicholson scheme which have good stability properties. Finally, spaces derivatives are discretized as in section 2.2.2.

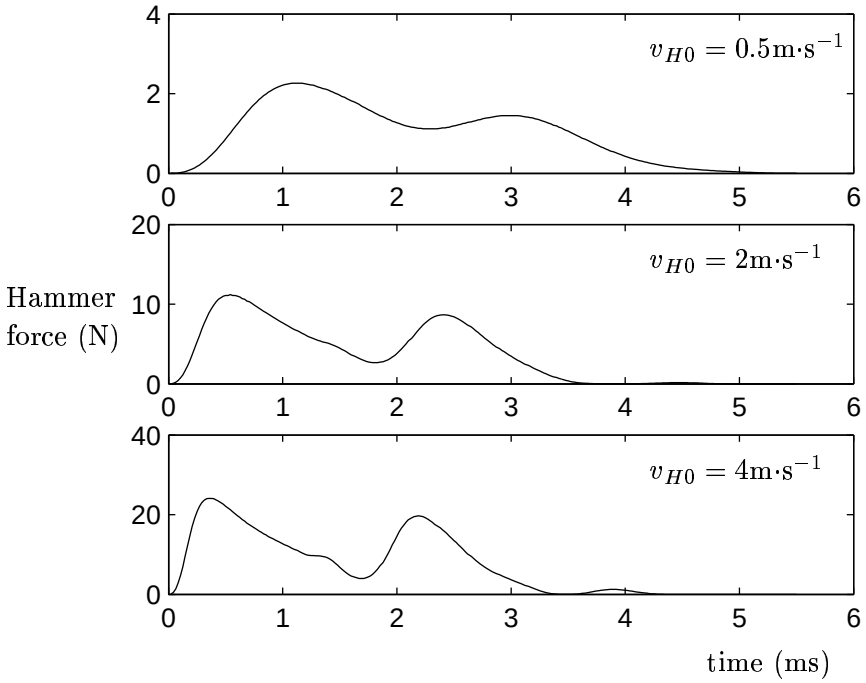

Figure 3.1: *Hammer force curves as a function of time, for a note C4 and three different hammer velocities. The hammer parameters and the striking positions are as given in [27].*

The hammer force behaves in accordance with what is forseen by the theory as shown on Figure 3.1. The hammer-string contact has a duration of approximatively 4 ms and decreases with the increase of the initial hammer velocity. The reflections of the waves propagating between the hammer and the agraffe can be seen on the hammer force shape. Figure 3.2 shows the force as a function of the compression of the felt. We can clearly see an hysteretic behavior (due to the term $K_H\mu(v_s - v_H)$ of the non linear force law) describing losses in the felt.

The displacement is computed at a location 9/10 of the string length. Its spectral modulus is plotted on Figure 3.3 for different initial hammer velocities. We clearly see a modification of the spectral shape, namely an increase of the amplitude of the high frequency partials (corresponding to an increase of the brightness of the sound).

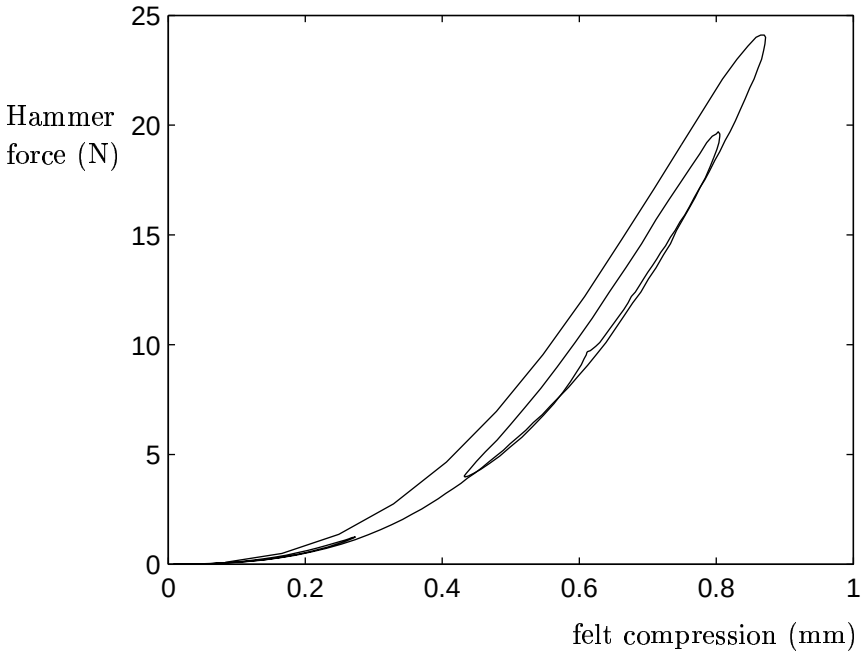

Figure 3.2: *Hammer force curve as a function of the felt compression, for a note C4 and a hammer velocity of 4* $m{\cdot}s^{-1}$. *The hammer parameters and the striking positions are as given in [27],* $\mu = 0.1$.

3.5 Calibration by optimization method

The estimation of the physical parameters of such a PDE model as described in Section 3.3 is not easy. Some of those parameters can be quite accurately measured on the experimental setup itself (the initial hammer velocity v_{H0}, the wave speed c ...) with a relative error. Other parameters, like the hammer stiffness K_H or the felt loss coefficient μ are very difficult to extract from measurements, especially under dynamic conditions. As a consequence, the synthesized signal should be different from the measured one. We propose here a method of estimating the parameters of the PDE model so that the resulting signal sounds like the original one. We use a classical optimization technique to calibrate the parameters of the system, referring to an experimental data set. To do so, gradients of the hammer and the string behaviors, with respect to physical parameters, are described as the solution of linear coupled systems of PDEs coming from Eq. (3.6).

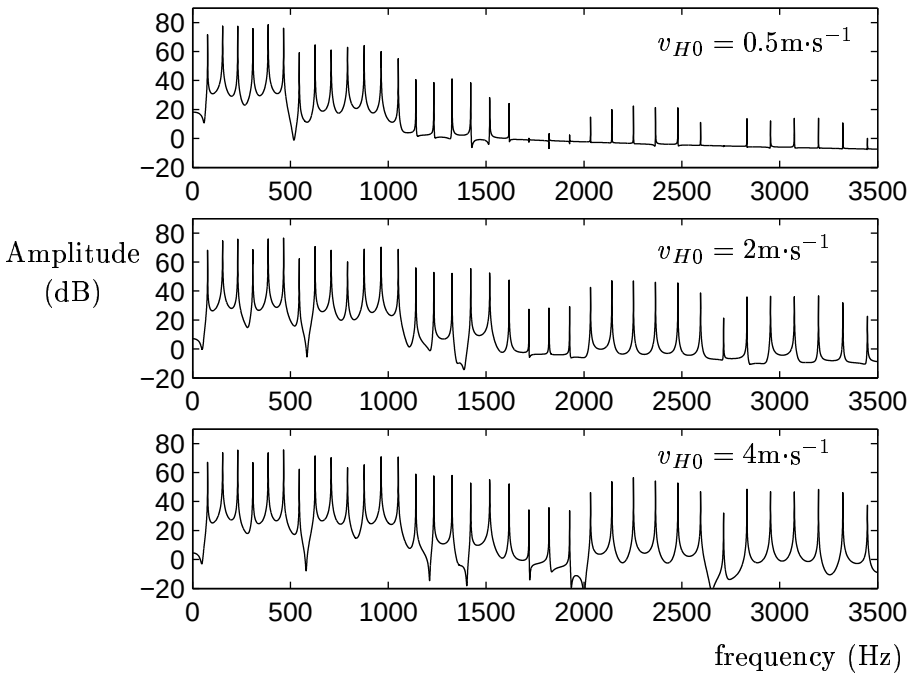

Figure 3.3: *Three spectra (of three velocity signals, for a sampling frequency of 24000 Hz and 2^{18} points) for three different hammer velocities for a note C4.*

3.5.1 Physical Parameter Derivatives

To get the gradient of the solution, the derivatives of the hammer and string displacements with respect to the physical parameters have to be calculated. (y_s, y_H) are time and space depending, but they may be seen as functions of physical parameters:

$$\begin{cases} y_s(x,t,(z_j)_{j=1,\cdots,9}) = y_s(x,t,(c,\kappa,b_1,b_2,K_H,\mu,M_H,p,v_{H0})) \\ y_H(x,t,(z_j)_{j=1,\cdots,9}) = y_H(x,t,(c,\kappa,b_1,b_2,K_H,\mu,M_H,p,v_{H0})) \end{cases} \tag{3.13}$$

We do not mention y_{s0}, v_{s0}, x_0, since we will not calculate the derivatives regarding those parameters. We suppose that $y_{s0} = 0$, $v_{s0} = 0$ and that x_0 can be precisely measured on the experimental setup.

We will only give the detailed calculation to obtain derivatives with respect to the parameter v_{H0}, but the same procedure could be applied to any of the physical parameters. We first write that the set $(y_s(v_{H0}), y_H(v_{H0}))$ is solution of the system (3.6) for a value v_{H0} of the initial hammer velocity and that the set $(y_s(v_{H0}+h), y_H(v_{H0}+h))$ is solution of the system (3.6) for a value $v_{H0}+h$ of the same parameter. By doing term by term the difference between the two corresponding systems, dividing by h and passing to the limit when h tends to zero, we obtain a linear system of equations, governing the derivatives (y'_s, y'_H)

$$
\begin{aligned}
&\frac{\partial^2 y'_s}{\partial t^2} = c^2\frac{\partial^2 y'_s}{\partial x^2} - \kappa^2\frac{\partial^4 y'_s}{\partial x^4} - 2b_1\frac{\partial y'_s}{\partial t} + 2b_2\frac{\partial^3 y'_s}{\partial x^2 \partial t} \\
&+\frac{g(x,x_0)}{\rho_L \int g(x,x_0)dx} K_H p(y'_H - y'_s)(y_H - y_s)^{p-1}(1+\mu\frac{\partial(y_H - y_s)}{\partial t}) \\
&+\frac{g(x,x_0)}{\rho_L \int g(x,x_0)dx} K_H \mu(y_H - y_s)^{\alpha}\frac{\partial(y'_H - y'_s)}{\partial t} \\
&y'_s(x,0) = 0,\ \frac{\partial y'_s}{\partial t}(x,0) = 0,\ y'_s(0,t) = y'_s(L,t) = 0,\ \frac{\partial^2 y'_s}{\partial x^2}(0,t) = \frac{\partial^2 y'_s}{\partial x^2}(L,t) = 0 \\
&M_H\frac{\partial^2 y'_H}{\partial t^2} = -K_H p(y'_H - y'_s)(y_H - y_s)^{p-1}(1+\mu\frac{\partial(y_H - y_s)}{\partial t}) - K_H\mu(y_H - y_s)^{\alpha}\frac{\partial(y'_H - y'_s)}{\partial t} \\
&y'_H(0) = 0,\ \frac{\partial y'_H}{\partial t}(0) = 1
\end{aligned}
\tag{3.14}
$$

where $u' = \dfrac{\partial u}{\partial v_{H0}}$.

We may generalize this procedure to any of the parameters. The derivatives $(y_s^{(j)}, y_h^{(j)})$ of (y_s, y_H) are then obtained by solving the following linear system of equations (supposing that (y_s, y_h) are known):

$$
\begin{aligned}
&\frac{\partial^2 y_s^{(j)}}{\partial t^2} = c^2\frac{\partial^2 y_s^{(j)}}{\partial x^2} - \kappa^2\frac{\partial^4 y_s^{(j)}}{\partial x^4} - 2b_1\frac{\partial y_s^{(j)}}{\partial t} + 2b_2\frac{\partial^3 y_s^{(j)}}{\partial x^2\partial t} + g(x,x_0)[f_H(y_H - y_s)]^{(j)} \\
&+(c^2)^{(j)}\frac{\partial^2 y_s}{\partial x^2} - (\epsilon c^2 L^2)^{(j)}\frac{\partial^4 y_s}{\partial x^4} - 2(b_1)^{(j)}\frac{\partial y_s}{\partial t} + 2(b_2)^{(j)}\frac{\partial^3 y_s}{\partial x_2 \partial t} \\
&y_s^{(j)}(x,0) = y_0^{(j)}(x),\ \frac{\partial y_s^{(j)}}{\partial t}(x,0) = v_0^{(j)},\ y_s^{(j)}(0,t) = y_s^{(j)}(L,t) = 0,\ \frac{\partial^2 y_s^{(j)}}{\partial x^2}(0,t) = \frac{\partial^2 y_s^{(j)}}{\partial x^2}(L,t) = 0 \\
&M_H\frac{\partial^2 y_H^{(j)}}{\partial t^2} = -g(x,x_0)[f_H(y_H(t) - y_s(x_0,t))]^{(j)} - M_H^{(j)}\frac{\partial^2 y_H}{\partial t^2} \\
&y_H^{(j)}(0) = y_s^{(j)}(x_0,0),\ \frac{\partial y_H^{(j)}}{\partial t}(0) = v_i^j
\end{aligned}
\tag{3.15}
$$

where $u^{(j)} = \dfrac{\partial u}{\partial z_j}$.

3.5.2 Optimization Process

We describe here the optimization tools and how they can be applied to our problem. We chose simple and classical algorithms known for their robustness. We used a simple gradient method with an optimal step. Consider a measured signal y_s^{obj} collected on the experimental setup described in Section 2.6.2 for which system Eq. (3.6) is assumed to be a good approximation. We recall that the measured signal corresponds to the displacement of one point of the string (measured using the laser vibrometer). The string is struck by the hammer link to the electronic key. The hammer velocity is measured just before the contact with the string

by the optical sensor.

We consider the following problem: is it possible to determine a set of physical parameters $(z_j)_{j=1,\cdots,9}$ so that the system (Eq. (3.6)) would give a solution as close as possible to the experimental signal? More precisely, considering a quadratic criterion $\mathcal{C}$ measuring the error between y_s and y_s^{obj}, $\mathcal{C}(y_s, y_s^{obj}) = \tilde{\mathcal{C}}(z_j)$, for which set $(z_j)_{j=1,\cdots,n}$, $\mathcal{C}$ is minimum? The knowledge of (y_s, y_H) obtained from Eq. (3.6) allows the computation of the derivatives $(y_s^{(j)}, y_H^{(j)})_{j=1,\cdots,n}$ and the gradient of solution (y_s, y_H) with respect to the physical parameters. Then, we construct an algorithm which insure the decreasing of the criterion $\tilde{\mathcal{C}}$ (we do not discuss here the uniqueness of such a minimum of $\tilde{\mathcal{C}}$). We start with initial parameters values measured on the experimental setup, or, for the hammer, found in the literature.

ϵ given; $(z)_{i=1,n} = (z_0)_{i=1,n}$ given
while $\mathcal{C} > \epsilon$ **do**
  Compute y_s solution of (Eq. (3.6)) with parameter set $(z)_{i=1,n}$
  for j=1,n **do**
    Compute $y_s^{(j)}$ solution of (Eq. (3.15))
    Compute $\tilde{\mathcal{C}}^{(j)}$ using $y_s^{(j)}$ { j^{th} component of $\nabla\tilde{\mathcal{C}}$}
  end for
  chose h such that $\tilde{\mathcal{C}} + \nabla\tilde{\mathcal{C}} \cdot h \leq \tilde{\mathcal{C}}$
  for j=1,n **do**
    $z_j \Leftarrow z_j + h_j$
  end for
end while

3.5.3 Choice of criterion

"Global" quadratic criterion

The optimization criteria have to be chosen with care to take into account perceptual phenomena. A criterion based on the quadratic error obtained by comparing the measured and computed temporal signals is too accurate. Indeed the physical model does not exactly describe the experiment (for example, beats due to the coupling between the two polarizations of the string, impedance of the bridge). Therefore, the waveforms of the original and synthetic signals may not be identical (slight phase modifications don't alter the perception of the sound). For this reason, we considered a criterion based on the spectral power density (without considering the phase) of the signal. More precisely, we ought to minimize the quadratic error $\mathcal{C}$ between the modulus of the Fourier transform Y_S^c of the string displacement y_S^c, the

solution of Eq. (3.6) and the modulus of the Fourier transform S^m of the measured signal s:

$$\mathcal{C} = \frac{1}{2}\int (|Y_S^c|^2 - |S^m|^2)^2 d\omega \tag{3.16}$$

One may easily compute the gradient $\nabla_z \tilde{\mathcal{C}}$ of the criterion $\tilde{\mathcal{C}}$ (∇_z denotes the gradient with respect to the physical parameters):

$$\nabla_z \tilde{\mathcal{C}} = 2\int (|Y_S^c|^2 - |S^m|^2)|Y_S^c|\nabla_z Y_S^c d\omega$$

Due to the linearity of the Fourier transform, one may deduce $\nabla_z Y_S^c$ from ∇y_s. We do not consider the phase of the Fourier transform in the minimization criterion. This phase contains information about the physical parameters. But this information is also contained in the spectral modulus. For example, the parameter κ, related to the dispersion of waves and consequently to the phase of the signal, is also correlated to the frequency of each partial of the spectrum. So we can estimate it by considering only the spectrum modulus. Since the most important part of the energy is localized in the low frequency region of the spectrum, this "global" criterion (power spectral density) hardly takes into account low frequency components, leading to a lack in brightness of the sound. As high frequency components are perceptually of great importance (namely for the first milliseconds), the attack of the sound obtained using such an optimization criterion appeared then to be too smooth compared to the original one.

Tristimuli based quadratic criterion

In order to overcome this high frequency loss, we propose here a quadratic criterion constructed from an adapted form of the so-called tristimuli [80] which was previously used for estimation purposes by Ystad st al.[118]. This approach consists of separating the spectral range into three spectral bands. The first band is centered on the fundamental frequency of the signal, while the second and the third bands respectively contain the partials two to ten and all the components above the tenth partial. This spectral split has several advantages:

- tracking of the fundamental ensures good estimation of the geometric and static parameters. This fundamental component is of importance for the sound since it is highly correlated with pitch. We estimate the wave speed c and the loss parameter b_1 over this band.

- components two to ten contain a large part of the energy and determine the damping behavior of the sound. Their frequencies lead to the sensation of inharmonicity. This band is used to estimate the stiffness coefficient κ.

- high frequency components are of great importance since they can dramatically change the timbre of the sound, even though their energy is low compared to the ones of the first

components. These high frequency components are related to the perception of brightness and mainly correlated with the velocity and with the stiffness of the hammer. They characterize the dynamics of the playing. We use this frequency band to estimate the remaining coefficients $(b_2, K_H, \mu, M_H, p, v_{H0})$.

3.5.4 Results

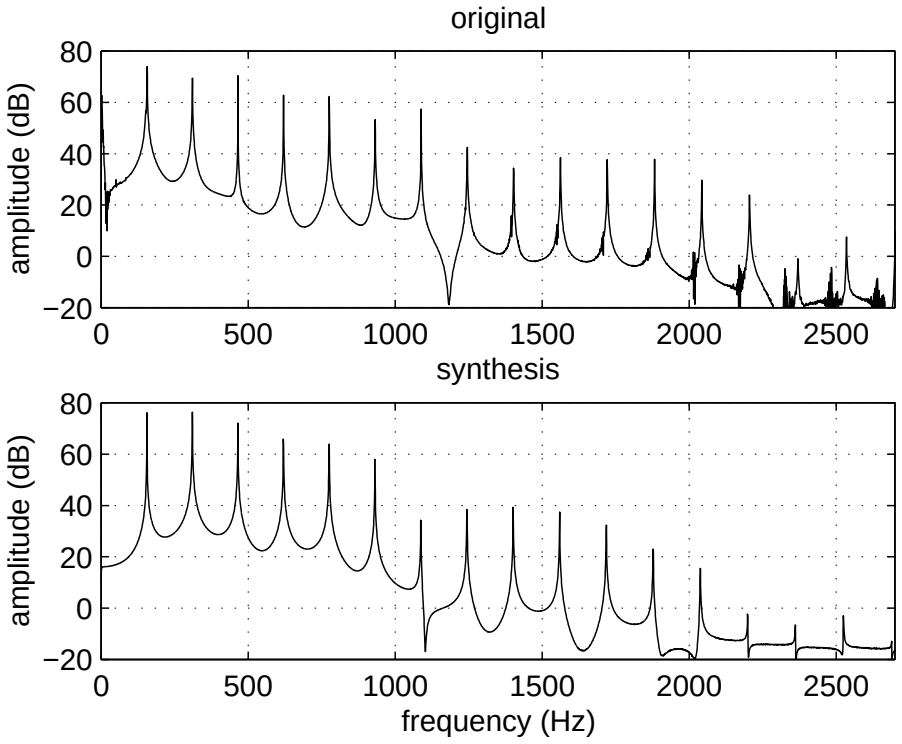

Figure 3.4: *Spectra of original (top) and synthesized (bottom) signals*

Starting from parameters measured on the experimental setup or, for the hammer, found in the literature the optimization process leads to a new set of physical parameters. Those parameters are then used to compute a synthesized signal (with Eq. (3.6)), closer to the original one from a perceptual point of view. In Figure 3.4, the spectral modulus of the original and synthesized signals are plotted. The spectral moduli are similar, but not exactly identical. There is namely a coupling between the horizontal and vertical polarizations of the string, leading to doublets of components. Those doublets are not reproduced by the model (since we only consider one polarization). In Figure 3.5.4 are plotted the waveshape for the real sound and the synthesized one. One may see the similarity between the signals.

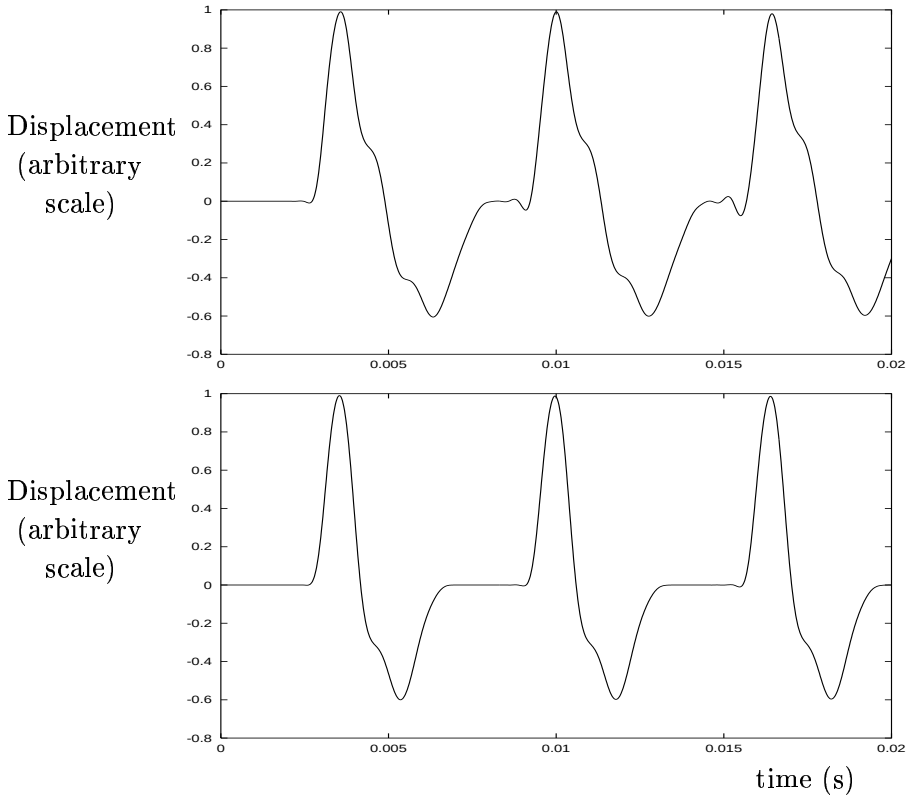

Figure 3.5: *Waveform of the displacement; top, original, bottom, syntesized using the optimized parameters.*

The quality of the resulting sound increases using the optimized parameters. But the tone still lacks realism. One explanation for this lack is that the hammer model is too simplistic to accurately simulate the non-linear behavior.

3.6 From mechanics to electricity

The hammer-string system Eq (3.6) and Eq (3.7) can be seen as a mass and a non-linear spring for the hammer and as a set of mass-spring oscillators for the string. Figure 3.6 shows such a model. M_H is the hammer mass and K_H the stiffness of the felt. The arrow on the hammer spring indicates its non-linear behaviour. Thanks to the analogy voltage/force and current/velocity, we obtain the corresponding electrical scheme of Figure 3.7. v_H, v_d, v_L and v_R are the velocities of the hammer mass, the hammer non-linear spring, the wave coming from the agraffe side and the wave coming from the bridge side of the string, respectively. Z_0 is the characteristic wave impedance of the string at the contact point ($Z_0 = \sqrt{\rho T}$ with the linear mass of the string and T the string tension). f_M, f_d, f_L and f_R are the forces exerted

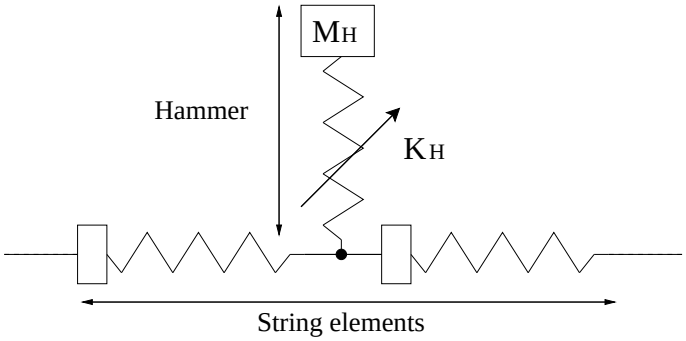

Figure 3.6: *Mechanical model of a non-linear hammer coupled to a string.*

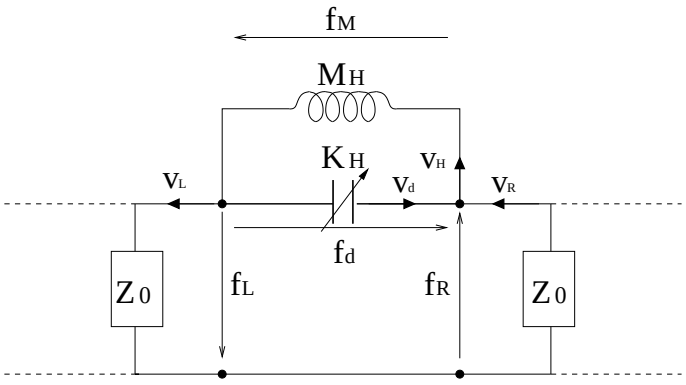

Figure 3.7: *Electrical model of a non-linear hammer coupled to a string.*

by the hammer mass, the non-linear hammer spring, the wave coming form the agraffe side and the wave coming from the bridge side, respectively.

