Helsinki University of Technology Laboratory of Acoustics and Audio Signal Processing
Espoo 2000 Report 54

Physics-Based Sound Synthesis of the Piano

Balázs Bank

Helsinki University of Technology Laboratory of Acoustics and Audio Signal Processing
Espoo 2000 Report 54

Physics-Based Sound Synthesis of the Piano

Master's thesis

Balázs Bank

Supervisors:

Dr. László Sujbert
Budapest University of Technology and Economics, Department of Measurement and Information Systems
Dr. Vesa Välimäki
Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing

Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing
P.O.Box 3000
FIN 02015-HUT
Tel. +358 9 4511
Fax. +358 9 460 224
E-mail lea.soderman@hut.fi

ISBN 951-22-5037-3
ISSN 1456-6303

Abstract

The present work is about the synthesis of piano sound based on the grounds of physical principles. For that, first the acoustical properties of the piano have to be understood, since the underlying physical phenomena establish the framework for the model-based sound synthesis. Therefore, the different parts of the piano were measured and analyzed.

The groundwork of the piano model lies in the digital waveguide modeling of the string behavior. Accordingly, the digital waveguide string model is thoroughly discussed and analyzed. The mathematical equivalence of the digital waveguide and the resonator bank is also presented.

The partition of the piano model follows the principles of the sound production mechanism of the real piano. The hammer is modeled by nonlinear interaction. The discontinuity problem arising when connecting the hammer to the string is investigated and new solutions for its avoidance are proposed. The instability problems of the hammer model are overcome by a novel multi-rate implementation. The possible use of a nonlinear damper model is also discussed. The string simulation is based on the digital waveguide. For beating and two-stage decay, a new parallel resonator bank structure is proposed. The soundboard model consists of a feedback delay network with shaping filters. A new technique is presented to reproduce the attack noise of the piano sound in an efficient and physically meaningful way. Concerning the implementation issues, a multi-rate piano model is proposed, which resolves the problem of different computational loads presented by the string models of the low and high register.

Additionally, the calibration of the piano model is described. A new loss filter design algorithm is presented for the calibration of the digital waveguide. The new technique minimizes the error of the resulting decay times and also ensures the stability of the feedback loop. For the one-pole filter as a special case, a novel filter design technique is proposed. It is founded on the new theoretical results of the Appendix concerning the decay times of a feedback loop containing the one-pole loop filter. A robust technique for the measurement of beating and two-stage decay is presented. This is used for the calibration of the parallel resonator bank.

The methods and techniques proposed here are described with the application to piano sound synthesis. Nevertheless, most of them can be exploited for the efficient synthesis of other musical instruments as well.

Keywords: digital signal processing, digital waveguide, musical acoustics, musical instruments, piano, sound synthesis

Preface

This work is a Master's Thesis for the Department of Measurement and Information Systems, Budapest University of Technology and Economics (Hungary). Most of the work has been carried out in the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology (Finland) during the academic year 1999–2000. The work is a part of the Sound Source Modeling Project financed by the Academy of Finland.

First of all, I wish to thank my supervisors, Dr. László Sujbert and Dr. Vesa Välimäki, for continuous help and support in all the aspects of my work.

I am grateful to Prof. Matti Karjalainen and Dr. Vesa Välimäki for providing me the honoring opportunity to conduct research at the Acoustics Laboratory of HUT. The understanding of the teachers at BUTE is also acknowledged. Especially, I would like to thank for the help of Dr. László Sujbert, who organized all my administration in Hungary, besides being my friend and supervisor.

I am much obliged to Prof. Matti Karjalainen and Dr. Tony Verma for carefully reading through my manuscript and giving supporting critics and comments.

I wish to thank all the people at the Acoustics Laboratory of HUT. Especially Mr. Tom Bäckström and Mr. Henri Penttinen for sharing with me the fresh slopes of Siberia, Mr. Poju Antsalo for helping in my measurements and for his special hammer, and Ms. Hanna Järveläinen for the annoying listening tests. I would like to thank Mr. Juha Merimaa for feedback about the sound of my piano model. I am thankful to Mrs. Lea Söderman for helping in my practical questions, and to Docent Unto K. Laine, who made me enthusiastic about the northern lights. I am particularly grateful for the LaTeX advice of Mr. Cumhur Erkut and Mr. Tero Tolonen, who helped me to make my thesis look nicer.

I am glad that I had the opportunity to discuss with Prof. Julius O. Smith from Stanford University about some parts of my piano model and other issues during the HUT Pythagoras seminar in December 1999.

Mr. János Márkus helped me in the piano measurements, which founded the basis of my work. I am also thankful to him for his MIDI-to-text converter and other useful programs written by him, which made it possible to try the capabilities of the piano model on real musical scores. I whish to thank Mr. Attila Nagy for writing a paper with me about the synthesis of piano and violin sound for a student paper contest. I also thank Dr. Fülöp Augusztinovicz for his advice and for providing the measurement equipment I needed.

I will always remember the Sunday dinners by the Peltomaa family and the

discussions around the table, and I thank them for teaching me cross-country skiing in the marvelous woods of Lapland. I am also grateful for all the Finnish, Hungarian, and foreign friends I have met here, who made me feel home in Finland. I thank my relatives and friends at home for their understanding in that I was away for almost a year.

I would like to express my gratitude to all the people who taught me. Particularly who opened up the fields of mathematics, physics, acoustics, and signal processing.

I am grateful to my brother for visiting me twice in Finland and for being not only my brother, but my best friend as well. Finally, I wish to thank my parents for encouraging me to study, and also to study abroad, even if they missed me a lot in these months. I am grateful to them for their constant care and support in every aspect of my life.

Otaniemi, Finland
May 30, 2000

Balázs Bank

Contents

List of symbols

a_1	filter coefficient
a_k	amplitude coefficient of the resonators
A	amplitude
$\mathbf{A}$	input coefficient matrix of the resonators, feedback matrix
b_1	filter coefficient
$b(t)$	beating signal
$\mathbf{b}$	input coefficient vector
B	inharmonicity coefficient
$\mathbf{B}$	output coefficient matrix of the resonators
c	wave velocity
c_1, c_3	coefficients of the time-derivates in the wave equation
$\mathbf{c}$	output coefficient vector
d	diameter of the string, output coefficient
D	phase delay
$\mathbf{D}$	delay matrix
e	approximation error
f	frequency
f_0	fundamental frequency
f_s	sampling frequency
f^+	right going traveling wave component
f^-	left going traveling wave component
F	force
g	real coefficient
$H(z)$	transfer function in z-domain
j	imaginary unit
K	hammer stiffness coefficient
l	length of the string
M	length of the digital waveguide
M_{in}	force input position of the digital waveguide
M_{out}	observation position of the digital waveguide
N	number of delays in the digital waveguide ($N = 2M$)
p	hammer stiffness exponent
p_k	poles of the resonators
Q	Young modulus

r_F	reflection coefficient for force waves
r_k	pole radius
r_v	reflection coefficient for velocity waves
$R(z)$	transfer function of the resonator
$\Re\{c\}$	real part of the complex variable c
$\mathbf{R}$	resonator matrix
$\mathbf{s}$	state vector
S	cross section of the string
$S(z)$	transfer function of a basic string model
t	time
t_n	time instant of the time-domain sampling step n
T	string tension
T_s	sampling period
v	velocity
x	position along the string
x_m	position of the spatial sampling step m
$x(t), x(n)$	input signal
$\mathbf{x}$	input vector
y	transversal displacement of the string
$y(t), y(n)$	output signal
$\mathbf{y}$	output vector
z	z-transform variable
z^{-1}	unit delay
Z	impedance
Z_0	characteristic impedance of the string
φ	phase in radians
κ	radius of gyration
μ	mass density
σ	decay rate
ϑ	angular frequency in radians
τ	decay time

Chapter 1

Introduction

The aim of this thesis was understanding and simulating the sound generation of the piano. Examining the physical principles of a musical instrument can be useful in many ways: a better insight in the sound production mechanism of the piano could lead to improvements of the instrument. On the other hand, it can be also applied for the design of sound synthesis algorithms, issuing in more realistic synthetic piano tones.

These two approaches yield different modeling levels. While for understanding physical phenomena complicated and accurate models are used, the efficient sound synthesis calls for simple and fast algorithms. Here we concentrate on the second approach. Accordingly, the proposed model captures only the most essential parts of the structure and sound generating behavior of the instrument.

The tasks for developing the structure of a physical model for an instrument are the following: first, the acoustical properties of the instrument have to be carefully investigated, then, a decision has to be made which features need to be simulated, the last step is finding efficient implementations for these features. After designing the structure, the parameters of the model are calibrated. This is based on the analysis of the real instrument. Thus, the design procedure both begins and ends with the analysis of the real world.

The physics-based approach has several advantages compared to traditional methods. The traditional techniques of sound synthesis simulate the resulting sound signal by means of abstract algorithms or by the modification of prerecorded samples. Consequently, they hide the underlying sound production mechanism. Moreover, their parameters have no meaning to the musician. This is similar to black-box modeling in system identification. On the contrary, physical modeling techniques concentrate on the internal structure, similar to white-box models in the field of system identification. This way, the parameters of the model will have a more direct interpretation in the real word. They can be, e.g., the length of a string or the stiffness of the hammer felt. Once a specific instrument is designed, the modification of the parameters will lead to meaningful results. For example, a modern grand piano can be turned to an old pianoforte by changing two or three parameters. The largest benefit of the physical modeling approach is that the interactions of the musician

are easily incorporated during sound synthesis.

Nonetheless, physical modeling does not mean that all the parts of the model have to be based on equations describing the behavior of the real piano. The general structure of the model originates from the real instrument, but some of the elements are modeled by non-physical algorithms. The aim is to incorporate the most advantageous methods for simulating the different parts of the instrument.

The drawbacks of the physical modeling technique are the loss of generality and the high implementation costs. While abstract algorithms or methods based on prerecorded samples can be used for the reproduction of many kinds of instruments, a physical model is valid only for a specific instrument, or instrument family. Generally, physical models need more computational time compared to traditional techniques. Most likely these are the reasons why physical modeling techniques have not gained significance in the commercial synthesis market. Nevertheless, the year by year increase in the computational power of digital signal processors can change this state. Multimedia applications are also a promising field for the physics-based modeling of musical instruments.

Since this work deals with the physics-based modeling of the piano, the acoustical properties of the instrument has to be discussed. This is done in Chapter 2, by reviewing the literature and presenting the results of own measurements. In Chapter 3, different sound synthesis algorithms are briefly outlined, concentrating on the application to piano sound. Chapter 4 discusses the principles of digital waveguide modeling, which is an efficient technique for simulating the string behavior. Based on the equations describing the digital waveguide, the equivalent resonator bank structure is presented. In Chapter 5, we concentrate on the structure of the piano model. After presenting the general structure, the different parts of the model are discussed and several new methods are proposed. The model structure follows that of the real piano. Namely, it consists of the hammer, string, and soundboard models. The chapter ends with the discussion of practical implementation issues. Chapter 6 discusses the calibration of the different parts presented in Chapter 5. These are aimed at determining the parameters of the model from measurements made on real pianos. The goal is to find a parameter set to the model which gives the most similar sound output to the real piano. New calibration techniques are also presented. Finally, Chapter 7 summarizes the results and gives the planned direction of future research.

Chapter 2

Acoustical properties of the piano

The piano belongs to the group of struck string instruments. Historically, it is the descendant of the harpsichord, but its excitation mechanism resembles that of the clavichord. The first piano was built as early as 1709 by Bartolomeo Cristofori, but it had to go through many changes to reach its modern form. This development was mostly due to the work of Henry Steinweg in the middle of the 19^{th} century. The comparison of a new and the 1720 Cristofori piano can be found in [Conklin 1996a,b,c]. The upright pianoforte was developed in the middle of the nineteenth century [Fletcher and Rossing 1998]. Here we concentrate on the grand piano, but most of the statements are valid for the upright piano as well. The description of the measurements can be found in the Appendix.

The general structure of the piano is the following: an iron frame is attached to the upper part of the wooden case and the strings are extended upon this in a direction nearly perpendicular to the keyboard. That end of the string which is closer to the keyboard is connected to the tuning pins on the bin block, and the other end, after crossing the bridge, is attached to the hitch-pin rail of the frame. The bridge is a thin wooden bar transmitting the vibration of the string to the soundboard, which can be found under the frame. This is displayed in Fig. 2.1.

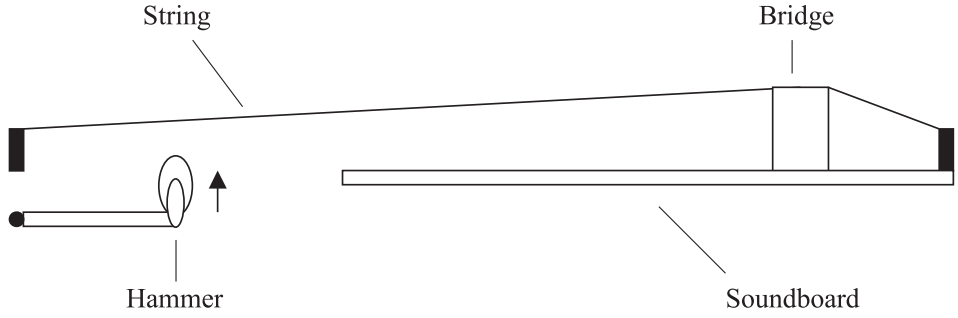

Figure 2.1: Schematic structure of the piano.

According to the above-mentioned parts, the sound-production mechanism of

the piano can be divided into the following steps:

1. The first is the excitation, the hammer strike. The piano action belongs to this part as well, since it transmits the kinetic energy taken in by the artist to the kinetic energy of the hammer, which, after bouncing to the string, transforms to vibrational energy.

2. This energy is stored by the string in its normal modes. One part of that is dissipated due to internal losses, the other gets to the soundboard through the bridge.

3. The soundboard converts the vibrational energy to acoustical energy, to the audible sound.

2.1 Piano action and hammer

2.1.1 The action

The action of the piano is an artwork of precision mechanics, and its operation is pretty complicated. This complexity is aimed at allowing the fastest possible repetition of a single note. In this way the repetition can be made before the hammer reaches its rest position. Roughly, the action can be considered as a lever system with a ratio of 1:5 between the movement of the key and the hammer. An important flavor of the action is that in the instant of hammer-string contact the hammer rotates freely. When the key is pressed down slowly, the hammer stops 4-6 mm beneath the strings. Under normal playing conditions the hammer is sent towards the string by the transferred kinetic energy. By pressing a key the corresponding damper is also lifted, which mutes the string by falling back after the key is released [Fletcher and Rossing 1998; Askenfelt and Jansson 1990, 1991].

By the examination of the timing of the action, Askenfelt and Jansson [1990] made some interesting observations. They concluded that the delay introduced by the action depends on the dynamic level to a great extent. In the piano-legato touch this delay can be as high as 100 ms, while at the forte-staccato touch from the hit of the key to the sound of the note elapses only 25 ms. Since this difference is audible, the skilled pianist must compensate for this actor. The key bottom contact can provide some mechanical feedback to the player, since the difference there is in the order of 15 ms, nevertheless skillful artists perform even more accurate timing.

The second article of the same series [Askenfelt and Jansson 1991] studies the dynamic properties of the action. There were contradicting standpoints between pianists and researchers concerning the possibilities of the pianist to influence the piano tone. Researchers claimed that the resulting sound is controlled by the final velocity of the hammer only. Nonetheless, pianists pay much attention to the "touch". According to them considerable variations can be made in the tone this way. Askenfelt and Jansson [1991] showed that the hammer exhibits various resonances and their level differs according to the type of touch. In legato playing, when

the action was accelerated continually, the amplitude of the hammer resonances was considerably lower than that of the staccato touch. The resonant frequencies were found at 50, 250 and 600 Hz with quality factor values between 15 and 30. It is still unclear if these vibrations can have an influence on the string vibrations or not. Thus, it calls for future investigation.

2.1.2 The hammer

Hammers have a great influence on the timbre of the piano, since they excite the strings. The hardwood cores of the hammers are covered by wool felt. The characteristics of the felt affect the resulting sound considerably. Harder felt results in stronger partials, i.e., a brighter tone. On the contrary, softer hammers produce less partials and a softer tone. The timbre of the piano can be influenced by "voicing". Hammers can be made softer by needling, or hardened by a hardening agent. Voicing is the last step of piano production, giving a "personality" to the instrument [Conklin 1996a].

The felt of the hammer is not homogeneous. Its hardness gradually changes from the outer part to the core. This is the main reason for the spectral differences at various dynamic levels: the high-frequency content of the tone increases with impact velocity [Conklin 1996a; Askenfelt and Jansson 1993]. Consequently, the felt can be considered as a nonlinear spring, with a stiffness increasing with compression. An approximate power-law model for the felt can be found in many articles [see, e.g., Boutillon 1988]:

$$F = K(\Delta y)^p \tag{2.1}$$

where Δy refers to the compression of the felt, K is the stiffness coefficient and p is the stiffness exponent. This may seem too simple for characterizing the complex behavior of the hammer, but in practice, with proper K and p values, it describes the hammer-string interaction at appropriate precision. These values (K and p) can be determined by curve-fitting from measured data [Boutillon 1988; Chaigne and Askenfelt 1994a]. Some papers suggest a hysteretic model for the hammer felt [Boutillon 1988; Stulov 1995]. An extensive theoretical and numerical analysis of the hammer-string interaction can be found in [Hall 1986, 1987a,b; Zhu and Mote 1994].

The process of the hammer-string interaction is the following: the hammer, accelerated by the action, hits the string, but it does not bounce back immediately, since its mass is not negligible compared to the string. The hammer is thrown back by the reflected pulses returning from the closer end of the string. The force experienced by the string (and the hammer) is a sequence of shock waves, which gives to the graph of the force a lumpy character [Fletcher and Rossing 1998].

The shape and smoothness (and thus the frequency content) of the force curve is influenced by p, K and the initial velocity as well. Increasing K or the initial velocity has the same effect, they both enlarge the high-frequency content of the excitation. The average duration of the hammer-string contact is determined by the ratio of the hammer and string mass. The heavier the hammer, the longer is

the contact time. With increasing initial velocities or K values the duration of the contact decreases to some extent, but the significance of this effect is much less than that of the mass ratio [Chaigne and Askenfelt 1994b; Fletcher and Rossing 1998].

The string is excited most effectively by the hammer when the contact time is equal to the half period of the tone. That's the case in the middle register. On the contrary, in the bass range the contact duration is much shorter, and in the treble much longer than the ideal. In contemporary pianos, hammers of gradually changing mass are used in order to reduce this effect [Conklin 1996a].

The spectra of the resulted sound depends on the striking point as well. This results in a comb filtering effect, since those modes of the string, which have a node near to the striking position cannot be excited effectively [Fletcher and Rossing 1998; Conklin 1996a].

The model built for synthesis purposes (details are described later) allows the investigation of the hammer-string interaction. The results of the simulation were very similar to those of the literature. The high-frequency content of the excitation force increased with initial velocity and the contact time decreased by a negligible amount. All other modifications on the excitation (hammer mass or stiffness parameter) gave the same results as those that can be found in the referred papers. This also justifies the correctness of the model.

2.1.3 Measurement of the hammer behavior

The movement of the hammer was examined by a small accelerometer. In this way we have an insight on the excitation signal, since one can estimate the shape of the force experienced by the hammer from the acceleration of the hammer. The correct value of the force could be calculated only by knowing the inertia of the hammer. Consequently, the results shown here differ from the force signal by a scaling factor. Ten different keys were measured at various dynamic levels using legato and staccato touch. The results resembled those found in the literature. Hammers in the middle register showed resonances at about 20 and 300 Hz. These values are presumably lower than those of normal conditions, since adding the mass of the accelerometer lowers the frequency of the resonances. The measured contact times were similar to those of the referred papers. In Fig. 2.2 one can examine the acceleration of the hammer at the hammer-string contact at piano and forte levels (the measured string: $A_4^\sharp$, 467 Hz). For clarity the sign of the acceleration was inverted (since the hammer accelerates downwards during bouncing back from the string).

It can be seen that at higher initial velocities not only the amplitude of the excitation force is changing but also its waveform. This is because of the nonlinear behavior of the hammer felt. By vertically magnifying and horizontally compressing the graph of the piano-level signal one can investigate the acceleration of the hammer before and after the impact (see Fig. 2.3). The first peak in the acceleration signal at around 20 ms corresponds to the beginning of the touch and the impact is now represented by a vertical line at about 140 ms. Here positive values correspond to upward acceleration.

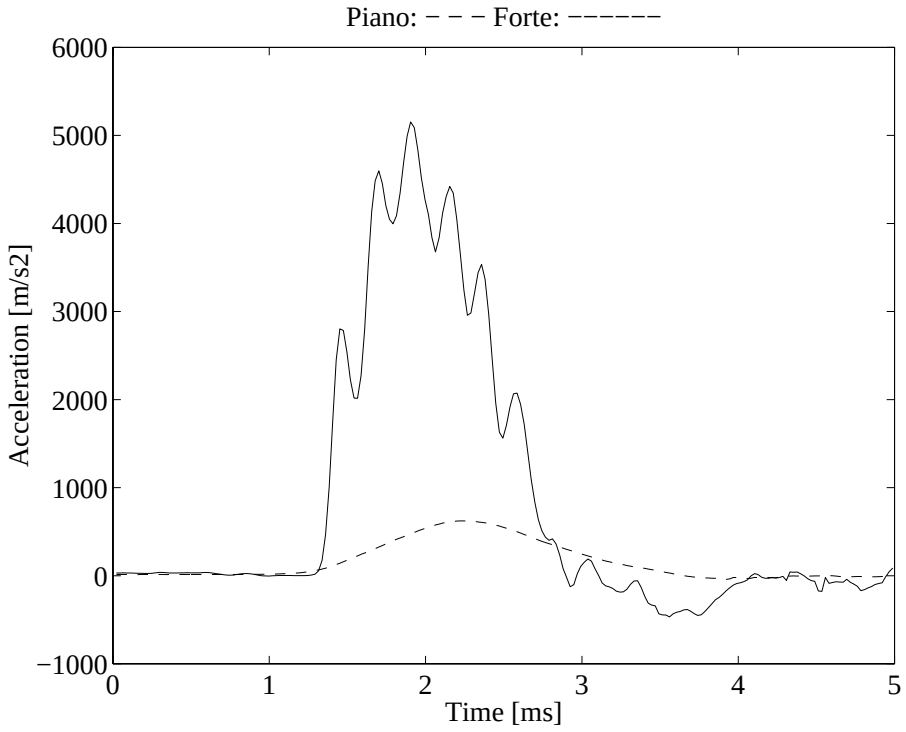

Figure 2.2: Acceleration of the hammer at the instant of the impact.

The figure shows that the hammer acceleration is not smooth, but it exhibits several resonances. The oscillation after the impact can be an important factor in the attack of the piano sound as well. The delay of the action is more than 100 ms in this case. At forte dynamic level it is about 20-30 ms.

2.2 Piano strings

The strings of the piano are made of steel wire. High efficiency requires high tension (about 700 N for each string), and accordingly the strings are strained at the 30-60% of their yield strength. In order to reach higher acoustic output, three strings for the same note are used (except for the lowest two octaves). These strings are not tuned in perfect unison, introducing beating and two-stage decay, two important characteristics of piano sound. The length of the strings is not exactly in inverse proportion to the fundamental frequency. If it were so, the lowest strings would be too long. Hence, the case should be larger than acceptable. This is avoided by increasing the mass of the bass strings. On the other hand, thicker string results in higher inharmonicity, since the string begins to behave as a stiff bar. The solution is winding the bass strings with one or two layers of copper wire, which reduces stiffness [Fletcher and Rossing 1998].

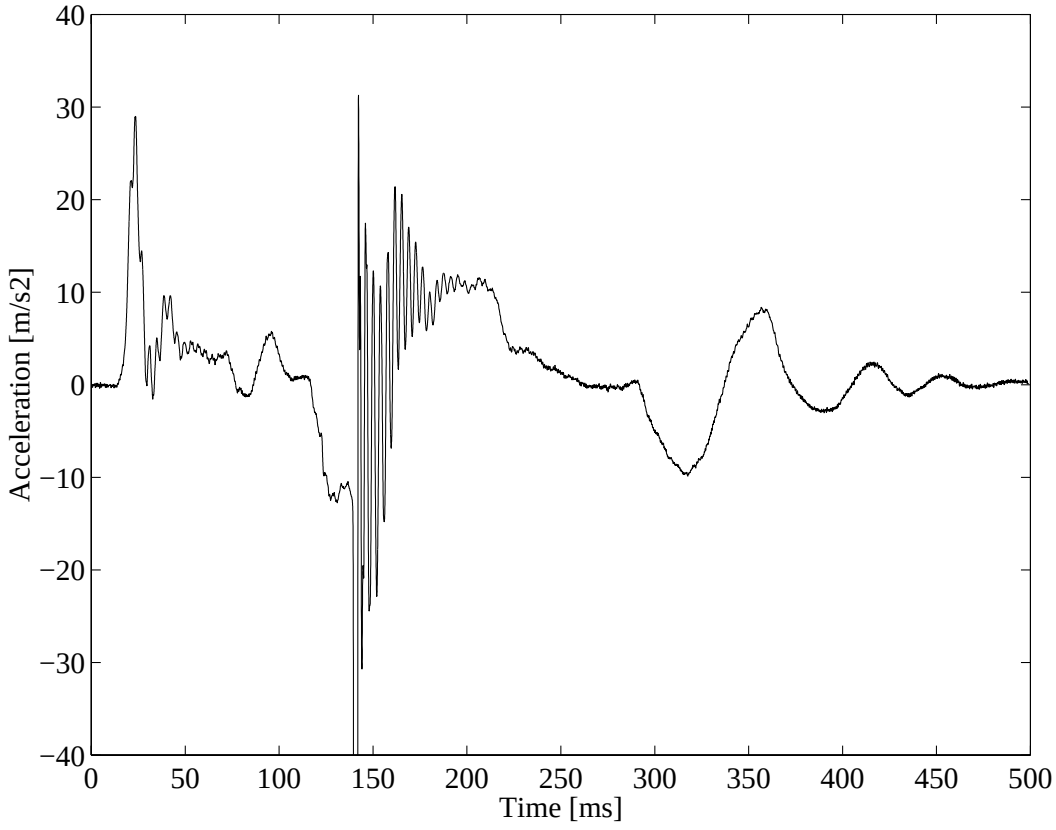

Figure 2.3: Acceleration of the hammer during the touch and after the impact.

2.2.1 Beating and two-stage decay

Having more strings for the same note not only increases the efficiency of the energy transmission towards the bridge but is also responsible for the characteristic piano sound. 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 the partials. In reality, these strings are not independent, since they are coupled to the bridge, giving rise to new modal frequencies. Coupling can be the reason for the two-stage decay as well. This means that in the early part of the tone the sound decays much faster than in the latter part [Weinreich 1977].

Weinreich [1977] studied the behavior of the coupling in the case of two one-polarization strings. He found that depending on the admittance of the bridge, different kind of normal modes arise. In the case of purely reactive bridge admittance, the frequencies of the two normal modes will be always different, even if the original uncoupled modal frequencies were the same for the two strings. When the bridge admittance is purely resistive, the behavior of the system falls to two regimes, depending on the amount of the uncoupled frequency difference of the two strings. If this frequency difference is greater than a certain value, both normal modes will have the same decay rate but will differ in frequency. In the other case,

the frequencies of the two normal modes will be the same, and their decay time will differ. If the admittance has both real and imaginary parts, as it is the case for a real soundboard, both the frequencies and decay times of the two normal modes can differ. This means that the beating can grow and decay. The beating will reach its maximum when the amplitudes of the two modes approximately equal.

There is another explanation for the two-stage decay: the behavior of the two different polarizations. Since the vertical polarization of the string is coupled more efficiently to the bridge than the horizontal polarization, the decay times differ significantly. The hammer excites the vertical polarization to a greater extent, and the energy transmission to the bridge is more effective in this direction as well. As a result, in the beginning of the tone the sound produced by the vertical polarization will be the dominant. As the vibration of the vertical polarization decays faster, the latter part of the tone will be determined by the horizontal polarization. However, these two polarizations are coupled to each other, so the description given here is rather a simplification of the real phenomenon. A detailed treatment on this subject can be found in [Weinreich 1977].

2.2.2 Inharmonicity

As mentioned earlier, the stiffness of the strings results in a slightly inharmonic sound. The wave equation of the stiff string is the combination of the equations for the ideal string (see Eq. 4.1) and ideal bar (for bending waves) [Morse 1948; Fletcher and Rossing 1998]:

$$\mu \frac{\partial^2 y}{\partial t^2} = T \frac{\partial^2 y}{\partial x^2} - QS\kappa^2 \frac{\partial^4 y}{\partial x^4} \tag{2.2}$$

where x is the position along the string, y is the transversal string displacement, t is the time, T refers to tension and μ to mass density. In Eq. (2.2) Q stands for Young's modulus, S is the cross-section area of the string and κ is the radius of gyration. As a result of the forth-power term, dispersion will appear, and waves with higher frequency will travel faster on the string, i.e., the wave velocity will not be constant any more. The results of group velocity measurements in good agreement to the theory can be found in [Podlesak and Lee 1988]. As a result of dispersion, higher modes will go back and forth on the string within a shorter time, so their frequency will be somewhat higher than that of the ideal string. Solving Eq. (2.2) for small stiffness the modal frequencies of the string will be the following [Fletcher et al. 1962; Fletcher and Rossing 1998; Morse 1948]:

$$f_n = n f_0 \sqrt{1 + Bn^2} \quad B = \frac{\pi^3 Q d^4}{64 l^2 T} \tag{2.3}$$

where f_0 is the fundamental frequency of the ideal string, n is the number of the partial, d is the diameter of the string and l is its length. The inharmonicity coefficient B can be calculated in this manner only for homogenous strings, but for wounded ones it is somewhat more difficult, the precise formula gives a lower value than this. The exact formula can be found in [Fletcher and Rossing 1998]. It can

be seen from the equation that inharmonicity increases with partial number. The B values are also increasing with the fundamental frequency of the string (since the lower strings are wound). On the contrary, when listening to piano sound, it seems to be more inharmonic in the lower register than in the higher one. One reason for this can be that the number of the partial which can be still heard is much higher for the low strings. The other is a psychoacoustic phenomenon: for higher fundamental frequencies the limit of perception for inharmonicity is higher than at low frequencies [Järveläinen et al. 1999]. The bandwidth of perceived inharmonicity was studied in [Rocchesso and Scalcon 1999].

It is an important question whether the inharmonicity is a desired factor or just an unavoidable feature. Conklin [1996c] states that pianists usually choose the largest piano available, which generally exhibits the lowest inharmonicity. The reason for this is that for low notes the amplitude of the fundamental is quite weak and the determination of the pitch is mainly based on the higher partials. Due to inharmonicity the frequency difference between partials is higher than the fundamental frequency and it increases with partial number. Accordingly, the definition of pitch becomes uncertain. Conklin [1996c] concludes that inharmonicity is an important factor of piano sound, but there should be as little of it as possible.

Inharmonicity also affects piano-tuning. Since tuning is based on beating between intervals (for fifths beating between the third partial of the lower note and the second of the higher), stretching of partials will cause stretching of fundamental frequencies as well. As a result, the lowest notes of the piano are 30 cent below, while the highest are 30 cent above the tempered values. There appears to be a psychological reason for stretched tuning too, since listeners often judge intervals as true octaves when their frequency ratio is slightly greater than 2:1 [Fletcher and Rossing 1998].

2.2.3 Measurement of inharmonicity

Several piano tones were recorded in this study with a microphone close to the soundboard. The results were similar to those of the literature. The graphs follow the theoretical curves quite well. In Fig. 2.4 the inharmonicity indices ($I_n = f_n/(nf_0)$) of the note C_2 (65 Hz) and $A_4^\sharp$ (466 Hz) are plotted. The inharmonicity coefficients calculated by curve fitting were 0.0001 for C_2 and 0.00075 for the $A_4^\sharp$ note. It can be noted that higher strings show higher inharmonicity, although perceptually they are considered to be less inharmonic. Besides the earlier mentioned reasons, another cause for this difference can be that the higher partials of the higher note decay much faster than those of the lower one making the inharmonicity less audible.

Based on the measurements, subjective experiments were made in order to explore whether it is needed or not to simulate inharmonicity in the synthesis model. Several piano tones were analyzed and resynthesized via additive synthesis. The partials had simple exponential decay curves fitted to measurements. The result sounded like a somewhat unsuccessful attempt to synthesize piano tones, but it was clear that it should be a piano sound. Then the dispersion was discarded, so all the

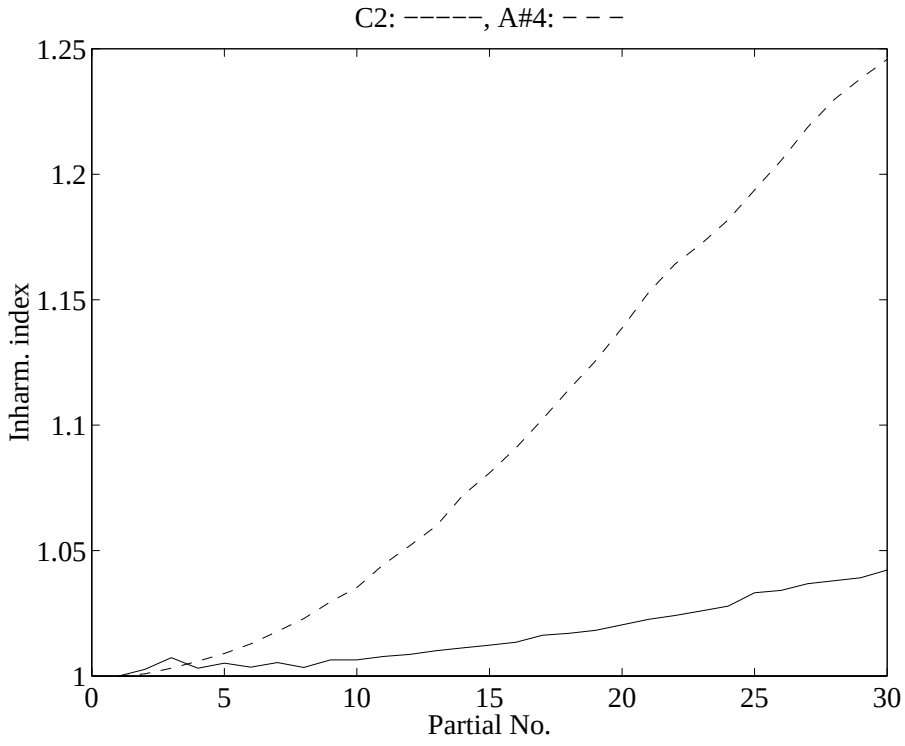

Figure 2.4: Inharmonicity indices of notes C_2 (65 Hz) and $A_4^\sharp$ (466 Hz).

partial frequencies were set to be perfectly harmonic with respect to the fundamental frequency. The difference for higher notes was negligible, it could be heard only in the attack part, since the higher harmonics fade away fast. For lower tones (below C_4) the effect was drastic: the synthetic sound was not a piano tone anymore, since it sounded rather harsh and disturbing. It should be also noted that the inharmonic synthesized tones for lower notes sounded more piano-like than the higher ones. As a conclusion, inharmonicity is an important factor and should be taken into account since it is essential in the distinction of the piano timbre.

