

Ph.D. thesis

Balázs Bank

Physics-based Sound Synthesis of String Instruments Including Geometric Nonlinearities

Ph.D. thesis

Balázs Bank
M.Sc.E.E.

Supervisor:

László Sujbert
Ph.D.



Budapest University of Technology and Economics
Department of Measurement and Information Systems
2006

© 2006 Balázs Bank

Budapest University of Technology and Economics
Department of Measurement and Information Systems
H-1117 Budapest, XI. Magyar tudósok körútja 2.
Homepage: <http://www.mit.bme.hu/~bank>
Email: bank@mit.bme.hu

This thesis and the corresponding sound examples can be downloaded from
<http://www.mit.bme.hu/~bank/phd>.

Elli néni

I dedicate this thesis to my piano teacher, Mrs. Schichtanz, or “Elli néni”, as her students called her. She left us in the days when I finished this thesis. With the experience of her more than one hundred year long, fulfilling life, she taught us more than piano playing. Her willpower and good humor should be an example to follow for all of us.

Nyilatkozat (Declaration of Authorship)

Alulírott Bank Balázs kijelentem, hogy ezt a doktori értekezést magam készítettem és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

A dolgozat bírálatai és a védésről készült jegyzőkönyv a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karának dékáni hivatalában elérhető.

Budapest, 2006. február 14.

Bank Balázs

Preface

This work is the summary of my research about physics-based sound synthesis of string instruments. Physics-based sound synthesis is a promising approach, which has several advantages over the commercially most often used sampling synthesizers. However, due to its significantly higher computational complexity and its lack of ability to model the important nuances of the instruments, the physics-based technique has only marginal commercial applications. My goal was to develop such methods that increase the competitiveness of the approach by either providing better sound quality or decreasing the computational cost compared to earlier physics-based methods. Most of the results have been developed with the application to the piano, but they can be directly used in the synthesis of other struck or plucked instruments, too.

The first part of the thesis concentrates on the development of parameter estimation techniques for the digital waveguide, the most often used string modeling technique, providing more accurate control over the decay of partials. This is followed by the development of multi-rate techniques that increase the efficiency of modeling the excitation, the beating and two-stage decay, and the filtering effect of the instrument body. However, certain important features of the sound, e.g., phantom partials, cannot be reproduced by linear string models. Therefore, the second part of the thesis concentrates on the development of nonlinear string models that are able to capture the characteristic subtleties of the tone generated by the nonlinear coupling of the transverse and longitudinal polarizations. For forming the basis of sound synthesis algorithms, a theoretical framework on the generation of longitudinal vibration is also developed. Thus, the structure of the thesis follows the progress of my research by proceeding from simpler linear models to more complicated nonlinear techniques.

Most of this work has been carried out at the Department of Measurement and Information Systems, Budapest University of Technology and Economics, Hungary. First of all, I wish to thank my supervisor, Dr. László Sujbert for his continuous help during our seven year long scientific relationship, and Prof. Gábor Péceli, the head of the department, for supporting my work. I am grateful for the support of all the other colleagues, especially for that of Dr. János Márkus, who has helped in the preparation of this thesis and in the organization of my Ph.D. defense.

The research work has been started in the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology (Finland), where I wrote my M.Sc. thesis during the academic year 1999-2000. I am grateful to Prof. Vesa Välimäki and Prof. Matti

Karjalainen for supervising my work and for providing me the opportunity to conduct research at their laboratory. Some parts of this research have been carried out at the Department of Electronics and Informatics, Padua University (Italy), during the period Nov. 2001–Apr. 2002. I wish to thank my Italian supervisor, Prof. Giovanni De Poli for his support. I'm grateful to all the other colleagues at Helsinki and Padua for their help and for the interesting scientific discussions I had during these visits.

I am much obliged to Dr. Cumhur Erkut, Dr. Federico Fontana, Péter Hussami, Dr. János Márkus, Jyri Pakarinen, Prof. Davide Rocchesso, and Prof. Vesa Välimäki for their detailed comments on my manuscript.

I am grateful to the first reviewers of my thesis, Dr. Fülöp Augusztinovicz and Prof. Julius O. Smith, for their supporting critics, which have improved the quality of the thesis significantly.

I am grateful to my family for supporting me in every aspects of my work and beyond. I'm especially grateful to my wife, Rita, for her assistance and for doing all the dishwashing during these busy days.

Budapest, February 14, 2006

Balázs Bank

Contents

List of Symbols	xiii
1 Introduction	1
1.1 Physics-based Modeling as a Sound Synthesis Technique	1
1.1.1 Signal-based Approach	2
1.1.2 Physics-based Approach	3
1.1.3 Instruments as Case Studies	4
1.2 The Benefit of Physics-based Sound Synthesis	4
1.3 Structure of the Thesis	5
2 Physical Modeling of String Instruments	7
2.1 Model Structure	7
2.2 String Equations	8
2.2.1 General Equations of String Vibration	8
2.2.2 Approximate Nonlinear Equations	10
2.2.3 Linear Equations	11
2.2.4 The Stiff and Lossy String	13
2.3 String Modeling Techniques	16
2.3.1 Finite-difference Modeling	16
2.3.2 Digital Waveguide Modeling	20
2.3.3 Modal-based Approach	27
2.4 Excitation Modeling	29
2.4.1 Struck Strings	29
2.4.2 Plucked Strings	31
2.4.3 Bowed Strings	32
2.5 Instrument Body Modeling	33
2.5.1 Physics-based Modeling	33
2.5.2 Post-processing Techniques	34
2.5.3 Commuted Synthesis	36
2.6 Conclusion	37
3 Loss Filter Design for the Digital Waveguide	39
3.1 The One-pole Loss Filter	39

3.1.1	Approximate Formulas for the Decay Times	39
3.1.2	Filter Design Based on Polynomial Regression	41
3.2	High-order Loss Filters	44
3.2.1	Weighting Function Based on the Taylor Series Approximation of Decay Times	45
3.3	Conclusion	47
4	Multi-rate Instrument Modeling Techniques	49
4.1	Excitation Modeling	49
4.1.1	The Multi-rate Excitation Model	51
4.2	Modeling Beating and Two-stage Decay	54
4.2.1	The Parallel Resonator Bank	54
4.3	Instrument Body Modeling	58
4.3.1	FIR Filters	59
4.3.2	IIR Filters	59
4.3.3	Multi-rate Body Modeling	60
4.4	Conclusion	63
5	Modeling of Geometric Nonlinearities	65
5.1	Classification of Nonlinear String Behavior	65
5.1.1	String Equations	66
5.1.2	Classification	67
5.1.3	Short Discussion of the Regimes of Nonlinearity	70
5.2	Spatially Uniform Tension	72
5.2.1	Theories and Experiments	72
5.2.2	Nonlinear Generation of Missing Modes	74
5.3	Modeling of Longitudinal Modes	76
5.3.1	Prior Work	76
5.3.2	Equations of Motion	77
5.3.3	Longitudinal Motion in the Case of Exponentially Decaying Trans- verse Modes	79
5.3.4	String Tension	82
5.3.5	Longitudinal Bridge Force	88
5.3.6	Possible Extensions and Limitations	88
5.3.7	Connections to Measurements	91
5.4	Modeling of Bidirectional Coupling	95
5.4.1	Equations of Motion	95
5.4.2	The Stabilization Effect	98
5.5	Conclusion	102
6	Sound Synthesis of Geometric Nonlinearities	105
6.1	Double Frequency Terms	105
6.2	Tension Modulation	106

6.3	Modeling of Longitudinal Vibrations for Sound Synthesis	106
6.3.1	Methods Proposed by Other Researchers	107
6.3.2	Finite-difference Modeling	108
6.3.3	Tension Decomposition	108
6.3.4	The Composite String Model	109
6.3.5	The Resonator-based String Model	111
6.3.6	Digital Waveguide and Longitudinal Resonators	113
6.3.7	Physically Informed Modeling Techniques	114
6.3.8	Comparison	119
6.4	Modeling of Bidirectional Coupling for Sound Synthesis	119
6.4.1	Finite-difference Modeling	120
6.4.2	Tension Decomposition	120
6.5	Conclusion	122
7	Summary	125
7.1	Further Research	126
7.2	New Scientific Results	126
	Bibliography	129
	Appendix	137
A.1	Parameter Dependence of Nonlinear Components	137
A.2	The Effect of Nonrigid String Terminations	140

List of symbols

a_k	real coefficient
A	amplitude
b_k	real coefficient
B	inharmonic coefficient
c_t, c_l	transverse (y) and longitudinal (ξ) wave velocity
c_k	real coefficient
C	complex coefficient
$d_y(x, t), d_\xi(x, t)$	excitation force densities of the string in the y and ξ polarization
D	phase delay
e	approximation error
E	Young's modulus
$\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$	unit vectors in the $x, y,$ and z direction
\mathbf{e}_s	unit vector pointing along the tangent of the string
f	frequency
$f_0, f_{\xi,0}$	transverse and longitudinal fundamental frequency
f_s	sampling frequency
F	force
$F_{y,k}(t), F_{\xi,k}(t)$	excitation forces of mode k in the y and ξ polarization
g	DC gain of the one-pole loss filter
g_k	magnitude specification of the loss filter
$H(z)$	transfer function in z-domain
I_k	inharmonic index
j	imaginary unit
K_h	hammer stiffness coefficient
L	length of the string
m_h	hammer mass
P_h	hammer stiffness exponent
p_k	poles of the resonators
r_F	reflection coefficient for force waves
r_v	reflection coefficient for velocity waves
$R(\omega)$	frictional resistance of the string
$R(z)$	transfer function of the resonator

S	cross section area of the string
t	time
t_n	time instant of the time-domain sampling step n ($t_n = n\Delta t$)
T	string tension
T_0	initial tension
$\bar{T}(t)$	spatially uniform part of tension
$\tilde{T}(x, t)$	space-dependent part of tension
T_p	time period
Δt	sampling period
v	velocity
x	position along the string
x_m	position of the spatial sampling step m ($x_m = m\Delta x$)
Δx	spatial sampling interval
$y(x, t)$	transverse displacement of the string
$y_{m,n}$	transverse displacement of the string ($y_{m,n} = y(x_m, t_n)$)
$y_k(t)$	instantaneous amplitude of transverse mode k
$y_{\delta,k}(t)$	impulse response of transverse mode k
Δy	compression of the hammer felt
z	z-transform variable
z^{-1}	unit delay
$z(x, t)$	transverse displacement of the string in the z direction
Z	impedance
Z_0	characteristic impedance of the string
β	wave number
φ	phase in radians
κ	radius of gyration
μ	mass density
σ	decay rate
ϑ	angular frequency in radians
τ	decay time
$\xi(x, t)$	longitudinal displacement of the string
$\xi_k(t)$	instantaneous amplitude of longitudinal mode k
$\bar{\xi}_k(t), \tilde{\xi}_k(t)$	amplitudes of long. mode k coming from the constant and dynamic response
$\xi_{\delta,k}(t)$	impulse response of longitudinal mode k
$\bar{\xi}_{\delta,k}(t), \tilde{\xi}_{\delta,k}(t)$	constant and dynamic impulse responses of longitudinal mode k
ω	analog angular frequency
Ω	complex angular frequency ($\Omega = \omega + j\sigma$)
$\text{Re}\{C\}$	real part of complex variable C
C^*	conjugate of complex variable C
$ C $	absolute value of real or complex variable C
$\varphi\{C\}$	phase of complex variable C
$a(t) * b(t)$	convolution of the functions $a(t)$ and $b(t)$
$\mathcal{L}\{a(t)\}$	Laplace transform of $a(t)$
$\mathcal{F}\{a(t)\}$	Fourier transform of $a(t)$
$Z\{a_k\}$	z-transform of a_k

Chapter 1

Introduction

Physics-based sound synthesis is on the borderline of acoustics and digital signal processing. Building a physics-based sound synthesis model has two steps: the first is to understand how the real instrument works and to construct a precise model that describes the physical reality. The second step is to implement this model on a computer or a dedicated hardware. This means the temporal and spatial discretization of continuous-time equations. Moreover, often some simplifications are necessary due to the limited computational resources. These simplifications are usually carried out in a way that the perceptually less significant effects are neglected. Naturally, the first step is skipped if the phenomenon that should be modeled is well understood. This often happens as physics and acoustics have a much longer tradition compared to developing efficient signal processing algorithms.

1.1 Physics-based Modeling as a Sound Synthesis Technique

Sound synthesis methods can be classified in many ways. Here we divide them into three groups, by unifying two groups of the classifications found in [Smith 1991; Tolonen et al. 1998].

The first group is the family of abstract methods. These are different algorithms which can easily generate synthetic sounds. Methods like frequency modulation [Chowning 1973] and waveshaping [Le Brun 1979; Arfib 1979] belong to this category. Modeling real instruments with these methods is fairly complicated as the relationship between the parameters of the technique and those of the real instruments cannot be easily formulated. As now the primary goal is to model the sound of acoustic instruments, we do not discuss this group any further.

The second group (signal modeling) is the one which models the sound of the musical instruments. In this case, the input to the model is only the waveform or a set of waveforms generated by the instrument and the physics of the sound generation mechanism is not examined in detail. Synthesis methods like PCM (Pulse Code Modulation) [Roads 1995] and SMS (Spectral Modeling Synthesis) [Serra and Smith 1990] belong to this category. The corresponding groups in the taxonomy of Smith [1991] are processing of pre-recorded samples and spectral models.

The third group (physical modeling) is the one which, instead of reproducing a specific sound of an instrument, models the physical behavior of the instrument. Usually, the physical system (such as a string on a violin or the skin of a drum) can be described with a set of difference equations and transfer functions. Given the excitation of the instrument (bowing, plucking, etc.), the difference equations can be solved (or the general solution can be applied for the given input), and the output of the model is expected to be close to the output of the real instrument. One well-known method in this category is the digital waveguide synthesis [Smith 1992], which efficiently models the vibration of a one-dimensional string, based on the solution of the wave-equation. A comprehensive review on the different physics-based modeling approaches can be found in [Välämäki et al. 2006].

Both of the two latter methods have their own advantages and disadvantages. Signal-based synthesis can be realized efficiently, and the generated sound is usually an accurate model of the response of an instrument for a given excitation. The method can be used if the excitation is nearly constant, or the instrument is linear to a good approximation. Its greatest advantage is the simple parameter estimation and that the same model structure can be used for many different instruments. In contrast, physical modeling is capable of the accurate modeling of nonlinear, transient responses of an instrument. It is able to respond dynamically to different excitations similar to a real instrument. However, usually physical models require more computational resources for real-time implementation. A strong disadvantage is that the parameter estimation of physical models is much more complicated, compared to signal models.

In the next subsections, the signal-based and the physics-based approach is compared, from the point of view of their applicability. The main features of the methods are listed in Table 1.1. Then the most important properties of the some example instruments are described, serving as a basis for the choice among the synthesis methods. We note that an exhaustive evaluation of many different sound synthesis methods can be found in [Tolonen et al. 1998].

1.1.1 Signal-based Approach

The signal-based approach models the sound of the instrument itself. Accordingly, it does not make any assumptions on the structure of the musical instrument, only that the generated sound is periodic. Therefore, it can model a wide range of instrument sounds, since they differ in their parameters only, not in the model structure, which is, e.g., a set of sinusoids for each instrument. As it is a general representation, its parameter estimation is simple, basically reduces to tracking partial envelopes, which can be easily automated. In general, a large amount of data is required to describe a given tone, but this specific tone from which the parameters originate, is almost perfectly reproduced. As the structure of the instrument is not modeled, the interaction of the musician cannot be easily taken into account, meaning that, e.g., for different bow forces or velocities in the case of the violin different parameter sets are required for resynthesis. In practice, this means that for a single note the analysis procedure has to be run for all the different playing styles that a player can produce, and a large amount of data has to be stored or transmitted. As it

Method	Signal modeling	Physical modeling
Assumptions on the structure	Poor	Yes
Generality	Yes	No
Parameter estimation	Simple	Complicated
Nature of parameters	Abstract	Meaningful
Number of parameters	Many	Few
Modeling a specific sound	Precisely	Approximately
Interaction of the musician	Hard to model	Modeled
Interaction of instrument parts	Hard to model	Modeled

Table 1.1: Main features of signal- and physics-based sound synthesis methods.

treats the notes separately, the interaction of the different notes, e.g., the coupled vibration of strings, cannot be modeled. Changing the parameters of the synthesis program directly is not user-friendly: dozens of parameters can be changed, which all influence the sound in a different way compared to musicians got used to it in the case of real instruments. The quality and the computational load of the synthesis is usually varied by changing the number of simulated partials, which is probably not the best way from a perceptual point of view.

1.1.2 Physics-based Approach

The physics-based approach models the functioning of the instrument, rather than the produced sound itself. It makes assumptions about the instrument it models, therefore, it loses generality. A piano model, e.g., cannot be used for violin modeling by just changing its parameters, since the excitation model is completely different for the two instruments. Consequently, the parameter estimation cannot be completely automated, at least the model structure has to be determined by the user. As the model structure already describes the main features of the instrument, only small number of parameters are needed, and modifications to these parameters produce perceptually meaningful results. For example, the user now controls the bow force, rather than the loudness of a single partial, and the instrument reacts in a way as a real violin would do. Therefore, only one parameter set is required, since the different playing styles according to the interaction of the musician are automatically modeled. As it describes the physical structure, the interaction of the different model parts are also taken into account, e.g., the string coupling on the piano is easily modeled. A drawback that none of the tones will be perfectly modeled: the model may sound as a piano, but will be always different from that piano where its parameters come from. The quality and the computational load is varied by, e.g., changing the accuracy of modeling losses and dispersion, rather than changing the number of simulated partials, which is less noticeable for the listener. These characteristics are summarized in Table 1.1.

Instrument	Organ	Piano	Violin
Number of partials	< 20	5-100	10–50
Number of playing parameters	0	Few	Many
Coupling between the instrument parts	Negligible	Present	Significant

Table 1.2: Main features of the different instruments, serving as a basis for choosing the proper synthesis approach.

1.1.3 Instruments as Case Studies

The choice between the signal-based or the physics-based approaches strongly depends on which instrument should be modeled. Here we will use the organ, the piano, and the violin as case studies. The features which are relevant from this viewpoint for these instruments are listed in Table 1.2. Naturally, other factors also influence the choice of the user, e.g., if automatic parameter estimation is required, the signal modeling approach should be chosen.

The sound of a specific organ pipe cannot be influenced by the player. Moreover, the coupling between the different pipes is negligible, therefore the different tones can be synthesized independently. As signal modeling models a specific sound almost perfectly, it is the best choice for organ synthesis. Its computational load is acceptable, since the number of partials is low in the case of the organ flue pipes.

As for the piano, the player can vary only one parameter for a given note, by changing the impact velocity of the hammer, thus, the timbre space of one note is one-dimensional. For a signal model, this would mean storing different parameter sets for a few hammer velocities, and interpolation could be used between sets. Although it is also possible with the signal model, the effect of the player is much easier modeled by the physics-based approach. Moreover, the strings of the piano are coupled when the damper pedal is depressed which is also controlled by the player: this can be modeled by the physics-based approach only.

For the violin, the freedom of the player is enormous: he can vary the bow force, velocity, position, and angle, the finger position and pressure, and decide on which string he plays the given note. Therefore, the timbre space of the violin is multi-dimensional: for signal-based synthesis many sounds along all these dimensions should be recorded and analyzed. Since the goal is not only to render the sound of a specific violin note, but to create a playable instrument, the only choice which remains is physical modeling. The inputs of the physical model are the real physical parameters (e.g., bow force and velocity), therefore the effect of the player is automatically taken into account.

1.2 The Benefit of Physics-based Sound Synthesis

The primary use of developing physics-based sound synthesis algorithms is that they lead to synthesizers with better sound quality. Naturally, synthesizers should not replace real instruments. However, for pianists for whom real pianos are too expensive, too large, or

too loud, practicing on a better digital piano could be very helpful. Better synthesis models could also increase the reality of games and multimedia applications. Here, the physics-based approach can be used as an intermediate coding level between MIDI (that stores the score and the instrument name) and perceptual coders (such as MPEG-1 Layer 3), as it is able to transmit the main features of the instrument by a low number of parameters. This kind of “structured audio coding” is supported by MPEG-4 [Scheier 1999].

Another benefit of physics-based sound synthesis algorithms is that they can be used for experimentation purposes. With a physical model it is possible to include or exclude a specific feature of the sound production mechanism and assess the importance of that phenomenon by listening to the result. Varying the physical parameters of the model could help instrument makers to estimate the sonic consequences of changing the geometry or material of the instrument. Moreover, the acoustic research and theory development triggered by the need of sound synthesis yields a better knowledge of real instruments.

1.3 Structure of the Thesis

Chapter 2 describes the most often used physics-based sound synthesis approaches and provides the necessary theoretical background for the rest of the thesis. In Chap. 3 new parameter estimation techniques are proposed for the loss-filters of digital waveguides (which are the most often used string models). Chapter 4 presents new algorithms to excitation, string, and body modeling applying the multi-rate approach, which result in a significantly lower computational cost compared to earlier methods. Chapter 5 is about the theory of geometric nonlinearities of musical instrument strings, presenting a modal model that not only founds the basis of sound synthesis algorithms but gives a qualitative insight to the physics of the phenomenon, too. In Chap. 6 the theoretical results on the geometric nonlinearities are applied for the development of efficient sound synthesis algorithms. Finally, Chap. 7 summarizes the results of this thesis and outlines the possible research directions.

Chapter 2

Physical Modeling of String Instruments

In this chapter the most often used modeling strategies are reviewed to provide the background for the new results of the following part of the thesis. Note that the notation and the derivations of the equations have been changed compared to the referenced literature, to be in coherence with the rest of the thesis.

Section 2.1 gives an outline of the model structure used in physics-based sound synthesis. Then, Sec. 2.2 provides the theoretical background of string vibrations, starting from the equations of motion. This is followed by the most often used string modeling approaches, namely, finite-difference modeling, digital waveguides and modal models in Sec. 2.3. Section 2.4 describes the modeling methods for the different excitation mechanisms (striking, plucking, and bowing), and finally, Sec. 2.5 summarizes the techniques for instrument body modeling.

2.1 Model Structure

Since the physical modeling approach simulates the sound production mechanism of the instrument, the parts of the model correspond to the parts of real instruments. In every string instrument, the heart of the sound production mechanism is the string itself. The string is excited by the excitation mechanism, which corresponds to the hammer strike in the case of the piano, to the pluck in the case of the guitar, or to the bow in the case of the violin. The string is responsible for the generation of the periodic sound by storing this vibrational energy in its normal modes. One part of this energy dissipates and another part is radiated to the air by the instrument body. The body can be seen as an impedance transformer between the string and the air, which increases the effectiveness of radiation significantly. The body provides a terminating impedance to the string, therefore it also influences the modal parameters of string vibration, i.e., partial frequencies, amplitudes, and decay times. The model structure is displayed in Fig. 2.1.

The modeling of the interaction between the string and the excitation can be both uni- and bidirectional. When, e.g., the plucking of the guitar is modeled by setting the initial

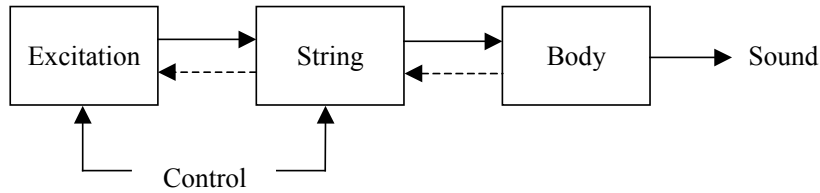


Figure 2.1: Model structure.

displacement of the string (see Sec. 2.4.2), then the string vibration has no influence on the excitation. On the other hand, if the hammer strike in the case of the piano is modeled by a nonlinear hammer model as in Sec. 2.4.1, the coupling is bidirectional, as the hammer force is a function of string shape. The same happens for a bowing type of excitation, where the periodic excitation force is the result of the continuous interaction of the string and the bow (see Sec. 2.4.3).

The interaction of the instrument body and the strings can also be treated in two ways. The physically precise solution is when the strings are continuously coupled to the instrument body, leading to the coupling of different strings and to a change of partial frequencies and decay times. However, often the instrument body is implemented as a post-processing unit, i.e., as a linear filter. In this case, the impedance effects of the body (that alter the frequencies and decay times of string modes) are implemented within the string model, and the body model reproduces only the radiation properties of the real instrument body.

Note that if both interactions are unidirectional (showed by solid lines in Fig. 2.1) and the building blocks are linear, the model reduces to a series of linear filters. In this case the model elements can be commuted, leading to large computational savings. This will be discussed in Sec. 2.5.3.

2.2 String Equations

The most important part of the string instrument is the string itself. This section provides the theoretical background for the string modeling methods of Sec. 2.3. Note that most of the findings of this section can be found in acoustic textbooks, such as [Morse 1948; Morse and Ingard 1968]. However, summarizing them here together with their derivations (which are slightly different from the ones in [Morse 1948; Morse and Ingard 1968]) should help the understanding of the results of Chaps. 5 and 6. In this section only the linear vibration of strings is investigated in detail, while the nonlinear phenomena will be discussed in Chap. 5.

2.2.1 General Equations of String Vibration

For perfectly describing the motion of the string, it should be treated as a prestressed thin rod by elasticity theory [Graff 1975]. However, it is sufficient for our purposes to suppose

that the string has negligible diameter, where simpler theory applies. Note that one effect of the finite diameter of real strings, the wave dispersion, will be included later in Sec. 2.2.4.

Let us suppose that a lossless and perfectly flexible string is stretched along the x axis with an initial tension T_0 (losses and stiffness will be covered later in Sec. 2.2.4). The string element that was, in equilibrium, at point $(x, 0, 0)$ will be at point $(x + \xi, y, z)$, where ξ refers to the longitudinal displacement (x direction), while y and z are the displacements of the two transverse polarizations (both perpendicular to the string). This part of the derivation is taken directly from [Morse and Ingard 1968, pp. 857–858]. The vector from the origin to point $(x + \xi, y, z)$ is

$$\mathbf{R}(x, t) = (x + \xi)\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z, \quad (2.1)$$

where \mathbf{e}_x , \mathbf{e}_y , \mathbf{e}_z are unit vectors along coordinate axes.

The length of the string element labeled x [the element that was originally at $(x, 0, 0)$], which was dx in equilibrium (under the initial tension T_0) will be

$$ds = |\mathbf{R}(x + dx, t) - \mathbf{R}(x, t)| = \left| \frac{\partial \mathbf{R}}{\partial x} \right| dx = \sqrt{\left(\frac{\partial \xi}{\partial x} + 1 \right)^2 + \left(\frac{\partial y}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial x} \right)^2} dx. \quad (2.2)$$

The tension $T = T(x, t)$ of the string is calculated from the relative elongation $(ds - dx_0)/dx_0$ according to the Hooke's law

$$T = ES \left(\frac{ds}{dx_0} - 1 \right) = ES \left(\frac{ds}{dx} \frac{dx}{dx_0} - 1 \right), \quad (2.3)$$

where E is the Young's modulus, S is the cross-section area of the string, and dx_0 is the length of the string element without tension ($T = 0$). Note that here we depart from Eq. 14.3.3 of [Morse and Ingard 1968] (which implicitly assumes $dx = dx_0$) by applying the equation of tension presented in [Kurmyshev 2003], leading to more precise results. By knowing that the initial tension is obtained as $T_0 = ES(dx/dx_0 - 1)$, dx_0 can be eliminated from Eq. (2.3) as

$$T = ES \left(\frac{ds}{dx} - 1 \right) + T_0 \frac{ds}{dx}. \quad (2.4)$$

If the string is perfectly flexible, the only force acting on the element x is the tension at its sides $T(x, t)$ and $T(x + dx, t)$. The direction of these force components is given by the unit vectors $\mathbf{e}_s(x, t)$ and $\mathbf{e}_s(x + dx, t)$ pointing along the tangent to the string

$$\mathbf{e}_s(x, t) = \frac{\frac{\partial \mathbf{R}}{\partial x}}{\left| \frac{\partial \mathbf{R}}{\partial x} \right|} = \frac{\left(\frac{\partial \xi}{\partial x} + 1 \right) \mathbf{e}_x + \left(\frac{\partial y}{\partial x} \right) \mathbf{e}_y + \left(\frac{\partial z}{\partial x} \right) \mathbf{e}_z}{\sqrt{\left(\frac{\partial \xi}{\partial x} + 1 \right)^2 + \left(\frac{\partial y}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial x} \right)^2}}. \quad (2.5)$$

The net force acting on the element x is the difference between the forces at the sides, which is related to the acceleration of the element according to the Newton's law

$$\mu dx \frac{\partial^2 \mathbf{R}}{\partial t^2} = T(x + dx, t) \mathbf{e}_s(x + dx, t) - T(x, t) \mathbf{e}_s(x, t), \quad (2.6)$$

where μdx is the mass of the element x and μ is the linear mass density (i.e., mass per unit length) of the stretched string.

Writing the difference in Eq. (2.6) by differentials gives

$$\mu \frac{\partial^2 \mathbf{R}}{\partial t^2} = \frac{\partial(T \mathbf{e}_s)}{\partial x} = \frac{\partial T}{\partial x} \mathbf{e}_s + T \frac{\partial \mathbf{e}_s}{\partial x}, \quad (2.7)$$

from which it follows that the net force arises from two reasons: because the tension T is different at the sides of the element and because it pulls at different directions \mathbf{e}_s at the two sides.

The above equations (Eqs. (2.4), (2.5), and (2.7)) completely characterize the motion of the flexible string, as no approximations were made so far. Note that the nonlinearity of Eqs. (2.4) and (2.5) comes from the geometry of the structure and not from the nonlinearity of the material (as we assumed that the Hooke's law holds). Therefore, it is often called "geometric nonlinearity".

2.2.2 Approximate Nonlinear Equations

Now let us assume that the relative elongation is small, i.e., $\partial \xi / \partial x$, $\partial y / \partial x$, and $\partial z / \partial x$ are small compared to unity. Moreover, we assume that ξ is in the order of y^2 and z^2 , which holds for metal strings. (For rubber-like strings ξ , y , and z are of same order [Kurmyshev 2003].) In this case Eq. (2.5) can be approximated as

$$\mathbf{e}_s(x, t) = \left[\left(\frac{\partial \xi}{\partial x} + 1 \right) \mathbf{e}_x + \left(\frac{\partial y}{\partial x} \right) \mathbf{e}_y + \left(\frac{\partial z}{\partial x} \right) \mathbf{e}_z \right] \left[1 - \frac{\partial \xi}{\partial x} - \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 - \frac{1}{2} \left(\frac{\partial z}{\partial x} \right)^2 \right]. \quad (2.8)$$

which was obtained by applying the approximations $(1+p)^2 \approx 1+2p$, $\sqrt{1+p} \approx 1+0.5p$, and $1/(1+p) \approx 1-p$.

Similar derivations yield the equation for tension from Eqs. (2.2) and (2.4)

$$T \approx T_0 + (ES + T_0) \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial z}{\partial x} \right)^2 \right]. \quad (2.9)$$

Substituting Eqs. (2.8) and (2.9) into (2.7) and neglecting some higher-order terms gives

$$\mu \frac{\partial^2 \xi}{\partial t^2} = (ES + T_0) \frac{\partial^2 \xi}{\partial x^2} + \frac{1}{2} ES \frac{\partial \left[\left(\frac{\partial y}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial x} \right)^2 \right]}{\partial x} \quad (2.10)$$

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} + ES \frac{\partial \left\{ \frac{\partial y}{\partial x} \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial z}{\partial x} \right)^2 \right] \right\}}{\partial x} \quad (2.11)$$

$$\mu \frac{\partial^2 z}{\partial t^2} = T_0 \frac{\partial^2 z}{\partial x^2} + ES \frac{\partial \left\{ \frac{\partial z}{\partial x} \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 + \frac{1}{2} \left(\frac{\partial z}{\partial x} \right)^2 \right] \right\}}{\partial x}. \quad (2.12)$$

Note that in Eqs. (14.3.7)–(14.3.9) of [Morse and Ingard 1968] $(ES - T_0)$ replaces the ES of Eqs. (2.10)–(2.12) of this thesis. Having $(ES - T_0)$ instead of ES would mean for

increasing T_0 that the nonlinear terms would decrease to zero as T_0 reaches ES , then raise again for higher T_0 values, which is unrealistic. Equations (2.10)–(2.12) do not show this problem, as the nonlinearity drops continuously and it can be considered zero for $T_0 \gg ES$ only. This difference is rather theoretical, as it is significant only for rubber-like strings, where $T_0 \approx ES$. In metal strings we have $ES \gg T_0$, thus, in the rest of this thesis we will assume $ES + T_0 \approx ES \approx ES - T_0$ anyway.

It can be seen that Eqs. (2.10)–(2.12) are three linear wave equations for the three polarizations with nonlinear forcing terms at their right-hand sides. By looking at their linear part, the transverse and longitudinal propagation speeds can be expressed as:

$$c_t = \sqrt{\frac{T_0}{\mu}} \quad c_l = \sqrt{\frac{ES + T_0}{\mu}}. \quad (2.13)$$

In metal strings we have $T_0 \ll ES$, therefore the longitudinal propagation speed can be approximated as $c_l = \sqrt{ES/\mu}$.

The amount of intermodal coupling in Eq. (2.10) is almost independent of string tension (as $ES + T_0 \approx ES$). On the other hand, in Eqs. (2.11) and (2.12) it depends on the ratio ES/T_0 , i.e., the lower the tension, the higher the effect of intermodal coupling is.

2.2.3 Linear Equations

If the displacement of the string is so small that the products and powers of the partial derivatives $\partial\xi/\partial x$, $\partial y/\partial x$, and $\partial z/\partial x$ are negligible compared to the terms themselves, Eqs. (2.10)–(2.12) reduce to three independent wave equations. As the equations for the three polarizations are of the same form, it is reasonable to perform the derivations for one polarization. The wave equation for the y polarization is

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2}, \quad (2.14)$$

or, by using Eq. (2.13)

$$\frac{\partial^2 y}{\partial t^2} = c_t^2 \frac{\partial^2 y}{\partial x^2}. \quad (2.15)$$

The Unterminated String

First we consider the infinite, unterminated string. We try to find the solution for Eq. (2.15) as the real part of the product of two exponential functions

$$y(x, t) = \text{Re}\{e^{j\omega t} e^{j\beta x}\}. \quad (2.16)$$

The substitution of Eq. (2.16) into Eq. (2.15) gives

$$\frac{\omega^2}{\beta^2} = c_t^2 \implies \frac{\omega}{\beta} = \pm c_t. \quad (2.17)$$

Writing this back into Eq. (2.16) yields

$$y(x, t) = \text{Re}\{e^{j(\omega t + \beta x)}\} = \text{Re}\{e^{j\omega(t \pm x/c_t)}\} \quad (2.18)$$

for arbitrary ω . Equation (2.18) shows that the ideal, unterminated string can vibrate at arbitrary angular frequencies. However, the wave number β and angular frequency ω are related by the propagation speed c_t . As the wave equation is linear, the superposition of the two cases of Eq. (2.18) is also a solution:

$$y(x, t) = \text{Re}\{C^+ e^{j\omega(t-x/c_t)} + C^- e^{j\omega(t+x/c_t)}\}, \quad (2.19)$$

where C^+ and C^- are the complex amplitudes of the two components.

By using the notation $p = t \pm x/c_t$, Eq. (2.18) becomes the basis function of the inverse Fourier transform. As the wave equation is linear, the superposition of the functions $e^{j\omega p}$ are also the solution of the wave equation. This means that arbitrary Fourier expandable functions

$$y^\pm(p) = \text{Re}\left\{\int_{\omega=-\infty}^{\infty} C^\pm(\omega) e^{j\omega p}\right\} \quad (2.20)$$

are solutions. As both $p = t + x/c_t$ and $p = t - x/c_t$ holds, the time-domain solution of the wave equation is written as the superposition of two functions

$$y(x, t) = y^+(t - x/c_t) + y^-(t + x/c_t), \quad (2.21)$$

where y^+ and y^- can be considered as two traveling waves, which retain their shape during their movement. The function y^+ is the wave going to the right and the function y^- is the wave going to the left direction. This is the “traveling wave solution” of the wave equation.

Infinitely Rigid Terminations

Rigidly terminating the string at $x = 0$, i.e., setting $y(0, t) = 0$, means that the two traveling waves should cancel each other at $x = 0$. This yields $y^-(p) = -y^+(p)$, meaning that two identical waves of opposite sign travel along the string in the two directions. The components y^+ and y^- can be considered as incident and reflected waves.

In reality the string is stretched between the supports $x = 0$ and $x = L$, where L is the string length. For ideally rigid terminations, this gives the constraint $y(0, t) = y(L, t) = 0$, leading to $y^+(t - L/c_t) = -y^-(t + L/c_t)$. As we have $y^+(p) = -y^-(p)$ from the termination at $x = 0$, we obtain $y^+(t - L/c_t) = y^+(t + L/c_t)$. This means that $y^+(p)$ is periodic with period length $T_p = 2L/c_t$. This corresponds to a fundamental frequency $f_0 = c_t/(2L)$, or an angular frequency $\omega_0 = c_t\pi/L$.

More interesting to us is the exponential form of Eq. (2.19), where $y(0, t) = 0$ gives $C^- = -C^+$, leading to

$$y(x, t) = \text{Re}\{C^+(e^{j\omega(t-x/c_t)} - e^{j\omega(t+x/c_t)})\} = \text{Re}\{2jC^+ \sin(\beta x) e^{j\omega t}\}, \quad (2.22)$$

which shows that if the string is rigidly terminated at $x = 0$ and is vibrating at a single angular frequency ω , a standing wave develops, i.e., the different points of the string vibrate in the same phase. Note that it is still true that arbitrary angular frequencies ω are allowed.

The additional constraint $y(L, t) = 0$ on Eq. (2.22) gives $\beta L = k\pi$, leading to

$$y(x, t) = \text{Re}\{2jC^+ \sin(\beta_k x) e^{j\omega_k t}\} = C \sin(\beta_k x) \cos(\omega_k t + \varphi), \quad (2.23)$$

where the allowed angular frequencies ω_k and wave numbers β_k are

$$\beta_k = \frac{k\pi}{L} \quad \omega_k = \frac{c_t k\pi}{L}. \quad (2.24)$$

This means that the terminated string can only vibrate at distinct angular frequencies ω_k , which form a perfect harmonic series. These frequencies ω_k belong to specific modal shapes $\sin(\beta_k t)$, where k is the mode number.

2.2.4 The Stiff and Lossy String

A wave equation for the stiff and lossy string is

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} - ES\kappa^2 \frac{\partial^4 y}{\partial x^4} - 2R(\omega)\mu \frac{\partial y}{\partial t} + d_y(x, t), \quad (2.25)$$

which is the Helmholtz equation extended by terms describing the stiffness and losses of the string. The stiffness of the string is characterized by $ES\kappa^2$, where κ is the radius of gyration. This term (fourth-order spatial derivative) is essentially the same as what can be found in the wave equation of bars and rods, but now it has a secondary role, as the main force on the string comes from the tension (second-order spatial derivative). Accordingly, the wave equation of the stiff string can be considered as a transition between the wave equation of the ideal string and that of the ideal bar [Morse 1948, p. 166].

The operator $R(\omega)$ is the frequency dependent frictional resistance [Morse 1948, p. 104]. The factor “2” before $R(\omega)$ in Eq. (2.25) is chosen in order to make the decay rate $\sigma(\omega)$ of the partial at the angular frequency ω equal to $R(\omega)$, as we will see later in Eqs. (2.27) and (2.37). External driving forces are included in the excitation force density $d_y(x, t)$, which has the dimension of force per unit length.

Unterminated Case

Let us first consider the case of infinite, unterminated string. Moreover, we assume that no external driving forces are present, i.e., $d_y(x, t) = 0$. The losses are assumed to be independent of frequency, that is, $R(\omega) = R$. The substitution of the trial function Eq. (2.16) into Eq. (2.25) gives

$$-\mu\omega^2 = -T_0\beta^2 - ES\kappa^2\beta^4 - j\omega 2R\mu, \quad (2.26)$$

which cannot be solved for real ω and β values. However, if we substitute ω with a complex variable $\Omega = \omega + j\sigma$ in Eq. (2.26), we obtain a pair of equations for the real and imaginary parts. The equation for the imaginary part of Eq. (2.26) becomes

$$-\mu 2j\omega\sigma = -j\omega 2R\mu, \quad (2.27)$$

which gives $\sigma = R$. The real part of Eq. (2.26) is

$$-\mu(\omega^2 - \sigma^2) = -T_0\beta^2 - ES\kappa^2\beta^4 + \sigma 2R\mu. \quad (2.28)$$

The substitution $\sigma = R$ gives

$$\omega^2 = \frac{T_0}{\mu}\beta^2 + \frac{ES\kappa^2}{\mu}\beta^4 - R^2, \quad (2.29)$$

where R^2 is negligible in comparison with the other two terms in the case of musical instrument strings.