We will now show that this representation corresponds to the system of partial differential equations Eq. (3.6) and Eq. (3.7) at the hammer-string contact point.

The voltage of the inductor is given by

$$f_M = M_H \dot{v}_H \tag{3.17}$$

And the current in the capacitor is

$$v_d = \frac{1}{K_H} \dot{f}_d \tag{3.18}$$

To take the non-linearity of the spring into account, we have to calculate an "equivalent" stiffness K_e which will be function of K_H and of the force f_d. This will be done in Section A,

with the use of digital wave filters. In this part we will use for seek of simplicity the relation

$$f_d = K_H y_d^p \tag{3.19}$$

where $\dot{y}_d = v_d$ and p is the hammer stiffness exponent.

The string is here modeled as a distributed structure such as a digital waveguide. The discrete velocity waves are propagating in the digital waveguide and the relation between the force and the velocity at the contact location is given by

$$\begin{cases} f_L = Z_0 v_L \\ f_R = Z_0 v_R \end{cases} \tag{3.20}$$

Applying Kirchoff's voltage laws (for a serial connection), we obtain for the connection between the hammer and the two parts of the string

$$\begin{cases} f_L + f_R - f_d = 0 \\ v_L = v_R = v_S \end{cases} \tag{3.21}$$

with v_S the velocity of the string at the contact location. Applying Kirchoff's current laws (for a parallel connection), we obtain for the junction between the hammer spring and hammer mass

$$\begin{cases} v_d + v_S - v_H = 0 \\ f_H = f_d = -f_M \end{cases} \tag{3.22}$$

where f_H is the hammer force.

Using Eqs. (3.17)- (3.22) we obtain the system

$$\begin{cases} f_H = 2Z_0 v_S \\ f_H = -M_H \dot{v}_H \\ f_H = K_H y_d^p \\ \dot{y}_d = v_H - v_S \end{cases} \tag{3.23}$$

This system of equation corresponds to the partial differential equations system Eqs. (3.6) and (3.7), except that the delay, loss and dispersion phenomena involved in the wave propagation in the string are now characterized using the corresponding impedance at the contact location. This system can also be written

$$\begin{cases} \dot{y}_d = v_H - \dfrac{1}{2Z_0} K_H y_d^p \\ \dot{v}_H = -\dfrac{1}{M_H} K_H y_d^p \end{cases} \tag{3.24}$$

We here exhibit the "delay free loop": to solve the first equation, we need the hammer velocity v_H, so we have to solve the second equation. But to solve the second equation, we need the felt compression y_d, so we have to solve the first equation. The K-method and the digital wave filters allow to cut this simultaneous dependancy.

3.7 Numerical solving using Digital Wave Filters

The numerical solving of a non-linear partial differential equation system leads to discrete time models exhibiting non computable loops. They are called *delay-free loops* since they involve instantaneous dependancy between inputs and outputs of the elements of the system. To solve this computability problem, many authors proposed to introduce an extra delay [44]. But the numerical error can leads to aberrant behaviour at low sampling rates and decrease the quality of the synthesized sound.

Implicit methods such as the K-method (developed by Borin et al. [19]) and the wave digital filters [37] [17] can take care of this computability problem. They allow to eliminate the delay-free loops in discrete time models of non-linear system. We describe in the next sections both methods.

Digital Wave Filters (DWFs) have been described in details by Stefan Bilbao in his dissertation [17], under the supervision of Julius O. Smith. We worked with both researchers to apply this simulation method to the hammer string interaction. The results are published in [18].

Since this model was not used in the final synthesis model, we will not detail the DWFs theory and calculation in this document. For more details, the reader may refer to the appendix A, thesis [17] and paper [18].

3.7.1 K-method

We applied here the K-method [19] to numerically solve the system of Eq. (3.23). We first have to transform the equations of the mechanical system to exhibits all non-computable loops involving non-linear maps (this is what we have done in Section 3.6). Then, the instantaneous dependencies are cut using geometrical transformations on the non-linearities. The formalism is straightforward, compared to the one of the digital wave filters. Nevertheless, wave filters allow an energy measure which guarantees numerical stability. Such a guarantee seems to be difficult to obtain using the K-method. At least, it has not been exhibited by its authors [19].

We first rewrite the system Eq. (3.23)

$$\begin{cases} \dot{y}_d = v_H - \frac{1}{2Z_0} f_H \\ \dot{v}_H = -\frac{1}{M_H} f_H \\ f_H = K_H y_d^p \end{cases} \tag{3.25}$$

We discretize the system by applying the Laplace transform to the first two equations

$$\begin{cases} s y_d(s) = v_H(s) - \frac{1}{2Z_0} f_H(s) \\ s v_H(s) = -\frac{1}{M_H} f_H(s) \end{cases} \tag{3.26}$$

We then applied a frequency mapping, the bilinear transform

$$s = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}}$$

where $z^{-1} = e^{-sT}$ which is interpreted as a unit delay of duration T. This mapping is used since it implies that stable, causal transfer function in s will be mapped to stable causal transfer functions in the discrete variable z^{-1} [17]. We obtain the discretize system

$$\begin{cases} y_d(n) = \frac{T}{2}\left(v_H(n) - \frac{1}{2Z_0} f_H(n)\right) + \frac{T}{2}\left(v_H(n-1) - \frac{1}{2Z_0} f_H(n-1)\right) + y_d(n-1) \\ v_H(n) = -\frac{T}{2M_H} f_H(n) - \frac{T}{2M_H} f_H(n-1) + v_H(n-1) \\ f_H(n) = K_H (y_d)^p(n)) \end{cases} \tag{3.27}$$

From the system of Eq. (3.27), we obtain the relation

$$f_H(n)\left(\frac{1}{2Z_0} - \frac{T}{2M}\right) + f_H^{1/p}(n) \frac{2}{T} K_H^{-1/p} = C(n-1) \tag{3.28}$$

$$C(n-1) = f_H(n-1)\left(-\frac{T}{2M_H} - \frac{1}{2Z_0}\right) + \frac{2}{T} y_d(n-1) + 2 v_H(n-1) \tag{3.29}$$

At the instant n, the term $C(n-1)$ is known. We then have to solve the non-linear function of the force f_H. This can be done using iterative procedures such as the Newton-Raphson's method. For more efficiency, Borin et al. [19] have precomputed solutions and stored it in a table.

3.7.2 Theory of wave digital filters

Concepts and formulae presented in this section are obtained from Stefan Bilbao thesis [17] and a co-signed article [18]. The reader can find more details in each of these documents.

Wave digital filter and numerical stability

The digital wave filter method [37] aims, like the K-method, at solving computability problems due to delay free loops. They are eliminated by changing variables of the system and by choosing, when possible, appropriates impedances for the scattering junctions between elements.

It is easy to surmise that regardless of any non-linearity in the hammer (typically the felt stiffness), the total energy, kinetic and potential, of the moving string and the hammer can never be greater than the energy provided initially by the hammer - in other words, the combined system is passive. For the sake of numerical stability, it is useful to have an energy measure (or Lyapunov function) for the discrete system, which mirrors the energetic behavior of the modeled physical system. Such a measure is rather difficult to obtain from time domain finite difference formulations (though techniques such as the energy method [53] might conceivably be employed). A scattering representation, in which the discretization is rewritten in terms of wave variables which are scattered at junctions, can be used to give an easily verifiable numerical stability condition. In the purely lossless case, all operations on the signals in the network are either shifts or scattering operations (orthogonal matrix multiplies) at each step in the recursion; the squared sum of all the signals in the network remains constant. Loss can be introduced simply by adding resistive drains at the scattering junctions or the terminating filters, and renders the system more generally passive.

Wave variables

In order to arrive at a scattering formulation, it is necessary to introduce wave variables [36, 37]. Due to the nonlinearities to be modeled, we make use of power-normalized wave variables. We write

$$\underline{a} = \frac{f + Rv}{2\sqrt{R}} \qquad \underline{b} = \frac{f - Rv}{2\sqrt{R}} \tag{3.30}$$

for an arbitrary positive number R, known as the port resistance.

Wave digital capacitor and inductor

We apply this change of variables to the case of a capacitor (a spring) and of an inductor (a mass). Let's consider the relation between the force f exerted by a spring and its compression velocity v.

$$v = \frac{1}{K}\dot{f}$$

with K the spring stiffness. By discretizing this equation using the bilinear transform in the linear case or the trapezoid rule of numerical integration [36] [37] [88] in the non-linear case,

we get the relation (see [17])

$$\underline{b}(n) = \underline{a}(n-1)$$

which is the standard form for a wave digital capacitor (or its mechanical counterpart, the spring) with a port resistance $R = TK/2$. Thus, the output $\underline{b}$ at the instant n, only depends on the input $\underline{a}$ at the instant $n-1$. Applying the same procedure to a mass M, we get the relation

$$\underline{b}(n) = -\underline{a}(n-1)$$

with a port resistance $R = 2M/T$.

Scattering junctions

In order to connect wave digital one-ports such as the spring and mass mentioned previously, it is necessary to make use of scattering junctions (or adaptors, as they are called in wave digital filter terminology). The equations for a scattering junction are simply restatements of Kirchhoff's parallel and series connection rules in terms of wave variables (see [17] for further details).

For a parallel combination of N one-ports, defined in terms of N velocities v_i and N forces f_i , $i = 1,,N$, where power-normalized wave variables a_i and b_i , $i = 1,,N$ are defined, as per Eq. (3.30). The resulting scattering equations are

$$\underline{b}_i = -\underline{a}_i + \frac{1}{\sqrt{R_i}} f_J \tag{3.31}$$

where the junction force f_J (common to all one-ports) is given by

$$f_J = \frac{2}{\sum_j \frac{1}{R_j}} \sum_j \frac{1}{\sqrt{R_j}} \underline{a}_j \tag{3.32}$$

The operation Eq. (3.31) produces N outputs $\underline{b}_i$ from N inputs $\underline{a}_i$, $i = 1,,N$, and can be viewed as an N by N matrix multiplication. It is possible to show that as long as the port resistances R_i are chosen positive, the matrix is orthogonal, and thus preserves an L^2 measure of the input signals through the output, i.e.,

$$\sum_{i=1}^{N} \underline{a}_i^2 = \sum_{i=1}^{N} \underline{b}_i^2 \tag{3.33}$$

This energy-preserving property is at the heart of the numerical stability offered by a scattering approach. Even for elements with a nonlinear dependence of the R_i on the state variables

(as for a nonlinear spring as we will discussed later), energy is still preserved.

Similarly, series connection are given by

$$\underline{b}_i = \underline{a}_i - \sqrt{R_i} v_J \tag{3.34}$$

where the junction velocity v_J (common to all one- ports) is given by

$$v_J = \frac{2}{\sum_j R_j} \sum_j \sqrt{R_j} \underline{a}_j \tag{3.35}$$

This scattering operation also corresponds to an orthogonal matrix multiplication.

Elimination of delay-free loops

A special case of the scattering junction occurs when one of the port resistances remains unspecified. For instance, if we choose, at port q,

$$R_q = \sum_{i=q} R_i \tag{3.36}$$

for the series junctions, the scattering equations Eqs. (3.34) above simplify to

$$\underline{b}_q = -\frac{1}{\sqrt{R_q}} \sum_{i=q} \sqrt{R_i} \underline{a}_i \tag{3.37}$$

Thus the output wave $\underline{b}_q$ at port q does not depend on the input wave $\underline{a}_q$. In other words, the port is reflection-free [36]. This special choice of a free port resistance is necessary in order to avoid delay-free loops, as we shall see in Section 3.7.3. The same operation can be done for parallel connections (see [17])

3.7.3 Scattering model for the hammer-string interaction

We describe in this section the scattering model for the hammer-string interaction and its connection to a digital waveguide model. The design of these model has been made in collaboration with Stefan Bilbao. Parts of this paragraphs can also be found in [18].

In terms of wave variables and scattering junctions, the hammer model defined by the scheme of Figure 3.7 corresponds to the scheme of Figure 3.8.

The upper parallel junction, when disconnected from the lower series junction, is an os-

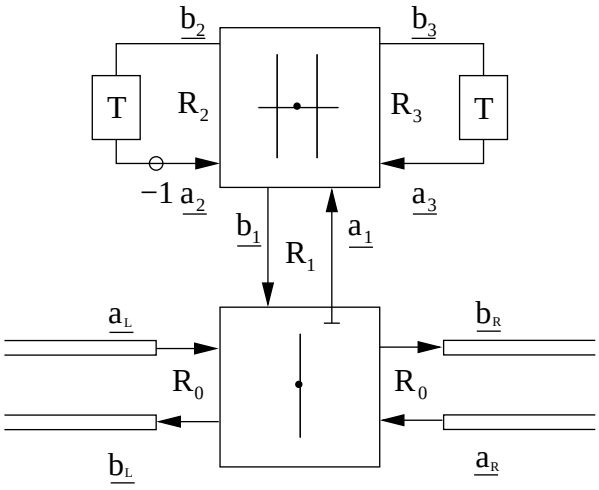

Figure 3.8: *A parallel connection of a mass and a nonlinear spring, top, representing a wave digital hammer, is connected in series with a digital waveguide, bottom.*

cillator (nonlinear), corresponding to the free vibration of a parallel mass-spring system (the hammer). Referring to Figure 3.8, the port resistances are given by

$$R_2 = 2M_H/T \qquad R_3 = TK_e/2 \tag{3.38}$$

where K_e is an equivalent stiffness, allowing to take into account the non-linear behavior of the felt. The expression of K_e and the detailed calculation are given in Appendix A.

The expression of the output normalized wave variables for the two wave digital one ports are

$$\underline{a_2} = -\underline{b_2} \tag{3.39}$$

$$\underline{a_3} = \underline{b_3} \tag{3.40}$$

Using the equations for the scattering junction, we get the relations

$$\underline{b_1} = -\underline{a_1} + \sqrt{\frac{1}{R_1}} f_H \tag{3.41}$$

$$\underline{b_2} = -\underline{a_2} + \sqrt{\frac{1}{R_2}} f_H \tag{3.42}$$

$$\underline{b_3} = -\underline{a_3} + \sqrt{\frac{1}{R_3}} f_H \tag{3.43}$$

with

$$f_H = \frac{2}{\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}}\left(\sqrt{\frac{1}{R_1}}\underline{a_1}+\sqrt{\frac{1}{R_2}}\underline{a_2}+\sqrt{\frac{1}{R_3}}\underline{a_3}\right) \tag{3.44}$$

for the parallel connection and

$$\underline{b_L} = \underline{a_L} - \sqrt{R_0}v_s \tag{3.45}$$

$$\underline{b_R} = \underline{a_R} - \sqrt{R_0}v_s \tag{3.46}$$

$$\underline{a_1} = \underline{b_1} - \sqrt{R_1}v_s \tag{3.47}$$

with R_0 the caracteristic impedance of the string and

$$v_s = \frac{2}{2R_0+R_1}(\sqrt{R_0}\underline{a_L}+\sqrt{R_0}\underline{a_R}+\sqrt{R_1}\underline{b_1}) \tag{3.48}$$

Connecting the series junction to the parallel junction poses a problem, due to the introduction of a delay-free loop; $\underline{a_1}$ is calculated from $\underline{b_1}$ and vice-versa. In order to surmount this difficulty, we note that the port resistance R_1 of the upper port of the series junction is left unspecified; we thus can set it to $2R_0$, as per Eq. (3.36), rendering the port reflection free. The expression of the output wave $\underline{a_1}$ is then

$$\underline{a_1} = -\frac{1}{\sqrt{2}}(\underline{a_L}+\underline{a_R}) \tag{3.49}$$

The order of operation in a single computational cycle is then:

1. Using Eq. (3.49), calculate $\underline{a_1}$ from $\underline{a_L}$ and $\underline{a_R}$, the two variables entering from the waveguides on either side of the series junction.

2. Calculate f_H, the junction force at the parallel junction from $\underline{a_1}$, $\underline{a_2}$ and $\underline{a_3}$ using Eq. (3.44).

3. Using Eqs. (3.41)-(3.43), calculate the scattered waves $\underline{b_1}$, $\underline{b_2}$ and $\underline{b_3}$ at the parallel junction.

4. Using Eq. (3.45)-(3.47), calculate the scattered waves $\underline{b_L}$ and $\underline{b_R}$ at the series junction.

5. Shift the scattered waves into the delay registers, and shift all samples in the waveguides.

Only step 2) above involves the hammer nonlinearity, and requires special care. Indeed, the definition of the junction force f_H , from Eq. (3.44), depends on R_3, which is a non-linear function of y_d, and thus, from Eq. (A.1), on f_H. This nonlinear equation must be solved before the scattering of step 3) above can occur. A simple iterative method such as Newton's Method suffices for this solution. Nevertheless, due to incomplete convergence of the iterative method, the calculated f_H is slightly different from the exact solution to the abovementioned equation, the resulting scattering operation is still orthogonal, and thus energy is preserved. This scattering hammer model has been connected with digital waveguides (see Section 2.3), simulating in the time domain the propagation of waves in each part of the string. This implementation of the digital waveguide in the time domain is described in Section 5.

3.7.4 Simulation by connecting the hammer model to a simple delay line

In this section, we present simulations of the piano string-hammer interaction. This has been done in collaboration with Stefan Bilbao. This paragraph can also be found in [18].

Here, perfectly reflecting string terminations, losslessness and stifflessness of the string are here assumed. The wave digital hammer is decoupled from the digital waveguide when contact is lost (i.e., when a negative value of y_d is calculated).

In Figure 3.9, three generated hammer force curves are shown, for the notes C2, C4 and C4; they agree well with figures published in [27].

More importantly, in Figure 3.10, we show the stored signal energy as a function of time for the C4 note, as well as its partition into hammer and string energies. The hammer energy is calculated from the sum of the squares of the signals in the two upper delay registers, and the string energy is simply the squared sum of all the signal values in the digital waveguides, on both sides of the scattering junction. Since there is no loss in the string, the stored energy is constant (reflecting perfect losslessness of the model system), to within errors due to machine precision.

3.7.5 Simulation by coupling with a lossy and stiff string

To get a more realistic model and consequently a better sound synthesis, we have coupled the hammer model to a stiff and lossy string. This has been done using digital waveguides, as described in section 2.3. The parameters of the loop filters have been estimated from the experimental setup using calibration methods as discussed in section 2.6. The corresponding string velocities are shown in Figure 3.11. The introduction of loss and dispersion phenomena in the string does not really modify the behavior of the hammer force. Nevertheless, it

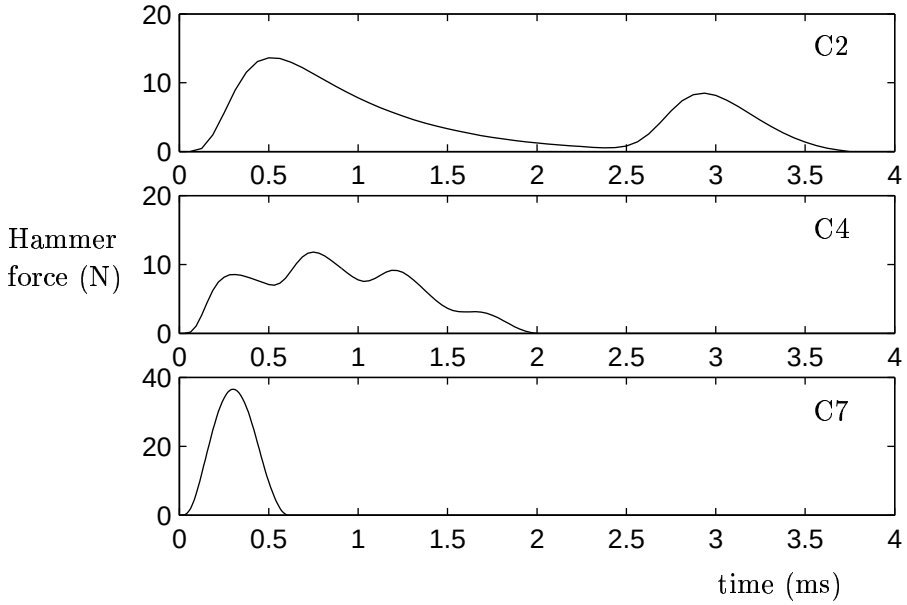

Figure 3.9: *Hammer force curves as a function of time, for C2, C4 and C7 lossless and stiffless piano strings. The hammer parameters and the striking positions are as given in [27].*

strongly affects the propagation of waves in the string and the resulting tone.

Synthesized sounds of struck strings obtained by coupling a hammer (modeled by digital wave filters) and a string (modeled using digital waveguides) are realistic, but differ from the original ones, measured on the experimental setup (they sound less bright). The filters of the waveguide model have been accurately estimated from the measured signal, but the parameters used for the hammer model are found in the literature [27]. The difficulties in measuring those parameters, as well as the relative simplicity of the physical hammer model we used (regarding to the physics of the phenomena) don't allow a perceptually satisfying resynthesis of the measured signal.

However, the DWFs method is extremely promising, thanks to its good stability properties and the possibilities of connection with digital waveguides. A more detailed work on the physical hammer model, will eventually allows a very good resynthesis of the original signal. The hammer model could be improved by taking into account some effects such as non-uniformity of hammer force over the contact area, delay effects in the felt expansion, hammer modes. We started this work by including the hysteretic behavior of the felt in the wave

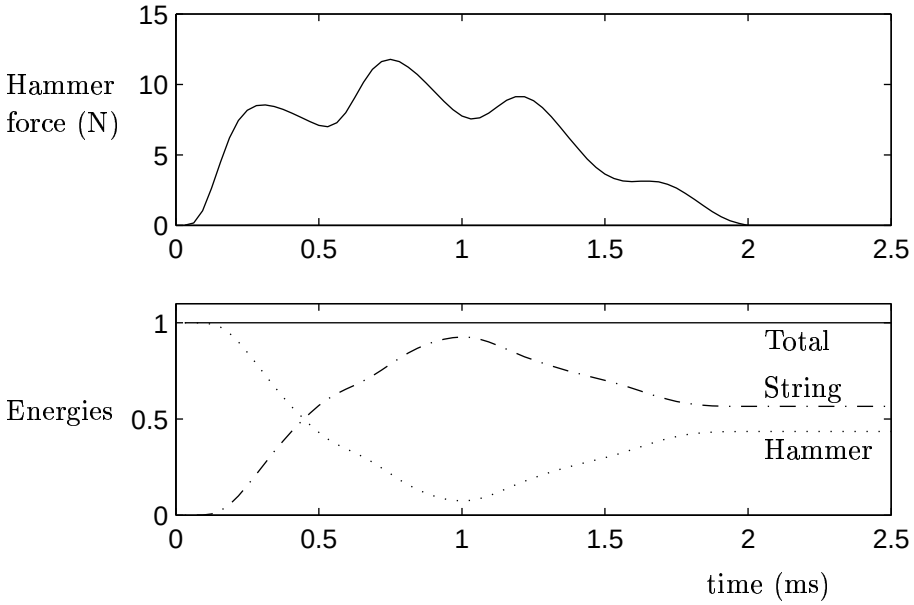

Figure 3.10: *At top, the hammer force curve for a C4 string, and at bottom, the energies of the hammer (dotted), the string (dashed) and their total (solid) as a function of time. The total string energy is normalized to unity.*

digital hammer model (presented in appendix B). However such a study necessitates a lot of physical measurements on the felt which are beyond the scope of this thesis.

3.8 Modeling the source using signal models

We return in this section the source-resonator approach used in Sections 2.5 and 2.6. As we have seen, the resonator we designed and calibrated accurately reproduces the physical phenomena in the strings and is invariant with hammer velocity. The source is related to the hammer-string interaction and takes into account the non-linear phenomena involved during this interaction. We first analyzed the behaviour of the excitation signals. They are obtained for several hammer velocities by the deconvolution between the signal collected on the experimental setup and the transfer function of the waveguide model. The excitation signal varies nonlinearly as a function of the velocity. Then, we have tried several signal methods to model this behaviour and calibrate their parameters. One drawback to this approach is that we do not take into account the feedback of the resonator on the source. As we have seen in Section 2.5.1, this is not of importance since the hammer string interaction is very short and not perceptible in itself. But when the hammer strikes an already vibrating

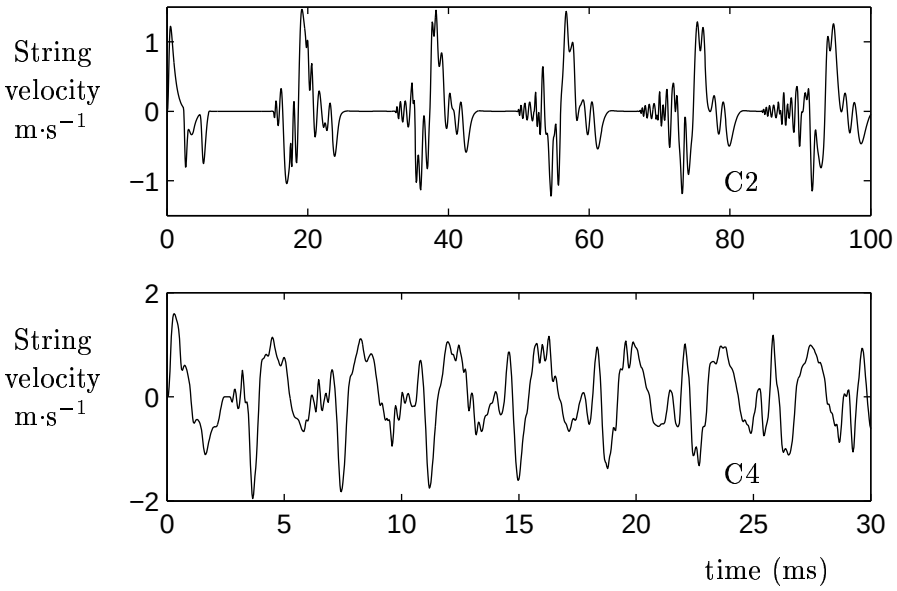

Figure 3.11: *The string velocity at the contact location between the hammer and the string. At top, for a note C2, and at bottom, for a note C4.*

string the excitation should be modified. This aspect, which has no great effect perceptually, is not taken into account here.