2.2.4 Nonlinear effects and longitudinal waves

When examining the sound of string instruments, those modes which have a node at the excitation point are not missing completely, in contrast to the theoretical model. Their envelope also differs from the others, they rise slowly for the first 0.1 second and just after that start to decay (although the referred paper [Legge and Fletcher 1984] does not say if it is the case for piano). It follows that these modes gain energy from other modes and it can be only due to a nonlinear coupling. Legge and Fletcher [1984] made both theoretical and experimental examinations. The starting-point of mode-coupling is the tension variation. As the shape of the string changes during its motion, its length changes as well, causing tension modulation proportional to

the elongation. By examining the lossy (first-order losses only), rigidly terminated string they concluded that the tension and the fundamental frequency decrease exponentially as the sound decays. A modulation in tension at a double frequency of the fundamental can be also found. In the case of completely rigid terminations, this tension variation can only affect the modes from which it originates. However, if one of the terminations has finite impedance perpendicular to the string, modes n and m can transfer energy to the mode $2n \pm m$. When the string passes the bridge at an angle, the tension modulation exerts a force to the bridge perpendicular to the strings. As the tension is modulated by twice the fundamental frequency, mode n can deliver energy to the mode $2n$ [Legge and Fletcher 1984].

In a more recent paper [Hall and Clark 1987] claim that the nonlinear generation of missing modes is not the case for the piano, since the support is much more rigid than that in Legge's and Fletcher's experiments. They explain that the small residual amplitude of the "missing" modes comes from the finite resistance of the bridge (they did not find the slow rise in the envelopes of these modes). The contradiction of the two papers has not been resolved yet, hence it calls for further research.

Conklin [1999] showed that the tension modulation is also audible in the acoustical output of the piano. The so called "phantom partials" were only about 10 dB lower in amplitude than the nearest "real" partials. These were found at frequencies two times that of the exciting mode ($2f_n$) and at sum and difference frequencies ($f_n \pm f_m$) of the exciting modes as well, contradicting the theoretical results of Legge and Fletcher [1984]. The significance of phantom partials is exceptionally high in the case of piano, since they produce beating with the inharmonic real partials. Note that in a perfectly harmonic string the effect would be only a change in the initial amplitude. This can be one of the reasons why pianists prefer pianos with less inharmonicity, since in that case the frequency difference between real and phantom partials is lower.

The tension modulation as nonlinear coupling can also transfer energy between the two transversal polarizations. As a result, the polarization ellipse does not remain still but it features a slow precession [Fletcher and Rossing 1998] Experimental results on the subject can be found in [Hanson et al. 1994].

According to Giordano and Korty [1994], tension modulation is the cause of the rise of longitudinal waves as well. Their starting-point was the observation that the musical sound is preceded by a precursor signal. This precursor has to be transferred by the longitudinal waves since they travel faster than the transversal ones. They are excited by the tension-variation exerted by the hammer strike.

The relative frequencies of the longitudinal modes also influence the quality of the piano sound. The difference between the first transversal and longitudinal modes is typically between 4200 and 5200 cents (42-52 semitones), and it is constant irrespectively of the tuning of the string. In order to reach good quality, the first longitudinal mode should be in tune with the piano. This means that it should coincide with the first transversal mode of any higher note [Conklin 1996c].

2.3 The bridge and the soundboard

The soundboard of modern pianos is generally made of assembling strips of solid softwood (such as Sitka or red spruce). These 5-15 cm wide strips are glued together. Since the stiffness of the wood is higher along the grain, the cross-grain stiffness is increased by adding ribs to the soundboard. This increases the travel velocity in that direction. Pianos of lower cost are generally built of laminated wood. The comparison of soundboards of different materials can be found in [Conklin 1996b].

The vibration of the strings is transmitted to the soundboard through the bridge. The bridge functions as an impedance transformer, presenting higher impedance to the string than that if the strings were directly connected to the soundboard. In the latter case decay times would be too short. By carefully designing the soundboard and the bridge the loudness and the decay time of the partials can be set, although these parameters are in an inverse ratio [Fletcher and Rossing 1998].

The impedance curve of the soundboard exhibits a high modal density. Many studies have been done on the low frequency behavior of the soundboard. These resonances are similar to the simple motion of a plate and can be easily observed by the Chladni method [Conklin 1996b]. The higher frequency region of the soundboard was investigated by Giordano [1998]. He found that the average of the impedance is constant up to about 3-7 kHz (between 1000 and 10000 kg/s) and after that it starts to decrease with frequency. This decline is due to the ribs. Although the stiffness of the soundboard differs in the different directions, without ribs it would not produce this result, since the modal density would remain constant. According to Giordano [1998] the ribs stiffen the soundboard at lower frequencies to a higher extent, as at higher frequencies the modes fit in between the ribs. This result was supported by numerical simulations as well [Giordano 1997].

2.3.1 Measurement of the soundboard

The soundboard was measured by an impact hammer and the movement of the bridge was recorded by an accelerometer. The nearfield pressure of the soundboard was also measured. The excitation signal (the force measured by the impact hammer) was of a lowpass character. It consisted spectral components at appropriate level until 5 kHz. Different measurements of the same position along the bridge showed good agreement until this frequency, but above 10 kHz they were very different because of measurement noise. Therefore, these measurements were considered to be relevant only up to 5 kHz. By changing the measurement point the curves varied a lot. This is also caused by the strings, although they were damped. The difference would be smaller if the strings were removed, but that would not equal the normal playing conditions. The impedance and force-pressure transfer function curves for the note $A_4^\sharp$ can be seen in Fig. 2.5.

By looking at the low-frequency part of the figures it is easy to notice that the impedance minima coincide with pressure maxima. At higher frequencies the interference and cancellation of sound from different areas of the soundboard becomes

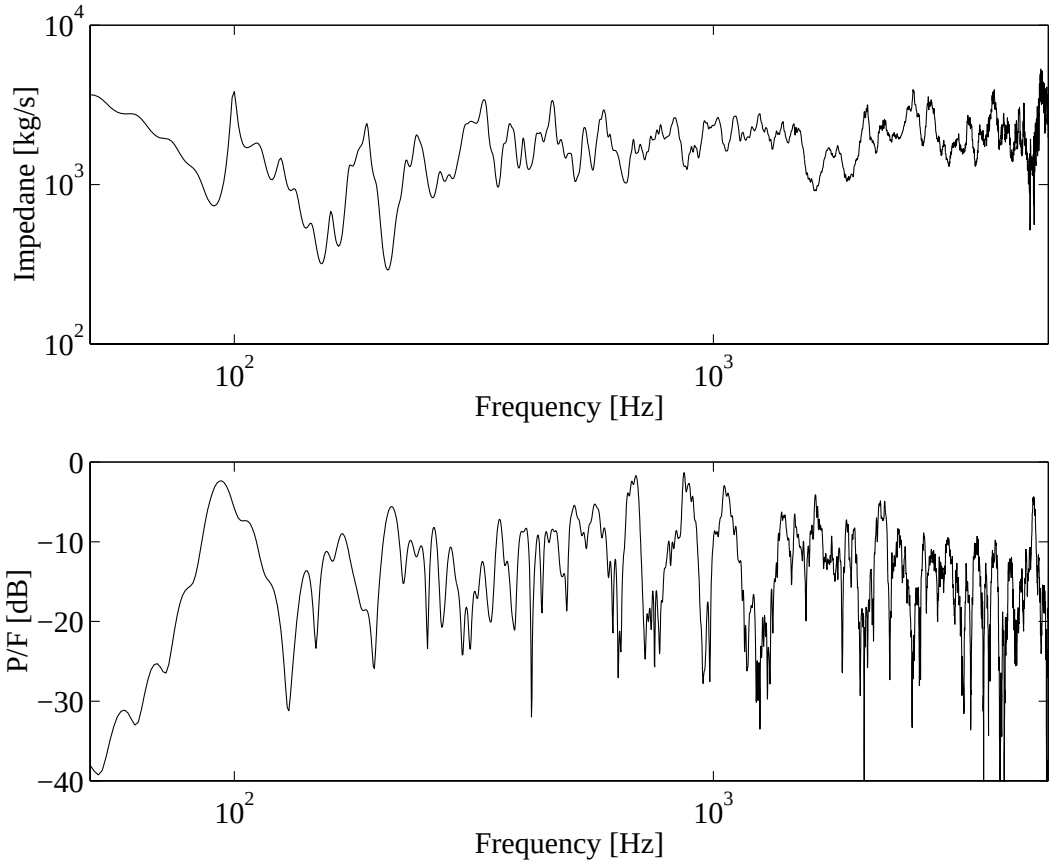

Figure 2.5: Impedance and force-pressure transfer function of the soundboard.

more significant, and the former coincidence cannot be found any longer. The average value of the soundboard impedance is similar to the results of Giordano [1998], although the decrease with frequency cannot be observed. To answer this question, new measurements are needed which can be used at least up to 10 kHz.

2.4 Pedals

Pianos have either two or three pedals. The most important one is the sustain pedal on the right, which lifts the dampers of all the strings. This sustains the struck keys on the one hand and also changes the character of the timbre on the other, since all the other strings can vibrate freely in sympathetic mode.

The effect of the sustain pedal was also measured. The bridge was hit by an impact hammer and the acceleration of the bridge and the near-field sound pressure were recorded. The impedance curve of the bridge did not change significantly compared to that without sustain pedal. The resonances became more peaky but the difference of the two curves stayed below one or two decibels. The main reason for the timbral change is not the change of the impedance but of the force-pressure transfer

function. The 300-400 ms long attack noise coming from the impulse response of the soundboard remains the same but it is superimposed on the slow (10-20 sec long) decay of string vibrations. The impulse response varies largely with the measurement position: hitting the bridge at the low register results in a softer timbre, hitting at the high stings in a brighter one.

The left pedal is the una corda pedal, which shifts the entire action sideways. Consequently, the hammers strike only two strings out of three in the treble, or one out of two in the midrange. This causes only about 1 dB reduction of the sound pressure level, but changes the timbre [Fletcher and Rossing 1998]. One of the reasons for this is that the third string gains energy from the vibration of the other two. The initial conditions of the coupled vibration are changed, resulting in a slower decay [Weinreich 1977]. Another reason for the spectral change can be that when the same hammer hits a smaller number of strings, it appears to be heavier with respect to the single strings, causing longer contact times and a softer timbre.

In upright pianos, the left pedal is a soft pedal. It moves the hammers closer to the string, resulting in lower striking force [Fletcher and Rossing 1998].

The middle pedal is the sustenuto pedal, which sustains only those notes that have been hit before depressing the pedal. In some uprights, this is a practice pedal, which lowers a piece of felt between the strings and the hammers [Fletcher and Rossing 1998].

2.5 Spectra of the piano sound

Fig. 2.6 shows the spectrum of the first 0.5 sec of the $A_4^\sharp$ note at forte dynamic level. The inharmonicity can be easily recognized: the distance between partials gets larger as the frequency increases. Between the real partials the phantom partials can also be found. Some are marked by $\times$ in Fig. 2.6. The frequencies of the marked phantom partials in terms of the real partial frequencies are: $2f_4$, f_4+f_5, $2f_5$, f_1+f_{10}, f_6+f_7, and f_7+f_8.

The upper part of Fig. 2.7 illustrates the STFT (Short Time Fourier Transform) of the first second of the same note up to 5 kHz. The length of the Hamming window was four times the pitch period. The envelopes of the distinct partials evolve differently as a result of the complex coupling. The basic phenomenon is the beating whose frequency is proportional to the partial number. The two-stage decay is not visible, since the displayed 1 sec time is too short for that.

The lower part of Fig. 2.7 shows the STFT diagram of the same note at piano level. It can be seen that not only the overall amplitude but also the level ratio of the partials changes. At lower hammer velocities the spectrum decays more steeply with frequency. The time-domain evolutions of the envelopes exhibit differences as well. The reason for this can be the nonlinear string behavior on one hand and the different excitation of the individual strings on the other. Because of the unevenness of the felt the hammer compliance for the distinct strings can be dissimilar, so the strings gain energy from the hammer with different ratios at different dynamic levels.

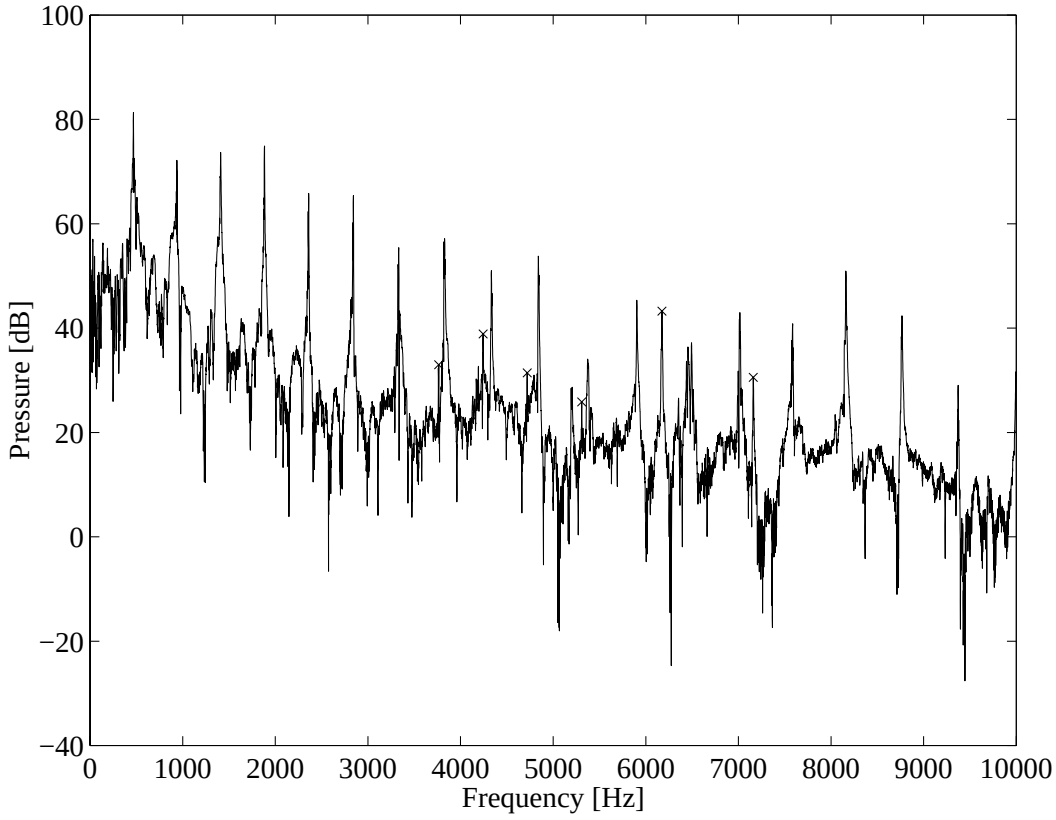

Figure 2.6: The spectrum of the first 0.5 sec of the $A_4^\sharp$ note.

Despite of this the average decay times remain the same.

The fundamental frequency of the tone decreases with respect to time by a negligible amount. At $A_4^\sharp$ forte note the difference is less than 0.2 Hz, which equals to 0.75 cent. As this change can be considered inaudible [Feldtkeller and Zwicker 1956; Rossing 1990], it can be stated that it has no effect on the timbre of the piano. The tension modulation can be still an important factor, since it is the cause of the phantom partials and longitudinal modes.

2.6 Conclusion

The timbre of the piano is determined by several factors. The main characteristic of the sound comes from the fact that the piano is a struck string instrument. It implies that the string is affected by an impulse excitation, resulting in a decaying amplitude. As the string is hit at about 1/7 of its length, a comb filtering effect occurs, where the amplitude of every 7^{th} partial is significantly decreased, although these partials may gain energy later by nonlinear coupling. The variations in the timbre and dynamic level come from the nonlinear behavior of the hammer. At higher initial velocities the spectrum gets wider and the peaks of the spectral envelope are shifted slightly

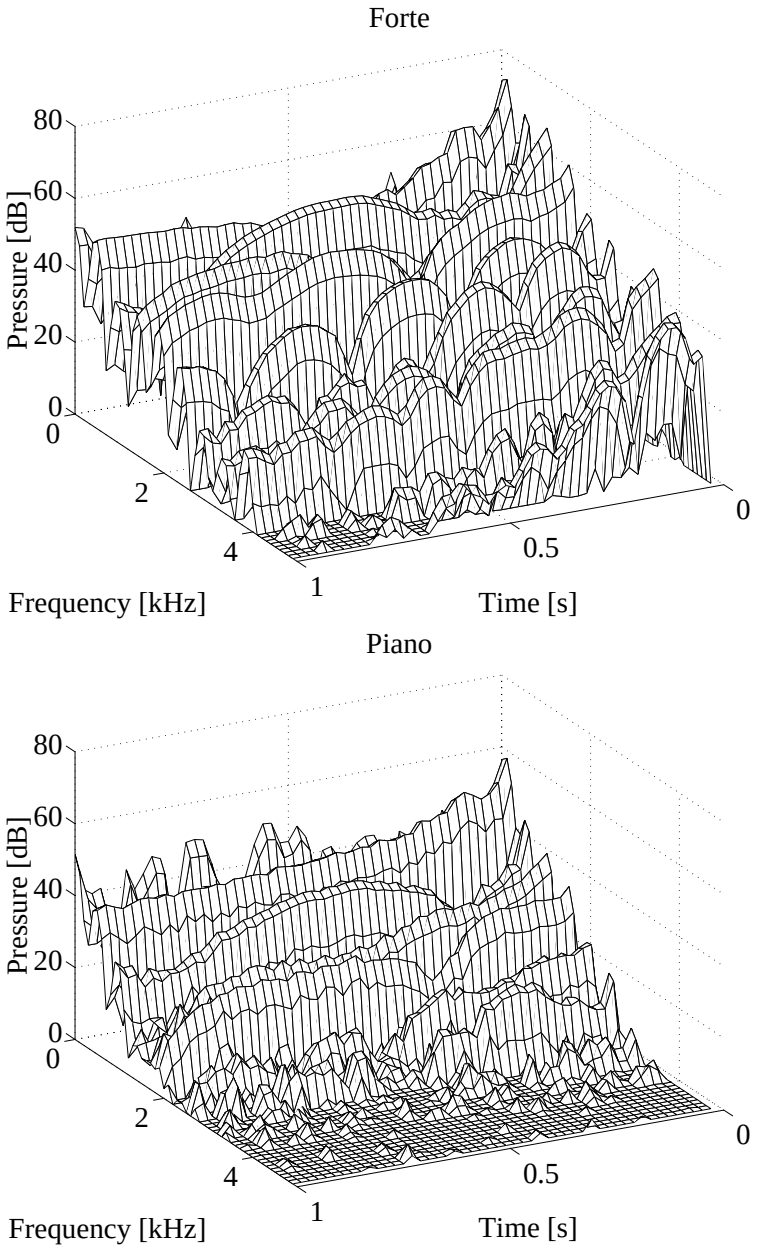

Figure 2.7: STFT diagram of the $A_4^\sharp$ note at forte and piano dynamic levels.

to the right.

The kinetic energy of the hammer transforms to the vibrational energy of the string. The main part of this energy is stored by the string, the second slowly gets to the soundboard and the third part dissipates. The string determines the fundamental frequency of the note and influences the decay of the partials. The decay and the transmitted energy also depend on the terminating impedance. At impedance maxima of the soundboard the partials can deliver less energy to the soundboard during the same amount of time. Therefore, their amplitudes will be lower and their decay times longer than those which are located at impedance minima. The stiffness of the string results in a high dispersion. There is no other western string instrument where the inharmonicity would be as high as in the case of the piano. Simulation results and informal listening experiments show that inharmonicity is especially important at the low register of the piano.

The different coupling of the two polarizations to the soundboard results in a two-stage decay. The envelope of the partials decays faster in the early part of the tone than in the latter. The slight frequency difference of the two or three strings causes beating. The resulting amplitude modulation is quite complicated because of the coupling of the strings. The tension modulation gives rise to longitudinal waves whose relative amplitude to the transversal modes increases with dynamic level. The appearance of phantom partials comes from the tension modulation as well. The three strings and their three different polarizations constitute a complex nonlinearly coupled system, issuing in a dynamic change of the timbre.

The soundboard, besides influencing the decay times, determines the overall spectrum significantly. Its behavior can be assumed to be linear, but its numerous resonances alter the timbre largely. The characteristic attack noise of the piano sound comes mainly from the impulse response of the soundboard, but also from the noise of the action.

Chapter 3

Overview of synthesis methods

Here we briefly overview the different techniques used for sound synthesis purposes, and their application to the piano. The details of the algorithms will not be given, for more thorough discussion see [Roads 1995; Tolonen et al. 1998]. The classification of different methods is based on the work of Smith [1991] and Tolonen et al. [1998].

3.1 Abstract algorithms

Here we discuss abstract algorithms. Their advantages are simplicity, generality, and the small number of control parameters. The drawback is that analysis procedures are complicated, making it almost impossible to simulate the sound of many real instruments. They are rather useful for creating synthetic, never heard sonorities.

The FM (frequency modulation) synthesis was presented by [Chowning 1973]. It is founded from the theory of frequency modulation used in radio transmission, but the frequencies of the modulating and carrier waves are in the same order here. Its advantage is that even a two-oscillator system can produce a rich spectrum. Inharmonic sounds can also be generated by this technique. If both the carrier and modulating waves are sinusoids, the amplitudes of the resulting harmonics can be calculated by the Bessel functions. Van Duyne [1992] used the FM synthesis for modeling inharmonic low piano tones.

The waveshaping synthesis was developed by Le Brun [1979] and Arfib [1979]. It is based on the nonlinear distortion of a simple input signal. An advantage is that the amplitude variation of the signal results in a large alteration of the output spectrum. When the input signal is sinusoidal, the amplitudes of the resulting harmonics are determined by the Chebyshev polynomials. This way, the shaping function can be analytically derived from the partial amplitudes. The method is capable to simulate harmonic spectra only.

The Karplus-Strong algorithm [Karplus and Strong 1983] is a simple algorithm based on the modification of the wavetable synthesis. In the wavetable synthesis, the signal is periodically read from a computer memory. In the Karplus-Strong algorithm the sample is modified after reading and written back to the same position. This way, the content of the wavetable will evolve with time. The algorithm has

been found very efficient in the simulation of plucked string and drum timbres. However, it was found soon that the Karplus-Strong algorithm is a special case of the technique now called digital waveguide modeling [Jaffe and Smith 1983; Smith 1983].

3.2 Processing of pre-recorded samples

The methods described here are based on recording and processing of real sounds. Accordingly, they are very accurate in reproducing the specific sound, which was recorded. They are not able to reproduce the changes in playing conditions, i.e., the simulation of not recorded sounds is not realistic. Another problem is that a large amount of data is needed for describing the instruments.

Sampling synthesis is the basic form of these algorithms. All the commercially available digital pianos use this technique. They store separate tones of the instrument in a memory and play it back when a key is depressed [Roads 1995]. The memory need is reduced by looping, that is, continuously repeating the steady state part of the tone. The amplitude and timbre evolution of the piano is simulated by a time-varying amplitude envelope and a time-varying filter. Different dynamic levels are taken into account by modifying these envelopes, hence a forte note results in a brighter and louder sound than a piano one. The method synthesizes the notes separately. Therefore, it cannot simulate the coupled vibrations of the strings or the restrike of the same string. Nevertheless, even if physical models theoretically could provide better results, the highest quality synthesized piano sounds are coming from devices using the sampling technique. Also an advantage is that the implementation of the algorithm is simple. Sampling synthesis can be successful in the synthesis of piano sound, because the impact velocity can be easily mapped to the amplitude and filter envelopes. Other instruments, such as the violin, where the musician has more influence on the sound, could not be realistically simulated.

Multiple wavetable synthesis is in a way similar to sampling synthesis. One musical tone is reproduced by many wavetables, whose content is interpolated [Roads 1995]. It also resembles additive synthesis since the wavetables have separate amplitude envelopes and their output is summed. In [Yuen and Horner 1997] the attack of sounds were synthesised by sampling, and for the decay part multiple wavetable synthesis was used.

Granular synthesis is based on composing the synthetic sounds from short sound elements in the time domain [Roads 1995]. The algorithms can be divided into asynchronous and pitch synchronous methods. The first method can rahter be used for creating synthetic sounds. For the reproduction of musical instrument sounds, usually the latter one is applied.

3.3 Spectral models

Spectral models try to approach the synthesis problem from the spectral domain, which is more similar to the human sound perception. Some of them also take psychoacoustic criteria into account. Their advantage comes from this fact. A drawback is that generally many parameters are needed for the description of an instrument. The simulation of transients is problematic, which makes them difficult to use for the synthesis of piano sound.

The simplest spectral technique is the additive synthesis. It is based on summing sinusoidal signals with different frequency and amplitude envelopes. It is popular since its mathematical background, the Fourier transform, is well developed. An advantage of the additive technique is the flexibility. Its drawback is the huge number of control parameters. [Fletcher et al. 1962] used this technique for the synthesis of piano sound by means of analog circuits.

Group additive synthesis is a combination of additive and wavetable synthesis techniques. It is motivated by the need for reducing the control stream of the additive synthesis. Therefore, some partials are grouped together and they will have common frequency and amplitude envelopes. These grouped partials are then combined to form wavetables. Lee and Horner [1999] used the method for modeling the sound of the piano. The grouping of the partials was determined with the help of genetic algorithms. In Zheng [1999], a critical-band based grouping was used.

Spectral modeling synthesis is a method which decomposes the sound into deterministic and stochastic components in the spectral domain [Serra and Smith 1990]. More recently, methods for deterministic, stochastic and transient decomposition were developed [Verma and Meng 1995]. Modeling the transients seems to be essential in synthesizing piano sound.

3.4 Physical modeling

The main feature of physical modeling approaches is that they model the origin of the sound phenomenon, and not the resulting signal. This issues in a realistic response to the interactions of the musician, since the input parameters of the model are analogous with the real world. Theoretically, the physical modeling approach could give the most realistic synthetic instrument sounds. Their drawback is the loss of generality and the high computational cost. The calibration of the parameters is not always trivial either.

The first method used in physical modeling was the discrete time solution of the finite difference equation [Hiller and Ruiz 1971a,b]. The finite difference method was applied for simulating the string of the piano in [Chaigne and Askenfelt 1994a,b]. Giordano [1997] used the technique for modeling the soundboard of the piano. A complete piano model based on the finite difference method was presented in [Hikichi and Osaka 1999]. The advantage of the approach is that the physical equations can be directly implemented as algorithms. Its drawback is the high computational cost,

especially if the structure is not one-dimensional.

Modal synthesis describes the structure with the linear combination of vibration modes [Adrien 1991]. Its main advantage is the generality, since the same kind of formulation can be used for the simulation of every vibrating structure. Analysis techniques for the determination of the modal parameters are already available, mostly used in the car and aircraft industry. However, these tools are rather expensive for academic research. The parameters of the model have physical meanings, but they are not intuitive. The visible parameters of the real structure, e.g., the length of the string, are hidden.

The exciter-resonator technique was used in [Laroche and Meillier 1994] for the synthesis of piano sound. A common excitation signal was derived which was filtered through the resonators. The drawback of the exciter-resonator approach is that it is not capable of simulating nonlinear interaction.

McIntyre et al. [1983] recognized that most of the instruments can be simulated with a nonlinearity and a linear element in a feedback loop. The linear element is the reflection function, whose main feature is the delay.

Smith [1983, 1987, 1992] proposed a computationally very efficient technique, the digital waveguide modeling. The method is based on the discretization of the time-domain solution of the wave equation. The efficiency of the digital waveguide lies in lumping the losses and dispersion of the structure to one point. Hence, the string or acoustic tube is simulated by a delay and a linear filter. In this way, it is similar to the approach of McIntyre et al. [1983]. The Karplus-Strong algorithm is a special case of digital waveguide modeling [Jaffe and Smith 1983]. The first piano model based on the digital waveguide was presented in [Garnett 1987]. There, the digital waveguides were connected to a common termination. In [Borin et al. 1997] a nonlinear hammer model was used. The computationally most efficient digital waveguide piano model is the commuted piano proposed in [Smith and Van Duyne 1995; Van Duyne and Smith 1995]. Since these are the approaches, which are most similar to the model presented in this thesis, their detailed description will be given in Chapter 5 with relation to the proposed model structure.

3.5 Conclusion

Until this time, the sampling synthesis has given the best results in synthesizing the piano tone. If all the notes of the piano are recorded at different dynamic levels, no other synthesis method can result in more realistic sound. The author believes that the only exception is the physical modeling, since it can take into account the interaction of the different parts, such as the coupling between different strings. The single weak point of sampling synthesis is that it treats the sounds separately. The only way to make a better synthesis method is to attack this weakness. However, the physical model described in this thesis does not completely fulfill this requirement yet. It can be only considered as a first step towards the high quality synthesis of piano sound.

Chapter 4

Principles of string modeling

In this chapter the mathematical background of digital waveguide modeling is presented. The formulation originates from the time-domain solution of the wave equation. After an overview of the ideal string, the effect of losses and dispersion is discussed. This chapter also introduces a new resonator bank interpretation and shows its equivalence to the digital waveguide model.

4.1 Modeling the ideal string

An effective approach for modeling a string and an acoustical tube was introduced by Smith [1987, 1992]. The method is based on the discretization of the time-domain solution of the wave equation.

By forming the wave equation of the ideal string, it is needed to make several simplifications: the length of the string is assumed to be infinite, its mass density μ and tension T is supposed to be homogenous and its displacement to be small with respect to string length, which means that its slope is very small ($dy/dx \ll 1$). Furthermore, only one transversal polarization of the string is taken into account. The result is the one-dimensional wave equation Eq. (4.1), which is similar to that of transmission lines or the longitudinal motion of bars. The derivation of this equation can be found in the literature [Morse 1948; Fletcher and Rossing 1998] et al.

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}, \qquad c = \sqrt{\frac{T}{\mu}} \tag{4.1}$$

In Eq. (4.1) x is the position along the string, y is the transversal displacement, t stands for time, T for the tension, μ for linear mass density and c for the wave velocity. The equation shows that the acceleration of a small section of the string is proportional to the curvature of the string at that section. Every traveling wave which retains its shape is a solution of the wave equation. Because of the linearity of the string, the general solution is a superposition of two traveling waves; one of them going to the right, the other to the left direction:

$$y(x,t) \;=\; f^{+}(ct - x) + f^{-}(ct + x)$$

$$v(x,t) = \frac{dy}{dt} = c\frac{df^+}{dt}(ct-x) + c\frac{df^-}{dt}(ct+x) \tag{4.2}$$

According to these equations, the shapes of the two traveling-waves f^+ and f^- are completely determined by the initial displacement $y(x,0)$ and velocity $v(x,0)$.

If the Nyquist condition [Oppenheim and Schafer 1975]is met, the temporal and spatial sampling of the string can be performed. In this case the continuous waveform can always be calculated via interpolation among the time-domain and spatial discrete points. If the sampling is done in a way that the traveling waves move one spatial sampling interval during one time-instant, it will lead to the digital waveguide model of the ideal string [Smith 1992]:

$$y(t_n, x_m) = y^+(n-m) + y^-(n+m) \tag{4.3}$$

This can be implemented by two parallel delay lines, where the transversal displacement of the string is calculated by adding up the output of the samples of the two delay lines at the same spatial coordinate. This is illustrated in Fig. 4.1.

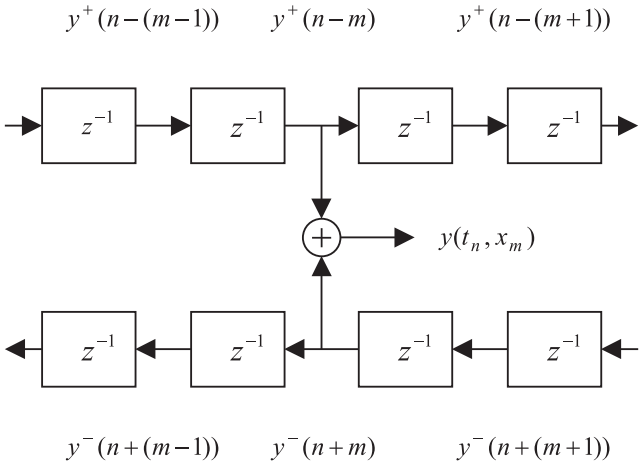

Figure 4.1: The principle of digital waveguide [Smith 1987, 1992].

As a result of the linear behavior of the digital waveguide, other variables can also be used instead of transversal displacement. These can be for example transversal velocity or acceleration, slope, curvature or force, since all of these satisfy the wave equation.

Nevertheless, it is worth turning our attention to the transversal velocity v and the force F, since they are proportional to each other. The characteristic impedance Z_0 of the string can be defined as follows (see e.g. [Fletcher and Rossing 1998]):

$$Z_0 = \frac{F^+}{v^+} = -\frac{F^-}{v^-}, \qquad Z_0 = \sqrt{T\mu} \tag{4.4}$$

where F^+ and v^+ are the force and velocity waves traveling to the right, and F^- and v^- to the left, respectively. Eq. (4.4) is valid at every position of the string and at every time instant. If a string with a characteristic impedance Z_0 is terminated by an impedance Z, the traveling waves will be reflected (except when $Z = Z_0$). This is identical to the termination of a transmission line. The equations for the reflection of force and velocity waves are the following:

$$r_v = \frac{v^-(x_{\text{term}}, t)}{v^+(x_{\text{term}}, t)} = \frac{Z_0 - Z}{Z_0 + Z}, \qquad r_F = \frac{F^-(x_{\text{term}}, t)}{F^+(x_{\text{term}}, t)} = -r_v = \frac{Z - Z_0}{Z_0 + Z} \tag{4.5}$$

An ideally rigid termination corresponds to an infinite terminating impedance $Z = \infty$. This implies that force waves reflect with the same amplitude and sign ($r_F = 1$), and velocity waves reflect with same amplitude but opposite sign ($r_v = -1$). In the following derivation the velocity v is used as the variable of the delay lines. The excitation force can be taken into account by adding $v_{\text{in}} = F_{\text{in}}/(2Z_0)$ to both delay lines at the position of the excitation. On the grounds of these equations the digital waveguide model of the ideal string can be formulated as shown in Fig. 4.2.

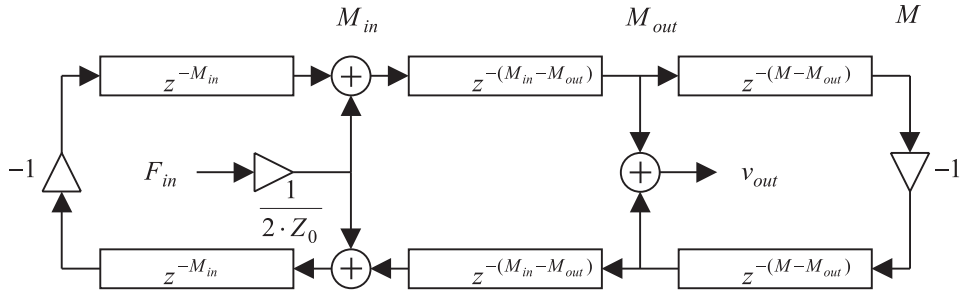

Figure 4.2: Digital waveguide model of the ideal string.