Accordingly, the final solution becomes

$$y = \operatorname{Re}\{e^{j\Omega t} e^{j\beta x}\} = e^{-\sigma t} \operatorname{Re}\{e^{j\omega(t \pm x/c_t(\beta))}\}, \quad (2.30)$$

which differs from Eq. (2.18) in two aspects: first, there is an additional term $e^{-\sigma t}$ meaning that the propagating waves decay exponentially. Second, the angular frequency ω and wave number β are no longer related to each other by a constant factor c_t as in Eq. (2.17), but now the propagating speed depends on the wave number:

$$c_t(\beta) = \frac{\omega}{\beta} = \sqrt{\frac{T_0}{\mu} + \frac{ES\kappa^2}{\mu}\beta^2}. \quad (2.31)$$

Equation (2.31) shows that waves at larger wave numbers β (i.e., at larger angular frequencies ω) propagate faster. Accordingly, the high frequency components of a wave will travel faster than the low frequency ones. As a result, traveling waves will no longer retain their shapes but disperse.

Termination and Driving Forces

As can be seen in Eq. (2.23), the ideal string rigidly terminated at $x = 0$ and $x = L$ can only vibrate at specific wave numbers β_k , and the spatial dependence is in the form of $\sin(\beta_k x)$. This property does not change by adding losses, dispersion, and external driving forces as in Eq. (2.25). However, the temporal dependence of a specific mode is no longer a cosine function, as was in Eq. (2.23). As the most general case, the string shape $y(x, t)$ is expressed by the Fourier-like series

$$y(x, t) = \sum_{k=1}^{\infty} y_k(t) \sin\left(\frac{k\pi x}{L}\right) \quad (2.32)$$

where $y_k(t)$ is the instantaneous amplitude of mode k .

Note that Eq. (2.32) completely characterizes $y(x, t)$ in the range $0 \leq x \leq L$ for each t if

$$\left. \frac{\partial^p y(x, t)}{\partial x^p} \right|_{x=0} = \left. \frac{\partial^p y(x, t)}{\partial x^p} \right|_{x=L} = 0 \quad (2.33)$$

holds for even p including $p = 0$. This constraint comes from the fact that the even derivatives of the sine functions in Eq. (2.32) are also sine functions, which are zero at $x = 0$ and $x = L$ for each n . This corresponds to hinged boundary conditions. The cosine functions are missing because Eq. (2.32) is considered as the Fourier series of the odd function $y(-x, t) = -y(x, t)$ with the period length $2L$ (however, the part $-L < x < 0$ is of no interest to us).

The solution of Eq. (2.25) can be separated for the different modes if Eq. (2.32) is substituted into Eq. (2.25), then multiplied by the modal shape $\sin(k\pi x/L)$ and integrated

over x from 0 to L . The resulting second-order differential equation covering the behavior of mode n is

$$\frac{d^2 y_k}{dt^2} + a_{1,k} \frac{dy_k}{dt} + a_{0,k} y_k = b_{0,k} F_{y,k}(t), \quad (2.34)$$

where

$$a_{1,k} = 2R(\omega_k) \quad (2.35a)$$

$$a_{0,k} = \frac{T_0}{\mu} \left(\frac{k\pi}{L} \right)^2 + \frac{ES\kappa^2}{\mu} \left(\frac{k\pi}{L} \right)^4 \quad (2.35b)$$

$$b_{0,k} = \frac{2}{L\mu} \quad (2.35c)$$

$$F_{y,k}(t) = \int_{x=0}^L \sin\left(\frac{k\pi x}{L}\right) d_y(x, t) dx. \quad (2.35d)$$

In Eq. (2.34) $F_{y,k}(t)$ can be considered as the excitation force of mode k , and it is computed as the scalar product of the excitation force density and the modal shape (see Eq. (2.35d)).

The solution of Eq. (2.34) for $F_{y,k}(t) = \delta(t)$ with zero initial conditions is an exponentially decaying sine function covered by

$$y_{\delta,k}(t) = A_k e^{-\frac{t}{\tau_k}} \sin(\omega_k t) \quad (2.36a)$$

$$A_k = \frac{b_{0,k}}{\omega_k} \quad (2.36b)$$

$$\tau_k = \frac{2}{a_{1,k}} \quad (2.36c)$$

$$\omega_k = \sqrt{a_{0,k} - \frac{a_{1,k}^2}{4}} \approx \sqrt{a_{0,k}}. \quad (2.36d)$$

The decay time $\tau_k = 1/\sigma_k$ of mode k is simply related to the frictional resistance by

$$\tau_k = \frac{2}{a_{1,k}} = \frac{1}{R(\omega_k)}. \quad (2.37)$$

This is in accordance with the result of Eq. (2.27), derived for the unterminated string. The approximation in Eq. (2.36d) is valid if the frictional resistance is small, i.e., it has a negligible effect on the angular frequency ω_k . In other words, if $a_{1,k}^2/4 = 1/\tau_k^2$ is negligible in comparison with ω_k^2 , which holds for the slowly decaying modes of strings. To this approximation the angular frequency is given by

$$\omega_k = \sqrt{a_{0,k}} = \sqrt{\frac{T_0}{\mu} \left(\frac{k\pi}{L} \right)^2 + \frac{ES\kappa^2}{\mu} \left(\frac{k\pi}{L} \right)^4} = \omega_0 k \sqrt{1 + Bk^2}, \quad (2.38)$$

where the the fundamental angular frequency of the string w_0 is

$$\omega_0 = \frac{\pi}{L} \sqrt{\frac{T_0}{\mu}} = c_t \frac{\pi}{L}. \quad (2.39)$$

The inharmonicity of the angular frequency series ω_k is determined by the inharmonicity coefficient B , which is computed as

$$B = \kappa^2 \frac{ES}{T_0} \left(\frac{\pi}{L} \right)^2. \quad (2.40)$$

Note that $B = 0$ corresponds to the case of perfectly flexible string, where the partial frequencies follow $\omega_k = k\omega_0$ (as predicted by Eq. (2.24)). The higher the stiffness of the string (or the smaller the tension), the larger is the difference with which the partials depart from the perfectly harmonic series.

As Eq. (2.36) computes the impulse response of the system characterized by Eq. (2.34), the response to the excitation force $F_{y,k}(t)$ is obtained by the time domain convolution

$$y_k(t) = y_{\delta,k}(t) * F_{y,k}(t). \quad (2.41)$$

From Eqs. (2.32), (2.36), and (2.41) the displacement of the string $y(x, t)$ as a response to the external force density $d_y(x, t)$ is computed as

$$y(x, t) = \frac{1}{\pi L \mu} \sum_{k=1}^{\infty} \sin\left(\frac{k\pi x}{L}\right) \left[\frac{1}{f_k} e^{-\frac{t}{\tau_k}} \sin(2\pi f_k t) * \int_{x=0}^L \sin\left(\frac{k\pi x}{L}\right) d_y(x, t) dx \right], \quad (2.42)$$

where the angular frequency was substituted by the frequency $f_k = \omega_k/(2\pi)$. Equation (2.42) means that first the scalar product of the excitation-force density and the modal shape has to be computed, then this has to be convolved with the time-domain impulse response of mode k , leading to the instantaneous amplitude of mode k . Finally, these modes are summed together multiplied by their modal shapes $\sin(k\pi x/L)$.

2.3 String Modeling Techniques

The following sections discuss the different string modeling approaches. Only the three most widely used modeling techniques are reviewed, namely, finite-difference, modal-based, and digital waveguide approaches, as these are the techniques that are applied in this thesis. For a more comprehensive review of physics-based string modeling techniques, see [Välimäki et al. 2006]. The first modeling technique is the finite-difference modeling, which is the direct numerical solution of the wave equation (such as Eq. (2.25)). The other two approaches are based on the discretization of the continuous time solutions of the wave equation: digital waveguide modeling presented in Sec. 2.3.2 discretizes the traveling wave solution of Eq. (2.21), while modal synthesis outlined in Sec. 2.3.3 implements the modal form of Eqs. (2.32) and (2.42).

2.3.1 Finite-difference Modeling

The first finite-difference string model was presented by Hiller and Ruiz [1971a,b], which was the first physics-based instrument model, too. Finite-difference modeling has become popular because it has a direct connection to the wave equation and also because it is straightforward to use it for two- or three-dimensional structures. Moreover, connecting

these different structures is a simple task. A drawback of the approach is the high computational complexity and the numerical dispersion, the latter meaning that the modal frequencies of the model will be different from that of the continuous time system.

The Ideal String

In finite-difference modeling, the solution of a partial differential equation is computed by substituting derivatives by finite differences. The usual way of discretizing Eq. (2.15) on a grid $x_m = m\Delta x$, $t_n = n\Delta t$ is

$$\left. \frac{\partial^2 y}{\partial x^2} \right|_{x_m, t_n} \approx \frac{y_{m-1,n} - 2y_{m,n} + y_{m+1,n}}{\Delta x^2} \quad (2.43a)$$

$$\left. \frac{\partial^2 y}{\partial t^2} \right|_{x_m, t_n} \approx \frac{y_{m,n-1} - 2y_{m,n} + y_{m,n+1}}{\Delta t^2}, \quad (2.43b)$$

where $y_{m,n} = y(x_m, t_n)$. The substitution of Eq. (2.43) into Eq. (2.15) gives

$$y_{m,n+1} = \frac{c_t^2 \Delta t^2}{\Delta x^2} (y_{m-1,n} - 2y_{m,n} + y_{m+1,n}) - y_{m,n-1} + 2y_{m,n}, \quad (2.44)$$

which computes the next value of the element at position m from the past and previous state of the string. To model infinitely rigid boundary conditions, $y_{0,n}$ and $y_{M,n}$ is set to zero for all n (where $M\Delta x = L$ is the string length). The system is numerically stable for $\Delta x/\Delta t \geq c_t$, i.e., when the waves do not move more than one spatial interval during one time step [Chaigne and Askenfelt 1994]. Due to the nature of discretization numerical dispersion arises on the string, which should not be confused with the dispersion caused by the stiffness of the string. The numerical dispersion stretches the partials in a way that they lie at a lower frequency compared to where they should be [see, e.g., Chaigne and Askenfelt 1994]. This means that the system Eq. (2.44) won't produce an exact copy of the analytical solution of Eq. (2.15). However, the difference is negligible for low frequencies and does not usually lead to the degradation of sound quality.

If Δx and Δt are chosen in a way that the wave travels exactly one spatial interval during one time step as suggested by Hiller and Ruiz [1971a], that is, $c_t = \Delta x/\Delta t$ in Eq. (2.44), we obtain

$$y_{m,n+1} = y_{m-1,n} + y_{m+1,n} - y_{m,n-1}. \quad (2.45)$$

Equation (2.45) is exact in the sense that there is no numerical dispersion in this case. It is easy to see that when Eq. (2.45) is excited by a unit impulse at the spatial position m , two unit pulses will develop that travel in the left and right direction without changing their shapes. This is a great advantage. Another benefit is that with this choice the finite-difference model can be easily connected to a digital waveguide model, as in digital waveguides we always have $c_t = \Delta x/\Delta t$ [Karjalainen and Erkut 2004].

However, the flexibility is lost: now Δx , Δt , and c_t are interdependent. Practically, this means that the fundamental frequency of the string model can only be set by changing the number of string elements, M , assuming that the sampling frequency $f_s = 1/\Delta t$ is given, since $f_0 = f_s/(2M)$. This leads to unnecessarily large M values for low notes. In the case

of the more general Eq. (2.44) the fundamental frequency can be easily varied by modifying c_t or $\Delta x = L/M$, where the number of string elements M can be set according to how many partials we wish to compute. This way, the accuracy/complexity of the model and the fundamental frequency of the tone become independent.

Losses and Dispersion

Modeling the stiff and lossy string requires the discrete-time implementation of Eq. (2.25). However, Eq. (2.25) has a frequency dependent parameter $R(\omega)$, which cannot be implemented directly. Hiller and Ruiz [1971a] and later Chaigne and Askenfelt [1994] suggested to use the formal substitution

$$R(\omega) \frac{\partial y}{\partial t} \Rightarrow b_1 \frac{\partial y}{\partial t} - b_3 \frac{\partial^3 y}{\partial t^3} \quad (2.46)$$

in Eq. (2.25) as the simplest way for implementing frequency dependent losses. This leads to the decay times

$$\tau_k = \frac{1}{R(\omega_k)} = \frac{1}{b_1 + b_3 \omega_k^2}. \quad (2.47)$$

This simplified formula has been found to match the decay times of musical instrument strings quite well [Chaigne and Askenfelt 1994]. Interestingly, it also coincides with the decay times of the digital waveguide string model if the most common loss filter is used (see Sec. 3.1).

However, the use of Eq. (2.46) yields a recurrence equation which may not be stable for all b_3 parameters. Therefore, a different way of implementing frequency dependent losses has been suggested by Bilbao in [Bensa et al. 2003]:

$$R(\omega) \frac{\partial y}{\partial t} \Rightarrow b_1 \frac{\partial y}{\partial t} - b_2 \frac{\partial^3 y}{\partial x^2 \partial t}. \quad (2.48)$$

This leads to a second-order differential equation for the individual modes of the same form as Eq. (2.34), but now $a_{1,k} = 2b_1 + 2b_2(k\pi/L)^2$, giving the decay times

$$\tau_k = \frac{2}{a_{1,k}} = \frac{1}{b_1 + b_2 \left(\frac{k\pi}{L}\right)^2}, \quad (2.49)$$

which is the same as Eq. (2.47) for harmonic (nondispersive) strings if $b_2 = c_t^2 b_3$ (see Eq. (2.24)). For inharmonic strings the decay times given by Eq. (2.49) will be somewhat higher than that of Eq. (2.47), as $c_t k\pi/L < \omega_k$ (see Eq. (2.38)). As the ear is relatively insensitive to differences in decay times [Tolonen and Järveläinen 2000], this difference won't influence sound quality. However, from the practical implementation viewpoint the form of Eq. (2.48) is more advantageous, as it is stable for arbitrary positive b_1 and b_2 , while requiring less memory and less computation than the third order system obtained by the use of Eq. (2.46) [Bensa et al. 2003].

Accordingly, we discretize the following wave equation

$$\frac{\partial^2 y}{\partial t^2} = c_t^2 \frac{\partial^2 y}{\partial x^2} - \frac{ES\kappa^2}{\mu} \frac{\partial^4 y}{\partial x^4} - 2b_1 \frac{\partial y}{\partial t} + 2b_2 \frac{\partial^3 y}{\partial x^2 \partial t} + \frac{1}{\mu} d_y(x, t), \quad (2.50)$$

which is the modification of Eq. (2.25) by applying the substitution Eq. (2.48). For the discretization, we can use the formulas

$$\left. \frac{\partial^4 y}{\partial x^4} \right|_{x_m, t_n} \approx \frac{y_{m-2,n} - 4y_{m-1,n} + 6y_{m,n} - 4y_{m+1,n} + y_{m+2,n}}{\Delta x^4} \quad (2.51a)$$

$$\left. \frac{\partial y}{\partial t} \right|_{x_m, t_n} \approx \frac{y_{m,n} - y_{m,n-1}}{\Delta t} \quad (2.51b)$$

$$\left. \frac{\partial^3 y}{\partial x^2 \partial t} \right|_{x_m, t_n} \approx \frac{(y_{m-1,n} - 2y_{m,n} + y_{m+1,n}) - (y_{m-1,n-1} - 2y_{m,n-1} + y_{m+1,n-1})}{\Delta x^2 \Delta t} \quad (2.51c)$$

together with Eq. (2.43). Inserting the approximations Eqs. (2.43) and (2.51) into Eq. (2.50) yields a recurrence equation that computes the next value of the string displacement $y_{m,n+1}$ from the past ($y_{m-1,n-1}$, $y_{m,n-1}$, $y_{m+1,n-1}$) and present ($y_{m-2,n}$, $y_{m-1,n}$, $y_{m,n}$, $y_{m+1,n}$, $y_{m+2,n}$) values. Note that the derivatives can be approximated in other ways as well, e.g., the temporal derivative $\partial y / \partial t$ can also be computed as $(y_{m,n+1} - y_{m,n}) / \Delta t$. Different approximations give different output, but these differences are generally negligible for sufficiently high sampling rates and large number of elements M (i.e., for small Δt and Δx).

For implementing the boundary condition, we need a further constraint besides $y_{0,n} = y_{M,n} = 0$, since $y_{-1,n}$ and $y_{M+1,n}$ are also required for the computation of the rightmost and leftmost points of the string (which have the spatial coordinates $m = 1$ and $m = M - 1$). Setting the second-order spatial derivative to zero at the boundaries gives $y_{-1,n} = -y_{1,n}$ and $y_{M+1,n} = -y_{M-1,n}$, which corresponds to hinged boundary conditions. More realistic string terminations can also be implemented by, e.g., computing the string force acting on the termination and calculating the velocity response by a model of the termination admittance. A termination model parameterized by a constant reflection coefficient is presented in [Hiller and Ruiz 1971a].

The force at the termination is of great interest even in the simplified models having ideally rigid string terminations, as this is the force which is transmitted to the body of the instrument. In other words, this is the output of the string model. This force is computed as

$$F_b(t_n) = -T_0 \left. \frac{\partial y}{\partial x} \right|_{x=L} \approx -\frac{T_0}{\Delta x} y_{M-1,n}. \quad (2.52)$$

It may sound paradoxical that the string termination is infinitely rigid, while it can still transfer energy to the instrument body. The resolution is that the termination moves at much smaller amplitude compared to the string, thus, it may be considered as motionless in the string model. From the instrument body side, however, this small movement is the excitation, i.e., it cannot be neglected.

Note that when the string is modeled as a series of masses connected with springs and dampers (which is already discrete system in space), and this is discretized with respect to time, the same or very similar recurrence equations are yielded as for the finite-difference model [Rowland and Pask 1999].

2.3.2 Digital Waveguide Modeling

Digital waveguide modeling introduced by Smith [1983, 1992] has been the most widely used string modeling technique. This is because the time-domain solution of the one-dimensional equation provides a very efficient implementation using DSP techniques. (This is unfortunately not true for structures of two or three dimensions.) Actually, the system reduces to a delay line and a filter in a feedback loop, similarly to what has been proposed by McIntyre et al. [1983]. The simplest variation is the Karplus-Strong algorithm, where the filter is a simple averaging operation [Karplus and Strong 1983; Jaffe and Smith 1983]. A nice feature of digital waveguide modeling is that while it is a simple filtering algorithm from the DSP point of view, it still retains the physicality of the system. Thus, the interaction of the different parts of the instrument (e.g., the coupling of different strings) is easily implemented. A comprehensive overview on digital waveguide modeling can be found in [Smith 2005].

Ideal String

Digital waveguide modeling is based on the spatial and temporal sampling of the traveling wave solution of the wave equation. Rewriting Eq. (2.21) with $x_m = m\Delta x$ and $t_n = n\Delta t$ gives

$$y(x_m, t_n) = y^+(n\Delta t - m\Delta x/c_t) + y^-(n\Delta t + m\Delta x/c_t). \quad (2.53)$$

This can be implemented by storing the samples of two traveling wave components in two vectors and moving their content into the left or right direction at every time step. In digital waveguide modeling we always have $\Delta x = c_t\Delta t$. Theoretically, we could have $\Delta x \neq c_t\Delta t$, but moving the content of the vectors by a fractional spatial sample would require the use of interpolation techniques. Accordingly, substituting $\Delta x = c_t\Delta t$ into Eq. (2.53) leads to a much simpler equation

$$y_{m,n} = y_{m-n}^+ + y_{m+n}^-, \quad (2.54)$$

where $y_{m-n}^+ = y^+(\Delta t(m-n))$ and $y_{m+n}^- = y^-(\Delta t(m+n))$. If the current shapes of the traveling waves are stored to the vectors $y_{m,n}^+$ and $y_{m,n}^-$, the wave propagation is simply implemented by shifting the contents of these vectors by one spatial sample at each time step. Consequently, the next values are computed by

$$\begin{aligned} y_{m,n+1}^+ &= y_{m-1,n}^+ \\ y_{m,n+1}^- &= y_{m+1,n}^- \end{aligned} \quad (2.55)$$

where $y_{m,n} = y_{m,n}^+ + y_{m,n}^-$. This is illustrated in Fig. 2.2, where z^{-1} stands for unit delays. These unit delays form two delay lines, which can be implemented very efficiently as circular buffers.

As a result of the linearity of the wave equation, other variables can also be used instead of displacement. These can be for example velocity, acceleration, slope, curvature or force. Nevertheless, it is worth turning our attention to the transverse velocity v and the force F ,

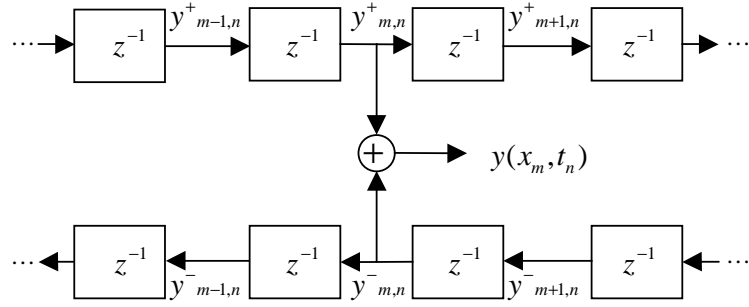


Figure 2.2: The principle of digital waveguide.

since they are proportional to each other. The characteristic impedance Z_0 of the string can be defined as follows [see, e.g., Smith 1992; Morse 1948, pp. 91–93]:

$$Z_0 = \frac{F^+}{v^+} = -\frac{F^-}{v^-}, \quad Z_0 = \sqrt{T\mu} \quad (2.56)$$

where F^+ and v^+ are the force and velocity waves traveling to the right, and F^- and v^- to the left, respectively. Eq. (2.56) 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 similar to the termination of a transmission line. The equations for the reflection of force and velocity waves are the following:

$$r_v = \frac{v^-(L, t)}{v^+(L, t)} = \frac{Z_0 - Z}{Z_0 + Z}, \quad r_F = \frac{F^-(L, t)}{F^+(L, t)} = -r_v = \frac{Z - Z_0}{Z_0 + Z} \quad (2.57)$$

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 ($r_v = -1$). The latter can also derived from $v(L, t) = v^+(L, t) + v^-(L, t) = 0$, which means motionless, i.e., perfectly rigid terminations.

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 M_{in} , as the excitation is acting on two pieces of string, which both have the impedance Z_0 . On the grounds of these equations the digital waveguide model of the ideal string can be formulated as shown in Fig. 2.3. The signs of the form “ z^{-D} ” stand for delay lines of length D .

Losses and Dispersion

In the case of the lossy and stiff string, the traveling waves no longer retain their shapes, but dispersed and attenuated, as it is shown in Eqs. (2.30) and (2.31). Frequency independent losses can be easily incorporated into the digital waveguide model of the ideal string by inserting constant gain factors between the delay elements in Fig. 2.2. Practically, this means that the vectors of the traveling waves are not simply shifted but attenuated by a

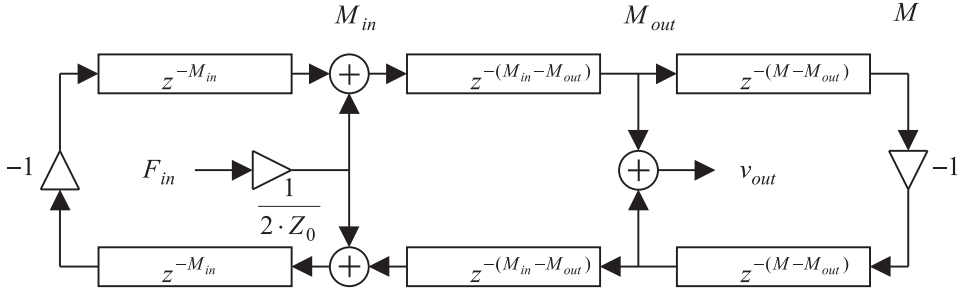


Figure 2.3: Digital waveguide model of the ideal string.

constant coefficient in every time step. Frequency dependent losses require the insertion of digital filters between all the elements. The effect of dispersion can also be implemented by digital filters that have larger phase delay at low frequencies than at high frequencies. This leads to virtually larger propagation speed at high frequencies, as required by Eq. (2.31). However, inserting filters between all the delay elements would be computationally demanding.

In fact, the efficiency of digital waveguide modeling lies in consolidating the effect of these filters into some specific points [Smith 1992]. Usually we are not interested in the precise motion of each point along the string: the model should only match the wave propagation between the excitation point and the output. All the filters between these two points can be moved anywhere in the chain, as the system is linear and time-invariant. The result is an ordinary delay line and a series of filter elements that we have collected together. The net phase and magnitude response of these filter elements can be implemented by a filter that requires significantly lower computation than the separate elements together.

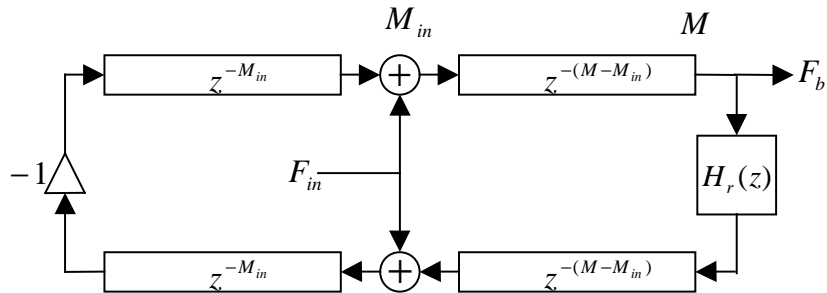


Figure 2.4: Digital waveguide model of the non-ideal string.

It is usual to consolidate the effect of losses and dispersion at one side of the digital waveguide, as depicted by $H_r(z)$ in Fig. 2.4. This has a physical interpretation, as the model of Fig. 2.4 is actually an ideal string terminated by a complex impedance Z . It follows from Eq. (2.57) that if Z is frequency dependent, r_v will also depend on frequency ($H_r(z)$ actually implements $-r_v$ in Fig. 2.4). From the signal processing viewpoint, an

impulse circulating along the string is no longer filtered at every time step but only when it travels through the termination.

The output of the string model is the force acting on the termination. From Eqs. (2.56) and (2.57) the force at the termination (i.e., bridge) F_b will be:

$$\begin{aligned} F^+ &= v^+ Z_0 \\ F^- &= r_F F^+ = r_F v^+ Z_0 \\ F_b &= F^+ + F^- = (1 + r_F) Z_0 v^+ \approx 2 Z_0 v^+, \end{aligned} \quad (2.58)$$

as $r_f \approx 1$. We again compute the bridge force by assuming infinitely rigid terminations, similarly to the case of Eq. (2.52) in finite-difference modeling.

Parameter Estimation

The behavior of the digital waveguide model is determined by the number of delay elements $N = 2M$ and the reflection filter $H_r(z)$. The digital waveguide model of Fig. 2.4 has the transfer function

$$\begin{aligned} H_{\text{wg}}(z) &= \frac{F_b}{F_{\text{in}}} = H_c(z) \frac{1}{1 - z^{-N} H_r(z)} \\ H_c(z) &= (1 - z^{-2M_{\text{in}}}) z^{-(M - M_{\text{in}})} \end{aligned} \quad (2.59)$$

where $H_c(z)$ is a comb filter and a delay depending on the position of the excitation along the string, and $H_r(z)$ is the reflection filter, or loop filter. The partial frequencies and decay times are determined by $H_r(z)$, while $H_c(z)$ influences the initial phases and amplitudes.

The modal frequencies of the digital waveguide can be estimated by finding the local maxima of the transfer function $H_{\text{wg}}(z)$, which are at those frequencies where the denominator is close to zero, that is, $z^{-N} H_r(z) \approx 1$. As the magnitude of the reflection filter $|H_r(z)|$ is close to unity, this condition is met when the phase of $z^{-N} H_r(z)$ is a multiple of 2π :

$$\varphi\{z^{-N} H_r(z)\} = \varphi\{e^{-j\vartheta_k N} H_r(e^{j\vartheta_k})\} = -N\vartheta_k + \varphi\{H_r(e^{j\vartheta_k})\} = -k2\pi, \quad (2.60)$$

which gives a digital angular frequency ϑ_k for each k . Accordingly, the analog partial frequencies are $f_k = [f_s / (2\pi)] \vartheta_k$, where f_s is the sampling frequency.

The decay time of mode k having the frequency f_k can be simply computed by knowing that mode k is attenuated by $|H_r(e^{j\vartheta_k})|$ each time it passes the reflection filter. As one period of mode k fits into the digital waveguide loop k times (see Eq. (2.60)), it is attenuated at a periodicity of k/f_k . This gives the following expression for the decay times:

$$\tau_k = -\frac{k}{f_k \ln |H_r(e^{j\vartheta_k})|}, \quad (2.61)$$

where $\vartheta_k = (2\pi f_k) / f_s$.

In the case of parameter estimation for the digital waveguide the partial frequencies f_k and decay times τ_k are known and the filter $H_r(z)$ has to be designed and the length

of the delay line, N , has to be set. The partial frequencies f_k and decay times τ_k are either computed from the physical parameters of the string (e.g., by the equations given in Sec. 2.2.4), or estimated from recorded instrument sounds. Estimation techniques include the work of [Välämäki et al. 1996; Karjalainen et al. 2002]. Substituting the partial frequencies and decay times into Eqs. (2.60) and (2.61) leads to a complex specification, which could be directly used for filter design. For this, one should be able to separate the phase and the amplitude errors 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 to simulate inharmonicity. The decay times of the partials depend on the amplitude response of the filter. It follows from Eq. (2.61) 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 complicated filter design algorithm would be needed, which could handle the magnitude and phase errors separately. Smith [1983] reviews a number of sophisticated filter design techniques, two of them are also discussed in [Laroche and Jot 1992]. However, these methods are rarely used in practice because of their complexity and their instability.

Probably this is the reason why it is common to divide the filter design procedure to designing the loss filter, dispersion filter and fractional delay filter parts [see, e.g., Jaffe and Smith 1983; Välämäki et al. 1996] and using these filters in series as $H_r(z) = H_l(z)H_d(z)H_{fd}(z)$. Note that dividing the filter design into different steps cannot give mathematically optimal results, since the separate parts of the filters have some constraints 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 it might be beneficial to develop robust algorithms that could design one complete reflection filter based on the previously mentioned amplitude and phase criteria. However, we have to note that this would only slightly diminish the computational load of the whole model while it would increase the complexity of parameter estimation.

Loss Filter Design

The specification of the loss filter can be computed by using the inverse of Eq. (2.61):

$$g_k = e^{-\frac{k}{f_k \tau_k}} \quad (2.62)$$

where τ_k is the decay time of partial k , 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 decay time error is a nonlinear function of the amplitude error. The stability of the digital waveguide loop is also hard to handle, since a small deviation from the specification can lead to a magnitude response larger than unity.

Because of the above mentioned problems and because it already provides good sound, it is widespread to use simple loss filters, such as second-order FIR filters [Borin et al. 1997] or first-order IIR filters, among those the one-pole loss filter [Välimäki et al. 1996; Välimäki and Tolonen 1998] being the most common one. The transfer function of such a filter is:

$$H_{1p}(z) = g \frac{1 + a_1}{1 + a_1 z^{-1}} \quad (2.63)$$

where $-a_1$ is the pole of the filter and g refers to the DC gain. The advantage of using a one-pole filter is that it is always of a lowpass character for $a_1 < 0$. Accordingly, keeping g below unity assures the stability of the waveguide loop. 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. Jaffe and Smith [1983] also discussed the use of the one-pole filter, but without the gain factor, i.e., $g = 1$. Interestingly, the decay times of the digital waveguide using the one-pole loss filter are very similar to the decay times of the simplest finite-difference string model, given by Eq. (2.47). This is discussed in Sec. 3.1.1. The different approaches for designing the loss filters will be outlined in Chap. 3 in detail, together with the proposed algorithms.

Dispersion Filter Design

The string dispersion is modeled by $H_d(z)$, which is an allpass filter with largely varying phase delay as a function of frequency. The phase delay specification for the dispersion filter can be obtained from Eq. (2.64) as

$$\varphi_k = N\vartheta_k - k2\pi - \varphi\{H_1(e^{j\vartheta_k})\}, \quad (2.64)$$

where φ_k is the prescribed phase response of the dispersion filter $H_d(z)$. Note that the phase response of the loss filter $H_1(z)$ is subtracted from the specification (last term of Eq. (2.64)). Van Duyne and Smith [1994] proposed an efficient method for simulating dispersion by cascading equal first-order allpass filters in the waveguide loop. However, the constraint of using equal first-order sections does not allow accurate tuning of inharmonicity.

Rocchesso and Scalcon [1996] proposed a design method based on [Lang and Laakso 1994]. Starting from a target phase response, K frequency points f_k are chosen corresponding to the partial frequencies. The filter order is $N_d < K$. For each partial k the method computes the quantities:

$$\alpha_k = -\frac{1}{2}(\varphi_k + N_d\vartheta_k), \quad (2.65)$$

Filter coefficients a_n are computed by solving the system

$$\sum_{n=1}^{N_d} a_n \sin(\alpha_k + n\vartheta_k) = -\sin(\alpha_k), \quad k = 1 \dots K \quad (2.66)$$

As Eq. (2.66) is overdetermined, it is solved in the least-squares sense. The error analysis shows that the phase error is weighted by the magnitude response of the denominator

$|H_d^{\text{den}}(z)|$. This can be compensated by the iterative application of an inverse weighting $1/|H_d^{\text{den}}(z)|$. It was shown by Rocchesso and Scalcon [1996] that several tens of partials can be correctly positioned for any piano key, with the allpass filter order not exceeding 20. Moreover, the fine tuning of the string is automatically taken into account in the design.

Since the computational load for $H_d(z)$ is heavy, it is important to find criteria for accuracy and order optimization with respect to human perception. Rocchesso and Scalcon [1999] studied the dependence of the bandwidth of perceived inharmonicity (i.e., the frequency range in which misplacement of partials is audible) on the fundamental frequency, by performing listening tests with decaying piano tones. The bandwidth has been found to increase almost linearly on a logarithmic pitch scale. Partial above this frequency band only contribute some brightness to the sound, and can be made harmonic without the degradation of sound quality.

Järveläinen et al. [2001] also found that inharmonicity is more easily perceived at low frequencies, even when coefficient B for bass tones is lower than for treble tones. This is probably due to the fact that beats are used by listeners as cues for inharmonicity, and even low B 's produce enough mistuning in higher partials of low tones. These findings can help in the allpass filter design procedure, although a number of issues still need further investigation.

Fractional Delay Filter Design

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. This can be accomplished by dispersion filter design, but it is more flexible to use a separate filter. Naturally, when the dispersion filter $H_d(z)$ is not implemented, the application of the fractional delay filter $H_{\text{fd}}(z)$ cannot be avoided. Once the loss and the dispersion filters are designed, the phase delay (defined as the phase φ divided by the digital angular frequency ϑ) of the delay line at the fundamental frequency $\vartheta_0 = 2\pi f_0/f_s$ should be

$$D_0 = N + D_{\text{fd}} = \frac{2\pi}{\vartheta_0} + \frac{\varphi\{H_1(e^{j\vartheta_0})\}}{\vartheta_0} + \frac{\varphi\{H_d(e^{j\vartheta_0})\}}{\vartheta_0}. \quad (2.67)$$

This was derived from Eq. (2.60) for the fundamental frequency, i.e., $k = 1$. The prescribed length of the waveguide D_0 is not an integer. The solution is to use fractional delay filters [Välimäki 1995; Laakso et al. 1996] in series with the delay line. The simplest choice is the first-order allpass filter, as proposed in [Smith 1983; Jaffe and Smith 1983]. The integer part of D_0 will be implemented as the delay line of the digital waveguide with a length of $N = \lfloor D_0 - 0.5 \rfloor$, and the fractional part will be realized with the first-order allpass filter with a phase delay of $D_{\text{fd}} = D_0 - N$ at ϑ_0 . The resulting fractional delay will be $0.5 \leq D_{\text{fd}} < 1.5$ 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 $a_1 = (1 - D_{\text{fd}})/(1 + D_{\text{fd}})$ [Smith 1983; Jaffe and Smith 1983; Välimäki 1995].

2.3.3 Modal-based Approach

Modal synthesis is based on the fact that any vibrating mechanical system can be decomposed into a set of mass-spring-damper models. This is analogous to decomposing the transfer function of the system into parallel second-order sections. This approach has been used by Adrien [1991]. A new variation of modal synthesis is the Functional Transformation Method [Trautmann and Rabenstein 1999, 2003], where the continuous-time impulse response of the vibrating system is computed by using the Laplace and the Sturm-Liouville transform, resulting in a set of exponentially decaying sinusoids for linear systems. This is then implemented by resonators in discrete time.

Here we will simply discretize the results of Sec. 2.2.4, where Eq. (2.34) describes the motion of the different string modes. Discretizing Eq. (2.34) gives an algorithm that can be directly applied to sound synthesis. Naturally, discretizing the impulse response Eq. (2.36) gives the same result. The only difference is that Eq. (2.34) is parameterized by the physical parameters of the strings, while in Eq. (2.36) the modal frequencies ω_k and decay times τ_k are the free parameters.

The discretization with respect to time can be done by various methods, but as the impulse response of the modes, Eq. (2.36), can be considered as a band-limited signal, it is sufficient to use the impulse invariant transform. However, we have to take care of not implementing any modes having the frequency which is near or above Nyquist rate $f_s/2$. Accordingly, we are looking for the discrete-time system that has the impulse response

$$y_{\delta,k}(t_n) = \frac{1}{\pi L \mu f_k} \frac{1}{f_s} e^{-\frac{t_n}{\tau_k}} \sin(2\pi f_k t_n), \quad (2.68)$$

where $t_n = n\Delta t$, $\Delta t = 1/f_s$ being the sampling interval. Equation (2.68) differs from Eq. (2.36) in that now frequencies $f_k = \omega_k/(2\pi)$ are used and in that it is scaled by a factor of $1/f_s$. This scaling is required because the discrete time dirac impulse has an area of $1/f_s$, while the analog dirac impulse has unity area.

In general, an exponentially decaying cosine with arbitrary initial amplitude A_k and phase φ_k is written as

$$c_k(t_n) = A_k e^{-\frac{t}{\tau_k}} \cos(2\pi f_k t_n + \varphi_k) = e^{-n\frac{1}{\tau_k f_s}} \frac{C_k e^{j2\pi n \frac{f_k}{f_s}} + C_k^* e^{-j2\pi n \frac{f_k}{f_s}}}{2} \quad (2.69)$$

where C_k is the complex initial amplitude, whose phase $\varphi\{C_k\}$ determines the initial phase φ_k and its absolute value $|C_k|$ equals the initial amplitude A_k of the decaying cosine function. The asterisk $*$ stands for complex conjugation. Taking the z-transform of $c_k(t_n)$ gives

$$Z\{c_k(t_n)\} = Z\left\{\frac{C_k}{2} p_k^n + \frac{C_k^*}{2} p_k^{*n}\right\} = \frac{1}{2} \left(\frac{C_k}{1 - p_k z^{-1}} + \frac{C_k^*}{1 - p_k^* z^{-1}} \right) \quad (2.70)$$

$$p_k = e^{j2\pi \frac{f_k}{f_s}} e^{-\frac{1}{\tau_k f_s}}, \quad (2.71)$$

where p_k and p_k^* are the poles of the two complex resonators. After some algebraic trans-

formations we obtain the second-order transfer function

$$H_{\text{res},k}(z) = \frac{b_{0,k} + b_{1,k}z^{-1}}{1 + a_{1,k}z^{-1} + a_{2,k}z^{-2}} \quad (2.72a)$$

$$b_{0,k} = \text{Re}\{C_k\} \quad (2.72b)$$

$$b_{1,k} = -\text{Re}\{C_k p_k^*\} \quad (2.72c)$$

$$a_{1,k} = -2\text{Re}\{p_k\} \quad (2.72d)$$

$$a_{2,k} = |p_k|^2. \quad (2.72e)$$

In the case of Eq. (2.68) the parameters C_k take the values

$$C_k = -j \frac{1}{\pi L \mu f_k} \frac{1}{f_s}. \quad (2.73)$$

It can be seen from Eqs. (2.73) and (2.72b) that $b_{0,k} = 0$, as the real part of C_k is zero. Accordingly, each mode is implemented by a two-pole no-zero discrete-time system.

The computation of the string response is as follows: first, the excitation force $F_{y,k}(t_n)$ of mode k is computed by the scalar product of Eq. (2.35d). If the excitation force $F_{\text{exc}}(t_n)$ is concentrated to a mathematical point along the string at x_{exc} , it only leads to a scaling by the modal shapes

$$F_{y,k}(t_n) = \sin\left(\frac{k\pi x_{\text{exc}}}{L}\right) F_{\text{exc}}(t_n). \quad (2.74)$$

These forces $F_{y,k}(t_n)$ are the input signals of the resonators $H_{\text{res},k}(z)$ calculating the instantaneous amplitudes $y_k(t_n)$ of the modes. Then, the displacement of the string can be computed at an arbitrary point $0 \leq x \leq L$ along the string by

$$y(x, t_n) = \sum_{k=1}^K y_k(t_n) \sin\left(\frac{k\pi x}{L}\right) \quad (2.75)$$

where K is the number of simulated modes. Note that the position of both the observation point x and the excitation x_{exc} are arbitrary, as the modal approach does not discretize the space variable, on the contrary to finite-difference and digital waveguide modeling already discussed in Secs. 2.3.1 and 2.3.2.