3.8.1 Theoretical expression of the source

Let's consider that the hammer strikes a string, fixed at its extremities, at a location x_0 with an impulse force P. As we have seen in Section 2.5.2, the signal measured at a given location can be expressed as a sum of exponentially decaying sinusoids (see Eq. (2.34)). According to Valette and Cuesta [112] the amplitudes of the partials of the waves traveling in the string are given by the relation

$$a_k = \frac{2P}{\sqrt{(}\rho_L T)} \frac{sin(k\pi x_0/L)}{k\pi\sqrt{1+k^2B}}$$

where ρ_L is the linear mass, T the string tension, L its length and B the inharmonicity factor. The excitation is related to the amplitude of the string partials (see Section 2.6.1). In the one string case, we have $E(\omega_k) = a_k D$ with D the delay of the elementary waveguide model. D is expressed as $D = 2L/c$ with c the fundamental wave speed, since D is the time the wave takes to go back and forth along the string. The excitation signal can be then expressed by

$$E(\omega_k) = \frac{4LP}{T} \frac{sin(k\pi x_0/L)}{k\pi\sqrt{1+k^2B}} \tag{3.50}$$

As the hammer force is not an impulse and is not located at a single point, this expression gives us a very approximate idea of the relation between some physical parameters and the excitation signal. The hammer velocity dependence of the excitation will be taken into account by the term P.

3.8.2 Experimental behavior of the source

The amplitudes of the excitation signals are estimated by deconvolution for each partial frequency (see Section 2.6). Those partials correspond to frequencies "selected" by the string, but the "actual" excitation spectra have no resonances. To get the "actual" continuous excitation, we interpolate its spectrum. The excitation signal sounds like an impact, short and broadband. Figure 3.12 shows the excitation signals extracted from the measurement

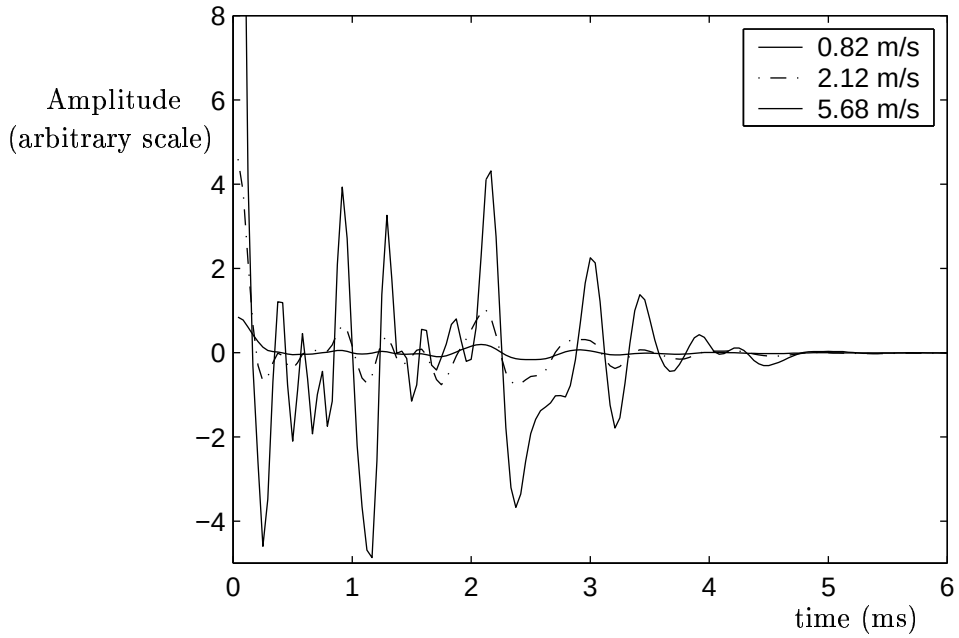

Figure 3.12: *Waveform of three excitation signals, corresponding to three different hammer velocities.*

of the vibration of a single string struck by a hammer for three velocities corresponding to the pianissimo, mezzo-forte and fortissimo musical playing. The excitation duration is about 4 ms, which is in accordance with the duration of the hammer-string contact [5].

Since this interaction is non-linear, the source behave non-linearly, basically becoming brighter for high hammer velocities. The spectra of several excitation signals obtained for a single string at different velocities regularly spaced between 0.8 and 4 $\mathrm{m.s^{-1}}$ are represented

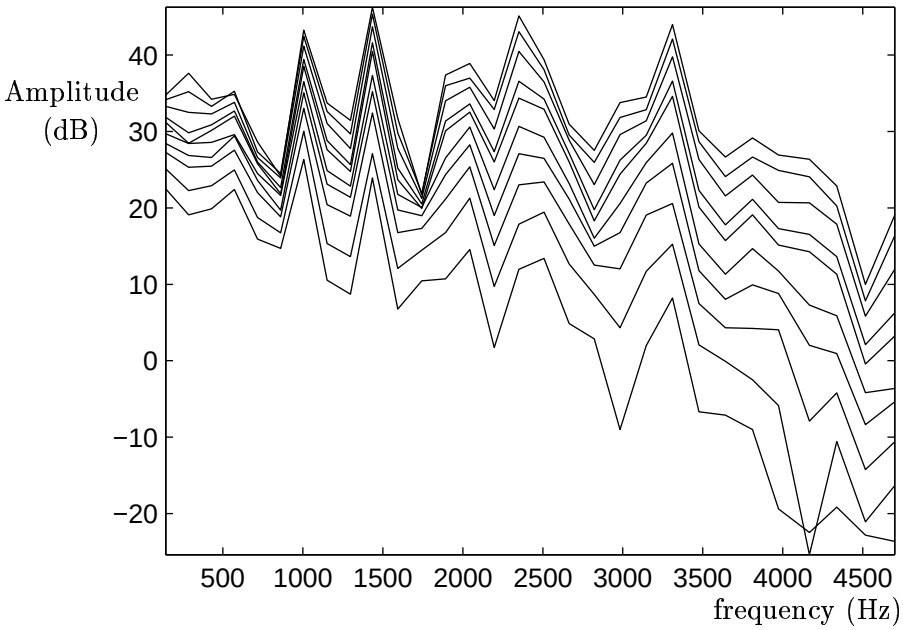

Figure 3.13: *Amplitude of the excitation signals for one string and several velocities, as a function of the frequency.*

in Figure 3.13. The excitation corresponding to fortissimo playing provides more energy than the ones corresponding to mezzo-forte and pianissimo. But this increase of the amplitude is frequency dependent: the higher partials increase more rapidly than the lower ones with the same hammer velocity. This increase of the high partials corresponds to an increase of brightness with respect to the hammer velocity. It can be better visualized by considering the spectral centroid (see Section 1.2.3 and [13]) of the excitation signals. Figure 3.14 shows the behaviour of this perceptual criteria as a function of the hammer velocity. We can clearly see for one, two or three strings an increase of the amplitude of the spectral centroid, showing an increase of the brightness of the sound. We can also notice on Figure 3.13 several irregularities among which appears a periodic modulation related to the location of the hammer impact on the string.

In the next sections, we try different signal methods to model the excitation behaviour.

3.8.3 Additive synthesis

One simple way to model the excitation is to use sinusoidal additive synthesis methods (see Section 1.3.1). The excitation signal should be continuous in the frequency domain, as the hammer produces a broadband spectrum. By using such a method, we generate an excitation

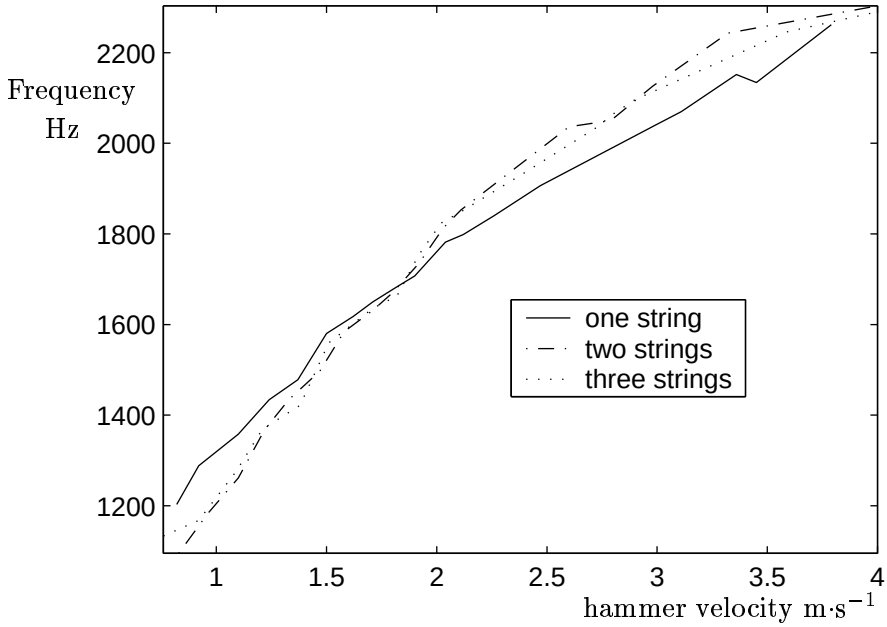

Figure 3.14: *The spectral centroid of the excitation signals for one (plain), two (dashed) and three strings (dotted).*

signal having modes with frequencies corresponding to the mode frequencies of the string. This is not a drawback as long as the excitation is tuned with the resonator.

Figure 3.15 shows the evolution of the first partials of the excitation spectra as a function of the hammer velocity. They all increase with respect to the hammer velocity. The mean slope of the curves depends on the rank of the component. We can efficiently fit their amplitude using a second order polynomial. We have identified 50 partials on the signal from our experimental setup. We then performed 50 fits and we got 50 sets of three polynomial coefficients. In order to reduce this amount of data, we considered the behaviour of each coefficient with respect to the partial number. Figure 3.16 shows that this behaviour is really irregular. It will then be very difficult to find a law, governing their evolution as a function of the partial index. Those irregularities are due to many phenomena, including the modulation of the spectrum due to the hammer position. We haven't found any relevant way to take it into account. For these reasons, we have put this model aside.

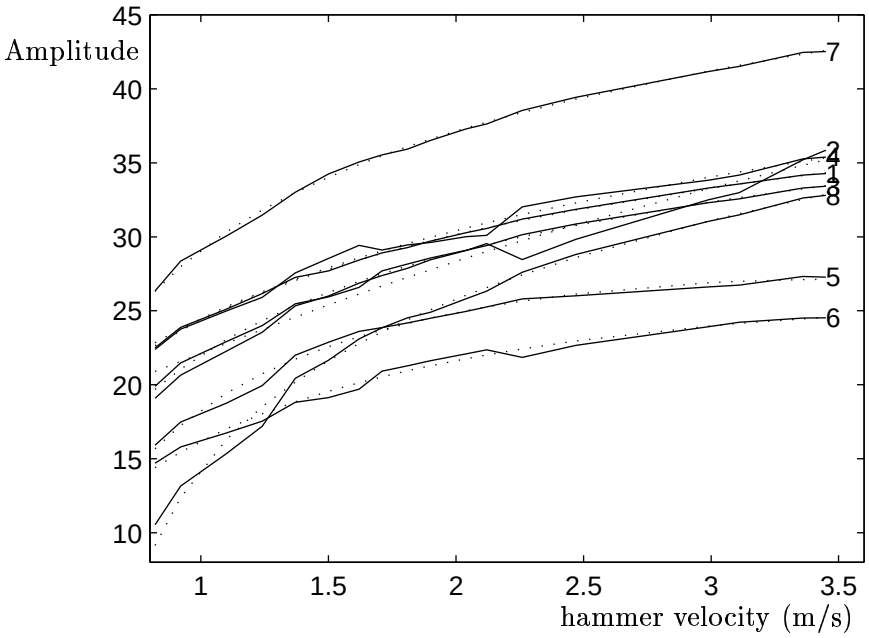

Figure 3.15: *Amplitude of the first ten partials of the excitation signal as a function of the hammer velocity for a single string.*

3.8.4 Wave shapping

The wave shapping method [22] is a non-linear synthesis method (see Section 1.3.1) allowing to generate complex signals. Ystad [118] used it to reproduce the behaviour of the source of a flute model as a function of the pressure. We have tried this method to model the behaviour of our excitation signal as a function of the hammer velocity. We first recall the mathematics of this method and then show its application to our problem.

Principle

The signal obtained by waveshaping is expressed as

$$s(t) = \gamma(I(t)cos(\omega_0 t)) \tag{3.51}$$

where $\gamma(x)$ is a non-linear function and $I(t)$ is the index of distorsion. The non-linear function can be decomposed into different basis. The first is Chebyshev's polynomials basis, frequently used because they allow the signal s to be written as an harmonic sum of sinusoids. We may

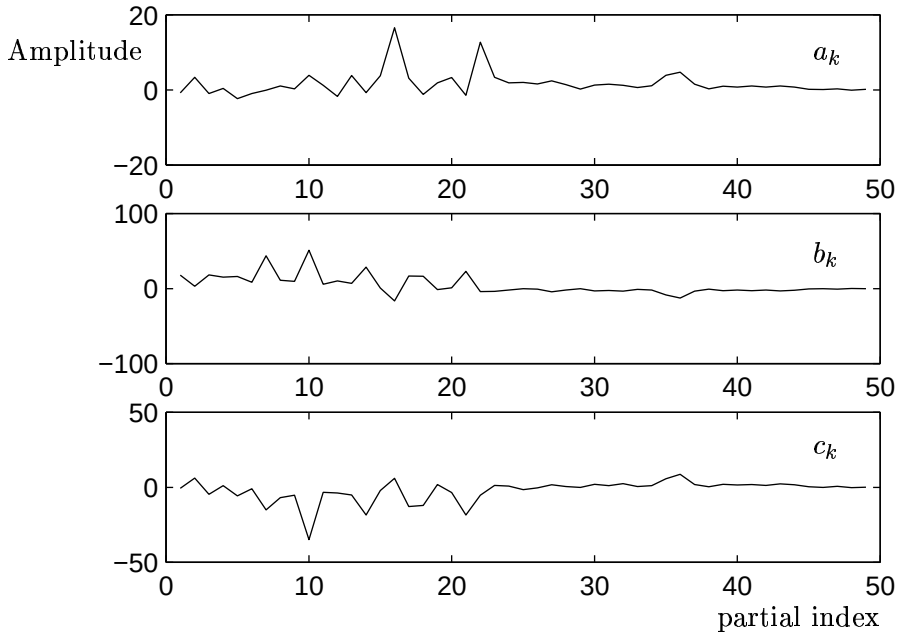

Figure 3.16: *Amplitude of the three polynomial coefficients as a function of the partial index.*

write

$$\gamma(x) = \sum_{k=0}^{\infty} a_k T_k(x) \tag{3.52}$$

with a_k being constant coefficients and T_k the Chebyshev's polynomials. These polynomials are defined by [50]

$$\begin{aligned} T_n(x) &= cos[n\ arccos(x)] \\ &= [(x + i\sqrt{1-x^2})^n + (x - i\sqrt{1-x^2})^n]/2 \end{aligned} \tag{3.53}$$

The Chebyshev's polynomials are orthogonal with respect to the measure

$$\frac{dx}{\sqrt{1-x^2}}, x \in [-1, 1]] \tag{3.54}$$

which means

$$\int_{-1}^{1} \frac{T_n T_m}{\sqrt{1-x^2}} dx = \begin{cases} \delta_{nm}\frac{\pi}{2} & m \neq 0 \\ \delta_{nm}\pi & m = 0 \end{cases} \tag{3.55}$$

When the index of distorsion is equal to 1, the output s is given by

$$s(t) = \gamma(cos(\omega_0 t)) = \sum_{k=0}^{\infty} a_k T_k(cos(\omega_0 t)) = \sum_{k=0}^{\infty} a_k cos(k\omega_0 t) \tag{3.56}$$

The synthesis method generates a signal with a fundamental frequency of f_0 and harmonics with amplitudes of a_k. For other values of the index, the spectrum of the generated signal varies according to

$$s(t) = \gamma(I(t)cos(\omega_0 t)) = \sum_{k=0}^{\infty} a_k T_k(I(t)cos(\omega_0 t)) \tag{3.57}$$

Application to the synthesis of the excitation signal

It is then possible to synthesize a harmonic spectrum and its evolution by acting on the waveshaping index $I(t)$. We can apply this method to model the evolution of the excitation, assuming that the synthesized harmonic spectrum will be "stretched" afterward since piano tones are inharmonic. As we model the evolution of the excitation as a function of the hammer velocity, the main perceptual feature we want to take into account is the increase in brightness. We then choose a waveshaping index allowing the centroid of the γ function to be as close as possible to the original.

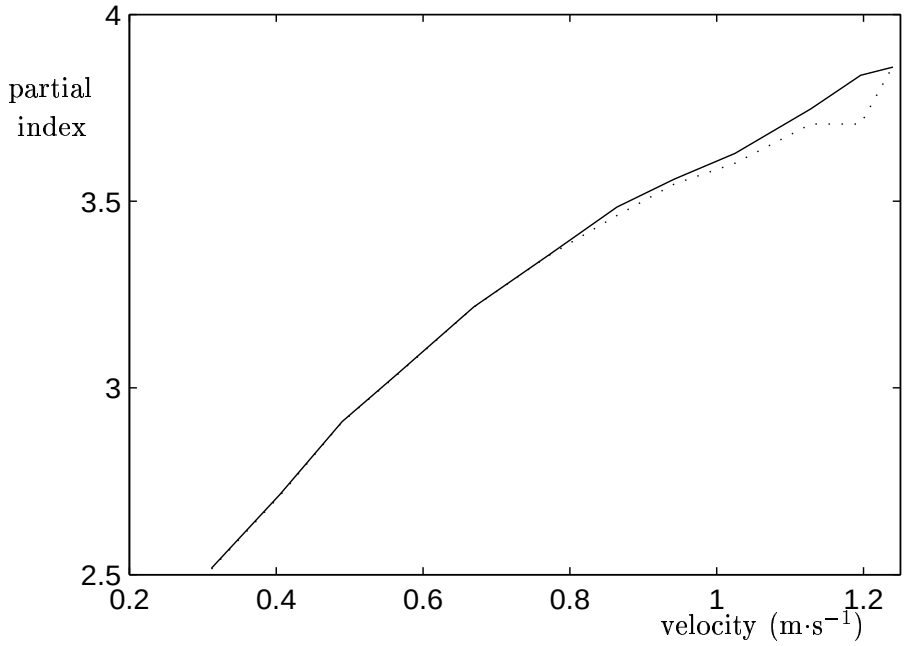

Figure 3.17: *Spectral centroid of the original signal (plain) and of the γ function (dotted) for twelve hammer velocities*

The first step consists of constructing a non-linear γ function so that the "richest" spectrum (corresponding to the highest hammer velocity) corresponds to a waveshaping index $I = 1$. For this value of the waveshaping index, the amplitudes of the synthesized signal

corresponds to the a_k coefficients of the γ function. The second step consists in estimating the wave-shaping index so that the spectral centroid of the γ function corresponds to the centroid of the original tone as a function of the hammer velocity. This was done for twelve hammer velocity in the range 0.3 to 3.4 m.s^{-1}. The corresponding index varied from 0.62 to 1. We can see on Figure 3.17 that the two centroids are very close, especially for low velocities. The corresponding excitation spectra are shown on Figure 3.18.

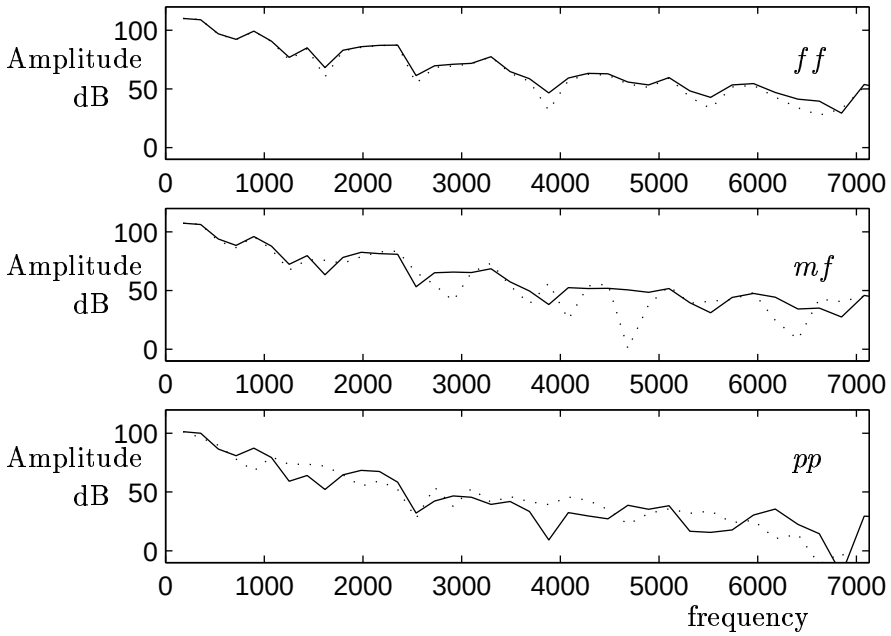

Figure 3.18: *Spectrum of the excitation for a three hammer velocities: original (plain) and modeled using wave-shaping (dotted).*

The global increase of brightness is well reproduced. When injected into the resonator model, these excitation signals give a sensation related to the dynamics of the final string sound. But when compared to the original tone, we notice that the synthetic tone is slightly different. In order to resynthesize the original sound, the amplitudes of each excitation partial must be controlled more accurately. Contrary to the flute [118], piano tones have a lot of partials and a global criteria, such as the centroid of the spectrum, is not relevant. Using an other criterion like the tristimulus criterion provides improvements to the synthesis, but the resulting string sound is still different from the original.

3.8.5 Subtractive synthesis

The excitation behaves regularly as a function of the hammer velocity. This causes an increase in the amplitudes of the spectral components. For high frequency components, this increase is more important than for low frequency components, leading to a flattening of the spectrum. Nevertheless, the general shape of the spectrum stays the same. Formants do not move and the modulation of the spectrum due to the hammer position on the string is visible at any velocity. These observations make us believe that the behaviour of the excitation could be well reproduced using a substractive synthesis model (Section 1.3.1). This work has been done in collaboration with researchers form DIKU university, Copenhaguen and has been published in [16].

The excitation signal can then be seen as an invariant spectrum shaped by a smooth frequency response filter the characteristics of which depends on the hammer velocity. The source model is shown on Figure 3.19.

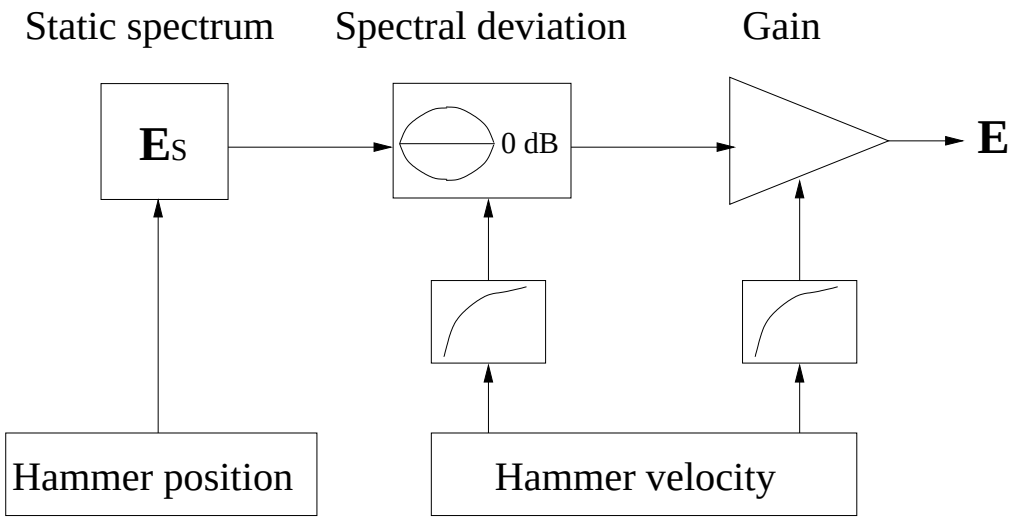

Figure 3.19: *Diagram of the substractive source model.*

It consists of the static spectrum, the spectral deviation and the gain. The static spectrum takes into account all the information that is invariant with respect to the hammer velocity. It is then a function of the hammer position and the strings characteristics, as we will see momentarily. The spectral deviation and the gain, shape the spectrum as function of the hammer velocity. The gain will model the global increase of amplitude and the spectral deviation will simulate the shifting of the energy to the high frequencies.

A similar approach has been developed by Smith and Van Duyne in the time domain

[104]. The hammer-string interaction force pulses were simulated using three impulses passed through three low pass filters which depends on the hammer velocity. In our case, we need a more accurate method to resynthesize the original excitation signal.

The static spectrum

We defined the static spectrum as the part of the excitation that is invariant with the hammer velocity. By considering the expression Eq. (3.50), the static spectrum E_s can be expressed as

$$E_s(\omega_k) = \frac{4L}{T} \frac{sin(k\pi x_0/L)}{k\pi\sqrt{1+k^2B}} \tag{3.58}$$

since the term P in Eq. (3.50) is related to the force and is consequently not invariant with the hammer velocity. We can easily measure the striking position x_0, the string length L and the inharmonicity factor B on our experimental setup. On the other hand, we only have an approximate estimation of the tension but we can calculate it using the relation Eq. (1.4), knowing the fundamental frequency and the linear mass. This formula will provide a good estimation of the static spectrum.

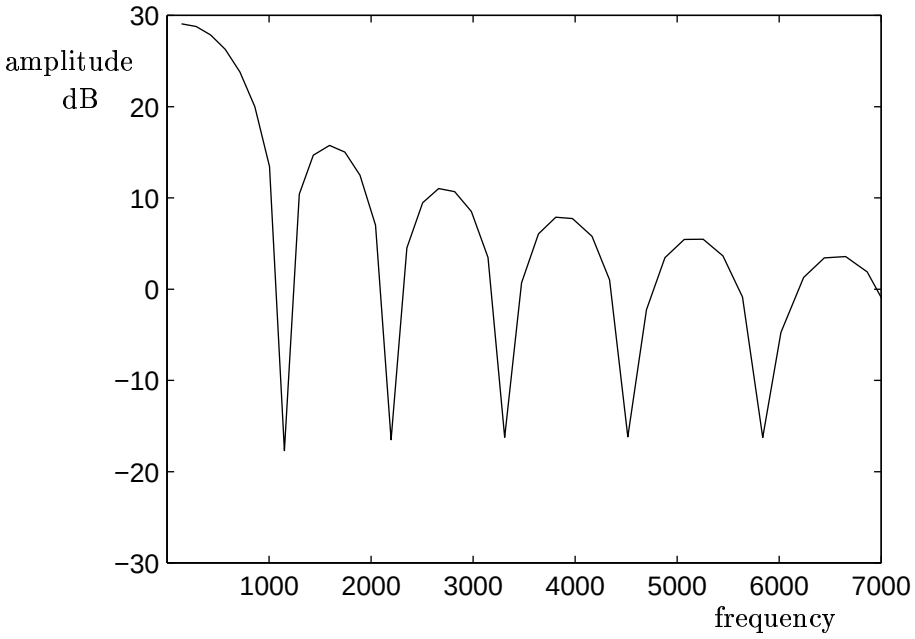

Figure 3.20: *Static spectrum.*

Figure 3.20 shows this static spectrum for a single string. Many irregularities are however

not taken into account. These are due to several reasons. We will see later their importance from a perceptual point of view.