If the force is introduced at the node M_{in} in Fig. 4.2 the velocity at the node M_{out} will be:

$$\begin{aligned} v_{\text{out}}(z) &= \frac{1}{1 - z^{-N}} \left(1 - z^{-2M_{\text{in}}}\right) \left(1 - z^{-2(M - M_{\text{out}})}\right) z^{-(M_{\text{out}} - M_{\text{in}})} \frac{F_{\text{in}}}{2Z_0} \\ v_{\text{out}}(z) &= \frac{1}{1 - z^{-N}} H_{\text{in}}(z) H_{\text{out}}(z) z^{-(M_{\text{out}} - M_{\text{in}})} \frac{F_{\text{in}}}{2Z_0} \end{aligned} \tag{4.6}$$

where $N = 2M$. After the fractional expansion of the first term one gets the following equation:

$$\begin{aligned} v_{\text{out}}(z) &= \frac{1}{N} \left\{ \frac{1}{1 - z^{-1} e^{j\vartheta_1}} + \ldots + \frac{1}{1 - z^{-1} e^{j\vartheta_N}} \right\} \frac{F_{\text{in}}}{2Z_0} \times \\ &\quad \times H_{\text{in}}(z) H_{\text{out}}(z) z^{-(M_{\text{out}} - M_{\text{in}})} \end{aligned} \tag{4.7}$$

where $\vartheta_k = (2k\pi)/N$. According to Eq. (4.7), the digital waveguide can be interpreted as a resonator bank nested in comb filters. This configuration is illustrated

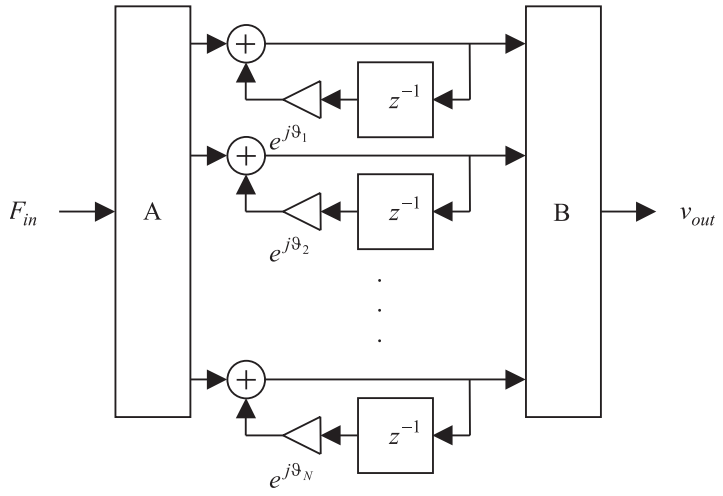

Figure 4.3: The equivalent resonator bank of the digital waveguide structure of Fig. 4.2.

in Fig. 4.3, where now A refers to $1/(2Z_0)H_{in}(z)$ and a junction, and B adds up the outputs of the resonators and implements $(1/N)z^{-(M_{out}-M_{in})}H_{out}(z)$.

It can be seen that the resonance frequencies are determined by the resonators. In the case of the ideal string the poles are equally distributed on the unit circle. The comb filters $H_{\rm in}$ and $H_{\rm out}$, besides introducing some delay, control the amplitude of the harmonics. They behave in a physically meaningful way: if the string is excited for example at the $7^{\rm th}$ of its length ($M_{\rm in}/M = 1/7$), the amplitude of every $7^{\rm th}$ mode will be zero. The amplitudes of the harmonics are influenced by the position of the observation point $M_{\rm out}$ in the same way.

Since all the poles lie on the unit circle, the impulse responses of the resonators are non-damping sinusoids. In this way the effect of the comb filters can be realized by multiplying the output of every resonator by the transfer function value of the comb filters at the corresponding frequency. Since the comb filters are of linear phase, the phase part of the transfer function is represented by the delay z^{-M}. Eq. (4.8) was derived by substituting $z = e^{j\vartheta}$.

$$\begin{aligned} H_{\rm comb}(z) &= \left(1 - z^{-2M_{\rm in}}\right)\left(1 - z^{-2(M-M_{\rm out})}\right) z^{-(M_{\rm out}-M_{\rm in})} \\ &= z^{-M}\left(e^{j\vartheta M_{\rm in}} - e^{-j\vartheta M_{\rm in}}\right)\left(e^{j\vartheta(M-M_{\rm out})} - e^{-j\vartheta(M-M_{\rm out})}\right) \\ &= -4\sin(\vartheta M_{\rm in})\sin(\vartheta(M - M_{\rm out}))z^{-M} \end{aligned} \tag{4.8}$$

So the transfer function of the whole string will be:

$$\begin{aligned} \frac{v_{\rm out}(z)}{F_{\rm in}(z)} &= \left\{\frac{a_1}{1 - z^{-1}e^{j\vartheta_1}} + \ldots + \frac{a_N}{1 - z^{-1}e^{j\vartheta_N}}\right\} z^{-M} \\ a_k &= -\frac{2}{Z_0 N}\sin(2k\pi\frac{M_{\rm in}}{N})\sin(2k\pi\frac{M - M_{\rm out}}{N}) \end{aligned} \tag{4.9}$$

When substituting the comb filters by multipliers at the output of the resonators, the output of the two systems represented by Eqs. (4.7) and (4.9) will be identical after the transient of the comb filters, which is $2M + M_{\text{in}} - M_{\text{out}}$ steps. The force-velocity impulse response of a digital waveguide and the equivalent resonator structure is illustrated in the upper and middle part of Fig. 4.4. In this example, $M = 8$, $M_{in} = 2$, $M_{out} = 7$, and $Z_0 = 1/2$.

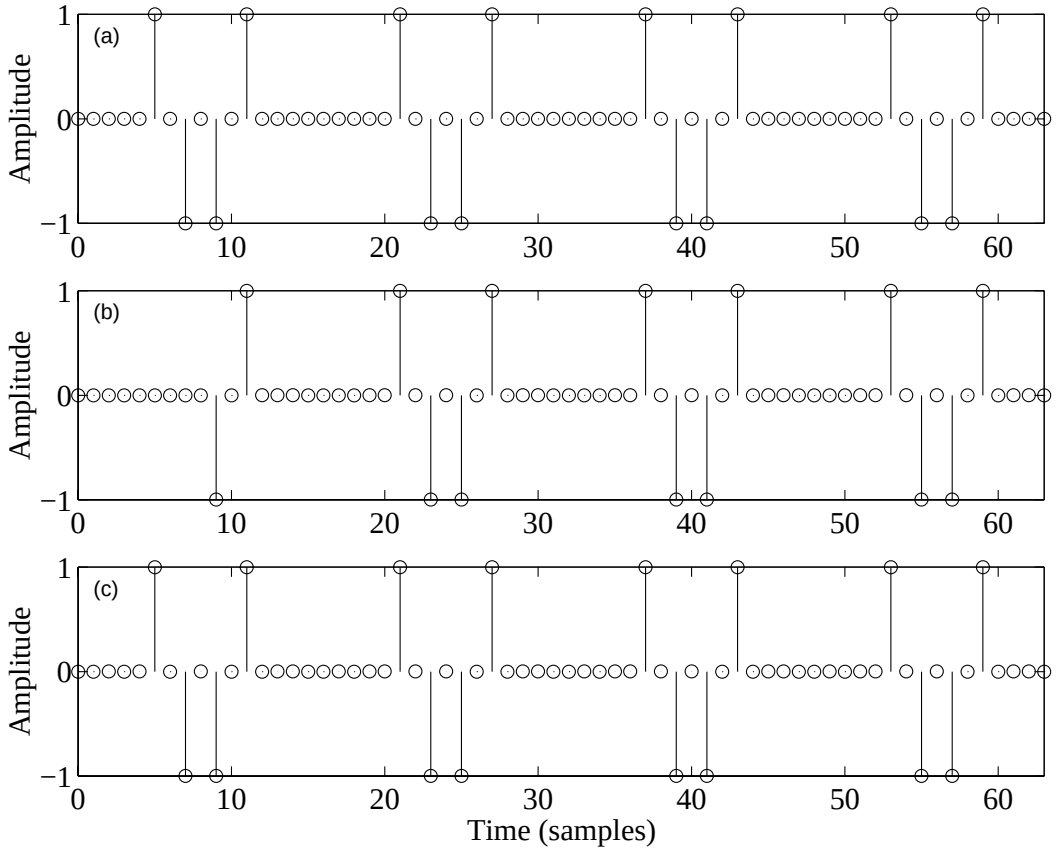

Figure 4.4: (a) The output of the digital waveguide, (b) the equivalent resonator structure with delay line, (c) and the resonators with circular delay.

As it can be seen in Fig. 4.4 (b), in the beginning of the response two impulses are missing, and therefore the output signal is not fully periodic. The periodicity can be retained when a "circular delay" is used. This means a circular convolution with a dirac function $\delta(n - M)$ in a space whose length is N. This can be done by changing the initial phases of the resonators with a value corresponding to a delay z^{-M}. The transfer function of the string takes the form:

$$\begin{aligned} \frac{v_{\text{out}}(z)}{F_{\text{in}}(z)} &= \left\{ \frac{a_1}{1 - z^{-1}e^{j\vartheta_1}} + \ldots + \frac{a_N}{1 - z^{-1}e^{j\vartheta_N}} \right\} \\ a_k &= -\frac{2}{Z_0 N} \sin(2k\pi \frac{M_{\text{in}}}{N}) \sin(2k\pi \frac{M - M_{\text{out}}}{N}) e^{-j\vartheta_k M} \end{aligned} \tag{4.10}$$

The impulse response of the system is shown in Fig. 4.4 (c). Now in Fig. 4.3 A will refer to a junction with input coefficients a_k to every branch and B reduces to a simple adder.

The impulse response of Eq. 4.10 can be formulated in the following way:

$$h(n) = \sum_{k=0}^{N-1} a_k e^{j\vartheta_k n} = \sum_{k=0}^{N-1} a_k e^{j\frac{2\pi}{N}kn} \tag{4.11}$$

which is the Inverse Discrete Fourier Transform of the a_k coefficients. Generally, when the poles of the resonators are equally distributed on the unit circle, the impulse response of the digital waveguide and the initial amplitudes and phases of the resonators are related by the Discrete Fourier Transform. This holds for the ideal string discussed above. This way the complex resonator coefficients a_k can be computed by taking the DFT of the first N samples of the impulse response.

In the digital waveguide the behavior of the string between the nodes can be calculated by interpolation [Smith 1992; Laakso et al. 1996]. The inverse operation of that is the deinterpolation [Välimäki et al. 1993; Välimäki 1995], where the excitation force is introduced by fractional delay filters between the nodes of the waveguide. In the resonator implementation of Eq. (4.9) there is no need for such operations since by choosing proper input and output coefficients for the resonators, M_{in} and M_{out} can be arbitrary rational numbers.

In the case of multiple inputs and outputs, the input and output signals form vectors $\mathbf{x}$ and $\mathbf{y}$, and the input and output coefficients can be arranged to matrices $\mathbf{A}$ and $\mathbf{B}$. The matrix $\mathbf{R}$ is diagonal and contains the transfer functions of the resonators. The output of the system is calculated in the following way:

$$\mathbf{y} = \mathbf{BRAx} \qquad R_{kk} = \frac{1}{1 - z^{-1}e^{j\vartheta_k}} \tag{4.12}$$

It is worth to note that in some cases the delay z^{-M} of Eq. (4.9) cannot be realized with a circular delay as shown in Eq. (4.10). Equation (4.10) is useful only when all the resonators are implemented. In the case of simulating some secondary effects of the piano sound, as it will be presented in Section 5.3, only some resonators are used. The solution for the problem will be discussed there.

In the simulation of string instruments, it is more important to know the force experienced by the termination (bridge) than the movement of the string. The reason for this is that the sound is mainly radiated by the body or the soundboard and not the string, since the impedance matching between the string and the air is inefficient. From Eqs. (4.4) and (4.5) the force at the bridge F_{br} will be:

$$\begin{aligned} F^+ &= v^+ Z_0 \\ F^- &= r_F F^+ = r_F v^+ Z_0 \\ F_{\text{br}} &= F^+ + F^- = (1 + r_F) Z_0 v^+ = 2 Z_0 v^+ \end{aligned} \tag{4.13}$$

The transfer function from the excitation point to the bridge is:

$$\frac{F_{\text{br}}}{F_{\text{in}}} = \frac{1}{1 - z^{-N}} \left(1 - z^{-2M_{\text{in}}}\right) z^{-(M - M_{\text{in}})} \tag{4.14}$$

The resonator coefficients a_k will take the form:

$$\begin{aligned} \frac{F_{\mathrm{br}}}{F_{\mathrm{in}}} &= \left\{ \frac{a_1}{1 - z^{-1} e^{j\vartheta_1}} + \ldots + \frac{a_N}{1 - z^{-1} e^{j\vartheta_N}} \right\} z^{-M} \\ a_k &= j \frac{2}{N} \sin(2k\pi \frac{M_{\mathrm{in}}}{N}) \end{aligned} \tag{4.15}$$

where $\vartheta_k = (2k\pi)/N$.

The output of Eqs. (4.14) and (4.15) will differ only in the first $M + M_{\mathrm{in}}$ taps. The delay z^{-M} can be eliminated from Eq. (4.15) in the same way as shown in Eq. (4.10). Then the impulse response of the resonators and of the digital waveguide will be the same from the first time instant.

4.2 Non-ideal termination

When the string is terminated by a finite impedance $Z > Z_0$, the absolute value of the reflection coefficient r_v will be somewhat lower than 1. If the termination is purely resistive (Z is real and frequency independent), r_v will be a real constant number. The waves traveling in the delay lines are multiplied by a constant factor every time they pass through the termination. Consequently, all the harmonics will decay exponentially with the same time coefficient τ. If one of the terminations is completely rigid and the other has an impedance of $Z > Z_0$, the decay time of the harmonics will be the following:

$$\tau = -\frac{1}{f_0 \ln r_F} \tag{4.16}$$

where r_F is the force reflection coefficient ($r_F = -r_v$), $f_0 = N/f_s$ is the fundamental frequency, f_s refers to the sampling frequency, and N to the total length of the delay line. The reflection coefficient r_F can be calculated using Eq. (4.5). Eq. (4.16) was presented with relation to the decay times of the Karplus-Strong algorithm in [Jaffe and Smith 1983]. If none of the terminations are completely rigid, the r_F used in Eq. (4.16) is the product of the reflection coefficients of the two sides. The equivalent resonator bank will change in the radius of the poles since they move slightly towards the origin:

$$\begin{aligned} \frac{F_{\mathrm{br}}}{F_{\mathrm{in}}} &= \left\{ \frac{a_1}{1 - z^{-1} r_1 e^{j\vartheta_1}} + \ldots + \frac{a_N}{1 - z^{-1} r_N e^{j\vartheta_N}} \right\} z^{-M} \\ r_1 &= \ldots = r_N = (r_F)^{\frac{1}{N}} \end{aligned} \tag{4.17}$$

where $\vartheta_k = (2k\pi)/N$, $N = 2M$ and r_k are the pole radii.

The approximate decay times can be calculated in the same way for frequency dependent rational impedance $Z(z)$ with the help of Eqs. (4.5) and (4.16) at the modal frequencies. If e.g. the impedance of the termination decreases with frequency, the decay times of the higher harmonics will be smaller than those of the lower ones.

If the slope of the function r_F is small, the modal frequencies will not differ from that of the infinitely rigid termination significantly.

By an arbitrary, complex termination the modal frequencies also change. A good approximation for these frequencies is:

$$\varphi(\vartheta_k) - \vartheta_k N = -2k\pi \tag{4.18}$$

where the first term $\varphi(\vartheta_k) = \arg(r_F)$ stands for the phase response introduced by the impedance, and the second term refers to the phase of the delay line. The physical meaning of Eq. (4.18) is that a standing wave of the frequency ϑ_k has to fit in the string $k/2$, or in the delay line of the digital waveguide k times.

If the amplitude of r_F is constant with respect to the frequency, Eq. (4.16) is exact. If not, the modal frequencies will be slightly influenced by the slope of $|r_F|$. In our cases this difference can be neglected, since $|r_F|$ is close to 1 and its slope is very small, thus it can be considered constant around the resonance frequencies of the digital waveguide.

If the impedance is not of an explicit form, the modal frequencies ϑ_k cannot be calculated directly. They can be estimated by the iterate use of Eq. (4.18) with a starting value of the ideal waveguide ($\varphi(\vartheta_k) = 0$). Precise results in determining the modal frequencies and decay times can be achieved by the fractional expansion of the whole transfer function. Because of the high order of the transfer function, numerical methods (such as the `residue` command in MATLAB) should be used, although these algorithms can easily give wrong results due to numerical instability.

4.3 The non-ideal string

In the general case, the wave equation of the string is similar to Eq. (4.1), but it has to be extended by higher-order spatial- and time-derivate terms. The odd-order time-derivates are responsible for the losses and the even-order spatial-derivates for the dispersion of the string [Smith 1993]. The latter means that the wave velocity of the stiff string will vary with the frequency. As a result, the partials will not be harmonic anymore.

Since we are observing the behavior of the string between two points (the input and the output), the losses and dispersion can be lumped to one point [Smith 1992]. As it was shown in the previous section, the influence of the termination can be easily taken into account by a frequency dependent reflection coefficient r_v. Hence, it is beneficial to use the reflection coefficient for implementing the dispersion and losses of the string as well. The reflection coefficients of the two sides can be transformed to one side of the string ($r_v = -r_{v\text{left}} \cdot r_{v\text{right}}$), therefore on the other side only a multiplication by -1 is needed (see Fig. 4.5). This new reflection coefficient r_v will contain the losses and dispersion of the string on the one hand and the effect of the terminations on the other. Our task now is to determine this reflection coefficient from measurement data or analytical expressions. If it succeeds, the digital waveguide will behave in the same way as the original system.

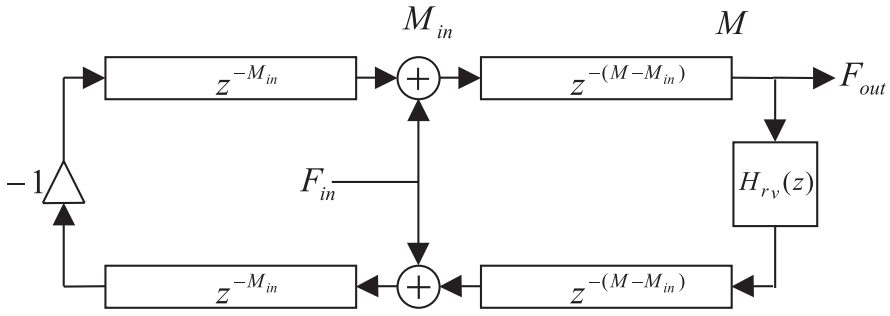

Figure 4.5: Digital waveguide model of the non-ideal string.

4.4 Equivalent resonators for lossy and dispersive waveguides

In the former sections, the resonators were presented rather for analytical reasons, making the behavior of the digital waveguide easier to understand. On the other hand, there are some instruments where the resonators can be more efficient. The resonators have been also used for sound synthesis purposes. For example Laroche and Meillier [1994] used them for the synthesis of piano sound, with a common excitation signal. The author of this thesis has found the digital waveguide more appropriate for the synthesis of piano sound, even for the high register of the piano. In the case of gongs and bells, where the inharmonicity is high, the case can be the contrary.

In the literature, the resonators are mostly treated as linear filters. The excitation signal is precomputed or given. On the other hand, the resonator structure presented here is capable of simulating nonlinear interaction as well. The nonlinear hammer model proposed in Section 5.2 could be connected to a resonator structure, since with the circular delay approach its impulse response is identical to the one of the digital waveguide. The semphasis is now on the delay, as the excitation mechanism of most instruments can be simulated with a nonlinear, time-invariant function and a feedback loop with delay [McIntyre et al. 1983]. It has to be noted that synthesizing the sound by means of resonators is somewhat similar to the modal synthesis approach presented in [Adrien 1991].

In the case of a general string model, where the poles of the transfer function are not on the unit circle, this delay can be still implemented by proper output coefficients for the resonators. With N second-order resonators, $2N$ samples of the impulse response can be determined by setting the initial amplitudes and phases. This is because the outputs of the resonators form linearly independent vectors. These $2N$ samples are typically the beginning of the impulse response, but could be anywhere else, although there are some constraints if the impulse response is periodic.

The analysis procedure for a general, nonlinear, resonator-based model can be

the following: first, the partial frequencies and decay times are determined. Then the impulse response of the acoustical resonator (string, bell, etc.) is estimated. By using the frequency and decay time parameters of the N resonators, together with $2N$ samples of the impulse response, $2N$ linear equations can be written. One resonator corresponds to two equations, since both the initial amplitude and phase are free parameters. These parameters can be determined by using resonators, e.g., with initial amplitude of 1 and initial phases of 0 and 90 degrees in the linear equations.

4.5 The digital waveguide as an approximation

By reversing the order of ideas, the digital waveguide can be viewed as the approximation of the resonator bank. If the fractions of the resonators are brought to a common denominator, the resulting system will be an exact copy of the resonator bank with respect to the input-output behavior. In the general case, the delay line disappears and the whole digital waveguide will become a reflection-coefficient filter with feedback. In this case the number of operations remains the same. On the contrary, if we let the specification for the modal frequencies and decay times to be somewhat loose, the computational complexity can be radically reduced. Nearly harmonic signals can be produced by using a delay line and, e.g., a 10^{th}-order reflection filter instead of using 100 resonators. While in the resonator-bank implementation the computational complexity depends on the number of the harmonics, in the digital waveguide it depends on the precision of the approximation. The human auditory system seems to be more tolerant for slight inaccuracies than for completely missing harmonics. Consequently, the digital waveguide is a more efficient tool for modeling nearly harmonic sounds. On the other hand, when the generation of a small number of partials is needed, the resonator bank can be a more efficient approach. This is the case for example when simulating the two-stage decay of the piano sound.

4.6 Conclusion

We overviewed a computationally very efficient technique for modeling the string behavior, the digital waveguide model. Its efficiency comes from the fact that losses and dispersion of the string-termination system are lumped to one point. Therefore, the string simulation loop consists of two delay lines and one reflection filter. By mathematically reformulating the transfer function of the digital waveguide, the equivalent resonator bank structure was presented. Methods for calculating the parameters of the resonator bank were also given. A new feature of this approach is that the resonator bank can be used for simulating nonlinear interaction, since its behavior is the same as that of the digital waveguide. This can be useful for the modeling of instruments with high inharmonicity, such as mallet percussion instruments, gongs, or bells.

Chapter 5

Model structure

First, the general structure of the piano model is presented. Then separate parts of the model are described in detail. A new multi-rate hammer model and a novel technique for simulating beating and two-stage decay are proposed. A simple technique for the realistic simulation of the attack sound of the piano is also outlined. The last part of the chapter deals with the issues of practical implementation and presents a multi-rate piano model. The calibration of the model and the estimation of the parameters are discussed in Chapter 6.

5.1 General structure of the piano model

Since the physical modeling approach tries to simulate the structure of the instrument, and not the sound itself, the piano model consists of the same parts as the real piano. The structure of the model is displayed in Fig. 5.1. The first step of the sound production mechanism is the excitation, which is the hammer strike in the case of the piano. The resulting signal propagates to the string, which determines the fundamental frequency of the tone. The periodic output signal is filtered through the radiator, covering the effect of the soundboard.

Figure 5.1 shows that the interaction between the string and the excitation is bidirectional, since the hammer force also depends on the string displacement [Fletcher and Rossing 1998]. On the other hand, there is no feedback from the soundboard to the string. Although the impedance of the bridge and the soundboard influence the decay times of the partials, this effect is taken into account in the string model. At this point, our model differs from the real piano: the two functions of the soundboard, namely, the determination of the decay times and forming the spectrum, are put to separate parts of the model. This way, the effect of the soundboard can be treated as a linear filtering operation [Smith 1983; Välimäki et al. 1996].

The hammer model solves the differential equation of the hammer in discrete time. This part of the model is responsible for shaping the initial spectrum, by providing the excitation to the string model. On the other hand, it has no influence on the time dependent evolution of the spectra in the decay part of the sound. Its

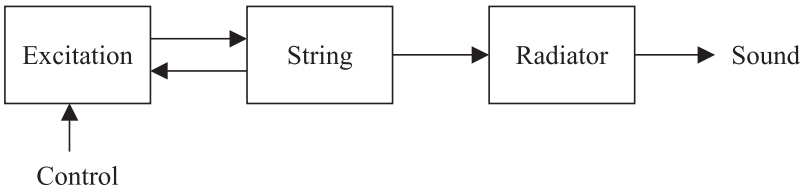

Figure 5.1: General structure of the piano model.

main role is varying the spectrum according to the dynamic level. This comes from its nonlinear behavior. It is an advantage that the input parameter of the hammer model is the impact velocity, which makes the control of the model straightforward and physically meaningful. In a nonphysical implementation the model parameters should be analyzed at every dynamic level and then used for resynthesis. In our case it is not needed, since the model gives the correct force input signal automatically.

The string is modeled by a digital waveguide, where the losses and the dispersion of the string are consolidated to one point [Smith 1987, 1992]. Its task is to control the frequencies of the partials and the decay times. It also influences the spectra by a comb filtering effect. The latter is a physically meaningful feature: those modes, which have a node at the excitation point, will be missing from the spectrum. It comes also from the principle of the digital waveguide that the displacement of the string can be computed at every point. This is crucial at the hammer position, since the displacement of the string is the feedback signal to the nonlinear hammer model. Lumping the losses and dispersion of the string has no physical sense, but it increases the computational efficiency of the model significantly [Smith 1987, 1992].

For secondary effects, such as beating and two stage decay, a resonator bank is used in parallel to the digital waveguide. This can be considered a linear filter, and it rather concentrates on the resulted signal and not on the structure. The string model is a transition between the physical and general signal models.

The simulation of the soundboard has no physical interpretation: it is simulated by a special linear filter, similar to that used for reverberation algorithms. By using a nonphysical model, the flexibility that the sound pressure generated by the soundboard is known at any point in the acoustic space, is lost, but the computational complexity can be reduced radically. Since the parameters of the soundboard do not change during playing, the nonphysical approach is adequate. The impulse response of the soundboard depends only on which string excites it, and therefore it can be considered a linear time-invariant multi-input, single-output system.

The model structure described here is a combination of different approaches. Physical modeling exhibits the most benefits in the realistic control of the parameters. Since in the case of the piano the only input parameter is the impact velocity, it is worthwhile to implement the hammer-string interaction as a physical model. For the other parts of the model, which have static parameters, the most effective implementation should be found, even if it is not analogous to the real structure of

the piano.

5.2 Modeling the hammer and the dampers

First, the different hammer models found in the literature are reviewed. Then, the problems for feeding the interaction force into the digital waveguide are discussed, and for avoiding these problems, new solutions are proposed. After that, the hammer and damper models are presented[1].

The hammer is modeled by a linear difference equation and an instantaneous nonlinearity. The difference equation calculates the movement of the hammer and the nonlinear part accounts for the felt characteristic. An advantage of the nonlinear model is the dynamic variation of the timbre, according to the impact velocity. A novel multi-rate hammer model is presented for assuring the stability of the model. The effect of dampers is simulated with a linear filter. A new nonlinear damper model is also proposed, which is based on the hammer model.

5.2.1 Overview of prior work

The hammer is generally considered as a mass connected to a nonlinear spring [see, e.g., Boutillon 1988]. As it has an initial velocity, it hits the string. The spring compresses, and the interaction force pushes the hammer away from the string. The most straightforward approach for modeling the hammer is discretizing the differential equation of the mass. The position of the hammer is obtained by multiplying the interaction force with a scalar and integrating twice with respect to time. The nonlinear felt characteristic (see Eq. 2.1) can be directly implemented in a discrete form. This approach was taken in [Chaigne and Askenfelt 1994a,b], where the hammer was connected to a finite-difference model of the string. Simulation results showed good agreement with the measured data. Borin et al. [1992] applied the same kind of power law model to the digital waveguide. An advantage of this technique is its simplicity, and, as it is a nonlinear model, the spectrum of its output varies dynamically according to the impact velocity. The problem with this straightforward approach is that for high stiffness and initial velocity values the hammer model can become unstable, mostly in the high note range of the piano [Borin and De Poli 1996; Borin et al. 1997].

A solution to this problem can be found in [Borin and De Poli 1996; Borin et al. 1997]. It is based on the separation of known and unknown terms, and solving an implicit equation. For integer stiffness exponents ($p = 2$, 3, or 4 in Eq. 2.1) an analytic solution of the implicit equation can be derived. They also introduce the hysteresis to their model by discretizing the equations of [Stulov 1995]. The drawback of the model is that the computation of the interaction force is quite complicated, even for a stiffness exponent value $p = 2$. The model cannot be used

[1] parts of this section are also published in [Bank 2000]

with noninteger p values, since the solution of an implicit equation would take too many computations.

An elegant solution for the hammer model based on the traveling wave decomposition of the mass-spring system was proposed in [Van Duyne et al. 1994]. There, a distributed hammer model was attached to the string by a scattering junction. The distributed model was derived for a linear spring. The nonlinear characteristic of the felt was taken into account by reading the stiffness coefficient from a lookup-table, according to the compression of the felt. Hysteresis was modeled by offsetting the pointer in the table, corresponding to the velocity of felt compression. The disadvantage of this technique lies in its complexity.

In [Smith and Van Duyne 1995; Van Duyne and Smith 1995], a linear hammer model was used. The hammer was modeled by a linear filter, whose parameters were determined by nonlinear simulation. The advantage of the linear approach is that the soundboard filter can be commuted through the string and the hammer [Karjalainen and Välimäki 1993; Smith 1993]. Hence, the soundboard need not be implemented as a filter, but as a wavetable, whose content is played into the string through the hammer filter. This reduces the computational complexity of the soundboard model largely. A drawback is that the hammer filter has to be designed for all of the strings and at every dynamic level. Accordingly, the nice feature of the nonlinear model, that it responds to the initial velocity in a physically meaningful way, is lost. Moreover, the restrike of the string cannot be simulated correctly.

5.2.2 The discontinuity problem: nonlinear interaction in the digital waveguide

There are two different tasks when connecting the hammer to the string. One is introducing the interaction force to the digital waveguide and the other is determining the displacement of the string, since that is the feedback signal to the hammer model. These can be done as shown in Fig. 4.2, where $M_{out} = M_{in}$ for this case. The displacement of the string can be determined by integrating the velocity signal v_{out}. However, there is a problem with this simple approach. Here we will deal with the problem thoroughly, since it has not been discussed in the literature.

Let us assume that the string is infinite, or terminated by a Z_0 impedance, i.e., there is no reflection from the terminations. There are two possibilities for connecting the hammer model to the digital waveguide, according to Fig. 4.2. One is first to read the string velocity from the delay lines and then to add the excitation signal to the cells. After this, the delay lines are shifted. In this case, the velocity which is read from the cells at the hammer position will be always zero. Consequently, all parts of the string will be moving, but not the excitation point. Since the feedback to the hammer model is coming from the displacement of the excitation point, the hammer will behave as if it was bouncing to a rigid wall. Obviously, this method is not appropriate.

The other approach is the opposite: first we add the excitation signal to the delay cells at the position of the hammer, then read the string velocity, and shift

the delay lines. Now the problem is that the string velocity and displacement at the excitation point will be twice the value of any other cells. This is illustrated in Fig. 5.2 (a), where $T_s = 1/f_s$ is the sampling period. Assume that a discrete unit impulse is added to the cells corresponding to the hammer position, and the delay lines of the digital waveguide are consecutively shifted. The velocity of the string, which is displayed in the figure, is calculated by summing the content of the delay cell pairs corresponding to the same spatial position. Note that the velocity of the excitation point will be twice that of the impulse. However, as this pulse travels further in the delay lines, the velocity values will be the same as the pulse amplitude. Integrating the string velocity in the time domain shows a discontinuity in the string displacement in Fig. 5.2.

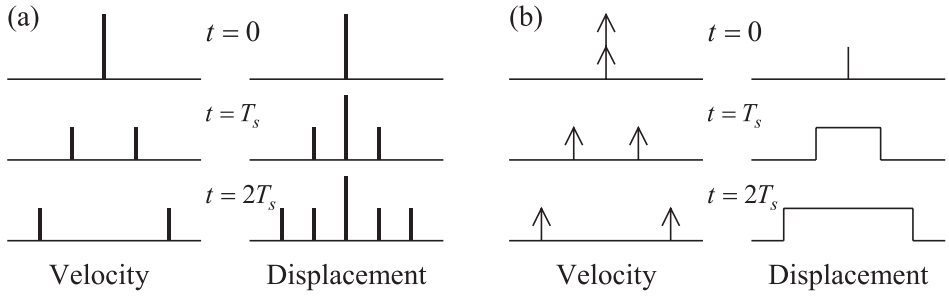

Figure 5.2: String velocities and displacements: (a) discrete and (b) continuous case.

The continuous string does not show this behavior: Fig. 5.2 (b) reveals the velocity and the displacement of a string excited with a continuous-time Dirac velocity impulse. The velocity and displacement values of the string are displayed for the same time instants as the discrete model, and the arrows refer to Dirac impulses. The difference comes from the fact that although there are two continuous Dirac pulses in the region corresponding to the spatial sample of the excitation position, these stay within this region only a half time step. They travel from the midpoint of this section to its borders within $T_s/2$. Thereafter, in every spatial section of the string there will be only one impulse, staying for T_s. Since in the discrete model an impulse cannot be in a cell for less than the sampling period, we have to diminish the amplitude of the impulse by a factor of two at the interaction position. Nevertheless, the amplitude of the impulse in the other cells should not be altered.

This can be done by first adding only half of the excitation to the delay lines at the hammer position M_{in}, and after one time sampling-interval, when the impulses moved further, we can add the remaining part. The corresponding structure is illustrated in Fig. 5.3. An equivalent and simpler method is to first add a pulse at position M_{in} to the upper delay line only, and one sampling period later to the lower one, but at the position $M_{in} - 1$.

Another solution can be keeping track of the string velocity at the excitation

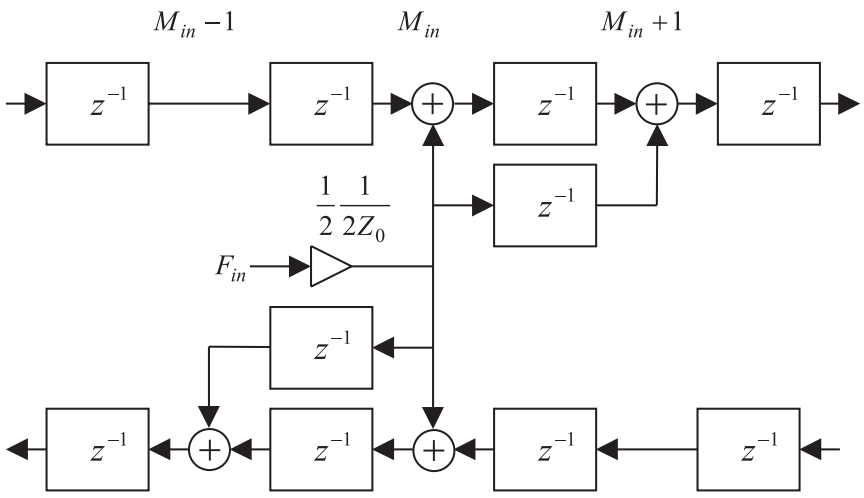

Figure 5.3: Proposed method for the correct simulation of force input.

point separately. This way, the displacement of the string at the excitation position is not calculated directly from the digital waveguide, but with the help of a separate variable.