The output of the string model is generally the force at the termination, which is computed by

$$F_b(t_n) = -T_0 \frac{\partial y}{\partial x} \Big|_{x=L} = -\frac{T_0 \pi}{L} \sum_{k=1}^K y_k(t_n) k (-1)^k. \quad (2.76)$$

Naturally, the input and output scaling factors can be incorporated into the $b_{1,k}$ parameter of the resonator. As an example, if the string is excited at the position x_{exc} by a force $F_{\text{exc}}(t_n)$ and its output is the termination force $F_b(t_n)$, then $b_{1,k}$ becomes

$$b_{1,k} = \text{Re}\{C_k p_k^*\} \frac{T_0 \pi}{L} k (-1)^k \sin\left(\frac{k\pi x_{\text{exc}}}{L}\right). \quad (2.77)$$

This simplifies the computation, as now all the resonators $H_{\text{res},k}$ have the same input signal $F_{\text{exc}}(t_n)$, and the force at the termination $F_b(t_n)$ is calculated by simply summing the outputs of the resonators. Note that now these outputs have no direct physical interpretation, as they are the scaled versions of $y_k(t_n)$.

A nice feature of the modal approach is that the modal frequencies, amplitudes and decay times can be controlled individually. This is advantageous when the goal is the perfect resynthesis of a specific instrument tone, as the measured data can be directly uploaded to the model. However, some of the physicality is lost. As an example, coupling of different strings is very straightforward with both finite-difference modeling and digital waveguides, but here it becomes more complicated. The computational complexity of the modal approach is somewhere between the finite-difference and digital waveguide models.

2.4 Excitation Modeling

The string and body models are of the same structure for the different string instruments, although they are parameterized in a different way for the various instruments. On the contrary, for modeling the excitation, different model structures have to be developed. This is because the excitation mechanisms of the instruments are completely different, and their precise implementation is essential for rendering the sonic characteristics of these instruments.

2.4.1 Struck Strings

As the simplest approach, striking can be modeled by setting the initial velocity of the string at the excitation point to a given value, while the initial displacement is zero [Smith 1992]. This is based on the assumption that the hammer is in contact with the string for an infinitesimally short duration, which could happen only if the hammer mass is negligible in comparison to the string mass. However, this is not the case in real instruments, so more elaborated models are needed. As the piano is the most important struck stringed instrument, we will concentrate on modeling the piano hammer.

The piano string is excited by a hammer, whose initial velocity is controlled by the player with the strength of the touch on the keys. The excitation mechanism of the piano is as follows: the hammer hits the string, the hammer felt compresses and feeds energy to the string, then the interaction force pushes the hammer away from the string. Accordingly, the excitation is not continuous, it is present for some milliseconds only. The hardwood core of the hammer is covered by wool felt, whose structure is not homogeneous. This is the reason why playing harder on the piano results not only in a louder tone, but also in a spectrum with stronger high frequency content [Fletcher and Rossing 1998, pp. 367].

The piano hammer is generally modeled by a small mass connected to a nonlinear

spring [Boutillon 1988]. The equations describing the interaction are as follows:

$$F_h(t) = F(\Delta y) = \begin{cases} K_h(\Delta y)^{P_h} & \text{if } \Delta y > 0 \\ 0 & \text{if } \Delta y \leq 0 \end{cases}, \quad (2.78a)$$

$$F_h(t) = -m_h \frac{d^2 y_h(t)}{dt^2}, \quad (2.78b)$$

where $F_h(t)$ is the interaction force, $\Delta y = y_h(t) - y_s(t)$ is the compression of the hammer felt, $y_h(t)$ is the position of the hammer, and $y_s(t)$ is the position of the string at the excitation point x_{exc} (i.e., $y_s(t) = y(x_{\text{exc}}, t)$). The hammer mass is referred by m_h , K_h is the hammer stiffness coefficient, and P_h is the stiffness exponent.

These equations can be easily discretized with respect to time. However, as seen from Eqs. (2.78a) and (2.78b), there is a mutual dependence between $F_h(t)$ and $y_h(t)$, i.e., for the calculation of one of these variables, the other should be known. This is generally overcome by the assumption that the hammer force changes a little during one time step, that is $F_h(t_n) \approx F_h(t_{n-1})$. Although it may lead to numerical instabilities for high impact velocities, the straightforward approach is often used in the literature [see, e.g., Chaigne and Askenfelt 1994]. The numerical instabilities can be avoided by rearranging the nonlinear equations to known and unknown terms [Borin et al. 2000], or, by the multi-rate method presented in Sec. 4.1.

We note that piano hammers are not fully characterized by the model of Eq. (2.78), as the felt has a hysteretic behavior. The hysteresis has been described by changing the parameter P_h in Eq. (2.78) by Boutillon [1988]. A more elaborated model was presented by Stulov [1995], which have shown good agreement with measurements when the hammer is bouncing into a rigid object. However, the model of Stulov [1995] does not provide accurate results when compared to real hammers striking piano strings [Giordano and Winans II 2000]. To date, no precise hammer models exist. However, it seems that the ear is less sensitive to these variations, mostly because the short duration (1–2 ms) of the hammer–string contact. Accordingly, implementing Eq. (2.78) already produces good sound and the addition of the hysteretic model of Stulov [1995] only introduces a perceptual effect comparable to lowpass-filtering. Nevertheless, hammer models based on [Stulov 1995] have been presented in [Borin and De Poli 1996; Giordano and Jiang 2004].

Although not being a physical approach, it is also possible to measure or precompute the temporal function of the excitation $F_h(t)$, store it in a memory, and lead this into the string model when a note is played. Naturally, this can also be done for other types of excitations, such as plucking. This kind of hybrid approach has been used in [Bensa et al. 2004], where the strings were modeled by coupled digital waveguides. The source (hammer force) was extracted by inverse filtering the measured string signal with the transfer function of the string model. Then, the excitation force was modeled by a static part and a filter that changes the frequency content of the excitation as a function of dynamic level. Note that the disadvantage of such an approach is that the features of a real physical model (e.g., it responds naturally to the restrike of a string still in vibration) is lost.

2.4.2 Plucked Strings

To the first approximation, when a string is plucked, it is pulled slowly by the finger or the plectrum at a certain point and then it is suddenly released. This means that plucking can be modeled by setting the initial displacement of the string to a triangle shape, which is zero at the terminations and has a maximum at the plucking point [Hiller and Ruiz 1971a]. If the string is plucked at the position $0 \leq x_{\text{exc}} \leq L$ with the amplitude A_p at the plucking point, the initial conditions can be written as

$$y(x, t) = \begin{cases} A_p \frac{x}{x_{\text{exc}}} & \text{if } 0 \leq x \leq x_{\text{exc}} \\ A_p \frac{L-x}{L-x_{\text{exc}}} & \text{if } x_{\text{exc}} < x \leq L \end{cases}, \quad (2.79a)$$

$$\left. \frac{\partial y(x, t)}{\partial t} \right|_{t=0} = 0. \quad (2.79b)$$

This is very simple from the sound synthesis point of view as there is no need for an additional dynamic system for excitation modeling. For finite-difference models and digital waveguides, Eq. (2.79) is implemented as is, while for modal models the Fourier series of the initial displacement is computed and the initial amplitudes of the modes are set accordingly. When the Fourier analysis is made, one finds that those modes that have a node at the plucking point are not excited, meaning that if L/x_{exc} is an integer, then every L/x_{exc} th mode is missing from the spectrum [see, e.g., Fletcher and Rossing 1998, p. 41]. Note that this is also true for other types of excitations. In the case of digital waveguides it is possible to approximate the plucking as a pair of impulses, if the wave variable is acceleration [Smith 1992]. However, for the plucking model of Eq. (2.79), the only two parameters are the excitation point x_{exc} (controlling the missing harmonics) and the plucking amplitude A_p (setting the overall amplitude), which is quite poor compared to the freedom what a guitar player has during plucking. Especially if we notice that the role of A_p is to set the overall amplitude of the modes, thus, it has no influence on the spectral content.

A more realistic model is a modified hammer model presented in [Borin et al. 1992], where the dynamics of the plectrum is written as Eq. (2.78b), while for the nonlinear plucking force Eq. (2.78a) is modified in a way that above a certain Δy level, the force $F(\Delta y)$ drops to zero. This corresponds to the point when the plectrum releases the string. The plucking model is implemented in the same way as the hammer model of Eq. (2.78), that is, it is discretized and connected the string model at the excitation point x_{exc} . A more elaborated plucking model has been presented in [Cuzzucoli and Lombardo 1999], where the finger is modeled as a dynamic system including its damping effect. By varying the model parameters and the shape of the external force (the force with which the player moves its finger), different kind of plucking sounds could be synthesized. For the harpsichord, a simple plucking model has been presented in [Giordano and Winans II 1999]. To conclude, an advantage of the plucking model approach that the model has more degrees of freedom, but it requires the additional computation of the excitation model. Note that this additional complexity is generally negligible compared to string modeling as these models are simple lumped systems and they run only the very first part of the tone (i.e., during

excitation).

2.4.3 Bowed Strings

In the case of bowed instruments the excitation comes from the friction between the string and the bow hairs. The bow, moving perpendicular to the string, grips the string (sticking phase). Due to the increasing displacement of the string, the elastic returning force is also increasing until its level reaches the sticking friction. At this point the bow releases the string, the string swings back (slipping phase) and then vibrates freely. This vibration is damped partly by the own losses of the string and partly by the slipping friction that develops between the string and the bow hairs. This state lasts as long as the bow grips the string again, which occurs only when the velocity of the bow and the string equals. In this case, their relative velocity is zero, the frictional force is maximal. This alteration of the stick and slip phases is the so-called Helmholtz motion. The excitation is periodical and generates a sawtooth shape vibration [see, e.g., Fletcher and Rossing 1998, p. 47].

In [Hiller and Ruiz 1971a] a simplified model of bowing was used assuming that the string either moves along with the bow or it slips away with negligible friction, i.e., vibrates freely. The switching between these two states was controlled by comparing the string slope $|\partial^2 y / \partial x^2|$ with a precomputed constant.

McIntyre et al. [1983] proposed a more physical algorithm later named the “MSW algorithm”, which is the most often used bowing model in the literature. In the MSW algorithm the dependence of the excitation force on the velocity difference of string and bow is described by a viciously nonlinear function. For zero velocity difference (sticking phase), the friction force can increase until a certain maximal value (the static friction), then the dynamic friction decreases as the velocity difference increases (slipping phase), similarly to what is found in reality. The method is well suited for the waveguide formalism, as it operates with incoming and reflected waves. It computes the required velocity correction of the string for a given incoming string velocity and bow velocity by finding the intersection of the nonlinear velocity-force curve of the interaction and the straight line determined by the impedance of the string. Smith [1986] presented a simplified implementation of the MSW algorithm by the application of a signal-dependent reflection function at the bowing point. The reflection coefficient is read from a look-up table as a function of the relative velocity of string and bow, eliminating the need of an equation solver.

The models based on a memoryless relation of the relative velocity and interaction force (such as the MSW algorithm) render the most important features of the bow-string interaction. However, for some secondary effects, such as hysteresis, more sophisticated models are needed. Numerous bow modeling techniques are reviewed in [Serafin 2004], including dynamic friction models where the contact force is modeled by a nonlinear differential equation.

2.5 Instrument Body Modeling

Here we treat the instrument body and the bridge (a wooden piece which transmits the string vibration to the body) as one subsystem. Thus, body modeling refers to modeling the bridge-body system of the instrument. The effect of the instrument body on the sound is twofold. Its main role is to lead the vibration of the string to the air, producing the sound pressure that can be heard. In most of the cases this effect can be characterized by a frequency-dependent transfer function, describing the sound pressure at a given location as a function of the string force acting on the instrument body. The other effect of the body is that it also influences the string vibration itself. This is because the bridge is not infinitely rigid, it vibrates according to the normal modes of the bridge-body system, providing a frequency-dependent terminating impedance to the string. This influences the decay times and the modal frequencies of the string vibration and produces a coupling between the different strings.

Two main approaches can be distinguished between body modeling techniques, depending on which one of these effects is implemented. In the physically accurate case, both the radiation and impedance characteristics are modeled. The input of the body model is the string force, and the outputs are the bridge velocity acting back on the string and the sound pressure at a given location in space. On the other hand, when the body is implemented as a post-processing technique, only the sound pressure is computed as a function of string force. Thus, the impedance effect (the feedback to the string) is not modeled with this approach but is taken into account in the string model. The advantage of the post-processing technique over the physics-based one is that the model structure becomes simpler. Moreover, the parameter estimation is also simplified as now the modal frequencies and decay times are influenced by the string model only. Conversely, the physics-based body modeling approach is closer to reality and automatically takes into account some interesting second-order effects, such as the coupling of different strings.

For the string and the excitation, one of the main motivation for using the physics-based approach is that it takes into account the interaction of the musician in a meaningful way as the input parameters are physical variables, such as string length or bow velocity. This factor has no importance in deciding whether a physics-based or a post-processing type body model should be used, as the player cannot influence the parameters of the instrument body in real instruments. Naturally, the physics-based approach might have its benefits for producing never heard synthetic sounds. As an example, such an approach can be made capable of producing a sound of a virtual instrument whose body is continuously changing its shape.

2.5.1 Physics-based Modeling

In physics-based body modeling, both the impedance and radiation properties of the bridge-body subsystem are modeled. A common feature of these models that they are generally parameterized by physical dimensions and material properties of the instrument body, and not by transfer function measurements. This way, these models can be used to predict how

the sound of a real instrument would change by having different shape or being of different material.

A straightforward choice for modeling the instrument body is finite-difference modeling, as here the continuous time equations can be directly implemented in discrete time. One nice example is the piano model of Giordano and Jiang [2004], where the piano soundboard is modeled by a finite-difference plate model. This model takes into account the effect of the ribs and the bridge by varying the thickness of the board. The soundboard is excited by the string force, and it acts back to the string through the bridge movement. An interesting feature of the model is that its output is taken from a three-dimensional finite-difference room model, which is excited by the soundboard movement. A finite-difference guitar model has been presented by Bader [2003]. The novelty of the approach is that the nonlinear coupling of the transverse and longitudinal waves in the guitar top-plate is also modeled, together with the vibration of the air cavity. Unfortunately, these finite-difference body models are computationally so demanding that they cannot be run in real-time.

The instrument body might be also modeled by the modal approach. Woodhouse [2004a,b] presented a guitar model where the vibration of the body has been simulated by a modal model, whose parameters were fitted from real measurements. In [Cuzzucoli and Lombardo 1999] a much simpler model has been used that describes the guitar body with three modes (the first resonance of the table and the air, and the Helmholtz resonance), yielding an approximation valid only for the low frequencies.

2.5.2 Post-processing Techniques

In this approach, the radiation effect of the instrument body is taken into account as a linear filtering operation upon the signal coming from the string model. Here the impedance effects of the body upon the string (the alteration of modal frequencies and decay times, the coupling of different strings) are implemented in other parts of the model (i.e., in the string). This way the modeling problem reduces to filter design. Unfortunately, the transfer function of real instrument bodies exhibits high modal density, making difficulties for standard filter design algorithms. For high quality sound, high-order filters are needed. Their computational complexity can be 10 or 100 times higher than that of needed for a digital waveguide based string model. As an example, the pressure-force transfer function of a piano soundboard is shown in Fig. 2.5. The soundboard was excited by hitting the bridge with an impact hammer. The excitation force and the sound pressure at 2 m distance from the piano were simultaneously recorded. The ratio of their spectra is depicted in this figure. Ideally, the last block of Fig. 2.1 (p. 8) should have similar transfer function to the one displayed here in Fig. 2.5. The different filtering techniques will be reviewed in Sec. 4.3.

A further difficulty of the filter-design based post-processing approach lies in the fact that the measurement of body impulse responses is a complicated task, especially at high frequencies (above 5–10 kHz). This is because impulse hammers generate an excitation pulse that is of lowpass character, leading to a measured response where the high frequency part is more noisy, i.e., less reliable. This might be avoided by using periodic excitations

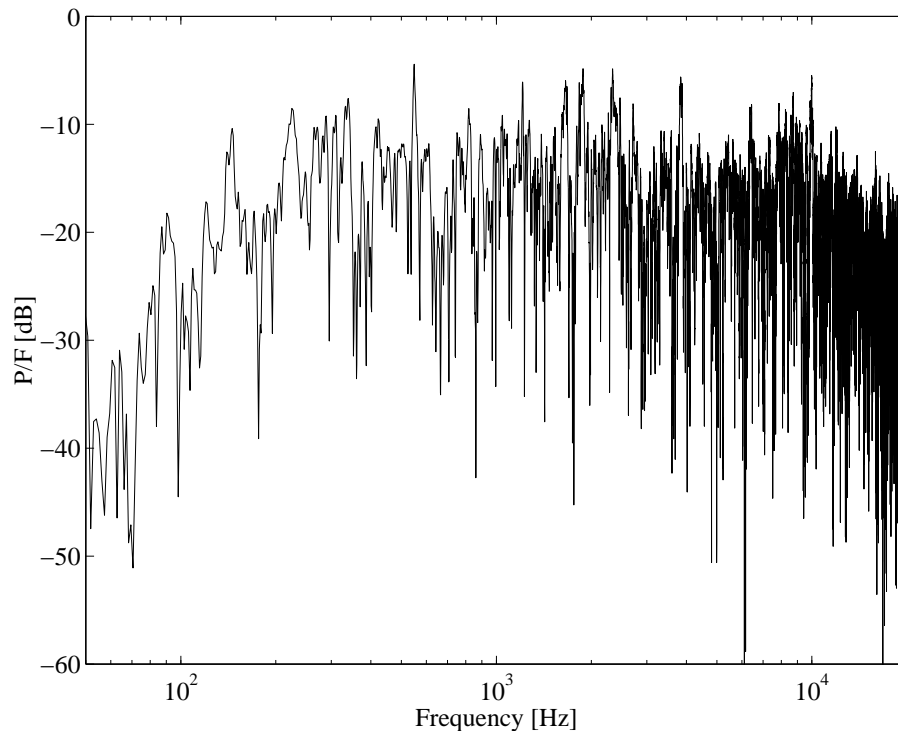


Figure 2.5: The force-pressure transfer function of a piano soundboard.

generated by shakers. However, it is much harder to properly fit a shaker to the instrument body than hitting the bridge by a hammer. The discussion of the measurement methods is out of the scope of this thesis. We only note that the impulse responses obtained by force-hammer excitation seem to provide perceptually acceptable responses. The most probable reason for this is that the less reliable high-frequency part of the response is excited by the string at a smaller amplitude (the string signal is also of lowpass character) and that the ear is relatively insensitive to the variation of modal parameters at high frequencies.

The computationally most efficient way of implementing the effect of the body filtering is to model it by a reverberation-like algorithm. As an example, in the model of Garnett [1987] several digital waveguides were coupled together to give a high-density impulse response. Borin et al. [1997] suggested the use of feedback delay networks for modeling the force-velocity response of the piano soundboard, which could be used for modeling the coupling of different strings, too. A difficulty of these reverberator-based approaches is that only the statistical distribution and the overall damping of the body normal modes can be set by the available parameter estimation techniques. Therefore, the sound quality of these algorithms is usually inferior to the filter-based techniques, but they require less computational power.

2.5.3 Commuted Synthesis

To avoid the problems of filter design and high computational demand, the commuted synthesis technique was presented in [Karjalainen and Välimäki 1993; Smith 1993]. This is based on the idea that if all the elements of Fig. 2.1 (p. 8) are linear and time-invariant, and the feedback from the string to the excitation and from the body to the string is neglected (i.e., the dashed arrows are deleted from Fig. 2.1), the system reduces to three linear filters in series. The order of these filters can be commuted, which results in a structure of Fig. 2.6.

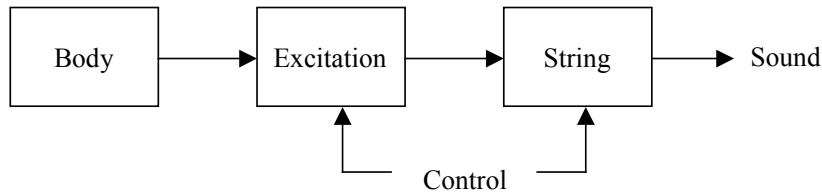


Figure 2.6: Commuted synthesis.

As the input signal of the model is a unit pulse, the body filter does not have to be implemented. Its impulse response can be stored in a wavetable whose content is simply fed to the string. The effect of excitation filtering can also be taken into account in the wavetable. In this case, different wavetables have to be stored for the different excitation types for all the notes. If the excitation table is computed by filtering the recorded sound with the inverse of the transverse function of the string model, then the output of the model equals the original up to the length of the excitation table. This is extremely useful when the goal is the precise resynthesis of the recorded instrument sound. Actually, in this case commuted synthesis can be considered as a tricky sampling algorithm, which reproduces the first part of the sound precisely, then it repeats one period of sound and filters it periodically to produce the approximate frequency and decay of the partials. Astonishing results have been achieved by this method in the case of the acoustic guitar at relatively low computational complexity [Välimäki et al. 1996; Välimäki and Tolonen 1998].

The drawback of the method is that the excitation model loses its physicality. Now the excitation model is either the part of the wavetable, or implemented as a linear filter. Thus, it is controlled by either switching between wavetables, or changing filter coefficients, rather than varying a physical parameter such as plucking force. The bidirectional interaction of the excitation and the string cannot be implemented anymore, therefore some effects, such as the restrike of a string, cannot be implemented this way. Moreover, complex nonlinear excitations, such as bowing, cannot be linearized. Although an efficient commuted violin model exists [Smith 1997], it models the main feature of violin excitation only. More importantly, when the string model itself is nonlinear (see Chap. 6), then commuted synthesis cannot be applied. Nevertheless, for such applications, where low complexity is more important than sound quality (multimedia, computer games), the commuted synthesis with linear string models is still the best option.

2.6 Conclusion

This chapter has outlined the most common strategies for modeling the different parts of string instruments.

For string modeling, three main approaches have been described. The first one, finite-difference modeling, is the direct discretization of the wave equation. Its main advantage is the direct connection to the continuous-time equation, making it possible to model such complicated situations as spatially nonuniform tension (this will be exploited in Sec. 6.4). On the other hand, it has the largest computational complexity among the three main string modeling approaches.

The discretization of the traveling-wave solution of the wave equation leads to digital waveguide modeling. The losses and dispersion of the string are lumped to one point, resulting in a delay line and a filter in a feedback loop. This allows very efficient implementation. Indeed, for linear string models, the digital waveguide seems to be the best choice as it results in the lowest computational complexity for a given quality requirement. Moreover, it has a strong connection to physical reality, as it operates with incident and reflected waves, making the coupling of the different strings easy to implement. On the other hand, it can model nonlinear string behavior with certain limitations (see Sec. 6.3.6).

The third technique is based on the discretization of the modal solution of the wave equation. Consequently, the string motion is computed by a set of second-order resonators, implementing the normal modes of the string. The computational complexity of the approach is between the finite-difference and digital waveguide techniques. The main advantage of the technique is that its parameters can be easily set in a way to mimic a recorded tone. On the other hand, it only has an indirect connection to physical reality, making some second-order effects (coupling of strings, temporal variation of tension) harder to implement. In any case, when only a small number of modes have to be implemented (as for modeling the longitudinal motion in Chap. 6), the modal approach is the most efficient.

The main difference between the various string instruments lies in the excitation mechanism. Therefore, the effects and modeling techniques of striking, plucking, and bowing have been outlined in Sec. 2.4. In general, the excitation is modeled by the discrete-time implementation of a one-dimensional, nonlinear differential equation.

Then, the most common techniques for modeling the instrument body have been described. The physics-based approach models the motion of the different parts of the instrument body and is generally parameterized by the geometric and material properties of the instrument. Therefore, these models are well suited for experimentation (e.g., how the shape of the body affects the sound), and they model the coupling of the instrument parts in a physically meaningful way.

On the other hand, the post-processing technique treats the body as a “black box” model, copying the sound pressure response of the original system for a given input force. Generally, post-processing techniques provide better sound quality as their parameters are estimated from the measurement of real transfer functions. Moreover, they can be efficiently implemented in DSPs, as they are implemented by digital filters.

A special case of instrument body modeling is the commuted synthesis technique, where

the impulse response of the body is stored in a wavetable, read sample by sample, and fed to the string. The commutation of body, excitation and string can be done in the case of linear string and excitation models. In this case, this is the most efficient body modeling technique, as the body model is reduced to reading samples from a memory. However, the technique cannot be applied for nonlinear excitation and string models. In that case the model had to be linearized by converting the time-invariant nonlinear elements to linear ones with varying parameters. As a result, some of the secondary effects of real instruments, such as the restrike of a string and the modulation of tension, cannot be modeled.

As we have seen, various approaches are available for physics-based sound synthesis of string instruments. The choice mainly depends on the quality requirement and the available computational power. Another factor is that whether the purpose is to experiment by changing the geometric and material properties of the instrument or to realize a sound synthesis model whose sound resembles to a given instrument. On the whole, no “perfect solution” exists, it is always a compromise between the given requirements.

Chapter 3

Loss Filter Design for the Digital Waveguide

This chapter presents new techniques for decay time-based loss filter design. First the case of the one-pole filter is discussed, which is the most commonly used loss filter in the literature. Approximate equations are given that relate the parameters of the one-pole filter to the coefficients of the lossy wave equation. Based on this relation, a simple and efficient filter design technique is presented, which applies weighted polynomial regression. In those cases where the goal is the precise resynthesis of a given sound, the one-pole filter might not provide sufficient accuracy. Therefore, a simple and robust design method is proposed for high-order loss filters. Both algorithms minimize the error of decay times instead of the error of the magnitude response, as suggested in [Smith 1983]. Another common feature is that both techniques are based on mean squares optimization, leading to robust and simple implementations.

3.1 The One-pole Loss Filter

As already mentioned in Sec. 2.3.2, the one-pole filter is the most commonly used loss filter in digital waveguide models. Despite its simplicity, 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. This was the main motivation to find the connection between the digital waveguide with a one-pole loss filter and the wave equation of the lossy string. After presenting the approximate equations for the decay times, a robust filter design technique is proposed.

3.1.1 Approximate Formulas for the Decay Times

The magnitude response of the one-pole filter $H_{1p}(z)$ of Eq. (2.63) is:

$$|H_{1p}(e^{j\vartheta})| = g \frac{1 + a_1}{\sqrt{a_1^2 + 1 + 2a_1 \cos \vartheta}} \quad (3.1)$$

From Eq. (2.61) decay times τ_k of the partials produced by a digital waveguide with such a filter are

$$\tau_k = -\frac{1}{f_0 \ln |H_{1p}(e^{j\vartheta_k})|} \approx \frac{1}{f_0(1 - |H_{1p}(e^{j\vartheta_k})|)}, \quad (3.2)$$

where f_0 is the fundamental frequency of the note, and for the logarithm function the first-order Taylor-series approximation was used. Note that here we assume $f_k = kf_0$ and neglect the effect of inharmonicity (originally included in Eq. (2.61)). The effect of inharmonicity is discussed at the end of this section.

From Eqs. (3.1) and (3.2) the decay rate $\sigma_k = 1/\tau_k$ is expressed as

$$\begin{aligned} \sigma_k &\approx f_0(1 - |H_{1p}(e^{j\vartheta_k})|) = f_0 \left(1 - g \frac{1 + a_1}{\sqrt{a_1^2 + 1 + 2a_1 \cos \vartheta_k}} \right) = \\ &= f_0 \frac{\sqrt{a_1^2 + 1 + 2a_1 \cos \vartheta_k} - g(1 + a_1)}{\sqrt{a_1^2 + 1 + 2a_1 \cos \vartheta_k}}. \end{aligned} \quad (3.3)$$

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

$$\sigma_k \approx f_0 \frac{\sqrt{(a_1 + 1)^2 - a_1 \vartheta_k^2} - g(1 + a_1)}{\sqrt{(a_1 + 1)^2 - a_1 \vartheta_k^2}} = f_0 \frac{\sqrt{1 - \frac{a_1}{(a_1 + 1)^2} \vartheta_k^2} - g}{\sqrt{1 - \frac{a_1}{(a_1 + 1)^2} \vartheta_k^2}}. \quad (3.4)$$

By noting that denominator is close to 1 and $\sqrt{1 + x} \approx 1 + x/2$ for $x \approx 0$ we obtain

$$\sigma_k \approx f_0 \left(\sqrt{1 - \frac{a_1}{(a_1 + 1)^2} \vartheta_k^2} - g \right) \approx f_0 \left((1 - g) - \frac{a_1}{2(a_1 + 1)^2} \vartheta_k^2 \right). \quad (3.5)$$

Finally, the decay times of the digital waveguide with a one-pole loss filter will be

$$\tau_k = \frac{1}{\sigma_k} \approx \frac{1}{c_1 + c_3 \vartheta_k^2} \quad (3.6a)$$

$$c_1 = f_0(1 - g) \quad (3.6b)$$

$$c_3 = -f_0 \frac{a_1}{2(a_1 + 1)^2}, \quad (3.6c)$$

where ϑ_k is the digital angular frequency. The approximation is 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].

Replacing ϑ_k with the angular frequency $\omega_k = f_s \vartheta_k$ gives

$$\tau_k = \frac{1}{\sigma_k} \approx \frac{1}{b_1 + b_3 \omega_k^2} \quad (3.7a)$$

$$b_1 = f_0(1 - g) \quad (3.7b)$$

$$b_3 = -\frac{f_0}{f_s^2} \frac{a_1}{2(a_1 + 1)^2}. \quad (3.7c)$$

Equation Eq. (3.7a) is similar to the decay time of a wave equation with the simplest frequency dependent losses, where the b_1 and b_3 coefficients correspond to the first- and third-order time derivatives of the wave equation (see Eqs. (2.46) and (2.47)).

Equation (3.7) can be reformulated to

$$\tau_k = \tau_0 \frac{1}{1 + \left(\frac{f_k}{f_{\tau/2}}\right)^2}, \quad (3.8)$$

where τ_0 is the decay time at DC (which is near to the decay time of the fundamental) and $f_{\tau/2}$ is the frequency where the decay time decreases to the half compared to τ_0 . The parameters of Eqs. (3.7) and (3.8) are related by $b_1 = 1/\tau_0$ and $b_3 = b_1/(2\pi f_{\tau/2})^2$. Equation (3.8) shows that the decay times τ_k evolve as a function of frequency f_k similarly to the magnitude response of a second-order filter, i.e., they fall at a rate of 12 dB/decade above $f_{\tau/2}$. It is interesting to note that a first-order filter leads to a second-order type behavior with respect to decay times.

The form of Eq. (3.8) is particularly advantageous for the real-time control of string decay, as τ_0 and $f_{\tau/2}$ are more meaningful than the filter coefficients g and a_1 or the loss parameters b_1 and b_3 of Eq. (3.7a). The overall decay time is set by τ_0 , while $f_{\tau/2}$ determines the frequency where the partial decay times start to fall.

So far we have assumed that the tone produced by the string model is perfectly harmonic, i.e., $f_k = kf_0$. If the model contains a dispersion filter $H_d(z)$ introducing significant inharmonicity, the derived equations Eqs. (3.2)-(3.7) have to be corrected: f_0 has to be replaced by f_k/k in every places. As a result, the decay times obtained by Eqs. (3.6)-(3.8) has to be multiplied by $(kf_0)/f_k$, which is the reciprocal of the inharmonicity index

$$I_k = \frac{f_k}{kf_0}. \quad (3.9)$$

The inharmonicity index I_k equals unity for perfectly harmonic sounds and it is around 1.05–1.25 for the 30th partial of the piano [see Bank 2000b, Fig. 2.4 on p. 25]. Accordingly, Eq. (3.6a) becomes

$$\tau_k = \frac{1}{\sigma_k} \approx \frac{1}{I_k(c_1 + c_3 v_k^2)}, \quad (3.10)$$

while c_1 and c_3 are still computed as in Eqs. (3.6b) and (3.6c). In the next section we will see how the “inverses” of Eqs. (3.6) and (3.10) can be used for filter design.

3.1.2 Filter Design Based on Polynomial Regression

In the literature, the two parameters of the one-pole loop filters are set by ad-hoc algorithms. In [Välämäki et al. 1996] the DC gain g was set according to the decay time of the first few partials. The pole of the filter was determined by continuously adjusting a_1 and searching for the minimum of the approximation error. The magnitude error was computed in a least squares sense, by using a weighting function $1/(1 - g_k)$ putting more emphasis on slowly decaying partials. However, the overall decay time is not always matched by this algorithm. This algorithm has been extended by a complicated nonlinear optimization in [Erkut et al. 2000], which is based on the amplitude envelope of the synthesized and the original signal.

Based on the results of the previous section, a simple and robust design method is presented that overcomes these problems. Equation (3.10) shows that the decay rate σ_k is a second order polynomial of ϑ_k , multiplied by the inharmonicity index series I_k . This is the basis of filter design: c_1 and c_3 are estimated by polynomial regression from the measured decay rates. Then, the g and a_1 coefficients of the one-pole filter can be easily calculated from c_1 and c_3 by the inverses of Eqs. (3.6b) and (3.6c).

However, it is perceptually more meaningful to minimize the mean-square error of the decay times instead of the error of decay rates. The expression of the decay-time error e_τ is:

$$e_\tau = \sum_{k=1}^K (\hat{\tau}_k - \tau_k)^2 = \sum_{k=1}^K \hat{\tau}_k^2 \tau_k^2 \left(\frac{1}{\hat{\tau}_k} - \frac{1}{\tau_k} \right)^2 = \sum_{k=1}^K \hat{\tau}_k^2 \tau_k^2 (\hat{\sigma}_k - \sigma_k)^2 \quad (3.11)$$

where $\sigma_k = 1/\tau_k$ are the prescribed, and $\hat{\sigma}_k = 1/\hat{\tau}_k$ are the approximated decay rates. Inserting Eq. (3.10) into Eq. (3.11) gives

$$e_\tau = \sum_{k=1}^K \hat{\tau}_k^2 \tau_k^2 [I_k(c_1 + c_3 \vartheta_k^2) - \sigma_k]^2 = \sum_{k=1}^K w_k \left(c_1 + c_3 \vartheta_k^2 - \frac{\sigma_k}{I_k} \right)^2 \quad (3.12)$$

where ϑ_k are the angular frequencies of the partials, $w_k = \tau_k^2 \hat{\tau}_k^2 / I_k^2$ are the weights, and e_τ is the approximation error which should be minimized with respect to the parameters c_1 and c_3 . Note that the weights w_k can be approximated by $w_k = \tau_k^2 \hat{\tau}_k^2$, which is neglecting the term $1/I_k^2$. This can be done because I_k has much smaller variation than τ_k , therefore it will only produce a slightly different weighting. However, I_k cannot be neglected from the prescription σ_k/I_k , as that would result in different decay rates (I_k times larger than desired).

The problem with Eq. (3.12) 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. (3.10). This iteration should be done until the error e_τ does not decrease significantly.

Differentiating Eq. (3.12) with respect to c_1 and c_3 , and setting $\partial e_\tau / \partial c_1 = 0$ and $\partial e_\tau / \partial c_3 = 0$ gives:

$$\begin{aligned} c_3 &= \frac{\mathcal{M}(w_k) \mathcal{M}(w_k \sigma_k \vartheta_k^2) - \mathcal{M}(w_k \sigma_k) \mathcal{M}(w_k \vartheta_k^2)}{\mathcal{M}(w_k) \mathcal{M}(w_k \vartheta_k^4) - \mathcal{M}^2(w_k \vartheta_k^2)} \\ c_1 &= \frac{\mathcal{M}(w_k \sigma_k) - c_3 \mathcal{M}(w_k \vartheta_k^2)}{\mathcal{M}(w_k)} \\ \mathcal{M}(x_k) &= \sum_{k=1}^K x_k \end{aligned} \quad (3.13)$$

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.

Once the coefficients c_1 and c_3 are in hand, the filter parameters g and a_1 are computed by the inverse of Eqs. (3.6b) and (3.6c):

$$g = 1 - \frac{c_1}{f_0} \quad (3.14a)$$

$$a_1 = -1 - \frac{f_0 - \sqrt{8c_3 f_0 + f_0^2}}{4c_3} \quad (3.14b)$$

Figure 3.1 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 100th iteration (solid line). Consequently, there is no need for using the polynomial regression of Eq. (3.13) more than twice to achieve good results. This has been found to be valid for other tones as well.

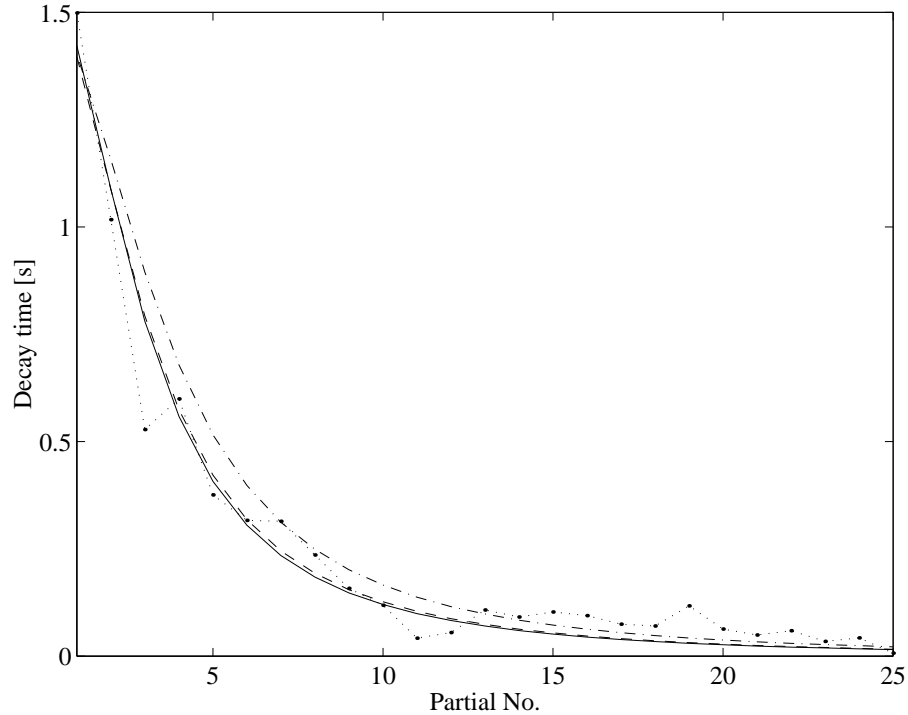


Figure 3.1: Prescribed (dotted line with black points), and approximated decay times for one (dash-dotted line), two (dashed line) and 100 (solid line) iterations by using the proposed one-pole filter design algorithm.

By looking at Fig. 3.1 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 relative error of decay times 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. It seems that the decay times of strings do not exactly follow the theoretical curve of the string with first- and third-order time derivatives. This is mainly because that simple model does not take into account the different sources of string losses [see Fletcher and Rossing 1998, pp. 53–56]. We have to note that this does not result in worse sound quality, mostly because the ear is relatively insensitive to the variation of decay times. The overall decay time of the tone can be varied from -25 to +40% without audible consequences [Järveläinen and Tolonen 2001]. Naturally, it is possible that smaller variations in the decay times of the individual partials are perceivable. Thus, further psychoacoustic research would be necessary to answer this question precisely.

On the whole, the novel algorithm gives good results for designing the one-pole loss filter. The string model with the one-pole loss filter is able to follow the general trend of the decay times, but it cannot capture the small variations coming from the effects of the string termination. 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: if one wishes to model the measured decay times more accurately, higher-order filters are needed.

3.2 High-order Loss Filters

The fact that the loss filter has only one pole is a hard constraint itself, since it restricts the sets of realizable decay times to a great extent. Moreover, 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 , e.g., by minimizing the mean squared magnitude error $\sum_{k=1}^K (|H_1(e^{j\vartheta_k})| - g_k)^2$. Nevertheless, problems arise because the decay times are a nonlinear function of the filter magnitude response. Therefore, as g_k approaches unity, the same amount of magnitude deviation $\hat{g}_k - g_k$ will correspond to a larger and larger difference in the decay time $\hat{\tau}_k - \tau_k$. Moreover, if the magnitude response of $H_1(z)$ exceeds unity at one of the partial frequencies, the digital waveguide becomes unstable, since the loop gain is larger than 1. Smith [1983] reviews many different filter designing techniques, but, most probably due to their complexity, they have not been widely used in practice.

In [Bank 2000b] a robust filter design method has been presented for high-order loss filters. The decay time error is computed from the expression of decay times (Eq. (2.61)) for harmonic strings ($f_k = kf_0$) by using the first-order Taylor series approximation for the logarithm function ($\ln x \approx x - 1$ for $x \approx 1$):

$$e_\tau = \sum_{k=1}^K (\hat{\tau}_k - \tau_k)^2 = \sum_{k=1}^K \left(\frac{1}{f_0 \ln \hat{g}_k} - \frac{1}{f_0 \ln g_k} \right)^2 \approx \frac{1}{f_0^2} \sum_{k=1}^K \left(\frac{1}{1 - \hat{g}_k} - \frac{1}{1 - g_k} \right)^2 \quad (3.15)$$

where g_k are the prescribed filter magnitudes at the partial frequencies ϑ_k and $\hat{g}_k = |H_1(e^{j\vartheta_k})|$ 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. Note that we again assumed $f_k = kf_0$.