The deviation with the dynamics

The spectral deviation and the gain take the dependency of the excitation signal on velocity into account. They are estimated by dividing the excitation signal with the static spectrum for all velocities:

$$d(\omega) = E(\omega)/E_s(\omega) \tag{3.59}$$

where E is the original excitation signal. Figure 3.21 shows this deviation for three hammer velocities. It effectively strengthens the fortissimo, in particular the medium and high partials. Its evolution with the frequency is quite regular and can successfully be fitted to a second order exponential polynomial (as shown in Figure 3.21)

$$\hat{d} = exp(af^2 + bf + g) \tag{3.60}$$

where $\hat{d}$ is the modeled deviation. The term g corresponds to the gain (since independent of the frequency) and the terms $af^2 + bf$ corresponds to the spectral deviation.

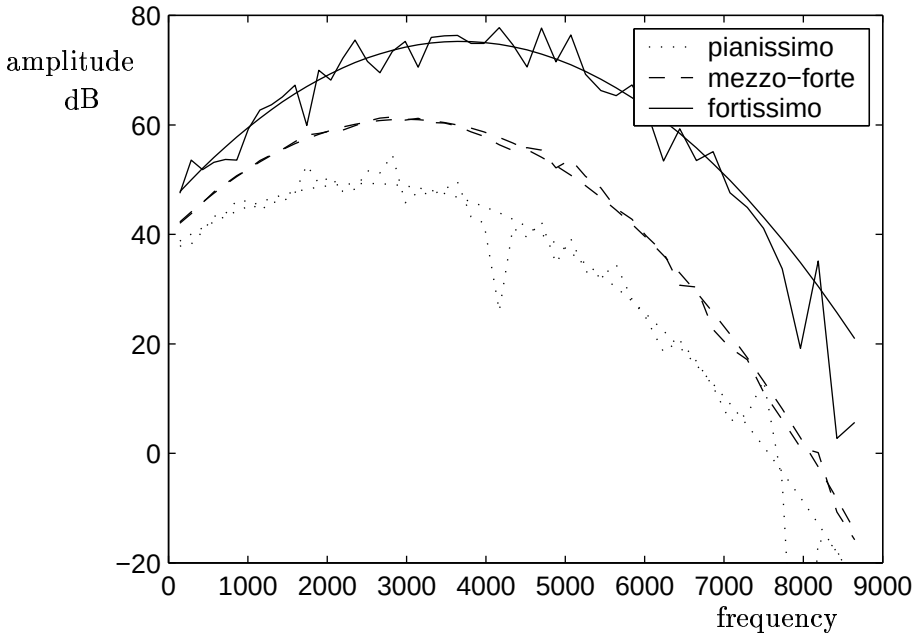

Figure 3.21: *Dynamic deviation.*

g, a and b depend on the hammer velocity. In order to get a usable source model, we must consider the parameter's behaviour with different dynamics. Figure 3.22 shows the three parameters for several hammer velocities. The model seems to be consistent since their behaviour is regular. But, the tilt is decreasing with the hammer velocity, showing an asymptotic and non-linear behaviour (at least for a and b). This observation can be directly related to the physics of the hammer. As we have seen, when the felt is compressed it becomes harder and thus gives more energy to high frequencies. But, for high velocities, the felt is totally compressed and its hardness is almost constant. Thus, the amplitude of the corresponding string wave increases but its spectral content is roughly the same. We have fitted this asymptotic behaviour by an exponential model (see Figure 3.22), for each parameter g, a and b

$$\begin{cases} g(v) = \alpha_g - \beta_g exp(-\gamma_g v) \\ a(v) = \alpha_a - \beta_a exp(-\gamma_a v) \\ b(v) = \alpha_b - \beta_b exp(+\gamma_b v) \end{cases} \tag{3.61}$$

where α_i $(i = g, a, b)$ is the asymptotic value, β_i $(i = g, a, b)$ is the deviation from the asymptotic value at zero velocity (the dynamic range), and γ_i $(i = g, a, b)$ is the velocity exponential coefficient, governing how sensible the attribute is to a velocity change. The parameters of this exponential model have been found using a non-linear curvefit.

Resynthesis of the excitation signal

For a given velocity, the excitation signal can now be recreated using Eqs. (3.59), (3.60), (3.61). This source model convoluted with the transfer function of the resonator leads to an acceptable sound of a string struck by a hammer. The increase of the brightness with the dynamic is well reproduced. But from a resynthesis point of view, this model is not satisfying. The reproduced signal is different from the original one, sounding too regular and monotonous. In order to understand this lack of our model, we have calculated the error we made by dividing the original excitation signal with the modeled one for each velocity. The corresponding curves are shown on Figure 3.23 for three velocities.

We can notice that this error term doesn't depend on the hammer velocity, meaning that our static spectrum model is too straightforward and does not take into account the irregularities of the original spectrum. Irregularities are due to many phenomena including the width of the hammer-string contact, hysteretic phenomena in the felt, non-linear phenomena in the string, mode resonances of the hammer, etc. To give a more realistic feature to our source model, we can include this error term in the static spectrum. The original and resynthesized signals are shown on Figure 3.24. The deviations of the resulting excitations are perceptually insignificant.

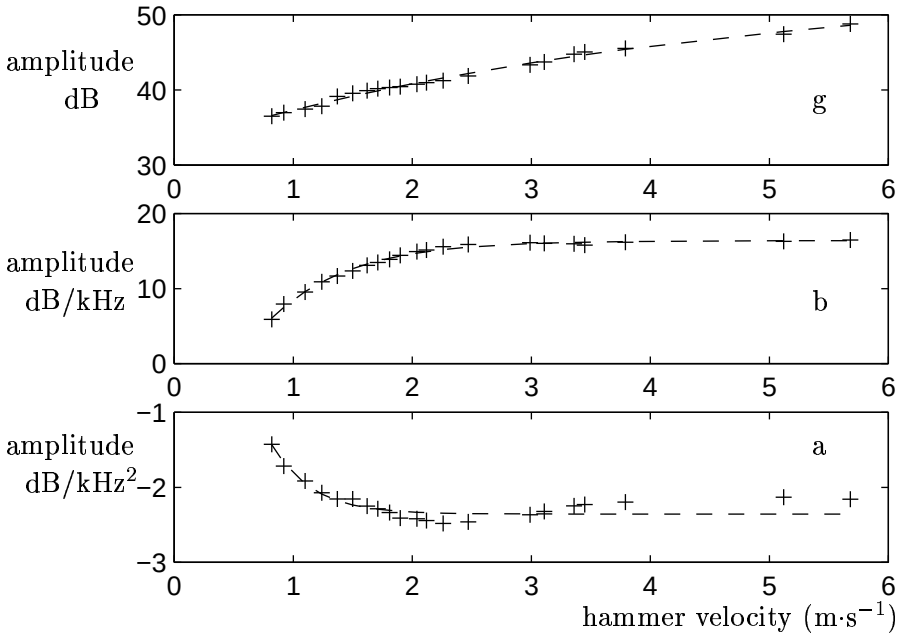

Figure 3.22: *Parameters g (gain), a and b (spectral deviation) as a function of the hammer velocity. Original(+) and modeled(dotted)*

3.9 Conclusion: physics versus resynthesis

We have described several methods allowing the modeling of the hammer-string contact and producing the corresponding sound. Purely physical models, like the coupling of a non-linear spring equation with the PDE of a stiff and frequency dependent lossy string, allow the reproduction of some of the phenomena, but still the sound lacks realism. A parametric calibration, using optimization methods, provides for a better perceptual effect, without being totally satisfying. Further, numerical methods using finite differences is expensive computationally and the construction of a real time polyphonic piano model currently seems difficult to achieve by this method.

The modeling of this interaction by the coupling of a hammer using wave digital filters to a string modeled by digital waveguide presents the double advantage of maintaining the close relationship with the physical parameters, all being less costly with respect to computation. But, still further, the hammer model seems too simple and its parameters difficult to calibrate to permit a quality resynthesis of the measured signal.

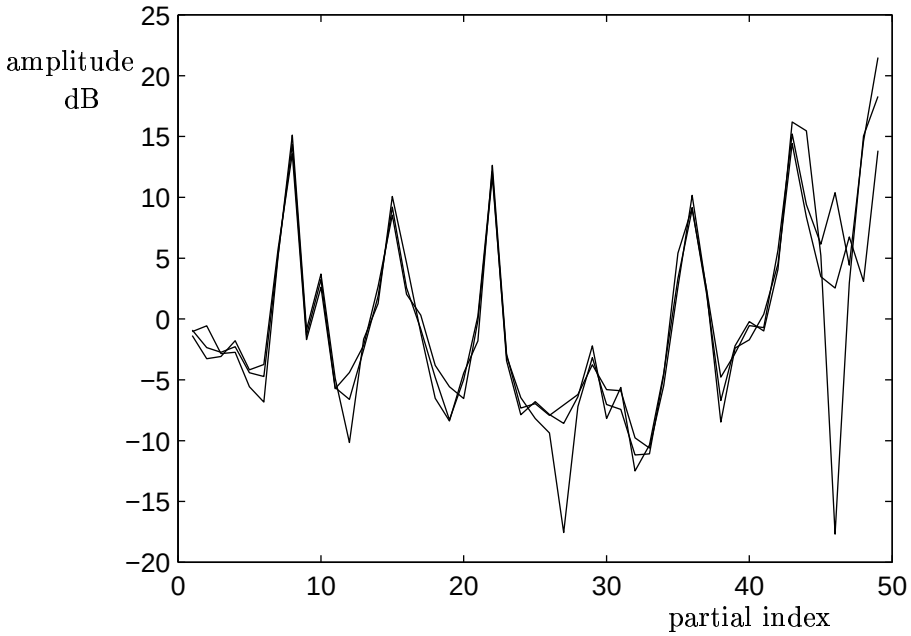

Figure 3.23: *Error spectrum.*

The signal models are, then, the most performant with respect to the quality of the resynthesis. But, the relationship between the parameters of these models and the physical parameters is difficult to establish. The subtractive synthesis model presented in Section 3.8.5 permits a very precise resynthesis of the original excitation. These parameters cannot be explicitly linked to the physical parameters, but we have shown that it is possible to establish empirically their relationship with the hammer velocity. We will choose, then, this method to synthesize the excitation signal and in so choosing will abandon the idea of a purely physical model of the piano in order to achieve a higher quality synthesis.

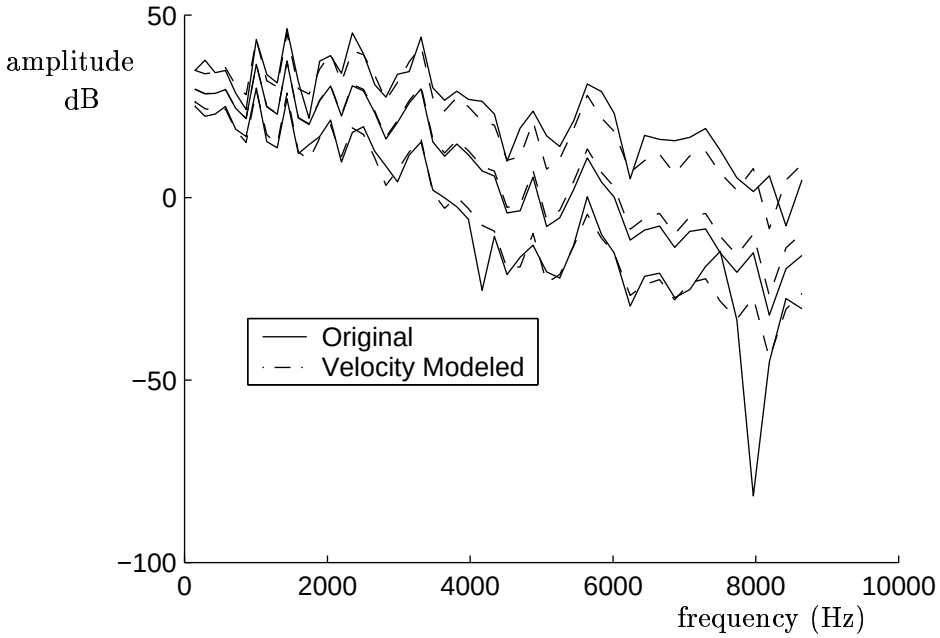

Figure 3.24: *Original and modeled excitation spectrum for three different hammer velocities.*

Chapter 4

The piano model

4.1 Introductory remarks and choice of the model

To take into account the necessary simplifications made in the physical description of the piano sounds, we have finally used a hybrid model. Hybrid models are obtained by combining physical and signal synthesis models [118]. The physical model simulates the physical behavior of the instrument while the signal model tends at simulating the perceptual effect produced by the instrument. It allows in perfectly reconstructing a given sound and in manipulating it in both a physical and a perceptual way.

The resonator has been modeled using a physically related model: the digital waveguide (Sections 2.3 and 2.5), while the source -the aim of which is to generate the initial condition for the string motion- has been modeled using a signal based non-linear model (Section 3.8.5). This approach was the best compromise we have found to reach as close as possible the objectives we have presented in Section 1.1.1. The advantages of such a hybrid model are numerous:

- it is simple enough so that the parameters can be estimated from the analysis of real sound,
- it takes into account some of the most relevant physical characteristics of the piano strings and permit the control with respect to the playing (the velocity of the hammer)
- it simulates the perceptual effect due to the non-linear behavior of the string-hammer interaction,
- it allows sounds transformation with both physical and perceptual approaches.

4.2 Piano setup

Because our piano model is intended for use in the context of musical sound synthesis, we had to estimate the behavior of all the parameters using measurements on a real piano. Data were collected on an experimental setup consisting of a Yamaha Disklavier C6 grand piano equipped with sensors. The Disklavier allows the piano to be played under computer control. Each key can be actionned at several velocities using a MIDI protocol. The actual MIDI velocity of the hammer can also be measured. As we wanted the hammer velocity in $\mathrm{m{\cdot}s^{-1}}$, we had to find the relation between the MIDI code and the "physical" hammer velocity. For several notes we measured both velocities, using a computer with a MIDI interface and an optical sensor pointing at the head of the hammer (see Figure 4.1). We deduce from those

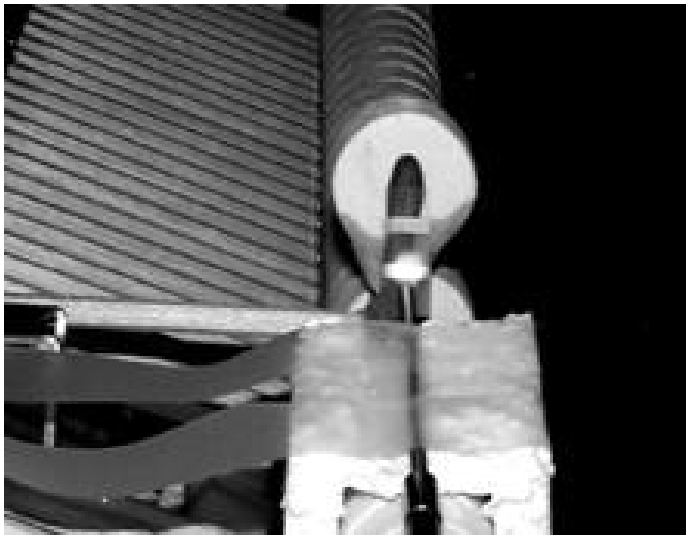

Figure 4.1: *Optical sensor used to measure the hammer velocity. The velocity is obtained through the duration corresponding to the travel time of the hammer between two reflectors placed on it.*

data the relation between the hammer velocity and the MIDI velocity

$$v_h = exp(0.0284 * v_{MIDI} - 1.4976) \tag{4.1}$$

where v_h is the hammer velocity in $\mathrm{m{\cdot}s^{-1}}$ and v_{MIDI} is the MIDI velocity (the MIDI velocity values varies from 0 to 127). The relation is the same for all the piano range.

After calibrating our experimental setup, we have measured at a first step, the string vibrations at the bridge level for one velocity and every notes. Those measurements will be usefull to estimate the resonator filters and their behavior as a function of the note. Since

we were interested in exciting a large portion of the frequency spectrum while remaining in the linear domain of vibration, we chose a medium (*mezzo-forte*) hammer velocity of 2.2 $\mathrm{m{\cdot}s^{-1}}$, which roughly corresponds to a MIDI value of 80. Such a hammer velocity allows the generation of about 140 spectral components for low frequency tones with a reasonable signal-to-noise ratio. The bridge vibration was measured using an accelerometer (BK 4374), as shown on Figure 4.2.

Figure 4.2: *Accelerometer at the bridge level and zooming supperposition*

For notes corresponding to double or triple sets of strings, the accelerometer was placed as close as possible to the strings. Due to the imprecision of MIDI coding, several measurements were made, until a target value of the hammer speed was obtained. We have deemed an uncertainty of ± 0.1 $\mathrm{m{\cdot}s^{-1}}$ for the hammer velocity to be acceptable, as the estimation of modal frequency and decay rates is relatively insensitive to such an error. The acceleration measured at the bridge level was then digitally recorded using 16 bits discretization and a sampling rate of 44.1 kHz, before being entered in the database.

As a second step, we performed the same measurements for several notes and several velocities. Typically, we measured the bridge acceleration every one octave and for ten hammer velocities. Those measurements will be usefull to verify the stability of the resonator and study the behavior of the source parameters.

4.3 Resonator model and calibration

4.3.1 Behavior of the resonator

The resonator is constituted by one digital waveguide for the one string notes, two and three-coupled digital waveguides for the two and three strings notes as presented in Sections 2.3 and 2.5. We have used the same estimation method we used to analyze the data provided by the experimental setup and calibrate the digital waveguide model.

As expected, the behavior of the resonator as a function of the hammer velocity and for a given note is similar to the one described in Section 2.6.4, for the signals measured on the experimental setup. The filters are quite similar with respect to the hammer velocity. Their modulus is close to one, but slightly weaker than previously, since it now takes into account the losses due to the acoustic field radiated by the soundboard. The resynthesis of the piano measurements using the resonator model and the excitation obtained by deconvolution is perceptively satisfying since the sound is almost undistinguable from the original one.

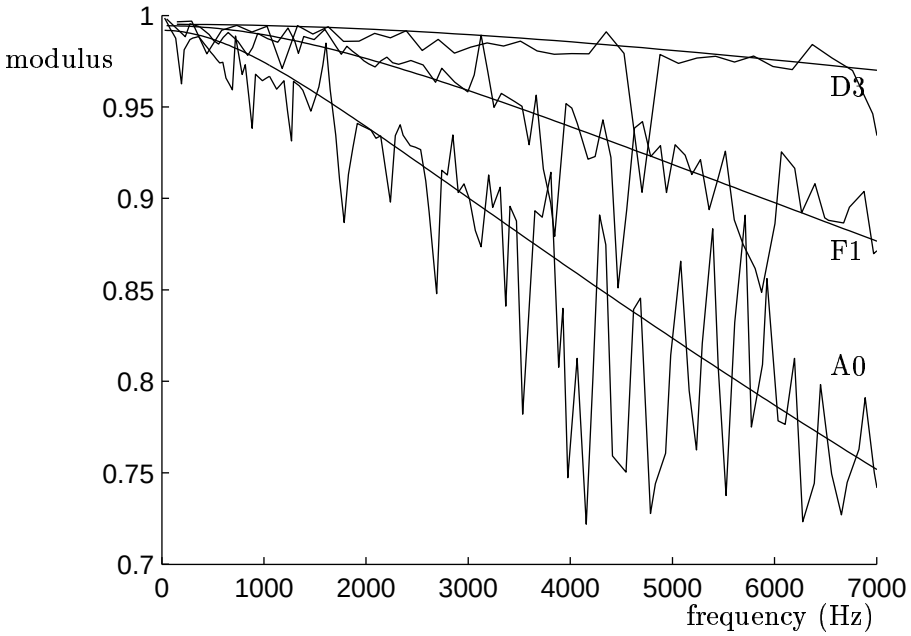

Figure 4.3: *Modulus of the waveguide filters for notes A0, F1 and D3: original and modeled.*

On the contrary, the shape of the filters is modified as a function of the note. Figure 4.3 shows the modulus of the waveguide filter F for several notes (in the multiple string case,

we have calculated an average filter using an arithmetic averaging). The modulus of the loop filter is related to the loss undergone by the wave over one period. We notice that this modulus increases with the fundamental frequency, indicating decreasing loss over one period as one approaches the treble range.

4.3.2 Model parameters calibration

We recall here the expression of the elementary loop filter (already presented in Section 2.4):

$$\begin{aligned} |\delta F(\omega)| &\simeq \exp\left(-\left[b_1 + \frac{b_2\pi^2\xi}{2BL^2}\right]\right) \\ \arg(\delta F(\omega)) &\simeq \omega - \omega_0\sqrt{\frac{\xi}{2B}} \end{aligned} \tag{4.2}$$

where

$$\xi = -1 + \sqrt{1 + 4B\omega^2/\omega_0^2}$$

L, the length of the string has been measured directly on the piano. ω_0 the fundamental frequency (rad/s) of the ideal string and B the inharmonicity coefficient, b_1 and b_2 the loss parameters will be estimated from the signal measured for each note. B and consequently ω_0 may be deduce from the mean frequency values estimated for each partial (see Section 2.6.3) using the relation Eq. (1.4). The inharmonicity factor B is plotted as a function of MIDI note number in Figure 4.4; B is an increasing function of the note number, except over the bass range, where the strings are double-wrapped (this behavior has also been measured by Conklin [32]).

Let's know estimate the loss parameters b_1 and b_2. For each note, we can calculate the damping of each component. In the case of multiple strings, we measured, the damping of the "immediate" sound which has the greatest decay rate and which approximately corresponds to the decay of the partials without coupling. Combining Eqs. (2.8) and (2.25), one obtains

$$\alpha(\omega) = -b_1 - b_2\left(\frac{\pi^2}{2BL^2}\left[-1 + \sqrt{1 + 4B(\omega/\omega_0)^2}\right]\right) \tag{4.3}$$

From the value of α obtained for each partial, b_1 and b_2 may be estimated for a given tone using a non-linear curve fit. The evolution of these parameters as a function of the fundamental frequency is shown in Figure 4.5.

We have also fitted extremely simple curves to the loss parameter data given by

$$b_1 = 4.4 \times 10^{-3} f_0 - 4 \times 10^{-2} \tag{4.4a}$$

$$b_2 = 1.0 \times 10^{-6} f_0 + 1 \times 10^{-5} \tag{4.4b}$$

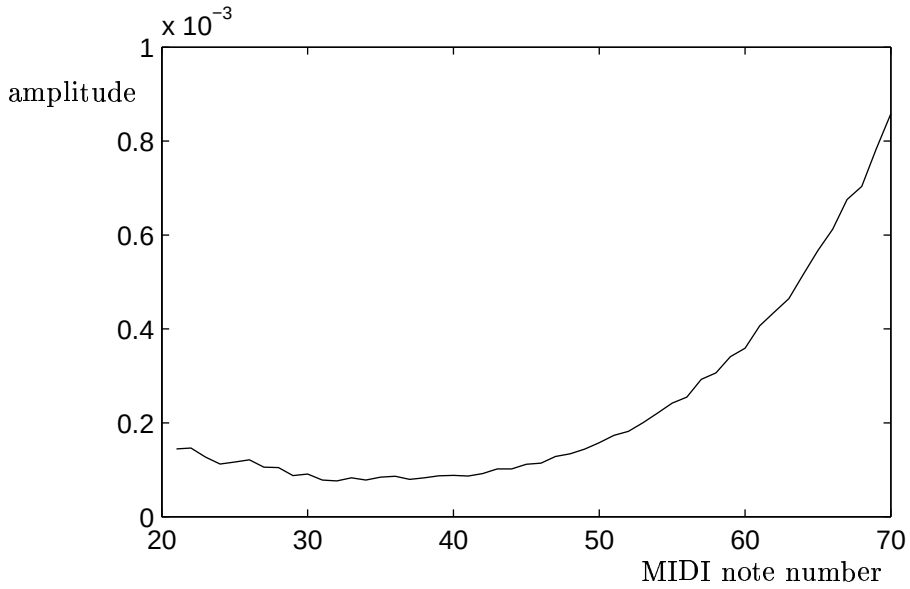

Figure 4.4: *The inharmonicity factor B as a function of the note number.*

These simple empirical descriptions of b_1 and b_2 allow the reproduction of piano tones whose damping will be close to that of the perceived acoustic note. We see that b_1 and b_2 are both increasing functions of the fundamental frequency. We give here a straightforward explanation. In a real piano, the loss is extremely important, as it is the mechanism by which energy is transferred to the soundboard, and, ultimately to the listener as a musical sound. Parameters b_1 and b_2 do not only take into account internal losses in the string, but also losses due to the energy transfer to the bridge. Each time the wave is reflected at the bridge, part of its energy will be given to the soundboard. As the strings are shorter in the treble, waves are able to complete more round-trip passages in a given time. Thus, the decay rate (which is correlated to the decrease of the amplitude per second) will be finally greater in the treble than in the bass range. In order to compare those losses, we divide b_1 and b_2 by the fundamental frequency, as shown on Figure 4.6. The loss coefficients obtained this way have a quite irregular behavior. The physical characteristics of the strings themselves (and consequently the internal losses) vary with the note, as well as the bridge impedance [46] (and consequently losses due to radiation). A more precise study of those curves would give us other information about the coupling between the strings and the soundboard but such a study is out of the scope of this work. Nevertheless, we notice that the increase of the modulus of the loop filter F (as observed on Figure 4.3) is due to the decrease of the coefficient b_2/f_0 (up to 300 Hz).