Note that for linear synthesis, when there is no feedback from the excitation position, the digital waveguide behaves properly, and therefore the procedures proposed for avoiding the discontinuity problem are not needed. Probably this is the reason why this question has not been dealt with in the literature. However, in the case of nonlinear interaction, the difference between the output of the earlier, simple method (Fig. 4.2, $M_{out} = M_{in}$) and of the proposed one (Fig. 5.3) is dramatic. This is shown in Fig. 5.4. The parameters of the hammer and the string were taken from [Chaigne and Askenfelt 1994a], C_4 note, and the impact velocity was set to 4 m/s. The solid line shows the interaction force of the finite-difference implementation, as discussed in [Chaigne and Askenfelt 1994a], but the stiffness of the string was set to zero. The dash-dotted line shows the hammer force when the same hammer model is connected to lossless, nondispersive digital waveguide in the traditional way. This differs from the force of the finite difference method largely. The dashed line displays the interaction force when the structure of Fig. 5.3 is used. The force curve is now close to that of the finite difference method.

The method based on discretizing the differential equation of the hammer works well for the low and middle range of the piano, but for the high notes with large impact velocities the model becomes unstable. This is because the assumption, that the interaction force changes only a little in one step, is no longer valid for those cases. This was also noted by Borin and De Poli [1996], but their solution to the problem, which was based on the separation of known and unknown terms and solving an implicit equation, seems to be too complicated.

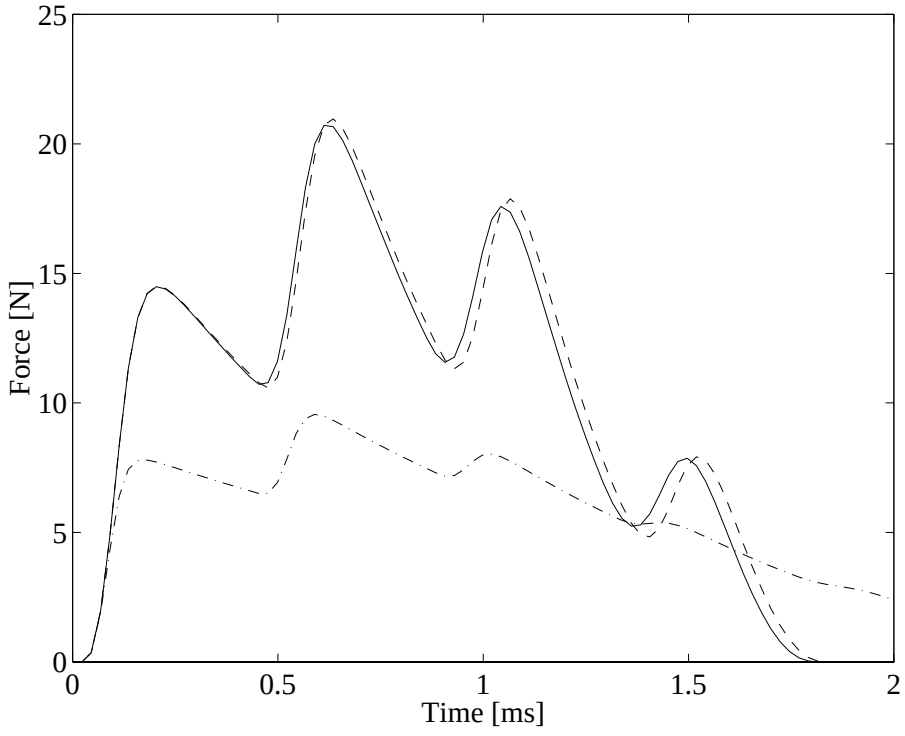

Figure 5.4: Interaction force for note C_4 calculated by finite difference method of Chaigne and Askenfelt [1994a] (solid line), by a digital waveguide using the traditional technique of Fig. 4.2 (dash-dotted line), and by using the proposed method of Fig. 5.3 (dashed line).

5.2.3 The multi-rate hammer

Here a novel multi-rate hammer model is proposed, which overcomes the stability problems. The idea comes from the fact that by increasing the sampling rate of the whole string model, the instability can be avoided. The hammer model is based on the discretization of a differential equation. It is stable when the variables change only a little in every sampling interval. The stability of such a system can be always maintained by choosing a sufficiently large sampling rate, assuming that the analog system was stable. When the sampling period converges to zero, the discrete system will behave as the original continuous differential equation.

Unfortunately, increasing the sampling rate by a factor of two of the whole string model would double the computation time as well. Nevertheless, if only the hammer model operates at a double rate, the computational complexity is raised by a negligible amount. Therefore, in the solution proposed here the string model operates at normal, but the hammer model runs at double sampling frequency. The process is the following: first the incoming string velocity signal has to be upsampled by a factor of two, then the hammer model is executed twice. The force input to the string is calculated by downsampling the output signal of the hammer model. Since

the hammer model needs the string displacement in between the time steps of the digital waveguide, the string displacement is kept track within the hammer model. This also overcomes the problem discontinuity discussed in the earlier section. The core of the hammer model is displayed in Fig. 5.5.

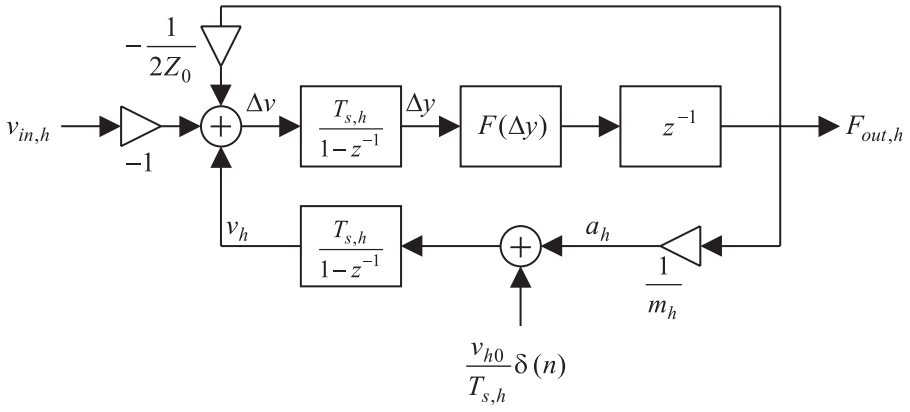

Figure 5.5: The core of the proposed hammer model.

The hammer model of Fig. 5.5 first computes the velocity difference of the string and the hammer $\Delta v = v_h - v_{in,h} - F_{out,h}/(2Z_0)$, where $v_{in,h}$ is the incoming string velocity, v_h is the velocity of the hammer, Z_0 is the string impedance, and $F_{out,h}$ is the force signal computed by the power law in the previous time instant. Then, the felt compression Δy is calculated by integrating Δv with respect to time. The integrators used here are obtained by the impulse-invariant transform of the continuous time integrator [Oppenheim and Schafer 1975]. The interaction force is computed by a power law of Eq. (2.1), whenever Δy is positive. For $\Delta y < 0$, that is, the hammer is below the string, the interaction force is zero. The delay z^{-1} is inserted for computation purposes. The velocity of the hammer v_h is calculated by integrating the hammer acceleration $a_h = F_{out,h}/m_h$, where m_h is the hammer mass. The initial velocity v_{h0} of the hammer is controlled by sending an appropriate acceleration pulse to the integrator, or by setting the initial value of the corresponding delay cell to v_{h0}.

Generally, the hammer model is a single-input single-output nonlinear filter. Alternatively, it can be seen as a linear filter with one signal dependent coefficient. Then, $F(\Delta y)$ is replaced with a multiplier with a coefficient $a(\Delta y) = F(\Delta y)/\Delta y$ in Fig. 5.5. It would be interesting to analyze the model from this point of view, in order to get some information about its stability.

When operated at a normal sampling rate, the model of Fig. 5.5 can be directly connected to the digital waveguide model of Fig. 4.2, by setting $v_{in,h} = v_{out}$, $F_{in} = F_{out,h}$, and $M_{out} = M_{in}$. The sampling rate of the hammer model now equals to that of the entire system, $T_{s,h} = T_s$. It works as follows: first the cells at position M_{in} are read and used as the input $v_{in,h}$ of the hammer model. Then, the hammer model

computes the force input F_{in} to the string, and this value is added to the delay lines. After that, the delay lines are shifted. Since the string velocity $v_{in,h} + F_{out,h}/(2Z_0)$ is computed inside the hammer model, the problems mentioned in the beginning of section 5.2.2 are avoided. However, this model is still running at normal sampling rate. An interface has to be made between the hammer and the string to be able to operate them at different sampling rates. This is done by up- and downsampling, as illustrated in Fig. 5.6.

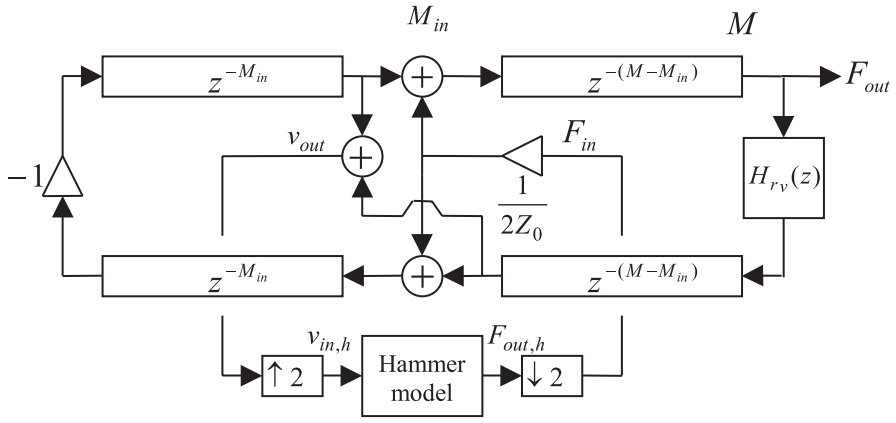

Figure 5.6: Connecting multi-rate hammer model to the digital waveguide.

In the proposed implementation, the core of the hammer model runs at a double sampling rate, that is, $T_{s,h} = T_s/2$. The simplest way to implement the upsampling ($\uparrow 2$ in Fig. 5.6) operation is by using zeroth-order interpolation, i.e., repeating the sample for the unknown time instant, $v_{in,h}(nT_s+T_s/2) = v_{in,h}(nT_s)$. A more accurate solution can be obtained by linear interpolation [Schafer and Rabiner 1973]. In this manner, the unknown samples will be the average of two consecutive known values. To be able to do this without introducing a delay, one should know the next incoming sample. This is easy in the case of the digital waveguide, since the upcoming values at the excitation point are already in the delay lines, exactly one time-step away (see Eq. 4.3 and Fig. 4.1). Hence, the input for the hammer model can be calculated using linear interpolation for upsampling by the following equations:

$$\begin{aligned} v_{in,h}(nT_s) &= v_{out}(nTs) = y^+(n, M_{in}) + y^-(n, M_{in}) \\ v_{in,h}(nT_s + T_s/2) &= \frac{v_{out}(nT_s) + v_{out}(nT_s + T_s)}{2} \\ &= \frac{y^+(n, M_{in}) + y^+(n, M_{in} - 1)}{2} + \\ &+ \frac{y^-(n, M_{in}) + y^-(n, M_{in} + 1)}{2} \end{aligned} \tag{5.1}$$

where $y^+(n, m)$ and $y^-(n, m)$ refer to the content of the upper and lower delay lines, at the time instant n and position m, respectively.

The force input for the string is computed by simply averaging the two output samples of the force model, i.e., $F_{in}(nT_s) = (F_{out,h}(nT_s) + F_{out,h}(nT_s + T_s/2))/2$.

The analysis of the new method

The multi-rate method has proven to be a robust solution for modeling the hammer. For note C_7 (2090 Hz), the hammer starts to be unstable for an impact velocity as much as 20 m/s, which is three times of the maximal velocity which can arise in a real piano. By using the single-rate model, it is already unstable at around 5 m/s.

In Fig. 5.7 the interaction force is shown for note C_5 (522 Hz). For the simulation, an ideal digital waveguide model was used, without any dispersion or losses. The parameters of the hammer were taken from [Chaigne and Askenfelt 1994a], C_4 hammer. The impact velocity was $v_{h0} = 6$ m/s. The dash-dotted line refers to the single-rate hammer model with $f_s = 44.1$ kHz. The solid line shows the force of the single-rate model, but the whole digital waveguide model is run at a double sample rate, that is, $f_s = 88.2$ kHz. This is our reference structure. The dashed line in Fig 5.7 is the force of the multi-rate implementation, by using $f_s = 44.1$ kHz for the waveguide model. It can be seen that the traditional technique operating at normal sampling rate goes unstable, while the output of the proposed multi-rate hammer model coincides with the output of the single-rate model operating at double sampling frequency.

Surprisingly, the multi-rate method is more stable than the single-rate operating at double sampling rate. For the same string and hammer parameters as in the previous example, the single-rate model at double sampling frequency goes unstable for $v_{h0} > 18$ m/s, the multi-rate model is still stable until $v_{h,0} = 42$ m/s. This is because the multi-rate model functions like the single-rate model running at double sampling frequency, but with lowpass-filtering at the input and output of the hammer model. This comes from the up- and downsampling: the hammer model operates at $2f_s$, but its input $v_{in,h}$ is upsampled from a signal v_{out} whose sampling rate is f_s. Consequently, the input cannot contain strong frequency components higher than $f_s/2$. In this way, the incoming high frequency components, which are mainly responsible for stability problems, are suppressed.

Hysteresis can be easily included in the model by discretizing the equations of [Stulov 1995]. This was done by the bilinear transform in [Borin and De Poli 1996]. The first-order IIR filter obtained that way is then cascaded with the instantaneous nonlinearity $F(\Delta y)$ in Fig. 5.5.

To conclude, the new multi-rate hammer model proposed here, overcomes the discontinuity problem discussed in Section 5.2.2. This is done by computing the string displacement at the excitation point within the hammer model. The stability of the hammer model is ensured by the multi-rate approach. This makes the model more stable since higher sampling rate yields a better discrete time approximation of the differential equation. It is also because the high frequency components of the input signal, which are mainly responsible for instability problems, are suppressed by the linear interpolation. Another benefit of the model is that it is simple, easy to

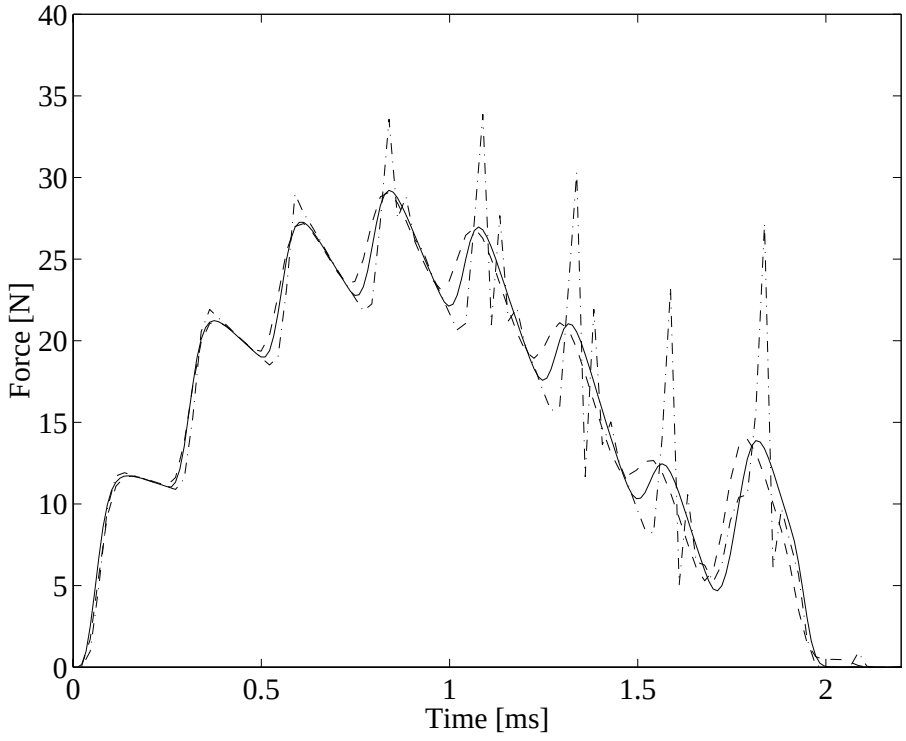

Figure 5.7: Simulated interaction forces for note C_5 (522 Hz), $v_{h0} = 6$ m/s, computed by the single-rate model (dash-dotted line), the single-rate model operating at double sampling frequency (solid line), and the multi-rate model (dashed line).

implement, and any kind of nonlinear function can be used as a felt characteristic $F(\Delta y)$.

5.2.4 Modeling the effect of the damper

Not much work can be found on modeling the piano dampers in the literature. A simple method, also suggested in [Van Duyne and Smith 1995], is cascading a real coefficient to the reflection filter, i.e., decreasing the gain of the loop filter of the string model. The author have found the value $c = 0.8$ a good approximation for all the notes. This ensures that the lower notes will damp more slowly. However, the damping, which arises when one uses such a simple method, is too clean. For the high range, where the damping is fast, it works well, but in the middle and low registers of the piano more refined methods are needed.

A development to this simple model can be using a separate damping filter, similarly to what was done for the guitar in [Erkut et al. 2000]. This filter can be implemented by cascading a new damping filter to the loop filter, or by changing the loop filter coefficients. However, changing the coefficients could introduce some transients. An easy way to overcome this problem is to run a damping filter in series

with the loop filter. When the key is released, one has to simply switch from the output of the loop filter to the output of the damping filter, whose state variable is already trained by the incoming samples [Välimäki and Laakso 1998]. The one-pole filter approach gives good results for the high and middle register, but the roughness of the damping of low piano strings is not reproduced.

5.2.5 The nonlinear damper model

The problem of these simple approaches, namely, changing the loop gain or cascading a damping filter, is that they are unable to simulate a characteristic feature of the damping. The feature is that every 7^{th} partial is damped inefficiently, since the damper cannot act well on a mode which has a node at the damper position (see Fig. 6.1). The difference can be heard especially at the lowest two octaves of the piano. By using the physical modeling approach, this feature could also be modeled.

The proposed nonlinear damper model operates with the same architecture of Fig. 5.5, although its parameters have to be changed. As the hammer hits the string from downwards, and the dampers fall on the string from upwards, the signs of the signals in the input and output of the hammer model have to be inverted. The gravity force, which pushes the damper to the string, can be simulated by adding a constant value to the hammer acceleration at every time instant. The m_h value and the $F(\Delta y)$ function have to be also changed. The initial velocity of the damper v_{h0} is set to zero. Since the main feature of the damper is dissipating energy, losses have to be introduced to the model. The easiest way to do that is adding hysteresis to the felt, by cascading a highpass filter with the output of the nonlinearity, as was done for the hammer in [Borin and De Poli 1996]. Another way can be to render the interaction force dependent on the velocity difference Δv of the string and the hammer. Then the force is calculated by $F(\Delta v, \Delta y)$.

The main benefit of the model is that the slower damping of every partial, which has a node at the damper position, is modeled automatically. This comes from the principle of the physical modeling approach. Another advantage is that now the hammers and the dampers are simulated by the same part of the piano model. A serious drawback is that the calibration of such a physical damper model can be problematic. Unfortunately, there has been no data in the literature concerning the damper characteristics of real pianos. This model has not been implemented in practice, due to the previous reason. However, after conducting some measurement on piano dampers, the author plans to experiment with this nonlinear model.

5.3 The string model

The behavior of a piano string is simulated by a digital waveguide, mostly because of practical reasons. The equivalent resonator structure presented in Chapter 4 could give the same sound output, but it is computationally more expensive. It will be shown later in Section 5.3.3 that the resonators can be useful in modeling some

details of the piano sound.

As discussed previously in connection with the hammer model, the digital waveguide has an important role in determining the excitation signal. Its other task is linear filtering: it controls the frequencies and decay rates of the partials and it introduces a comb-filtering effect, thus also affects the short-time spectrum. If the samples of the excitation signal are recorded and then fed to the digital waveguide, the same output will arise. This way the two functions of the digital waveguide, namely, determining the excitation and linear filtering, can be separated. In this chapter, for didactical reasons, the digital waveguide is considered as a linear filter and the excitation signal is assumed to be given.

5.3.1 The basic string model

The basic string model consists of a digital waveguide shown in Fig. 4.5 in Chapter 4. The fundamental frequency of the note and the frequency distribution of the partials are determined by the delay line length M and the phase delay of the reflection filter H_{r_v}. The decay times of the partials are controlled by the amplitude transfer function of the reflection filter. The estimation of these parameters and methods for reflection filter design will be discussed in the next chapter. The basic string model is capable of simulating pure exponentially decaying envelopes. Therefore it is unable to reproduce the beating and two-stage decay found in real piano tones. The physics of these two phenomena were discussed in Chapter 2.

5.3.2 Methods for beating and two-stage decay simulation

Here we review the methods previously presented in the literature.

By using two digital waveguides in parallel for one note, the beating and two-stage decay can be simulated in a simple way. If the fundamental frequencies of the two digital waveguides are different, beating results, and if their decay times are different, two-stage decay will occur in the sound. The two digital waveguides can be coupled with real coefficients [Karjalainen et al. 1998]. Figure 5.8 shows this configuration, where $S_v(z)$ and $S_h(z)$ are the two string models. The g parameters control the input and output amplitudes of the digital waveguides and determine the amount of coupling. The problem with this method is that the envelopes of the different partials will be similar, which is not the case for real piano tones. It also increases the computational complexity by a factor of two compared to the basic string model.

Another approach was taken by Smith [1993]. In his approximation, two ideal strings are coupled to the same termination, and all the losses are lumped to the bridge impedance Z. The structure is shown in Fig. 5.9. This comes from the assumption that all the losses come from the bridge, which is a rough approximation. By using one coupling filter H_c for calculating the velocity of the bridge, the reflected velocity waves v_1^- and v_2^- of the strings can be computed by subtracting the incoming waves v_1^+ and v_2^+ from the bridge velocity v_{br}. The allpass filters H_{ap_1} and H_{ap_2} still

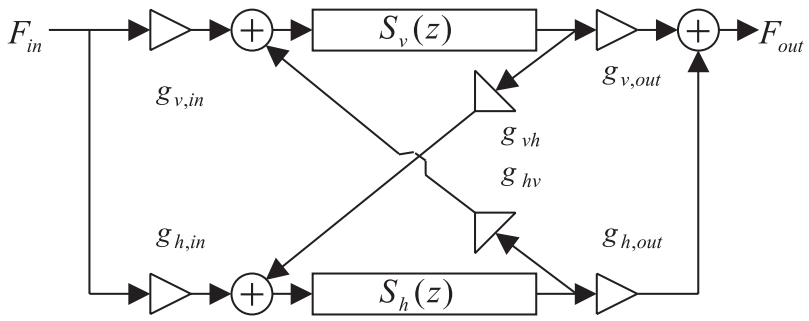

Figure 5.8: Realization of beating and two-stage decay with two coupled string models, presented in [Karjalainen et al. 1998].

have to be implemented separately. The advantage of this technique is that only one loss filter H_c is needed, whose transfer function can be determined from the decay times of the partials. The drawback is that the decay times and the coupling of the modes are not independent. The reason for this is that shorter decay times correspond to lower terminating impedance. Thus, the terminating impedance will decrease with frequency to a great extent resulting in stronger coupling for the higher partials. On the contrary, the average impedance of real piano bridges is nearly constant.

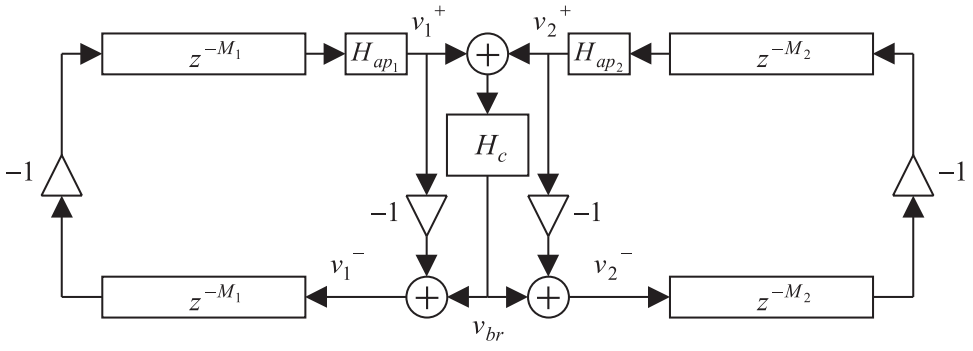

Figure 5.9: Coupling filter implementation of beating and two-stage decay by Smith [1993].

A more general method was presented in [Daudet et al. 1999]. The approach is similar to that of Fig. 5.8, but the coupling between the two string models is complex and frequency dependent. As displayed in Fig. 5.10, two basic string models $S_v(z)$ and $S_h(z)$ are coupled to each other with two coupling filters $H_{v,h}(z)$ and $H_{h,v}(z)$. In this manner the evolution of the partials are controlled by four independent filters. The filter parameters were computed from the amplitude envelopes. The use of four independent filters increases the degrees of freedom, and many different envelopes

can be simulated. However, it makes the filter design procedure quite complicated and boosts the computational complexity heavily.

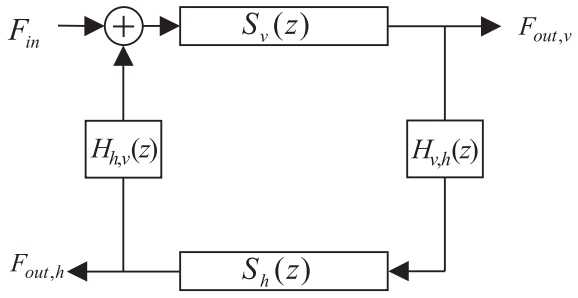

Figure 5.10: Coupled waveguides model suggested by Daudet et al. [1999].

Somewhat similar but computationally less expensive implementation was presented in [Borin et al. 1997]. There second-order loss filters were used for the different strings and the strings of all notes were coupled to the same impedance. This structure is illustrated in Fig. 5.11, where $v_1^+ \ldots v_n^+$ are the incoming velocity waves from the digital waveguides, and $v_1^- \ldots v_n^-$ are the reflected ones. $Z_{0,1} \ldots Z_{0,n}$ refer to the characteristic impedance of the strings $1 \ldots n$, respectively. The method for designing such a common coupling filter $H_c(z)$ was not discussed there. Probably the easiest way is to use the data of impedance measurements of the bridge, and not the envelopes of the partials. The problem with their approach is that the behavior of the different strings are not independent, since only one coupling filter is used. Furthermore, all the strings are coupled to each other in the same proportion, unlike in the real piano. In reality, those strings that are closer to each other are coupled more strongly. The advantage of the method is its low computational cost.

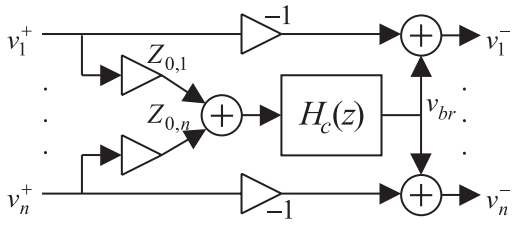

Figure 5.11: Coupling n strings to one terminating impedance, as presented in Borin et al. [1997].

5.3.3 Novel resonator bank implementation for beating and two-stage decay

As was shown by Weinreich [1977] a two mode model with two exponentially damping sinusoids can give a good approximation for the beating and two-stage decay of the piano sound. All the frequencies, decay times, initial amplitudes and phases of the two modes can be different. These differences determine the characteristics of beating and two-stage decay.

A new method presented here[2] is based on the equivalent resonator structure of the digital waveguide presented in Chapter 4. The transfer function of the digital waveguide can be written similarly to Eq. (4.17):

$$\frac{F_{out}}{F_{in}} = \frac{1}{N}\left\{\frac{a_1}{1-z^{-1}p_1} + \ldots + \frac{a_N}{1-z^{-1}p_N}\right\} z^{-M}$$
$$p_k = r_k e^{j\vartheta_k} \tag{5.2}$$

where M is the length of the string in samples shown in Fig. 4.5. N equals to $2M$, a_k are the complex amplitudes of the resonators, ϑ_k refer to the pole frequencies and r_k to the pole radii. Since all complex zeros and poles of Eq. (5.2) form conjugate pairs, the transfer function of Eq. (5.2) can be realized with second-order resonators $R(z)$ connected in parallel. The transfer function of these resonators is as follows:

$$R(z) = \frac{2\Re\{a_{2k}\} - 2\Re\{a_{2k}p_{2k}^*\}z^{-1}}{1 - 2\Re\{p_{2k}\}z^{-1} + p_{2k}p_{2k}^* z^{-2}}$$
$$p_{2k+1} = p_{2k}^* \qquad a_{2k+1} = a_{2k}^* \tag{5.3}$$

where p_{2k} and p_{2k+1} are the complex conjugate pole-pairs, a_{2k} and a_{2k+1} are the corresponding complex amplitude coefficients, $\Re\{\}$ refers to taking the real part of its argument, and * denotes complex conjugation.

It is proposed here that the beating and two-stage decay should be modeled by using one digital waveguide $S_v(z)$ as a basic string model and connecting some resonators $R_1(z) \ldots R_K(z)$ in parallel, instead of using a second digital waveguide. This is illustrated in Fig. 5.12. The excitation signal is common for the digital waveguide and the resonators, and it is computed via nonlinear interaction by the hammer model presented in section 5.2. The efficiency of this structure comes from the fact that only those partials are simulated precisely, where the beating and two-stage decay is prominent. For the others, no resonators are used, and they will have simple exponential decay determined by the basic string model. By using about five or ten resonators, good results can be achieved. In this way the initial amplitudes of those modes which correspond to a resonator will be higher than those simulated by only the digital waveguide. This amplitude difference is not significant, since it is much less than the one introduced by the peaks of the soundboard model.

[2] also published in [Bank et al. 2000]

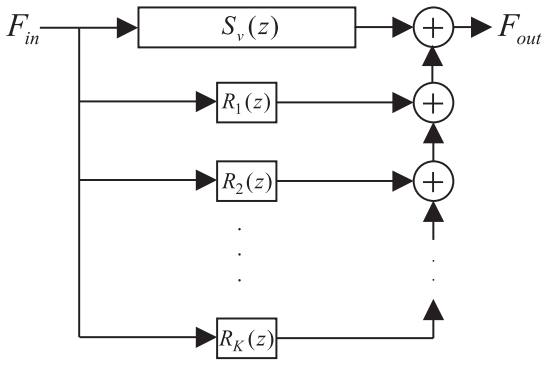

Figure 5.12: The proposed realization of beating and two-stage decay with a digital waveguide and a parallel resonator bank.

Since not all of the resonators of Eq. (5.2) are implemented, the delay z^{-M} cannot be taken into account in the a_k coefficients, on the contrary to the case of Chapter 4. If it was done so, the parallel resonators would sound earlier than the digital waveguide model, resulting in an unnatural attack.

The implementation of the delay z^{-M} in Eq. (5.3) can be still avoided if needed: the output of the digital waveguide should be moved to the force input position M_{in} of Fig. 4.5. In this way the delay of the digital waveguide will be $z^{-M_{in}}$, but there will be some output from the digital waveguide already at the first time instant. Now the resonators and the basic string model start to sound together. Although there are still some differences from the ideal case in the first $2M_{in}$ samples, these were found to be inaudible. However, when setting the initial phases of the resonators, the delay $z^{-M_{in}}$ of the digital waveguide has to be taken into account.

The new structure shown in Fig. 5.12 can be used for generating many different kind of envelopes, such as two-stage decay, constant, temporally increasing or decreasing beating, and some other special effects. Figure 5.13 shows some examples. For that the structure of Fig. 5.12 with four resonators was used. In Fig. 5.13 envelope No. 5 is a pure exponential decay, as it would come from a basic string model, since for that partial no parallel resonator is used. No. 1 is a two stage decay and No. 3 is a beating with decreasing amplitude. The envelope No. 2 and 4 exhibit interesting features: No. 4 is a simple exponential decay but has a hole at about one second, and No. 2 first increases and then decays. In the latter the initial phase difference between the digital waveguide and the resonator was 180 degrees. This can be used for the simulation of the generation of missing modes, often found in other string instruments [Legge and Fletcher 1984].

There are still some open questions concerning this structure which have to be dealt with: these are the re-strike of the string, damping, and the effect of the una corda pedal. If the string is struck again while it is still vibrating, the structure can be still used. The force signal computed by the digital waveguide during the hammer-string contact simply has to be led to the resonators, just as in the case

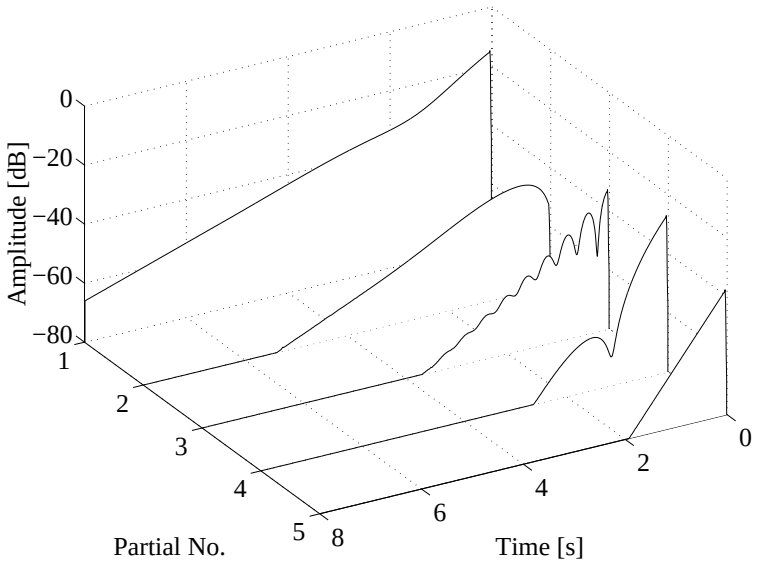

Figure 5.13: Different envelopes generated by the parallel resonator structure of Fig. 5.12

of a normal strike. The effect of una corda pedal can be taken into account by feeding the resonators with the output of the digital waveguide through a small coupling coefficient. The weight of the hammer in the hammer model should be set accordingly, as discussed in Section 6.11.. The implementation of the damper is more difficult: even if a nonlinear damper model is used, which produces a force input to the string, feeding this to the resonators will not lead to their damping, since they differ in phase from the modes of the basic string model. Changing the coefficients of the resonators to shorter decay times would introduce unwanted transients in the sound. A solution can be multiplying the sum of the output of the resonators with an exponential decay. The decay time should be set according to the decay of the basic string model.

To conclude, the benefits of the new structure are that the beating and two-stage decay of the partials can be set separately and it is computationally less expensive than all the methods using two parallel digital waveguides discussed in Section 5.3.2. Computational savings are most advantageous in the case of low piano tones, where the implementation of an another high-order dispersion filter can be avoided, although for the simulation of higher notes less resonators are needed. Estimated computational costs can be found in [Bank et al. 2000].

5.4 Modeling the soundboard

After reviewing the methods used for soundboard simulation, the soundboard model is presented. The model is based on feedback delay networks [Jot and Chaigne

1991]. The principles of feedback delay networks are also outlined. A new method for modeling the attack noise of the piano is proposed.