A transformed filter $H_{tr}(z)$ is designed by a least squares filter design algorithm (e.g., `invfreqz` in MATLAB) by using a transformed specification $g_{k,tr}$. This minimizes the error

$$e_\tau = \frac{1}{f_0^2} \sum_{k=1}^K w_k \left(H_{tr}(e^{j\vartheta_k}) - g_{k,tr} \right)^2, \quad g_{k,tr} = \frac{1}{1 - g_k} \quad (3.16)$$

where ϑ_k refers to the frequency of partial k . The loss filter $H_1(z)$ can be computed from the transformed filter $H_{tr}(z)$ by the inverse transformation

$$H_1(z) = 1 - \frac{1}{H_{tr}(z)}. \quad (3.17)$$

The loss filter is either directly implemented as $H_1(z)$ or as a parallel structure having unity gain at one branch and $-1/H_{tr}(z)$ in the other. Implementing $H_1(z)$ directly leads to a simpler structure, but it might lead to higher computational complexity (e.g., when $H_{tr}(z)$ is an FIR filter). The physical interpretation and the stability analysis of the method is covered in [Bank 2000b] in detail.

A different approach was taken by Erkut [2001]. The idea is similar to the one-pole filter design of Sec. 3.1.2 but now a high-order polynomial is fit to the decay rates $\sigma_k = 1/\tau_k$, which contains terms of even order only. Then, instead of analytically computing the filter coefficients, a magnitude specification is calculated from the decay rate curve defined by the polynomial and this magnitude response is used as a specification for minimum-phase filter design. A difficulty of the approach is that the decay rates given by the higher-order polynomials might take unrealistic (e.g., negative) values. This is dealt with parameter shrinkage algorithms, increasing the complexity of filter design. Recently, two papers have been published that apply special filter structures (sparse filters) in [Lehtonen et al. 2005; Rauhala et al. 2005], which use the weighting-based method of Sec. 3.2.1 as a reference (which has been published in [Bank and Välimäki 2003]).

3.2.1 Weighting Function Based on the Taylor Series Approximation of Decay Times

A simple yet efficient filter design method is based on the first-order Taylor series approximation of the decay-time error. Here again the error with respect to the time constants is minimized. For example, in the mean squares sense the error is $e_\tau = \sum_{k=1}^K (\hat{\tau}_k - \tau_k)^2$ where τ_k are the prescribed and $\hat{\tau}_k$ are the approximated decay times. The decay times of the synthetic tone can be computed from the magnitude of the designed filter $\hat{g}_k = |H_1(e^{j\vartheta_k})|$ by Eq. (2.61). Writing Eq. (2.61) as a function gives

$$\hat{\tau}_k = \tau(\hat{g}_k) = -\frac{k}{f_k \ln \hat{g}_k}. \quad (3.18)$$

If the function $\tau(\hat{g}_k)$ is approximated by the first-order Taylor polynomial around the specification g_k , we obtain:

$$e_\tau = \sum_{k=1}^K (\tau(\hat{g}_k) - \tau(g_k))^2 \approx \sum_{k=1}^K (\tau'(g_k)(\hat{g}_k - g_k))^2 = \sum_{k=1}^K w_k (\hat{g}_k - g_k)^2, \quad (3.19)$$

which is a simple mean squares minimization with weights $w_k = (\tau'(g_k))^2$, where $\tau'(g_k)$ is the derivative of the function τ with respect to the specified magnitude g_k . Similar derivations can be performed for other error criteria (e.g., minimax). Note that now the weights depend on the magnitude specification and not on the frequencies, which is more common in digital filter design.

The first derivative of $\tau(g_k)$ is

$$\tau'(g_k) = \frac{k}{f_k g_k (\ln g_k)^2} = \frac{1}{I_k f_0 g_k (\ln g_k)^2}, \quad (3.20)$$

where $I_k = f_k/(k f_0)$ is the inharmonicity index. For $g_k = 1 - \epsilon$, $0 < \epsilon \ll 1$, which is generally the case, $\tau'(g_k)$ can be approximated by $\tau'(g_k) \approx 1/(I_k f_0 (g_k - 1)^2)$. This comes from the first-order Taylor series approximation of $\ln g_k$. Since $1/f_0$ does not depend on k , it can be omitted from the weighting function. Hence, the weighting function becomes

$$w_k = \frac{1}{I_k^2 g_k^2 (\ln g_k)^4} \approx \frac{1}{I_k^2 (g_k - 1)^4}. \quad (3.21)$$

A similar weighting function was obtained for the L^∞ norm minimization of the decay-time error in [Smith 1983, p. 182–183]. If the inharmonicity is moderate, which is the case for string instruments including the piano, I_k^2 may be neglected from the weighting function. This is not the case for simulating the sound of bars.

The approximation of Eq. (1) is accurate only for $\hat{g}_k \approx g_k$, which means that the magnitude of the designed filter is close to the specification. In many cases the measured decay times have a great variance, which cannot be followed by filters of reasonable order ($N < 20$). Therefore, it is worthwhile to smooth the decay time data τ_k , e.g., by convolving them with a short window function before computing the specification g_k . This way, the condition $\hat{g}_k \approx g_k$ can be assured.

The Phase

The magnitude specification g_k and the weights w_k can be directly used for linear-phase FIR filter design. However, by doing so, half of the degrees of freedom are wasted for demanding the impulse response to be symmetric. In practice, it is not necessary to have an exactly linear-phase loss filter, since a nonlinear phase response corresponds to a slightly inharmonic tone, which does not corrupt the sound quality.

Designing minimum-phase filters is a pleasant choice, since then the phase specification can be easily computed from the logarithm of the magnitude specification by Hilbert transform [Oppenheim and Schaffer 1975]. Note that the Hilbert transform needs magnitude data for the entire digital frequency band and on a linear frequency scale. The missing data points in the high frequency region are calculated by designing a one-pole filter for the specification, e.g., by the method proposed in Sec. 3.1.2. Then, the magnitude response of the one-pole filter is used as a specification for the high frequencies. This is reasonable since the loss filter behavior in the high frequency region has no significant influence on the resulted tone and such a simple specification is easily fulfilled by the filter design. At the original data points of the highest specified frequencies a crossfade is applied to avoid discontinuities.

Design Example

Examples are presented for IIR filter design. For the examples, the weighted least squares method implemented in MATLAB's `invfreqz` function is used [Mathworks 1996]. The decay time data are smoothed by convolving them with a triangular window $[0.25, 0.5, 0.25]$. Here, the last five data points of the measured specification are linearly mixed to the magnitude response of the designed one-pole filter. The magnitude response on a dense linear grid is calculated by using third-order polynomial interpolation, which was found to be accurate enough. The phase response on this dense grid is computed by the Hilbert transform and then resampled at the frequencies of the original specification points.

The decay time data used for this example was calculated from an F_2^\sharp piano tone ($f_0 = 92.2$ Hz), near-field recording. The decay rate of the partials up to 6.57 kHz were measured, which yielded data for 64 partials. The sampling frequency is $f_s = 22.05$ kHz. The smoothed decay times are displayed with points in Fig. 3.2 (c). The filter magnitude specification calculated from the smoothed decay times has been plotted with points in Fig. 3.2 (a) and (b). IIR filters of order 2, 8, and 16 were designed. The magnitude responses are depicted in Fig. 3.2 (a), and Fig. 3.2 (b) shows the same curves magnified for the most relevant frequency and magnitude region. Fig. 3.2 (c) shows the corresponding decay times. Figure 3.2 reveals that the magnitude error is smaller where the specification g_k is closer to unity, which is necessary for the equal accuracy in decay times. Similar results have been obtained with several cases of piano and guitar data. The magnitude response of the designed filters never exceeded unity, that is, the digital waveguide loop remained always stable.

3.3 Conclusion

In this chapter different loss filter design techniques have been presented, all of them based on minimizing the decay-time error. In Sec. 3.1.2 a simple technique have been described for designing one-pole loss filters, which applies the analytical expression of the decay times of such a filter [published in Bank 2000b]. Section 3.2.1 has proposed a simple and robust technique for high-order loss filter design, based on the first-order Taylor series of the decay-time error and suggested the use of smoothing the target response and designing a minimum-phase loss filter [published in Bank and Välimäki 2003]. The choice between the one-pole or the high-order loss filters mainly depends on whether the synthetic tone should highly resemble to a specific recorded one or it should only reproduce the characteristic sound of a given instrument. As an example, piano sounds synthesized by the one-pole loss filter still sound piano like. Actually, they do not sound worse than the recorded ones with respect to tone decay, they just sound different. This difference can be minimized by applying high-order loss filters. We have to note that the computational complexity of the loss filter is negligible compared to other parts of the model (e.g., the body, or the dispersion filter in the case of inharmonic strings), so there is no real drawback of applying a high-order loss filter with respect to computational complexity. However, the parameters of the high-order loss filter cannot be tuned intuitively by the user, unlike for the one-pole loss-filter, where τ_0 and $f_{\tau/2}$ in Eq. (3.8) have a physical and perceptual meaning.

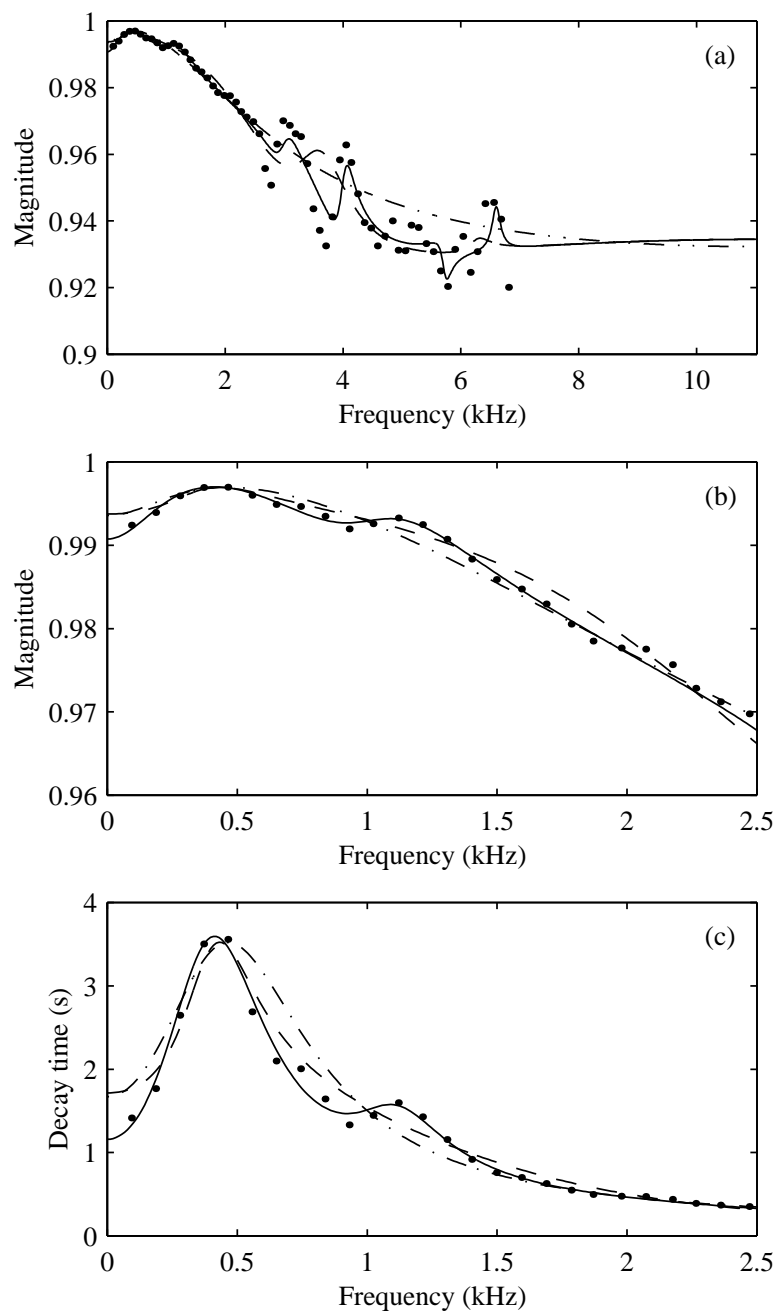


Figure 3.2: IIR loss filter magnitude responses (a) for the full frequency band and (b) at low frequencies up to 2.5 kHz, and the corresponding decay times (c) for filter orders of 2 (dash-dotted line), 8 (dashed line), and 16 (solid line). The specification is depicted with dots in each case.

Chapter 4

Multi-rate Techniques for Efficient String Instrument Modeling

In this chapter multi-rate techniques are applied for increasing the computational efficiency of string instrument modeling. First, a multi-rate excitation model is presented in Sec. 4.1 which solves the problem of noncomputable loops in a simple and efficient way. This is followed by a multi-rate resonator bank in Sec. 4.2, which is able to model the beating and two-stage decay of strings at a fraction of the computational cost needed by earlier methods. In Sec. 4.3 a multi-rate filtering technique is proposed which decreases the computational cost of instrument body modeling by an order of magnitude compared to previous filtering methods. These techniques have been developed for piano modeling, but they can also be used for other string instruments. The excitation model and the beating and two-stage decay model can be used for impulsively excited (plucked or struck) string models. The body model is the most general in this sense, as it can be used for bowed strings, too.

Note that whenever the methods are compared to each other with respect to computational cost, it is in terms of the numbers of multiplications and additions, and we do not deal with the necessary overhead coming from loops, conditional jumps, etc., as they depend on the implementation.

4.1 Excitation Modeling

Let us take the example of a hammer model implementing Eq. (2.78). Here we repeat the equation

$$F_h(t) = F(\Delta y) = \begin{cases} K_h(\Delta y)^{P_h} & \text{if } \Delta y > 0 \\ 0 & \text{if } \Delta y \leq 0 \end{cases} \quad (4.1a)$$

$$F_h(t) = -m_h \frac{d^2 y_h(t)}{dt^2}. \quad (4.1b)$$

As already noted in Sec. 2.4.1, the problem of discretizing Eq. (4.1) lies in the fact that there is a mutual dependence between the excitation force F_h and the hammer position y_h . Thus, for computing F_h by Eq. (4.1a), y_h should be known, which is, in turn, can be

computed only by Eq. (4.1b) if we know F_h . When the block diagram of this system is drawn, this fact leads to a noncomputable delay-free loop. Note that the problem is the same for other types of nonlinear excitation, as they all have a similar form: one equation describes a nonlinear dependence of force on the position of the exciter (which is not always memoryless as in Eq. (4.1a)) and the other characterizes the motion of the exciter as a function of force, which is usually a linear dynamic system.

The straightforward solution to the problem that we compute y_h and F_h in an interleaved fashion, meaning that for computing $y_h(t_n)$, we use the force computed in the previous time instant, $F_h(t_{n-1})$. In theory, this means inserting a delay element in the delay-free loop. This element is called “fictitious” in Borin and De Poli [1996], but actually it is a real one, which is implemented unintentionally. If the sampling rate is sufficiently high in comparison with the variation of hammer force (i.e., $F_h(t_n) \approx F_h(t_{n-1})$), this does not present a practical problem. However, if this is not true, numerical instabilities may arise.

The theory of wave digital filters addresses the problem of noncomputable loops in terms of wave variables. Every component of a circuit is described as a scattering element with a reference impedance, and delay-free loops between components are treated by “adapting” reference impedances. Van Duyne et al. [1994] presented a “wave digital hammer” model, where wave variables are used. The 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. In this model the “fictitious” delay element appears when the stiffness coefficient is read from the lookup-table.

Borin and De Poli [1996] have proposed a general strategy named “K method” for solving noncomputable loops in a wide class of nonlinear systems. The method is fully described in [Borin et al. 2000] along with some application examples. Here only the basic principles are outlined. Whichever the discretization method, the hammer compression $\Delta y(t_n)$ can be written as:

$$\Delta y(t_n) = p(t_n) + KF_h(t_n), \quad (4.2)$$

where $p(t_n)$ is the linear combination of past values of the variables, namely, F_h , y_h , and y_s , where y_s refers to the string displacement at the position of the excitation. The value of coefficient K depends on the numerical method in use. The interaction force F_h at time instant t_n , computed by Eq. (4.1a), is therefore described by the implicit relation $F_h(t_n) = F(p(t_n) + KF_h(t_n))$. The K method uses the implicit function theorem to solve this implicit relation:

$$F_h = F(p(t_n) + KF_h(t_n)) \xrightarrow{K_{meth.}} F_h = h(p(t_n)). \quad (4.3)$$

The new nonlinear map h defines $F_h(t_n)$ as a function of $p(t_n)$, hence instantaneous dependencies across the nonlinearity are dropped. The function h can be precomputed and stored in a look-up table for efficient implementation. However, for different hammer parameters different look-up tables have to be stored.

4.1.1 The Multi-rate Excitation Model

A simpler approach for avoiding the numerical instability is the multi-rate excitation model. The idea is that stability of the discretized excitation model with a “fictitious” delay can always be maintained by choosing a sufficiently large sampling rate f_s if the corresponding continuous-time system is stable. As $f_s \rightarrow \infty$, the discrete-time system will behave as the original differential equation. In practice, doubling the sampling rate of the whole string–excitation system would double the computation cost as well. However, if only the excitation model operates at double rate, the computational complexity is raised only by a negligible amount. This is because the excitation models usually require much less computation compared to the string model. Thus, running the excitation model twice does not increase the computational load of the whole instrument model significantly.

We stay by the example of hammer modeling. As the hammer model runs at double sampling rate, the string displacement at the hammer position y_s should be computed at double sampling rate, too, as the felt compression Δy is computed as $y_h - y_s$. Therefore, the string displacement y_s is kept track within the hammer model. The core of the hammer model is displayed in Fig. 4.1. The hammer model of Fig. 4.1 first computes the velocity

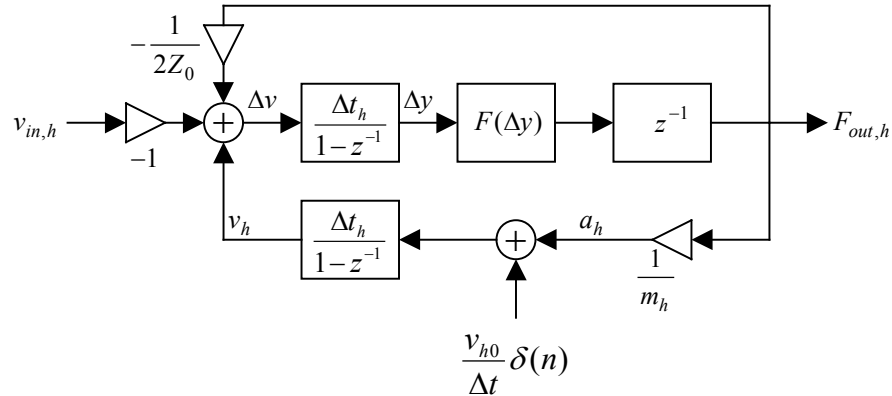


Figure 4.1: The core of the proposed hammer model.

difference of the string and the hammer $\Delta v = v_h - v_s$, where v_h is the hammer velocity and v_s is the string velocity. The string velocity is computed as $v_s = v_{in,h} - F_{out,h}/(2Z_0)$, where $v_{in,h}$ is the incoming string velocity (the velocity of the string without excitation), Z_0 is the string impedance, and $F_{out,h}$ is the force signal computed by the power law in the previous time instant (z^{-1} refers to the “fictitious” delay element). 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 Schaffer 1975], and the sampling time of the hammer is referred as Δt_h . The interaction force is computed by the law of Eq. (4.1a). 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} .

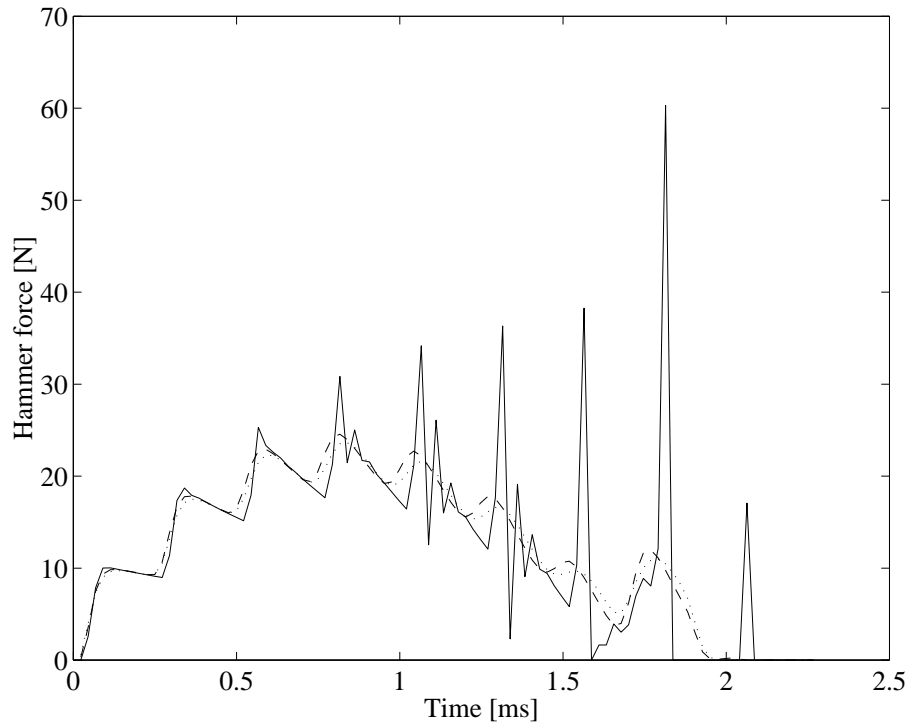


Figure 4.3: The interaction force of note C_5 (522 Hz) with $f_s = 44.1$ kHz, and hammer velocity $v = 5$ m/s, computed by inserting a fictitious delay element (solid line), with the K method (dotted line), and with the multi-rate hammer (dashed line).

at string terminations. The K method [Borin et al. 2000] and the multi-rate hammer proposed here produce very similar force signals. On the other hand, inserting a fictitious delay element drives the system towards instability (the spikes are progressively amplified). In general, the multi-rate method provides comparable output to the K method for hammer parameters realistic for pianos, while it does not require the use of precomputed look-up tables and leads to a simpler implementation. However, when low sampling rates (e.g., $f_s = 11.025$ kHz) or extreme hammer parameters are used (i.e., the stiffness of the hammer is increased ten times), its stability cannot be maintained by upsampling by a factor of 2. In such cases, either the upsampling factor has to be increased, or the K method should be used.

Note that the multi-rate method can be used for increasing the numerical stability of other kind of excitation models, as only the excitation model of Fig. 4.1 has to be replaced by the model of the new excitation. The only important thing is that the string velocity v_s or displacement y_s at the excitation position $x_{exc} = L(M_{in}/M)$ has to be computed within the excitation model. If this has been already included in the development of the straightforward, single-rate model, then it can be a direct replacement of Fig. 4.1.

The method has been presented in connection with a digital waveguide string model. However, it can be used with finite-difference string models, too. Here the input of the

excitation model $v_{\text{in,h}}(n\Delta t)$ is the velocity of the string at the position of the hammer $v_{\text{out}}(n\Delta t)$, that is already known. The next value of the string velocity $v_{\text{out}}(n\Delta t + \Delta t)$ required for computing $v_{\text{in,h}}(n\Delta t + \Delta t/2)$ is calculated by running the string (actually only the corresponding string element) for the next time-step, which gives the next string velocity value without the excitation force.

4.2 Modeling Beating and Two-stage Decay

In string instruments, the coupling of the two transverse polarizations (y and z) through the bridge leads to beating and two-stage decay, the first referring to an amplitude modulation overlaid on the exponential decay, and the second meaning that the tone decays faster in the early part than in the latter. These phenomena were studied by Weinreich [1977]. This is even more complicated in the case of the piano, where two or three slightly mistuned strings of the same note are sounded together when a single piano key is pressed (except for the lowest octave), leading to a four or six vibrating modes for one partial.

In the digital waveguide string modeling paradigm, the simplest way for modeling beating and two-stage decay is to use two digital waveguides in parallel for a single note. Varying by the type of coupling used, many different solutions have been presented in the literature. In [Karjalainen et al. 1998] the two digital waveguides were coupled by constant coefficients, while in [Smith 1993] a frequency-dependent termination impedance was applied, and the loss filters of the strings were omitted. A more refined method was presented in [Aramaki et al. 2002; Bensa 2003] that connects the two or three coupled waveguides by frequency dependent filters. This latter method provides accurate resynthesis of the coupling phenomena, but requires significantly larger computational cost compared to a single string model. Here a different approach is presented that combines the advantages of digital waveguides and modal modeling.

4.2.1 The Parallel Resonator Bank

In the resonator bank approach, presented first in [Bank 2000b; Bank et al. 2000], second-order resonators $R_1(z) \dots R_K(z)$ are connected to the basic string model $S_v(z)$ (e.g., digital waveguide) in parallel. This is displayed in Fig. 4.4. The idea comes from the observation that the behavior of the coupled y and z polarizations can be described by a pair of exponentially damped sinusoids [Weinreich 1977]. Although more modes would be required for the perfect reconstruction of piano partial envelopes (due to the unison triplets), it was found that a two mode model can reproduce the main features of the phenomenon (see Fig. 4.5).

In this model, one sinusoid of the mode-pair is simulated by one partial of the digital waveguide and the other one by one of the resonators $R_k(z)$. These resonators can be realized by using Eq. (2.72). The advantage of the structure is that the resonators $R_k(z)$ are implemented only for those partials whose beating and two-stage decay are significant. The others will have simple exponential decay, determined by the digital waveguide model $S_v(z)$. Five to ten resonators have been found to be enough for high quality sound synthesis

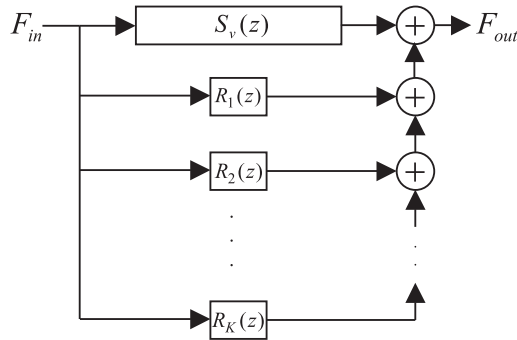


Figure 4.4: Modeling beating and two-stage decay by a digital waveguide $S_v(z)$ and a parallel resonator bank $R_1(z) \dots R_K(z)$.

for the piano. As a radical example, only one resonator has been implemented for the harpsichord model of Välimäki et al. [2004].

The choice of those partials where the phenomenon is most prominent can be automated by computing the energy (i.e., mean square value) of the beating modes and selecting the ones with the largest energy value. This results in a synthesized tone that is most similar to the original one in a mean squares sense. Naturally, a psychoacoustic selection of the dominant beating modes that incorporates the masking effect similarly to perceptual coders could result in a better sound quality, but would also complicate the parameter estimation significantly. After the beating partials are selected, the parameters of the resonator bank are determined by first fitting an exponential decay on the amplitude envelopes of the partials. This is done by linear regression in the logarithmic amplitude scale, as proposed in [Välimäki et al. 1996]. Then, the relative deviation of the real amplitude envelope from this simple exponential decay is computed as a ratio of the two signals. Then, an exponentially increasing or decaying sinusoid is fitted to this deviation signal. The parameters of this “beating sinusoid” determine the parameters of the resonators. The procedure is outlined in [Bank 2000b, Sec. 6.4] in detail. Note that methods based on ARMA modeling [Karjalainen et al. 2002; Bensa 2003] could also be used for the parameter estimation of mode pairs.

In Fig. 4.5 (a) the first 8 partial envelopes of a recorded A_4^\sharp note (466Hz) are displayed. Figure 4.5 (b) shows the output of the synthesis model with one digital waveguide and five resonators. It can be seen in Fig. 4.5 (b) that the characteristics of beating and two-stage decay are well preserved for the first five partials (where the resonators are implemented), and the other partials have simple exponential decay determined by the digital waveguide. Note that the initial amplitudes of the partials for the original and synthesized tones are different. This is because the initial amplitudes are determined by a physics-based string and excitation model, where the goal is to simulate a piano-like behavior, and not the reproduction of a given tone. On the other hand, the beating and two-stage decay are modeled by a signal model, whose parameters are determined by the analysis of a specific tone, therefore, they can be more similar to the original.

The general shapes of the partial envelopes are well reproduced, while some details are

missing in Fig. 4.5. This is because now we are fitting a lower order model to a higher order system, since the A_4^\sharp note has three strings both with two transverse polarizations (i.e., a total number of six modes). When more precise results are desired, more resonators should be used for one partial. Note that for all the other stringed instruments (guitar, lute, harp, harpsichord, etc.), where one string belongs to one note, the two-mode model can perfectly resynthesize the string decay for the two transverse polarizations (here we neglect the effect of longitudinal vibration).

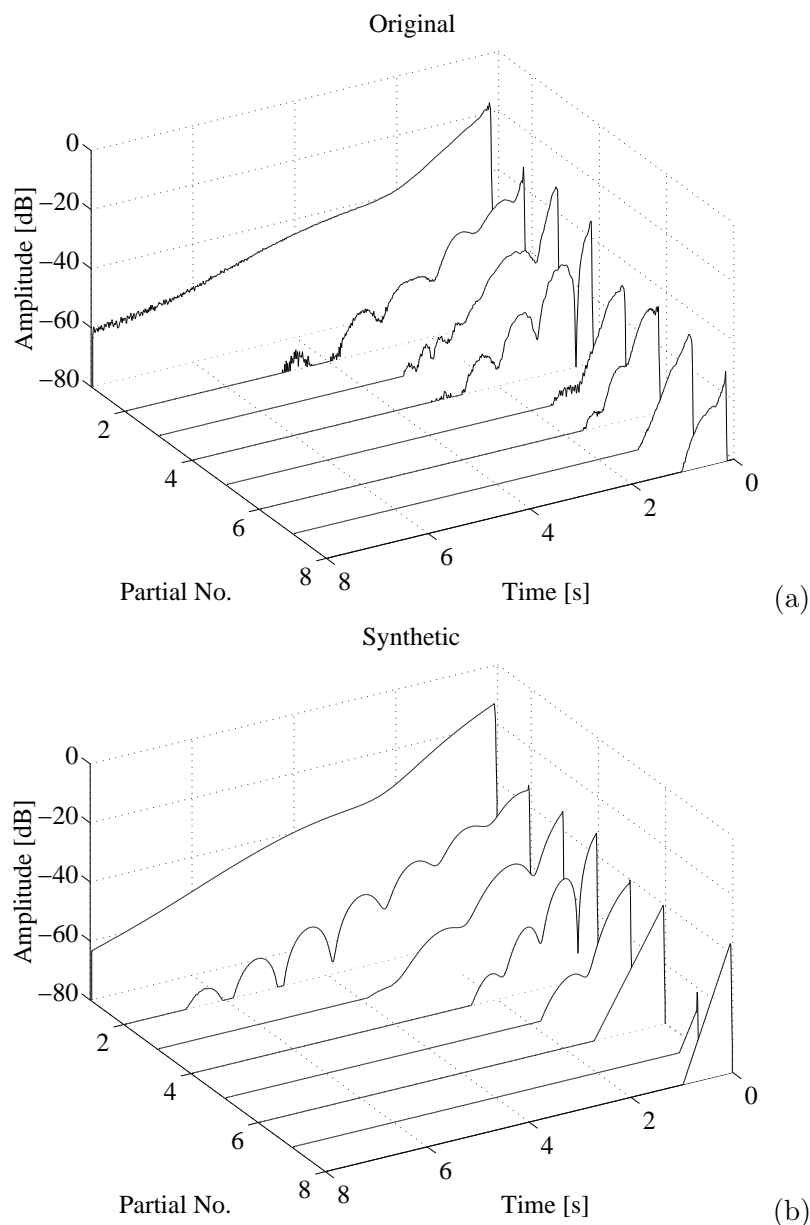


Figure 4.5: Partial envelopes of an A_4^\sharp piano tone: original (a) and synthesized by a digital waveguide and five resonators in parallel (b).

Multi-rate Modeling

The computational complexity can be even more decreased if the resonator bank is implemented by the multi-rate approach, running the resonators at a much lower sampling rate, e.g., the 1/8 or 1/16 part of the original sampling frequency. The structure is depicted in Fig. 4.6, where first the force signal coming from the hammer model is downsampled, filtered by the second-order resonators, and then upsampled to the original sampling rate. Since, for one note, resonators with different sampling rates are used, it is beneficial to implement the multi-rate system by cascading half-band downsampling and upsampling filters. This also simplifies the filter design. The sign $\downarrow 2$ in Fig. 4.6 refers to the downsampling operation with prior antialiasing filtering, and the sign $\uparrow 2$ stands for upsampling operation with an interpolation filter. For simplicity, Fig. 4.6 shows only one resonator at every downsampled sampling rate, but in practice, many resonators are connected in parallel within the same branch.

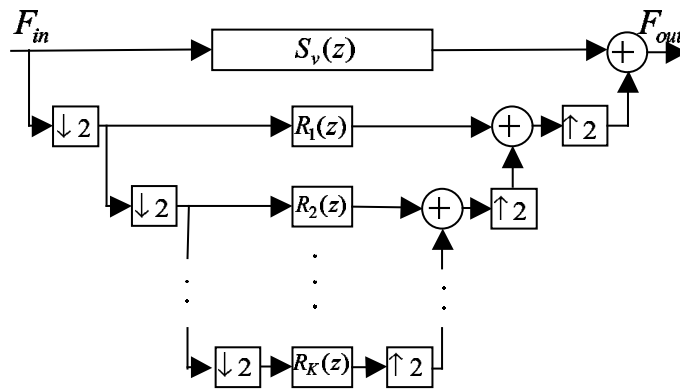


Figure 4.6: The multi-rate resonator bank.

In the filter design we can take the advantage that the downsampled signal is imposed to filtering by a second order resonator, which has a very narrow amplitude response. The input signal of the resonator (the excitation force) is a short pulse, and its only role is to set the initial amplitude and phase of the resonator. This means that a small aliasing after downsampling is acceptable. Having 20 dB stopband attenuation leads to an amplitude change of 0.8 dB, which has been found to be inaudible in practice. The upsampling filters cannot be simplified this way, there 60 dB stopband attenuation is needed to avoid audible aliasing. On the other hand, all the output signals having the same sampling rate can be summed before upsampling, therefore the same interpolation filters can be used for all the notes (this is not shown in Fig. 4.6). The computational complexity can be reduced further by using filters having less tight specification in the passband, leading to lower filter orders. This can be done because the amplitude and phase errors of the downsampling and upsampling filters can be corrected by changing the amplitudes and phases of the resonators. The total transfer function error E_k coming from

the downsampling and upsampling filters for the k th resonator is computed as follows:

$$E_k = \prod_{n=1}^N H_{\text{dn}} \left(\frac{2\pi f_k 2^{n-1}}{f_s} \right) H_{\text{up}} \left(\frac{2\pi f_k 2^{n-1}}{f_s} \right) \quad (4.6)$$

where $H_{\text{dn}}(\vartheta)$ and $H_{\text{up}}(\vartheta)$ are the transfer functions of the downsampling and upsampling filters, and they are computed by substituting $z = e^{j\vartheta}$. The total downsampling factor is 2^N , i.e., the resonator runs at $f_s/2^N$. The estimated amplitude A_k and phase φ_k parameters of the resonators have to be modified by the magnitude and phase of $1/E_k$.

In practice, for a passband of $0 < \vartheta < 0.4\pi$ and stopband of $0.6\pi < \vartheta < \pi$, the `remez` algorithm in MATLAB [Mathworks 1996] gives a third-order FIR filter with 5 dB passband ripple and 20 dB stopband attenuation. For the interpolation filters 60 dB stopband attenuation is required, leading to 13th-order filters. Because of the passband goes only up to 0.4π , in the region $0.4\pi < \vartheta < 0.5\pi$ no resonators should be implemented. In this way, we loose 20% of the downsampled frequency range but this leads to low-order downsampling and upsampling filters. Note that linear phase FIR filters are only used for simplicity of their design, as the phase errors of the downsampling and upsampling filters are easily corrected by changing the phases of the resonators.

The computational complexity is reduced compared to having a parallel waveguide even if the resonators are running at the sample rate of the digital waveguide string model, as only a few of them (5–10) have to be implemented. Realizing the resonators in a multi-rate fashion decreases the computational load significantly. The average computational cost of the method for one note is around ten multiplications per sample. This is because the highest computational load is the first downsampling filter (four tap FIR filter), then the second one runs at every second time instant (two operations per sample), the third at every fourth (one operation per sample), etc. A second-order resonator $R_k(z)$ requires two multiplications, but as it runs at $f_s/8$ or $f_s/16$, its computational load is negligible. So is the load of the 14 tap upsampling filters, as they are shared by all the notes played. The complexity is even lower (around five multiplications per sample) if the first or first few downsampling filters are omitted. This can be done because the input signal (excitation force) is of a lowpass character. In this case, the method requires an order of magnitude smaller computational resources compared to implementing a second digital waveguide. Moreover, the parameter estimation becomes simpler, since only the parameters of the mode-pairs have to be found by one of the methods presented in [Bank 2000b; Karjalainen et al. 2002; Bensa 2003], and there is no need for coupling filter design. Although the method have been presented in connection with a digital waveguide string model, it can be directly used to augment finite-difference or modal-based string models.

4.3 Instrument Body Modeling

The different techniques for modeling the instrument body have been already outlined in Sec. 2.5. The most realistic sound can be achieved by filtering-type body models, where the body filter is designed with the help of recorded impulse responses. This approach is

generally applied as a post-processing technique, modeling the radiation properties of the body only, while the impedance properties are treated within the string model. Here we will review the possible modeling methods.

4.3.1 FIR Filters

The most straightforward approach for implementing the body response is to use the windowed and truncated version of the measured impulse response as an FIR filter. This technique is simple and capable to provide the best sound quality from a given measurement. On the other hand, high filter orders are required to reach high quality sound.

In the case of the acoustic guitar, it was found that filter orders lower than 1000 do not produce satisfactory sound [Karjalainen et al. 1999]. For modeling the piano soundboard at a sampling rate of $f_s = 44.1$ kHz, the present author have found that the sound changes only slightly when the filter length is raised over 2000 tap. Under filter order 1000, the sound starts to lose its character. Consequently, having a filter order between 1000 and 2000 seems to be a reasonable choice. These results coincide with the ones presented in [Karjalainen et al. 1999] for the guitar.

The transfer function of a 2000 tap filter is displayed in Fig. 4.8 (b), implementing the transfer function of Fig. 4.8 (a) (which is the same as presented in Fig. 2.5, p. 35). It is easy to notice in the low frequencies that the resonances have been smeared. This is because the long decaying low modes have been truncated. In practice, this does not alter the sound significantly. However, for precise synthesis even larger impulse responses are required, as even the 2000 tap filter cannot reproduce the knocking sound of high piano tones, because the “knock” is much longer (0.2–0.5 sec, requiring ca. 10000–20000 taps).

The computational requirements can be somewhat reduced if the lowest resonances of the instrument body are factored out from the FIR filter and implemented as second-order resonators [Karjalainen et al. 1999]. In the case of the acoustic guitar, this resulted in 500 tap FIR filters.

4.3.2 IIR Filters

For modeling the violin body, many different filter design techniques were compared by Smith [1983]. The final choice was an eighth-order IIR filter designed by minimizing the Hankel norm on Bark scale. Nowadays the implementation of higher-order filters have become possible, but the quality requirements have also been increased.

In [Karjalainen et al. 1999], two IIR filter design methods were compared for modeling the guitar body. It has turned out that IIR filters perform almost the same as FIR filters with the same computational cost. Similar results have been found for the piano, i.e., filter orders less than 500 do not produce appropriate sound. As also noted by Karjalainen et al. [1999], minimum-phase equalization is not a good option, since it destroys the reverberant character of the response. Moreover, in the case of the piano, when experiments were made with minimum-phase filters, it has turned out that they ruin the characteristic attack of the piano and result in an unnatural sound.

Frequency warping can also be used in filter design to give more emphasis to the psychoacoustically more important low frequencies. For the violin model of Smith [1983], this approach was used. In [Karjalainen and Smith 1996], warped filter design was proposed for modeling the body response of the acoustic guitar. By this technique, the required filter orders can be reduced significantly. However, either these warped filters need special structures for implementation, or they have to be converted to conventional structures. During conversion, numerical instabilities may arise, especially when high filter orders are used.

4.3.3 Multi-rate Body Modeling

Among the filter design approaches, the FIR filter is capable to produce the best sound quality for a given transfer-function measurement. This is because that it preserves not only the overall magnitude response of the instrument body, but also the phase information. Having accurate time-domain response seems to be crucial for the realistic attack of synthesized sounds.