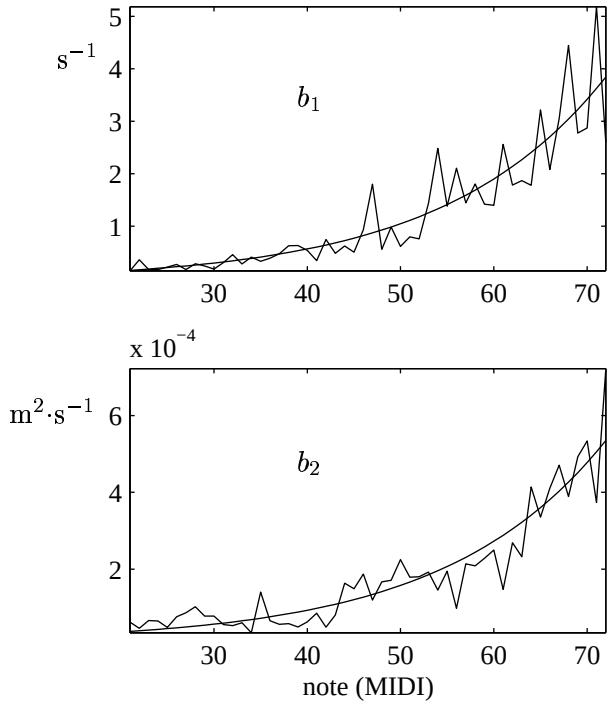

Figure 4.5: *Values of b_1 and b_2 for a grand piano, fitted from measured data, as a function of MIDI note number*

The determination of b_1, b_2, B and ω_0 for each note allows for an explicit expression of the behavior of the filter F as a function of note number. The modeled modulus is close to the original one, as shown in Figure 4.3. The original modulus has many "artifacts" that are not taken into account in the model. This will not be perceptually of great importance from a synthesis point of view.

Several modulus and phases of the elementary filter model are plotted on Figure 4.7 for the notes A0, A1, A2 and A3. In order to understand the general behavior of the elementary modulus and the phase of the loop filter and then to be able to control it in a synthesis situation, we expand their expressions (Eqs. (4.2)) to third order for $4B(\omega/\omega_0)$ near zero. We obtain

$$|\delta F(\omega)| \simeq \exp(-[b_1 + b_2\frac{\omega^2}{c^2}]) \tag{4.5a}$$

$$\arg(\delta F) \simeq \frac{\pi\omega^3 B}{2\omega_0^2} \tag{4.5b}$$

The decrease of the modulus is mainly due to the increase in b_1. But the parameter

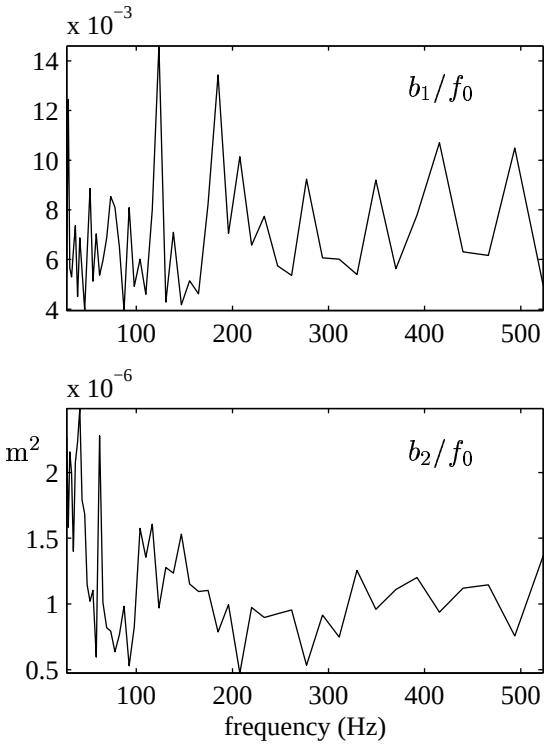

Figure 4.6: *Values of b_1/f_0 and b_2/f_0 for a grand piano as a function of the fundamental frequency f_0*

b_2, which allows for frequency dependent loss is also increasing, leading to a modification of the slope of the modulus versus frequency for different note numbers. We note that this behavior is slightly different for the wrapped strings (A0, A1) than for the other strings (A2, A3). Moreover, the behavior of the normalized filter modulus with the note is different from the one of the non-normalized filter (see Figure 4.3. In the first case we actually consider losses over a time interval of one second and over one period in the second case. We also note that although B is mainly an increasing function of the note number, the phase of the filters δF grows less rapidly for the bass tones than for higher tones. This is due to the fact that the phase of the filter depends not only on B, but also on the fundamental frequency, as evidenced by Eq. (4.5b). Though this expression is meaningful only for the first few partials, it is clear that phase dispersion decreases as a function of fundamental frequency.

The relations Eqs. (4.2), relating the physical parameters to the waveguide parameters, allow the control of the resonator in a relevant physical way. We can either change the length of the strings, the inharmonicity, or the loss. But in order to be in accordance with the physical system, we have to take into account the inter-dependence of some of the parameters.

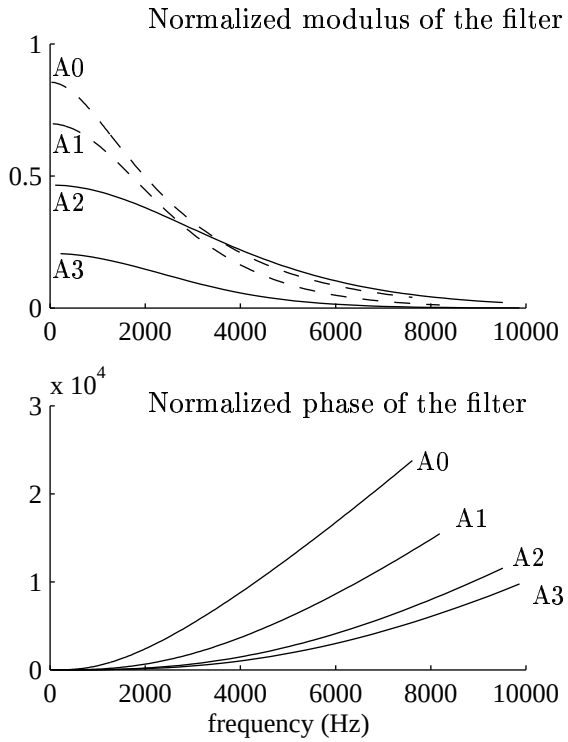

Figure 4.7: *Normalized modulus and phase of the filter F for selected tones.*

For instance, the fundamental frequency is obviously related to the length of the string, as well as the tension or the linear mass. If we modify the length of the string, we also have to modify, for instance, the fundamental frequency, considering that the tension and the linear mass are unchanged. This aspect will be taken into account in the implementation of the model.

4.4 Source model and calibration

The model we have finally chosen for the excitation signal is the substractive signal model presented in Section 3.8.5. The source model parameters have been calculated for a subset of the data for the piano, namely the notes A0, F1, B1, G2, C3, G3, D4, E5 and F6. Each note has approximately ten possible velocities, from around 0.4 m/s to 3.5 m/s. Figure 4.8 shows the normalized amplitude (divided by the delay D) of several excitation spectra for a medium velocity. The amplitude increases slightly as the frequency of the note increases.

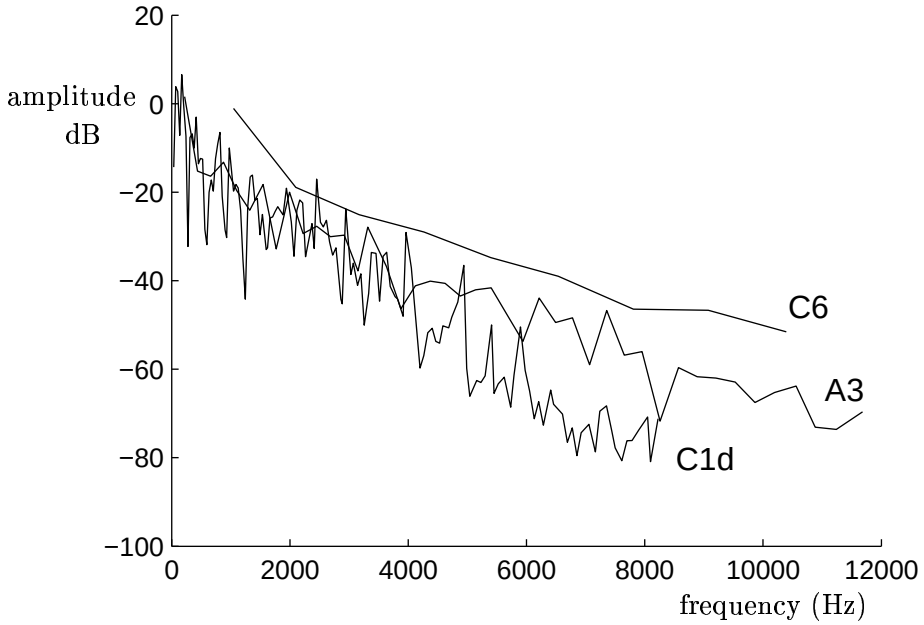

Figure 4.8: *Normalized excitation signal (divided by the delay D) for three different notes.*

4.4.1 The static spectrum as a function of the note

The static spectrum is analytically calculated (see Eq. 3.58). It is different for each note, since its behavior depends on the inharmonicity coefficient B, on the striking point x_0, on the length of the string L and on the string tension T. x_0 and L have been measured on our previous piano setup. B has been estimated on the measured signal (see Section 4.3). It is not possible to measure precisely the string tension. Moreover, we can not deduce it from the estimation of the fundamental frequency (as we did for the experimental setup), since we don't precisely know the linear mass of the strings. This linear mass is different for each note and we didn't get tables containing such information from the Yamaha Company. The tension will therefore be removed from the expression of the static spectrum and will be implicitly incorporated in the expression of the gain g.

This static spectrum model is, again, too straightforward to allow an efficient resynthesis of the final sound. The error term we added previously to the static spectrum (see Section 3.8.5) is still necessary to obtain perfect a resynthesis of the original sound. This error term is independent of the hammer velocity for a given note, but depends on the note played. Its behavior as a function of the note doesn't exhibit any coherent form and will be consequently impossible to model. As we have measured the string vibration for every notes and for a

medium velocity, we are then able to identify this error term for every note, afterwards (we first model the spectral deviation as a function of the note and then, we calculate the error term by comparing the synthesized signal with the original one).

4.4.2 The dynamic deviation as a function of the note

The source extracted from the signals measured on the piano, behaves as the experimental data for all notes with respect to the hammer velocity. The dynamic deviation is well modeled by the gain g and the spectral deviation parameters a and b.

As shown in Section 3.8.5, their behavior as a function of the velocity is well fitted using an asymptotic exponential curve. For more clarity, we recall here their expressions:

$$\begin{cases} g(v) = \alpha_g - \beta_g exp(-\gamma_g v) \\ a(v) = \alpha_a - \beta_a exp(-\gamma_a v) \\ b(v) = \alpha_b - \beta_b exp(+\gamma_b v) \end{cases} \tag{4.6}$$

For sake of a playable piano model, we have to consider the behavior of the asymptotic value α_i $(i = g, a, b)$, the deviation from the asymptotic value at zero velocity (the dynamic range) β_i $(i = g, a, b)$, and the velocity exponential coefficient γ_i $(i = g, a, b)$, with respect to the note. Figure 4.9 shows the evolution of the gain parameter as a function of the fundamental frequency of the note. The curves are smooth enough to be fitted by a second order polynomial for the gain g and also for the parameters a and b. This last fit leads to a complete piano model allowing us to resynthesize the string vibration at the bridge location as a function of the velocity and the note.

4.4.3 Control of the source

From a perceptual point of view, an increase of the hammer velocity corresponds to an increase of the intensity of the sound as well as a relative increase in high frequencies leading to a brighter tone. Equations Eqs. (4.6) allow for the resynthesis of the excitation signal for a given note and hammer velocity. However, the parameters g, a, b, used in the modeling, are linked in a complex way to the two important perceptual features of the tone which are intensity and brightness. Thus, without a thorough knowledge of the model, the user will not be able to adjust the parameters of the virtual piano to obtain a satisfying tone. To get an intuitive control of the model, we therefore need to provide the user with perceptual parameters. We have therefore given the user access to the energy and the spectral centroid parameters. The energy En is directly correlated to the perception of intensity and the spectral centroid Ba to the perception of brightness [13]. These parameters are given by:

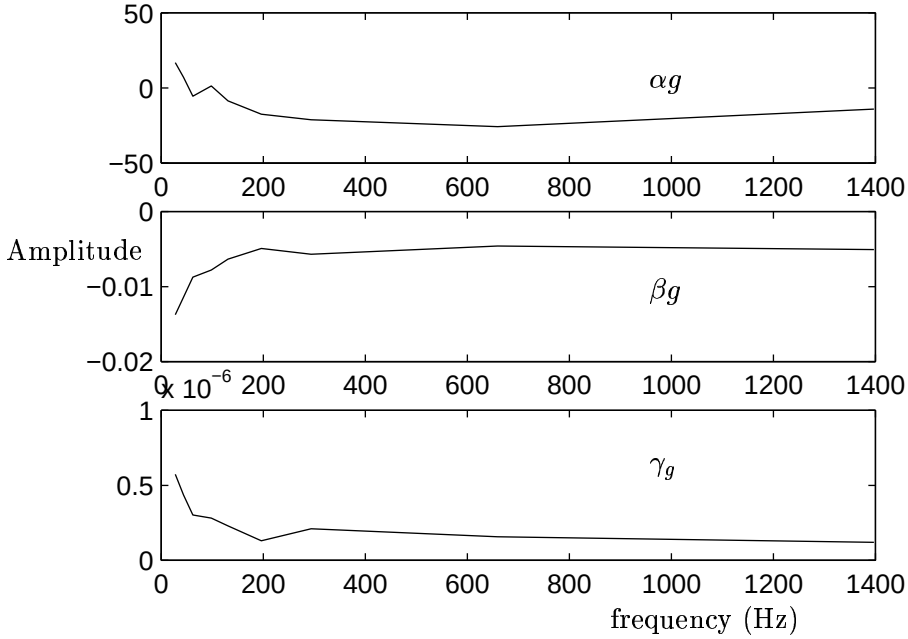

Figure 4.9: *Parameters α_g, β_g and γ_g as a function of the fundamental frequency.*

$$En = \frac{1}{T}\int_0^{Fs/2} E^2(f)df \tag{4.7}$$

$$Ba = \frac{\int_0^{Fs/2} E(f)f df}{\int_0^{Fs/2} E(f)df} \tag{4.8}$$

Where f is the frequency and Fs the sampling frequency.

In order to synthesize an excitation signal having a given energy and spectral centroid, we must express parameters g, a, b as functions of Ba and En. The centroid actually depends on two parameters, a and b:

$$Ba = \frac{\int_0^{Fs/2} Es(f)e^{af^2+bf} f df}{\int_0^{Fs/2} Es(f)e^{af^2+bf} df} \tag{4.9}$$

Thus, a given centroid corresponds to several couples of variables (a, b), and is therefore not unique. In order to solve this problem, we choose one of the two parameters as the control

parameter. The other parameter will then be determined by the corresponding relation Eq. (4.6). We have chosen the parameter b as a control parameter, since it determines the tilt of the excitation spectrum and is therefore intuitively related to the brightness. Since we can not find an analytical expression for b as a function of Ba, we numerically calculate the solution and store it in a table.

Knowing b and a, we can calculate g from the energy En.

$$g = \frac{1}{2} log\left(\frac{E_n T}{\int_0^{Fs/2} Es^2(f) e^{2af^2+2bf}}\right) \tag{4.10}$$

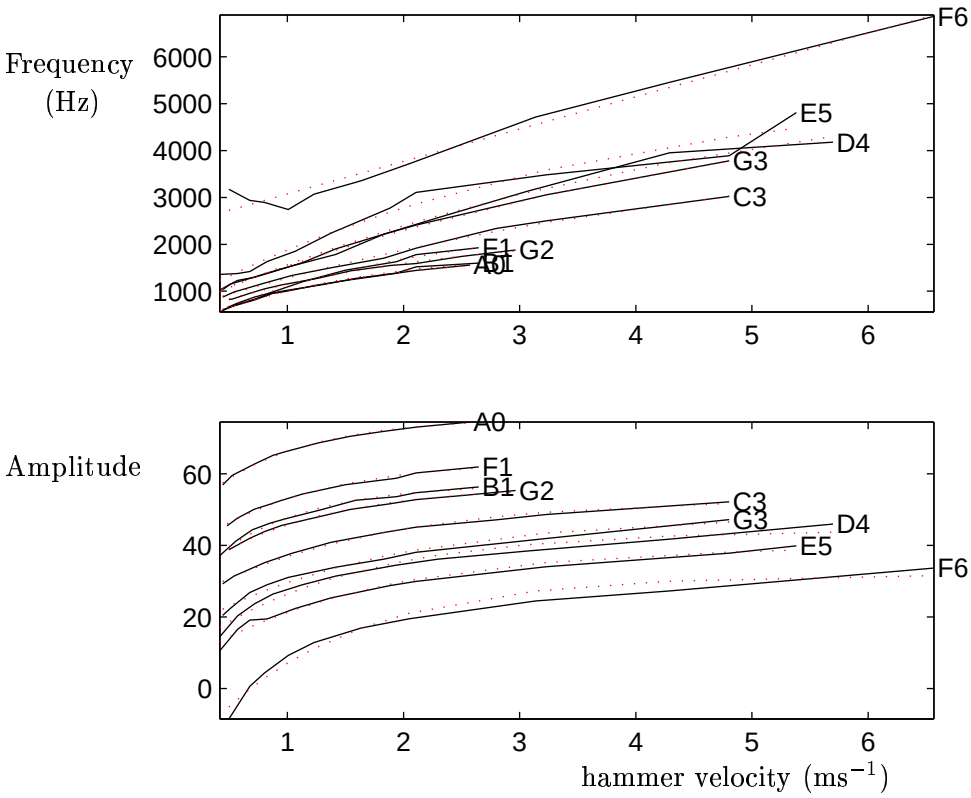

Figure 4.10: *Spectral centroid (top) and energy (bottom) for several notes as a function of the hammer velocity. Original (plain) and modeled (dotted).*

The behavior of Ba and En as a function of the hammer velocity will determine the dynamic range of the instrument and must be defined by the user.

Since the evolution of the parameters a as a function of the note is small, we choose the same relationship for all notes to model its behavior. Figure 4.10 shows the behavior of the spectral centroid and the energy for several notes. The curves have similar behavior

and only differ by a multiplicative constant. We have fitted their asymptotic behaviour by an exponential model, similarly to what have been done with functions Eqs. (4.6). These functions will be applied to the synthesis of each excitation signal and will then characterize the dynamic range of the virtual instrument. It will be easy for the user to change the dynamic range of the virtual instrument, which may be modified by the user by changing the shape of these functions.

Calculating the excitation signal is then done as follows: to a given note and velocity, we associate a spectral centroid Ba and an energy En (using the asymptotic exponential fit) and a coefficient a (thanks to relation Eq. (3.61)); b is then obtained from the spectral centroid (from a table)and g from the energy (Eq. (4.10)). One finally gets the spectral deviation which, multiplied by the static spectrum allows the excitation signal to be calculated.

4.5 Damper model

The dampers allow for a fast decay of the vibration after the release of the key (see Section 1.2.5). The damping is related to the string/damper mass ratio. It is weaker for bass strings than for treble ones. This damping also depends of the partial for a given note. Figure 4.11 shows the behavior of the amplitude of the first sixth partials of the note A0, before and after the fall of the damper. The number indicates the estimated damping coefficient.

We note that this coefficient varies depending on the partial index. This difference is related to the position of the damper on the string. The damper can not efficiently damp a partial having a node of vibration at the damper position. Moreover, we see that non-linear phenomena appear at the damper-string contact. When falling, the damper gives energy to some modes, causing the corresponding partials (the second, third and fourth on Figure 4.11) to temporarily increase before being damped.

To take the damping increase into account, we have cascaded another filter to the loop filter. We have estimated this filter from the analysis of experimental data. The dashed curves on Figure 4.11 correspond to the modeled damping. But, to reduce the computation time, we finally chose a real coefficient for this filter. The damping will then be the same at every frequencies. This real coefficient varies from 0.75 to 0.9 for the note A0 to A5. This simple model is accurate enough from a perceptive point of view. This ensures that the lower notes will damp more slowly.

The non-linear phenomena at the beginning of the decay can be well modeled by injecting a new excitation in the loop filter. This excitation can be obtained by deconvolution between

the original damped sound and the synthesized one. Nevertheless, this method has not been used in the real-time implementation of the model, since the phenomena is not of great importance from a perceptual point of view.

4.6 Soundboard model

The soundboard has not been modeled in this study. We have mainly worked on the strings and their interaction with the hammer. Nevertheless, the filtering of the soundboard greatly modifies the timbre of the instrument. We therefore need to take it into account in our piano synthesis model. Studies presented by Aramaki et al. [4] show that it is possible to simulate the soundboard contribution using a set of linear filters. Those filters have been identified for each note from experiments on a real piano. The Disklavier grand piano was placed in an anechoic room. The string vibration at the bridge level and the radiated acoustic pressure at the location of the ears of the pianist were simultaneously measured, using an accelerometer and a dummy head. The authors have then calculated the transfer functions which describe the wave propagation from the bridge, through the soundboard and air, to the right and left ears of the pianist. We will use an average of both transfer functions for each note, to simulate the perceptual effect of the soundboard.

An efficient way to compute the effect of the resonating body (i.e. the soundboard) was proposed by Smith [99] and is called "commuted synthesis". The string and the body are considered as linear time-invariant (LTI) system and they can be consequently commuted. The filter corresponding to the soundboard can be convoluted with the excitation signal to create a wave table of "aggregate excitations". The resulting signal is then injected in the waveguide model to synthesize the sound final sound. This technique reduces the computational costs since the aggregate excitation can be computed a priori. We have used this method in Section 5 to pre-calculate the aggregate excitation signals corresponding to all notes and all MIDI velocities.

4.7 Conclusion

The model we presented in this chapter allows an accurate resynthesis of piano tones. We used a source-resonator synthesis model to simulate the string vibration and the hammer string-interaction. The resonator constitutes a set of digital waveguide filters. Their parameters have been estimated from measurements on a real piano. The behavior of those parameters as a function of the note has also been modeled. The source is simulated using a substractive synthesis signal model allowing to reproduce the excitation signal for a given note and a

given velocity. To get an intuitive control of the source model, we exhibited two perceptual parameters: the energy and the spectral centroid.

Thus, we have shown that it is possible to synthesize the tone of a given piano, using a hybrid model. However, at this stage, the algorithm is not real-time compatible. It is computed in the frequency domain, by multiplying the excitation signal by the transfer function of the resonator. To make a digital piano, the previous model has to be adapted for a real-time implementation. We show in the next Chapter that for this purpose, it is necessary to implement in the time domain.

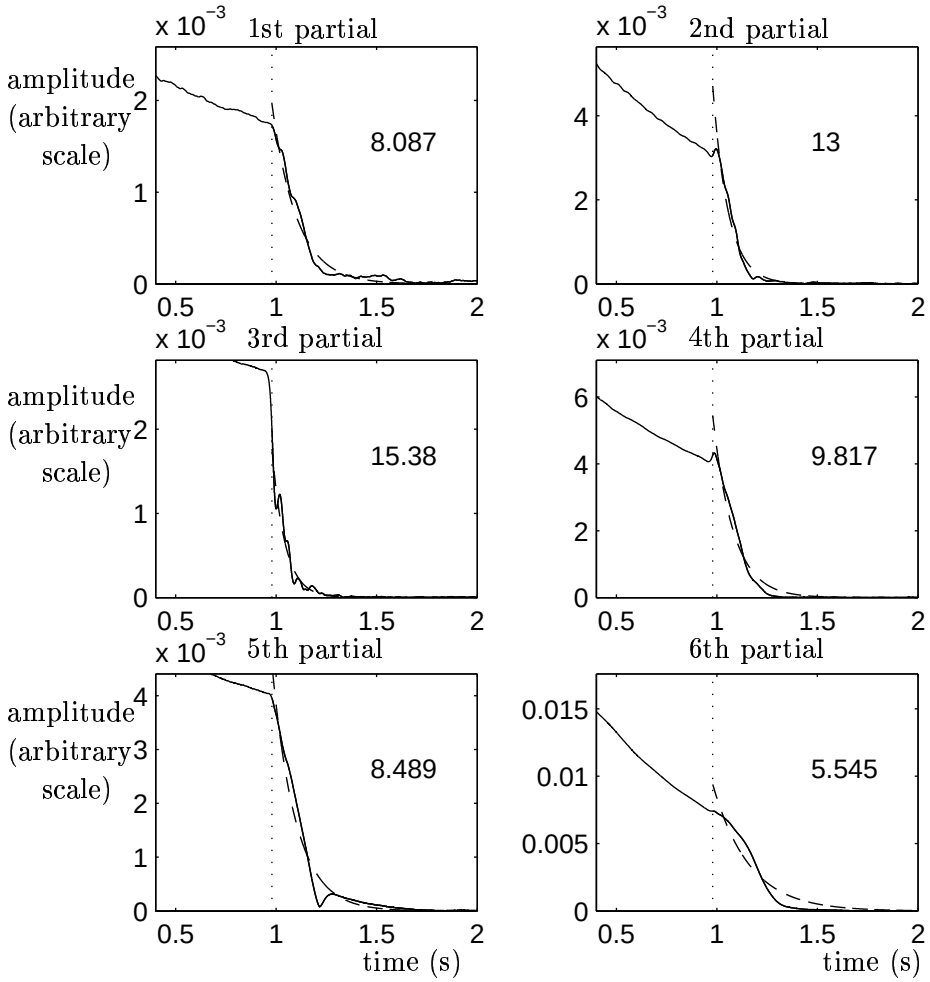

Figure 4.11: *Amplitude of the first 6 partials of the note A0 as a function of the time. The vertical dotted line indicates instant of the lift of the damper. The dashed line corresponds to the modeled damping. The number corresponds to the damping coefficient, after the lift of the damper.*

Chapter 5

Real-time implementation

5.1 Introductory remarks

In this section, we present the methods used for implementing the real-time piano model presented in the previous section. The complexity of the model, it's control and the calculation limitations of computers have constrained us to make a certain number of choices and simplifications, to render this implementation possible. In doing this, we were required to leave the context of analysis-resynthesis for that of synthesis. We have however, at each stage of this implementation, compared the synthesized sound with the original sound in order to stay true to the latter.