5.4.1 Earlier work on soundboard modeling

There are several different approaches in the literature, concerning the simulation of the soundboard. The computationally most efficient one is the commuted piano of Smith and Van Duyne [1995]; Van Duyne and Smith [1995]. It is based on commuted synthesis [Karjalainen and Välimäki 1993; Smith 1993], where the soundboard filter is commuted with the digital waveguide and the hammer model. For this, the whole system has to be linear, hence the hammer model as well. The advantage of the method is that the soundboard model does not have to be implemented as a high order filter, but as a wavetable, whose content is simply played to the string. No doubt, this is the computationally most efficient soundboard modeling technique so far. However, it is a serious drawback that it cannot be used with a nonlinear hammer model (see Section 5.2.1).

The other end of the scale is the finite difference soundboard model presented in [Giordano 1997]. It is physically meaningful and allows the simulation of different features of the soundboard, such as the effect of ribs. However, the solution of the two dimensional finite difference equation takes a long computational time. Therefore, the finite difference technique can be considered an analytical tool, and not a soundboard model candidate for sound synthesis.

A simple and efficient piano soundboard model was presented in [Garnett 1987]. The strings were modeled by digital waveguides, connected to the same termination. The soundboard was simulated by connecting six additional digital waveguides to this common termination. Each of these interconnected waveguides contained a lowpass filter with a large damping factor. An advantage of the technique is that it can provide high modal density for low computational cost. The disadvantage is that it simulates a general soundboard, modeling of the soundboard of a specific piano is quite difficult with this method.

A similar approach, suggested, e.g., in [Rocchesso and Smith 1997], is based on feedback delay networks. Although this approach has not been used for modeling the soundboard of the piano, but rather for room response simulation. The principles of the feedback delay networks are outlined in the next section.

5.4.2 Feedback delay networks

Feedback delay networks can be considered the generalization of the state space representation of a discrete-time linear system [Rocchesso and Smith 1997]. Here, the elementary unit delays z^{-1} are replaced with delay lines of arbitrary length z^{-m_k}. The equations in the z domain are:

$$\begin{aligned} \mathbf{s}(z) &= \mathbf{D}(z)\left(\mathbf{A}\mathbf{s}(z) + \mathbf{b}x(z)\right) \\ y(z) &= \mathbf{c}^T\mathbf{s}(z) + dx(z) \end{aligned} \tag{5.4}$$

where $x(z)$ and $y(z)$ are the input and output signals, $\mathbf{s}(z)$ is the state vector, $\mathbf{D}(z)$ is the diagonal delay matrix $D_{kk} = z^{-m_k}$, and $\mathbf{A}$ is the feedback matrix. In Eq. (5.4) $\mathbf{b}$ is the input coefficient vector, $\mathbf{c}$ is the output coefficient vector, and d is the output coefficient.

It can be shown that the technique used in [Garnett 1987], i.e., connecting digital waveguides to the same termination, is a special case of feedback delay networks [Rocchesso and Smith 1997].

An important issue is the stability of the system. The feedback delay network is stable when the corresponding state-space representation with $z^{-1} = z^{-m_k}$ for each delay element is stable [Rocchesso and Smith 1997]. It was proven in [Rocchesso and Smith 1997] that as long as $\|\mathbf{A}\|^n$ decays exponentially with n, the stability is assured. A straightforward approach for controlling the properties of the system is making a lossless prototype and then introducing losses to the system [Jot and Chaigne 1991; Rocchesso and Smith 1997]. For such a prototype, we need that all the poles lie on the unit circle. This can be assured when the feedback matrix $\mathbf{A}$ is unitary [see, e.g., Jot and Chaigne 1991; Rocchesso and Smith 1997]. Since we generally deal with systems with real coefficients, the equivalent constraint is the orthogonality ($\mathbf{A}^{-1} = \mathbf{A}^T$) of the feedback matrix.

When the size of $\mathbf{A}$ is sufficiently large (about 12 rows and columns) and the total delay length is at least 1 second, the impulse response of the lossless prototype can be compared to a pseudo-random noise generator [Jot 1992].

A new class of feedback delay networks was presented in [Rocchesso and Smith 1997]. By choosing the feedback matrix to be circular, the distribution of eigenvalues is easily controlled. The eigenvalues of the system and the first row of the circular feedback matrix $\mathbf{A}$ are related through the Discrete Fourier Transform. Losslessness can be achieved by choosing eigenvalues of unit length, and calculating the feedback matrix by means of Inverse Discrete Fourier Transform.

Losses are introduced to the system by connecting filters $H_k(z)$ to the output of the delay lines [Jot and Chaigne 1991; Rocchesso and Smith 1997], that is, substituting z^{-m_k} with $z^{-m_k}H_k(z)$ in the delay matrix $\mathbf{D}$. Then, to maintain flat magnitude response, a tone corrector filter has to be used in series to the feedback delay network [Jot and Chaigne 1991; Jot 1992].

5.4.3 The soundboard model

Feedback delay networks have been successfully used for room reverberation simulation. This is because the late part of the impulse response of rooms can be considered as a nonstationary Gaussian process with a time-variant power spectral density. By properly choosing the feedback matrix $\mathbf{A}$ and the loss filters $H_k(z)$, similar output arises from the feedback delay network [Jot and Chaigne 1991]. The underlying metaphor and the connections of the response of a rectangular room and the feedback delay network are discussed in [Rocchesso 1995].

The use of feedback delay networks for simulating the body of an instrument was suggested in [Rocchesso and Smith 1997]. This is motivated by the fact that

the exact position and height of resonances are perceptually insignificant. As long as the modal density and the decay times are close to the original one, the sound will be similar. In reality, when listening to the impulse response of the soundboard, it sounds like a colored, decaying noise.

However, for soundboard modeling other considerations are needed than those for the room response simulation. The density of the pulses in the output of feedback delay network is rather small in the beginning of the impulse response. This well coincides with the behavior of real rooms, where the late reverberation slowly builds up. These early pulses in the response of the feedback delay network can be considered as the early reflections in the room response. The colorless reverberation can be ensured by using delay lines of comparable length [Rocchesso and Smith 1997].

On the contrary, the impulse response of the piano soundboard does not show this feature. It is rather a sum of damping modes, the rise in the attack cannot be observed. If the soundboard is simulated with a delay line lengths of more than two or three hundred samples, the resulted sound will be rather cello-like, since the sharp attack of the piano sound disappears. Therefore, we need energy in the early region of the impulse response as well. This can be done by applying short delay lines together with the long ones. The author has successfully used an eight delay line feedback delay network with gradually increasing delay lengths (37, 87, 181, 271, 359, 492, 687, 721) for soundboard simulation. These are relative primes, in order to avoid harmonic structure. For the losses, one-pole filters were used in series to the delay lines. The parameters of the model are given in Section 6.5.

The feedback delay network model was found to be a good approximation of the soundboard impulse response. The main benefit of such a structure against traditional FIR or IIR filters comes from the efficient implementation of the delay lines. This way, very high modal density can be obtained with relatively small number of operations. The drawback is that the resonances cannot be controlled separately. Consequently, this method is not capable to reproduce the exact copy of a real soundboard impulse response, but it is good enough to capture its main features.

From this point of view, its relation to general IIR filters is similar to that of the digital waveguide to the resonator bank. In IIR filters the computational complexity depends on the number of realized peaks in the transfer function. On the contrary, in feedback delay networks it comes from the accuracy of the approximation. In other words, the feedback delay network is capable to simulate many modes for low computational cost, but their place and height cannot be precisely controlled.

If the soundboard is simulated by the feedback delay network only, more natural result occurs, when the tone corrector filter of [Jot and Chaigne 1991; Jot 1992] is not used. A better approximation has been obtained by whitening the impulse response with the tone corrector [Jot and Chaigne 1991; Jot 1992], and then filtering with a shaping filter whose magnitude response is similar to that of the real soundboard. The author has used a 100 tap FIR filter with good results. The advantage of the FIR filter that it can be easily designed by means of Inverse Discrete Fourier

Transform. Possibly, warped filter design could give better results [Karjalainen and Smith 1996].

There can be different levels in modeling the soundboard impulse response. The most efficient one is by using a feedback delay network without a shaping filter. The second step is using the feedback delay network in series with a shaping filter, to match the power frequency response of the measured soundboard.

However, the impulse response of the soundboard differs largely when different points of the bridge are excited. This can be simulated by separate shaping filters for different regions of the soundboard. Note that there is no need to implement separate feedback delay networks for the different parts. This comes from the fact that when the bridge is hit at different positions, the same modes of the soundboard are excited, they only differ in level, and not in their frequency or decay time. This way, the string signals are filtered through two or three shaping filters, corresponding to the regions of the soundboard. Then, the output of the shaping filters are summed and sent to the feedback delay network.

A more efficient implementation is proposed here, illustrated in Fig. 5.14. It is based on using only one shaping filter $H_{sh}(z)$ (e.g., the one designed for the midrange) connected in series with the feedback delay network $H_{FDN}(z)$. To simulate the differences in the low and high range, low order IIR filters $H_{low}(z)$ and $H_{high}(z)$ are used.

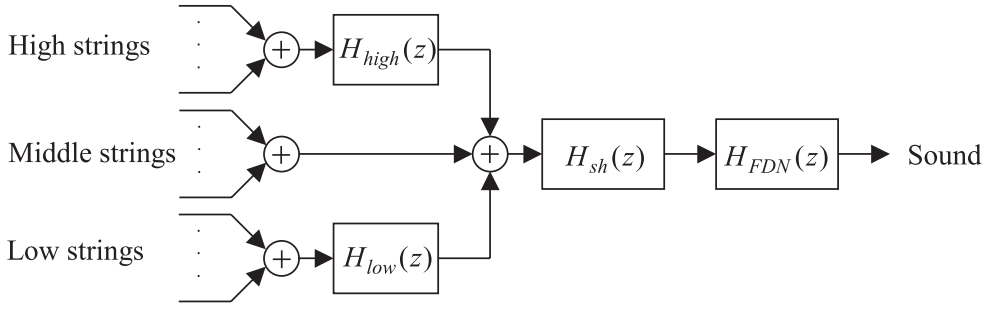

Figure 5.14: Proposed method for soundboard modeling.

For real-time applications, the complicated shaping filter $H_{sh}(z)$ of Fig. 5.14 can be omitted. Instead, second or third order IIR shaping filters are applied for all the different regions of the soundboard.

5.4.4 Efficient method for modeling the attack noise

The model of Fig. 5.14 gives good results for the low and the midrange of the piano, but the attack noise of the high notes is not reproduced. This is because the high frequency signal coming from the high strings cannot excite the low modes of the soundboard model. Nevertheless, in real pianos, a knock can be heard when listening

to high tones. This attack noise has been found to be a distinctive character of the sound of pianos, hence its simulation is necessary.

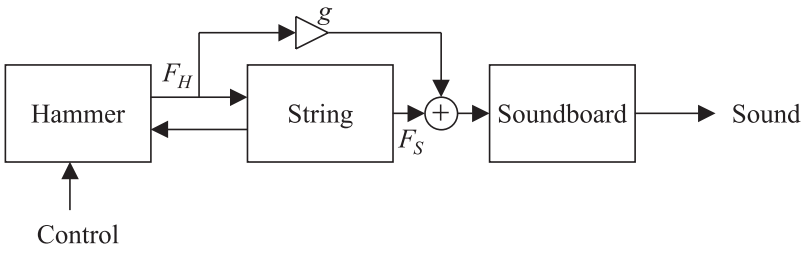

Figure 5.15: New method for simulating the attack noise of the piano.

A simple model is proposed here to simulate the attack noise of high piano tones. The idea is displayed in Fig. 5.15. Now the interaction force calculated by the hammer model F_H is led to the input of the soundboard model through a real coefficient g (compare to the earlier methods in Fig. 5.1). The value $g = 0.2$ was found to be good in practice. The force input of the soundboard is the weighted sum of the force output of the string and the hammer, i.e., $F_S + gF_H$. This way, the force signal can efficiently excite the low modes of the soundboard as well. It also automatically ensures the different coloration of the attack noise for different keys. This is because for low notes the impulse of the interaction force is longer than for the high ones. Consequently, the force signals of low notes introduce more lowpass filtering to the impulse response of the soundboard model than that of the high ones. Another advantage is that the amplitude and the tone of the attack noise varies with impact velocity. Higher impact velocities correspond to a force signal with higher amplitude and stronger high frequency content, and therefore result in a louder and brighter attack noise. The physical meaning of the new method is that in real pianos the interaction force experienced by the hammer can excite the soundboard through the action and the case [Chaigne and Askenfelt 1994b]. The earlier piano model of the present author was described to be too clean in the high range by many experts of the field. This simple technique partly solved the problem and made a large improvement in the perceptual quality of the model. What is still missing is the simulation of the coupled vibration of the other undamped high strings.

5.5 Real-time implementation issues

Here we discuss the practical aspects related to efficient real-time implementation. The model described previously exists in the form of MATLAB code. The author is planning to implement the model in a real-time environment, since that could really show the advantages and drawbacks of the proposed methods. Experimenting with some parameters would also be more convenient that way. Therefore, the techniques

discussed here are the sketches of the planned realization.

First, the idea of the multi-rate piano model is presented. Then, the considerations of real-time implementation are outlined for the hammer, string, and soundboard model separately. The last part of the section roughly discusses the effect of quantization on the separate parts of the model.

5.5.1 The multi-rate piano

The computationally most expensive part of the model is the dispersion filter needed for the low piano strings. However, the output of these strings does not contain significant frequency components above 10 kHz. Therefore, it is worth to operate the lowest three octaves of the piano at half sampling rate. This will reduce the computational load in two ways: one is that there has to be only one sample computed instead of two during the same amount of time. The other is that the specification for the dispersion filter will be easier to fulfill, since the total length of the digital waveguide will diminish by a factor of two. This will result in a lower order dispersion filter. For notes higher than C_4, informal listening tests show that the implementation of dispersion is not necessary, especially, if the beating is simulated. Accordingly, the string models for notes in the middle and high range operate at the output sampling rate (e.g., $f_s = 44.1$ kHz), but the dispersion filters are omitted. The sampling rate f_s refers to the sampling rate of the output device, i.e., the digital-to-analog converter. This way, the computational load of the high and low notes can be made approximately equal. This is shown in Fig. 5.16. The inputs from the strings contain the force signal coming from the digital waveguide, the resonators, and the hammer model. These parts are not displayed.

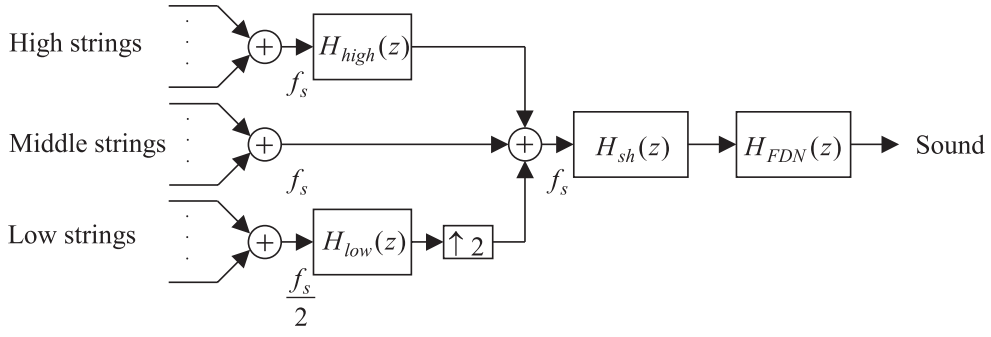

Figure 5.16: New method for simulating the attack noise of the piano.

Note that the hammer models always run at a double sampling rate compared to the corresponding string model. Hence, for the high notes, the sampling rate of the digital string model, that is, the digital waveguide and the resonators, is f_s. For the hammers it equals to $2f_s$. In the case of the low notes, the sampling frequency of the string model is $f_s/2$ and f_s for the hammer model. As a next step, the lowest

octave of the piano can be operated at the fourth of the normal sampling rate $f_s/4$, but that could lead to some degradation in sound quality.

The shaping filters are operated at the same sampling rate as the strings connected to them. Then, the signal of the low register with sampling rate $f_s/2$ is upsampled. Probably a 10^{th} order FIR filter will be appropriate for this purpose. After upsampling, the output signal of the low strings is added to that of the high ones and fed to the shaping filter and the feedback delay network.

5.5.2 Hammer and dampers

The implementation of the hammer model is quite straightforward. This is because it can be considered a linear filter connected in series with an instantaneous nonlinearity.

The only problem here is the realization of the $F(\Delta y) = K(\Delta y)^p$ function (see Fig. 5.5). One way is to use a lookup table with the values of $(\Delta y)^p$ for Δy values that normally occur and for about ten p values between 2 and 3. The not stored Δy values can be determined by means of linear interpolation.

A simpler solution, also proposed in [Borin and De Poli 1996], can be to approximate $K(\Delta y)^p$ with $K_2(\Delta y)^2 + K_3(\Delta y)^3$. Alternatively, when determining the parameters of the hammer model, this polynomial form could be used for optimization (see Section 6.1.1). An advantage of this approach is that no lookup-table is needed. The interpolation between different felt characteristics is straightforward, because the linear interpolation of K_2 and K_3 parameters give meaningful results. Note that in the case of a force model $K(\Delta y)$ it is not obvious how to interpolate between force characteristics of different p values.

The fact, that the hammer runs at the double sampling frequency, does not lead to any complication in the implementation. From outside, the double sampling rate is not visible. It only means that the same kind of operations are executed twice in the program code.

The dampers can be realized as one-pole filters connected to the digital waveguide in series. The implementation of a nonlinear damper needs similar considerations as the hammer model. The problem is the lack of data about real damper behavior. However, a real-time implementation could be useful in this sense, since the small number of parameters of the damper model could be tuned just by listening to the result.

5.5.3 String model

The efficiency of the digital waveguide lies in the delay lines: they can be implemented as circular buffers. By commuting the -1 signs in Fig. 4.5, the figure can be redrawn with two delay lines, without any sign change in the feedback loop [Smith and Van Duyne 1995]. This is similar to the case when the variables of the waveguide are force waves. The difference is up to a constant factor Z_0. From a practical point of view it is worth to use force signals in the delay lines instead of the velocity.

This is because both the input and the output of the digital waveguide are forces. By doing so, the connections of the hammer model and the digital waveguide have to be changed.

Now that there is no sign change in the delay lines, the whole digital waveguide can be implemented as one circular buffer with four pointers. Four pointers are needed to compute the input of the hammer (see Eq. 5.1) and two of them are used for inserting the force input to the digital waveguide. The output of the digital waveguide is taken from the position $y^+(n, M_{in})$. Then, this cell is overwritten by the filtered version of the output signal. This realizes the reflection filter $H_{r_v}(z)$. After that, all the four pointers are shifted by one step.

The implementation of the resonator bank is rather straightforward, since it consists of parallel second order sections.

5.5.4 The soundboard

The soundboard is implemented as a feedback delay network with correction and shaping filters (see Fig. 5.14).

The delay lines of the feedback delay network are implemented as simple circular buffers. The most time-consuming operation of the feedback delay network is the matrix multiplication (see Eq. 5.4). Nevertheless, by the careful choice of the feedback matrix $\mathbf{A}$, the matrix multiplication can be eliminated [Rocchesso and Smith 1997]. For example, in the case of the circular feedback matrix presented in Section 6.5.1, all the output signals of the delay lines after the loss filter are summed, multiplied by $-1/4$ and fed into the input of all the delay lines. This corresponds to the $-1/4$ elements of the feedback matrix $\mathbf{A}$. The consecutive delay lines are also connected with gain factors of 1. The latter implements the elements $1 - 1/4$ in the feedback matrix.

5.5.5 Finite wordlength effects

For practical implementation, it is interesting to examine how sensitive the different parts of the system are for roundoff errors. The problem with the hammer model lies in the digital integrators, since there a very small value is added to the past output. Some of the problems can be avoided by rescaling the signal levels in Fig. 5.5, i.e., using nonphysical variables. However, their realization seems to be problematic in 16-bit arithmetic.

The reflection filter of the digital waveguide is sensitive to quantization. Its magnitude response is close to 1, which means very small coefficients. Also the dispersion filter can have large difference in the values of its coefficients, if implemented in direct form. Therefore, to avoid instabilities, the reflection filter should be split into second order sections. If the quantization of the architecture is very rude, even the use of an FIR loss filter should be considered. The second order resonators of the resonator bank can also face with stability problems, since the implemented quality

factor values are very large. Interpolated low frequency resonators could overcome this problem [Välimäki and Tolonen 1998].

On the contrary, the soundboard model is quite robust against quantization errors. This is because there the magnitude of the loss filters is not as close to unity as that of the string model. The shaping filter is an FIR filter, which is stable by principle.

To sum up, it seems that a 16-bit fix point arithmetic is not enough for the piano model. At least 24-bit fix point arithmetic should be used. The best solution would be the use of a floating point architecture. Then the physical units of the variables could be directly used, i.e., they would not need any rescaling.

5.6 Conclusion

We discussed the structure of the piano model. After outlining the general structure of the piano, the hammer and damper models were presented. The discontinuity problem of the string was analyzed and methods were suggested for taking in the interaction force to the digital waveguide in a correct way. A new multi-rate hammer model was presented, which is stable even for extremely high impact velocities. The advantage of this method compared to those of the literature lies in its simplicity. The same structure can also be used for the simulation of the effect of dampers, albeit the calibration of such a damper model has not been solved yet. A novel resonator bank based structure was presented for the simulation of beating and two-stage decay. Although it simulates the harmonics separately, it responds to the variation of the interaction force automatically. The soundboard is simulated by a feedback delay network, a technique which has been found very successful in room response modeling. It can be viewed as a statistical model of the impulse response, since only the general properties of the soundboard are matched. A simple technique for modeling the attack sound of the piano was presented. The last part of the chapter dealt with practical implementation issues. There, a multi-rate piano model was presented, which balances the computational load of the high and low strings of the model.

Chapter 6

Model calibration

In the previous chapter the structure of the piano model was presented. Here we discuss the calibration of the different parts of the model. This is the inverse operation of the simulation, since in simulation the inputs are the model parameters and the output is the audible sound. Here the case is the contrary: we wish to find parameters for our model which give the most similar sound output to the sound of the real piano.

Theoretically, the easiest approach would be a nonlinear optimization process for all the parameters of the model. It is the easiest from the point of view of the scientist, and not of the computer, of course. This optimization could be done by applying genetic algorithms or neural networks for searching in the parameter space and maximizing a fitness function between the original and the synthesized signal. However, this would lead to instabilities in the parameter estimation, not to mention the need for a lot of computational power. Finding an appropriate fitness function could also be problematic.

The basic idea of this chapter is to keep the parameter estimation as simple as possible, with respect to the necessary computations. The building blocks of the procedures should be well known and easily realizable operations. The calibration process should be straightforward, without any feedback from the output of the model, to avoid any instabilities. This kind of approach needs more analytical work, since the behavior of the model structure should be well understood.

Two new filter design algorithms are presented for the digital waveguide, one for the general case and another for the one-pole loop filter. The idea of both methods is to minimize the error of the resulting decay times, and not the magnitude response of the filter. An advantage of these techniques is that the stability of the filters can be easily assured. In Section 5.3.3, a new technique was proposed for the simulation of beating and two-stage decay. A novel method had to be developed for the calibration of this structure, which was found to be robust and reliable for all kind of piano sounds.

6.1 Parameters of hammer and damper models

The parameters for the hammer model were taken from the literature, because of the lack of time. Nevertheless, a simple method for calibrating the hammer is proposed. The calibration of the linear damper model is similar to that of the digital waveguide, which is discussed in Section 6.2.

6.1.1 Calibration of the hammer

For the implemented piano model, the parameters of [Chaigne and Askenfelt 1994a] were used. The parameters of the three hammers measured in [Chaigne and Askenfelt 1994a] were interpolated for other notes. The hammer mass parameter of the model was set in a way that the ratio of hammer and string mass was the same as for the real piano. This means that the mass of a hammer which hits a triplet, has to be divided by three, since in the model it is connected to one digital waveguide only. For the same hammer, if the una corda pedal is depressed, the hammer mass in the model will be the half of the real hammer mass.

However, there were only three hammers measured in [Chaigne and Askenfelt 1994a], which obviously does not cover the entire keyboard. It would be also beneficial to make hammer measurements on old pianofortes, since their simulation could be even more interesting than the modern ones.

A method proposed here for estimation of hammer parameters is based on the work of [Boutillon 1988]. A miniature accelerometer is attached to the measured hammer. Then, the key is hit several times at different dynamic levels. The acceleration signal is recorded for analysis. The impact velocities of the hammer can be computed by integrating the acceleration signal. Then, a digital waveguide model with the proposed hammer model is run, and the acceleration of the hammer is computed. By using standard optimization algorithms, the difference of the two signals can be minimized with respect to the hammer parameters. The initial guesses for the optimization can be obtained from analytic expressions of the hammer behavior (see Section 2.1.2 for references). The hammer mass is computed by subtracting the mass of the accelerometer from the outcome of the optimization. An advantage of the method is that the piano action does not have to be disassembled, and it does not make any harm in the hammer felt.

6.1.2 Calibration of the dampers

The calibration of the damper is similar to that of [Erkut et al. 2000], and it is based on recorded piano tones. The analyzed signal can be sound pressure or the acceleration of the bridge. The procedure is as follows: first, the beginning of the damping has to be determined. This was done manually, but it could be made automatic by looking for a jump in the time derivate of the amplitude envelope. This time instant will be the beginning of the analysis signal. The end is where the amplitude of the partials falls under the noise level. After determining the

beginning and the end of the analysis signal, the decay times of the partials are estimated. From the estimated decay times a one-pole loop filter is designed. The details of the analysis and filter design procedures will be presented in Section 6.2.

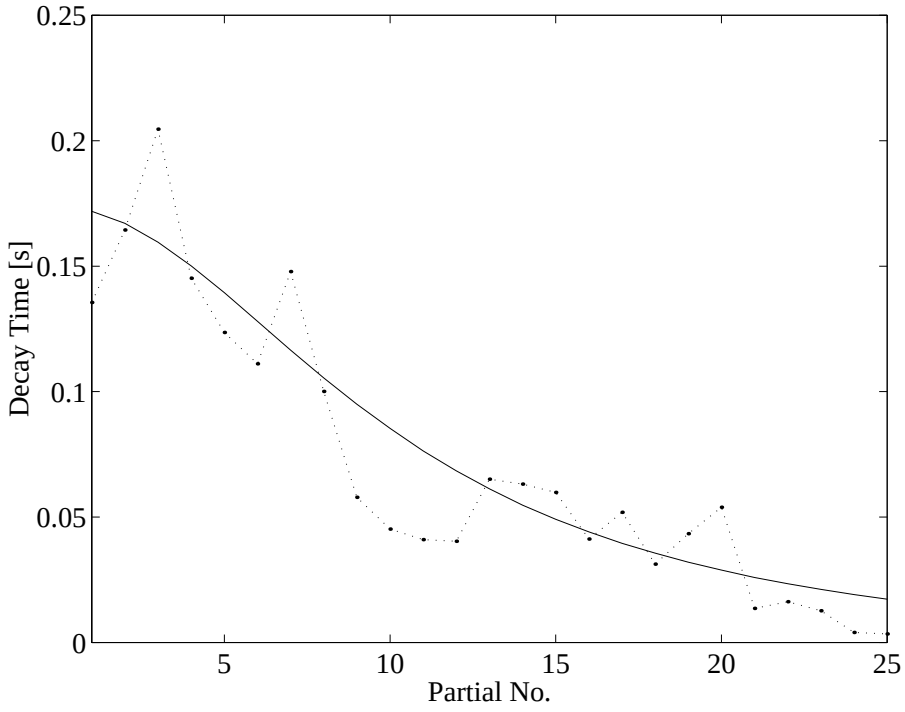

Figure 6.1: Prescribed (dotted line with black points), and approximated decay times (solid line) for the linear damper model, note $A_4^\sharp$ (466 Hz).

Fig. 6.1 shows the decay times of damping of note $A_4^\sharp$ and the approximation using a one-pole filter. It can be seen that the filter follows the general trend well, but the important feature of real dampers, that every 7^{th} partial is damped inefficiently, is lost. This is even a bigger problem for the low strings, were the difference is more audible, because of the long damping times. A nonlinear hammer model would overcome this problem.

The calibration of the nonlinear damper model could be made similarly to what is proposed for the hammer, although it is more complicated. To obtain good results, not only the acceleration of the damper should be measured, but the velocity of the string as well. This could be done by optical, magnetic, or electrostatic sensors. Unfortunately, this kind of measurements are more problematic than measuring with an accelerometer. Therefore, finding robust methods for measuring the damper behavior are of interest. This will be also a part of future research.

6.2 Parameter estimation for the string model

Since the model used for simulating the string behavior consists of two different parts, namely, the digital waveguide and the resonator bank, the calibration of these needs different considerations. Accordingly, the description of the string calibration is split to three sections. The first two discuss the calibration of the digital waveguide, namely, signal estimation and filter design, and the third concentrates on the calibration of the resonator bank.

The calibration of the digital waveguide consists of designing the reflection filter $H_{r_v}(z)$ and setting the length of the digital waveguide M in Fig. 4.5. One approach could be the impedance measurement of the bridge and the analytic derivation of string losses and vibration. This could give an insight to the physical behavior of the string-soundboard system. The separation of the effects of the string and the termination could be helpful in determining the parameters of a model displayed on Fig. 5.10 or 5.11. Smith [1983] computed the specification for the reflection filter $H_{r_v}(z)$ by means of linear prediction and by a system identification method. However, these did not give a robust solution. A more straightforward approach taken here, also suggested by Välimäki et al. [1996], is based on the measurement of partial envelopes. The loss and dispersion filters can be designed from the frequencies and the decay times of the partials. This way, the effects introduced by the string and the termination cannot be separated, but it is not needed for a single-string digital waveguide model. The interest is now only on designing a digital waveguide whose behavior is similar to the measured string.

6.2.1 Overview of signal estimation methods

For the analysis of the partials, basically two types of methods are used in the literature. One is based on the Short Time Fourier Transform [Allen and Rabiner 1977]. The signal is first multiplied by a window function with a length corresponding to integer number multiples of the pitch period. Then the DFT of the windowed data is computed. As this is done for several time instants, the result is the temporal evolution of the short time spectrum. By finding and following the peaks of these spectra, the frequencies and the decay times of the partials can be calculated. The method has been described in [Välimäki et al. 1996].

The other approach, also taken here, uses heterodyne filtering [Moorer 1977]. The observed partial is brought to around zero in the frequency domain by multiplying the signal with a complex exponential corresponding to that frequency. The frequency components of other partials are suppressed by lowpass filtering. The amplitude of the partial is calculated by taking the absolute value of the lowpass filtered signal [Välimäki and Tolonen 1998; Ding and Qian 1997]. The signal is examined only at the partial frequencies, and this speeds up the analysis procedure compared to the DFT based approach.

6.2.2 Heterodyne filtering for signal estimation

Both the sound pressure and bridge acceleration of several piano tones were recorded for analysis in this study. The bridge acceleration signals were found to be more useful than the sound pressure, since they contained stronger high harmonics. This is because the pressure signal is rather proportional to the bridge velocity, i.e., its spectrum falls at a rate of 20 dB/decade compared to acceleration. The frequencies and decay times of the partials are the same for the pressure and acceleration signals. Their amplitudes differ but the initial amplitudes are irrelevant in the filter design.

The heterodyne technique was chosen for the analysis of the measured signal, mostly because of its computational simplicity. The steps of the algorithm are illustrated in Fig. 6.2. The fundamental frequency of the tone can be calculated by some pitch estimation methods, for example based on the autocorrelation function [Rabiner 1977; Välimäki et al. 1996], or by finding peaks in the DFT of the signal. This is not needed here, because the fundamental frequency of the analyzed note is approximately known, since it comes from which key of the piano was depressed.

The analyzed signal $x(n)$ is multiplied by a complex exponential $e^{-j\hat{\vartheta}_1 n}$ corresponding to the fundamental angular frequency estimate $\hat{\vartheta}_1$. Then the resulted complex signal is lowpass filtered. There is a tradeoff between the time domain and frequency domain resolution of the analysis: if the lowpass filter H_{lpf} has a low cut-off frequency, the envelopes (especially in the attack part) will be somewhat smeared, but it will suppress the measurement noise appropriately. On the contrary, high cut-off frequency leads to more noise, making the analysis procedure uncertain. The cutoff frequency of the filter obviously has to be below the fundamental frequency of the tone. Here, a fourth-order Butterworth filter was used, designed by the MATLAB `butter` command. To avoid phase distortion, the data was filtered in both forward and backward directions, by using the MATLAB `filtfilt` procedure [MAT 1996]. In this way, the resulted filter had an order of eight. The cutoff frequency was set to half of the fundamental for the very low notes and to one fourth for the others.

The amplitude envelope of the partial is computed by simply taking the absolute value of the filtered complex signal. Note that this does not give the amplitude of the sinusoidal representation, but the amplitude of the complex exponential. To obtain the amplitude of the corresponding sinusoid, the value should be multiplied by two. However, we are not interested here in the amplitudes, but the decay times. But before computing those, the time-span of valid data should be found. For the beginning of the note, a good estimate is the time instant with maximal amplitude, since the piano sound is decaying by nature. Finding the end is somewhat more complicated. If the measured data is long enough, its very last part consists of measurement noise only, whose level is then computed. The note is assumed to be terminated when its level falls below a certain limit, which was set to 10 dB above the measurement noise. The decay times τ_k are estimated by taking the logarithm of the amplitude envelope and applying linear regression to that, similarly to [Välimäki et al. 1996]. The slope of the fitted line determines the decay time of the partial.

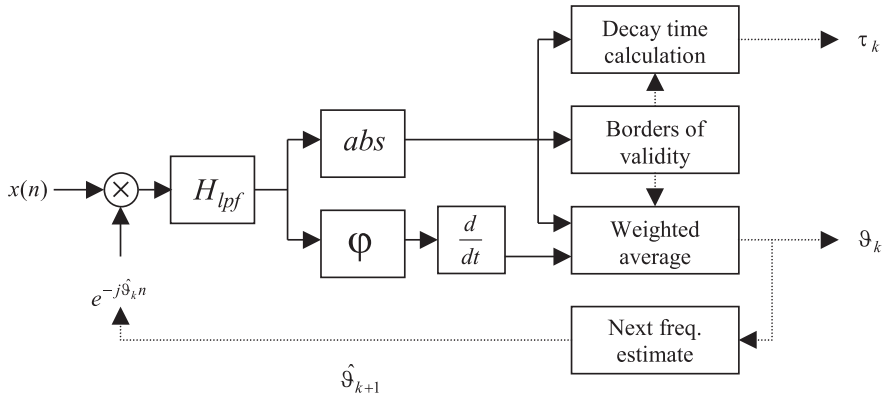

Figure 6.2: Analysis procedure for partial envelopes.

The calculation of frequency deviation from the frequency estimate is based on calculating and unwrapping the phase of the complex lowpass filtered signal. By differentiating the phase with respect to time, the instantaneous frequency deviation can be determined. The frequency ϑ_1 of the first partial is then computed by adding up the frequency estimate $\hat{\vartheta}_1$ and the weighted average of the instantaneous frequency deviation. The weighting is based on the amplitude envelope of the partial and gives less significance to the weaker, and therefore more noisy parts of the input data.