Here, a multi-rate approach is presented to avoid the high computational cost of the FIR filter, while still maintaining its benefit of preserving the sound characteristics. As shown in Fig. 4.7, the string signal F_{string} is split into two frequency bands. The lower is filtered by a long FIR filter $H_{\text{low}}(z)$ running at a considerably lower sampling rate ($f'_s = f_s/8$), precisely synthesizing the body impulse response up to 2 kHz. This means that the same impulse response length (in ms) consumes only the 1/64 part of computation compared to a single-rate filter. This is because the filter length is reduced by a factor of 8 compared to a single-rate FIR filter with the same length in ms, and this shorter filter is run at every eighth time instant.

In the high frequency band only the overall magnitude response of the body is modeled, using a low-order filter $H_{\text{high}}(z)$ running at the sampling rate of the system ($f_s = 44.1$ kHz). This part of the signal flow is delayed by N samples to compensate for the delay of the downsampling and upsampling operations. The simplification in the high frequency region is motivated by the fact that here the human ear has been found to be less sensitive to the position of the modes.

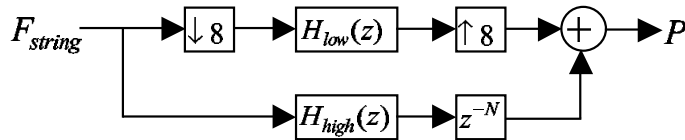


Figure 4.7: The multi-rate body model.

For the decimation and interpolation filters, a polyphase FIR filter $H_{\text{di}}(z)$ has been used. Note that the interpolation and decimation filters could be different in principle, but here the same filter $H_{\text{di}}(z)$ is used for both operation. In a general multi-rate system, large stopband attenuation and small passband ripple would be required, but this would result in long interpolation and decimation filters. However, here having ca. 5 dB passband

ripple have been found to be an appropriate solution, as it can be corrected by changing the magnitude response of the low frequency body filter $H_{\text{low}}(z)$. For stopband attenuation, 80 dB is sufficient in practice. The filter $H_{\text{di}}(z)$ is designed by the `remez` algorithm implemented in MATLAB [Mathworks 1996] giving a filter order of 192, which results in 24 operations per cycle in polyphase implementation.

The body filters $H_{\text{low}}(z)$ and $H_{\text{high}}(z)$ can be designed from the measurement of real instruments. The target impulse response $H_t(z)$ is a 2000 tap FIR filter obtained by truncating the measured impulse response. This is lowpass-filtered by a linear-phase FIR filter and then decimated by a factor of 8. This results in a 250 tap FIR filter $\tilde{H}_{\text{low}}(z)$. Now the passband errors of the decimation and interpolation filters have to be corrected. This is computed as follows:

$$H_{\text{low}}(e^{j\vartheta}) = \tilde{H}_{\text{low}}(e^{j\vartheta}) \frac{1}{H_{\text{di}}^2(e^{j\frac{\vartheta}{8}})} e^{-j\vartheta N}, \quad (4.7)$$

where we have used the substitution $z = e^{j\vartheta}$. The multiplication by $e^{-j\vartheta N}$ comes from the fact that we have already compensated the delay of the interpolation and decimation filters by delaying the other signal flow in Fig. 4.7. Neglecting this term from Eq. (4.7) would lead to a noncausal filter. The filter $H_{\text{low}}(z)$ corresponds to the body filter to be implemented in Fig. 4.7.

Now the impulse response of the low frequency chain of Fig. 4.7 is known. The remaining high frequency part can be easily calculated by subtracting this low frequency response from the target impulse response $H_t(z)$. This way, a 2000 tap FIR filter arise containing energy mainly at frequencies above 2 kHz. This response is then made minimum-phase, e.g., by the `rceps` function in MATLAB [Mathworks 1996]. This concentrates the energy to the beginning of the impulse response. Then this minimum-phase response is truncated to a length of 50 tap to form the high frequency body filter $H_{\text{high}}(z)$.

As an example, the magnitude response of a multi-rate piano soundboard model is depicted in Fig. 4.8 (c). The magnitude response of the target FIR filter $H_t(z)$ is depicted in Fig. 4.8 (b) for comparison. It can be seen from the figures that the magnitude response is accurately preserved up to 2 kHz. Although not shown, but so is the phase response. Above 2 kHz, only the overall magnitude response is retained.

The model consumes around 130 operations per cycle. This is because $H_{\text{low}}(z)$ requires 250 operations in every eighth cycle, while the downsampling and upsampling filters require 2×192 . The high-frequency filter $H_{\text{high}}(z)$ is an 50 tap filter running in each cycle. This gives the average load $(250 + 2 \times 192)/8 + 50 = 129.25$. Despite its significantly lower computational cost, the model has a very similar spectral character to the 2000 tap target filter $H_t(z)$. This is because the main features of the instrument body are reproduced. One such feature is the proper overall magnitude response, which seems to be accomplished. Another important feature of the instrument body is providing the proper attack of the sound. This is partly fulfilled, since up to 2 kHz the time-domain response of the model equals that of the target response. The attack of high notes sounds sharper compared to the 2000 tap target filter. This is because the energy of the soundboard response is

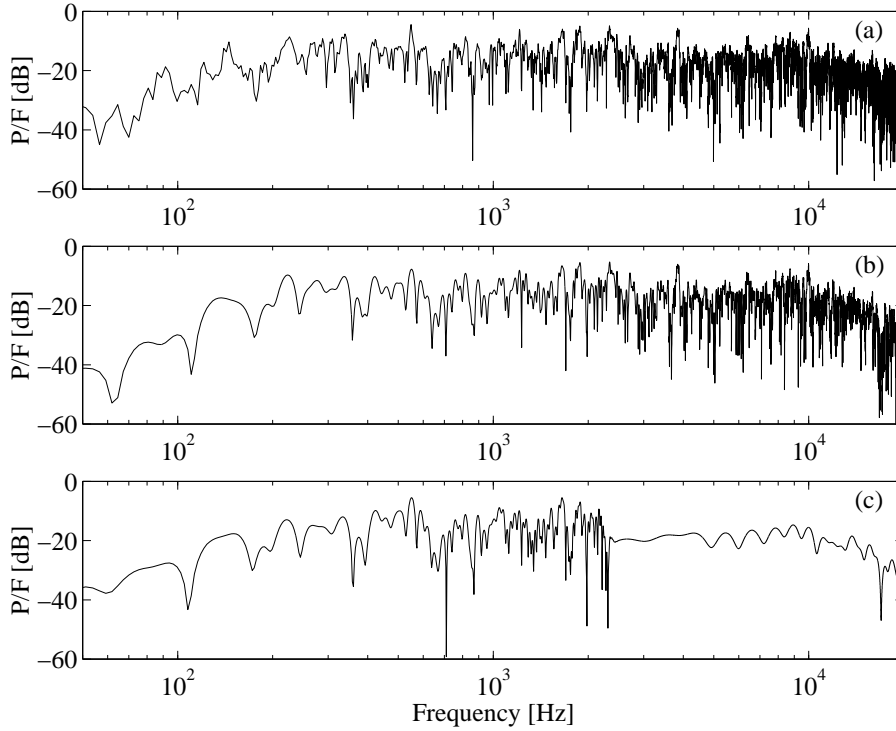


Figure 4.8: The magnitude transfer function of the piano soundboard (a) implemented by a 2000 tap FIR filter (b) and by the multi-rate body model (c).

concentrated to the first 1.1 ms above 2 kHz. Therefore, the attacks of high frequency partials are not smoothed enough by the body filter.

A different way of utilizing the idea of multi-rate filtering is increasing the quality, rather than decreasing the computational complexity. When 2000 operations per sample is acceptable, the strategy is as follows: the signal below 4.4 kHz is downsampled by a factor of 4 and filtered by a 4000 tap (meaning 360 ms length at $f_s/4$) FIR filter $H_{\text{low}}(z)$. The signal above 2.2 kHz is filtered by a 1000 tap (ca. 20 ms at $f_s = 44.1$ kHz) FIR filter, $H_{\text{high}}(z)$. The filter $H_{\text{high}}(z)$ is computed by subtracting the impulse response of the low frequency chain from the target response $H_t(z)$ (which now has the length of 16000 taps) providing a residual response containing energy above 4.6 kHz. This residual response windowed to a shorter length (1000 tap). Note that the residual response does not have to be made minimum-phase prior to windowing, as in this case the largest part of the energy in the impulse response is contained within the first 1000 tap. Now the model reproduces both the overall magnitude and the transient behavior of the tone even for the higher partials, as the attack time is 20 ms instead of the 1.1 ms of the low complexity solution. The characteristic “knock” of the high notes are also reconstructed by the long (0.36 s) low frequency response. Indeed, the sound produced by this model is indistinguishable from that calculated by a 16000 tap FIR filter directly implementing the soundboard impulse response, while it requires “only” about 2000 operations per sample.

This computational load is now acceptable for string modeling, as it consumes about 10 % of an average personal computer when the code is written in C++. However, for multimedia applications where the sound has a secondary role, the low complexity solution is preferred.

4.4 Conclusion

In this chapter the multi-rate approach has been utilized for string instrument modeling. The multi-rate excitation model of Sec. 4.1 [published in Bank 2000a; Bank et al. 2003] provides a comparable solution to the K method for avoiding numerical instabilities in excitation models, but it results in a simpler structure. This is because it only requires that the existing excitation model runs at a higher sampling rate, and there is no need for the rearrangement of the model to known and unknown terms. Moreover, the need for a lookup-table is also avoided, which is particularly advantageous when the excitation parameters are varied in real-time.

The multi-rate resonator bank of Sec. 4.2 [published in Bank 2001; Bank et al. 2003] models beating and two-stage decay by augmenting the basic string model with second-order resonators. An advantage of the technique is that the second-order resonators are required only for those partials that are dominated by the effect (typically 5–10). The computational complexity is further reduced by running the resonators at a lower sampling rate. The decimation and interpolation filters can be of low order, as their passband errors are corrected by changing the amplitudes and phases of the resonators.

Section 4.3 presented a multi-rate filtering technique for instrument body modeling [published in Bank et al. 2002, 2003]. The low frequency part of the body response is modeled by a long FIR filter running at a lower sampling rate, while the high frequency part is modeled by a shorter filter running at normal sampling rate. Similarly to the multi-rate resonator bank, the decimation and interpolation filters can be of low order as their passband ripple is corrected by the filter implementing the low frequency part of the body response. The multi-rate technique decreases the computational requirement by an order of magnitude compared to previous filtering methods for the same impulse response length. This can be utilized either for computational savings or for increasing the response length leading to higher sound quality.

Chapter 5

Modeling of Geometric Nonlinearities

This chapter is about the physics behind the nonlinear behavior of musical instrument strings. These theoretical and experimental results form the basis of sound synthesis methods presented in Chap. 6. Here we deal with geometric nonlinearities, and the nonlinearities arising from the string material are neglected. The reason for the geometric nonlinearities is that above a certain amplitude of vibration the length of the string cannot be assumed constant. The change of string length varies the tension of the string, leading to the generation of longitudinal motion and new transverse components.

The classification of the different nonlinear phenomena has not been given in the literature. Section 5.1 fills this gap by discussing the main features of geometric nonlinearities and presenting a “nonlinearity map” that helps to estimate the nature of string behavior as a function of material properties and the amplitude and bandwidth of vibration. Section 5.2 reviews the literature describing the phenomenon with spatially uniform tension. This is followed by the main scientific contribution of this chapter, namely, the investigation of the motion of longitudinal modes in Sec. 5.3. A new theoretical model is presented that is based on the modal formulation of the transverse and longitudinal vibration, and on the assumption that the longitudinal to transverse coupling can be neglected. Due to this approximation it is possible to analytically compute the longitudinal components for arbitrary transverse vibrations. The results show good agreement with the measurements of other authors and give the theoretical explanation of the experiments. Section 5.4 presents some findings about the bidirectional coupling of transverse and longitudinal modes. The results of Sec. 5.4 are preliminary, as this research is still going on.

5.1 Classification of Nonlinear String Behavior

This section investigates the factors that influence the significance of nonlinear behavior. The goal is to give a classification of the geometric nonlinearity of the string. For that, we estimate the relative significance of the nonlinear components in the bridge force compared to the standard linear components.

5.1.1 String Equations

For simplicity, it is assumed in this section that the string is vibrating in one plane, i.e., one transverse and one longitudinal polarization are present. Losses and dispersion are neglected and we assume that $T_0 \ll ES$, which holds for metal strings. For rubber-like strings, the following derivations are not accurate. However, they still show the qualitative behavior of such strings.

From Eqs. (2.4) and (2.9), with the above assumptions in mind, the string tension is approximated as

$$T = T_0 + ES \left(\frac{ds}{dx} - 1 \right) = T_0 + ES \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 \right]. \quad (5.1)$$

From Eq. (2.10) with $ES + T_0 \approx ES$ the equation for the longitudinal displacement $\xi(x, t)$ is

$$\mu \frac{\partial^2 \xi}{\partial t^2} = ES \frac{\partial^2 \xi}{\partial x^2} + \frac{1}{2} ES \frac{\partial \left(\frac{\partial y}{\partial x} \right)^2}{\partial x}. \quad (5.2)$$

We recall that Eq. (5.2) is a standard one-dimensional wave equation with an additional force term nonlinearly depending on the transverse vibration $y(x, t)$.

From Eq. (2.11) the wave equation for the transverse motion can be written as

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} + ES \frac{\partial \left\{ \frac{\partial y}{\partial x} \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 \right] \right\}}{\partial x}, \quad (5.3)$$

which is again a one-dimensional wave equation with an additional force term depending on the product of the transverse slope and the tension variation.

From the musical acoustics point of view it is more important to know what the force is at the termination (e.g., bridge of the instrument), as the radiated sound pressure is proportional to this force. Note that we assume that the string termination cannot exchange energy between the two polarizations. The bridge force in the longitudinal direction can be approximated by the tension variation at the termination of the string ($x = L$) as

$$F_l(t) = -[T(L, t) - T_0] = -ES \left[\frac{\partial \xi}{\partial x} \Big|_{x=L} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \Big|_{x=L} \right)^2 \right], \quad (5.4)$$

showing that the force $F_l(t)$ depends not only on the longitudinal motion but on the transverse vibration as well. Note that T_0 has been subtracted from $T(L, t)$ because it only acts as a constant strain on the instrument body, which does not appear in the radiated sound.

The transverse force $F_t(t)$ at the bridge is the product of the string slope $\partial y / \partial x$ and the tension $T(x, t)$:

$$F_t(t) = -T(L, t) \frac{\partial y}{\partial x} \Big|_{x=L} = -T_0 \frac{\partial y}{\partial x} \Big|_{x=L} - ES \left[\frac{\partial \xi}{\partial x} \Big|_{x=L} \frac{\partial y}{\partial x} \Big|_{x=L} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \Big|_{x=L} \right)^3 \right], \quad (5.5)$$

again showing that the transverse force at the bridge depends on both the transverse and longitudinal string motion. However, for small vibration amplitudes (linear behavior) only the first term is significant.

Note that Eqs. (5.1)–(5.5) become more complicated in the case of rubber-like strings, where the assumption $ES \gg T_0$ does not hold. The correct equations could be simply derived from Eqs. (2.9)–(2.11), but then we would also sacrifice the simplicity of Eqs. (5.7) and (5.8).

5.1.2 Classification

It can be seen from Eqs. (5.2)–(5.5) that the character of string vibration depends not only on the physical properties, but also on the amplitude of vibration. As musical instrument strings are generally excited in the transverse polarization, we will concentrate on the effect coming from the variation of the transverse slope $\partial y/\partial x$. The magnitude of the transverse force at the termination ($x = L$) is computed by the Euclidean norm (root mean square value), and is referred as $\|\partial y/\partial x\|$.

By looking at Eqs. (5.2)–(5.5), the amplitude of the different nonlinear components can be expressed as a function of the amplitude of transverse slope and the physical parameters of the string. The derivation and some simulations are included in the Appendix A.1, here only the results are presented. From Eq. (5.5) it follows that the linear transverse component (the component which would arise if the string was ideal) of the bridge force $F_{t,\text{lin}}$ has the magnitude

$$\|F_{t,\text{lin}}\| = T_0 \left\| \frac{\partial y}{\partial x} \right\|. \quad (5.6)$$

From Eqs. (5.2) and (5.4) the magnitude of longitudinal force at the bridge $\|F_l\|$ is approximately described by

$$\|F_l\| \approx C_1 ES \left\| \frac{\partial y}{\partial x} \right\|^2, \quad (5.7)$$

where C_1 is a constant in the order of unity, which depends on the type of string excitation. From Eqs. (5.3) and (5.5) the magnitude of the nonlinear transverse component can be approximated as a third order function of transverse slope:

$$\|F_{t,\text{nonlin}}\| \approx C_t ES \left\| \frac{\partial y}{\partial x} \right\|^3, \quad (5.8)$$

where C_t is a constant in the order of unity.

Let us assume that the longitudinal force F_l is significant, if its Euclidean norm $\|F_l\|$ reaches the 10% (–20 dB) of the transverse linear component $\|F_l\| = 0.1\|F_{t,\text{lin}}\|$ in Eqs. (5.6) and (5.7), giving

$$\frac{0.1}{C_1} = \left(\sqrt{\frac{ES}{T_0}} \right)^2 \left\| \frac{\partial y}{\partial x} \right\|. \quad (5.9)$$

Similarly, the parameter values where the nonlinear transverse component is 20 dB lower than the linear one are on the line

$$\frac{0.1}{C_t} = \left(\sqrt{\frac{ES}{T_0}} \right)^2 \left\| \left\| \frac{\partial y}{\partial x} \right\| \right\|^2. \quad (5.10)$$

Note that the choice of 0.1 (−20 dB) is almost arbitrary, as far as it is lower than unity. It was mainly motivated by the fact that the theoretical curves of Eqs. (5.7) and (5.8) are reasonable approximations when the nonlinear transverse component is smaller than the linear transverse one. It does not refer to a definite value where the specific component starts to be audible. However, we can still say that these are a kind of “underestimates”, meaning that nonlinear components under this limit are most probably masked by the transverse components, thus, are inaudible. Unfortunately, no research results exist in the literature that could estimate the perceptual significance of these components precisely.

In Eqs. (5.9) and (5.10) the parameter dependence is written as a function of $\sqrt{ES/T_0}$, as $\sqrt{ES/T_0}$ equals the ratio of the longitudinal and transverse fundamental frequencies

$$\frac{f_{\xi,0}}{f_0} = \sqrt{\frac{ES}{T_0}}, \quad (5.11)$$

where f_0 is the transverse and $f_{\xi,0}$ is the longitudinal fundamental frequency (this follows from Eq. (2.13) with $ES \gg T_0$). The change of $f_{\xi,0}$ may change the string behavior significantly, as discussed in the next subsection.

For musical instruments, $f_{\xi,0}/f_0 = \sqrt{ES/T_0}$ values around 3–5 are typical for nylon strings, while this value is around 10–20 for metal strings. Note that $\sqrt{ES/T_0}$ values in the order of 100 correspond to loosely stretched strings, which are often used in experimental setups, as in this case the nonlinearity is larger, i.e., more easily observable. As for the slope, the value $\|\partial y/\partial x\| = 10^{-2}$ corresponds to a fortissimo hammer strike (5 m/s hammer velocity) in the case of a piano string.

The curve of Eq. (5.9) is plotted by a solid line in Fig. 5.1 and the function of Eq. (5.10) by a dashed line. Above these lines the longitudinal and nonlinear transverse components are considered to be significant. The longitudinal component is generated by the transverse to longitudinal coupling, while the nonlinear transverse one by the longitudinal to transverse coupling mechanism (see the arrows in the right hand side of Fig. 5.1).

Spectral Content of the Excitation

Another important factor that influences the nature of nonlinear vibration is the spectral content of the transverse vibration. If all the longitudinal modes are excited under their resonance frequency, the tension can be considered uniform along the string, as will be discussed in Sec. 5.2. Here we only note that in this case the term $\mu(\partial^2 \xi/\partial t^2)$ in Eq. (5.2) equals zero, i.e., the inertial effects of longitudinal modes are negligible. The string tension can be computed from the elongation of the string, which depends on $y(x, t)$ directly. Accordingly, the longitudinal motion does not have to be computed to obtain the transverse

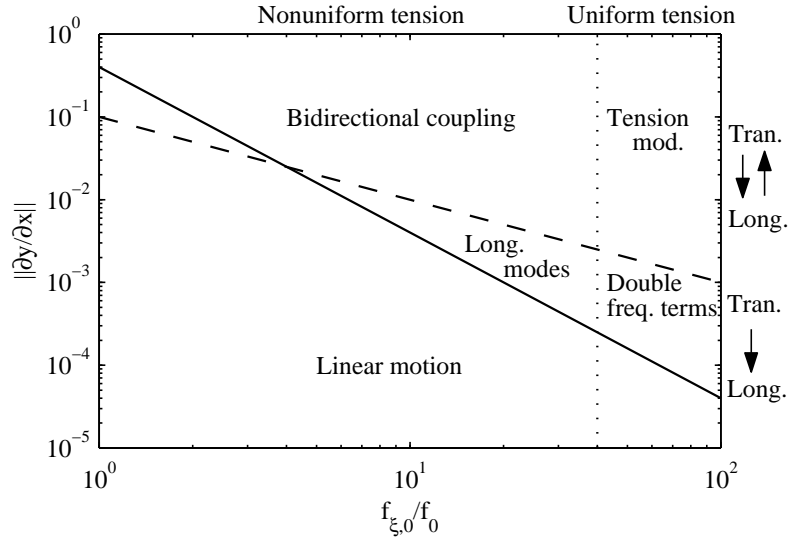


Figure 5.1: Classification of the nonlinear string behavior. At parameter values above the solid line the longitudinal component becomes significant compared to the linear transverse one. Above the dashed line the nonlinear transverse component starts to appear. The arrows between “Tran.” and “Long.” show which directions of coupling are significant. On the right-hand side of the dotted line, the tension can be considered spatially uniform along the string (assuming 10 significant transverse partials).

and longitudinal bridge force, leading to a large simplification from the sound synthesis point of view.

The bandwidth of the force exciting the longitudinal modes is the double of that of the transverse motion, as it is generated by a second-order nonlinearity (see Eq. (5.2)). Thus, the excitation bandwidth is $2Nf_0$ for harmonic strings, if all the transverse modes are significant up to the mode number N . The longitudinal components responsible for the inertial effects are assumed to be negligible if all the longitudinal modes are excited below the half of their resonance frequency (this will be outlined in Sec. 5.3.4 in detail). Consequently, we can suppose that the tension is uniform along the string if the bandwidth of the excitation, $2Nf_0$, is lower than the half of the lowest longitudinal resonance frequency, $f_{\xi,0}/2$, leading to

$$4N < \frac{f_{\xi,0}}{f_0}. \quad (5.12)$$

This is indicated in Fig. 5.1 as a dotted line for a transverse vibration containing the first 10 partials. Thus, for $f_{\xi,0}/f_0 = \sqrt{ES/T_0}$ values higher than 40 the string tension can be considered uniform. This dashed line should be shifted to the right or to the left depending on whether there are more or less significant partials in the transverse vibration. Again, this line is a rough limit, as the real significance of the inertial effects depends both on the relative amplitudes of transverse modes and on the decay times of longitudinal modes.

It is interesting to note that while increasing $\sqrt{ES/T_0}$ complicates the string motion by increasing the effect of the nonlinear terms, it also changes the nature of string behavior

by raising the longitudinal modal frequencies. Above a certain value the tension becomes spatially uniform along the string, leading to a motion which can be explained by simpler equations.

Dependence on Inharmonicity

The perceptual significance of the nonlinear components also depends on the inharmonicity. When B is negligible, the partial series is almost harmonic, thus, the nonlinearly generated frequencies, which are generally the sum or difference of transverse modal frequencies, appear at the same frequencies as the original modes, thus, their effect is less audible. If B is large, these nonlinear sum and difference components appear between the transverse ones, leading to beating, thus, to a more audible effect. On the other hand, the nonlinear energy exchange between the transverse modes might be more efficient for small B values, as in that case the excitation frequencies are close to the resonant frequencies. For homogenous, unwrapped strings the inharmonicity coefficient is computed by Eq. (2.40). As can be seen from Eq. (2.40) the inharmonicity is a second order function of $\sqrt{ES/T_0}$. This means that the increasing inharmonicity coefficient B reinforces the perceptual effect of the nonlinear components which increase with $\sqrt{ES/T_0}$ anyway. In any case, the nonlinear behavior of the string and the inharmonicity are interdependent, which is not included in Fig. 5.1.

5.1.3 Short Discussion of the Regimes of Nonlinearity

Figure 5.1 can be used to estimate whether a specific nonlinear phenomenon is significant for a given parameter set of the string. The borders separating the different regimes are at approximate positions, which may vary from instrument to instrument. In any case, the topology of Fig. 5.1 should be similar for all string instruments.

This section outlines the most important properties of the different regimes displayed in Fig. 5.1. These are also summarized in Table 5.1, showing whether the transverse to longitudinal and the longitudinal to transverse coupling is significant for a class of Fig. 5.1 or not. The last column signs if the longitudinal inertial effects are significant, i.e., if the longitudinal modes have to be modeled or not. If not, then the tension is spatially uniform along the string and can be simply computed from the transverse displacement. The detailed descriptions of the different classes and the corresponding modeling methods will be given later in the next sections.

Linear Motion

When $\sqrt{ES/T_0}$ and $\partial y/\partial x$ are small, the string obeys the standard linear wave equation. In this case the transverse and longitudinal polarizations are independent. Therefore, if only the transverse polarization is excited (which is generally the case), the longitudinal motion is negligible. This kind of motion has been discussed in Sec. 2.2.3.

	Tran. to long. coupl.	Long. to tran. coupl.	Long. inertial effects
Linear motion			
Double freq. terms	×		
Tension modulation	×	×	
Longitudinal modes	×		×
Bidirectional coupling	×	×	×

Table 5.1: Main features of the different regimes of string behavior according to Fig. 5.1. The “×” sign means that the specific feature of vibration is significant, i.e., it has to be included in the model.

Double Frequency Terms

In this case the longitudinal modes are excited under their resonant frequency, hence the string tension varies with time but spatially uniform along the string. This means that the tension can be computed directly from the transverse displacement. The tension variation is significant compared to the transverse force at the bridge, but it is negligible compared to the initial tension T_0 , so it cannot excite any “nonlinear” transverse modes. The longitudinal force component will include terms having double the frequency of transverse modes. This type of motion will be covered in Sec. 5.2.

Tension Modulation

This case is similar to the previous one in a way that the tension is spatially uniform along the string, but now the temporal variation of the tension is no longer negligible in comparison with T_0 . The temporal modulation of tension leads to the nonlinear excitation of transverse modes, nonplanar motion, and pitch glide. This regime of string motion is studied in Sec. 5.2.

Modeling of Longitudinal Modes

In this case the frequencies of the excitation terms in Eq. (5.2) are around or above the longitudinal modal frequencies. As a result, the tension varies with both time and space along the string. This leads to the appearance of odd and even phantom partials and the free motion of longitudinal modes. As the tension variation is significant compared to the transverse bridge force but small compared to T_0 , the longitudinal motion has a significant contribution to the sound but does not influence the transverse vibration. For modeling, the largest difference from the previous two cases is that now the motion of longitudinal modes also has to be computed. On the other hand, the longitudinal to transverse coupling does not have to be implemented. This type of string behavior is discussed in Sec. 5.3.

Bidirectional Coupling

Here neither the tension is uniform along the string, nor its variation is negligible in comparison with T_0 . This is the most complex situation, since odd and even phantom

partials and the longitudinal free modes also appear and they influence the transverse motion by generating new transverse components. A model describing such a motion should thus include the precise computation of longitudinal motion and the implementation of the longitudinal to transverse coupling. This is further investigated in Sec. 5.4.

5.2 Spatially Uniform Tension

In this section we will review the findings of the literature about the ‘‘Tension modulation’’ regime of Fig. 5.1. For the ‘‘Double frequency terms’’, the tension variation is computed in the same way but the transverse vibration is calculated by assuming $T = T_0$, as the tension variation is so small that it has a negligible effect on the transverse vibration. Note that it can still be audible as it provides a longitudinal force variation at the instrument body.

5.2.1 Theories and Experiments

This kind of nonlinear string motion has been already investigated by Kirchhoff in the late 19th century, then revisited by Carrier in the middle of the last century. Therefore, the governing equations are referred as Kirchhoff-Carrier equations by some authors.

Oplinger [1960] states that when $T_0/ES \ll 1$, then the inertial effects of longitudinal modes can be neglected, the tension is spatially uniform along the string and can be directly computed from the elongation of the string according to the Hooke’s law

$$\overline{T} = T_0 + ES[(L' - L)/L], \quad (5.13)$$

where L' is the actual length of the string and L is the minimum length at equilibrium. The overline in \overline{T} emphasizes that the tension is spatially uniform along the string. The length L' equals the length of the curve $y(x, t)$ for a given t and $0 \leq x \leq L$ and is given by

$$L' = \int_0^L \sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2} dx \approx L + \frac{1}{2} \int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx. \quad (5.14)$$

The substitution of Eq. (5.14) into Eq. (5.13) gives

$$\overline{T} = T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx. \quad (5.15)$$

It is an interesting result that this is the same as spatially averaging (i.e., integrating from 0 to L along x and dividing by L) the space-dependent equation of tension of Eq. (5.1).

The approximate equation describing the transverse motion is

$$\mu \frac{\partial^2 y}{\partial t^2} = \frac{\partial \left(\overline{T} \frac{\partial y}{\partial x} \right)}{\partial x}. \quad (5.16)$$

Inserting Eq. (5.15) into Eq. (5.16) yields

$$\mu \frac{\partial^2 y}{\partial t^2} = \left[T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx \right] \frac{\partial^2 y}{\partial x^2} \quad (5.17)$$

which is the Kirchhoff-Carrier equation used by most of the papers as a starting point (although sometimes extended to the z polarization). Oplinger [1960] has given the frequency response curves of the string based on the solution of Eq. (5.17), showing the warped peaks typical of nonlinear forced vibrations. This type of response curve gives rise to the jump phenomenon, where the string can suddenly shift from one state of behavior to another at a given frequency. For example, it can shift from the first to the second mode, resulting in an amplitude jump. This has been also observed experimentally in [Oplinger 1960].

The same equation has been derived by Anand [1969] for the three dimensional case by assuming $\partial^2 \xi / \partial t^2 = 0$ in Eq. (5.2). Integrating the such modified Eq. (5.2) with respect to x twice gives a memoryless relation of ξ and y , meaning that the longitudinal displacement $\xi(x, t)$ for a given t can be directly computed when $y(x, t)$ is known. Inserting the equation for $\xi(x, t)$ into Eq. (5.1) gives the same result as Eq. (5.17). The three dimensional version of Eq. (5.17) takes the form

$$\mu \frac{\partial^2 y}{\partial t^2} = \left\{ T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left[\left(\frac{\partial y}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial x} \right)^2 \right] dx \right\} \frac{\partial^2 y}{\partial x^2} \quad (5.18a)$$

$$\mu \frac{\partial^2 z}{\partial t^2} = \left\{ T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left[\left(\frac{\partial y}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial x} \right)^2 \right] dx \right\} \frac{\partial^2 z}{\partial x^2} \quad (5.18b)$$

As for the applicability of Eq. (5.18), Anand [1969] states that the condition $ES/T_0 \gg 1$ given by Oplinger [1960] is sufficient only if the order of the transverse modes is small compared to $\sqrt{ES/T_0}$. Anand [1969] investigates the nonplanar motion of the string and finds that the nonlinearity leads to intermodal coupling but does not generate modes that are not present at the initial instant. He also suggests that the planar vibration of an undamped string is not stable.

Narasimha [1968] also arrives at the same equations through a different demonstration, but adds the effect of damping. He investigates the critical amplitudes and frequencies where the planar motion becomes unstable, finding that a slight damping raises the critical amplitudes and frequencies. As a result, planar motion of a damped string may be stable, as opposed to the undamped string, where the critical amplitude is zero at the fundamental frequency of the string.

The analysis of Gough [1984] (based again on Eq. (5.18) with damping) confirms the prediction of Anand [1969] that the nonlinearity leads to a precession of the orbital motion. He shows both theoretically and experimentally that the precessional frequencies can be simply related to the geometric properties of the orbital motion.

A more sophisticated analysis is given by Watzky [1992], describing the damped free vibration of the stiff string. Its main novelty is the inclusion of torsional vibration as coupling term between the transverse modes. He suggests that the effect of longitudinal modes can be neglected (i.e., the tension can be considered spatially uniform) if all the transverse modes are far from half of a longitudinal resonance frequency. (We will see in Sec. 5.3.4 that the right condition is a bit more complicated). In any case, the inertial effects are neglected in [Watzky 1992], too.

Hanson et al. [1994] have made thorough experiments on the vibration of a loosely

stretched red brass harpsichord wire. Most of their results are in agreement with the theoretical studies quoted above. One exception is that they suggest that there is no critical frequency for the onset of motion perpendicular to the driving force (which they call “ z motion”). They have found “ z motion” quite far below resonance, which continuously grows as the driving frequency approaches the resonance frequency, and that this “ z motion” does not necessarily mean nonplanar vibration.

5.2.2 Nonlinear Generation of Missing Modes

The literature quoted in the previous section investigates only the first (or the first few) transverse modes. As in a musical instrument string dozens of transverse modes vibrate, the above works cannot be directly applied in musical acoustics. There is one paper, however, which deals with the problem from this point of view. The purpose of Legge and Fletcher [1984] was to investigate the intermodal coupling of strings. That is, how a specific transverse mode can gain energy from another transverse mode.

We reproduce the derivation of tension in [Legge and Fletcher 1984] in a more general formulation, by writing the transverse displacement in its modal form

$$y(x, t) = \sum_{n=1}^{\infty} y_n(t) \sin\left(\frac{n\pi x}{L}\right), \quad (5.19)$$

where $y_n(t)$ is the instantaneous amplitude of the transverse mode n , as introduced in Sec. 2.2.4. Inserting Eq. (5.19) into Eq. (5.15) gives

$$\begin{aligned} \bar{T}(t) &= T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left[\sum_{n=1}^{\infty} y_n(t) \frac{n\pi}{L} \cos\left(\frac{n\pi x}{L}\right) \right]^2 dx = \\ &T_0 + \frac{\pi^2 ES}{4L^3} \int_0^L \left\{ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} m n y_m(t) y_n(t) \left[\cos\left(\frac{(m+n)\pi x}{L}\right) + \cos\left(\frac{(m-n)\pi x}{L}\right) \right] \right\} dx. \end{aligned} \quad (5.20)$$

After performing the integration, all the terms cancel out, except the rightmost cosine term for $m = n$, giving

$$\bar{T}(t) = T_0 + \frac{\pi^2 ES}{4L^2} \sum_{n=1}^{\infty} n^2 y_n^2(t) \quad (5.21)$$

Writing the instantaneous amplitudes as in [Legge and Fletcher 1984], i.e., in the form of exponentially decaying sinusoidal functions

$$y_n(t) = A_n \sin(\omega_n t + \varphi_n) e^{-\frac{t}{\tau_n}}, \quad (5.22)$$

yields the expression for tension

$$\bar{T}(t) = T_0 + \frac{\pi^2 ES}{8L^2} \sum_{n=1}^{\infty} n^2 A_n^2 [1 - \cos(2\omega_n t + 2\varphi_n)] e^{-\frac{2t}{\tau_n}}. \quad (5.23)$$

The first time-dependent part of Eq. (5.23) is a quasistatic increase of tension

$$T_{\text{qs}} = \frac{\pi^2 ES}{8L^2} \sum_{n=1}^{\infty} n^2 A_n^2 e^{-\frac{2t}{\tau_n}}, \quad (5.24)$$

which decays slowly. This leads to a proportional increase in modal frequencies, giving a relative change of $\sqrt{(T_0 + T_{\text{qs}})/T_0}$. This shift decreases as a function of time, leading to a pitch glide effect. That is, the initial pitch of the string is higher than in the latter part, due to the decaying amplitude of vibration.

The second part contains the double frequency terms

$$T_{\text{df}} = -\frac{\pi^2 ES}{8L^2} \sum_{n=1}^{\infty} n^2 A_n^2 \cos(2\omega_n t + 2\varphi_n) e^{-\frac{2t}{\tau_n}}, \quad (5.25)$$

leading to a continuous modulation of tension, built up of sinusoidal functions having double the frequencies of transverse modes. The amplitude of this modulation decays exponentially, and the decay times of its components are the half compared to that of the originating transverse modes.

Substituting Eq. (5.23) into Eq. (5.16) and concentrating on the effects of double frequency terms leads us to the observation that the different transverse modes cannot efficiently exchange energy if the string is rigidly terminated. This can be explained when we consider the motion of transverse mode m . The driving force $T\partial^2 y/\partial x^2$ for mode m will contain the frequencies $2\omega_n \pm \omega_m$ for all n . However, for effective excitation the excitation frequency should be near to the resonant frequency ω_m , which is satisfied only in the case of $n = m$. Thus, mode m can only act on itself and the other modes cannot influence its motion. As a practical result, if a mode is missing from a vibrating string, it cannot be generated by nonlinear coupling in the case of infinitely rigid terminations [Legge and Fletcher 1984].

If the bridge is not infinitely rigid, but has the admittance $Y(\omega)$ at $x = L$, then a different situation occurs. The most important difference is, besides a change in the transverse modal frequencies and decay times, that the spatial distribution of modes also changes. Practically, L has to be replaced by $L + \delta L_n$ in Eq. (5.19). The string behaves as a lossy string rigidly supported at 0 and $L + \delta L_n$. This virtual change of length can be different for all the modes, and approximately computed as

$$\delta L_n \approx \frac{T_0}{\omega_n} \text{Im}\{Y(\omega_n)\}. \quad (5.26)$$

In this case the force acting on the bridge contains the frequencies $2\omega_n \pm \omega_m$. However, a significant difference arises from the rigid bridge because now the string modes are virtually supported at $L + \delta L_n$, but excited by the bridge movement at $x = L$, thus, all the modes can gain energy from the bridge motion. Strong coupling arises when the excitation and resonance frequencies are near, thus, the mode p can gain energy from modes $p = 2m \pm n$, as $\omega_p \approx 2\omega_m \pm \omega_n$ [Legge and Fletcher 1984].

For a more realistic bridge, when the string passes the bridge at an angle, the tension Eq. (5.23) directly appears in the bridge movement. This means that the double frequency

terms $2\omega_n$ can directly excite any of the transverse modes. Naturally, effective excitation will arise when $p = 2n$, as in this case the excitation frequency $2\omega_n$ will be close to the resonance frequency ω_p of mode p [Legge and Fletcher 1984].

Under these circumstances, it is possible that an initially unexcited mode can gain energy from the other transverse modes. This has been justified by experiments in [Legge and Fletcher 1984]. Moreover, they note that things get even more complicated when the modal frequencies of the string are not harmonic, as in that case the different coupling mechanisms ($2\omega_m \pm \omega_n$ and $2\omega_n$) excite mode p at slightly different frequencies, which both are a bit different from ω_p , resulting in beating and amplitude fluctuations.

5.3 Modeling of Longitudinal Modes

In Sec. 5.2 the tension was considered spatially uniform along the string and could be directly computed from the transverse displacement, as the dynamics (inertial effects) of longitudinal modes were negligible. However, this assumption is no longer acceptable for the class “Longitudinal modes” of Fig. 5.1 as there the longitudinal modes are excited near or above their resonance frequency. Therefore, in this section we investigate how the longitudinal vibration is excited by the transverse string motion. As the tension variation is negligible in comparison to T_0 , the longitudinal to transverse coupling is neglected. As a result, both the transverse and longitudinal wave equations can be considered as linear systems, where the excitation of the latter is a nonlinear function of the transverse string shape. Thus, the equations of Sec. 2.2.3 can be directly applied, leading to an analytical solution. The tension will be decomposed into a spatially uniform part (for what the results outlined in Sec. 5.2 hold) and to a space-dependent part, at which we look in more detail. We assume that the string is vibrating in one plane, i.e., one transverse polarization is present.

5.3.1 Prior Work

As for the theories and numerical simulations, the motion of longitudinal modes has been investigated numerically in [Leissa and Saad 1994] by the Galerkin method, showing non-periodic regimes of motion for low strain (i.e., loosely stretched strings) or/and large vibration amplitudes. Leamy and Gottlieb [2000] describe the whirling motion of strings, including longitudinal vibrations and material nonlinearities for rubber-like strings. Kurmyshev [2003] investigates rubber-like strings too, showing that the single-mode motion of a rubber-like string may evolve to multi-mode motion due to the coupling of the different polarizations. We note that for rubber-like strings $T_0 \approx ES$, meaning that the longitudinal fundamental frequency is not much higher than the transverse one. In this case nonlinearity arises only at extremely large amplitudes, i.e., when the string displacement is in the order of string length. Although nylon strings of musical instruments are almost rubber-like, in those cases the nonlinearity is negligible because of much smaller string displacement.