The control of the synthesis model is done in two successive stages. In the first stage, in non-realtime, an interface permits us to change the physical parameters of the strings, in a global or detailed (individual) way, as well as the perceptual parameter related to the source, resulting in the hammer string interaction. Regarding the issue of the choice in parameters, the modulus and phase of the filters as well as the excitations are calculated. For reasons that we have developed in Section 5.3.1, it was necessary to implement the piano model presented in the previous sections in the frequency domain, in the time domain. It is necessary to calculate the FIR and IIR filters approximated by modulus and phase given in the frequency domain.

In the second step, the software MAX/MSP uses these quantities in real time, to calculate the operation necessary to the synthesis (convolution, delay lines etc.). The control parameters are therefore, the number of the note played, its begin and end time, as well as it's velocity. The instrumentalist uses a traditional MIDI keyboard, and sends this playing information to the software MAX. This then synthesizes the corresponding sound starting from the synthesis model, and restores it with the help of an amplifier and a loudspeaker.

5.2 Control of the resonator and the source parameters in the frequency domain

5.2.1 Resonator interface

The model of the resonator, presented in Section 2.5 and 4.3, is extremely parameter rich. Given a note, the resonator can consist of one, two, or three coupled digital waveguides. Each digital waveguide is calculated according to the length L of the string (or the fundamental frequency ω_0), the tension T, the loss coefficients b_1 and b_2, the inharmonicity factor B and finally the parameters of the coupling filters in the case of multiple strings. We have approximated the modulus and phase of the filters by simple linear curves as a function of frequency. From a perceptual point of view, this was sufficient to simulated the beating and double decay of the modes. Finally, the control of the model involves 1727 parameters.

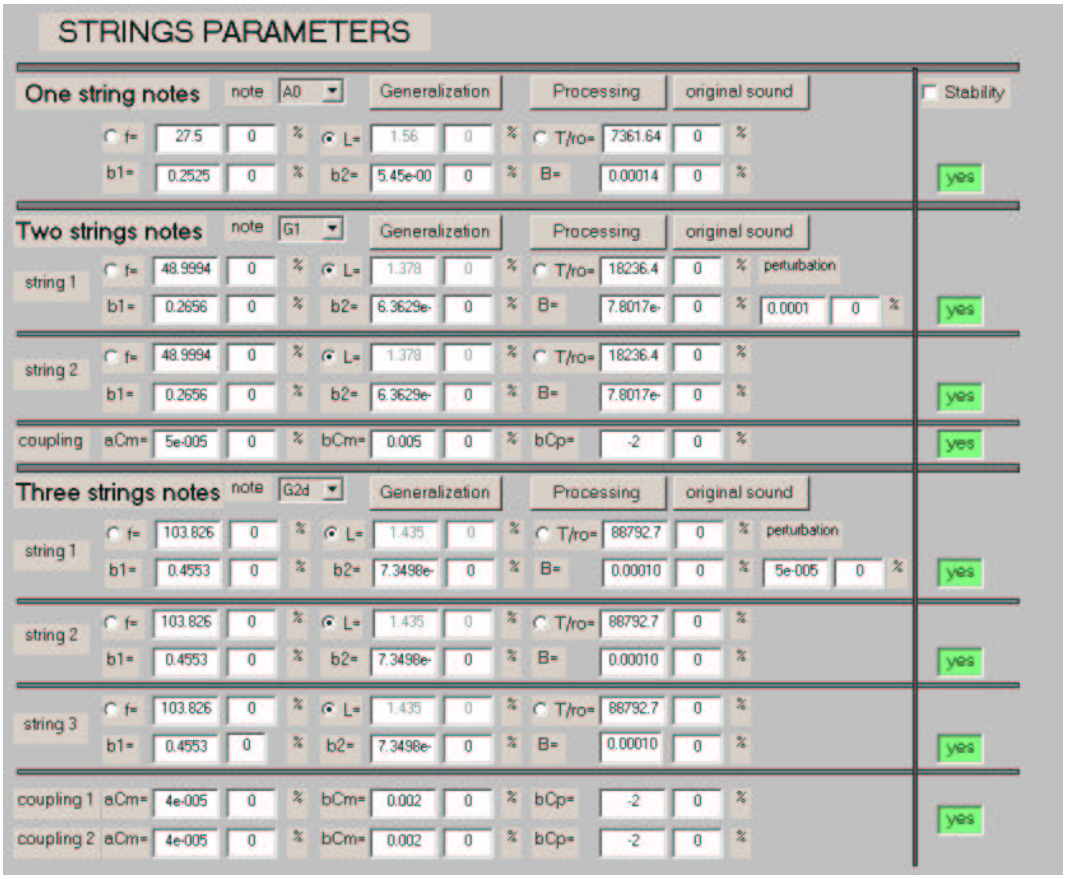

Figure 5.1: *Interface allowing the control of the resonator parameters*

Constructing ones own piano could involve adjusting one by one each of these parameters. This operation is extremely difficult to realize, and global modifications of the parameters should therefore be possible. We may, for example, want to change the decay time or the inharmonicity of the whole keyboard range with one sole operation. In order to resolve this problem of control, we have developed an interface (represented in Figure 5.1), using the software Matlab, that permits the user to choose the parameters of the resonator, in either a global way or on a individual basis.

It is first possible to change the parameters of an existing resonator. Many pianos are available according to the type of timbre that is desired. The keyboard range is divided into three sections corresponding to notes with one, two and three strings. The user selects from a section a note whose parameters he wishes to modify. The interface then calculates in the frequency domain the parameters of the corresponding digital waveguide and plays the obtained sound for a medium excitation (or given by the control interface of the source, see Section 5.2.2). He can also listen to the original sound measured on a grand piano, to compare them. Once he is satisfied with the obtained sonority, he can either chose to modify another note, or generalize the modifications to all the notes in the section. These modifications can then be saved to a file. Excessive modification of the parameters can produce sounds that are very far from the original, and very unrealistic in the case of the piano. Certain modifications can change the causality and the stability of the model. In this case, the interface indicates to the user that the chosen parameters produce an unstable model, but will still calculate the corresponding sound.

5.2.2 Source interface

In the same way, we have developed an interface that permits modification of the source. For each note, it is possible to adjust the brightness and the energy of the source. Global modifications, by section, are equally possible. The interface calculates the corresponding excitation signal (in the frequency domain) and multiplies it by the transfer function of the resonator (obtained by the resonator interface, see Section 5.2.1). When the user is satisfied by the obtained timbre, he can generalize the modifications and calculate the group of excitation signals for the whole keyboard range.

5.3 Time domain synthesis

5.3.1 why realize synthesis in real-time?

The digital wave guide model can be implemented in either the time or frequency domain. It consists of convolving (or multiplying) an excitation signal with an impulse response of long

duration (or by a transfer function). The wave guide model parameters are estimated in the frequency domain. It would be advantageous to realize this synthesis in the frequency domain and then obtain the time domain signal by discrete Fourier transform. This would allow the preservation of high accuracy resynthesis, as much as possible, in the frequency domain, to choose a loop filter modulus and phase as complex as is needed. But, the implementation of the model in the frequency domain also presents significant difficulties. The Max-MSP environment provides, for real-time synthesis, the objects " fft" and " ifft," which can process buffers of 4096 points maximum, which, in general, are too few to process the duration of an entire note. It is necessary, then, to perform this convolution in several blocks. Only one buffer of data at a time can be modified while processing, so the parametric modifications can only be performed at time intervals corresponding to the duration of this buffer. To obtain a relatively high parameter refresh rate, the data buffer size must be reduced, which lessens the frequency resolution of the buffer. There exist algorithms for dividing long signals into shorter blocks to be processed sequentially ("overlap-add") and even to remove the inherent enter-exit latencies in block processing. However, after having tested certain among these algorithms, there were weaknesses, notably in phase continuity. Therefore, we have decided to realize the synthesis in the time domain. The loop filter, then, being estimated in the frequency domain, the first difficulty encountered was to find some approximate expressions for the modulus and phase in the time domain. We have chosen to treat separately the modulus and phase filter allowing the possibility of controlling them independently. We will see that this separation is no more costly in calculation time, because it allow for the reduction of the number of coefficients of each of the filters used. The modulus is then approximated by a linear phase FIR filter, while the phase is approximated by a cascade of all-pass IIR filters.

5.3.2 Modulus approximation

The amplitude of the loop filter is simulated using an FIR filter whose phase is linear. In order to reduce the cost of real-time computation, the number of filter's coefficients needs to be the smallest possible. To find optimal filter's coefficients, we used the $firls$ function from Matlab. For a given filter, this function allows to find the coefficients of a linear phase FIR filter whose amplitude is closest to the one of the original (in the least mean square sense, using the norm $\mathcal{L}^2$) and for a number of coefficients which are predetermined. In order to increase precision at high frequencies we also added a weight function. The number of chosen coefficients depends on the note played. We used 17 coefficients for the first 10 notes and 9 coefficients for the following. The firls function imposes an odd number of coefficients, in order to have amplitude zero at Nyquist. Figure 5.2 shows the original amplitude of the loop

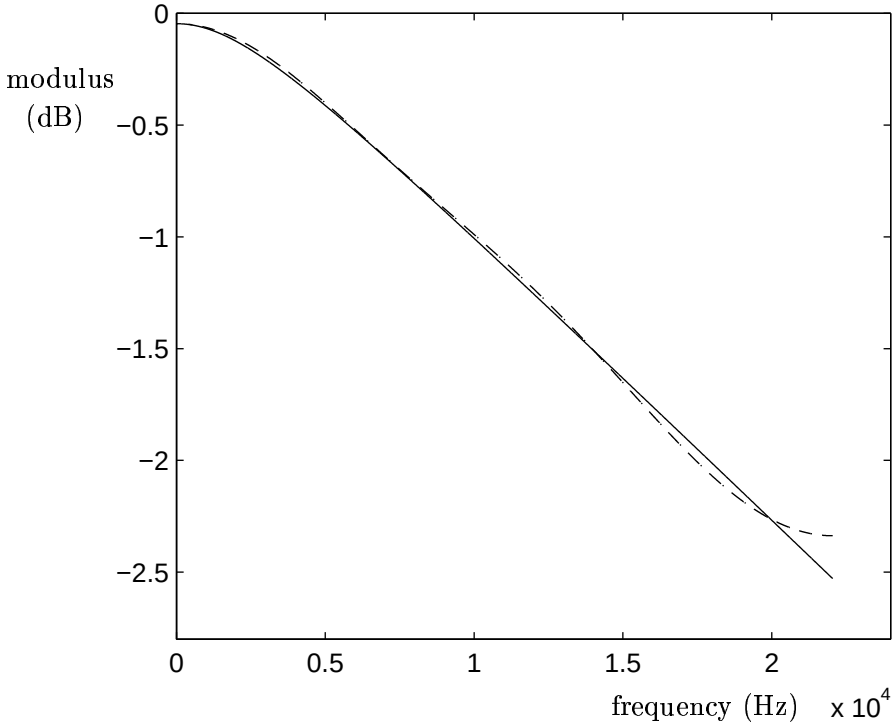

Figure 5.2: *Modulus of the loop filter original (plain) and fitted using a FIR filter (dotted).*

filter and the approximated one using the FIR filter.

Since the phase is linear, their impulse response is symmetric. It is therefore possible to reduce the number of operations. Given the following impulse response:

$$H(z) = \sum_{k=0}^{N} a_k z^{-k} \tag{5.1}$$

Since the impulse response is symmetric, the coefficients are redundant. For N even, we can write $a_k = a_{N-k}$ with $k \neq N/2$ so:

$$H(z) = \sum_{k=0}^{\frac{N}{2}-1} a_k (z^{-k} + z^{-(N-k)}) + a_{\frac{N}{2}} z^{-\frac{N}{2}} \tag{5.2}$$

The number of MACS (Multiplication Accumulation) becomes from N+1 for Eq. (5.1) to N/2+1 for Eq. (5.2). For $z = exp(i\omega)$, Eq. (5.2) can be written:

$$H(e^{i\omega}) = e^{-i\omega\frac{N}{2}} \left(\sum_{k=0}^{\frac{N}{2}-1} 2a_k cos\omega(k - \frac{N}{2}) + a_{\frac{N}{2}} \right) \tag{5.3}$$

Notice that the filter FIR creates a phase delay τ_{FIR} of N/2 samples when N is even, and N/2+1 when N is odd. It is necessary to account for this delay when calculating the total filter.

5.3.3 Phase approximation

Thick piano strings create a dispersion when waves propagate, leading inharmonic partials. In terms of waveguides, this inharmonicity is accounted for in the phase of the loop filter. As seen in section 2.4, this phase is non linear. In order to reproduce in the temporal domain the phase of the loop filter (which comes from Eq. (2.27a)), we use a cascade of allpass filters. The allpass filter allows to approximate a given phase, without modifying the amplitude which is already accounted for by the FIR filter. In real-time simulations we need to limit the number of filters' coefficients. If it is possible to simulate inharmonicity in guitars, violin and cello which is rather weak, it is hard, in the case of the piano, to obtain the exact filters' phase for a given number of coefficients. Using a cascade of allpass filters allow to simplify the optimization process. The original phase is divided by the number of allpass in the cascade. So the phase behavior is more linear, allowing the optimization procedure to work more efficiently in a reduced number of coefficients.

In order to find the coefficients of the allpass filter that better approximate the phase, we use an optimization procedure proposed by M. Lang [70] and already used by Serafin in [90]. Calling k_{AP} the number of allpass coefficients in cascade, and given an N-order all-pass filter whose transfer function is:

$$H(z) = \frac{z^{-N}P(z^{-1})}{P(z)} \tag{5.4}$$

with $P(z) = \sum_{k=0}^{N} a_k z^{-k}$ and $a_0 = 1$. The phase response is given by:

$$\phi_{AP} = arg(H(e^{i\omega})) = -N\omega + 2\ arctan\ \frac{\sum_{k=0}^{N} a_k sin(k\omega)}{\sum_{k=0}^{N} a_k cos(k\omega)} \tag{5.5}$$

As proposed by Lang, we minimize (using a Remez algorithm) the L_∞-norm of the weighted phase error between the phase ϕ_F of the original loop filter and the phase ϕ_{AP} of the all-pass filter

$$min_{a_k,\tau_0} \| W(\omega)[\phi_F/k_{AP} - \tau_0\omega - \phi_{AP}] \|_\infty \tag{5.6}$$

$W(\omega)$ is the weighted function, allowing us to give more importance to low frequencies. $\tau_0\omega$ is the pure delay that has to be removed from ϕ_F/k_{AP} to fit the phase of the all-pass filter. Figure 5.3 shows phases ϕ_F and $k_{AP}(\phi_{AP} + \tau_0\omega)$ for a note F4#, using three all-pass filters of third degree. Figure 5.4 shows the corresponding frequency error.

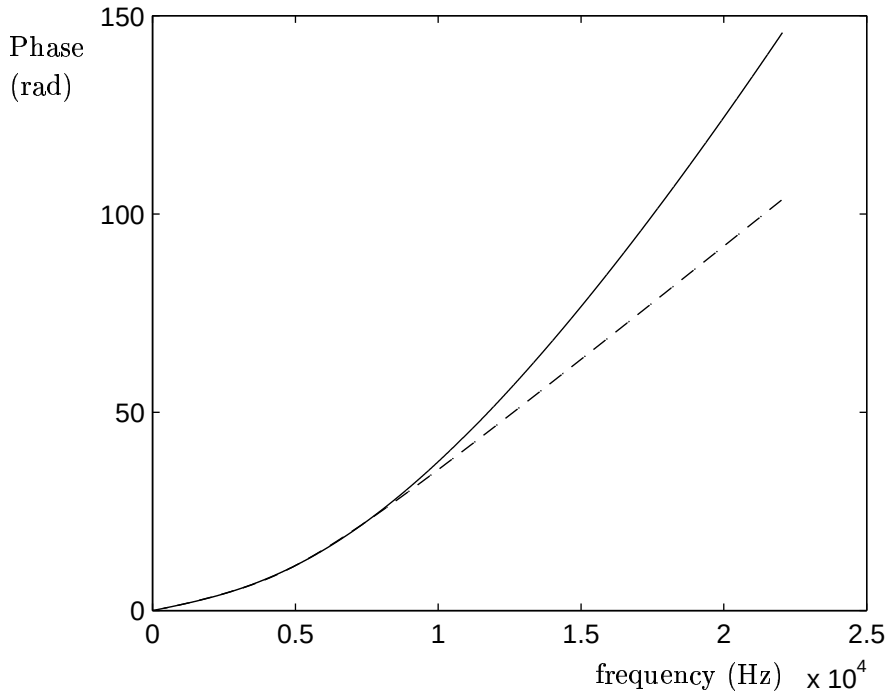

Figure 5.3: *Phase of the loop filter original (plain) and fitted using a cascade of IIR filter (dotted).*

We notice that the phase is well approximated only for low frequencies. For middle and high frequencies, the inharmonicity of the synthesized sound will be much weaker than the one of the original signal. To get a better approximation, we could use more all-pass filters or increase their degree. But for real-time application, we have to use as few filters of the lowest degree as possible. We have then compared the synthesized signal to the original one and proceeded to informal perceptual tests. We concluded that for the first 10 notes of the keyboard range (corresponding to the one string notes) we can use 4 all-pass filters of the third order to get a satisfying perception of the inharmonicity. For other notes, we use 3 all-pass filters of third order.

We also need to take into account the delay introduced by the filters' cascade. For a

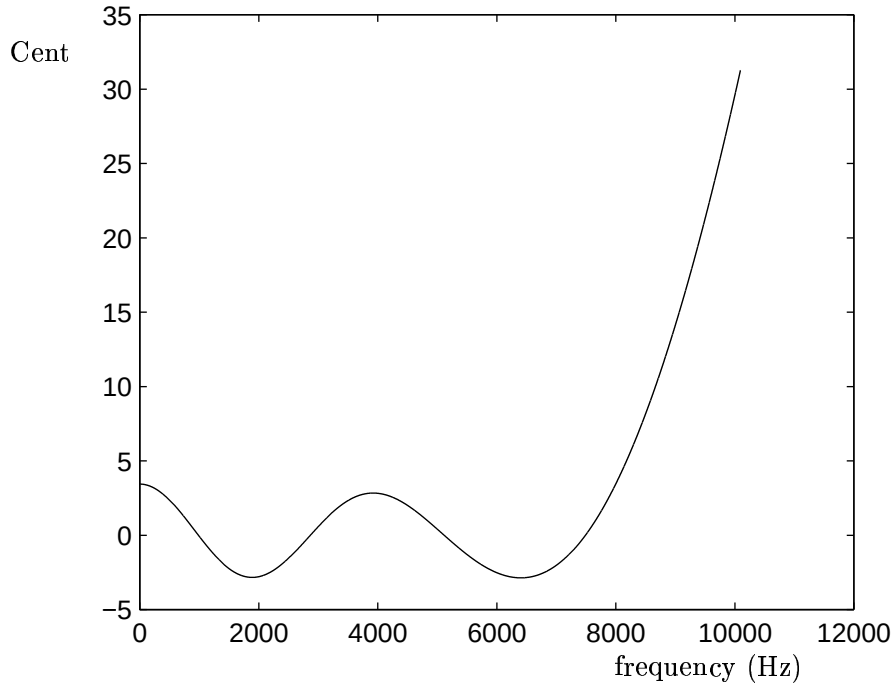

Figure 5.4: *Frequency error between original and approximated phases.*

cascade of k_{AP} allpass filters, the delay introduced for the fundamental frequency ω_0 is given by: $\tau_{AP} = k_{AP}\Phi_{AP}(\omega_0)/\omega_0$. Using Eq. (5.5), we have

$$\tau_{AP} = -k_{AP}N + \frac{2}{\omega_0}\, arctan\, \frac{\sum_{k=0}^{N} a_k sin(k\omega_0)}{\sum_{k=0}^{N} a_k cos(k\omega_0)} \tag{5.7}$$

As in the case of the FIR filter with linear phase, we show that it is possible for the allpass filter to minimize the number of operations. Equation Eq. (5.4) can be written as:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{z^{-N} + \sum_{k=1}^{n} a_k z^{k-N}}{1 + \sum_{k=1}^{N} a_k z^{-k}} \tag{5.8}$$

The corresponding time equation becomes:

$$y(n) = x(n-N) + \sum_{k=1}^{N} a_k (x(n-N+k) - y(n-k)) \tag{5.9}$$

Using this formulation, it is necessary to perform N+1 MACS instead of 2(N+1).

5.3.4 Excitations calculation

The excitation is obtained using the mean spectrum, filtered with respect to the hammer velocity. We may compute this filtering in the time domain, in real-time. Nevertheless, for more efficiency, the excitation is not computed in real-time due to processor's capacity.

The MIDI protocole defines 128 different velocities. For each note, 128 excitations were calculated and stored in a table. When necessitating, those excitations are loaded, depending on the note and the dynamic and finally convoluted with the digital waveguide.

5.4 Real-time implementation

5.4.1 The real-time software environment, Max/MSP

High quality digital signal processing and sound synthesis at audio sampling rates (using 16 bit, 44100 Hz samples), once reserved for specialized DSP processors, is now easily done on the home computer. Many environments for the development of such applications are now available. The software MAX, that we used for the real-time implementation of our model, was developed by Miller Puckette and David Zicarelli for the development of control applications for instruments and MIDI synthesizers on the Macintosh platform [82]. It consists of an environment in which the applications are constructed by assembling elementary functions or objects in their graphical form: functions are represented by boxes with inlets and outlets and the data flow is given by patch chords connecting the outputs and the inputs. The maximum real-time data flow rate (on the order of 3 kbytes per second) was compatible with the computational power of microprocessors of the time. The rise of microprocessor power has permitted real-time data flow rates at digital audio rates (on the order of 100 kbytes per second per channel). An extension of the software MAX was developed by David Zicarelli [119]. This extension (Max Signal Processing, or MSP) consists of sixty objects dedicated to synthesis and digital signal processing. Similar environments have since appeared for other platforms, notably under UNIX or Linux systems (PD [81]). The functionality is identical, and the problems are of the same nature.

5.4.2 Processing of samples using blocks

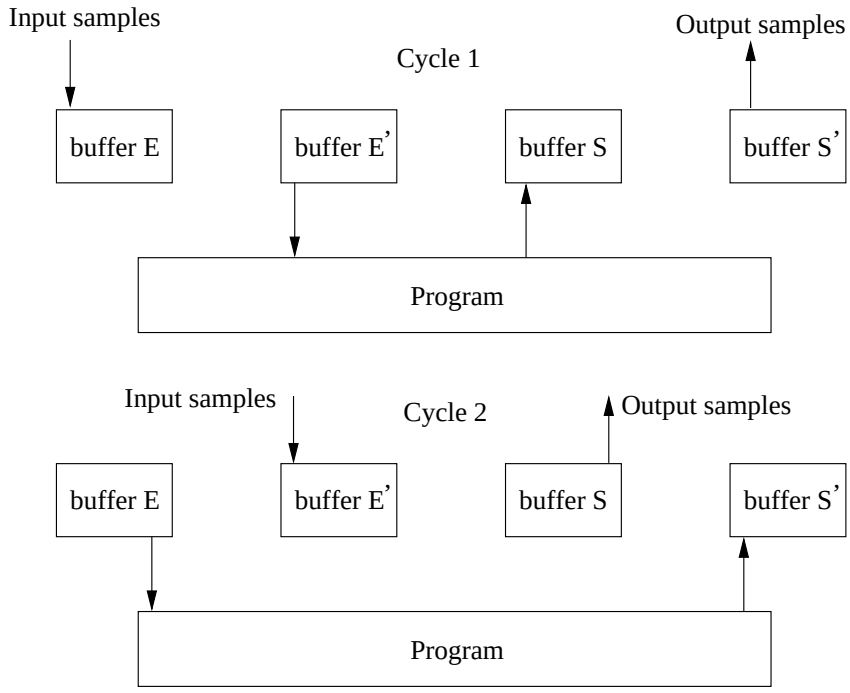

Figure 5.5: *Double buffer principle.*

Real-time synthesis refers not only to the algorithm, but to the manipulations involved in the reading of input data and the writing of output data. The first real-time synthesis machines (IRCAM's 4X, GRM/Digilog's Syter, and early DSPs, etc.) operated in a simple form of synchronization with the sampling frequency. During each sampling period (which lasts some 20 microseconds), an input sample is read from the analog/digital converter, and a sample value is sent to the digital/analog converter. The same set of operations is thus carried out at each sampling interval, a mode of operation which is far from optimal. For example, variable initialization is carried out at each cycle, resulting in a loss of valuable computational power.

An improvement over these simple schemes is known as double-buffering (see Figure 5.5). Here, operations recur every two sampling periods, and involves two sets of buffers, one for storing input samples, and the other for reading output samples, at the sampling frequency. At the same time, the synthesis program uses two other buffers, one in which input samples are read, and another in which output samples are written. When the former two buffers

are entirely full and read, respectively, the two sets of buffers switch roles with the latter set. The microprocessor is thus free to devote more of its time to system operation tasks; the only constraint is the synchronization of the buffer exchange, so as to maintain constant input and output data streams. We will not discuss the details and relative advantages of such a method, which is used in the MAX/MSP synthesis environment.

Double buffering can create problems when the algorithm to be implemented involves recursive feedback structures, such as digital waveguides. Let us consider, as an example, the waveguide model shown in Figure 5.6, constructed from basic delay and filtering objects.

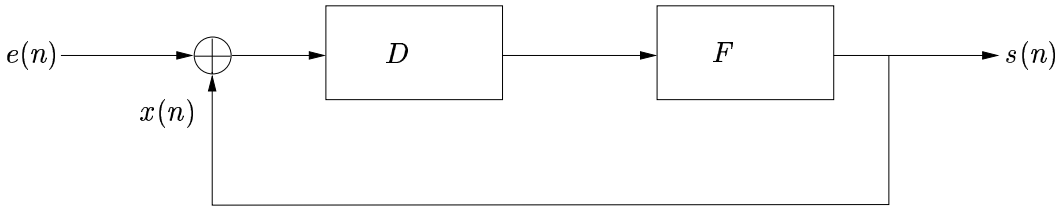

Figure 5.6: *Elementary digital waveguide.*

The calculation of output samples $s(n)$ requires knowledge of input samples $e(n)$, as well as samples $x(n)$, which are not known until the entire vector has been processed. A delay line of duration less than the vector length can thus not be realized. There is a resulting upper bound on the fundamental frequency which can be attained in real time: for a minimum vector size of 64 samples, this bound will be 690 Hz, at a sampling frequency of 44100 Hz. For the piano, encompassing a range of fundamental frequencies up to approximately 4000 Hz, this bound is not respected. The problem can thus not be solved except by employing a recursive structure inside an elementary object, where one has access at each sampling instant. We have thus created a "waveguide object" (in the C language), which is used directly by MAX-MSP.