In the case of most string instruments, the tone can be considered harmonic. Accordingly, the frequency estimates for the partials can be simply the multiples of the fundamental frequency ($\hat{\vartheta}_k = k\vartheta_1$) [see Välimäki et al. 1996]. On the contrary, the inharmonicity of the piano tone calls for a more refined method based on the estimation of the inharmonicity coefficient B. Once the fundamental partial of the tone is analyzed, the second frequency estimate will be twice of this frequency ($\hat{\vartheta}_2 = 2\vartheta_1$). Then the analysis procedure is run again, giving the exact frequency ϑ_2 of the second partial. The frequency estimates for the next partials are computed by fitting an inharmonicity curve (see Eq. 2.3) on the already measured partial frequencies $\vartheta_1 \ldots \vartheta_k$ and calculating the next frequency estimate $\hat{\vartheta}_{k+1}$ from that expression.

6.3 Filter design

The output of the analysis procedure is a set of partial frequencies and decay times, which can be used for designing the reflection filter of the digital waveguide ($H_{r_v}(z)$ in Fig. 4.5). This filter has to contain the effect of the losses and dispersion. Its other task is the fine-tuning of the length of the digital waveguide. Since the digital waveguide contains an integer number of delays, varying only that would cause large quantization errors in the fundamental frequency, similarly to the case of the basic

Karplus-Strong algorithm [Karplus and Strong 1983; Jaffe and Smith 1983].

The analysis data results in a complex specification, which could be directly used for filter design. For this, one should be able to separate the phase and the amplitude error of the approximation, since for those different constraints are needed. The phase response of the reflection filter has to be very accurate at the fundamental frequency of the note. On the contrary, at other frequencies it is enough if it follows the general trend of the prescription. If the shape of the phase delay is similar to what is needed to simulate the inharmonicity, the sound output will be still piano-like, even if there are some inaccuracies. The decay times of the partials depend on the amplitude response of the filter, as shown in Eq. (4.16). It follows from the equation, that the closer the amplitude response is to 1, the larger error will arise in decay times for the same amplitude difference. The amplitude response should never be larger than 1, since it would make the feedback loop unstable. Consequently, a very complicated filter design algorithm would be needed, considering the different kind of approximation errors, in order to give good results.

Smith [1983] reviews a number of sophisticated filter design techniques, two of them are also discussed in [Laroche and Jot 1992]. However, these methods suffer from instabilities. The papers rather suggest the division of the filter design procedure to the loss filter, dispersion filter and fractional delay filter parts [Jaffe and Smith 1983; Välimäki et al. 1996]. By using allpass filters for simulating the dispersion and controlling the fundamental frequency of the note, the amplitude response of the reflection filter will depend only on the loss filter, simplifying the filter design procedure largely. The slight phase difference caused by the loss filter can be neglected when compared to the phase response of the dispersion filter.

Note that dividing the filter design into different steps cannot give mathematically optimal results, since the separate parts of the filters have some constrains on the filter coefficients, like being allpass for the dispersion filter. Now the simplicity of the analysis will lead to computationally less efficient implementation. It follows that in the future it is of great interest to find a robust algorithm which could design one complete reflection filter based on the previously mentioned amplitude and phase criteria.

6.3.1 Review of loss filter design algorithms

The specification for the loss filter can be computed by using the inverse of Eq. (4.16):

$$g_k = e^{-\frac{1}{f_0 \tau_k}} \tag{6.1}$$

where τ_k is the decay time of partial k, f_0 is the fundamental frequency of the note and g_k is the desired amplitude value of the loss filter at the angular frequency ϑ_k of partial k. Fitting a filter to g_k coefficients is not trivial, even if the phase part of the transfer function is not considered. This is because of the previously mentioned nature of the loop filter: the error in the decay time is a nonlinear function of the amplitude error. The stability of the digital waveguide loop is also hard to

handle. Unfortunately, there exists quite a small number of papers in the literature concerning this subject. The methods reviewed in [Smith 1983] were found to be too complicated for practical implementation. The stability constraint of the digital waveguide $|H_{r_v}(z)| < 1$ cannot be incorporated in these methods either.

In [Välimäki et al. 1996; Välimäki and Tolonen 1998] the problem was solved by using a one-pole filter. The transfer function of such a filter is:

$$H_{1p}(z) = g\frac{1 + a_1}{1 + a_1 z^{-1}} \tag{6.2}$$

where $-a_1$ is the pole of the filter and g refers to the DC gain. In [Välimäki et al. 1996; Välimäki and Tolonen 1998] such a filter was found to be adequate for simulating the acoustic guitar and other plucked string instruments. The DC gain g was set according to the decay time of the first partial ($g = g_1$), or the average of the first two or three g_k coefficients. The pole of the filter was determined by continuously adjusting a_1 and searching for the minimum of the analytical expression of the approximation error. The error was computed in a least squares sense, by using a weighting function putting more emphasis on slowly decaying partials. Jaffe and Smith [1983] also discussed the use of the one-pole filter, but without the gain factor, i.e., $g = 1$ in Eq. (6.2).

The advantage of using a one-pole filter is that it is always of a lowpass character for $a_1 < 0$. Since g is less than unity by the principle of calculating it, the waveguide loop will be always stable. As we discuss in the Appendix, such a simple constraint can be found for the first-order FIR filter as well. The problem of the technique of [Välimäki et al. 1996] comes from the separation of the g and a_1 parameters, since first g is set and then a_1 is optimized assuming g to be given. Nonlinear optimization can overcome this problem, as presented in [Erkut et al. 2000]. There the previously mentioned procedure was used calculating the initial estimates for g and a_1. Then a nonlinear optimization was carried out based on the amplitude envelope of the synthesized and the original signal. This is motivated by the fact that the overall decay time of the synthesized sound has a large perceptual significance. On the other hand, the timbral evolution of the sound was not taken into account in the nonlinear optimization loop, still not leading to an optimal solution. However, it is a question, how precise the approximation must be. Listening tests show that the threshold for recognizing differences in the overall decay time is rather high [Tolonen and Järveläinen 2000].

6.3.2 The transformation method: a novel loss filter design technique

The fact that the loss filter has only one pole is also a hard constraint itself, since it restricts the sets of realizable decay times to a great extent. As we shall see, in the case of the piano the dispersion filter is computationally much more expensive than the loss filter. Hence, applying e.g., a fourth-order loss filter instead of a first-order one would not increase the computational costs significantly. A straightforward

solution for designing such a filter is to use standard filter design methods for g_k with a weighting function, e.g., such as presented in [Välimäki et al. 1996]. Unfortunately this does not lead to useful results, since in most of the cases the amplitude response of the resulted filter goes above unity and the decay times calculated from the designed filter are nothing similar to those of the specification. This is because the dependence of the decay times on the filter magnitude is highly nonlinear. Note that weighting would give a solution to that only if this relation could be assumed linear. Until this time, no general and robust solution to the problem has been presented in the literature.

The new approach taken here is based on a perceptually meaningful criterion. Rather than minimizing the amplitude error of the filter, it optimizes the resulting decay times. The error e_τ of the decay times used in the following derivations is defined in the mean-square sense, but it could also be minimized according to minimax or absolute error criteria. Consequently, the results obtained here for least-squares optimization are valid for the other criteria as well, although when weights $w_k \neq 1$ are used, some corrections are necessary. The squared decay time error is defined as :

$$e_\tau = \sum_{k=1}^{K} (\hat{\tau}_k - \tau_k)^2 \tag{6.3}$$

where τ_k are the prescribed, and $\hat{\tau}_k$ are the approximated decay times of partial k. K refers to the number of partials with which the filter design is carried out. By substituting the expression of decay times (see Eq. 4.16) and using the first-order Taylor series approximation for the ln function ($\ln x \approx x - 1$ for $x \approx 1$) one obtains:

$$e_\tau = \sum_{k=1}^{K} \left(\frac{1}{f_0 \ln \hat{g}_k} - \frac{1}{f_0 \ln g_k} \right)^2 \approx \frac{1}{f_0^2} \sum_{k=1}^{K} \left(\frac{1}{1-\hat{g}_k} - \frac{1}{1-g_k} \right)^2 \tag{6.4}$$

where g_k are the prescribed filter magnitudes at the partial frequencies $|vartheta_k$ and $\hat{g}_k = |H_{g_k}(e^{j\vartheta_k})|$ are the corresponding values of the approximation. Since the g_k and $\hat{g}_k$ values are close to 1, the approximation for the error is very accurate.

Based on Eq. (6.5) a transformed filter $H_{tr}(z)$ is designed. This can be done by any least squares filter design algorithm (e.g., `invfreqz` in MATLAB) by using a transformed specification $g_{k,tr}$. This minimizes the error:

$$e_{LS} = \sum_{k=1}^{K} w_k \left(H_{tr}(e^{j\vartheta_k}) - g_{k,tr} \right)^2, \qquad g_{k,tr} = \frac{1}{1-g_k} \tag{6.5}$$

where ϑ_k refers to the frequency of partial k and w_k is the corresponding weight, which is now equal to 1 for all k. The error e_{LS} differs from the right hand side of Eq. (6.4) by a constant factor only, which means that if Eq. (6.5) has a minimum, so do Eqs. (6.4) and (6.3). This implies that the deviation from the prescribed decay times is minimized. The loss filter $H_{g_k}(z)$ can be computed from the transformed filter $H_{tr}(z)$ by an inverse transformation:

$$H_{g_k}(z) = 1 - \frac{1}{H_{tr}(z)} \tag{6.6}$$

Minimizing the error of the decay times puts more emphasis on the lower, slowly decaying partials than on the higher ones, since the same absolute difference will give larger relative error for shorter decay times. The relative error

$$e_{\tau,rel} = \sum_{k=1}^{K} \left(\frac{\hat{\tau}_k}{\tau_k} - 1 \right)^2 \tag{6.7}$$

can also be minimized by using the same procedure described above, but a weighting of $w_k = 1/\tau_k^2$ or the corresponding $w_k = (1 - g_k)^2$ should be used in the least squares filter design algorithm (see Eq. 6.5).

In practice, the approach of Eq. (6.3) was found to be more successful than of Eq. (6.7), since it automatically gives better approximation for the lower harmonics than for the higher ones, and it thus preserves the original decay time. Although simply minimizing the error of the decay times gave good results, it would be still interesting to experiment with other weighting functions, based on some psychoacoustic criteria on the audibility of decay time deviations of the partials. The audibility of the overall decay time deviations were studied in [Tolonen and Järveläinen 2000]. Unfortunately, the research was conducted for the one-pole loop filter case, and hence the results cannot be directly used for designing a weighting function. One interesting outcome of the study is that variations of 25 to 40 % in the overall decay time are inaudible.

There is still one point which has to be considered: calculating the g_k coefficients by Eq. (6.1) is precise only if the inharmonicity of the digital waveguide can be neglected, that is, the partial frequencies are multiples of the fundamental frequency ($f_k = kf_0$). In the case of the piano string model, when the inharmonicity is simulated by a dispersion filter, the effective length of the digital waveguide will be smaller for the higher harmonics than for the lower ones. Consequently, components of high frequency will travel around the digital waveguide more times during the same amount of time than in the harmonic case. This will result in more damping than desired. The audibility of the error caused by this is rather small, since the relative error depends on the corresponding inharmonicity index $I_k = f_k/(kf_0)$, which is close 1 for the perceptually most significant low partials. For the upper ones, some error is introduced, but their amplitude and decay time is small anyway. However, this error can be avoided easily by using a variation of Eq. (6.1), where f_0 is replaced with f_k/k:

$$g_k = e^{-\frac{k}{f_k \tau_k}} \tag{6.8}$$

where f_k is the frequency of the partial k.

The optimization for the decay time deviation can be done by using Eqs. (6.5) and (6.8) with a weighting of $w_k = (k/f_k)^2$ or, the same with angular frequencies, $w_k = (k/\vartheta_k)^2$. This can be easily confirmed by solving Eq. (6.8) for τ_k and substituting into Eq. (6.3).

Analysis of the new method

It is interesting to note that the designed transformed filter $H_{tr}(z)$ has a physical interpretation. When all losses are assumed to come from the termination, it comes from Eq. (6.6) that the inverse of the transformed filter $1/H_{tr}(z)$ is the bridge filter, which calculates the velocity of the bridge v_{br} from the incoming velocity wave v^+. The reflected velocity wave is then computed by a subtraction ($v^- = v_{br} - v^+$). This is similar to the two string case of Fig. 5.9, but now only one string is connected to the termination (see Fig. 6.3).

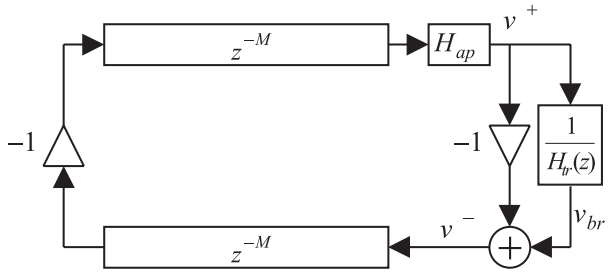

Figure 6.3: Physical interpretation of the transformed filter $H_{tr}(z)$.

By using Eq. (4.5), the transformed filter can be expressed with the following physical parameters:

$$H_{tr}(z) = \frac{v^+(z)}{v_{br}(z)} = \frac{v^+(z)}{v^+(z) + v^-(z)} = \frac{1}{1 + r_v(z)} = \frac{Z(z) + Z_0}{2Z_0} \tag{6.9}$$

where Z_0 is the characteristic impedance of the string, $Z(z)$ is the impedance of the termination and r_v is the velocity reflection coefficient.

Note that in the case of equal string impedances, the coupling filter $H_c(z)$ of Fig. 5.9 can be computed from the designed $H_{tr}(z)$ in a simple way ($H_c(z) = 1/(H_{tr}(z) + 1/2)$). This kind of approach would lead to better results than designing the coupling filter from a one string loss filter prototype, which was suggested in [Van Duyne and Smith 1995], since $H_{tr}(z)$ is optimal in the least squares sense considering the decay times. Mathematically this approach is equivalent to designing the loss filter $H_{g_k}(z)$ from $H_{tr}(z)$ by Eq. (6.6) and then applying the equations of [Van Duyne and Smith 1995] for determining the coupling filter $H_c(z)$ from the single string loss filter $H_{g_k}(z)$.

Practically, the loss filter does not have to be implemented as shown in Fig. 6.3. It can be transformed back by Eq. (6.6), and then the structure of Fig. 4.5 can be used. After this inverse transformation, the resulting filter will always have at least the same amount of zeros as poles. As a special case, if H_{tr} has only poles, the reflection filter H_{g_k} will be an FIR filter.

The stability of the designed filter H_{g_k} is still an issue, since the standard filter design algorithms (such as `invfreqz` in MATLAB) do not always give stable results.

The filter can made stable by reflecting the unstable poles to the unit circle. This will change the phase response only. Since in the phase response of the complete reflection filter the dispersion filter will be dominant, the effect of this change can be neglected. The magnitude response remains the same up to a constant factor, which can be easily corrected [Oppenheim and Schafer 1975].

The stability of the overall digital waveguide loop is another question. The constraint for the stability of the waveguide is that the modulus of the loss filter cannot exceed unity at the resonance frequencies of the string loop:

$$|H_{g_k}(e^{j\vartheta_k})| < 1 \Longleftrightarrow |1 - \frac{1}{H_{tr}(e^{j\vartheta_k})}| < 1 \tag{6.10}$$

This means that $1/H_{tr}(e^{j\vartheta_k})$ must be within a circle of radius 1 centered at $z = 1$. From that comes the stability constraint for $H_{tr}(z)$:

$$\Re\{H_{tr}(e^{j\vartheta_k})\} > \frac{1}{2} \tag{6.11}$$

which should hold for every partial frequency ϑ_k. Since the specification g_k is real and close to one, the transformed specification $g_{k,tr}$ of Eq. 6.5 is much larger than $1/2$. As a result, there is a large margin for the filter design inaccuracies, where the loop still remains stable.

Note that when designing the filter from g_k using traditional methods, the specification g_k is very close to 1, therefore already small error in the approximation makes the loop unstable, on the contrary to the new method. The new method proposed here was used with specifications of piano tones in the entire range of the keyboard. The designed filters never had an amplitude response greater than unity at the prescribed frequency points. On the other hand, they sometimes exceeded unity between the specification points, particularly for high filter orders. The author has not found any cases when it would have led to the instability of the digital waveguide loop, but this still should be avoided. Especially if the number of the prescribed points is small, it can happen that the loop gain exceeds unity at one of the nonspecified resonant frequencies of the digital waveguide. This can be avoided by increasing the number of points in the specification, e.g., by interpolating between the known g_k coefficients.

Fig. 6.4 and 6.5 show the output of the filter design algorithm for the note $A_4^\sharp$ (466 Hz), for IIR filter of orders of 2, 4 and 12. In Fig. 6.4 the specification g_k and the magnitude response of the designed filters at the partial frequencies $\hat{g}_k = |H_{g_k}(e^{j\vartheta_k})|$ are displayed. Note that the match is more precise where the specification is closer to unity. Fig. 6.5 illustrates the decay times of the specification and of the different approximations. This was calculated from $\hat{g}_k$ by Eq. (4.16). The accuracy of the approximation is about the same for every point, since the method minimizes the mean square error of the decay times. The graph of the transformed specification $g_{k,tr}$ is not shown here, as it is very similar to that of the decay times (Fig. 6.5).

The figures reveal that already a second-order IIR filter (solid line) can follow the trend of the decay times, and a fourth-order filter (dashed line) is appropriate

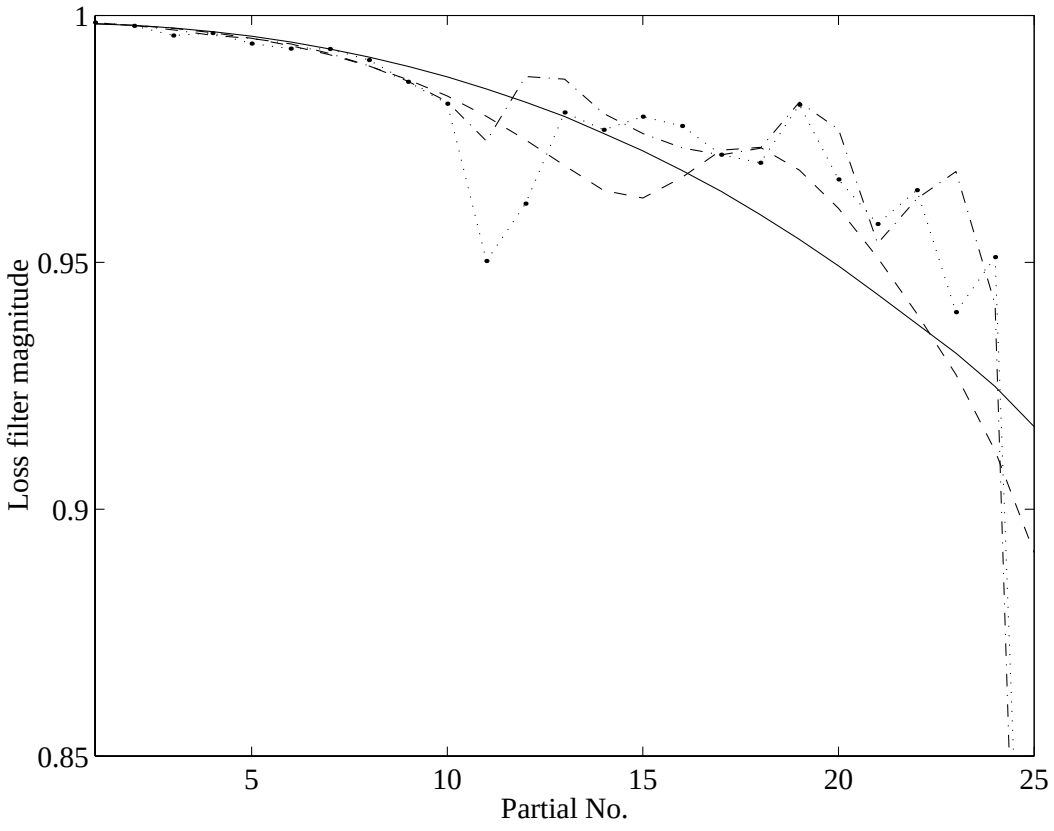

Figure 6.4: Prescribed (dotted line with black points) and approximated loss filter magnitudes for filter orders of 2 (solid line), 4 (dashed line), and 12 (dash-dotted line) using the novel method.

for our purposes (compare with Fig. 6.6 of the one-pole filter). The results of the 12^{th}-order filter (dash-dotted line) design are presented just to show the stability of the method. In practice there is no use for such a large filter, since simulation of the small deviations in the decay times gives only slightly better sound.

6.3.3 New method for designing the one-pole filter

For real-time applications, it is still useful to consider the one-pole filter of Eq. (6.2). It will be also used in the calibration of the soundboard model. Therefore, a simple and robust method is proposed for its design. As proven in the Appendix, the decay times of the digital waveguide with a one-pole loop filter can be approximated with the following formulas:

$$\begin{aligned} \tau &\approx \frac{1}{c_1 + c_3 \vartheta^2} \\ c_1 &= f_0(1-g) \end{aligned} \tag{6.12}$$

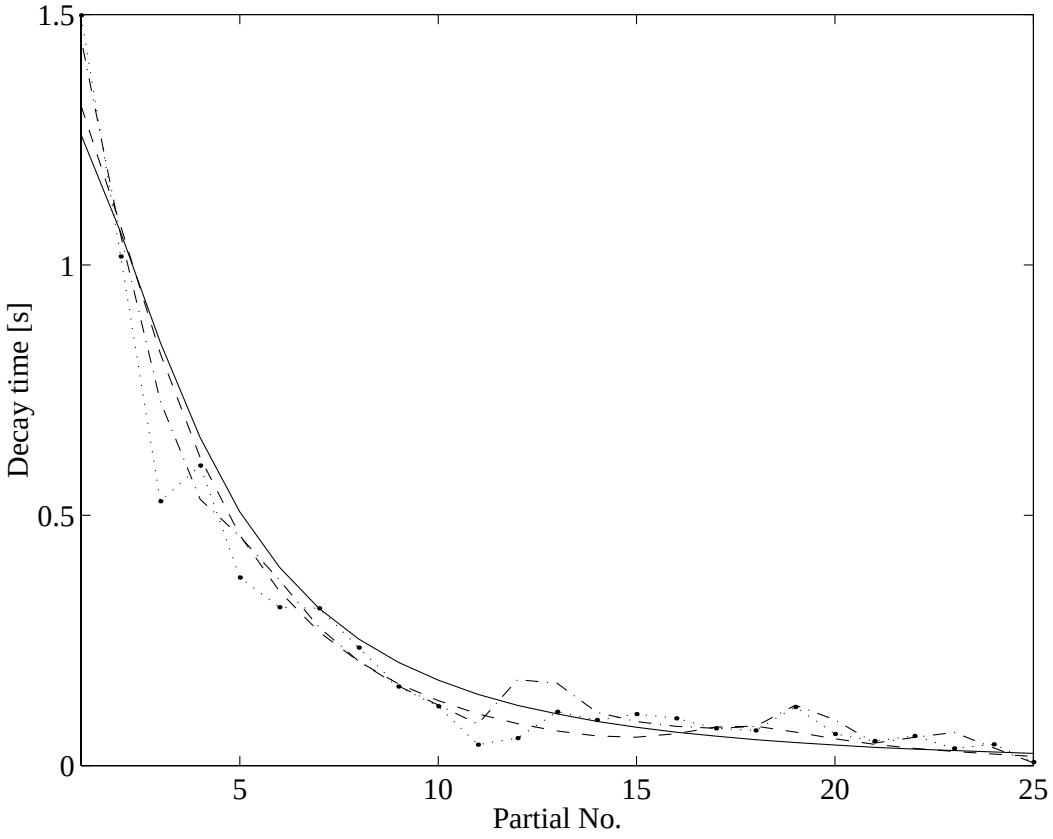

Figure 6.5: Prescribed (dotted line with black points) and approximated decay time for filter orders of 2 (solid line), 4 (dashed line), and 12 (dash-dotted line).

$$c_3 \quad = \quad -f_0 \frac{a_1}{2(a_1+1)^2}$$

where ϑ is the angular frequency in radians, and f_0 is the fundamental frequency of the digital waveguide in Hz. See Appendix for more details and comparison to the one-zero FIR filter.

The results of Eq. (6.12) can also be used in a reverse order: if one knows the c_1 and c_3 parameters, the g and a_1 coefficients of the one-pole filter can be easily calculated.

As it was the approach of Eq. (6.3), the goal is to minimize the mean-square error of the decay times, since that has been found to be a perceptually adequate criterion. The expression of the error e_τ is:

$$e_\tau = \sum_{k=1}^{K} (\hat{\tau}_k - \tau_k)^2 = \sum_{k=1}^{K} \hat{\tau}_k^2 \tau_k^2 \left(\frac{1}{\hat{\tau}_k} - \frac{1}{\tau_k} \right)^2 = \sum_{k=1}^{K} \hat{\tau}_k^2 \tau_k^2 (\hat{\sigma}_k - \sigma_k)^2 \tag{6.13}$$

where $\sigma_k = 1/\tau_k$ are the prescribed, and $\hat{\sigma}_k = 1/\hat{\tau}_k$ are the approximated decay rates.

It can be noted from Eq. (6.12) that the decay rate $\sigma = 1/\tau$ is a second-order polynomial of ϑ. This means that its parameters c_1 and c_3 can be readily computed by means of polynomial regression. The corresponding equation:

$$e_\tau = \sum_{k=1}^{K} w_k (c_1 + c_3 \vartheta_k^2 - \sigma_k)^2 \tag{6.14}$$

where ϑ_k are the angular frequencies of the partials, $w_k = \tau_k^2 \hat{\tau}_k^2$ are the weights, and e_τ is the approximation error which should be minimized with respect to the parameters c_1 and c_3.

The problem with Eq. (6.14) lies in the weights w_k: the approximated decay times $\hat{\tau}_k$ are not known beforehand. This can be solved by first using $w_k = \tau_k^4$ and then running the polynomial regression algorithm again, now computing $\hat{\tau}_k$ from the c_1 and c_3 values by applying Eq. (6.12). This iteration should be done until the error e_τ does not decrease significantly.

Differentiating Eq. (6.14) with respect to c_1 and c_3, and setting $de_\tau/dc_1 = 0$ and $de_\tau/dc_3 = 0$ gives:

$$\begin{aligned} c_3 &= \frac{\mathcal{M}(w_k)\mathcal{M}(w_k\sigma_k\vartheta_k^2) - \mathcal{M}(w_k\sigma_k)\mathcal{M}(w_k\vartheta_k^2)}{\mathcal{M}(w_k)\mathcal{M}(w_k\vartheta_k^4) - \mathcal{M}^2(w_k\vartheta_k^2)} \\ c_1 &= \frac{\mathcal{M}(w_k\sigma_k) - c_3\mathcal{M}(w_k\vartheta_k^2)}{\mathcal{M}(w_k)} \\ \mathcal{M}(x_k) &= \sum_{k=1}^{K} x_k \end{aligned} \tag{6.15}$$

The advantage of polynomial regression is that it is fast to compute and it does not need any iteration or nonlinear approximation technique. However, the polynomial regression should be run at least twice, since the weights w_k can be computed accurately only this way.

Figure 6.6 shows the results of the novel one-pole filter design algorithm for the note $A_4^\sharp$ (466 Hz). It can be seen that already the first approximation ($w_k = \tau_k^4$) gives good results, but it is biased towards the high decay times (dash-dotted line). The second approximation was calculated by using the $\hat{\tau}_k$ values from the output of the first approximation for $w_k = \tau_k^2 \hat{\tau}_k^2$ (dashed line). It is very close to the graph of the 100^{th} iteration (solid line). Consequently, there is no need for using the polynomial regression of Eq. (6.15) more than twice to achieve good results. This was valid for other piano tones as well.

By looking at Fig. 6.6 one notes that the decay times of the first 10 partials are matched quite well but the higher ones are much smaller in the approximation than in the prescription. This is because of the nature of the one-pole filter. A more precise approximation for the higher partials, by e.g., minimizing the error of Eq. (6.7) by applying $w_k = \tau_k^2$, would lead to a large error in the decay times of the first partials, and thus in the overall decay time of the note. Comparing

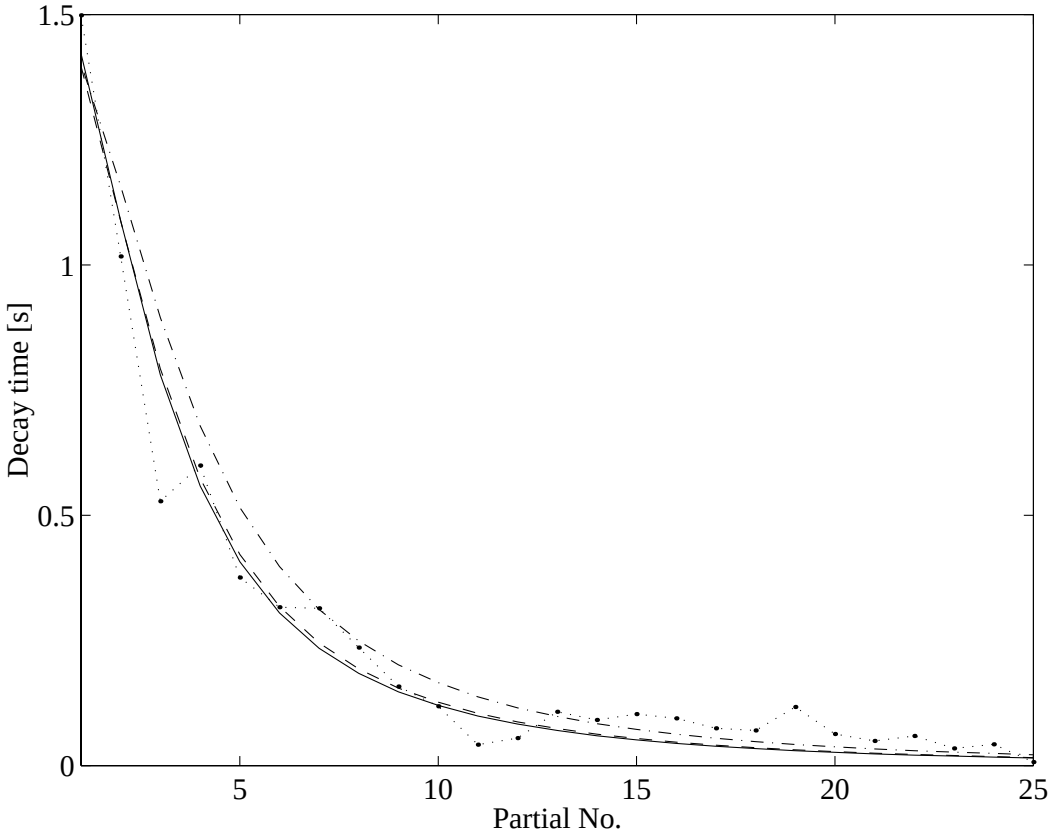

Figure 6.6: Prescribed (dotted line with black points), and approximated decay times for one (dash-dotted line), two (dashed line) and 100 (solid line) iterations by using the new one-pole filter design algorithm.

Fig. 6.6 to Fig. 6.5, it can be seen that a fourth-order IIR filter can give a solution to this problem. It seems that the decay times of piano strings do not exactly follow the theoretical curve of the string with first- and third-order losses. This is mainly because that simple model does not take into account the string termination.

On the whole, the novel algorithm presented here gives good results for designing the one-pole loss filter. The error criterion optimizing for the decay times seems to be appropriate, but any other kind of weighting can be used, if necessary. However, it will not improve the approximation significantly: for simulating the decay times of the piano string higher-order filters are needed. By using derivation similar to that in Eqs. (A.2)-(A.5) for higher-order filters, i.e., expressing the decay rate σ as the polynomial of ϑ, the same filter design method could be used. For example, a two-pole filter with real poles will lead to a polynomial of $\hat{\sigma} \approx c_1 + c_3\vartheta^2 + c_5\vartheta^4$, which is the decay time of the string with first-, third- and fifth-order time derivates in the wave equation. Using this polynomial regression based approach could give better results than the transformation of the specification Eq. (6.5), since the design is now fully under control and no standard filter design algorithm need to be used. The

disadvantage of this technique is that quite much analytical work is needed, making it unrealizable for high filter orders. On the contrary, the transformation method can be used for those cases easily. The author suggests the polynomial regression based approach for low filter orders, such as one or two, and the transformation method for higher ones.

6.3.4 Designing the dispersion filter

The dispersion of the string is implemented by an allpass filter. This ensures that the decay times will be controlled by the loss filter and the stability of the delay loop will depend on the loss filter only. However, these are influenced by the dispersion filter, since it sets the resonance frequencies of the digital waveguide. By reformulating Eq. (4.18) the effective length D_{eff} of the digital waveguide should be:

$$D_{eff}(f_k) = \frac{f_s k}{f_k} = D_{wg} + D_{r_F}(f_k) \tag{6.16}$$

where f_s is the sampling frequency, f_k is the frequency of the partial k, $D_{wg} = N$ is the phase delay introduced by delay lines of the digital waveguide, and D_{r_F} is the phase delay of the reflection filter.

Since the phase delay of the loss filter can be neglected when designing the dispersion filter, the specification for the dispersion filter will be:

$$D_{disp}(f_k) = \frac{f_s k}{f_k} - D_{wg} \tag{6.17}$$

When the inharmonicity indices $f_k/(kf_0)$ increase with frequency, which is the case for real strings, this will lead to a monotonically decreasing specification. The delay of the digital waveguide D_{wg} is a free parameter during the filter design.

Here the approach suggested by Van Duyne and Smith [1994] is taken. The dispersion filter is a series of first-order allpass filters, whose transfer function is as follows:

$$H_{1ap}(z) = \frac{a_1 + z^{-1}}{1 + a_1 z^{-1}} \tag{6.18}$$

The phase delay of such a filter is similar to that what is needed for simulating the inharmonicity of real strings. Unfortunately, one filter is not enough to implement the necessary phase delay. Therefore more first-order allpass filters are connected in series [Van Duyne and Smith 1994].

In [Van Duyne and Smith 1994] using the same a_1 coefficients for all the first-order allpass filters was motivated by computational reasons. They suggested a multiplier-free implementation with a_1 values being a power of 2 or 1 minus a power of two. Here we do not restrict ourselves to these kind of coefficients, a_1 can be any rational number. However, using the same a_1 for all the filters in the chain simplifies the filter design process considerably.

The filter coefficient a_1 was determined by minimizing the least squares error in the phase delay. A weighting was used which gave more emphasis to the lower

partials. The number of first-order allpass filters in the chain was a given parameter. For the lowest notes of the piano, 16 first-order allpass filters were connected in chain. For the higher ones, this number was 8 or 4. This was found to give satisfactory sonic results, but it is not optimal in any sense. More work has to be done concerning this part of the model.

Rocchesso and Scalcon [1996] suggested a more refined method by using the allpass filter design technique of Lang and Laakso [1994]. The analysis of that technique and the sonic comparison to the simple method applied here will be a part of future research. The results of listening tests by Rocchesso and Scalcon [1999] and Järveläinen et al. [1999] could be also incorporated in the filter design.

6.3.5 Setting the fundamental frequency

As mentioned earlier, the phase delay of the digital waveguide loop should be very accurate at the fundamental frequency of the note. If it is not so, the corresponding note will be out of tune. Once the loss and the dispersion filters are designed, the length of the digital waveguide $D(f_0)$ at the fundamental frequency f_0 of the tone should be:

$$D(f_0) = D_{wg} + D_{fd} = \frac{f_s}{f_0} - D_{g_k}(f_0) - D_{disp}(f_0) \tag{6.19}$$

where D_{g_k} and D_{disp} are the phase delay of the loss and dispersion filters, respectively.