It is a general feature of the studies cited above that they investigate only the first few modes and also consider longitudinal to transverse coupling. Therefore, their results

cannot be directly applied for the present purposes. This is because the present problem is more complex in the sense that not only the first few but 20 to 100 transverse modes have to be taken into account in the case of musical instruments. On the other hand, it is simpler in the way that the longitudinal to transverse coupling and the coupling of different transverse modes are not investigated, as the primary interest is now on the longitudinal vibration itself. Therefore, in this section a new modal model is developed that computes the spectrum of the longitudinal vibration in the case of arbitrary transverse modal frequencies.

From the musical acoustics point of view, the importance of longitudinal vibration of piano strings was recognized long ago by piano builders. Conklin [1996] demonstrated that the pitch relation of the transverse and longitudinal component strongly influences the quality of the tone and described a method to tune these components. Giordano and Korty [1996] found that the amplitude of the longitudinal vibration is a nonlinear function of the amplitude of transverse vibration, confirming the assumption that the longitudinal component is generated by the nonlinearity of the string and not by the “misalignment” of the hammer.

Nakamura and Naganuma [1993] found a second series of partials in piano sound spectra having one-fourth of inharmonicity compared to the main partial series. They attributed these to the horizontal polarization of the string, but they have actually found the series that was later named “phantom partials” by Conklin [1999]. Conklin has pointed out that the phantom partials are generated by nonlinear mixing and their frequencies are the sum or difference of transverse modal frequencies. He named “even phantoms” those having double the frequency ($2f_n$) of a transverse mode, and “odd phantoms” those which appear at the sum $f_m + f_n$ or difference $f_m - f_n$ frequencies of two transverse modes. Conklin’s measurements have shown that odd phantoms generally originate from adjacent parents, i.e., can be found at $f_5 + f_6$ rather than at $f_4 + f_7$. Phantom partials have also been found in the spectrum of guitar tones [Conklin 1999]. In a recent paper about guitar transients, Woodhouse [2004b] states that the amplitude of phantom partials seems to be modulated according to the longitudinal modal frequencies. This section gives a theoretical explanation for these experimental findings and provides some new measurement results.

5.3.2 Equations of Motion

The wave equation for the longitudinal motion can be derived from Eq. (5.2) by adding losses similarly to Eq. (2.25):

$$\mu \frac{\partial^2 \xi}{\partial t^2} = ES \frac{\partial^2 \xi}{\partial x^2} - 2R_\xi(\omega) \mu \frac{\partial \xi}{\partial t} + \frac{1}{2} ES \frac{\partial \left(\frac{\partial y}{\partial x} \right)^2}{\partial x}, \quad (5.27)$$

where $R_\xi(\omega)$ in Eq. (5.27) is the frequency dependent frictional resistance of the longitudinal polarization. Here we assume that the longitudinal polarization is not affected by the excitation directly, but gains energy from the transverse polarization. By comparing with Eq. (2.25) it can be noticed that the two equations are of the same form, and they differ only in their parameters: T_0 is substituted by ES , and the dispersion term (fourth order

spatial derivative) is missing. The external force density $d_y(x, t)$ is replaced by

$$d_\xi(x, t) = \frac{1}{2}ES \frac{\partial \left(\frac{\partial y}{\partial x} \right)^2}{\partial x}. \quad (5.28)$$

Note that we again assume that the string is metallic, i.e., $ES \gg T_0$. This is reasonable, as in the case of rubber like strings the longitudinal motion is negligible for transverse slopes typical in musical instruments.

The formal similarity to Eq. (2.25) means that the results of Sec. 2.2.4 can be directly used. By assuming infinitely rigid terminations, the longitudinal displacement can be written in its modal form

$$\xi(x, t) = \sum_{k=1}^{\infty} \xi_k(t) \sin \left(\frac{k\pi x}{L} \right). \quad (5.29)$$

Accordingly, the instantaneous amplitude $\xi_k(t)$ of the longitudinal mode k is obtained as

$$\xi_k(t) = F_{\xi,k}(t) * \xi_{\delta,k}(t), \quad (5.30a)$$

$$F_{\xi,k}(t) = \int_0^L d_\xi(x, t) \sin \left(\frac{k\pi x}{L} \right) dx, \quad (5.30b)$$

$$\xi_{\delta,k}(t) = \frac{2}{L\mu} \frac{e^{-\sigma_{\xi,k}t}}{\omega_{\xi,k}} \sin(\omega_{\xi,k}t), \quad (5.30c)$$

where the $*$ sign denotes time-domain convolution and $F_{\xi,k}(t)$ is the excitation force acting on the longitudinal mode k . The time-domain impulse response of longitudinal mode k is denoted by $\xi_{\delta,k}(t)$, where $\omega_{\xi,k} = 2\pi f_{\xi,k}$ and $\sigma_{\xi,k} = 1/\tau_{\xi,k}$ stand for the angular frequency and decay rate of the longitudinal mode k . Note that the subscript ξ in $f_{\xi,k}$, $\omega_{\xi,k}$, $\sigma_{\xi,k}$, and $\tau_{\xi,k}$ is used to distinguish the longitudinal variables from the transverse ones.

For small frictional resistance, the longitudinal modal frequencies are

$$f_{\xi,k} = kf_{\xi,0} = \frac{k}{2L} \sqrt{\frac{ES}{\mu}}, \quad \text{or,} \quad \omega_{\xi,k} = k\omega_{\xi,0} = \frac{\pi k}{L} \sqrt{\frac{ES}{\mu}}, \quad (5.31)$$

where $f_{\xi,0} = f_{\xi,1}$ is the fundamental frequency of the longitudinal vibration. The decay rates and decay times are simply written as

$$\sigma_{\xi,k} = R(\omega_{\xi,k}) \quad \text{and} \quad \tau_{\xi,k} = \frac{1}{R(\omega_{\xi,k})}. \quad (5.32)$$

The first step in calculating the longitudinal motion is the computation of the excitation force $F_{\xi,k}(t)$ by Eq. (5.30b), which is the scalar product of the excitation-force density $d_\xi(x, t)$ and the longitudinal modal shape. Expressing the transverse string shape in its modal form

$$y(x, t) = \sum_{n=1}^{\infty} y_n(t) \sin \left(\frac{n\pi x}{L} \right) \quad (5.33)$$

and inserting Eq. (5.33) into Eq. (5.28) gives

$$d_\xi(x, t) = \frac{1}{2}ES \frac{\partial \left[\sum_{n=1}^{\infty} y_n(t) \frac{n\pi}{L} \cos \left(\frac{n\pi x}{L} \right) \right]^2}{\partial x}, \quad (5.34)$$

which, after some derivations, takes the following form:

$$d_{\xi}(x, t) = -ES \frac{\pi^3}{4L^3} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} y_m(t) y_n(t) m n \times \\ \times \left[(m+n) \sin\left(\frac{m+n}{L} \pi x\right) + (m-n) \sin\left(\frac{m-n}{L} \pi x\right) \right]. \quad (5.35)$$

Note that the indices m and n belong to variables of transverse modes throughout the chapter. The variables of longitudinal modes are indexed by k .

From Eqs. (5.35) and (5.30b) it follows that $F_{\xi,k}(t)$ is nonzero for $m+n=k$ and $|m-n|=k$ only, since in all other cases the spatial distribution of the excitation $d_{\xi}(x, t)$ is orthogonal to the modal shape of mode k , which is $\sin(k\pi x/L)$. In other words, a longitudinal mode with mode number k is excited by such transverse mode pairs, for which either the sum or the difference of their mode numbers equal to k .

The two cases can be computed separately by defining $F_{\xi,k}(t)$ as a sum of two components, i.e., $F_{\xi,k}(t) = F_{\xi,k}(t)^+ + F_{\xi,k}(t)^-$. The component originating from the transverse modes that satisfy $m+n=k$ is

$$F_{\xi,k}(t)^+ = -ES \frac{\pi^3}{8L^2} \sum_{n=1}^{k-1} y_{k-n}(t) y_n(t) k(k-n)n. \quad (5.36a)$$

The component coming from $|m-n|=k$ becomes

$$F_{\xi,k}(t)^- = -2ES \frac{\pi^3}{8L^2} \sum_{n=1}^{\infty} y_{k+n}(t) y_n(t) k(k+n)n. \quad (5.36b)$$

The factor of 2 in Eq. (5.36b) comes from the fact that there are two equal series $m=k+n$ and $n=k+m$, since both satisfy $|m-n|=k$.

If the instantaneous amplitudes $y_n(t)$ of the transverse modes are known, the longitudinal string displacement can be directly computed by the use of Eqs. (5.36), (5.30) and (5.29).

5.3.3 Longitudinal Motion in the Case of Exponentially Decaying Transverse Modes

For the freely vibrating, dispersive, lossy, and rigidly terminated string the transverse displacement for a given position $0 \leq x \leq L$ and time $t \geq 0$ can be written in the following form:

$$y(x, t) = \sum_{n=1}^{\infty} y_n(t) \sin\left(\frac{n\pi x}{L}\right) = \sum_{n=1}^{\infty} A_n \cos(\omega_n t + \varphi_n) e^{-\sigma_n t} \sin\left(\frac{n\pi x}{L}\right), \quad (5.37)$$

where ω_n is the angular frequency, σ_n is the decay rate, A_n is the initial amplitude, and φ_n is the initial phase of the transverse mode n . This form is of particular interest since the motion of the struck or plucked strings can be described similarly after the excitation (which, for example, lasts for 1–2 ms in the case of the piano). The steady state vibration

of a continuously excited (e.g., bowed) string can also be approximated by Eq. (5.37) with $\sigma_n = 0$.

In Eq. (5.36) all the terms are products of two transverse modal amplitudes:

$$\begin{aligned} y_m(t)y_n(t) &= A_m \cos(\omega_m t + \varphi_m) e^{-\sigma_m t} A_n \cos(\omega_n t + \varphi_n) e^{-\sigma_n t} = \\ &= \frac{1}{2} A_m A_n e^{-t(\sigma_m + \sigma_n)} [\cos((\omega_m + \omega_n)t + \varphi_m + \varphi_n) + \cos((\omega_m - \omega_n)t + \varphi_m - \varphi_n)]. \end{aligned} \quad (5.38)$$

It can be seen from Eq. (5.38) that two sinusoidal terms arise with the sum and difference frequencies of the originating modes. Both terms have the same amplitude $0.5A_m A_n$ and the same decay rate, which is the sum of σ_m and σ_n .

The excitation force of mode k can be computed by inserting Eq. (5.38) into Eq. (5.36). This will lead to a sum of exponentially decaying sinusoidal terms

$$F_{\xi,k}(t) = \sum_{p=1}^P A_p e^{-\sigma_p t} \cos(\omega_p t + \varphi_p), \quad (5.39)$$

where the parameters A_p , σ_p , ω_p , and φ_p can be computed from Eqs. (5.36) and (5.38).

As the equation describing the longitudinal modes is linear (see Eq. (5.30)), its response can be computed for the individual terms of Eq. (5.39) separately.

The Laplace transform of the impulse response of longitudinal mode k (covered by Eq. (5.30c)) is

$$\mathcal{L}\{\xi_{\delta,k}(t)\} = \frac{2}{L\mu} \frac{1}{s^2 + 2\sigma_{\xi,k} s + \sigma_{\xi,k}^2 + \omega_{\xi,k}^2}, \quad (5.40)$$

The system characterized by Eq. (5.40) has a free response (or, the homogenous solution of the corresponding differential equation) that is an exponentially decaying sinusoidal function with the angular frequency $\omega_{\xi,k}$ and decay rate $\sigma_{\xi,k} = 1/\tau_{\xi,k}$. Its forced response (the particular solution of the differential equation) to the terms of Eq. (5.39) is again an exponentially decaying sinusoidal function with the frequency f_p and decay time τ_p .

If we assume that $\tau_p \gg \tau_{\xi,k}$, then we can compute the amplitude and phase of the forced response easily, since the transfer function derived from Eq. (5.40) (with the substitution $s = j\omega_p = j2\pi f_p$) describes the amplitude and phase change of the excitation signal when it passes through the system of Eq. (5.40). It can be seen from Eq. (5.40) that the longitudinal modes can be considered as second-order lowpass filters with a resonance around $f_{\xi,k}$. Excitation frequencies ω_p around $\omega_{\xi,k}$ are largely amplified, while for $\omega_p \ll \omega_{\xi,k}$ there is a small gain, and for $\omega_p \gg \omega_{\xi,k}$ there will be no output signal. Example frequency responses for the first two longitudinal modes of a piano string are displayed in Fig. 5.2 with dashed lines, giving an idea on how much the excitation terms are emphasized depending on their relation to the resonance frequency.

Note that the assumption $\tau_p \gg \tau_{\xi,k}$ usually holds for stringed instruments, as the decay time of the longitudinal modes are an order of a magnitude lower than that of the transverse modes. If one wishes to compute the amplitude and phase of the forced response more precisely, it can be done in an analytical way by taking the Laplace transform of Eq. (5.39), multiplying by Eq. (5.40) and performing the inverse Laplace transform.

Excitation Frequencies

For qualitative understanding of the longitudinal components, it is useful to look at the spectra of the excitation force series $F_{\xi,k}(t)$. The most important question is where the frequency peaks can be found.

To the first approximation, the instantaneous amplitudes $y_n(t)$ are decaying sinusoidal functions with the frequencies $f_n = \omega_n/(2\pi)$, as described by Eq. (5.37). By observing Eqs. (5.36) and (5.38), the frequencies of the mixing terms in $F_{\xi,k}(t)$ can be calculated as

$$\begin{aligned} &\text{Frequencies} \\ &\text{in } F_{\xi,k}(t)^+ : \begin{cases} f_n + f_{k-n} \approx f_k, \\ f_n - f_{k-n} \approx f_{|2n-k|}, \end{cases} \\ &\text{in } F_{\xi,k}(t)^- : \begin{cases} f_n + f_{k+n} \approx f_{2n+k}, \\ f_n - f_{k+n} \approx f_k, \end{cases} \end{aligned} \quad (5.41)$$

where the form f_a refers to the frequency of the transverse mode with mode number a .

Harmonic Transverse Vibration

The approximations in Eq. (5.41) become equalities if the transverse frequencies f_n are perfectly harmonic, i.e., $f_n = nf_0$, which is the case for string instruments having negligible inharmonicity. In this case, there is a strong peak at the frequency f_k , and a series of partials at f_{2n-1} for odd k and at f_{2n-2} for even k , with $n = 1, 2, \dots$. This means that the odd longitudinal modes are excited by components having the same frequencies as the odd transverse modes. Similarly, the even longitudinal modes are excited at the even transverse modal frequencies.

These frequencies form the inputs of the time-domain impulse responses $\xi_{\delta,k}(t)$, which can be considered as second-order resonators. As we have noted earlier, the output of a resonator has two types of components: one component is the free response, which is a decaying sinusoid at the frequency $f_{\xi,k}$. The other component is the forced response consisting of the frequency series f_{2n-1} or f_{2n-2} with $n = 1, 2, \dots$. The amplitudes of these spectral lines are amplified around the peak of the resonator $f_{\xi,k}$. As the responses of all the longitudinal modes are summed together, the output becomes similar to having formants on a rich harmonic spectrum. The forced longitudinal components are indistinguishable from the transverse ones since they are exactly at the same frequencies.

Inharmonic Transverse Vibration

For stiff strings, as derived in Sec. 2.2.4, the transverse partial frequencies do not form a perfect harmonic series but are described by the equation

$$f_n = f_0 n \sqrt{1 + Bn^2}, \quad (5.42)$$

where B is the inharmonicity coefficient and $f_0 \approx f_1$ is the fundamental frequency of the string.

In this case the terms $f_n + f_{k-n}$ and $f_n - f_{k+n}$ do not have the frequency of f_k but form a bunch of peaks around f_k . The peaks at the frequencies $f_n - f_{k-n}$ lie somewhat higher compared to $f_{|2n-k|}$ and the frequencies $f_n + f_{k+n}$ are lower than f_{2n+k} . This means that these peaks depart from the transverse partials in a rate determined by the inharmonicity coefficient B and the longitudinal mode number k . However, it is still true that odd longitudinal modes are excited by an odd-like partial series, while even longitudinal modes are excited by an even-like one.

The force exciting the first longitudinal mode $F_{\xi,1}(t)$ is displayed in Fig. 5.2 (a) by a solid line, computed by the discrete-time implementation of the modal model described by Eqs. (5.36) and (5.30) (see Sec. 6.3.5). The simulation example is a G_1 piano string. Note that the excitation force has an odd-like partial series and a lower inharmonicity compared to the spectrum of the transverse bridge force, which is displayed by dots to show the transverse modal frequencies as a reference. The dashed line indicates the Fourier transform of the impulse response of the first longitudinal mode $\xi_{\delta,1}(t)$, amplifying the frequencies around 690 Hz. Figure 5.2 (b) shows the excitation-force spectrum of the second longitudinal mode for the same example. It can be seen that here the excitation spectrum contains even partials only and that the peak of the longitudinal mode (dashed line) is located at a higher frequency (1380 Hz in this case).

The longitudinal motion is the sum of the motion of different modes. This means that spectra similar to Figs. 5.2 (a) and (b) should be superimposed with slightly shifted excitation frequencies and very different longitudinal modal frequencies. The result is similar to formants on a quasi-harmonic spectrum but here the peaks are somewhat smeared as they are made up of many close frequencies. The most important difference from the case of harmonic transverse vibration is that these smeared peaks appear between the transverse ones and therefore they can be easily distinguished.

5.3.4 String Tension

The longitudinal component of string vibration is transmitted to the instrument body by the force acting on the bridge in the longitudinal direction. To the first approximation, this equals the tension at the termination $F_1(t) = -T(L, t)$.

The string tension is computed according to Eq. (5.1). Equation (5.1) is made up of three terms. By defining T_0 as the tension at rest, $T_1(x, t)$ as the tension component proportional to the longitudinal slope, and $T_t(x, t)$ as the tension component proportional to the square of the transverse slope, the total tension can be written as

$$T(x, t) = T_0 + T_1(x, t) + T_t(x, t). \quad (5.43)$$

The tension component coming from the longitudinal motion is

$$T_1(x, t) = ES \frac{\partial \xi}{\partial x} = ES \frac{\pi}{L} \sum_{k=1}^{\infty} \xi_k(t) k \cos\left(\frac{k\pi x}{L}\right), \quad (5.44)$$

which was obtained by substituting Eq. (5.29) into Eq. (5.1).

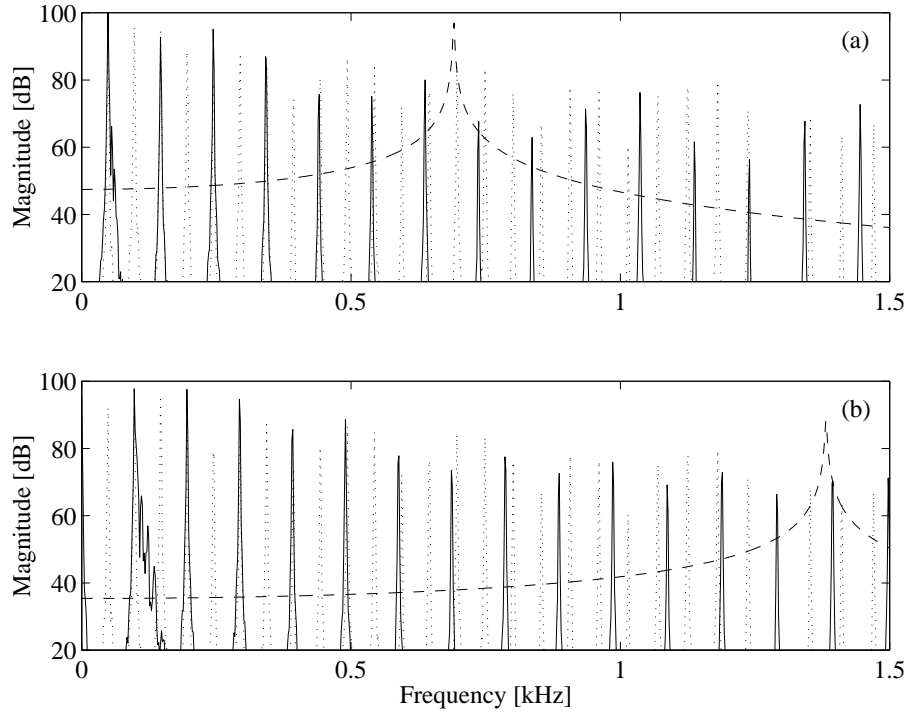


Figure 5.2: The force spectrum exciting the first (a) and the second (b) longitudinal modes ($F_{\xi,1}(t)$ and $F_{\xi,2}(t)$) of a simulated G_1 piano string (solid line). The transverse bridge force (dotted line) is displayed to show the transverse modal frequencies. The dashed line shows the frequency response of the first (a) and the second (b) longitudinal modes. The relative levels of the signals are arbitrary.

Inserting Eq. (5.33) into Eq. (5.1) gives the tension component coming from the transverse vibration:

$$\begin{aligned}
 T_t(x, t) &= ES \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 = ES \frac{1}{2} \left[\frac{\pi}{L} \sum_{n=1}^{\infty} y_n(t) n \cos \left(\frac{n\pi x}{L} \right) \right]^2 = \\
 &= ES \frac{\pi^2}{4L^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} y_m(t) y_n(t) m n \left[\cos \left(\frac{m+n}{L} \pi x \right) + \cos \left(\frac{m-n}{L} \pi x \right) \right]. \quad (5.45)
 \end{aligned}$$

Calculating the string tension $T(x, t)$ by Eqs. (5.44) and (5.45) is quite complicated and do not provide a qualitative insight to the phenomenon. Therefore in the next subsections we decompose the tension into a spatially uniform part $\bar{T}(t)$ and into a space-dependent part $\tilde{T}(x, t)$. This is advantageous because the effects of the spatially uniform tension variation are well described in the literature (see Sec. 5.2).

Decomposing the String Tension

It can be seen from Fig. 5.2 that the longitudinal modes have a constant gain under their resonance frequency. Moreover, we know from the literature (see Sec. 5.2) that if all the

longitudinal modes are excited under resonance, the tension will be spatially uniform along the string. Therefore, it seems reasonable to decompose the longitudinal modal response $\xi_{\delta,k}(t)$ into a constant part $\bar{\xi}_{\delta,k}(t)$ representing this constant gain and into a dynamic part $\tilde{\xi}_{\delta,k}(t)$ corresponding to the dynamics of longitudinal modes. We will see that the constant part $\bar{\xi}_{\delta,k}(t)$ of the longitudinal response will be responsible for generating the spatially uniform part of the tension $\bar{T}(t)$, while the space-dependent part of the tension $\tilde{T}(x,t)$ is coming from the dynamic response $\tilde{\xi}_{\delta,k}(t)$.

Accordingly, the impulse response $\xi_{\delta,k}(t)$ is written as

$$\xi_{\delta,k}(t) = \bar{\xi}_{\delta,k}(t) + \tilde{\xi}_{\delta,k}(t), \quad (5.46)$$

where $\bar{\xi}_{\delta,k}(t)$ can be computed from Eq. (5.40) by $s \rightarrow 0$ for $\omega \ll \omega_{\xi,k}$:

$$\bar{\xi}_{\delta,k}(t) = \frac{2}{L\mu} \frac{1}{\omega_{\xi,k}^2 + \sigma_{\xi,k}^2} \delta(t) = \frac{2L}{ESk^2\pi^2} \delta(t). \quad (5.47)$$

From Eqs. (5.40) and (5.47) the Laplace transform of the “dynamic response” $\tilde{\xi}_{\delta,k}(t)$ (the correction term containing the system dynamics) is

$$\begin{aligned} \mathcal{L}\{\tilde{\xi}_{\delta,k}(t)\} &= \mathcal{L}\{\xi_{\delta,k}(t) - \bar{\xi}_{\delta,k}(t)\} = -\frac{2}{L\mu} \frac{s^2 + 2\sigma_{\xi,k} s}{(\sigma_{\xi,k}^2 + \omega_{\xi,k}^2)(s^2 + 2\sigma_{\xi,k} s + \sigma_{\xi,k}^2 + \omega_{\xi,k}^2)} = \\ &= -\frac{2L}{ESk^2\pi^2} \frac{s^2 + 2\sigma_{\xi,k} s}{s^2 + 2\sigma_{\xi,k} s + \sigma_{\xi,k}^2 + \omega_{\xi,k}^2}, \end{aligned} \quad (5.48)$$

which corresponds to a second-order highpass filter with a resonance at $\omega_{\xi,k}$. Example frequency responses of $\tilde{\xi}_{\delta,k}(t)$ are displayed in Fig. 5.3 by solid line, where the parameters of the longitudinal modes are the same as in Fig. 5.2, but now the frequency axis is in a logarithmic scale. The dashed line shows the constant response corresponding to $\bar{\xi}_{\delta,k}(t)$, while their sum (which is the complete response $\xi_{\delta,k}(t)$) is displayed by a dotted line. Note that the phase of the dynamic response is the opposite of that of the static response at high frequencies, therefore they cancel out.

Similarly to the decomposition of the impulse response, we may also decompose the modal amplitude $\xi_k(t)$ into a part coming from the static response, referred as $\bar{\xi}_k(t)$, and into a part coming from the dynamic response, referred as $\tilde{\xi}_k(t)$. We may also do the same with the longitudinal tension component by writing $T_1(x,t) = \bar{T}_1(x,t) + \tilde{T}_1(x,t)$.

Tension Coming From the Constant Response

From Eqs. (5.47) and (5.30a) the static part of the instantaneous amplitude $\bar{\xi}_k(t)$ of longitudinal mode k is obtained as

$$\bar{\xi}_k(t) = \frac{2L}{ESk^2\pi^2} F_{\xi,k}(t). \quad (5.49)$$

From Eqs. (5.44) and (5.49), the tension component that comes from the longitudinal motion takes the form

$$\bar{T}_1(x,t) = -\frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1}{k} F_{\xi,k}(t) \cos\left(\frac{k\pi x}{L}\right). \quad (5.50)$$

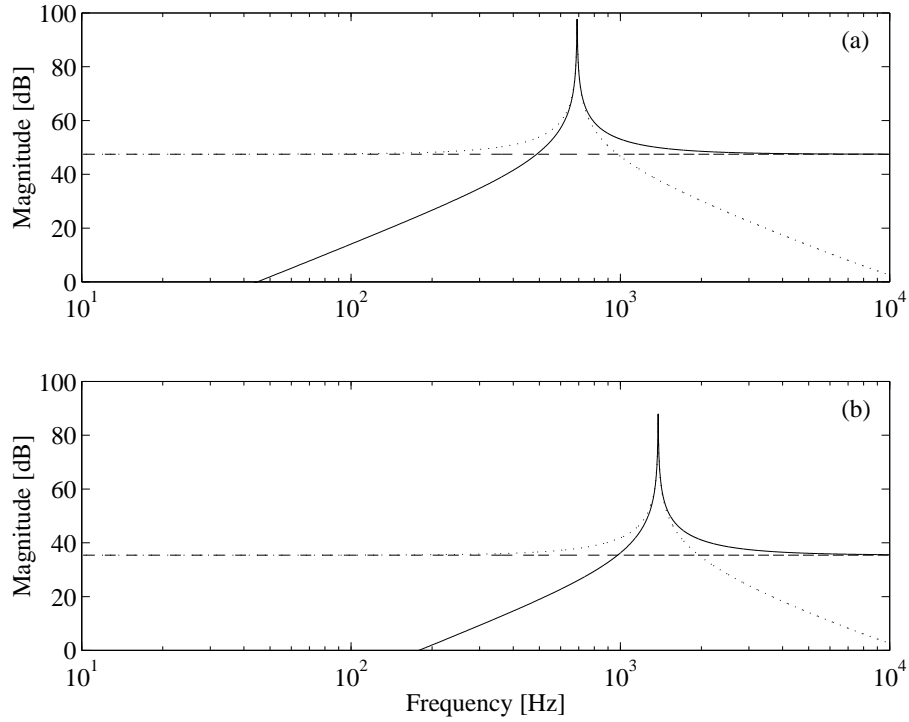


Figure 5.3: Frequency responses of the first (a) and the second second (b) longitudinal modes of a G_1 piano tone. The solid line shows the Fourier transform of the dynamic response $\tilde{\xi}_{\delta,k}(t)$, while the constant response $\bar{\xi}_{\delta,k}(t)$ is displayed by a dashed line. The complete response $\xi_{\delta,k}(t) = \bar{\xi}_{\delta,k}(t) + \tilde{\xi}_{\delta,k}(t)$ is displayed by a dotted line.

Calculating the excitation force $F_{\xi,k}(t) = F_{\xi,k}(t)^+ + F_{\xi,k}(t)^-$ with the help of Eq. (5.36) and eliminating k by substituting $m + n = k$ and $|m - n| = k$ gives

$$\begin{aligned} \bar{T}_1(x, t) = & -ES \frac{\pi^2}{4L^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} y_m(t) y_n(t) m n \cos\left(\frac{m+n}{L} \pi x\right) - \\ & - ES \frac{\pi^2}{4L^2} \sum_{m=1}^{\infty} \sum_{\substack{n=1 \\ n \neq m}}^{\infty} y_m(t) y_n(t) m n \cos\left(\frac{m-n}{L} \pi x\right), \end{aligned} \quad (5.51)$$

where $n \neq m$ in the second term comes from the fact that the longitudinal mode number $k = |m - n|$ cannot be zero. Note that there is no such constraint for the first term as $k = m + n$ in that case.

Comparing Eq. (5.51) with (5.45) shows that they contain the same terms but with opposite sign. The only difference is the $n \neq m$ under the last sum of Eq. (5.51). Indeed, if Eqs. (5.45) and (5.51) are substituted into Eq. (5.43), all the terms cancel out, except some with $m = n$ giving

$$\bar{T}(t) = T_0 + \bar{T}_1(x, t) + T_t(x, t) = T_0 + ES \frac{\pi^2}{4L^2} \sum_{n=1}^{\infty} y_n^2(t) n^2, \quad (5.52)$$

where the tension does not depend on x , i.e., it is uniform along the string. This is the same as Eq. (5.21) obtained by computing the tension from the elongation of the string by Eq. (5.15). The temporal variation of this spatially constant tension $\bar{T}(t)$ in the case of sinusoidally decaying transverse modes have been already discussed in Sec. 5.2.2.

Tension Coming From the Dynamic Response

From Eq. (5.44), by writing $\xi_k(t) = \bar{\xi}_k(t) + \tilde{\xi}_k(t)$, the tension coming from the longitudinal motion is obtained as

$$T_1(x, t) = ES \frac{\pi}{L} \sum_{k=1}^{\infty} [\bar{\xi}_k(t) + \tilde{\xi}_k(t)] k \cos\left(\frac{k\pi x}{L}\right) = \bar{T}_1(x, t) + \tilde{T}_1(x, t), \quad (5.53)$$

where $\bar{T}_1(x, t)$ is the tension coming from the constant response and can be computed by Eq. (5.51), and $\tilde{T}_1(x, t)$ is the tension component originating from the dynamics of longitudinal modes, and it is obtained as

$$\tilde{T}_1(x, t) = ES \frac{\pi}{L} \sum_{k=1}^{\infty} [\tilde{\xi}_{\delta,k}(t) * F_{\xi,k}(t)] k \cos\left(\frac{k\pi x}{L}\right), \quad (5.54)$$

where the substitution $\tilde{\xi}_k(t) = \tilde{\xi}_{\delta,k}(t) * F_{\xi,k}(t)$ means that the dynamic part of the motion of longitudinal mode k is computed by convolving the dynamic impulse response $\tilde{\xi}_{\delta,k}(t)$ with the excitation force acting on mode k .

The total tension is given by

$$T(x, t) = T_0 + T_t(x, t) + \bar{T}_1(x, t) + \tilde{T}_1(x, t) = \bar{T}(t) + \tilde{T}_1(x, t), \quad (5.55)$$

where $\bar{T}(t)$ is the spatially uniform tension (both transverse and longitudinal) computed by the static response of longitudinal modes, as given in Eq. (5.52). Accordingly, the total tension is made up from a spatially uniform part $\bar{T}(t)$ which comes from the transverse motion and the constant response of longitudinal modes, and a spatially non-uniform part $\tilde{T}_1(x, t)$ coming from the dynamics of longitudinal modes.

Let us turn our attention to this dynamic part: according to Eq. (5.48), the dynamic impulse response $\tilde{\xi}_{\delta,k}(t)$ corresponds to a second-order highpass filter. This filter has a free response that is an exponentially decaying sinusoid at the longitudinal modal frequency $\omega_{\xi,k}$ and a forced response, whose amplitude and phase can be computed from Eq. (5.48) by $s = j\omega_p$. The excitation terms of $F_{\xi,k}(t)$ have been already discussed in Sec. 5.3.3. Note that the excitation signal of $\xi_{\delta,k}(t)$, $\bar{\xi}_{\delta,k}(t)$, and $\tilde{\xi}_{\delta,k}(t)$ are all the same (that is, $F_{\xi,k}(t)$).

From Eq. (5.48) and from Fig. 5.3 we can conclude that for $\omega_p \ll \omega_{\xi,k}$ the forced response is zero. For $\omega_p \gg \omega_{\xi,k}$, the forced response is small, as its contribution to the tension is comparable to that coming from the constant response. However, for $\omega_p \approx \omega_{\xi,k}$, the longitudinal mode is excited around resonance, leading to a strong component of the frequency ω_p in the tension, having a spatial distribution of the form $\cos(k\pi x/L)$.

The free response of longitudinal mode k is an exponentially decaying sinusoidal function with the frequency $\omega_{\xi,k}$ and decay rate $\sigma_{\xi,k}$, which adds a further component to the

tension, again with the spatial distribution of $\cos(k\pi x/L)$. The amplitude of the free response is in the order of that of the forced response. This means that when forced response is negligible, so is the free response.

Some Notes on the Uniform Tension Approximation

We have already reviewed the conditions given by the literature when the tension can be considered uniform along the string in Sec. 5.2. However, the decomposition of the string tension into a spatially uniform part $\bar{T}(t)$ and into a space-dependent part $\tilde{T}_1(x, t)$ helps us to understand this phenomenon better. Actually, the uniform tension approximation of Sec. 5.2 can be applied if the spatially nonuniform part of the tension $\tilde{T}_1(x, t)$ is negligible compared to the spatially uniform part $\bar{T}(t)$. This happens when the dynamic responses of all the longitudinal modes (which generate the space-dependent part $\tilde{T}_1(x, t)$) are significantly lower than the static responses (which determine the spatially uniform part $\bar{T}(t)$). That is, the longitudinal modes are excited at only those frequencies, where the solid lines are significantly below the dashed ones in Fig. 5.3. This happens if all the longitudinal modes are excited by frequencies that are considerably smaller than the corresponding longitudinal modal frequency $f_{\xi, k}$. As the dynamic responses (solid lines in Fig. 5.3) correspond to second-order high-pass filters, it is reasonable to require that all the longitudinal modes are excited below the half of their modal frequencies.

Thus, the validity of the uniform tension approximation should be evaluated by comparing $f_{\xi, k}$ with the excitation frequencies calculated by Eq. (5.41) for each k . Having small order transverse vibrations in comparison to the ratio of longitudinal and transverse propagation speeds [see, e.g., Anand 1969] is a sufficient, but not a necessary, condition for the applicability of the uniform tension approximation.

As a special case, if the transverse vibration contains only one mode, the uniform tension approximation can always be applied. This is because the transverse mode n excites longitudinal mode $k = 2n$ at the frequency $2f_n \approx 2nf_0$, and the modal frequency $f_{\xi, 2n} \approx 2nf_{\xi, 0}$ of the $2n^{\text{th}}$ longitudinal mode is much larger than $2f_n$ (as $f_{\xi, 0}/f_0 \gg 1$ holds for metal strings). Note that this is true for all the transverse modes, and not only for the first few.

Another special case occurs when the transverse partials are present up to a mode number N . Here the assumption of the uniform tension can be applied only if the excitation force does not contain significant components around and above $f_{\xi, 0}/2$. The excitation forces $F_{\xi, k}(t)$ have approximately double the bandwidth ($\approx 2Nf_0$) compared to the bandwidth of the transverse vibration ($\approx Nf_0$). This gives the constraint $2Nf_0 < f_{\xi, 0}/2$, which is equivalent to Eq. (5.12).

5.3.5 Longitudinal Bridge Force

The longitudinal bridge force can be approximated as the tension at the termination ($x = L$):

$$F_1(t) = -T(L, t) = -[\bar{T}(t) + \tilde{T}_1(L, t)] = - \left[T_0 + ES \frac{\pi^2}{4L^2} \sum_{n=1}^{\infty} y_n^2(t) n^2 + ES \frac{\pi}{L} \sum_{k=1}^{\infty} \{ \tilde{\xi}_{\delta, k}(t) * F_{\xi, k}(t) \} k \cos(k\pi) \right]. \quad (5.56)$$

In the case of exponentially decaying sinusoidal transverse modes, the first part $\bar{T}(t)$ is made up of a slowly decaying average tension and of a tension variation having double frequency terms (see Sec. 5.2.2). The slow tension variation of Eq. (5.24) yields a change in the strain of the instrument body, but it is not radiated to the air, i.e., it has no importance in this case (if we were considering longitudinal to transverse coupling, it would have). On the other hand, the sinusoidal terms of Eq. (5.23) are effectively radiated by the instrument body. This means that terms having double the frequency of transverse modes will appear in the sound generated by the instrument.

The part covering the dynamics of longitudinal modes $\tilde{T}_1(L, t)$ can be computed by taking the Laplace transform of the excitation Eq. (5.39), multiplying by Eq. (5.48), and performing an inverse Laplace transform. However, for qualitative understanding it is enough to notice that $\tilde{T}_1(L, t)$ contains the free and forced response of those modes which are excited around or above their resonance frequency. The dominant components in $\tilde{T}_1(L, t)$ will come from those excitation terms which are near resonance.

Note that the components originating from the excitation of longitudinal modes have much higher amplitude than the double frequency terms. This is because these “resonance terms” are originating from the dynamic response of longitudinal modes (solid line in Fig. 5.3), where the gain is 30–40 dB larger compared to the static response (dashed line in Fig. 5.3), which generates the double frequency terms.

5.3.6 Possible Extensions and Limitations

Nonrigid String Terminations

The termination of musical instrument is not perfectly rigid, contrary to the assumptions made here. As the impedance of the bridge is usually much larger compared to the impedance of the string, its main effect is a change in the transverse partial frequencies f_n and decay times τ_n , which can be easily incorporated in Eq. (5.37). The modal shapes also change slightly, which can be taken into account by substituting L with $L + \delta L_n$ in Eqs. (5.33) and (5.37), where δL_n is computed by Eq. (5.26), meaning that none of the longitudinal modal shapes are completely orthogonal to the modal shapes of the excitation-force density. However, it is still true that the dominant force components are those computed by Eq. (5.36).

It is proven in Appendix A.2 that the equations derived by assuming infinitely rigid string terminations are good approximations to the reality for longitudinal modal numbers

in the order of 1-10, which are actually the most important ones from the musical acoustics point of view (these are the ones that can be heard). This has also been confirmed by finite-difference simulations showing only a small change in the output when a more realistic termination model is applied.

The termination of musical instrument strings can also contribute to the energy transfer between the transverse and longitudinal motion, if the string passes the bridge at an angle [Legge and Fletcher 1984]. As such a coupling is linear, it does not introduce new terms by itself. The transverse frequencies can appear in the longitudinal motion, and conversely, the longitudinal frequencies may turn up in the transverse vibration. However, those transverse and longitudinal components that have the same frequency cannot be distinguished in the sound pressure. The coupling through the bridge in combination with the transverse to longitudinal coupling along the string could produce new terms, but they are of fourth order in the amplitude of the transverse vibration.

Extension to Two Transverse Planes

Real strings vibrate in two transverse polarizations. The modal frequencies for these polarizations can be different for the same modes, mostly because of the direction-dependent termination impedance. This produces beating and two-stage decay in piano sound [Weinreich 1977]. The three-dimensional version of Eq. (5.27) is

$$\mu \frac{\partial^2 \xi}{\partial t^2} = ES \frac{\partial^2 \xi}{\partial x^2} - 2R_\xi(\omega) \mu \frac{\partial \xi}{\partial t} + \frac{1}{2} ES \left[\frac{\partial \left(\frac{\partial y}{\partial x} \right)^2}{\partial x} + \frac{\partial \left(\frac{\partial z}{\partial x} \right)^2}{\partial x} \right], \quad (5.57)$$

where z is the string displacement in the direction perpendicular to the already considered transverse y and longitudinal x directions.