5.4.3 Determining the delay value

The fundamental frequency of the signal generated by a waveguide model is linked to the pure delay of the loop. This loop makes use of several elements in series, each of which introduces a delay. It is necessary to take into account the delay τ_{FIR} introduced by the FIR filter, the delay τ_{AP} introduced by the all-pass cascade, and the delay D of the delay filter which allows for tuning. If D_t is the total delay of the loop, the following relation can be obtained:

$$D_t = \tau_{FIR} + \tau_{AP} + D \tag{5.10}$$

where $D_t = F_s/f_0$ with F_s is the sampling frequency and f_0 is the fundamental frequency of the string. We get

$$D = D_t - \tau_{FIR} - k_{AP}\tau_{AP} \tag{5.11}$$

5.4.4 Implementation of the fractional delay

The implementation of a delay line is very simple when the delay is made of a n integer number of samples. In this case, the delay line is made of a memory block with a read and a write pointer (using the "circular memory" technique). At each new sampling period, the two pointers are increased. The distance between the two pointers gives the delay time in samples. The delay obtained in this situation is an integer multiple of the sampling period. We therefore obtain a discretization of the frequencies which becomes audible in the middle and high registers. In order to obtain a continuous variation of the frequencies in order to use a musical scale, we need to build fractional delays.

For sampled signals, the implementation of a constant delay can be considered as the approximation of an ideal allpass filter, with an amplitude equal to one and a constant delay given by: D, $H(\omega) = e^{-i\omega D}$ so the impulse response becomes:

$$h(n) = \frac{sin(\pi(n-D))}{\pi(n-D)} \tag{5.12}$$

since the signal are sampled, therefore bandlimited. This impulse response is represented in Figure 5.7.

We notice that when the delay corresponds to an integer number of samples, the response becomes unit for $n = D$. Indeed, the cardinal sinus is zero for $n = kD$ with k integer different from 1, which justify the implementation using a circular memory. For non integer delay values, the impulse response is an infinite cardinal sinus, sampled and shifted. A non integer delay must therefore consider all the samples of the signal, each one weighted by a value of the cardinal sinus function. This function being infinite and non causal, this gives a problem. The practical solution consists of building FIR or IIR filters which approximate the ideal case. Laakso et al. [69] propose 12 methods to build FIR filters and 6 methods to build IIR. The IIR filters that show the best approximation are the allpass, i.e. with unitary amplitude response in [0, fe/2]. Usually an FIR filter does not have this property. Moreover, an IIR filter allows to obtain the desired approximation (constant delay in a frequency band) with less operations (multiplications) than an FIR filter. The implementation of a fractional delay using an allpass IIR seems the best choice, in terms of computational cost and artifacts on the amplitude of the frequency response.

We can therefore build a fractional delay using the cascade of allpass filters used to approximate the phase of the loop filter. We therefore decided to obtain the fractional delay

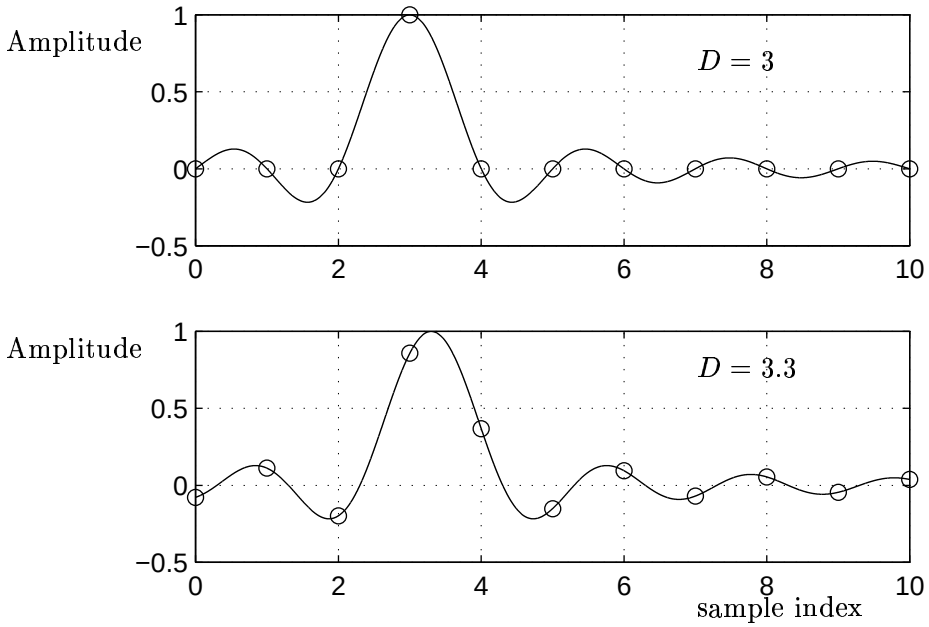

Figure 5.7: *Impulse response of an ideal linear phase all-pass filter for an integer number of samples (top) and for a fractional number of samples (bottom)*

using another allpass filter (which implies more calculations) for two reasons. First of all we want to be able to change the tuning of the waveguide, eventually in real-time, without recalculating the filters' coefficient of the allpass filters of the cascade. The second reason is that the optimization algorithm that calculated the coefficients of the cascade of allpass filters is made in the all frequency band which gives an error at all frequencies. In order to tune the loop delay at the fundamental frequency, we need to use another allpass with linear phase.

The method for calculating the coefficients ak (for a filter defined by Eq. (5.4)) of this all-pass filter depends on the type of approximation desired. The simplest method is due to Thiran [109]. It attempts to find a constant delay over the largest possible spectral region. The error is nearly zero on an interval between the DC frequency and a frequency fmax which depends on filter order, and then increases with frequency. In addition, this method leads to simple calculation, useful in the real-time context. The best approximation is obtained for a phase delay $D = N + d$ (with N integer and d fractional) and the values of the filter

coefficients of order N given by:

$$a_k = (-1)^k \begin{pmatrix} N \\ k \end{pmatrix} \prod_{n=0}^{N} \frac{D-N+n}{D-N+k+n} \quad \text{avec} \quad \begin{pmatrix} N \\ k \end{pmatrix} = \frac{N!}{k!(N-k)!} \tag{5.13}$$

The Figure 5.8 represents the value of the fractional part of the delay as a function of frequency for different all-pass filter orders (1, 2, 3, 40, 50). One remarks that the bandwidth over which the delay is constant increases more and more slowly with filter order. But increasing the filter order also increases the integer portion of the delay of the filter, and thus the maximal fundamental frequency and the calculation cost (one must recalculate N coefficients and realize an IIR filter with N+1 coefficients). It is necessary to determine the minimal order acceptable for a correct waveguide delay. We have shown, in Figure 5.8 the maximal fundamental frequency for a piano (4186 Hz). It is worth noting that the tuning of the fundamental frequency can be accomplished using a filter of order 3.

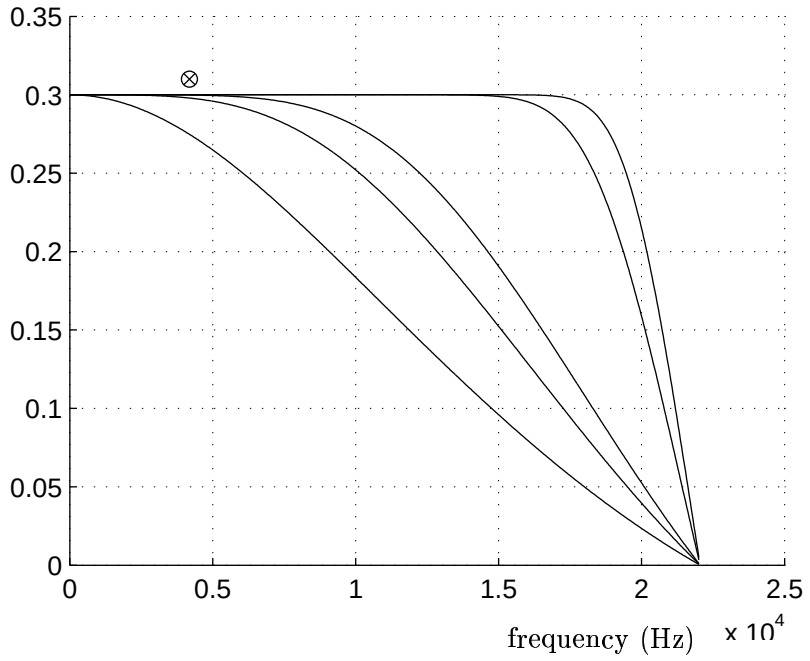

Figure 5.8: *Fractional delay for all-pass filters of orders 1, 2, 3, 20 et 40. The $\times$ in a circle indicates the maximal fundamental frequency for a piano(4186 Hz).*

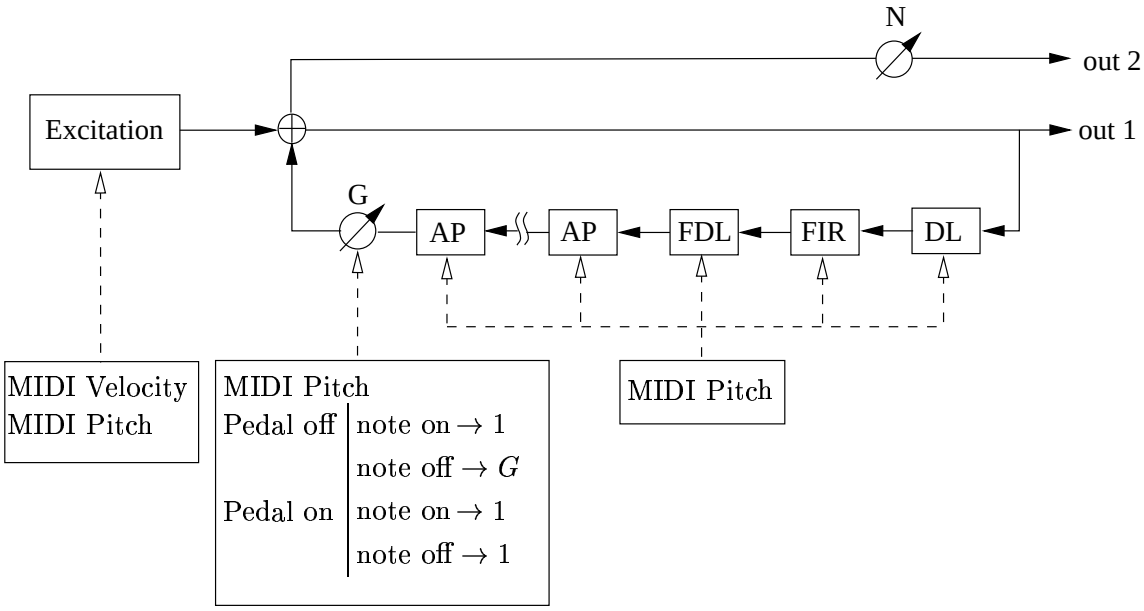

Figure 5.9: *The implemented elementary digital waveguide*

5.4.5 Implementation for a simple waveguide model

The scheme for a simple waveguide in a real-time implementation is shown in Figure 5.9. The retroactive loop makes use of a delay line to simulate the integer part of the delay, a linear-phase FIR filter simulates loss during a period (see Section 5.3.2), an IIR filter to take care of the fractional part of the delay (see Section 5.4.4), a cascade of all-pass filters to simulate dispersion during a period (Section 5.3.3) and finally a simple gain G which allows the global real-time modification of the decay rates of the partials, simulating the effect of the dampers. The filters and the delay line depend on MIDI pitch. The gain G depends also on MIDI pitch, but also on the state of the pedal. If the pedal is not in use, the gain is 1 when the note is being played, and G otherwise (see Section 4.5). If the pedal is in use, the gain is set to 1.

The waveguide shown in Figure 5.9 has an excitation which depends on MIDI velocity, and on MIDI pitch (see Section 5.3.4). There are two outputs. The output "out1" corresponds to the signal measured at the bridge, and the output "out2" is in fact the excitation signal, amplified or attenuated by the gain N. This second output allows the percussive reinforcement of the attack, and thus the ability to reproduce the key impact noise (see Section 1.2.1). The outputs "out1" and "out2" corresponding to each note are then added.

We have represented here the implementation of a simple waveguide, in the case of a single string. In the case of two or three strings, the waveguides are coupled as shown in Section 2.5.3. The coupling filters are implemented using simple phase-linear FIR filters.

5.4.6 Sympathetic resonances

In order to make the synthesized sound realistic in playing conditions, the energy transfer between the strings of different notes need to be taken into account. If they are, from a quantitative point of view, less important that the energy transfers between the strings of a single note, they may be perceptually significant. A simple, but relatively costly solution is to send the global output of the model, that is, the sum of the outputs of all the waveguides, into the set of resonators. In order to avoid stability issues, it is necessary, for a given note, to filter the output corresponding to this note. It is necessary, thus, before to sum the various outputs, to multiply them by a square matrix of size 88, with zeros only on the diagonal. The other values of the matrix will be functions of the distance between strings, at the bridge.

This type of implementation is, however, much too computationally costly, and thus impossible to use in real time. We thus choose a slightly different solution. Instead of using the matrix in order to not send the output of a waveguide into its own input, we inject this output a second time, but in phase opposition, canceling the effect of the first signal.

5.4.7 Model complexity and computation time

Since we would like a real-time implementation, the number of operations (multiplications) per second (flops) should be suitable for the computer we are using. By considering that the set of waveguides in the piano range are simultaneously in use, and coupled as described in 5.4.6, without taking into account the filtering calculations for the soundboard, we estimate that the algorithm requires slightly more than 200 million flops, for a sampling frequency of 44100 Hz. This is a large figure, and can be reached with difficulty on todays computers. We have thus made some simplifications. Only waveguides corresponding to notes which are not damped will be permanently active, as these notes resonate sympathetically, even if they are not directly played. For the other notes, only played notes will be active. This can cause problems when the pedal is down: all the strings are susceptible to sympathetic vibration. In this case, we activate the waveguides whose modes are closest to those of other notes, in other words, those which are most coupled.

5.5 Conclusion

We have presented in this chapter the real-time implementation of the model, in the time domain. The filters have been approximated by a set of FIR and IIR filters which reproduce, at least for the first half of the spectrum, the magnitude and phase characteristics of the loop filters in the waveguides. We have coupled these waveguides in such a way as to reproduce, in real-time, beating and double-decay phenomena. We have also taken into account the damping of partials due to the dampers, sustain pedal effects and sympathetic resonance, by using a set of filters identified with measurements (see 4.6 and [2]).

The necessary approximations for a real-time implementation forced us to abandon analysis/synthesis in favor of synthesis. It was necessary to limit the number of filters and their order, so that the model is able to function on today's computers. The synthesized sound remains piano-like, but differs from the original sound recorded on a grand piano, and which we have tried to maintain in this work. It is clear that as computational power grows, higher filter precision and model complexity will make for a greatly improved quality of sound synthesis.

Chapter 6

Conclusion

In this research we have studied diverse synthesis models of piano sound. We have developed a collection of methods for the analysis and resynthesis of real signals starting from physical and semi-physical models, allowing for the a high quality reproduction of piano sounds.

In the first chapter we precisely described our objectives: to build a synthesis model that is strongly related to the physics of the piano and to achieve a synthesis that is as close as possible to the recorded sound while still being computable in real-time. We have then briefly described the principle physical phenomena that come into play during the production of sound. Many that were perceptually important where incorporated into the model (frequency dependent loss, inharmonicity, non-linear behaviour of the hammer string contact...). Others were neglected because they had less influence on the timbre and were too computationally expensive for real-time implementation. Likewise, we have described the principle physical models or signals found in the literature that successful reproduce the piano sound.

The second chapter was devoted to the study of models simulating the propagation of transverse waves in the strings. A partial differential equation slightly different from that of Ruiz and having only two propagating solutions was proposed. The digital waveguide model, in which a filter simulates the delay, the attenuation and the dispersion undergone by the transverse waves during their propagation was described. We have explicitly given the relationship between the parameters of the digital waveguide model and the physical parameters of the EDP. Calculating these expressions has required, for the sake of simplicity, that certain approximation be made, justified for the piano strings. If the relationship between this EDP and this digital waveguide model are valid for certain values of damping and dispersion, the waveguide model, once implemented in the frequency domain, can simulate all types of laws for damping or dispersion. Unlike a finite difference solution, the waveguide model does not introduce digital dispersion or dissipation. Its drawback however is that the string vibration can only be calculated at one point (one may, however, get additional

output points along the string at the price of one small digital filter), whereas in a finite difference solution we have access to all the discrete points along the string. We have then showed that by coupling several digital waveguides, we are capable of simulating beating and double decay phenomenon. The originality of our approach lies in the explicit resolution of the inverse problem: calculate the parameters of the coupled waveguide model from modal parameters of a recorded signal. This very fine calibration of the model allowed us to precisely resynthesize the vibration of coupled strings.

In the third chapter, we focused on the contact between hammer and string. We coupled a nonlinear hammer model with the EDP proposed in chapter 2. This system which has no analytical solution was resolved using finite difference methods. Since some of the physical parameters of the model are not easily measurable we proposed a method of calibration by optimization: an iterative process allowing the values of each of the parameters to be adjusted so that the spectrum of the synthesis will match as close as possible that of the recorded signal.

We also proposed another method for resolving the system of equations using "digital wave filters" which provides a knowledge, and thus a control, of the distribution of energy in the system. The stability of this method is assured by its construction—regardless of the sampling rate. In this method the physical model is decomposed into mass (inductance), spring (capacitance) and dampers (resistance) elements which are arranged either in parallel or in series. The appropriate change of variable allows the relationship between force and velocity for each element to be written in terms of wave variables and characteristic impedance. The expressions obtained for the relationships allow the different elements obtained using Kirchoff laws to be connected. The "digital wave filters" also connects to the waveguide model. The sound obtained from this method was not entirely satisfying. It seems that the hammer model is too straightforward to accurately simulate the hammer string interaction.

Another method consisted of using a signal model to simulate the excitation of the digital waveguide. This method presupposes that the resonator (consisting of the string and modeled using a digital waveguide) and the source (the result of the interaction between hammer and string) can be decoupled. This hypothesis is evidently false as seen in the measurement where the contact between the hammer and string lasts for several milliseconds. However this interaction is too short to be perceivable in itself. Thus, if we consider only the vibration of the transverse waves (and not the shock of the impact from touching the frame nor the longitudinal waves), and we delete the first milliseconds of the signal, the perception of the attack is not changed. It is sufficient to simulate the result of this interaction, and not the the interaction itself, to obtain the same perceptual result. By using a subtractive synthesis model, we were able to reproduce the behaviour of the excitation of the digital waveguide in

a way that provides a synthesis that is perceptually identical to the recorded piano.

Chapter 4 presents the final model, a collection of digital waveguide models and a signal model simulating the excitation. To calibrate the model, we conducted an experiment on a real piano which involved measuring the chain of elements that produce the piano sound, from the pianist's finger to the vibration that s/he perceives. The piano was placed in an anechoic chamber so that measurements were not affected by the geometry and reflections of the room. Measurements from the keyboard elements were made using a piano that could be controlled using MIDI. The string vibrations were measured using a laser vibrometer and the bridge behavior using an accelerometer. An artificial head was also placed in the position of the performer to measure the sound at that point radiating from the soundboard. These measurements allowed us to follow the evolution of the model's parameters as a function of the keyboard range and therefore to resynthesize the timbre of each note for different velocities.

In Chapter 5 we show how this model can be implemented in real-time in the time domain. The filter module of the loop was approximated using zero-phase infinite impulse response (IIR) filters. The phase of the filter loop was reproduced using a cascade of all-pass filters. The more the number of coefficients of the filters is important, the more the sound of the time-domain synthesis approached that obtained in the frequency domain. However, increasing the number of coefficients also increased the number of operations and therefore the time required to complete the simulation. It is necessary therefore to choose the optimal number of coefficients. It is here that we see a limitation in the digital waveguide for simulating loss and dispersion. The parameters of the synthesis model must be controllable from a simple interface such as a MIDI keyboard. The model was therefore implemented using the software Max/MSP. This software allows for real-time processing of a large number of inputs. The computer receives input from the MIDI controller—the excitations are calculated as a function of the speed with which the hit was pressed—and they are sent to the corresponding waveguide. Depending on the note that was played, the software simulates a transfer of energy between the strings, and calculates the resulting sound from the sound producing elements.

The "concrete" result of my thesis with therefore the implementation of a digital piano, which, unlike other commercial synthesizers, does not simply play a recorded sound, but simulated some of the physical phenomena in real-time and synthesized the corresponding vibration. It offers everyone the possibility of choosing and modifying the sound of the virtual instrument by modifying the model's parameters. This piano "simulator" also provides a useful tool to piano manufacturers.

To further improve the quality of the synthesis model and the realistic sound control, we

propose several research directions. For the real-time model, with an increase in processor computational power, some of the simplifications and approximations that we did here will no longer be necessary. The synthesis will therefore be improved. Consider the model of the piano alone. The hybrid model which associates excitation by signal model and resonator by waveguide allows for an excellent resynthesis of the recorded sound but doesn't give access to the physical parameters of the hammer. One solution would be to define another physical model of a hammer that more accurately simulates behaviour of the force on the string and then finally calibrate these parameters by optimization methods. Additionally, we did not keep track of a certain number of phenomenon in the modeling. The models simulate linear phenomenon intervening in the string and the nonlinear behaviour of the hammer-string interaction. However, a precise comparison of example signals with those of the synthesis show that some components of the spectrum are not reproduced. If we only consider the linear phenomenon, the spectrum should contain a simple series of partials, lightly inharmonic. However, the spectrum contains other components due to nonlinear phenomenon occurring in the strings and the soundboard. Listening attentively to the synthesis signal and the recorded signal brings us to think of the these components occurring essentially in the attack of the sound. It would therefore be worthwhile to simulate them to improve the synthesis.

Appendix A

The non-linear hammer

The physical hammer model we use is the same as previously, i.e. a mass and a non-linear spring obeying the power law

$$f_H = K_H y_d^p \tag{A.1}$$

But to get a standard form of a wave digital capacitor (spring) allowing to simulate the non-linear behaviour of the felt, we have to define an "equivalent" stiffness K_e which depends on the force f_H. To insure the stability of the model, we first define an energy measure for the mass and the spring and then we calculate the equivalent stiffness K_e

Springs and masses: energetic properties

Consider a spring-like object defined by the following equation:

$$f(x) = Kx \tag{A.2}$$

Here, f is a force, and x is position; both are functions of time t. $K > 0$ is the spring constant. Taking a time derivative of both sides of this equation gives

$$\frac{df}{dt} = K\frac{dx}{dt} = Kv \tag{A.3}$$

where v is velocity, and thus we may write

$$K^{-1/2}v = \frac{d}{dt}(K^{-1/2}f) \tag{A.4}$$

Defining power-normalized variables $\underline{v} = K^{1/2}v$ and $\underline{f} = K^{-1/2}f$, we then obtain

$$\underline{v} = \frac{d\underline{f}}{dt} \tag{A.5}$$

The spring constant is thus absorbed into the definitions of the variables themselves. Multiplying both sides of Eq. (A.5) by $\underline{f}$ gives:

$$p = vf = \underline{v}\underline{f} = \underline{f}\frac{d\underline{f}}{dt} = \frac{d}{dt}(\frac{1}{2}\underline{f}^2) \tag{A.6}$$

The left side of the equation can be interpreted as the instantaneous power supplied to the spring, and the right side as the rate of change of the energy

$$E = \frac{1}{2}\underline{f}^2 = \frac{1}{2}Kx^2 \tag{A.7}$$

stored in the spring. Thus all energy supplied to the spring is stored; the spring is lossless, according to such an energy measure. An important point here is that the energy E is a positive definite function of the state variables in this case, a quadratic.

Suppose, now, that we take Eq. (A.4) as our definition of a spring. We can see that all of the manipulations above still follow, even if $K = K(x, v, f, t, ..)$, as long as K is a positive function. Thus Eq. (A.4) can be viewed as the definition of a nonlinear lossless spring, if the energy is defined as per Eq. (A.7). The relation between force and velocity for a simple mass (the dual to the spring) may be defined by

$$m^{-1/2}f = \frac{d}{dt}(m^{-1/2}v) \tag{A.8}$$

for a possibly non-linear mass m.

Expression of the equivalent stiffness

Using the definition of a spring given by Eq. (A.4) we will define in this section the equivalent hammer felt stiffness K_e.

Let us return to the definition of the elastic behavior of the nonlinear felt in the piano hammer, which is

$$f_H = K_H y_d^p \tag{A.9}$$

for $y_d > 0$ and $f_H = 0$ otherwise.

Here, $K_H > 0$ and p are constants. When $y_d > 0$ (i.e., the hammer is in contact with the string), we take time derivatives to get

$$\frac{df_H}{dt} = K_H p y_d^{p-1} v_d \tag{A.10}$$

and using Eq. (A.9) gives

$$v_d = \frac{1}{p} K_H^{-1/p} f_H^{(1-p)/p} \frac{df_H}{dt} \tag{A.11}$$

Developping Eq. (A.4) in the case of a hammer stiffness K_e which depends on the force f_H, we obtain

$$v = K_e^{-1}(1 - \frac{1}{2} K_e^{-1} f_H \frac{dK_e}{df_H}) \frac{df_H}{dt} \tag{A.12}$$

Posing $C = 1/K_e$, we get by identification between Eq.(A.11) and Eq.(A.12) the non-linear differential equation

$$C(f_H) + \frac{1}{2} f_H \frac{dC(f_H)}{df_H} - \frac{1}{p} K_H^{-1/p} f_H^{\frac{1-p}{p}} \tag{A.13}$$

Using the change of variable $C = C_0 f_H^{\frac{1-p}{p}}$ this equation can be rewritten after some manipulation in the form of Eq. (A.4), with

$$K_e = \frac{1+p}{2} K_H^{1/p} f_H^{(p-1)/p} = \frac{1+p}{2} K_H y_d^{p-1} \tag{A.14}$$

which reduces to $K_e = K_H$ in the linear case, i.e., when $p = 1$. The stored energy in the nonlinear spring is thus

$$E = \frac{1}{2} \frac{1+p}{2} K_H y_d^{p+1} \tag{A.15}$$

which is again a positive function of x, when $x > 0$.