The prescribed length of the waveguide $D(f_0)$ is not an integer. On the contrary, only an integer phase delay can be implemented with delay lines. The solution is to use fractional delay filters [Välimäki 1995; Laakso et al. 1996] in series with the delay line. Here a first-order allpass filter was used for this purpose, as proposed in [Smith 1983; Jaffe and Smith 1983], whose transfer function is described by Eq. (6.18). The integer part of $D(f_0)$ will be implemented as the delay line of the digital waveguide with a length of $D_{wg} = \lfloor D(f_0) - 0.5 \rfloor$, and the fractional part will be realized with a first-order allpass filter with a phase delay of $D_{fd}(f_0) = D(f_0) - D_{wg}$ at the fundamental frequency f_0. The resulting fractional delay will be $0.5 \leq D_{fd}(f_0) < 1.5$ which is the optimal range for the first-order allpass filter [Välimäki 1995]. The a_1 coefficient of the allpass filter can be approximately calculated as $a_1 = (1 - D_{fd}(f_0))/(1 + D_{fd}(f_0))$ [Smith 1983; Jaffe and Smith 1983; Välimäki 1995].

6.4 Calibration of the resonator bank

This section will deal with the calibration of the novel resonator bank implementation for beating and two-stage decay. Since the resonator based approach for modeling these effects is new, no analysis techniques could be found in the literature. The problem is similar to what is needed for designing the filters for the model of Fig. 5.10, but the corresponding paper [Daudet et al. 1999] only says that the authors used parametric modeling. The method presented here has the advantage

that there is no nonlinear optimization needed and it is robust for real piano tones. Consequently, this new approach could also be useful for calibrating the model of Daudet et al. [1999].

The structure was described in Section 5.3.3, and now the task is to determine the poles p_k and the amplitude coefficients a_k for the resonators (see Eqs. (5.2) and (5.3)). To be able to do that, first we have to investigate what kind of amplitude envelopes can arise from two exponentially damping sinusoids.

6.4.1 Analysis of the two-mode model

Our starting-point is the two-mode model, with two exponentially damping sinusoids, whose initial amplitude, phase, decay time and fundamental frequency parameters can all differ. This simplified model was found useful in describing the main features of beating and two stage decay [Weinreich 1977]. The physics of these two phenomena was described in Section 2.2.1. Let us assume that the parameters of the two modes are already known. Then the sum of the two modes is:

$$x(t) = A_1 e^{-\frac{t}{\tau_1}} \sin(2\pi f_1 t + \varphi_1) + A_2 e^{-\frac{t}{\tau_2}} \sin(2\pi f_2 t + \varphi_2) \tag{6.20}$$

where A_1 and A_2 are the initial amplitudes, φ_1 and φ_2 are the initial phases, f_1 and f_2 are the frequencies, and τ_1 and τ_2 are the decay times of the two modes.

To simplify the derivation, the following notation will be used:

$$\begin{aligned} a &= A_1 e^{-\frac{t}{\tau_1}} \\ b &= A_2 e^{-\frac{t}{\tau_2}} \\ \alpha &= 2\pi f_1 t + \varphi_1 \\ \beta &= 2\pi f_2 t + \varphi_2 \end{aligned} \tag{6.21}$$

Accordingly, the sum of the two modes can be written as:

$$\begin{aligned} x(t) &= a\sin\alpha + b\sin\beta = \frac{a+b}{2}(\sin\alpha + \sin\beta) + \frac{a-b}{2}(\sin\alpha - \sin\beta) \\ &= \left((a+b)\cos\frac{\alpha-\beta}{2}\right)\sin\frac{\alpha+\beta}{2} + \left((a-b)\sin\frac{\alpha-\beta}{2}\right)\cos\frac{\alpha+\beta}{2} \end{aligned} \tag{6.22}$$

Since $f_1 \approx f_2$, Eq. (6.22) can be seen as a sum of an amplitude modulated sine and an amplitude modulated cosine function. We are not interested here in the instantaneous phase or frequency of the resulted signal, but in the evolution of the modulating wave. The amplitude envelope $A(t)$ of the signal $x(t)$ will be:

$$\begin{aligned} A(t) &= \sqrt{(a+b)^2\cos^2\frac{\alpha-\beta}{2} + (a-b)^2\sin^2\frac{\alpha-\beta}{2}} \\ &= \sqrt{a^2+b^2+2ab\left(\cos^2\frac{\alpha-\beta}{2} - \sin^2\frac{\alpha-\beta}{2}\right)} \\ &= \sqrt{a^2+b^2+2ab\cos(\alpha-\beta)} \end{aligned} \tag{6.23}$$

The geometrical interpretation of this equation is the length of the sum of two vectors in the x-y plane, whose lengths are a and b, and their angles to the x-axis are α and β.

By substituting the original variables according to Eq. (6.21) we obtain:

$$A(t) = \sqrt{A_1^2 e^{-\frac{2t}{\tau_1}} + A_2^2 e^{-\frac{2t}{\tau_2}} + 2A_1 A_2 e^{-t\left(\frac{1}{\tau_1}+\frac{1}{\tau_2}\right)} \cos(2\pi(f_1 - f_2)t + \varphi_1 - \varphi_2)} \quad (6.24)$$

This is the amplitude envelope of the signal $x(t)$. It covers not only the beating, but also the decay of the sound. We can have better insight to beating if we reformulate Eq. (6.20):

$$\begin{aligned} x(t) &= A_1 e^{-\frac{t}{\tau_1}} \left(\sin(2\pi f_1 t + \varphi_1) + \frac{A_2}{A_1} e^{-t\left(\frac{1}{\tau_2}-\frac{1}{\tau_1}\right)} \sin(2\pi f_2 t + \varphi_2) \right) \\ &= A_1 e^{-\frac{t}{\tau_1}} x_{rel}(t) \end{aligned} \quad (6.25)$$

Now we have factored out the general exponential decay and normalized the initial amplitudes. The amplitude envelope $b(t)$ of the normalized signal $x_{rel}(t)$ will be:

$$\begin{aligned} b(t) &= \sqrt{1 + A_b^2 e^{-\frac{2t}{\tau_b}} + 2A_b e^{-\frac{t}{\tau_b}} \cos(2\pi f_b t + \varphi_b)} \\ & A_b = \frac{A_2}{A_1} \\ & \tau_b = \frac{1}{\tau_2} - \frac{1}{\tau_1} \\ & f_b = f_1 - f_2 \\ & \varphi_b = \varphi_1 - \varphi_2 \end{aligned} \quad (6.26)$$

We see from Eq. (6.26) that for $A_b \geq 0$ the upper and lower limits for the beating signal $b(t)$ are:

$$|1 - A_b e^{-\frac{t}{\tau_b}}| \leq b(t) \leq 1 + A_b e^{-\frac{t}{\tau_b}} \quad (6.27)$$

From Eq. (6.27) follows that the beating is the largest when $A_b e^{-t/\tau_b} = 1$, which is when the amplitude envelope of the two decaying sinusoids cross. Before this point, the decay grows, and after that decays. Note that the largest beating does not refer to the maximal $b(t)$ value, but to the maximal relative difference between the lower and upper bounds of $b(t)$, which is when the lower bound equals to zero.

For those cases where $0 \leq A_b e^{-t/\tau_b} \leq 1$, the beating $b(t)$ can be approximated with an exponentially rising or decaying sinusoid:

$$b(t) \approx 1 + A_b e^{-\frac{t}{\tau_b}} \cos(2\pi f_b t + \varphi_b) \quad (6.28)$$

The smaller the $A_b e^{-t/\tau_b}$ value, the more accurate the approximation. It comes form the constraint $A_b e^{-t/\tau_b} \leq 1$ that Eq. (6.28) handles only either the decaying, or the growing part of the beating. If we separate these parts during the analysis of real amplitude envelopes, and fit this kind of curve on one of them, we can get all the parameters which characterize the two-mode model. However, the synthesis model will behave as in Eq. (6.26).

6.4.2 Estimation of beating and two stage decay

The analysis is based on the theory described previously. The goal is first to remove the general exponential decay from the amplitude envelope $A(t)$ and then to fit a signal model of Eq. (6.28) on the data. The analysis should be restricted only either to the growing, or to the decaying part of the beating, since the simplified model is valid only for this case. The separation can be done by finding the minimum of the normalized beating signal $b(t)$ after the general exponential decay is removed from the amplitude envelope $A(t)$.

Figure 6.7 shows the parts of the algorithm. First, the algorithm of Fig. 6.2 is applied to the signal, which calculates the amplitude envelopes $A(t)$ of the partials and estimates the initial amplitude and decay time of the exponential decay $A_1 e^{-t/\tau_1}$ for every single partial. The exponential trend is removed from $A(t)$ by division, yielding $b'(t) = A(t)/(A_1 e^{-t/\tau_1})$. Since we try to fit a model of Eq. (6.28) to the beating, we need a signal whose mean is 1. Therefore, $b'(t)$ has to be normalized to get the beating signal $b(t) = b'(t)/Mean\{b'(t)\}$. Then the DC component of $b(t)$ is removed, $s(t) = b(t) - 1$. The frequency f_b and initial phase φ_b values of the beating are found by taking the DFT of $s(t)$ and picking the frequency and phase values where the magnitude of the DFT is the largest.

Computing the amplitude A_b and decay time τ_b of the beating is more complicated. Theoretically, the same procedure could be used as in Fig. 6.2, since the envelope of an amplitude modulated sinusoidal signal $s(t)$ has to be determined. The difference is that here the frequency of the signal f_b is very low, it is in the order of 1 Hz. Accordingly, the lowpass filter of Fig. 6.2 should be very narrow, which would lead to an unacceptably bad time resolution. The solution here is based on recognizing that the signal contains only one dominant frequency component, and therefore no filtering is needed. The amplitude envelope is calculated by taking the absolute value of the analytic signal of $s(t)$, which is $|s(t) + j\mathcal{H}\{s(t)\}|$, where $\mathcal{H}$ denotes the Hilbert transform [Oppenheim and Schafer 1975]. The initial amplitude A_b and decay time τ_b is estimated by means of linear regression in the logarithmic domain, in the same way as it was mentioned in Section 6.2.1.

Note that all the operations performed here concern only one particular partial of the analyzed signal. This procedure should be run for every partial separately.

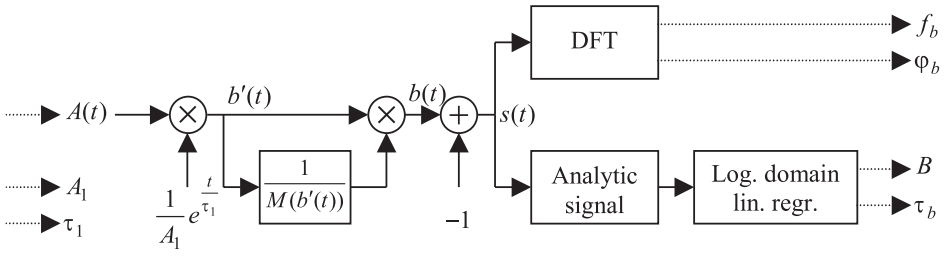

Figure 6.7: The proposed analysis procedure for beating and two-stage decay.

Figures 6.8 and 6.9 show the fitting procedure on the first and second partial of the note $A_4^\sharp$, respectively. It can be seen that the first partial exhibits two-stage decay (Fig. 6.8), and the second one is featured by a growing beating (Fig. 6.9).

In both figures, (a) solid line shows the amplitude envelope of the partial and the dashed line is the fitted exponential decay $A_1 e^{-t/\tau_1}$. The normalized beating signal $b(t)$ is displayed in (c), solid line. The magnitude of the analytic signal of $s(t)$, i.e., the amplitude of the beating is revealed in (b), solid line. The dashed line in (b) refers to the fitted exponential envelope $A_b e^{-t/\tau_b}$. The dash-dotted line in (c) is the approximated $b(t)$ signal by using the model of Eq. (6.28), and the dash-dotted line in (d) shows the approximated amplitude envelope.

The dashed lines of (c) and (d) display the real $b(t)$ and $A(t)$ signals, which will arise if the parameters determined by this model are implemented with two exponentially damping sinusoids. These signals were calculated by applying Eq. (6.26). Note that the dash-dotted and dashed lines in (c) and (d) are close to each other. This shows that the model of Eq. (6.28) is a good approximation.

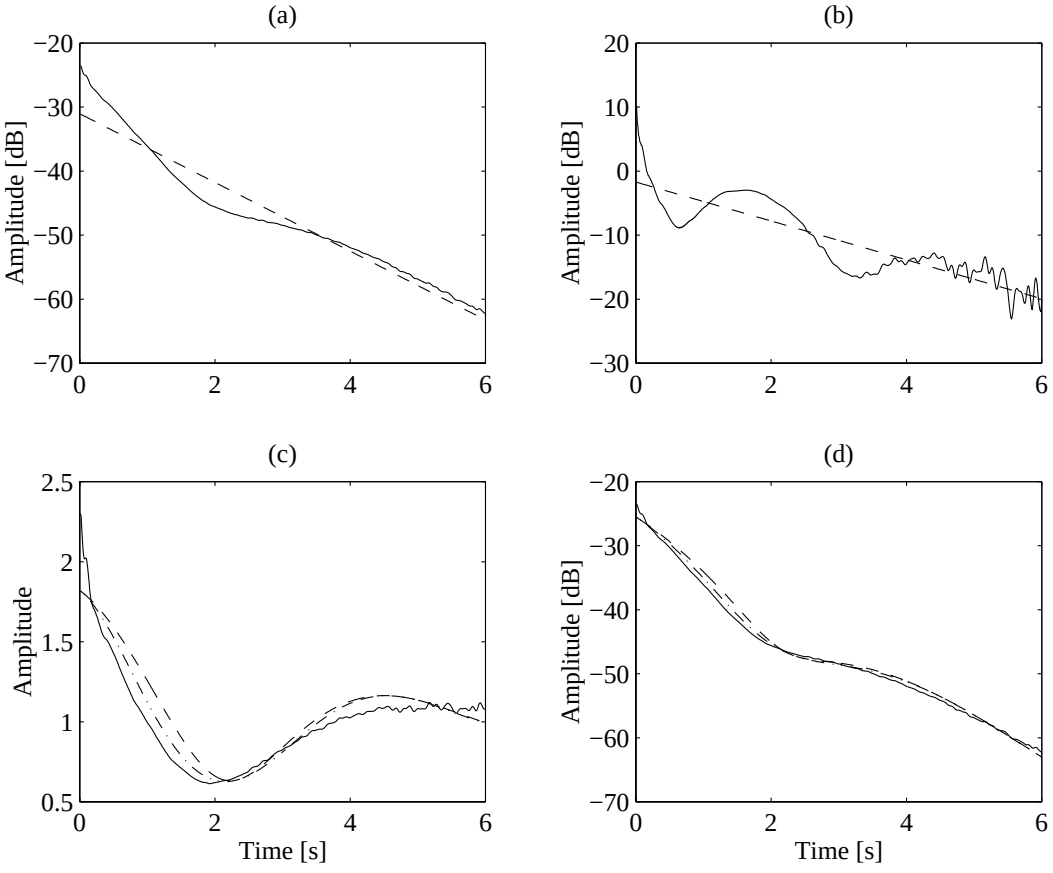

Figure 6.8: Analysis of the first partial of the $A_4^\sharp$ piano tone.

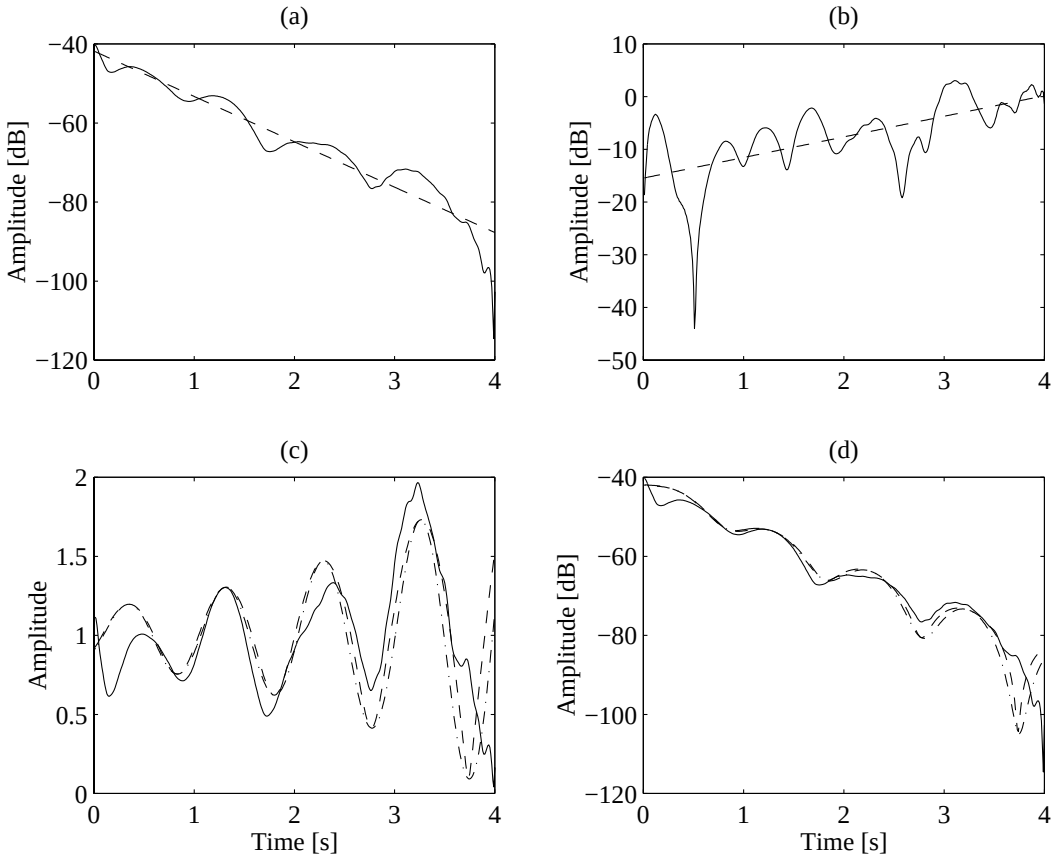

Figure 6.9: Analysis of the second partial of the $A_4^\sharp$ piano tone.

6.4.3 The parameters of the resonator bank

Now that the f_b, φ_b, A_b, and τ_b parameters of the beating are known, the parameters of the resonators can be calculated by the help of Eq. (6.26). However, the A_1 and τ_1 parameters determined during the analysis cannot be used for this purpose, since the digital waveguide gives only a rough approximation of these. Consequently, the features of the resulted beating would be different. Thus, the parameters of the resonators have to be determined relatively to the digital waveguide. As a result, the characteristics of the beating are preserved, but the decay times and the initial amplitudes of the partials will differ from those of the original signal. This is perceptually acceptable, although the synthetic envelopes will not match the real ones precisely. The steps of the calibration process are displayed in Fig. 6.10.

First the digital waveguide of the string model is designed. Then its behavior is analyzed. One way is simulating the impulse response of the corresponding string model by sending an impulse to the force input, without using the hammer model. Then this signal can be analyzed in the same way as it was described in Section 6.2.1, in order to obtain the frequencies, initial amplitudes and decay times of the partials. The initial phases can also be determined from the phase of the lowpass filtered signal

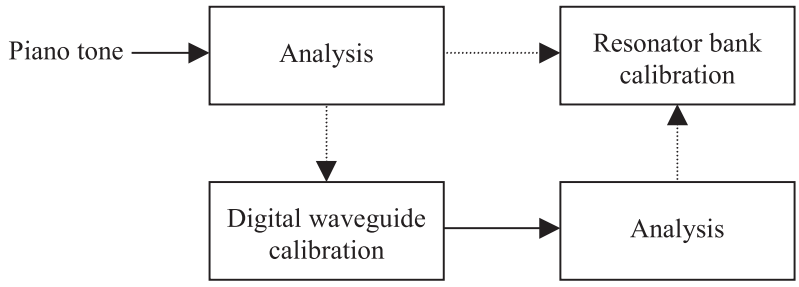

Figure 6.10: The calibration process for the resonators.

in Fig. 6.2. Another solution is the analytic determination of these parameters by using an approach similar to Chapter 4, Eqs. (4.15)-(4.18). Nevertheless, this needs more careful considerations, since, e.g., the length of the waveguide N in Eq. (4.15) varies with frequency, as an effect of the dispersion filter.

Once the A_1, τ_1, f_1, and φ_1 parameters of one partial of the digital waveguide are known, the A_2, τ_2, f_2, and φ_2 parameters of the corresponding resonator are determined by Eq. (6.26).

The benefit of the new structure of Section 5.3.3 is that the accuracy of the approximation can be adjusted by the number of the realized resonators. This gives a new task: since not all the resonators are implemented, a decision has to be made which partials should have their beating and two-stage decay simulated. The others will have simple exponential decay determined by the digital waveguide. The best method would be applying psychoacoustic criteria, but these were not found in the literature. The method taken here is based on the energy of the resonators and it is quite simple: the resonators with the K largest energy value are implemented. This approach tries to minimize the energy difference between the implementation with K resonators and the imaginary case when resonators for all partials are used. This works pretty well for tones in the high and middle range, where the number of partials is small anyway, but for the lower ones the sonic result is not satisfactory. The resonators chosen by this algorithm are generally next to each other, and probably a broader distribution in frequency would be more appropriate. Choosing the resonators manually gave better results after some experiments. Therefore, in the future a better criterion has to be found for this purpose.

In the upper part of Fig. 6.11 the first 8 partial envelopes of the $A_4^\sharp$ note (466Hz) are displayed. The lower part of Fig. 6.11 shows the output of the synthesis model including all parts of the model, hence the hammer and the soundboard as well. For the string model, the structure of Fig. 5.12 with 5 resonators was used. The parameters of the resonator bank were determined by the method described in this section. It can be noted that the characteristics of beating and two-stage decay are well preserved, but the initial amplitudes and the decay times of the partials are different for the original and synthetic signals. This is by the nature of physical

modeling techniques, since they model the sound generating behavior, and not the signal, therefore only the main features of the sound are preserved. However, we are not interested here to copy one specific piano tone, but to build a model which sounds like a piano. The beating and the two-stage decay can be more accurate because they are simulated by separate resonators.

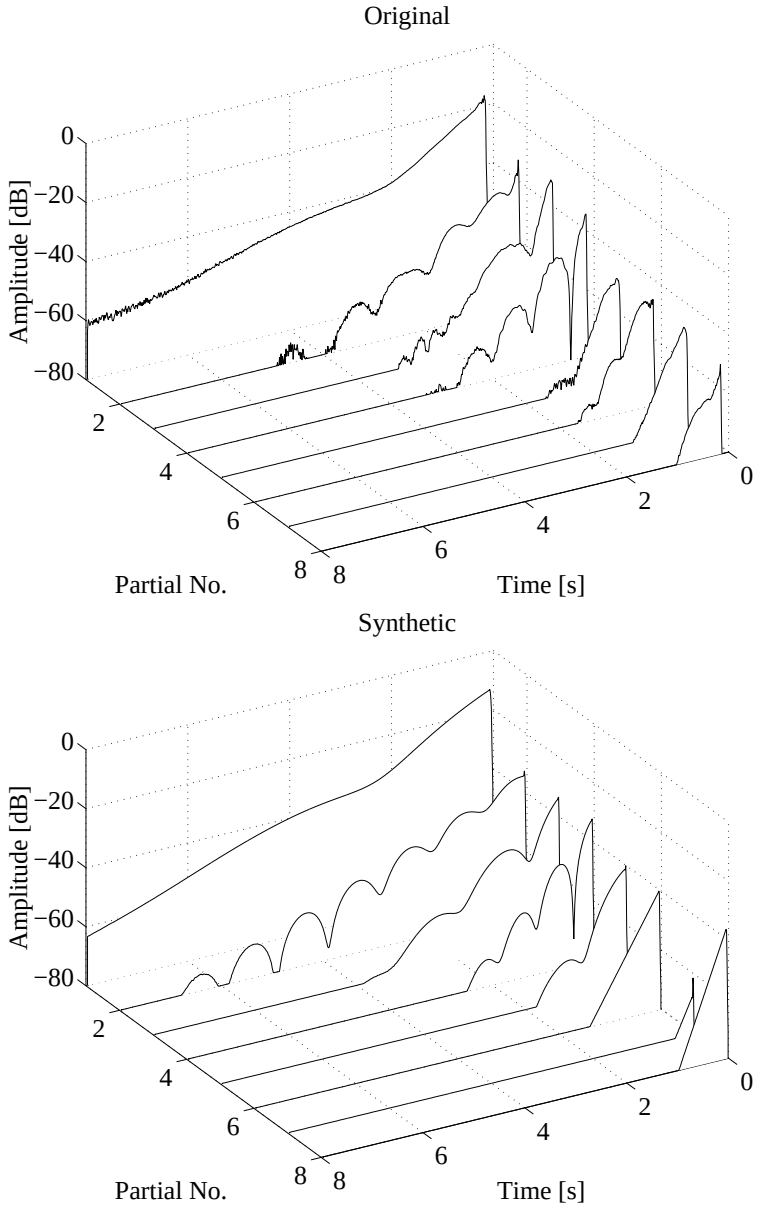

Figure 6.11: Partial envelopes of the original and synthesized $A_4^\sharp$ note.

To conclude, the new algorithm described here provides a robust solution for calibrating the resonator bank. Although it contains many steps, and may therefore seem to be complicated, these steps are simple operations. An advantage is that

there is no need for any iteration or nonlinear optimization, and this ensures the stability of the algorithm. The method was found to be reliable for all analyzed piano tones.

The problem of the technique described here is that it should be done for all piano notes. The decay times for reflection filter design can be interpolated from one or two measurements in every octave. On the contrary, the characteristics of beating and two-stage decay are so different from note to note that these cannot be determined by interpolation between the measured notes. By measuring the impedance at some points along the bridge and calculating the parameters from that could possibly overcome this problem. Then the impedance curve should be sampled at the modal frequencies of the string and the parameters of the beating could be calculated using the equations described in [Weinreich 1977]. This probably would not give the same features as of the original tones, but it would give physically meaningful behavior. The robustness of such a method is a question, since the noise in impedance measurement can make the whole process unstable. Therefore, the author suggests measuring all piano tones and using the method based on partial envelopes. The analysis process is automatic. The only time-consuming operation is the recording of the 88 piano notes.

6.5 Calibration of the soundboard model

The calibration of the soundboard model consists of two parts: determining the parameters of the feedback delay network and designing the shaping filter.

6.5.1 Parameters of the feedback delay network

The feedback delay network is described by the variables of Eq. (5.4) and by the loss filters $H_k(z)$ connected to the delay lines in series. Choosing the length of the delay lines was already discussed in Section 5.4.3.

The feedback matrix $\mathbf{A}$ was first chosen to be a special case of a circular matrix, where all the eigenvalues of $\mathbf{A}$ are at 1 expect one, which is at -1. This results in a feedback matrix, where the elements are the same expect the ones in the diagonal [Rocchesso and Smith 1997]. This leads to a very efficient implementation. However, when the number of delay lines is higher than four, the magnitude of the elements in the diagonal will be larger than the others. Consequently, the delay lines will get more input from their own output than desired, and this results in more harmonic timbre. Better solutions can be obtained by shifting the feedback matrix to the right with one step. This spreads the eigenvalues of the feedback matrix around the unit circle [Rocchesso and Smith 1997]. Thus, the first row of the circular feedback matrix in our case is: $[-1/4, 1-1/4, -1/4, -1/4, -1/4, -1/4, -1/4, -1/4]$.

The input and output coefficient vectors $\mathbf{b}$ and $\mathbf{c}$ were chosen as suggested in [Rocchesso and Smith 1997] to avoid coloration in artificial reverberators. However, the equations in [Rocchesso and Smith 1997] are valid when the lengths of the delay

lines are comparable. This is not the case here, hence the coloration cannot be completely avoided. Note that this is not a problem, since real piano soundboards color the sound as well.

The loss filters were designed from the decay times of the different frequency regions of the impulse response, in a way suggested in [Jot and Chaigne 1991]. The bridge of the piano was excited with an impulse hammer and the resulting sound pressure was recorded. Then the temporal energy evolution of the response was calculated in octave bands. The curves were smoothened by reverse-time integration [Jot and Chaigne 1991]. The decay times were determined by applying linear regression to the logarithm of the smoothed envelopes, as discussed in Section 6.2.2.

One-pole filters were used for loss filters, and their parameters were determined by the novel one-pole filter design algorithm of Section 6.3.3. It is done as follows: first the c_1 and c_3 parameters of Eq. (6.12) are calculated by using the method of Section 6.3. These will be the same for all the loss filters. Then the g and a_1 parameters of the one-pole filter are computed by using the inverse of Eq. (6.12) with $f_0 = f_s/m_k$, where f_s is the sampling frequency and m_k is the length of the delay line to which the loss filter $H_k(z)$ is connected in series.

The decay times of the real piano soundboard (dotted lines with black dots) and of the feedback delay network (solid line) are shown in Fig. 6.12. For measurement, the bridge was excited at the string $A_4^\sharp$ with an impulse hammer.

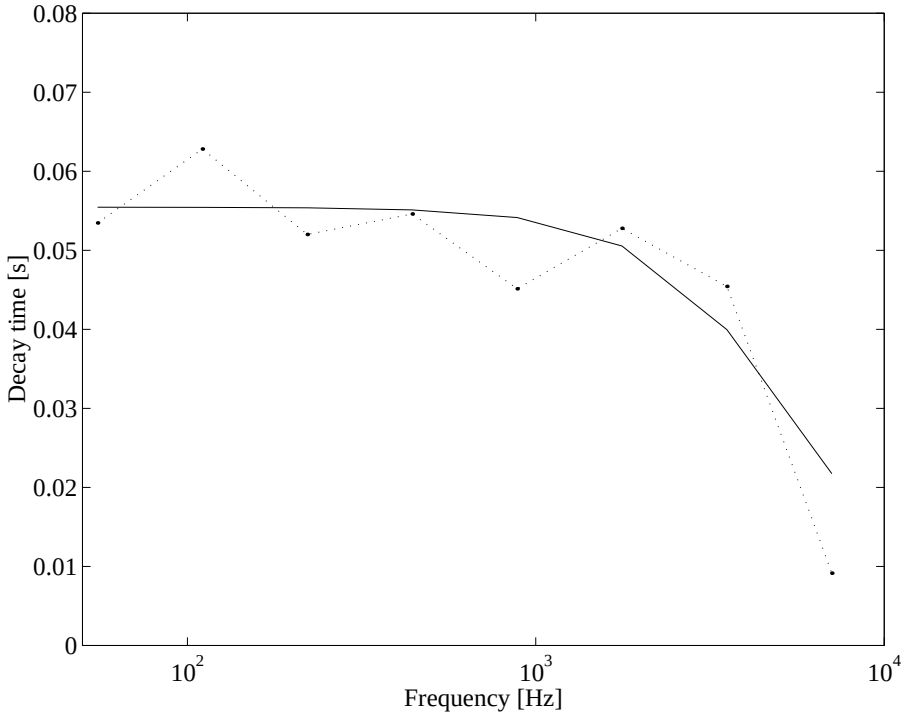

Figure 6.12: Original (dotted line with black dots) and synthesized (solid line) decay times of a piano soundboard.

It can be seen from Fig. 6.12 that the one-pole filter can follow only the main trend of the decay times. However, the perceptual quality of the soundboard model with such a simple filter is already quite high. It seems that there is no need for very complicated loss filters to reproduce the general characteristics of a piano soundboard.

6.5.2 Shaping filter design

For shaping filter design, the bridge of the soundboard was excited by an impulse hammer and the force and sound pressure signals were synchronously recorded. Then their spectra were computed by the Discrete Fourier Transform. The transfer function was obtained by the frequency domain division of the two signals. Unfortunately, the noise level of both signals were high. This is especially critical in the force signal, since its spectrum is in the denominator. Where the force spectrum approaches zero, the transfer function of the soundboard has high peaks and this leads to useless results.

This can be avoided by noting that the information content of the force signal is concentrated only to the first hundred samples, since the excitation is of an impulse type. After the impulse, when the force hammer has lost contact to the bridge, all the values of the force signal were set to zero. This resulted in a large improvement in the noise level. The reduction is proportional to the length of the force impulse and of the analyzed record, which comes from the length of the pressure signal. For typical measurements, the noise energy in the force signal was reduced to the $1/50$ by this technique.

The impulse response of the soundboard was obtained by the Inverse Discrete Fourier Transform of the transfer function. The shaping filter $H_{sh}(z)$ in Fig. 5.14 is the windowed version of the impulse response. The length of the FIR filter used in the model was 100 taps. Using IIR filters could lead to a more efficient implementation. However, if the filter order is high, the filter design algorithms easily run into stability problems.

The specifications for the correction filters $H_{high}(z)$ and $H_{low}(z)$ in Fig. 5.14 for the high and low range of the keyboard, respectively, are obtained by dividing the transfer functions of the soundboard at the high and low range by the transfer function of the shaping filter $H_{sh}(z)$. Then this frequency-domain data is smoothed and the filters are designed by standard filter design methods. Typically, second or third-order IIR filters give good results.

6.6 Conclusion

In this chapter the calibration of the piano model was outlined. For the loss filter of the digital waveguide, a new design algorithm was presented. It is based on the transformation of the specification and optimizes the resulting decay times. The method has been found stable even for high filter orders. A fast and robust algo-

rithm was presented for the design of the one-pole loss filter, founded on polynomial regression. A new algorithm was developed for the measurement of beating and two-stage decay, which can be used for the calibration of the resonator bank. Methods for determining the parameters of the soundboard model were also described.

Chapter 7

Summary and future directions

7.1 The results

The thesis presented a model for the synthesis of piano sound. The approach is grounded on physical principles, because the structure of the model follows that of the real piano. However, the implementations of the different parts are not necessarily physical. The aim of the author was to find the most efficient fusion of the different sound synthesis approaches.

The acoustical properties of the piano were discussed by both reviewing the results of the literature and presenting the outcomes of measurements. Implications were given for which features of the piano are responsible for the distinctive character of the timbre. This part of the work forms the basis for the sound synthesis of the piano.

This was followed in Chapter 4 by the discussion of an efficient string modeling technique: the digital waveguide modeling. The equations describing the digital waveguide were given. The fractional expansion of the transfer function resulted in a new approach in considering the behavior of digital waveguide. These equations allowed the formulation of the equivalent resonator bank structure. When the poles of the digital waveguide are equally distributed along the unit circle, the impulse response of the waveguide and the amplitude coefficients of the resonators are related by the Discrete Fourier Transform. For lossy and dispersive digital waveguides, the resonator coefficients can be determined by solving a set of linear equations. Because of being capable of implementing delay, the resonator bank can be also used for simulating nonlinear interaction. This can be beneficial for modeling largely inharmonic instruments, such as, mallet percussion or bells.