It follows from Eq. (5.57) that the excitation-force density $d_\xi(x, t)$ is the superposition of the excitation-force densities computed for the two transverse planes separately. After performing similar derivations as for Eq. (5.35), the excitation force density of the longitudinal vibration becomes

$$d_\xi(x, t) = -ES \frac{\pi^3}{4L^3} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [y_m(t)y_n(t) + z_m(t)z_n(t)] m n \times \\ \times \left[(m+n) \sin \left(\frac{m+n}{L} \pi x \right) + (m-n) \sin \left(\frac{m-n}{L} \pi x \right) \right], \quad (5.58)$$

where z_n is the instantaneous amplitude of the mode n in the z direction:

$$z(x, t) = \sum_{n=1}^N z_n(t) \sin \left(\frac{k\pi x}{L} \right). \quad (5.59)$$

By replacing $y_m(t)y_n(t)$ with $[y_m(t)y_n(t) + z_m(t)z_n(t)]$ in Eq. (5.36), the excitation force of longitudinal mode k is obtained. The component originating from the transverse modes

that satisfy $m + n = k$ is

$$F_{\xi,k}(t)^+ = -ES \frac{\pi^3}{8L^2} \sum_{n=1}^{k-1} [y_{k-n}(t)y_n(t) + z_{k-n}(t)z_n(t)] k(k-n)n. \quad (5.60a)$$

The component coming from $|m - n| = k$ becomes

$$F_{\xi,k}(t)^- = -2ES \frac{\pi^3}{8L^2} \sum_{n=1}^{\infty} [y_{k+n}(t)y_n(t) + z_{k+n}(t)z_n(t)] k(k+n)n. \quad (5.60b)$$

The total excitation force of mode k is the sum of the two components, i.e., $F_{\xi,k}(t) = F_{\xi,k}(t)^+ + F_{\xi,k}(t)^-$. As the force exciting the longitudinal modes is known, the longitudinal motion can be readily computed by Eq. (5.30). The equations covering the string tension could also be obtained by slightly modifying the equations of Sec. 5.3.4. Actually, terms of the form $y_m(t)y_n(t)$ have to be replaced by $[y_m(t)y_n(t) + z_m(t)z_n(t)]$ in all the places.

It is useful to take a short look at the excitation frequencies. If two modes vibrate in two planes perpendicular to each other, their sum and difference frequencies do not appear in the excitation force. In reality the vibrating planes of the different modes are not perfectly perpendicular to each other because of the direction- and frequency-dependent terminating impedance, meaning that mode m vibrating in one plane will mix with mode n vibrating in a different plane.

The modal frequencies $f_{n,1}$ and $f_{n,2}$ of the two transverse polarizations are slightly different for the same mode number n . Thus, the excitation components coming from the transverse modes m and n consist of four different frequencies. For example, the sum-frequency components have the frequencies $f_{m,1} + f_{n,1}$, $f_{m,1} + f_{n,2}$, $f_{m,2} + f_{n,1}$, and $f_{m,2} + f_{n,2}$. The difference-frequency components can be expressed similarly.

The Effect of Neglecting the Longitudinal to Transverse Coupling

The assumption of neglecting the longitudinal to transverse coupling is valid until the longitudinal vibration is small compared to the transverse one. However, if one of the excitation frequencies of the longitudinal mode k (see Eq. (5.41)) is very close to the resonant frequency $f_{\xi,k}$ of that mode, the longitudinal motion can have extremely large amplitude. This does not happen in reality since the longitudinal motion diminishes the amplitude of those transverse modes from which it originates (the total energy of transverse and longitudinal vibrations cannot increase). This stabilizing effect is not included in this model. Some of its aspects will be covered in Sec. 5.4.2.

The longitudinal to transverse coupling would also introduce some terms of third order in the amplitude of the transverse vibration, but their contribution is less significant in this regime of vibration, as already discussed in Sec. 5.1.2.

Accordingly, the frequencies predicted by the model of Sec. 5.3 should be in quantitative agreement with the dominant peaks found in real string spectrum. The amplitude behavior is described properly for those peaks that do not coincide with the resonant frequency of the excited longitudinal mode. (This holds for most of the peaks.)

Note that the coincidences of longitudinal modal and excitation frequencies have a small practical significance from the sound synthesis viewpoint since they produce an unpleasant ringing sound even when computed with full bidirectional coupling, e.g., by the model of Sec. 6.4.2. Therefore, when synthesizing string sounds with this paradigm, the longitudinal modal frequencies have to be tuned in a way that they are not too close to their excitation frequencies (see Sec. 6.3.4 for details). This fact implies that these annoying coincidences should also be avoided in real musical instruments (mostly in pianos) by careful string- and scale design.

5.3.7 Connections to Measurements

In this section the theoretical results of Sec. 5.3 are related to the measurements of other authors, namely Nakamura and Naganuma [1993]; Conklin [1999]; Giordano and Korty [1996]. On the one hand, this confirms the theoretical model developed here. On the other hand, it helps to understand the theoretical reasons underlying the findings of these experimental studies.

Parentage of Phantom Partial

From the theoretical point of view, phantom partials are coming from the forced motion of longitudinal vibrations. An interesting property of odd phantom partials discovered by Conklin [1999] is that they originate from adjacent parents, i.e., they can be found at frequencies $f_m + f_n$ where $|m - n| = 1$.

By looking at $F_{\xi,k}(t)^-$ in Eq. (5.41) it turns out that the frequencies $f_n + f_m = f_n + f_{k+n}$ are quite close to each other for $m + n = p$ (they would actually coincide in the case of a perfectly harmonic transverse vibration, and would have the frequency f_{2n+k}). The question is which $f_m + f_n$ combination has the largest amplitude in the resulting sound. It follows from Eq. (5.41) that the different $f_m + f_n$ components belonging to the same smeared peak (i.e., $m + n = p$) excite different longitudinal modes. Namely, the frequency $f_m + f_n$ excites the longitudinal mode having the mode number $k = |m - n|$. Accordingly, that $f_m + f_n$ component results in the largest longitudinal motion which excites the longitudinal mode having a modal frequency $f_{\xi,k}$ close to the frequency $f_m + f_n$. In other words, if the frequency of a phantom partial group is close to the frequency $f_{\xi,k}$ of the k th longitudinal mode, it mainly originates from parents having mode number difference of k .

The lower odd phantom partials, which were measured by Conklin [1999], most probably have frequencies to which the first longitudinal mode is the nearest. In this case the $f_m + f_n$ terms satisfying $|m - n| = 1$ dominate, which actually originate from adjacent parents f_n and f_{n+1} .

Similar considerations apply for even phantoms, that is, they are generated by parents having mode number difference of 2, 4, 6, etc., depending on the frequency of the phantom partial. However, there is an important difference that double frequency terms $2f_n$ also occur in the spectrum. These $2f_n$ components would arise even when the bandwidth of transverse components was significantly lower than the frequency of the first longitudinal mode, i.e., when the tension was uniform along the string (see Sec. 5.3.4). Actually, these

are the only components that can be explained by the uniform tension approximation of Sec. 5.2, while for the sum- and difference frequency components the inertial effects of longitudinal modes have to be included in the model. The double frequency terms have a lower amplitude compared to the phantom partials originating from adjacent parents, as the latter are generated by the resonance of longitudinal modes, while the $2f_n$ components correspond to the excitation of a specific longitudinal mode under resonance (see Secs. 5.3.4 and 5.3.5).

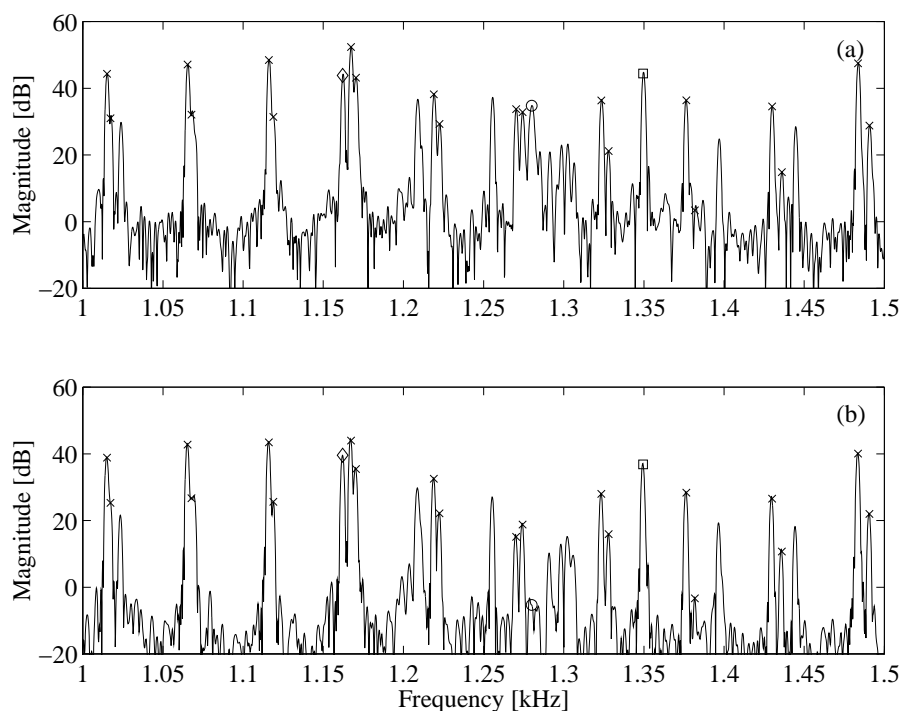


Figure 5.4: Spectrum of the first (a) and the second second (b) of a F_1 piano tone. Transverse partials are marked by crosses and the second longitudinal mode is marked by a circle. Two prominent phantom partial groups are indicated by a square and a diamond (the latter is magnified in Fig. 5.5).

The spectrum of a recorded F_1 piano tone (having only one string) is displayed in Fig. 5.4. Transverse partials are indicated by crosses. Note that two crosses belong to one partial, as there are two transverse polarizations with slightly different modal frequencies. The free response of the second longitudinal mode is marked by a circle. The remaining peaks are the forced response of the longitudinal motion, i.e., they are the phantom partials. Figure 5.4 (a) shows the first second of the tone and Fig. 5.4 (b) displays the second, giving an insight to the evolution of the spectrum. Note that the free response of the longitudinal mode (circle) disappears fast in the noise (the decay time is ca. 0.15 s), while the phantom partials remain significant and their decay rate is comparable to that of the transverse partials (1–2 s). In general, it has been found that the highest nontransverse peaks in the long-term spectrum are phantom partials amplified by a longitudinal mode

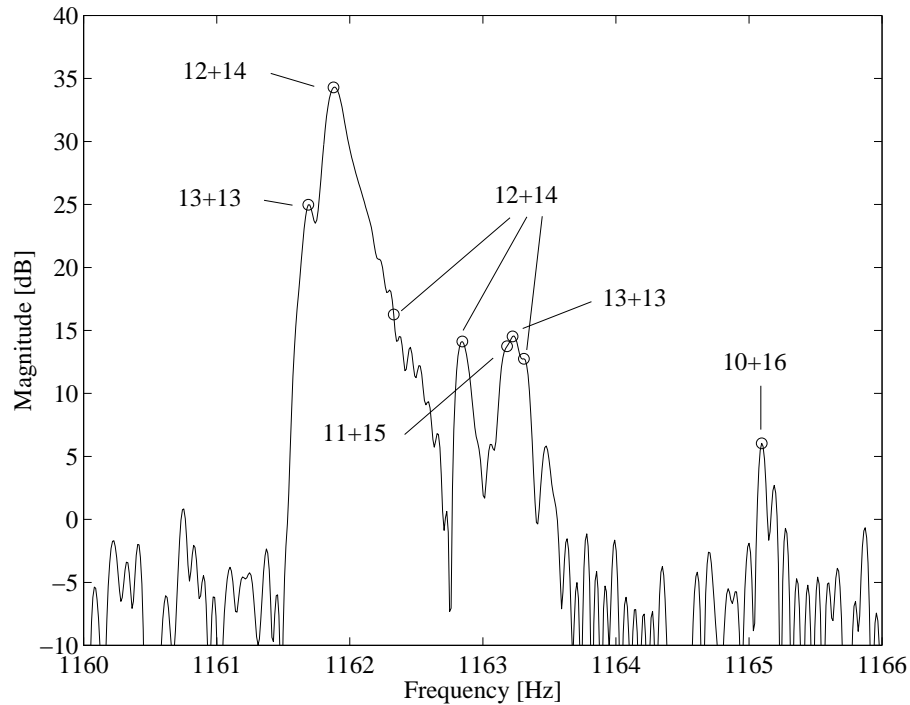


Figure 5.5: The spectrum of an even phantom partial group in the F_1 piano tone of Fig. 5.4. Sum frequencies of transverse modes $f_m + f_n$ are marked by circles, and the mode numbers of the parent partials are labeled in the form of $m + n$. The phantom group is displayed by a diamond in Fig. 5.4.

(one prominent example is marked by a square). This suggests that the forced response of the longitudinal motion may have a larger perceptual significance than the free response itself. Most probably the pitch of the longitudinal component is determined by these amplified phantom partials (like the one marked by a square in Fig. 5.4) and not by the fast decaying free response. The interested reader may listen to the sound examples demonstrating the relative significance of these components [Sound examples n.d.].

The “single” phantom partial marked by a diamond in Fig. 5.4 becomes a group of partials when plotted at a higher frequency resolution in Fig. 5.5. In this case the data length is 16 seconds (705 600 samples at $f_s = 44.1$ kHz), which was zero padded to 2^{22} samples after applying a Hanning window. The most prominent peaks of the phantom group are marked by circles. The label “ $m + n$ ” beside a circle indicates that the circle is located at the sum frequency of the transverse modes m and n (i.e., at $f_m + f_n$). The frequencies of the transverse modes were determined by finding peaks in the spectrum. Note that the same $m + n$ combinations can be found at several peaks: the reason is that the two different frequencies $f_{m,1}$ and $f_{m,2}$ of the two transverse polarizations of mode m mix with the two different frequencies $f_{n,1}$ and $f_{n,2}$ of mode n , as predicted in Sec. 5.3.6.

It can be seen in Fig. 5.5 that the highest peak comes from the 12th and 14th transverse modes and not from the 13th mode itself, although the amplitude of the latter is only 10 dB smaller. Other even phantoms show the same phenomenon: they principally originate

from parents having mode-number difference of 2, 4, etc., and not from a single mode by frequency doubling. This contradicts the findings of Conklin [1999] but confirms the analysis of Sec. 5.3.5.

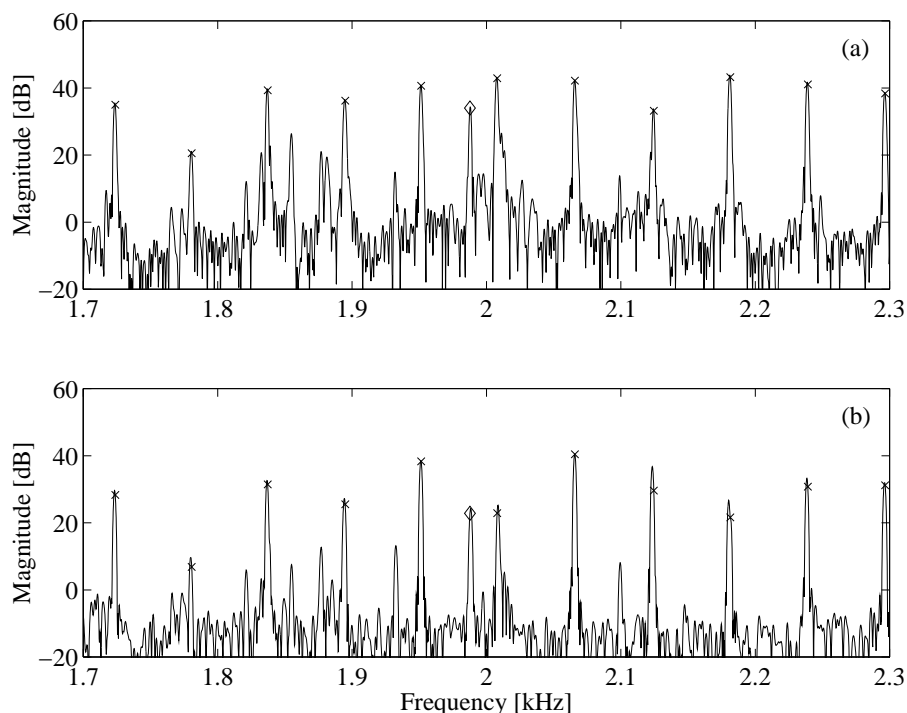


Figure 5.6: Spectrum of the first (a) and the second second (b) of an A_1 harpsichord tone. Transverse partials are marked by crosses. All the other peaks correspond to phantom partials (the one with the largest amplitude is marked by a diamond).

The spectrum of the first and second second of an A_1 harpsichord tone¹ is displayed in Fig. 5.6, showing that phantom partials can be found in other instruments, too. The displayed part of the spectrum has the largest nonlinear components, meaning that in other parts of the spectrum the nonlinearity is less prominent. Note that even in this region the nonlinear components have lower amplitude compared to those in the piano, displayed in Fig. 5.4. Other example spectra of instruments with phantom partials can be found in [Conklin 1999].

Inharmonicity of Phantom Partial

Nakamura and Naganuma [1993] found that the inharmonicity of phantom partials (which they call “lower series”) is one-fourth of that of normal transverse partials. This can be explained by knowing that phantom partials are mainly produced by parents with mode numbers close to each other. This means that even phantoms have an approximate frequency of $f_p = 2f_n$, where $p = 2n$ is the “mode number” of the phantom partial. (See

¹Sound sample provided by the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology.

Fig. 5.5 as an example, where $f_{12} + f_{14} \approx 2f_{13}$.) Writing $f_p = 2f_n$ according to Eq. (5.42) and expressing the frequencies by the phantom mode number $p = 2n$ gives

$$f_p \approx 2f_n = 2f_0n\sqrt{1 + Bn^2} = f_0p\sqrt{1 + \frac{1}{4}Bp^2}. \quad (5.61)$$

For even phantoms, the expression is quite accurate. For odd phantoms, $n = p/2$ is not an integer number. However, as the inharmonicity curve is a smooth function, the frequencies of odd phantom partials are also close to the ones predicted by Eq. (5.61).

Amplitude of Longitudinal Vibration

Giordano and Korty [1996] found that the amplitude of the longitudinal vibration in recorded piano sound is a nonlinear function of the amplitude of the transverse component. They noted that the nonlinear curve is faster than a simple quadratic function.

Equation (5.35) shows that a peak in the excitation spectrum of a longitudinal mode is a quadratic function of the overall amplitude of the generating transverse modes m and n . However, the amplitude of longitudinal motion is mostly determined by parents having sum frequencies $f_m + f_n$ around the longitudinal modal frequencies $f_{\xi,k}$. The amplitude of these parents (with mode numbers around 10–20 in practice) is a nonlinear function of the overall amplitude of the transverse vibration. This is because of the nonlinear nature of the hammer–string interaction [see, e.g., Fletcher and Rossing 1998, p. 367]. The presence of this second kind of nonlinearity explains why Giordano and Korty [1996] could not measure a second-order relationship.

5.4 Bidirectional Coupling of the Transverse and Longitudinal Polarizations

In this section we will briefly overview the bidirectional coupling of longitudinal and transverse polarizations in the case of spatially nonuniform tension. The papers that consider the bidirectional coupling are those already mentioned in Sec. 5.3.1. It is a general feature of these studies [Leissa and Saad 1994; Leamy and Gottlieb 2000; Kurmyshev 2003] that they investigate the first few modes of vibration, thus, they cannot be used directly. We will rather concentrate on the basic equations, which can be formulated similarly to Sec. 5.3, and give a qualitative explanation to the stabilizing effect, namely, that the amplitude of a longitudinal mode cannot grow to infinity, even if it is excited at its resonance frequency.

5.4.1 Equations of Motion

The equations presented in Sec. 5.3 gave analytical solutions as there the nonlinear coupling between two linear subsystems was unidirectional. Here the coupling is bidirectional, which cannot be treated analytically for arbitrary number of modes. Therefore, we will examine the nonlinear transverse force assuming a specific transverse and longitudinal vibration of the string. In reality, this nonlinear transverse force acts back and modifies the transverse

motion and the longitudinal motion through the transverse one. Therefore, only qualitative results are expected.

The equation of motion for the longitudinal polarization is described by Eq. (5.27). From Eqs. (2.25) and (5.3), the wave equation for the transverse motion (y polarization) is written as

$$\mu \frac{\partial^2 y}{\partial t^2} = T_0 \frac{\partial^2 y}{\partial x^2} - ES\kappa^2 \frac{\partial^4 y}{\partial x^4} - 2R(\omega)\mu \frac{\partial y}{\partial t} + ES \frac{\partial \left\{ \frac{\partial y}{\partial x} \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 \right] \right\}}{\partial x}, \quad (5.62)$$

where the right-most term is the nonlinear excitation-force density

$$d_y(x, t) = ES \frac{\partial \left\{ \frac{\partial y}{\partial x} \left[\frac{\partial \xi}{\partial x} + \frac{1}{2} \left(\frac{\partial y}{\partial x} \right)^2 \right] \right\}}{\partial x} = \frac{\partial \left\{ [T(x, t) - T_0] \frac{\partial y}{\partial x} \right\}}{\partial x} = \bar{d}_y(x, t) + \tilde{d}_y(x, t). \quad (5.63)$$

Similarly to decomposing the tension $T(x, t) = \bar{T}(t) + \tilde{T}_1(x, t)$ (see Sec. 5.3.4), the excitation-force density $d_y(x, t)$ is decomposed into two parts. The part that comes from the spatially uniform part of the temporal modulation of the tension is

$$\bar{d}_y(x, t) = \frac{\partial \left\{ [\bar{T}(t) - T_0] \frac{\partial y}{\partial x} \right\}}{\partial x} = [\bar{T}(t) - T_0] \frac{\partial^2 y}{\partial x^2}. \quad (5.64)$$

Note that T_0 is subtracted from $\bar{T}(t)$ as it is already included in the term $T_0(\partial^2 y/\partial x^2)$ of Eq. (5.62). The part $\bar{d}_y(x, t)$ can be considered as the excitation force coming from the temporal variation of the spatially uniform part of the tension.

The part that comes from the space-dependent component of the tension is

$$\tilde{d}_y(x, t) = \frac{\partial \left(\tilde{T}_1(x, t) \frac{\partial y}{\partial x} \right)}{\partial x}. \quad (5.65)$$

As the total nonlinear excitation force is the superposition of $\bar{d}_y(x, t)$ and $\tilde{d}_y(x, t)$, we will take a look at the effect of these two components separately.

Excitation Force Coming From the Temporal Variation of the Spatially Uniform Tension

By using Eq. (5.52) and writing the transverse vibration in its modal form, the first excitation-force density component $\bar{d}_y(x, t)$ becomes

$$\bar{d}_y(x, t) = \{\bar{T}(t) - T_0\} \frac{\partial^2 y}{\partial x^2} = - \left(ES \frac{\pi^2}{4L^2} \sum_{n=1}^{\infty} y_n^2(t) n^2 \right) \left(\frac{\pi^2}{L^2} \sum_{n=1}^{\infty} y_n(t) n^2 \sin \left(\frac{n\pi x}{L} \right) \right). \quad (5.66)$$

The excitation-force $\bar{F}_{y,m}(t)$ acting on the transverse mode m is given by the scalar product of the excitation force density and the modal shape

$$\bar{F}_{y,m}(t) = \int_{x=0}^L \bar{d}_y(x, t) \sin \left(\frac{m\pi x}{L} \right) dx = -ES \frac{\pi^4}{8L^3} y_m(t) m \sum_{n=1}^{\infty} y_n^2(t) n^2. \quad (5.67)$$

As this part of excitation is coming from the spatially uniform part of the tension, its effects are discussed in Sec. 5.2 and the references cited therein. Here we only recall that the quasistatic decaying part of tension modulation leads to a pitch glide, while the double-frequency terms lead to the excitation frequencies $2\omega_n \pm \omega_m$ for all n , where effective excitation happens if $m = n$. It is more interesting to turn our attention to the effect of tension variation that comes from the dynamics of longitudinal modes.

Excitation Force Coming From the Space-dependent Part of the Tension

The excitation-force density $\tilde{d}_y(x, t)$ is given as

$$\tilde{d}_y(x, t) = \frac{\partial \left(\tilde{T}_1(x, t) \frac{\partial y}{\partial x} \right)}{\partial x} = \frac{\partial^2 y}{\partial x^2} \tilde{T}_1(x, t) + \frac{\partial y}{\partial x} \frac{\partial \tilde{T}_1(x, t)}{\partial x}, \quad (5.68)$$

which, by using Eq. (5.54) and writing the transverse and longitudinal displacement in the modal form, becomes

$$\begin{aligned} \tilde{d}_y(x, t) &= -ES \frac{\pi^3}{L^3} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} n^2 k y_n(t) \tilde{\xi}_k(t) \sin\left(\frac{n\pi x}{L}\right) \cos\left(\frac{k\pi x}{L}\right) - \\ &\quad - ES \frac{\pi^3}{L^3} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} n k^2 y_n(t) \tilde{\xi}_k(t) \cos\left(\frac{n\pi x}{L}\right) \sin\left(\frac{k\pi x}{L}\right) = \\ &= -ES \frac{\pi^3}{2L^3} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} n k y_n(t) \tilde{\xi}_k(t) \left[(n+k) \sin\left(\frac{n+k}{L}\pi x\right) + (n-k) \sin\left(\frac{n-k}{L}\pi x\right) \right] \end{aligned} \quad (5.69)$$

It is a nice example of the symmetry in nature that Eq. (5.69) has exactly the same form as the excitation force of the longitudinal motion Eq. (5.35), with the substitutions $m \rightarrow n$, $n \rightarrow k$, $k \rightarrow m$, $y_m(t) \rightarrow y_n(t)$, and $y_n(t) \rightarrow \tilde{\xi}_k(t)$. Accordingly, the excitation force $\tilde{F}_{y,m}(t)$ acting on transverse mode m can be directly written from Eq. (5.36), by defining $\tilde{F}_{y,m}(t) = \tilde{F}_{y,m}^+(t) + \tilde{F}_{y,m}^-(t)$. The component originating from $m = n + k$ is

$$\tilde{F}_{y,m}(t)^+ = -ES \frac{\pi^3}{4L^2} \sum_{k=1}^{m-1} y_{m-k}(t) \tilde{\xi}_k(t) m(m-k)k. \quad (5.70a)$$

The component coming from $m = |n - k|$ becomes

$$\tilde{F}_{y,m}(t)^- = -2ES \frac{\pi^3}{4L^2} \sum_{k=1}^{\infty} y_{m+k}(t) \tilde{\xi}_k(t) m(m+k)k. \quad (5.70b)$$

The factor of 2 in Eq. (5.70b) comes from the fact that there are two equal series $n = m + k$ and $k = m + n$, since both satisfy $|n - k| = m$. It follows from Eq. (5.70) that transverse mode m is excited effectively by such longitudinal and transverse mode pairs, for which either the sum $n + k$ or the difference $|n - k|$ of their mode numbers is equal to m .

If we assume that the linear transverse component is built up by exponentially decaying sinusoidal functions, as in Eq. (5.37), and so is the longitudinal motion (but including the

forced response of longitudinal modes, too), it is possible to find the excitation frequencies of transverse mode m similarly to Eq. (5.41), given as

$$\begin{aligned} \text{Frequencies} \\ \text{in } \tilde{F}_{y,m}(t)^+ : & \begin{cases} f_{\xi,k,p} + f_{m-k}, \\ f_{\xi,k,p} - f_{m-k}, \end{cases} \\ \text{in } \tilde{F}_{y,m}(t)^- : & \begin{cases} f_{\xi,k,p} + f_{m+k}, \\ f_{\xi,k,p} - f_{m+k}, \end{cases} \end{aligned} \quad (5.71)$$

where the form $f_{m\pm k}$ refers to the frequency of the transverse mode with mode number $m \pm k$, and the form $f_{\xi,k,p}$ refers to the p th frequency component of longitudinal mode k , corresponding to the frequency of the free response $f_{\xi,k}$ and to all the frequencies of the forced response, given by Eq. (5.41), which can be summarized as $f_q \pm f_{k\pm q}$, for $q = 1, 2, \dots$

Concentrating on the forced response of the longitudinal modes (the phantom partials), i.e., letting $f_{\xi,k,p} = f_q \pm f_{k\pm q}$ in Eq. (5.71) gives the excitation frequencies of mode m as

$$f_q \pm f_{k\pm q} \pm f_{m\pm k} \quad (5.72)$$

for all q and k . For harmonic transverse modal frequencies, the excitation frequencies given by 5.72 will coincide with the modal frequencies f_n , but for inharmonic strings they will provide a wideband excitation spectra. Naturally, those terms will have the largest influence on mode m in both cases, whose frequency equals the modal frequency of mode m .

From the free response of the longitudinal modes, which means $f_{\xi,k,p} = f_{\xi,k}$ in Eq. (5.71), the excitation frequencies are

$$f_{\xi,k} \pm f_{m\pm k}. \quad (5.73)$$

These frequencies are different from the transverse modal frequencies even in the case of perfectly harmonic transverse vibrations. Efficient excitation of transverse mode m occurs if this frequency is near f_m , which happens, e.g., if the sum $f_m + f_{m\pm k} \approx f_{2m\pm k}$ of two transverse modes with the distance k is near to the longitudinal modal frequency $f_{\xi,k}$.

5.4.2 The Stabilization Effect

We know from Sec. 5.3.6 that in a model where the longitudinal to transverse coupling is neglected, the amplitude of longitudinal motion can grow to unrealistically large value when the longitudinal mode is excited at its resonance. This does not happen in reality, as due the longitudinal to transverse coupling the rising longitudinal mode will influence the motion of its transverse parent modes significantly. In this section we aim at the qualitative explanation of this ‘‘stabilizing effect’’. First we investigate whether this growing forced component of longitudinal motion (which is actually a phantom partial) can act on its parents. For that, the nonlinear excitation force of the parent modes should contain a term at their natural frequency. Then, we check whether this excitation term decreases or increases the amplitude of the parent by comparing the phase of the parent and its

excitation force. As it can be seen in Eq. (5.67), the excitation force coming from the spatially uniform tension does not depend on the longitudinal displacement. Thus, the “stabilization effect” must originate from the excitation force $\tilde{F}_{y,m}(t)$ coming from the space-dependent tension $\tilde{T}(x, t)$.

Excitation Frequency

Let us suppose that we have two transverse modes m and n with the frequencies f_m and f_n , with $m \neq n$. It follows from Eq. (5.41) that these modes will excite the longitudinal modes $k = |m - n|$ and $l = m + n$. The excitation terms of both longitudinal modes will have the frequencies $f_m + f_n$ and $f_m - f_n$. As the fundamental frequency of the longitudinal vibration is much larger than that of the transverse vibration, the longitudinal mode $l = m + n$ with the modal frequency $f_{\xi, m+n}$ will not be effectively excited by the frequencies $f_m \pm f_n$. On the other hand, it is possible that longitudinal mode $k = |m - n|$ has a resonance frequency $f_{\xi, |m-n|}$ that is near to $f_m + f_n$, leading to an efficient excitation.

In Eq. (5.71) the only significant longitudinal component is now $f_{\xi, k, p} = f_m + f_n$ with $k = |m - n|$. For the transverse components, $f_{m-k} = f_{m-(m-n)} = f_n$ is present, while $f_{m+k} = f_{m+(m-n)} = f_{2m-n}$ is not, as we have assumed that the only significant transverse modes are m and n (we exclude the coincidences $m = 2m - n$ or $n = 2m - n$, as they would only occur if $m = n$). Accordingly, we only have to concentrate on $\tilde{F}_{y,m}(t)^+$ in Eq. (5.71), where the two excitation frequencies are $(f_m + f_n) - f_n = f_m$ and $(f_m + f_n) + f_n = 2f_n + f_m$. The latter have no significant effect, but the term with the frequency f_m will largely influence the motion of the transverse mode m . To sum up, one of the parents of a phantom partial together with the phantom partial itself create a mixing term in the excitation force which influence the other parent. The phantom partial have the sum frequency $f_m + f_n$ of the two transverse modes, thus, subtracting the frequency of either parents from this sum will give the frequency of the other parent.

Increase or Decrease

It has been clarified that a phantom partial can act back to its parent transverse partial by mixing with the other parent. However, it is still not clear whether this “feedback” increases or decreases the amplitude of the parent mode.

Let us assume that the two transverse modes m and $n < m$ have the instantaneous amplitudes $y_m = A_m \cos(\omega_m t + \varphi_m)$ and $y_n = A_n \cos(\omega_n t + \varphi_n)$. From Eq. (5.36), the excitation force of longitudinal mode $k = m - n$ will have the form

$$\begin{aligned}
 F_{\xi, k}(t) &= F_{\xi, k}(t)^- = -2ES \frac{\pi^3}{8L^2} A_m A_n y_m(t) y_n(t) (m - n) mn = \\
 &= -2ES \frac{\pi^3}{8L^2} (m - n) mn A_m A_n \cos(\omega_m t + \varphi_m) \cos(\omega_n t + \varphi_n) = \\
 &= -ES \frac{\pi^3}{8L^2} (m - n) mn A_m A_n [\cos((\omega_m + \omega_n)t + \varphi_m + \varphi_n) + \cos((\omega_m - \omega_n)t + \varphi_m - \varphi_n)],
 \end{aligned} \tag{5.74}$$

where only the first term with the frequency $\omega_m + \omega_n$ will effectively excite longitudinal mode $k = m - n$, as the other has a frequency which is much lower than $\omega_{\xi,k}$, where the dynamic response of the longitudinal mode is negligible (see Fig. 5.3). Note that $F_{\xi,k}(t) = F_{y,k}^-(t)$ in Eq. (5.74) because $\tilde{F}_{y,k}^+(t)$ is zero, since the two transverse modes that are present on the string have mode numbers m and n that does not satisfy $k = m + n$ (as $k = m - n$ and $m \neq n$).

Accordingly, the first term ($\omega_m + \omega_n$) of Eq. (5.74) is filtered by the transfer function of the dynamic response of longitudinal mode k , which is given in Eq. (5.48). By substituting $s = j\omega$ in Eq. (5.48), one obtains

$$\mathcal{F}\{\tilde{\xi}_{\delta,k}(t)\} = -\frac{2L}{ESk^2\pi^2} \cdot \frac{-\omega^2 + j2\sigma_{\xi,k}\omega}{-\omega^2 + 2j\sigma_{\xi,k}\omega + \sigma_{\xi,k}^2 + \omega_{\xi,k}^2}. \quad (5.75)$$

For $\omega \approx \omega_{\xi,k}$, i.e., around resonance, this gives

$$\mathcal{F}\{\tilde{\xi}_{\delta,k}(t)\} = \frac{L}{ESk^2\pi^2} \cdot \frac{\omega_{\xi,k}^2}{j\sigma_{\xi,k}\omega_{\xi,k} + (\omega_{\xi,k}^2 - \omega^2)/2}, \quad (5.76)$$

which was obtained by assuming $\omega \approx \omega_{\xi,k} \gg \sigma_{\xi,k}$.

For $\omega = \omega_{\xi,k}$, by using Eq. (5.31), we get

$$\mathcal{F}\{\tilde{\xi}_{\delta,k}(t)\}|_{\omega=\omega_{\xi,k}} = \frac{L}{ESk^2\pi^2} \cdot \frac{\omega_{\xi,k}}{j\sigma_{\xi,k}} = -j \frac{1}{k\sigma_{\xi,k}\mu\sqrt{ES}}. \quad (5.77)$$

The more ω departs from $\omega_{\xi,k}$, the less is the gain. For $\omega < \omega_{\xi,k}$, the phase shift is smaller than $-\pi/2$, while for $\omega > \omega_{\xi,k}$, it is larger.

If the longitudinal mode is excited by the sum term ($\omega_m + \omega_n$) of Eq. (5.74) exactly at resonance, the forced response of the longitudinal motion of mode k (or, the full motion after the transient response has died out) will be

$$\tilde{\xi}_k(t) = \frac{\pi^3\sqrt{ES}}{\sigma_{\xi,k}8L^2\mu} mnA_mA_n \cos[(\omega_m + \omega_n)t + \varphi_m + \varphi_n + \pi/2]. \quad (5.78)$$

Now it is possible to calculate the force acting on mode m originating from the motion of longitudinal mode k and transverse mode n by the help of Eq. (5.70a)

$$\begin{aligned} \tilde{F}_{y,m}(t) &= \tilde{F}_{y,m}(t)^+ = -\frac{ES\pi^3}{4L^2} \tilde{\xi}_k(t) y_n(t) \quad m \ n \ k = \\ &= -\frac{ES\pi^3}{4L^2} \frac{\pi^3\sqrt{ES}}{\sigma_{\xi,k}8L^2\mu} m^2 n^2 k A_m A_n^2 \cos((\omega_m + \omega_n)t + \varphi_m + \varphi_n + \pi/2) \cos(\omega_n t + \varphi_n) = \\ &= \frac{\pi^6(ES)^{1.5}}{\sigma_{\xi,k}64L^4\mu} m^2 n^2 k A_m A_n^2 [\cos((2\omega_n + \omega_m)t + 2\varphi_n + \varphi_m - \pi/2) + \cos(\omega_m t + \varphi_m - \pi/2)], \end{aligned} \quad (5.79)$$

where the second term with the frequency ω_m will have a strong effect on mode m , as it excites mode m exactly at resonance.

The velocity of transverse mode m is

$$v_m(t) = \frac{dy_m(t)}{dt} = A_m \omega_m \cos(\omega_m t + \varphi_m + \pi/2), \quad (5.80)$$

which can be used to write $\tilde{F}_{y,m}$ as a function of $v_m(t)$

$$\tilde{F}_{y,m}(t) = - \left(\frac{\pi^6 (ES)^{1.5}}{\sigma_{\xi,k} 64 L^4 \mu} m^2 n^2 k A_n^2 \omega_m \right) v_m(t), \quad (5.81)$$

where we have neglected the term $2\omega_n + \omega_m$ of Eq. (5.79), as it cannot effectively act on mode m .

Equation (5.81) shows that the force acting on mode m is directly proportional to the instantaneous velocity of mode m (as all the terms in the bracelets are constants), but it has an opposite sign. This means that the phenomenon will indeed decrease the amplitude of the parent mode m . The effect is similar to a damping term coming from friction. It is interesting to see that this damping term is proportional to the square of the amplitude A_n^2 of the other parent. This means that if there are two parents with different amplitudes, the “stabilization effect” will mostly act on the smaller one. Then, as the amplitude of the two parents gets smaller, so does the amplitude of the forced longitudinal component, decreasing this “stabilizing force” until a kind of energy balance is obtained.

Simulation Results

The motion of a G_1 piano string was computed by the discrete-time implementation of the equations of Sec. 5.3. The transverse vibration is calculated by finite-difference modeling, and the motion of longitudinal modes are computed by a modal model (see Sec. 6.4.2 for implementation details).

Figure 5.7 (a) solid line shows the evolution of the amplitude envelope of $\xi_1(t)$ for a G_1 piano string, where the coupling from the transverse to the longitudinal polarization is unidirectional, i.e., the longitudinal to transverse coupling is neglected. The transverse vibration is excited by the initial condition

$$y(x, 0) = 0, \quad (5.82)$$

$$\left. \frac{\partial y}{\partial t} \right|_{t=0} = 10 \left[\sin\left(\frac{m\pi x}{L}\right) + \sin\left(\frac{n\pi x}{L}\right) \right] \frac{\text{m}}{\text{s}}, \quad (5.83)$$

with $m = 9$ and $n = 10$, meaning that only modes m and n are present in the transverse motion. The amplitude envelopes of these two modes are displayed by dashed and dotted lines, respectively. The frequency of the first longitudinal mode $f_{\xi,1}$ is set in a way that it almost coincides with $f_m + f_n$ (the difference is less than 5 Hz). It can be seen in Fig. 5.7 (a) that the transverse modes have a simple exponential decay, while the amplitude of the first longitudinal mode increases as it is excited near resonance.

Figure 5.7 (b) shows the same example but now the longitudinal to transverse coupling is also included in the model, leading to the “stabilization effect”. Here the increase of the longitudinal motion leads to a faster decay of the transverse modes. As predicted

by the theory, the effect is stronger for that mode which has smaller initial amplitude (dotted line). The amplitude decrease of the transverse modes leads to the decay of the longitudinal motion. Note that after the amplitude of longitudinal motion reaches a certain smaller value, the transverse modes start to decay at their linear decay rate, which is the same as in Fig. 5.7 (a). It is interesting to see that two stage decay can occur in the amplitude envelopes of transverse modes not only because of the linear coupling of two transverse polarizations, but also because of the nonlinear coupling of one transverse and the longitudinal polarization. It is probable that even more complicated envelopes would arise if the other transverse (z) polarization was also taken into account in the nonlinear model.