Appendix B

A non-linear and lossy wave hammer

As we have seen in Section 3.2, the hammer felt relaxation is not immediate, leading to losses and hysteretic behavior of the hammer force. In order to get a more realistic model, we can take this behavior into account using the relation Eq. (3.4) of Hunt and Crossley [63]. The hammer model has to be modified and the corresponding diagram is shown on Figure B.1. The force is now given by the relation

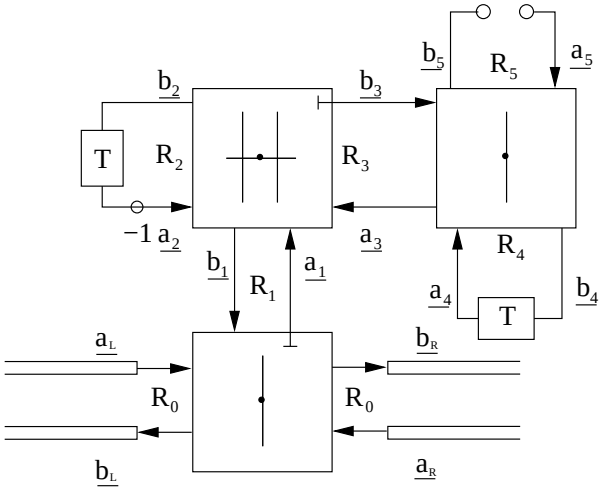

Figure B.1: *A parallel connection of a mass and a nonlinear lossy spring, top, representing a wave digital hammer, is connected in series with a digital waveguide, bottom.*

$$F_H = f_H + f_r = K_H y_d^p (1 + \mu \dot{y_d}) \tag{B.1}$$

There is then two force terms. One which is the same as previously:

$$f_H = K_H y_d^p \tag{B.2}$$

And another one which introduces loss in the spring:

$$fr = K_H \mu y_d^p \dot{y_d} \tag{B.3}$$

Using the "equivalent" spring stiffness K_e defined by Eq. (A.14), we rewrite the two force terms:

$$f_H = K_e(f_H) y_d \tag{B.4}$$

$$fr = \mu f_H \dot{y_d} = R_5(f_H) \dot{y_d} \tag{B.5}$$

The wave digital port corresponding to this non-linear lossy spring is then a capacitor with port resistance $R_4 = TK_e/2$ connected in series with a resistor with the port resistance $R_5 = \mu f$. In order to eliminate a new delay-free loop, the port admittance G_3 has to be set to

$$G_3 = G_1 + G_2 \tag{B.6}$$

$G_1 = 1/R_1$ and $G_2 = 1/R_2$ are the same as previously.

The expressions of the output normalized wave variables for the three wave digital one ports are

$$\underline{a_2} = -\underline{b_2} \tag{B.7}$$

$$\underline{a_4} = \underline{b_4} \tag{B.8}$$

$$\underline{a_5} = 0 \tag{B.9}$$

Using the equations for the scattering junction, we get for the parallel connection the relations

$$\underline{b_1} = -\underline{a_1} + \sqrt{\frac{1}{R_1}} F_H \tag{B.10}$$

$$\underline{b_2} = -\underline{a_2} + \sqrt{\frac{1}{R_2}} F_H \tag{B.11}$$

$$\underline{b_3} = \sqrt{R_3}\left(\sqrt{\frac{1}{R1}}\underline{a_1} + \sqrt{\frac{1}{R2}}\underline{a_2}\right) \tag{B.12}$$

with

$$F_H = \frac{2}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}} \left(\sqrt{\frac{1}{R_1}}\underline{a_1} + \sqrt{\frac{1}{R_2}}\underline{a_2} + \sqrt{\frac{1}{R_3}}\underline{a_3}\right) \tag{B.13}$$

For the "lower" series junction, the scattering equations are unchanged:

$$\underline{b_L} = \underline{a_L} - \sqrt{R_0} v_s \tag{B.14}$$

$$\underline{b_R} = \underline{a_R} - \sqrt{R_0} v_s \tag{B.15}$$

$$\underline{a_1} = -\frac{1}{\sqrt{2R_0}} (\sqrt{R_0}\underline{a_L} + \sqrt{R_0}\underline{a_R}) \tag{B.16}$$

with

$$v_s = \frac{2}{2R_0 + R_1} (\sqrt{R_0}\underline{a_L} + \sqrt{R_0}\underline{a_R} + \sqrt{R_1}\underline{b_1}) \tag{B.17}$$

and for the "upper" series junction, we can write

$$\underline{a_3} = \underline{b_3} - \sqrt{R_3} v_J \tag{B.18}$$

$$\underline{b_4} = \underline{a_4} - \sqrt{R_4} v_J \tag{B.19}$$

$$\underline{b_5} = \underline{a_5} - \sqrt{R_5} v_J \tag{B.20}$$

with

$$v_J = \frac{2}{R_3 + R_4 + R_5} (\sqrt{R_3}\underline{b_3} + \sqrt{R_4}\underline{a_4} + \sqrt{R_5}\underline{a_5}) \tag{B.21}$$

Bibliography

[1] J.-M. Adrien. *The missing link: modal synthesis.* Representations of Musical Signals, The MIT Press, Cambridge, Massachusetts, USA, 1991.

[2] M. Aramaki. *Analyse-synthése de sons impulsifs: approches physique et perceptive.* PhD thesis, Université de la Méditerranée, France, 2003.

[3] M. Aramaki, J. Bensa, L. Daudet, P. Guillemain, and R. Kronland-Martinet. Resynthesis of coupled piano string vibrations based on physical modeling. *Journal of New Music Research*, 30(3):213–26, 2002.

[4] M. Aramaki, Ph. Guillemin, N. Hoeb, and A. Roure. From piano string vibrations to the acoustic field radiated by the soundboard. In *Proceedings International Computer Music Conference, Havana*, pages 111–114, Cuba, 2001. Computer Music Association.

[5] A. Askenfelt, editor. *Five Lectures on the Acoustics of the Piano.* Royal Swedish Academy of Music, Stockholm, 1990. Lectures by H. A. Conklin, Anders Askenfelt and E. Jansson, D. E. Hall, G. Weinreich, and K. Wogram. Sound example CD included. Publication number 64. Available online at http://www.speech.kth.se/music/5_lectures/.

[6] A. Askenfelt and E. V. Jansson. From touch to string vibrations. part I. timing in the grand piano action. *J. Acoust. Soc. Amer.*, 88(1):52–63, 1990.

[7] A. Askenfelt and E. V. Jansson. From touch to string vibrations. part II. the motion of the key and hammer. *J. Acoust. Soc. Amer.*, 91(5):2383–2393, 1991.

[8] A. Askenfelt and E. V. Jansson. From touch to string vibrations. part III. string motion and spectra. *J. Acoust. Soc. Amer.*, 93(4):2181–2195, 1993.

[9] F. Avanzini, B. Bank, G. Borin, G. De Poli, and D. Rocchesso. Musical instrument modeling: the case of the piano. In *Proc. of the worshop on current research directions in computer music*, pages 124–125. MOSART Reaserach training network, 2001.

[10] F. Avanzini and D. Rochesso. Modeling collision sounds: non linear contact force. In *Proceedings of the COST G6 Conference on Digital Audio Effects*, 2001.

[11] R. Bacon and J. Bowsher. A discrete model of a struck string. *Acustica*, 41:21–7, 1978.

[12] B. Bank. Physics-based sound synthesis of the piano. Master's thesis, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, 2000.

[13] J. Beauchamp. Synthesis by spectral amplitude and "brightness" matching of analyzed musical instrument tones. *J. Audio Eng. Soc.*, 30(6):396–406, 1982.

[14] J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. O. Smith III. The simulation of piano string vibration: from physical model to finite difference schemes and digital waveguides. *To be published in J. Acoust. Soc. Amer.*

[15] J. Bensa, F. Gibaudan, K. Jensen, and R. Kronland-Martinet. Note and hammer velocity dependance of a piano string model based on coupled digital waveguides. *Proceedings International Computer Music Conference, Havana*, pages 95–98, 2001.

[16] J. Bensa, K. Jensen, R. Kronland-Martinet, and S. Ystad. Perceptual and analytical analysis of the effect of the hammer impact on piano tones. *Proceedings International Computer Music Conference, Berlin*, pages 58–61, 2000.

[17] S. Bilbao. *Wave and scattering methods for the numerical integration of partial differential equations.* PhD thesis, Stanford University, 2001.

[18] S. Bilbao, J. Bensa, R. Kronland-Martinet, and J. O. Smith III. The wave digital piano hammer: a passive formulation. In *Proceedings of the 144th meeting of the Acoustical Society of America*, Cancun, Mexico, 2002. CD-ROM.

[19] G. Borin, G. De Poli, and D. Rochesso. Elimination of delay-free loops in discrete-time models of nonlinear acoustic system. *IEEE Transaction on Speech and Audio Processing*, 8(5):597–605, 2000.

[20] X. Boutillon. Model for piano hammers: Experimental determination and digital simulation. *J. Acoust. Soc. Amer.*, 83(2):746–754.

[21] X. Boutillon. Le piano: Modelisation physiques et developpements technologiques. In *Congres Francais d'Acoustique Colloque C2,supplement au numero 2 Tome 51*, pages 811–820, 1990.

[22] M. Le Brun. Digital waveshaping synthesis. *J. Audio Eng. Soc.*, 27(4):250–266, 1979.

[23] C. Cadoz, A. Luciani, and J.-L. Florens. Responsive input devices and sound synthesis by simulation of instrumental mechanisms: the cordis system. *Computer Music Journal*, 8(3):60–73, 1983.

[24] S. Cavaliere and A. Piccialli. *Granular synthesis of musical signals.* Musical Signal Processing, Swets Zeitlinger, Lisse, the Netherlands, 1997.

[25] A. Chaigne. On the use of finite differences for musical synthesis. Application to plucked stringed instruments. *J.Acoust.*, 5(2):181–211, 1992.

[26] A. Chaigne and A. Askenfelt. Numerical simulations of struck strings. I. A physical model for a struck string using finite difference methods. *J. Acoust. Soc. Amer.*, 95(2):1112–8, Feb. 1994.

[27] A. Chaigne and A. Askenfelt. Numerical simulations of struck strings. II. Comparisons with measurements and systematic exploration of some hammer-string parameters. *J. Acoust. Soc. Amer.*, 95(3):1631–40, Mar. 1994.

[28] A. Chaigne and V. Doutaut. Numerical simulations of xylophones.I. time-domain modeling of the vibration bars. *J. Acoust. Soc. Amer.*, 101(1):539–557, January 1997.

[29] J. Chowning. The synthesis of complex audio spectra by means of frequency modulation. *J. Audio Eng. Soc.*, 21(7):526–34, 1973.

[30] H. A. Conklin Jr. Design and tone in the mechanoacoustic piano. part I. Piano hammers and tonal effects. *J. Acoust. Soc. Amer.*, 99(6):3286–3296, 1996.

[31] H. A. Conklin Jr. Design and tone in the mechanoacoustic piano. part II. Piano structure. *J. Acoust. Soc. Amer.*, 100 (2):695–708, août 1996.

[32] H. A. Conklin Jr. Design and tone in the mechanoacoustic piano. part III. Piano strings and scale design. *J. Acoust. Soc. Amer.*, 100(3):1286–1298, 1996.

[33] H. A. Conklin Jr. Generation of partials due to non-linear mixing in stringed instrument. *J. Acoust. Soc. Amer.*, 105(1):536–545, 1999.

[34] J. le Rond d'Alembert. Investigation of the curve formed by a vibrating string, 1747. In R. Bruce Lindsay, editor, *Acoustics: Historical and Philosophical Development*, pages 119–123. Dowden, Hutchinson & Ross, Stroudsburg, 1973.

[35] S. A. Van Duyne. Low piano tones: Modeling nearly harmonic spectra with regions of FM. In *Proceedings International Computer Music Conference, San Jose*, pages 427–428. Computer Music Association, 1992.

[36] A. Fettweis. Digital filters related to classical structures. In *AEU: Archive für Electronik und Übertragungstechnik*, volume 25, pages 79–89, 1971.

[37] A. Fettweis. Wave digital filters: theory and practice. In *Proc. IEEE*, volume 74, pages 270–327, 1986.

[38] H. Fletcher, E. D. Blackham, and R. Stratton. Quality of piano tones. *J. Acoust. Soc. Amer.*, 34(6):749–61, 1961.

[39] N. Fletcher and T. Rossing. *The physics of musical instruments.* Springer-Verlag, New York, 1991.

[40] N. H. Fletcher and T. D. Rossing. *The physics of musical instruments*, page 756. Springer-Verlag, New York, USA, 1998.

[41] D. Gabor. Theory of communication. *J.Inst.Elec.Eng.*, 93:429–57, 1946.

[42] W. G. Gardner. Efficient convolution without input-output delay. In *presented at the 97th convention of the Audio Engineering Society*, San Francisco. Preprint 3897.

[43] G. E. Garnett. Modeling piano sound using digital waveguide filtering techniques. In *Proceedings International Computer Music Conference, Urbana,Illinois*, pages 89–95. Computer Music Association, 1987.

[44] B. Gazengel, J. Gilbert, and N. Amir. Time domain simulation of single reed wind instruments. From the measured impedance to the synthesis signal. Where are the traps? *Acta Acustica*, 3:445–472, 1995.

[45] M. Ghosh. *Indian J. Phys.*, 7:365–382, 1932.

[46] N. Giordano. Mechanical impedance of a piano soundboard. *J. Acoust. Soc. Amer.*, 103 (4):2128–2133, April 1998.

[47] N. Giordano. Sound production by a vibrating piano soundboard: Experiment. *J. Acoust. Soc. Amer.*, 104 (3):1648–1653, septembre 1998.

[48] N. Giordano and A. J. Korty. Motion of a piano string: Longitudinal vibrations and the role of the bridge. *J. Acoust. Soc. Amer.*, 100(6):3899–3908, December 1996.

[49] N. Giordano and J. P. Millis. Hysteretic behavior of piano hammers. In *Proceedings International Symposium of Musical Acoustics*, pages 237–240.

[50] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products.* Academic press, 1980.

[51] K. Graff. *Wave Motion in Elastic Solids.* Prentice-Hall, Englewood Cliffs, NJ, third edition, 1974.

[52] Ph. Guillemain and R. Kronland-Martinet. Characterization of acoustic signals through continuous linear time-frequency representations. *IEEE Special Issue on Time-Frequency and Time-Scale Analysis*, 84(4):561–585, April 1996.

[53] B. Gustaffson, H.-O. Kreiss, and J. Oliger. *Time dependent problems and difference methods.* John Wiley and Sons, New York, 1986.

[54] D. E. Hall. Piano string excitation in the case of small hammer mass. *J. Acoust. Soc. Amer.*, 79(1):141–147, January 1986.

[55] D. E. Hall. Piano string excitation II: General solution for a hard narrow hammer. *J. Acoust. Soc. Amer.*, 81(2):535–546, February 1987.

[56] D. E. Hall. Piano string excitation III: General solution for a soft narrow hammer. *J. Acoust. Soc. Amer.*, 81(2):547–555, February 1987.

[57] D. E. Hall and A. Askenfelt. Piano string excitation V: Spectra for real hammers and strings. *J. Acoust. Soc. Amer.*, 83(6):1627–1638, April 1988.

[58] D. E. Hall and P. Clark. Piano string excitation IV: The question of missing modes. *J. Acoust. Soc. Amer.*, 82(6):1913–1918, December 1987.

[59] E. Hayashi, M. Yamane, and H. Mori. Behavior of piano-action in a grand piano. part I. analysis of the motion of the hammer prior to string contact. *J. Acoust. Soc. Amer.*, 105(6):3534–3544, 1999.

[60] H. Helmholtz, editor. *On the Sensations of Tone.* Dover Publications, New-York, 1954. trans. Alexander Ellis from 4th German ed. of 1877.

[61] T. Hikichi and N. Osaka. An approach to sound morphing based on physical modeling. In *Proceedings International Computer Music Conference, Beijing*, pages 108–111. Computer Music Association, 1999.

[62] L. Hiller and P. Ruiz. Synthesizing musical sounds by solving the wave equation for vibrating objects. *J. Audio Eng. Soc.*, 19:462–72, 542–51, 1971.

[63] K. H. Hunt and F. R. E. Crossley. Coefficient of restitution interpreted as damping in vibroimpact. *ASME J. Appl. Mech.*, pages 440–445.

[64] D. A. Jaffe and J. O. Smith III. Extensions of the Karplus-Strong plucked-string algorithm. *Computer Music Journal*, 7(2):56–69, 1989.

[65] H. Järveläinen, V. Välimäki, and M. Karjalainen. Audibility of inharmonicity in strings instruments sounds, and implications to digital sound synthesis. In *Proceedings International Computer Music Conference, Beijing*, pages 359–362. Computer Music Association, 1999.

[66] M. Karjalainen, V. Välimäki, and Z. Jánosy. Towards high-quality sound synthesis of the guitar and string instruments. In *Proceedings International Computer Music Conference, Tokyo*, pages 56–63. Computer Music Association, Sept. 10–15 1993. available online at http://www.acoustics.hut.fi/~vpv/publications/icmc93-guitar.htm.

[67] K. Karplus and A. Strong. Digital synthesis of plucked string and drum timbres. *Computer Music Journal*, 2(7):43–55, 1983.

[68] R. Kronland-Martinet, Ph. Guillemain, and S. Ystad. Modelling of natural sounds by time-frequency and wavelet representations. *Organised Sound*, 2(3):179–91, 1997.

[69] T.I. Laakso, V. Välimäky, and M. Karjalainen. Splitting the unit delay. *IEEE Signal Processing Magazine*, 1996.

[70] M. Lang. Allpass filter design and applications. *IEEE Transactions on Signal Processing*, 46(9), 1998.

[71] J. Laroche and J. L. Meillier. Multichannel excitation/filter modeling of percussive sounds with application to the piano. *IEEE Speech and Audio Processing*, 2(2):329–344, April 1994.

[72] K. Lee and A. Horner. Modeling piano tone with group synthesis. *J. Audio Eng. Soc.*, 47(3):101–111, 1999.

[73] K. A. Legge and N. H. Fletcher. Nonlinear generation of missing modes on a vibrating string. *J. Acoust. Soc. Amer.*, 76(1):5–12, 1984.

[74] J. D. Markel and A. H. Gray. *Linear prediction of Speech.* Springer-Verlag, New York, 1976.

[75] B. C. J. Moore, R. W. Peters, and B. R. Glasberg. Thresholds for the detection of inharmonicity in complex tones. *J. Acoust. Soc. Amer.*, 77(5):1861–1867, May 1985.

[76] J.A. Moorer. *Signal processing aspects of computer music: a survey.* Digital Audio Signal Processing:An Anthology, William Kauffmann, Inc, 1985.

[77] P. M. Morse. *Vibration and sound.* McGraw-Hill, Reprint, (1^{st} *ed.* 1936), 1948.

[78] A.V. Oppenheim and R.W. Schafer. *Digital signal processing.* Prentice-Hall, Engelwood Cliffs, NJ, 1975.

[79] G. De Poli and A. Piccialli. *Granular synthesis of musical signals.* Representations of Musical Signals, The MIT Press, Cambridge, Massachusetts, USA, 1991.

[80] H.F. Pollard and Jansson E.V. A tristimulus method for the specification of musical timbre. *Acustica*, 51:162–171, 1982.

[81] M. Puckette. Pure data. In *Proceedings International Computer Music Conference, San Francisco.* Computer Music Association, 1996.

[82] M. Puckette and D. Zicarelli. *Max — An interactive graphic programming environment.* Opcode Systems, 1985.

[83] Kronland-Martinet R. *Digital subtractive synthesis of signals based on the analysis of natural sounds.* In *Etat de la Recherche Musicale* Ed. A.R.C.A.M., Aix en Provence, France, 1991.

[84] J. W. Strutt Lord Rayleigh, editor. *The Theory of Sound.* Dover Publications, New-York, 1954. 2nd ed.

[85] A. Reinholdt, E. Jansson, and A. Askenfelt. Analysis and synthesis of piano tone. *J. Acoust. Soc. Amer.*, 81(S61), 1987. Suppl. 1.

[86] C. Roads, editor. *The Computer Music Tutorial.* The MIT Press, Cambridge, Massachusetts, USA, 1995.

[87] P. M. Ruiz. *A technique for simulating the vibrations of strings with a digital computer.* PhD thesis, University of Illinois, 1970.

[88] A. Sarti and G. De Poli. Generalized adaptators with memory for nonlinear wave digital structures. In *Proc. EUSIPCO-96, VII European Signal Processing Conference*, volume 3, pages 191–194, 1995.

[89] P. Schaeffer. *Traité des objets musicaux*, pages 219–220. Edition du seuil, Paris, France, 1966.

[90] S. Serafin and J. O. Smith. Impact of string stiffness on digital waveguide models of bowed strings. *Catgut Acoustical Society Journal*, 4(4), 2001.

[91] X. Serra and J. O. Smith. Spectral modeling synthesis: a sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music* , 14(4):12–24, 1990.

[92] J. O. Smith. *Techniques for Digital Filter Design and System Identification with Application to the violin.* PhD thesis, Stanford University, California, USA, 1983.

[93] J. O. Smith III. *Techniques for Digital Filter Design and System Identification with Application to the Violin.* PhD thesis, Elec. Dept., Stanford University (CCRMA), June 1983. available as CCRMA Technical Report STAN-M-14. Portions available online at http://www-ccrma.stanford.edu/~jos/(try a search there).

[94] J. O. Smith III. A new approach to digital reverberation using closed waveguide networks. In *Proceedings International Computer Music Conference*, pages 47–53. Computer Music Association, 1985. also available in [96].

[95] J. O. Smith III. Efficient simulation of the reed-bore and bow-string mechanisms. In *Proceedings International Computer Music Conference, The Hague*, pages 275–280. Computer Music Association, 1986. Also available in [96].

[96] J. O. Smith III. Music applications of digital waveguides. Technical Report STAN–M–39, CCRMA, Music , Stanford University, 1987. a compendium containing four related papers and presentation overheads on digital waveguide reverberation, synthesis, and filtering. CCRMA technical reports can be ordered by calling (650)723-4971 or by sending an email request to info@ccrma.stanford.edu.

[97] J. O. Smith III. Viewpoints on the history of digital synthesis. In *Proceedings International Computer Music Conference, Montreal*, pages 1–10. Computer Music Association, 1991.

[98] J. O. Smith III. Physical modeling using digital waveguides. *Computer Music* , 16(4):74–91, Winter 1992. Special issue: Physical Modeling of Musical Instruments, Part I. Available online at http://www-ccrma.stanford.edu/~jos/pmudw/.

[99] J. O. Smith III. Efficient synthesis of stringed musical instruments. In *Proceedings International Computer Music Conference, Tokyo*, pages 64–71. Computer Music Association, 1993. available online at http://www-ccrma.stanford.edu/~jos/cs/.

[100] J. O. Smith III. Physical modeling synthesis update. *Computer Music* , 20(2):44–56, Summer 1996. available online at http://www-ccrma.stanford.edu/~jos/pmupd/.

[101] J. O. Smith III. *Digital Waveguide Modeling of Musical Instruments.* http://www-ccrma.stanford.edu/~jos/waveguide/, May 9, 2002. (printed version forthcoming).

[102] J. O. Smith III and P. Gossett. A flexible sampling-rate conversion method. In *Proceedings International Acoustics, Speech, and Signal Processing, San Diego*, volume 2, pages 19.4.1–19.4.2, New York, March 1984. IEEE Press. An expanded tutorial based on this paper and

associated free software are available online at The Digital Audio Resampling Home Page: http://www-ccrma.stanford.edu/~jos/resample/.

[103] J. O. Smith III and D. Rocchesso. Aspects of digital waveguide networks for acoustic modeling applications. *http://www-ccrma.stanford.edu/~jos/wgj/*, December 19, 1997.

[104] J. O. Smith III and S. A. Van Duyne. Commuted piano synthesis. In *Proceedings International Computer Music Conference, Banff*, pages 319–326. Computer Music Association, 1995.

[105] K. Steiglitz and L. E. McBride. A technique for the identification of linear systems. *IEEE Trans. Automatic Control*, 10:461–464, 1965.

[106] J. Strikwerda. *Finite Difference Schemes and Partial Differential Equations.* Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, Calif., 1989.

[107] A. Stulov. Hysteretic model of the grand piano hammer felt. *J. Acoust. Soc. Amer.*, 97(4):2577–2585, April 1995.

[108] H. Suzuki. Model analysis of a hammer-string interaction. *J. Acoust. Soc. Amer.*, 82(4):1145–1151, October 1987.

[109] J.P. Thiran. Recursive digital filters with maximally flat group delay. *IEEE Trans. Circuit Theory*, CT-18(6), Nov. 1971.

[110] T. Tolonen, C. Erkut, V. Välimäki, and M. Karjalainen. Simulation of plucked strings exhibiting tension modulation driving force. In *Proceedings International Computer Music Conference, Beijing*, pages 5–8. Computer Music Association, 1999.

[111] T. Tolonen, V. Välimäki, and M. Karjalainen. Evaluation of modern sound synthesis methods. Technical Report 48, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland, 1998. URL:http://www.acoustics.hut.fi/ ttolonen/sound_synth_report.html.

[112] C. Valette and C. Cuesta. *Mécanique de la corde vibrante.* Traité des nouvelles technologies. Série Mécanique, ed. Hermès., 1993.

[113] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Zoltan Jánosy. Physical modeling of plucked string instruments with application to real-time sound synthesis. *Audio* , 44(5):331–353, May 1996.

[114] S. A. Van Duyne and J. O. Smith III. Developments for the commuted piano. In *Proceedings International Computer Music Conference, Banff*, pages 335–343. Computer Music Association, 1995.

[115] G. Weinreich. Coupled piano strings. *J. Acoust. Soc. Amer.*, 62(6):1474–84, Dec 1977. Also contained in [5]. See also *Scientific American*, vol. 240, p. 94, 1979.

[116] T. Yanagisana and K. Nakanura. Dynamic compression characteristics of piano felt. *J. Acoust. Soc. Jpn*, 40.

[117] R. W. Young. Inharmonicity of plain wire piano strings. *J. Acoust. Soc. Amer.*, 24:267–273, 1952.

[118] S. Ystad. *Sound modeling using a combination of physical and signal models.* PhD thesis, Université de la Méditerranée, France, 1998.

[119] D. Zicarelli. An extensible real-time signal processing environment for max. In *Proceedings International Computer Music Conference, Michigan.* Computer Music Association, 1998.