After presenting the general principles of string modeling, Chapter 5 concentrated on modeling a specific instrument, the piano. The piano model is a combination of different approaches. The physical modeling approach is most advantageous for those parts of the model which are controlled by the musician. In the case of the piano, the input from the outside world is the movement of the hammer and the dampers. Consequently, the hammer is modeled by discretizing the physical equations of real hammers. Since the force signal experienced by the string is influenced

by the movement of the string at the excitation point, it is essential to implement the string as a physical model. The basic string behavior is modeled by a digital waveguide. A feature of these two parts of the piano model is that all the variables present are physical units. The beating and two-stage decay of the piano sound are simulated by a parallel resonator bank. The frequencies and the decay times of the resonators are determined through the analysis, however, the initial amplitudes are controlled by the excitation signal, which is computed by the hammer model and the digital waveguide. This way, even if it is treated in the analysis phase as a signal model, it responds to the variation of impact velocity in a meaningful way. The soundboard of the piano is modeled by a feedback delay network. This can be considered as a statistical model of the impulse response of the soundboard since it matches only its main features. The physical parameters of the soundboard, such as size, shape, etc., are not visible in the model. Since these parameters cannot be influenced by the musician, the nonphysical approach is adequate here. The message of the thesis is that the secret of efficient sound synthesis lies in the combination of different approaches.

During the description of the piano structure, some new inventions were also presented. The thesis pointed out the discontinuity problem, which arises when an interaction force is taken into the string. Then, solutions to the problem were proposed. A multi-rate hammer was presented in order to overcome the stability problem of the hammer model. The use of the same structure for modeling the effect of dampers was outlined. A resonator bank based technique was introduced for modeling the beating and two-stage decay. For the realistic simulation of the characteristic attack noise of piano sounds a simple technique was proposed. To resolve the difference between the computational load presented by the string models of the high and low register, a multi-rate piano model was proposed.

Novel techniques for calibrating the model were also developed. For designing the loss filter of the digital waveguide, a transformation based technique was presented. The method optimizes the resulting decay times, which has a perceptual sense. The stability of the designed filter is also assured. Its another advantage is simple implementation. In the Appendix, the link between the decay times of a lossy string and a digital waveguide with a one-pole filter is described. The equivalence of the one-zero filter to the one-pole filter is also outlined. Based on these results, a simple and reliable algorithm is presented for designing the one-pole filter. A technique for the robust estimation of beating and two-stage decay parameters was also presented.

Most of the novelties of the thesis can be applied not only for modeling the piano, but for the synthesis of other musical instruments as well. Some ideas (e.g., about filter design) may even find applications in other fields of digital signal processing.

7.2 The future

The piano model described in the thesis gives good results.[1] Nevertheless, there is still much to do in making the model more efficient and the sonic outcome more realistic.

Physical damper models have to be implemented. For their calibration, new measurements are needed. This would be useful also from the analytical point of view, since the effect of dampers in real pianos has not been discussed in the literature.

For the dispersion filter of the digital waveguide, a better approximation should be found, because this is computationally the heaviest part of the simulation of low piano strings. It would be even better to develop a reliable filter design algorithm, which could design the reflection filter as a whole. This could lead to more efficient implementation than separating the different parts of the reflection filter. Experiments have to be made, whether the nonlinear effects of piano strings should be simulated.

The data of soundboard measurements used in this study were noisy, and therefore the calibration of the soundboard model could not be accurately done. New measurements have been already made, but they still have to be processed. It would be also worth to experiment with loss filters exhibiting different decay times for the different delay lines in the feedback delay network. Using other type of shaping filters than simple FIR filters would lead to more efficient simulation of the piano soundboard. The soundboard model could be made more realistic by modeling the coupled vibrations of the undamped highest strings. An efficient method for simulating the effect of sustain pedal also has to be developed.

A part of the plans is to implement the piano model in a real-time environment. It would be beneficial in many ways: it could allow the efficiency evaluation of different parts of the model, and experimenting with the model parameters would be more convenient that way. However, the real motivation is much more personal. The author wishes to play his own instrument...

[1] Sound examples and the electronic version of this thesis are available at:
`http://www.acoustics.hut.fi/~bbank`

Bibliography

Adrien, J.-M. 1991. The Missing Link: Modal Synthesis, *in:* G. D. Poli, A. Piccialli and C. Roads (eds), *Representations of Musical Signals*, The MIT Press, Cambridge, Massachusetts, USA, pp. 269–297.

Allen, J. B. and Rabiner, L. R. 1977. A Unified Approach to Short-Time Fourier Analysis and Synthesis, *Proceedings of IEEE*, pp. 1558–1564.

Arfib, D. 1979. Digital Synthesis of Complex Spectra by means of Multiplication of Nonlinear Distorted Sine Waves, *Journal of the Audio Engineering Society* **27**(10): 757–768.

Askenfelt, A. and Jansson, E. V. 1990. From Touch to Vibrations. I: Timing in the Grand Piano Action, *Journal of the Acoustical Society of America* **88**(1): 52–63.

Askenfelt, A. and Jansson, E. V. 1991. From Touch to Vibrations. II: The Motion of the Key and the Hammer, *Journal of the Acoustical Society of America* **90**(5): 2383–2393.

Askenfelt, A. and Jansson, E. V. 1993. From Touch to Vibrations. III: String Motion and Spectra, *Journal of the Acoustical Society of America* **93**(4): 2181–2196.

Bank, B. 2000. Nonlinear Interaction in the Digital Waveguide with the Application to Piano Sound Synthesis, *Proceedings of the International Computer Music Conference*, Berlin, Germany.

Bank, B., Välimäki, V., Sujbert, L. and Karjalainen, M. 2000. Efficient Physics-Based Sound Synthesis of the Piano Using DSP Methods, *Proceedings of the 10$^{\text{th}}$ European Signal Processing Conference*, Tampere, Finland.

Borin, G. and De Poli, G. 1996. A Hysteretic Hammer-String Interaction Model for Physical Model Synthesis, *Proceedings of the Nordic Acouctical Meeting*, Helsinki, Finland, pp. 399–406.

Borin, G., De Poli, G. and Sarti, A. 1992. Sound Synthesis by Dynamic System Interaction, *in:* D. Baggi (ed.), *Readings in Computer Generated Music*, IEEE Computer Society Press, Los Alamitos, USA, pp. 139–160.

Borin, G., Rocchesso, D. and Scalcon, F. 1997. A Physical Piano Model for Music Performance, *Proceedings of the International Computer Music Conference*, Thessaloniki, Greece, pp. 350–353.

Boutillon, X. 1988. Model for Piano Hammers: Experimental Determination and Digital Simulation, *Journal of the Acoustical Society of America* **83**(2): 746–754.

Chaigne, A. and Askenfelt, A. 1994a. Numerical Simulations of Piano Strings. I. A Physical Model for a Struck String Using Finite Difference Methods, *Journal of the Acoustical Society of America* **95**(2): 1112–1118.

Chaigne, A. and Askenfelt, A. 1994b. Numerical Simulations of Piano Strings. II. Comparisons with Measurements and Systematic Exploration of some Hammer-String Parameters, *Journal of the Acoustical Society of America* **95**(3): 1631–1640.

Chowning, J. M. 1973. The Synthesis of Complex Audio Spectra by Means of Frequency Modulation, *Journal of the Audio Engineering Society* **21**(7): 526–534. Reprinted in C. Roads and J. Strawn, eds. 1985. *Foundations of Computer Music.* Cambridge, Massachusetts: The MIT Press. pp. 6-29.

Conklin, H. A. 1996a. Design and Tone in the Mechanoacoustic Piano. Part I. Piano Hammers and Tonal Effects, *Journal of the Acoustical Society of America* **99**(6): 3286–3296.

Conklin, H. A. 1996b. Design and Tone in the Mechanoacoustic Piano. Part II. Piano Structure, *Journal of the Acoustical Society of America* **100**(2): 695–708.

Conklin, H. A. 1996c. Design and Tone in the Mechanoacoustic Piano. Part III. Piano Strings and Scale Design, *Journal of the Acoustical Society of America* **100**(3): 1286–1298.

Conklin, H. A. 1999. Generation of Partials due to Nonlinear Mixing in a Stringed Instrument, *Journal of the Acoustical Society of America* **105**(1): 536–545.

Daudet, L., Guillemain, P. and Kronland-Martinet, R. 1999. Resynthesis of Piano Strings Vibrations Based on Physical Modeling, *Proceedings of the International Computer Music Conference*, Beijing, China, pp. 48–51.

Ding, Y. and Qian, X. 1997. Processing of Musical Tones Using a Combined Quadratic Polynomial-Phase Sinusoid and Residual (QUASAR) Signal Model, *Journal of the Audio Engineering Society* **45**(7/8): 571–584.

Erkut, C., Välimäki, V., Karjalainen, M. and Laurson, M. 2000. Excitation of Physical and Expressive Parameters for Model-Based Sound Synthesis of the Classical Guitar, *Proceedings of the 108$^{\text{th}}$ AES convention, Preprint No. 5114*, Paris, France.

Feldtkeller, R. and Zwicker, E. 1956. *Das Ohr als Nachrichtenempfänger*, S. Hirzel Verlag, Stuttgart, Germany, p. 91.

Fletcher, H., Blackham, E. D. and Stratton, R. 1962. Quality of Piano Tones, *Journal of the Acoustical Society of America* **34**(6): 749–761.

Fletcher, N. H. and Rossing, T. D. 1998. *The Physics of Musical Instruments*, Springer-Verlag, New York, USA, p. 756.

Garnett, G. E. 1987. Modeling Piano Sound Using Digital Waveguide Filtering Techniques, *Proceedings of the International Computer Music Conference*, Urbana, Illinois, USA, pp. 89–95.

Giordano, N. 1997. Simple Model of a Piano Soundboard, *Journal of the Acoustical Society of America* **102**(2): 1159–1168.

Giordano, N. 1998. Mechanical Impedance of Piano Soundboard, *Journal of the Acoustical Society of America* **103**(4): 2128–2133.

Giordano, N. and Korty, A. J. 1994. Longitudinal Vibrations and the Role of the Bridge, *Journal of the Acoustical Society of America* **96**(3): 1549–1556.

Hall, D. E. 1986. Piano String Excitation in the case of Small Hammer Mass, *Journal of the Acoustical Society of America* **79**(1): 141–147.

Hall, D. E. 1987a. Piano String Excitation II: General Solution for a Hard Narrow Hammer, *Journal of the Acoustical Society of America* **81**(2): 535–546.

Hall, D. E. 1987b. Piano String Excitation III: General Solution for a Soft Narrow Hammer, *Journal of the Acoustical Society of America* **81**(2): 547–555.

Hall, D. E. and Clark, P. 1987. Piano String Excitation IV: The Question of Missing Modes, *Journal of the Acoustical Society of America* **82**(6): 1913–1918.

Hanson, R. J., Anderson, J. M. and Macomber, H. K. 1994. Measurements of Nonlinear Effects in a Driven Vibrating Wire, *Journal of the Acoustical Society of America* **96**(3): 1549–1556.

Hikichi, T. and Osaka, N. 1999. An Approach to Sound Morphing based on Physical Modeling, *Proceedings of the International Computer Music Conference*, Beijing, China, pp. 108–111.

Hiller, L. and Ruiz, P. 1971a. Synthesizing Musical Sounds by Solving the Wave Equation for Vibrating Objects: Part 1, *Journal of the Audio Engineering Society* **19**(6): 462–470.

Hiller, L. and Ruiz, P. 1971b. Synthesizing Musical Sounds by Solving the Wave Equation for Vibrating Objects: Part 2, *Journal of the Audio Engineering Society* **19**(7): 542–550.

Jaffe, D. A. and Smith, J. O. 1983. Extensions of the Karplus-Strong Plucked-String Algorithm, *Computer Music Journal* **7**(2): 56–69. Also published in Roads C. (ed). 1989. *The Music Machine*, pp. 481–494. The MIT Press. Cambridge, Massachusetts, USA.

Jot, J.-M. 1992. An Analysis/Synthesis Approach to Real-Time Artificial Reverberation, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 2, San Francisco, California, USA, pp. 221–224.

Jot, J.-M. and Chaigne, A. 1991. Digital Delay Networks for Designing Artificial Reverberators, *Proceedings of the 90$^{\text{th}}$ AES convention, Preprint No. 3030*, Paris, France.

Järveläinen, H., Välimäki, V. and Karjalainen, M. 1999. Audibility of Inharmonicity in String Instrument Sounds, and Implications to Digital Sound Synthesis, *Proceedings of the International Computer Music Conference*, Beijing, China, pp. 359–362.

Karjalainen, M. and Smith, J. O. 1996. Body Modeling Techniques for String Instrument Synthesis, *Proceedings of the International Computer Music Conference*, Hong Kong, pp. 232–239.

Karjalainen, M. and Välimäki, V. 1993. Model-Based Analysis/Synthesis of the Acoustic Guitar, *Proceedings of the Stockholm Music Acoustic Conference*, Stockholm, Sweden, pp. 443–447.

Karjalainen, M., Välimäki, V. and Tolonen, T. 1998. Plucked-String Models: from Karplus-Strong Algorithm to Digital Waveguides and Beyond, *Computer Music Journal* **22**(3): 17–32.

Karplus, K. and Strong, A. 1983. Digital Synthesis of Plucked-String and Drum Timbres, *Computer Music Journal* **7**(2): 43–55. Also published in Roads C. (ed). 1989. *The Music Machine*. pp.467-479. The MIT Press. Cambridge, Massachusetts.

Laakso, T. I., Välimäki, V., Karjalainen, M. and Laine, U. K. 1996. Splitting the Unit Delay—Tools for Fractional Delay Filter Design, *IEEE Signal Processing Magazine* **13**(1): 30–60.

Lang, M. and Laakso, T. I. 1994. Simple and Robust Method for the Design of Allpass Filters Using Least-Squares Phase Error Criterion, *IEEE Transactions on Circuits and Systems–II: Analog and Digital Signal Processing* **41**(1): 40–48.

Laroche, J. and Jot, J.-M. 1992. Analysis/Synthesis of Quasi-Harmonic Sounds by Use of the Karplus-Strong Algorithm, *Proceedings of the 2$^{\text{nd}}$ French Congress on Acoustics*, France.

Laroche, J. and Meillier, J.-L. 1994. Multichannel Excitation/Filter Modeling of Percussive Sounds with Application to the Piano, *IEEE Transactions on Speech and Audio Processing* **2**(2): 329–344.

Le Brun, M. 1979. Digital waveshaping synthesis, *Journal of the Audio Engineering Society* **27**(4): 250–266.

Lee, K. and Horner, A. 1999. Modeling Piano Tones with Group Synthesis, *Journal of the Audio Engineering Society* **47**(3): 101–111.

Legge, K. A. and Fletcher, N. H. 1984. Nonlinear Generation of Missing Modes on a Vibrating String, *Journal of the Acoustical Society of America* **76**(1): 5–12.

MAT 1996. MATLAB 5 Manual. Mathworks, Inc. Natick, Massachusetts, USA.

McIntyre, M. E., Schumacher, R. T. and Woodhouse, J. 1983. On the Oscillations of Musical Instruments, *Journal of the Acoustical Society of America* **74**(5): 1325–1345.

Moorer, J. A. 1977. Signal Processing Aspects of Computer Music: A Survey, *Proceedings of IEEE*, pp. 1108–1137.

Morse, P. M. 1948. *Vibration and Sound*, McGraw-Hill, p. 468. Reprint, (1st ed. 1936).

Oppenheim, A. V. and Schafer, R. W. 1975. *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, USA, p. 585.

Podlesak, M. and Lee, A. R. 1988. Dispersion of Waves in Piano Strings, *Journal of the Acoustical Society of America* **83**(1): 305–317.

Rabiner, L. R. 1977. On the Use of Autocorrelation Analysis for Pitch Detection, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **25**: 24–33.

Roads, C. 1995. *The Computer Music Tutorial*, The MIT Press, Cambridge, Massachusetts, USA, p. 1234.

Rocchesso, D. 1995. The Ball within the Box: A Sound-Processing Methaphor, *Computer Music Journal* **19**(4): 47–57.

Rocchesso, D. and Scalcon, F. 1996. Accurate Dispersion Simulation for Piano Strings, *Proceedings of the Nordic Acoustical Meeting*, Helsinki, Finland, pp. 407–414.

Rocchesso, D. and Scalcon, F. 1999. Bandwith of Perceived Inharmonicity for Physical Modeling of Dispersive Strings, *IEEE Transactions on Speech and Audio Processing* **7**(5): 597–601.

Rocchesso, D. and Smith, J. O. 1997. Circulant and Elliptic Feedback Delay Networks for Artificial Reverbiration, *IEEE Transactions on Speech and Audio Processing* **5**(1): 51–63.

Rossing, T. D. 1990. *The Science of Sound*, Addison-Wesley, USA.

Schafer, R. W. and Rabiner, L. R. 1973. A Digital Signal Processing Approach to Interpolation, *Proceedings of the IEEE* **61**(6): 692–702.

Serra, X. and Smith, J. O. 1990. Spectral Modeling Synthesis: a Sound Analysis/Synthesis System Based on Deterministic plus Stochastic Decomposition, *Computer Music Journal* **14**(4): 12–24.

Smith, J. O. 1983. *Techniques for Digital Filter Design and System Identification with Application to the Violin*, PhD thesis, Stanford University, California, USA, p. 260.

Smith, J. O. 1987. Music Applications of Digital Waveguides, *Technical Report STAN-M-39*, CCRMA, Dept. of Music, Stanford University, California, USA, p. 181.

Smith, J. O. 1991. Viewpoints on the History of Digital Synthesis, *Proceedings of the International Computer Music Conference*, MOntreal, Canada, pp. 1–10.

Smith, J. O. 1992. Physical Modeling using Digital Waveguides, *Computer Music Journal* **16**(4): 74–91. URL: `http://www-ccrma.stanford.edu/~jos/wg.html`.

Smith, J. O. 1993. Efficient Synthesis of Stringed Musical Instruments, *Proceedings of the International Computer Music Conference*, Tokyo, Japan, pp. 64–71. URL: `http://www-ccrma.stanford.edu/~jos/cs/cs.html`.

Smith, J. O. and Van Duyne, S. A. 1995. Commuted Piano Synthesis, *Proceedings of the International Computer Music Conference*, Banff, Canada, pp. 335–342. URL: `http://www-ccrma.stanford.edu/~jos/cs.html`.

Stulov, A. 1995. Hysteretic Model of the Grand Piano Felt, *Journal of the Acoustical Society of America* **97**(4): 2577–2585.

Tolonen, T. and Järveläinen, H. 2000. Perceptual Study of Decay Parameters in Plucked String Synthesis, *Proceedings of the 109th AES convention.*

Tolonen, T., Välimäki, V. and Karjalainen, M. 1998. Evaluation of Modern Sound Synthesis Methods, *Technical Report 48*, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland, p. 114. URL: `http://www.acoustics.hut.fi/~ttolonen/sound_synth_report.html`.

Van Duyne, S. A. 1992. Low Piano Tones: Modeling Nearly Harmonic Spectra with Regions of FM, *Proceedings of the International Computer Music Conference*, San Hose, USA, pp. 427–428.

Van Duyne, S. A. and Smith, J. O. 1994. A Simplified Approach to Modeling Dispersion Caused by Stiffness in Strings and Plates, *Proceedings of the International Computer Music Conference*, Aarhus, Denmark, pp. 407–410.

Van Duyne, S. A. and Smith, J. O. 1995. Developments for the Commuted Piano, *Proceedings of the International Computer Music Conference*, Banff, Canada, pp. 319–326. URL: `http://www-ccrma.stanford.edu/~jos/cs.html`.

Van Duyne, S. A., Pierce, J. R. and Smith, J. O. 1994. Traveling Wave Implementation of a Lossless Mode-Coupling Filter and the Wave Digital Hammer, *Proceedings of the International Computer Music Conference*, Aarhus, Denmark, pp. 411–418.

Verma, T. S. and Meng, T. H. Y. 1995. An Analysis/Synthesis Tool for Transient Signals that Allows a Flexible Sines+Transients+Noise Model for Audio, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Seattle, WA, USA, pp. 12–15. URL: `http://www.acoustics.hut.fi/~tverma/publications/index.html`.

Välimäki, V. 1995. *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*, PhD thesis, Helsinki University of Technology, Espoo, Finland, p. 193. URL: `http://www.acoustics.hut.fi/~vpv/publications/vesa_phd.html`.

Välimäki, V. and Laakso, T. I. 1998. Suppression of Transients in Variable Recursive Digital Filters with a Novel and Efficient Cancellation Method, *IEEE Transactions on Signal Processing* **46**(12): 3408–3414.

Välimäki, V. and Tolonen, T. 1998. Development and Calibration of a Guitar Synthesizer, *Journal of the Audio Engineering Society* **46**(9): 766–778.

Välimäki, V., Huopaniemi, J., Karjalainen, M. and Jánosy, Z. 1996. Physical Modeling of Plucked String Instruments with Application to Real-Time Sound Synthesis, *Journal of the Audio Engineering Society* **44**(5): 331–353.

Välimäki, V., Karjalainen, M. and Laakso, T. I. 1993. Fractional Delay Digital Filters, *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 1, Chicago, USA, pp. 355–358.

Weinreich, G. 1977. Coupled Piano Strings, *Journal of the Acoustical Society of America* **62**(6): 1474–1484.

Yuen, J. and Horner, A. 1997. Hybrid Sampling-Wavetable Synthesis with Genetic Algorithms, *Journal of the Audio Engineering Society* **45**(5): 316–331.

Zheng, H. 1999. *Analysis and Critical-Band-Based Group Synthesis of Piano Tones*, Master's thesis, University of Illinois, Urbana, USA, p. 70. URL: `http://pierre.music.uiuc.edu/~zheng/G/work/index.html`.

Zhu, W. D. and Mote, C. D. 1994. Dynamics of the Pianoforte String and Narrow Hammers, *Journal of the Acoustical Society of America* **96**(4): 1999–2007.

Appendix

A.1 The secrets of the one-pole loss filter

The one-pole loss filter has been found to be a good approximation for many string instruments [Välimäki et al. 1996; Välimäki and Tolonen 1998]. The pattern of the decay times, which arises when one uses such a filter, matches the decay of a real string perceptually well. The author was motivated by this fact to find the connection between the digital waveguide with a one-pole loss filter and the lossy string. The magnitude response of the one-pole filter $H_{1p}(z)$ of Eq. (6.2) is:

$$|H_{1p}(e^{j\vartheta})| = g\frac{1+a_1}{\sqrt{a_1^2+1+2a_1\cos\vartheta}} \tag{A.1}$$

The decay time τ of the digital waveguide with such a filter:

$$\tau = -\frac{1}{f_0 \ln|H_{1p}(e^{j\vartheta})|} \approx \frac{1}{f_0(1-|H_{1p}(e^{j\vartheta})|)} \tag{A.2}$$

where f_0 is the fundamental frequency of the note, and for the ln function the first-order Taylor-series approximation was used. Note that here we apply Eq. (4.16) for computing the decay times and neglect the effect of the dispersion filter (this was discussed previously in relation with Eq. (6.8)).

Now we derive the decay rate $\sigma = 1/\tau$, instead of the decay times:

$$\begin{aligned}\sigma \;\approx\; f_0(1-|H_{1p}(e^{j\vartheta})|) &= f_0\left(1-g\frac{1+a_1}{\sqrt{a_1^2+1+2a_1\cos\vartheta}}\right) = \\ &= f_0\frac{\sqrt{a_1^2+1+2a_1\cos\vartheta}-g(1+a_1)}{\sqrt{a_1^2+1+2a_1\cos\vartheta}}\end{aligned} \tag{A.3}$$

using the second-order Taylor-series approximation for the cos function ($\cos x \approx 1-x^2/2$ for $x \approx 0$):

$$\sigma \approx f_0\frac{\sqrt{(a_1+1)^2-a_1\vartheta^2}-g(1+a_1)}{\sqrt{(a_1+1)^2-a_1\vartheta^2}} = f_0\frac{\sqrt{1-\frac{a_1}{(a_1+1)^2}\vartheta^2}-g}{\sqrt{1-\frac{a_1}{(a_1+1)^2}\vartheta^2}} \tag{A.4}$$

since the denominator is close to 1 and $\sqrt{1+x} \approx 1 + x/2$ for $x \approx 0$:

$$\sigma \approx f_0 \left(\sqrt{1 - \frac{a_1}{(a_1+1)^2}\vartheta^2} - g \right) \approx f_0 \left((1-g) - \frac{a_1}{2(a_1+1)^2}\vartheta^2 \right) \tag{A.5}$$

It follows that the decay times of the digital waveguide with a one-pole loss filter will be:

$$\begin{aligned} \tau &= \frac{1}{\sigma} \approx \frac{1}{c_1 + c_3\vartheta^2} \\ c_1 &= f_0(1-g) \\ c_3 &= -f_0 \frac{a_1}{2(a_1+1)^2} \end{aligned} \tag{A.6}$$

which is the decay time of a string with the simplest frequency dependent losses, where the c_1 and c_3 coefficients correspond to the first- and third-order time derivates of the wave equation [see Chaigne and Askenfelt 1994a]. The approximation is very accurate for $g = 1 - \epsilon_g$ and $a_1 = -\epsilon_a$, where ϵ_g and ϵ_a are small positive numbers. This holds for loop filters used in practice. Example values are available, e.g., in [Välimäki and Tolonen 1998].

The same derivation can be done for the first-order FIR filter, whose transfer function is:

$$H_{1z}(z) = g \frac{1 + b_1 z^{-1}}{1 + b_1} \tag{A.7}$$

where $-b_1$ is the zero of the filter and g is the DC gain. If $|g| < 1$ and $b_1 > 0$, then the filter will be of a lowpass character and the stability of the digital waveguide loop is guaranteed.

The decay times of the digital waveguide with a first-order FIR filter:

$$\begin{aligned} \tau &= \frac{1}{\sigma} \approx \frac{1}{c_1 + c_3\vartheta^2} \\ c_1 &= f_0(1-g) \\ c_3 &= f_0 \frac{b_1}{2(b_1+1)^2} \end{aligned} \tag{A.8}$$

By comparing Eqs. (A.6) and (A.8), the b_1 parameter of the FIR filter can be calculated from the a_1 coefficient of the one-pole filter in the following way:

$$b_1 = -\frac{(a_1+1)^2 + 2a_1 - (a_1+1)\sqrt{(a_1+1)^2 + 4a_1}}{2a_1} \tag{A.9}$$

The expression under the square-root has to be nonnegative ($(a_1+1)^2 + 4a_1 \geq 0$). This gives a constraint for the realizable FIR filter: Eq. (A.9) is valid only for $a_1 \geq -3 + 2\sqrt{2} \approx -0.1716$. For $a_1 = -3 + 2\sqrt{2}$ the b_1 coefficient of the FIR filter equals to 1, which corresponds to the first-order moving average filter, also used in the Karplus-Strong algorithm [Karplus and Strong 1983]. This is the maximum

lowpass filtering that can be achieved by a first-order FIR filter. The corresponding coefficient in Eq. (A.8) will be $c_3 \leq f_0/8$. For the one-pole filter there is no such constraint.

Figure A.1 solid line shows the decay times of a digital waveguide with a one-pole filter. The parameters are: $f_0 = 500$ Hz, $f_s = 44.1$ kHz, $g = 0.999$, and $a_1 = -0.05$. The dash-dotted line illustrates the decay times calculated by the approximation of Eq. (A.6), where $c_1 = 0.5$ and $c_3 = 13.85$. The dashed line represents the decay times of the equivalent FIR filter implementation, using Eq. (A.9), where $b_1 = 0.0625$. It can be seen from the figure that the polynomial approximation for the decay rates is very accurate. So is the one-zero filter approximation. Whenever the FIR filter is realizable, it is equivalent to the one-pole filter concerning the decay times.

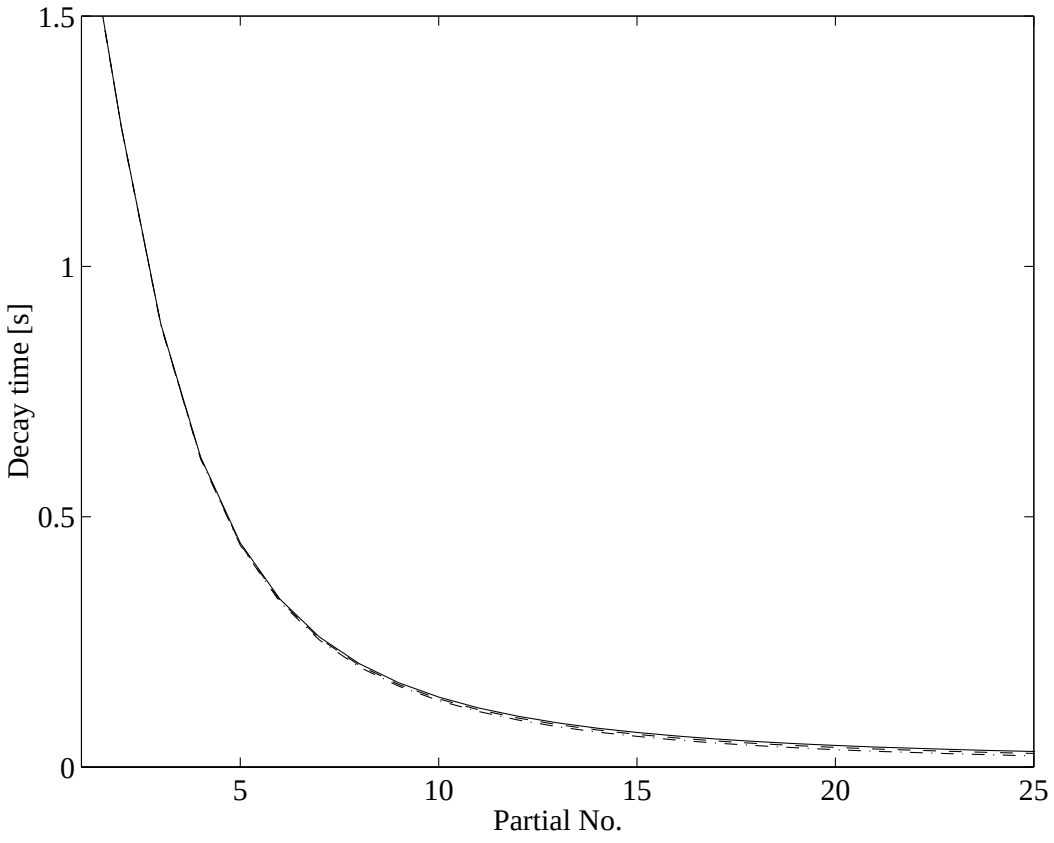

Figure A.1: Decay times of the digital waveguide with a one-pole filter (solid line), the decay times approximated by Eq. (A.6)(dash-dotted line), and the decay times of the equivalent one-zero filter according to Eq. (A.9)(dashed line).

To sum up, the one-pole and a one-zero filter behave very similarly in terms of the resulting decay times, the difference is that there is a limit for the FIR filter for what kind of decay times can be realized. This means that the one-pole filter is more flexible, it is more useful for our purposes. The advantage of the FIR filter comes into play when changing the loop filter coefficients during playing. This can

be the case for the linear damper model in the piano or, more importantly, some other effects in other instruments, e.g., in the guitar or in the violin. However, for those c_3 values, which can be still implemented as a FIR filter, the a_1 parameter of the one-pole filter is small, and so is the transient of the filter when changing coefficients. Consequently, there is no real need for the one-zero implementation.

A.2 Measurement of the piano

Measurement characteristics

- The grand piano: August Förster Mod. 170. SN 135178.
- Place: BUTE Schönherz Zoltán dormitory, Budapest, Hungary
- Date: July 21–23, 1999

Measurement equipment

description	*type*	*sensitivity* (EU/V)
accelerometer No. 1	PCB 353B13	1966
accelerometer No. 2	PCB 353B13	1869
microphone	TMS13P10/TMS130B10	32.01
impact hammer	PCB 086C03	441
8 channel amplifier	PCB F482A20	
stereo soundcard	Gravis Ultrasound P&P	
personal computer		

Sound card

frequency response	20 Hz – 20 kHz (−3 dB)
crosstalk	≤ -50 dB
harmonic distortion (THD)	≤ 0.01
signal-to-noise ratio (SNR)	≥ 75 dB
nominal input level	-5 dBV$_{\mathrm{rms}}$

Description of the measurements

The measurement equipment provided the opportunity of two channel measurements. The transducers were connected to the input of the multichannel amplifier. The amplified signals were recorded by a personal computer with a sound card in *wav* format. The gain of the amplifier was set to 10 in all the cases, and the accelerometer No. 1 was always used. The only exception was the measurement of hammer acceleration. Then, accelerometer No. 2 was employed, and the corresponding gain factor of the amplifier was set to unity.

Accelerometer No. 1 was attached to the bridge by wax. Accelerometer No. 2 was fixed on the hammer by the help of thin wire.

The microphone was placed above the string F_4 at a distance 12 cm from the soundboard, and 24 cm from the tuning pins.

The bridge was hit by the impact hammer next to the accelerometer.

Performed measurements

- Examination of the bridge and the soundboard by hitting the bridge with the impulse hammer

  - without sustain pedal (strings damped):

measured parameters:	force at the bridge – bridge acceleration
	force at the bridge – sound pressure
	bridge acceleration – sound pressure
measured notes:	C_1, G_1, C_2, $F_2^\sharp$, $C_3^\sharp$, $F_3^\sharp$, $A_4^\sharp$, $G_5^\sharp$, $F_6^\sharp$, $F_7^\sharp$
number of measurements/note:	20

  - with sustain pedal (strings undamped):

measured parameters:	force at the bridge – bridge acceleration
	force at the bridge – sound pressure
	bridge acceleration – sound pressure
measured notes:	C_2, $F_3^\sharp$, $A_4^\sharp$, $F_6^\sharp$
number of measurements/note:	20

- Examination of single string vibration with piano hammer excitation (the other strings of the doublets and triplets were damped with felt and the una corda pedal was depressed):

  - without sustain pedal:

measured parameters:	bridge acceleration – sound pressure
measured notes:	C_1, G_1, C_2, $F_2^\sharp$, $C_3^\sharp$, $F_3^\sharp$, $A_4^\sharp$, $G_5^\sharp$, $F_6^\sharp$, $F_7^\sharp$
number of measurements/note:	forte: 5, piano: 5

  - with sustain pedal:

measured parameters:	bridge acceleration – sound pressure
measured notes:	C_2, $F_3^\sharp$, $A_4^\sharp$, $F_6^\sharp$
number of measurements/note:	forte: 3

- Movement of the hammer

  - without sustain pedal:

measured parameters:	hammer acceleration – bridge acceleration
measured notes:	C_1, G_1, C_2, $F_2^\sharp$, $C_3^\sharp$, $F_3^\sharp$, C_4, $A_4^\sharp$, $G_5^\sharp$, $F_6^\sharp$
number of measurements/note:	legato piano: 2, mezzoforte: 2 forte: 2
	staccato piano: 2, mezzoforte: 2, forte: 2

- Sound of the piano (in natural circumstances)

  - without sustain pedal:

measured parameter:	sound pressure
measured notes:	C_1, G_1, C_2, $F_2^\sharp$, $C_3^\sharp$, $F_3^\sharp$, C_4, $A_4^\sharp$, $G_5^\sharp$, $F_6^\sharp$, $F_7^\sharp$
number of measurements/note:	legato piano: 2, mezzoforte: 2, forte: 2
	staccato piano: 2, mezzoforte: 2, forte: 2

– with sustain pedal:

measured parameter:	sound pressure
measured notes:	C_1, G_1, C_2, $F_2^\sharp$, $C_3^\sharp$, $F_3^\sharp$, C_4, $A_4^\sharp$, $G_5^\sharp$, $F_6^\sharp$, $F_7^\sharp$
number of measurements/note:	forte: 1