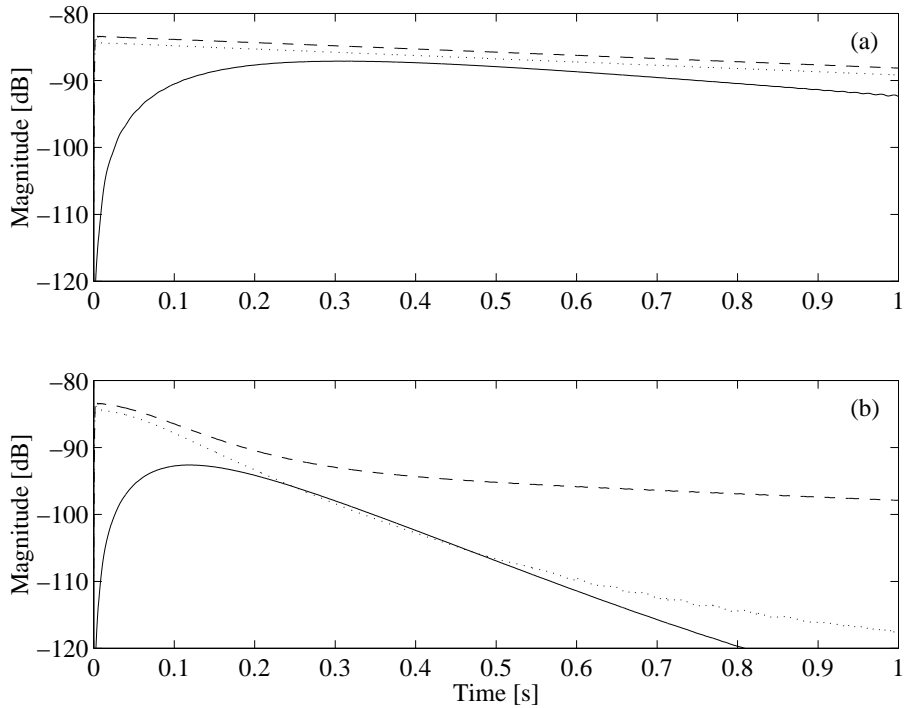


Figure 5.7: Amplitude of the first longitudinal mode (solid line), the ninth transverse mode (dashed line) and the tenth transverse mode (dotted line) with unidirectional (a) and bidirectional (b) coupling.

5.5 Conclusion

In this chapter the geometric nonlinearities of strings have been investigated. Section 5.1 has given the classification of the phenomenon by proposing a nonlinearity map as a function of longitudinal/transverse fundamental frequency ratios, the amplitude of string vibration, and the number of significant transverse modes. Four nonlinear regimes have been separated depending on whether the longitudinal to transverse coupling is significant and whether the tension can be considered spatially uniform on the string. These results have

been published in [Bank and Sujbert 2005a].

The two classes with spatially uniform tension have been covered in Sec. 5.2 by describing the results of the literature. Some nonlinear phenomena, such as the appearance of phantom partials, cannot be described by the spatially constant tension approximation. Therefore, a new theoretical model had to be developed.

The main results of this chapter are connected to the regime with spatially nonuniform tension but with negligible longitudinal to transverse coupling. In Sec. 5.3, a modal model has been proposed, which computes the longitudinal motion with an arbitrary number of transverse modes. The applied assumption, that the longitudinal to transverse coupling is negligible, is valid for most of the string instruments (piano, guitar, harpsichord, etc.). For both helping qualitative understanding and increasing the efficiency of sound synthesis models of Chap. 6, the tension computed by the modal model has been decomposed into a spatially uniform and a space-dependent part. The spatially uniform part equals the tension computed from the elongation of the string. Therefore, the spatially uniform approximation can be considered as a special case of the modal model. The model has been justified by showing good agreement with the measurement results. It has been shown that, for explaining phantom partials, the dynamic response (inertial effects) of longitudinal modes cannot be neglected. It also comes from the presented theory that longitudinal modes are continuously excited by the transverse motion, and not only during the excitation (i.e., hammer–string contact). As the model with rigid termination describes the phenomenon, it was found that the string termination has less significance. We note that the accuracy of the model has also been verified by the discrete-time implementations of the equations (this will be presented in Sec. 6.4.2). Most of these findings have been published in [Bank and Sujbert 2003, 2005b].

Section 5.4 has presented some recent results on the bidirectional coupling of the transverse and longitudinal polarizations. Particularly, the “stabilization effect” has been investigated, showing the qualitative reason why the longitudinal modes cannot grow infinitely even if they are excited at their resonance.

Chapter 6

Sound Synthesis of Geometric Nonlinearities

This chapter is about the sound synthesis of geometric nonlinearities of musical instrument strings, applying the physical principles presented in Chap. 5. In Secs. 6.1 and 6.2 sound synthesis methods based on the spatially uniform tension approximation are discussed. As the main part of this chapter, Sec. 6.3 presents various novel techniques for modeling the motion of longitudinal modes. The basic idea of the models is that the longitudinal motion is computed by nonlinearly excited second-order resonators. The methods have been developed with the applications to the piano, as the perceptual effect of the longitudinal component is the strongest for that instrument. However, they can be directly used to improve the sound quality of the models of other string instruments, such as the harpsichord or the guitar. Finally, Sec. 6.4 presents an efficient technique for modeling the bidirectional coupling of transverse and longitudinal polarizations.

6.1 Double Frequency Terms

In this case the string tension varies with time but spatially uniform along the string. Consequently, the longitudinal force component includes terms having double the frequency of transverse modes (see Sec. 5.2.2). These are called even phantoms in the notation of Conklin [1999]. As we have noted in Sec. 5.1.3, the tension variation is negligible compared to the initial tension T_0 , so it cannot excite any “nonlinear” transverse modes.

A straightforward modeling approach would be a linear string model for the transverse polarization, from which the tension is computed at every time instant by Eq. (5.15). Note that the longitudinal motion $\xi(x, t)$ does not need to be computed, since the temporal variation of the longitudinal bridge force is simply obtained as the tension variation $F_1(t) = -[T(L, t) - T_0] = -[\bar{T}(t) - T_0]$.

A computationally much more efficient approach was proposed by Karjalainen et al. [1993] for modeling the Kantele, a traditional Finnish instrument. There the output of the transverse string model (implemented by a digital waveguide) is lead to a second-order nonlinearity and a lowpass filter, and the result is mixed with the string output. The

nonlinearity adds the required double-frequency components, but some unwanted sum- and difference frequencies too. In [Karjalainen et al. 1993] the main motivation was to add a reinforcement to the second harmonic. As a result, efficient lowpass filtering could be used after the nonlinearity, suppressing the unwanted peaks. This model can be considered as a very efficient perception based approach, which works for the Kantele, but it is likely that it would give less agreeable results for other instruments, especially for inharmonic strings.

6.2 Tension Modulation

In this modeling paradigm the tension is spatially uniform along the string as in Sec. 6.1, but the temporal variation of the tension is no longer negligible in comparison with T_0 . This leads to the nonlinear excitation of transverse modes. The theory of this regime of string motion has been reviewed in Sec. 5.2. For modeling, the tension is computed by the discretization of Eq. (5.15) and then fed back to the transverse string model. The longitudinal force at the bridge equals the tension variation, similarly to the case of Sec. 6.1.

The most efficient approach for modeling the temporal modulation of tension is based on digital waveguides. In this case the effect of tension variation can be taken into account by varying the delay line length, which is done by a variable allpass filter at the termination [Tolonen et al. 2000; Erkut et al. 2002]. A computationally more demanding, but more accurate method is distributing the variable length delays between the delay elements [Pakarinen et al. 2003, 2005b]. Recently, an energy conserving variation of the technique have been presented in [Pakarinen et al. 2005a; Välimäki et al. 2006].

It is straightforward to implement tension modulation in finite-difference modeling, as the tension T_0 is an independent parameter of the model (see Sec. 2.3.1), which can be varied according to the tension computed from the string length. Bilbao [2004a] has presented such a model with energy-conserving property, which is beneficial as the stability of the model is guaranteed. Finite-difference schemes applying the constraint $\Delta x = c\Delta t$ are generally not well suited for modeling the modulation of tension, as the tension and the number of string elements are interdependent (see Sec. 2.3.1). However, Pakarinen et al. [2005b] have proposed a somewhat complicated solution to the problem, where the tension modulation is implemented by interpolating the string state between consecutive samples.

A modal-based tension modulation string model have been presented by Trautmann and Rabenstein [2000] that apply the Functional Transformation Method. Bilbao [2004b] has provided an energy conserving modal method for computing the response of tension modulated strings.

6.3 Modeling of Longitudinal Vibrations for Sound Synthesis

In this case the frequencies of the terms exciting the longitudinal modes are around or above the longitudinal modal frequencies. As a result, the tension varies with both time and space along the string. As the tension variation is small compared to T_0 , the longitudinal motion

does not influence the transverse vibration. For modeling, the largest difference from the cases of Secs. 6.1 and 6.2 is that now the motion of longitudinal modes also have to be computed.

In this Section, after the review of others work, the theoretical findings of Sec. 5.3 are used for sound synthesis purposes. As the goal is efficient sound synthesis, the theoretical model is simplified in a way that the most important features of the phenomenon are retained, while the less significant ones are neglected. We proceed from computationally more complex, physics-based approaches to more efficient, but less physical models.

6.3.1 Methods Proposed by Other Researchers

The first attempt for modeling the longitudinal component of the piano tone was made by Borin [2001] in the mid-nineties. In that model the transverse vibration of the string was computed by a digital waveguide. The hammer-string interaction force computed by the transverse string model was lead to an auxiliary digital waveguide, tuned at a much higher frequency, aimed at simulating the longitudinal components. A shortcoming of this model is that it is based on the assumption that the longitudinal string motion is excited during the hammer-string contact only. (We know from Sec. 5.3 that the longitudinal motion is continuously excited during the whole string motion.) As a result, only the free response of the longitudinal motion is simulated, while the forced response (phantom partials) is neglected. This reproduces some of the features of the longitudinal component, but the longitudinal component sounds as a separate tone, unlike in real pianos. Informal listening tests show that the simulation of phantom partials is even more important than modeling the free response (see Sec. 5.3.7).

A very efficient method for modeling the phantom partials was proposed by Bensa and Daudet [2004]. In that model some of the transverse modes of the string output are filtered by narrow bandpass filters, then these components are multiplied by the transverse string signal itself. As a result, sum- and difference frequency components arise between the transverse peaks, simulating the perceptual effect of phantom partials. The shortcoming of the approach is that it is based on a wrong physical model. In [Bensa and Daudet 2004] the tension is considered to be uniform along the string (which is not the case for the piano), and the main source of tension variation is a linear coupling from the transverse motion. As a result, the frequencies of these mixing terms are different from that found in real pianos. Although it is physically not correct, this is the most efficient approach for modeling the perceptual effect of phantom partials.

In [Caramaschi 2004] the idea of the composite string model of [Bank and Sujbert 2004] (presented here in Sec. 6.3.4) have been applied for the digital waveguide. That is, the transverse string displacement has been computed by a digital waveguide string model instead of the finite-difference scheme. Although being more efficient, this way of computing the excitation force of the longitudinal modes produces audible artifacts. This is discussed in Sec. 6.3.6.

6.3.2 Finite-difference Modeling

A straightforward approach for modeling the vibration of musical instrument strings is implementing the simultaneous differential equations (5.27) and (5.62) by finite-difference approach. Actually, this approach is capable of simulating the bidirectional coupling of the transverse and longitudinal polarizations, so it “knows more” than required (it rather belongs to Sec. 6.4). In an earlier work [Bank and Sujbert 2003] such a model was developed. The computational demand of such an approach is large because high sampling frequency ($f_s \approx 500$ kHz) is required due to the higher propagation speed in the longitudinal direction. Another reason of waste of resources is that the longitudinal displacement is computed at each point, while this is not necessary, as here we assume that the feedback from the longitudinal to the transverse vibration is negligible. All we need to know is the longitudinal bridge force, which is the tension at the termination. (Note that the transverse displacement has to be known at each point along the string for the computation of the longitudinal excitation force by the scalar product of Eq. (5.30b).) Still, this approach can be very useful for experimental purposes. A commercial computer program based on a finite-difference string model was written by Stopper [2003] helping piano tuners in scale design.

6.3.3 Tension Decomposition

An equally precise but computationally less demanding way of computing the longitudinal force at the bridge is using the idea of tension decomposition presented in Sec. 5.3.4.

In this model the transverse vibration is computed by any of the standard techniques, e.g., by finite-difference modeling. The spatially uniform part of the tension is computed from the elongation of the string by Eq. (5.15). This is followed by calculating the excitation forces $F_{\xi,k}(t)$ of the longitudinal modes by Eqs. (5.28) and (5.30b). Then, the dynamic response $\tilde{\xi}_k(t)$ of the first few (e.g., 5–20) longitudinal modes is computed by feeding these excitation forces into second-order highpass filters, which implement Eq. (5.48). By knowing $\tilde{\xi}_k(t)$, the bridge force (which equals the tension at the termination) is obtained as

$$F_l(t) = -T(L, t) = - \left[T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left(\frac{\partial y}{\partial x} \right)^2 dx + ES \frac{\pi}{L} \sum_{k=1}^K k \tilde{\xi}_k(t) (-1)^k \right]. \quad (6.1)$$

Naturally, the derivatives and the integration in Eq. (6.1) have to be substituted by finite-differences and summation, respectively. This also holds for the rest of this chapter, i.e., when a continuous-time equation appears, it has to be discretized for implementation.

The method provides almost the same output as the full finite-difference model (see Sec. 6.4) while its computational cost is around the 20% of the full finite-difference model (the computational cost is estimated by the number of additions and multiplications). This is because now the longitudinal modes are computed by second-order highpass filters and not by a finite-difference scheme, which eliminates the need for high sampling rates. The sampling rate is now a function of the transverse string model, where 44.1 kHz is usually

sufficient even for a finite-difference transverse string model. However, it is possible to decrease the complexity even more by some perceptual simplifications.

6.3.4 The Composite String Model

The composite string model is a simplification of the tension decomposition model of Sec. 6.3.3. The model structure is depicted in Fig. 6.1. The transverse deflection $y(x, t)$ is computed by a finite-difference string model running at audio sampling rate (e.g., $f_s = 44.1$ kHz), which implements the differential equation of Eq. (2.50). The discretization is done as discussed in Sec. 2.3.1. A finite-difference hammer model (see Secs. 2.4.1 and 4.1) is also attached to the string. The initial velocity of the hammer is denoted by v_0 in Fig. 6.1.

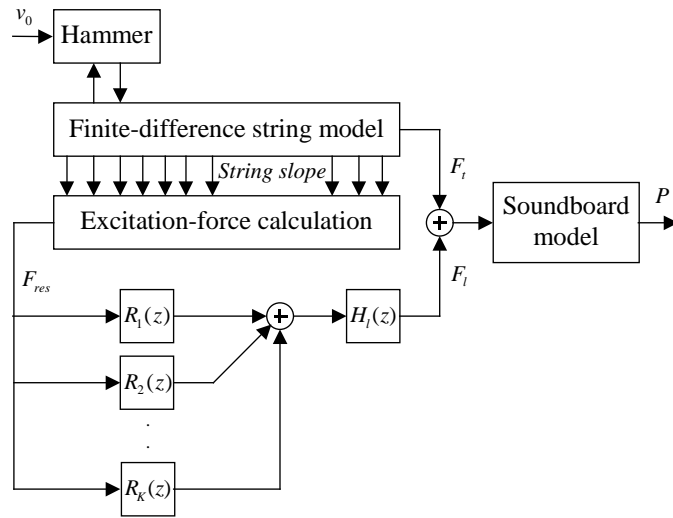


Figure 6.1: The composite string model applying finite-difference modeling and second-order resonators R_1, \dots, R_K .

The excitation-force density of the longitudinal motion, $d_\xi(x, t)$, is computed according to Eq. (5.28) from the transverse displacement calculated by the finite-difference transverse string model. Then the excitation force $F_{\xi,k}(t)$ of each longitudinal mode k is computed by a scalar product with the longitudinal modal shape as in Eq. (5.30b). The instantaneous amplitudes $\xi_k(t)$ of the longitudinal modes are calculated according to Eq. (5.30a), which is implemented by second-order resonators (R_1, \dots, R_K in Fig. 6.1). The computationally heavy part of longitudinal-vibration modeling lies in Eqs. (5.28) and (5.30b). Especially the load of Eq. (5.30b) is heavy, as the force input $F_{\xi,k}(t)$ is computed by scalar products for all the modes ($N \approx 10$ in practice) separately. Therefore, further simplifications are necessary.

The excitation spectrum (the Fourier transform of $F_{\xi,k}(t)$) of all the odd and all the even longitudinal modes are very similar, respectively. It can be seen in Fig. 6.2 that the only difference is that the frequency peaks are slightly shifted as a function of mode number k because of the inharmonicity of the string. The amplitudes are also somewhat different, but the general envelopes are of quite similar structure. Therefore, it is a reasonable choice

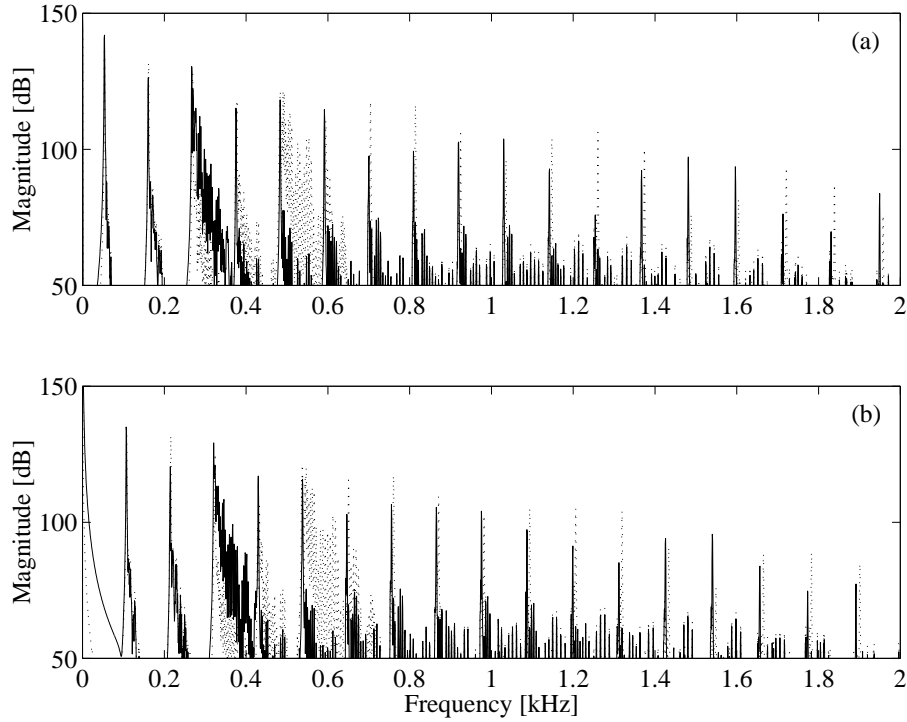


Figure 6.2: The spectrum of the excitation forces of longitudinal modes. (a): The excitation force of longitudinal mode 5 (solid line) and 9 (dotted line). These modes contribute to odd phantoms. (b): The excitation force of longitudinal mode 6 (solid line) and 10 (dotted line). These modes give a rise to even phantoms.

to substitute the excitation force $F_{\xi,k}(t)$ of all the odd longitudinal modes by the excitation force of one odd longitudinal mode (e.g., $F_{\xi,k}(t) = F_{\xi,5}(t)$ for odd k). The same can be done for the even longitudinal modes. It follows from Eq. (5.41) that odd phantoms arise from the vibration of odd longitudinal modes and even phantoms from the vibration of even ones. Therefore, it is important to incorporate at least one odd and one even modal shape. Having only one modal shape in the model would lead to an excitation spectrum with odd or even harmonics only. Accordingly, the model can be simplified by computing the force input for two modes (e.g., $F_{\text{res}} = F_{\xi,5} + F_{\xi,6}$, but any other odd and even mode would do) and using this as a common excitation for all the resonators. This way the multiple (or smeared) peak of a real phantom partial (as depicted in Fig. 5.5) is substituted by a single exponentially decaying sinusoid. The double-frequency terms are also neglected in this model. However, it seems that the ear is insensitive to this small variation, as the simplified model leads to almost identical perceptual results compared to the full model of Sec. 6.3.3, but requires a significantly lower computational power.

As already noted in Sec. 5.3.6, if the longitudinal modes are excited at resonance, the longitudinal motion can have unrealistically high magnitude, leading to unrealistic sound. This would not happen in reality, as the feedback from longitudinal to transverse motion would diminish the amplitude of the transverse parents, thus, the longitudinal mode as well

(see Sec. 5.4.2). However, this feedback is not included in the model, as it would increase the computational complexity, and its effect is only significant in this situation. It is much simpler to set the frequencies $f_{\xi,k}$ of the resonators R_1, \dots, R_K in a way that they do not coincide with the peaks of their excitation signal $F_{\text{res}}(t)$. This can be done automatically by computing the excitation frequencies of the longitudinal modes by Eq. (5.41) and slightly shifting those longitudinal modal frequencies, which are excited near resonance.

The force signals $F_t(t)$ and $F_l(t)$ in Fig. 6.1 coming from the transverse and longitudinal polarizations are sent to the soundboard model, which computes the sound pressure $P(t)$. The soundboard is modeled by a multi-rate filtering algorithm approximating the measured impulse response of a transversely excited piano soundboard (see Sec. 4.3). The soundboard responds slightly differently to a longitudinal force than to a transverse one. This difference can be modeled by the filter $H_1(z)$ in the longitudinal force path. However, it was found that a constant coefficient $H_1(z) = C_1$ in the range of 0.2–0.5 produces adequate sound.

The sound pressure spectrum of the first second of a synthesized G_1 note is displayed in Fig. 6.3 (a). The phantom partials are clearly visible between the transverse modes, which are emphasized around the longitudinal free mode at 1450 Hz. The circle indicates the component coming from the longitudinal free response, while the crosses show the transverse modal frequencies. It can be seen that the spectrum is similar to that of a real piano tone displayed in Fig. 5.4. The composite string model produces the same sound quality as the methods of Secs. 6.3.2 and 6.3.3, while its computational requirements are reduced significantly.

6.3.5 The Resonator-based String Model

In the resonator-based string model the string displacement is represented by its modal form (see Eqs. (5.33) and (5.29)) for both the transverse and longitudinal polarizations and the instantaneous amplitudes $y_n(t)$ and $\xi_k(t)$ are computed by second-order resonators. Therefore, it can be considered as a variation of the composite model of Sec. 6.3.4, where the finite-difference transverse string model is replaced by a modal model.

The string is excited by a hammer in the transverse polarization. The hammer is modeled in the same way as in the case of finite-difference string (see Secs. 2.4.1 and 4.1). The string response to the hammer force is calculated by a set of second-order resonators, which have input and output coefficients depending on the hammer position, as discussed in Sec. 2.3.3. The outputs of these resonators correspond to the instantaneous amplitudes $y_n(t)$ of the transverse vibration, which can be directly used to compute the excitation force $F_{\xi,k}(t) = F_{\xi,k}(t)^+ + F_{\xi,k}(t)^-$ of the longitudinal modes by using Eq. (5.36). From this point, the approach is the same as taken in Sec. 6.3.4: the excitation force of one even and one odd longitudinal mode is calculated and summed (e.g., $F_{\text{res}}(t) = F_{\xi,5}(t) + F_{\xi,6}(t)$). This signal is then fed to the resonators calculating the instantaneous amplitudes $\xi_k(t)$ of the longitudinal modes. The efficiency can be further increased if those components of the excitation signal $F_{\text{res}}(t)$ are not computed where the gain of the longitudinal resonator bank is small.

This model is capable of producing the same sound quality as the composite string

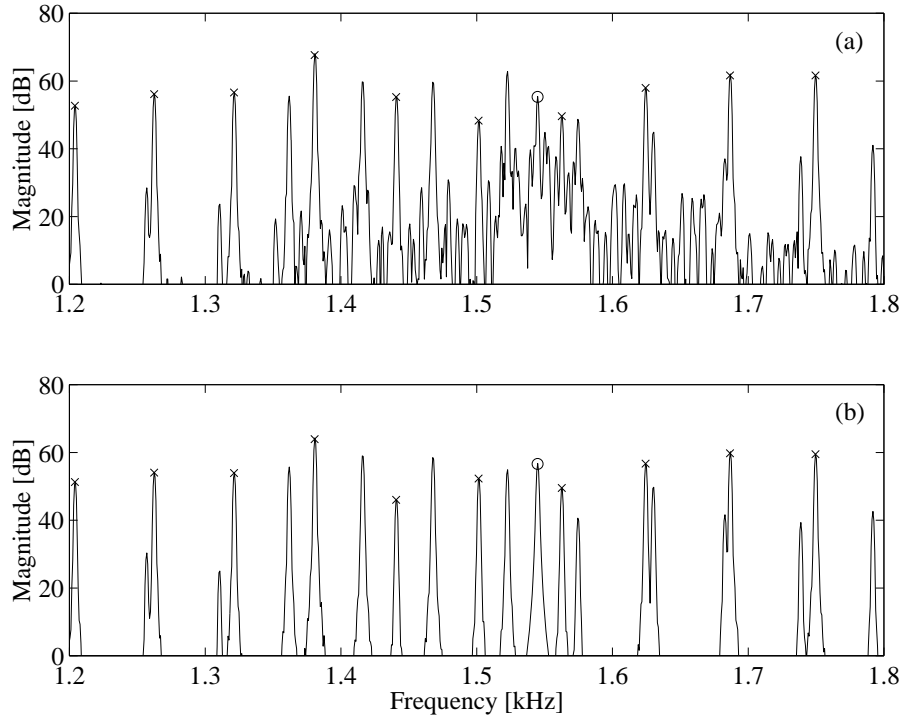


Figure 6.3: The sound pressure spectrum of the first second of a synthesized G_1 piano tone computed by the composite string model of Sec. 6.3.4 (a) and by the resonator-based string model of Sec. 6.3.5 (b). Crosses indicate transverse partials and the second longitudinal mode is marked by a circle in both figures. All the other peaks correspond to phantom partials. To be compared with Fig. 5.4 (a).

model of Sec. 6.3.4 when the number of resonators implementing the transverse modes equals the number of string elements in the finite-difference model. Figure 6.3 (b) displays the sound pressure spectrum of the first second of a G_1 piano tone synthesized by the resonator-based string model. It can be seen in Fig. 6.3 that the resonator-based model produces a similar output compared to the composite model of Sec. 6.3.4 when the string and hammer parameters are set to be the same. The only difference is that the composite string model generates noise-like peaks between the dominant partials due to computational inaccuracies. However, this is not considered as an advantage because the difference between the output of the two models is almost inaudible [Sound examples n.d.].

An advantage of the resonator-based approach is that the computational complexity is reduced by a factor of two. Moreover, this method is particularly advantageous when the goal is to reproduce a tone which is similar to that of a particular instrument since the measured partial frequencies f_n and decay times τ_n can be directly implemented in the model. On the other hand, the resonator-based model is less physical in the sense that the physical parameters of the string (such as string mass and tension) have only indirect connection to the model.

6.3.6 Digital Waveguide and Longitudinal Resonators

Based on the composite model of Sec. 6.3.4, it seems to be a straightforward simplification to compute the transverse string shape by a digital waveguide model instead of finite-difference modeling. This is advantageous because digital waveguide string models need less than 10% computational power compared to the finite-difference approach.

However, while a finite-difference transverse string model is made up of 100 elements (the transverse displacement is computed at 100 points along the string), the digital waveguide model of a low piano string may consist of 1000 unit delays. This means that the spatially discretized implementation of excitation-force calculation (Eqs. (5.28) and (5.30b)) should be computed at 1000 points, instead of 100. This increase can be overcome by spatially downsampling the string shape computed by the digital waveguide model, similarly to what has been done in the case of tension modulation modeling [Tolonen et al. 2000]. This leads to spatial aliasing, unless some spatial lowpass filtering is used. Unfortunately, the computational complexity of such a filtering is high. Therefore, the best option seems to avoid lowpass filtering, and hoping that the difference cannot be heard, as it was also done in [Tolonen et al. 2000].

A more intrinsic problem with the digital waveguide is that it is intended to compute the string motion at specific points only. Actually, its efficiency comes from the fact that the effects of dispersion and losses can be moved anywhere between the excitation and observation point. In linear string models there is usually one observation point (the bridge of the instrument), therefore all the losses and dispersion are realized as a single filter in the delay loop. However, in our case the transverse string displacement should be known at each point, or, if we spatially downsample as suggested above, at 50–100 different points. This means 50–100 observation points, requiring 100–200 small filters between them. Unfortunately, this leads to a computational complexity comparable to the finite-difference method or the modal-based approach.

It is interesting to see what happens if the losses and dispersion are still lumped into one filter (although we know from the previous paragraph that this is not correct), and calculate the excitation force of the longitudinal modes by Eqs. (5.28) and the scalar product of (5.30b). In this case the spectrum of the excitation force $F_{\xi,k}(t)$ is not quasi-harmonic as depicted in Fig. 5.2 but contains a lot of inharmonic and noise-like peaks. This is shown in Fig. 6.4 (b), while the excitation force of a harmonic waveguide is displayed in Fig. 6.4 (a) for comparison. Figure 6.5 (a) displays the 1.5–2 kHz region of Fig. 6.4, giving a better insight to the erroneous peaks between the desired ones. Unfortunately this leads to an unsatisfactory sound. When the dispersion filter is distributed between the observation points, these peaks have much smaller amplitude, as shown in Fig. 6.5. In this case 128 first order allpass filters were implemented. Note that not only the noise-like peaks disappear, but so does every second partial, giving the spectrum an even-like character, as in Fig. 5.2.

The reason for this difference is that the modal shapes of the string computed by the lumped dispersion filter are not orthogonal to each other, since some parts of the modal shapes are “within” the dispersion filter. As the dispersion filter has different delay at

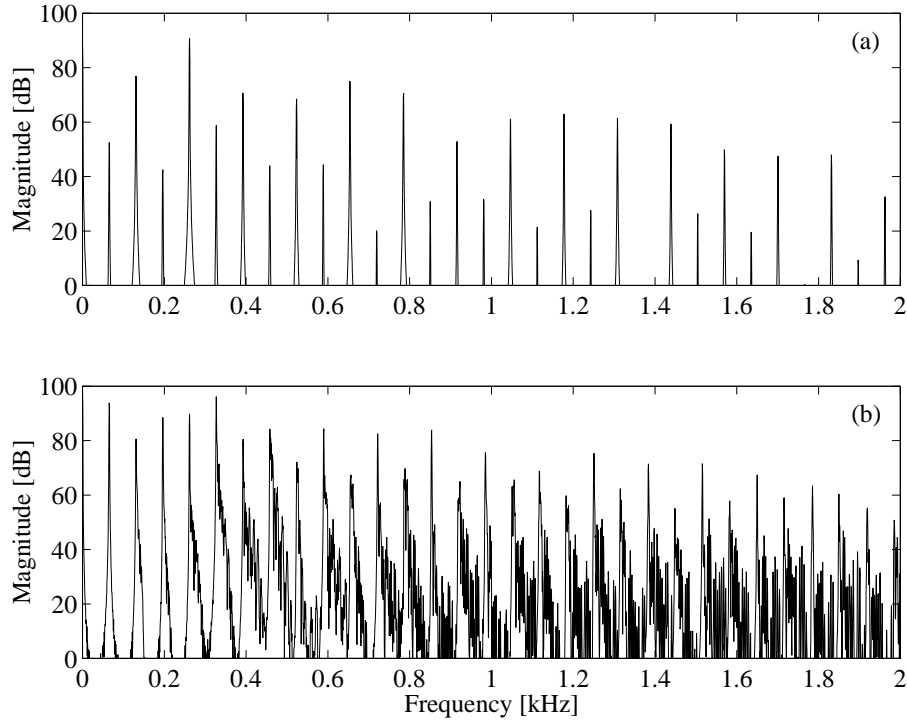


Figure 6.4: Spectrum of the longitudinal excitation force $F_{\xi,6}(t)$ of a C_2 piano tone synthesized by a harmonic (a) and an inharmonic (b) digital waveguide with lumped dispersion filter.

different frequencies, the lengths of the modes will be different. In this case the L in the modal form of Eq. (5.33) should be substituted by L_n , where L_n is different for all n . On the other hand, if the inharmonicity is negligible, this approach works precisely as can be seen in Fig. 6.4 (a). The lumped loss filter changes the modal shapes only slightly, which is in the order of the effect that comes from the nonrigid termination of real instruments, discussed in Sec. 5.3.6.

As a result, if the string dispersion is not implemented, it is beneficial to compute the transverse string displacement by a digital waveguide model. Then the transverse displacement should be spatially downsampled to 50–100 samples, followed by the steps outlined in Sec. 6.3.4 for computing the longitudinal modal response. On the other hand, for highly dispersive strings, there is no particular advantage in using digital waveguides.

6.3.7 Physically Informed Modeling Techniques

A very efficient way of modeling the longitudinal components is implementing them as a signal-based model. This means that the longitudinal force at the bridge $F_1(t)$ is synthesized as a signal, without any concern about its origin. The approach is called physically informed, as the parameters of the signal are calculated from the physical parameters of the string off-line by the equations presented in Sec. 5.3, or by analyzing the longitudinal component computed by a physics-based model.

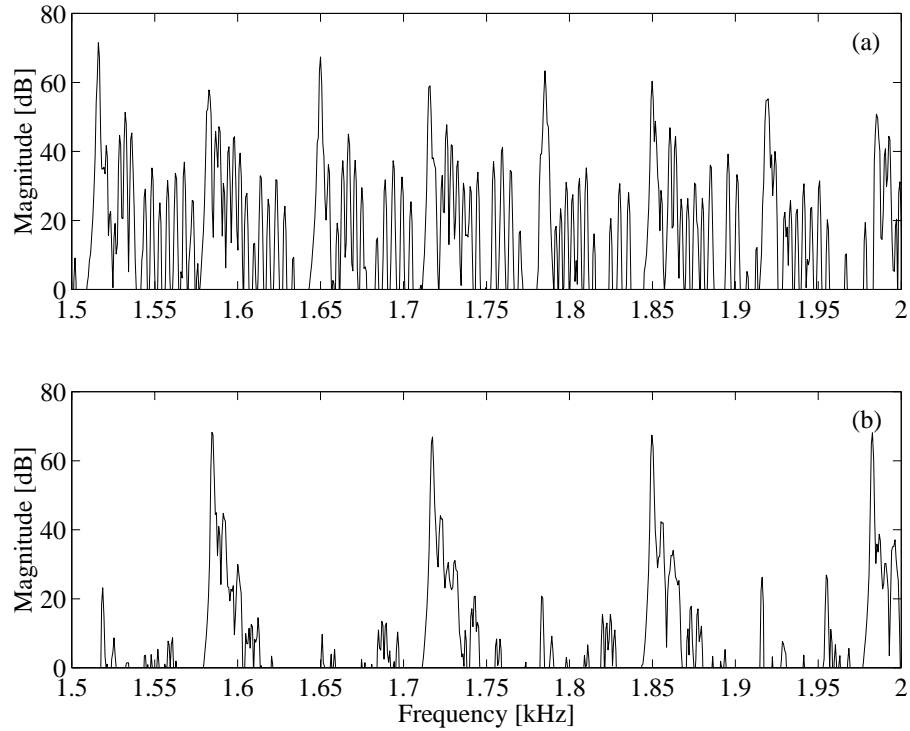


Figure 6.5: Spectrum of the longitudinal excitation force $F_{\xi,6}(t)$ of a C_2 piano tone synthesized by an inharmonic waveguide with lumped (a) and distributed (b) dispersion filters. Note that (a) is the same example as in Fig. 6.4 (b), displayed at a higher resolution.

An example can be first computing the longitudinal component by the composite model of Sec. 6.3.4. Then, this longitudinal component is analyzed and resynthesized by additive synthesis. This means that the longitudinal signal is built up by exponentially decaying sinusoids, implemented by second-order resonators. Note that separate sinusoids are needed for the forced and for the free response of the longitudinal modes, i.e., these resonators are not the same as $R_1..R_K$ in Fig. 6.1. Next time when the note is played, there is no need to compute the longitudinal response by the physical model, as it can be synthesized more efficiently by adding up the corresponding sinusoids. Naturally, one parameter set is valid for one specific note at a specific dynamic level. Therefore, the parameters of the decaying sinusoids for all the notes and playing levels should be stored or calculated before each note is sounded. The gain is that computational complexity is decreased significantly, and that digital waveguide models can be used to compute the transverse vibration even in the case of dispersive strings as now the longitudinal modes are computed separately. On the other hand, the flexibility of the physical modeling is lost and huge amount of parameters has to be stored.

The Square of the Excitation Force as an Input Signal

A simple but efficient way of avoiding the parameter dependence of the longitudinal model on dynamic level is to implement the longitudinal component as second-order resonators with a special input. The input of these resonators is the square of the excitation force (hammer force for pianos) computed by the transverse string model. This way, the amplitude of the longitudinal component will be proportional to the square of the amplitude of the transverse component. As an effect, the longitudinal component will be relatively more prominent at forte playing levels, which is also the case in reality. The model structure is depicted in Fig. 6.6. The resonators of the longitudinal component are excited only during the hammer-string contact. This may seem contradictory to the discussion in Sec. 6.3.1, where it is stated that the longitudinal modes are excited during the entire string motion, and not only during the hammer-string contact. However, now the resonators of the longitudinal components are not only those corresponding to the free response, but the forced components are also represented. The only role of exciting the resonators by the square of the hammer force is to set the initial amplitudes of the resonators. With this simple trick we have avoided the need of storing the resonator amplitudes at different dynamic levels. As a result, only one parameter set is required for each note.

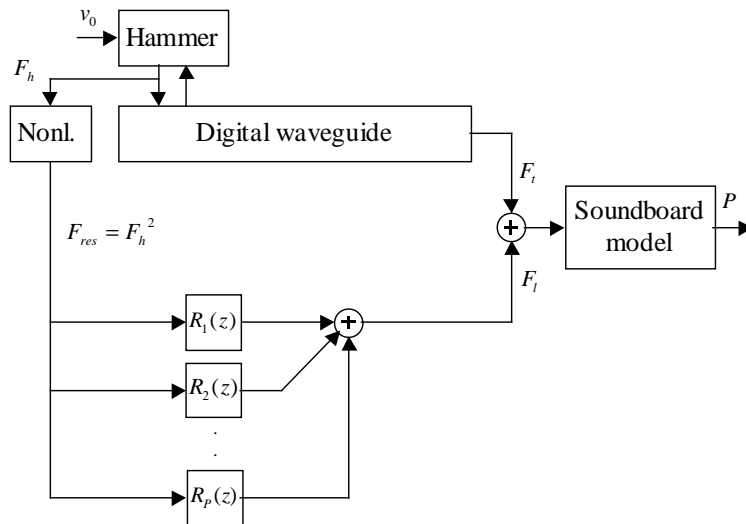


Figure 6.6: An example for the physically informed synthesis of longitudinal motion.

With this model it is possible to reach the sound quality of the models presented in Secs. 6.3.4 and 6.3.5 at 10% of their computational cost when the transverse vibration is calculated by a digital waveguide, but some of the nice features of physical modeling are lost. For example, when the string is struck again while it is still in vibration, the longitudinal component coming from the model will be different from the one that would arise from physics-based models, although this difference usually does not mean worse sound quality.

Other Alternatives

Along these lines, many different alternatives are possible. For example, one might compute the input signal of the longitudinal modes $F_{\text{res}}(t)$ of Fig. 6.1 by a signal model. This is advantageous as $F_{\text{res}}(t)$ is quasi harmonic, i.e., it can be computed by a dispersive digital waveguide model with one fourth of inharmonicity compared to the main one. Then, this is lead to the resonators R_1, \dots, R_K implementing the modal response of the longitudinal modes, whose parameters now can be the same as in the case of Sec. 6.3.4. The only problem with this approach is that the initial phases of the partials coming from the digital waveguide are all the same, leading to a spiky excitation signal, which produces a somewhat unnatural sound. This might be avoided by randomizing the phase with allpass filters, or by simply fading in the excitation signal $F_{\text{res}}(t)$ after the first spikes have been filtered by the allpass filter in the waveguide loop.

Naturally, it is also possible to substitute the longitudinal resonators R_1, \dots, R_K by a digital waveguide. In this case, the system is made up of three digital waveguide models: The first digital waveguide computes the transverse vibration. The second digital waveguide provides the excitation force for the longitudinal modes that are implemented by the third digital waveguide.

For those instruments where the longitudinal component decays faster, it might also be possible to store it in a wavetable. In that case the transverse vibration is synthesized by a physics-based solution, while the longitudinal component is simply played back from memory.

Commuted Synthesis

One important property of the physically informed approaches is that they can be easily linearized. That is, instead of squaring the hammer force, it is lead to the model of the longitudinal component through a constant coefficient, whose value is changed according to dynamic level. As a result, the whole system becomes linear, thus, its elements can be simply commuted, as discussed in Sec. 2.5.3. This is displayed in Fig. 6.7. The body response is stored in a wavetable. This wavetable is read sample by sample and the corresponding signal is sent to a filter implementing the effect of the hammer. The filtered signal is lead to the transverse string model. The same signal is also lead to the model of the longitudinal component, but through a constant gain, which varies with dynamic level. The outputs of the transverse and longitudinal models are simply added and produce the system output. This way, a nonlinear system with constant parameters has been transformed into a linear one having a parameter that depends on dynamic level.

For those instruments, where the longitudinal components decay fast (e.g., in the case of the guitar), this might not even be needed, if the body response is computed by inverse filtering, as for example in [Tolonen 1998; Erkut et al. 2000]. This is because the inverse-filtered version of the longitudinal component of the tone is already contained in the body response, and when filtered by the transverse string model it is perfectly reproduced. Actually, this form of commuted synthesis reproduces the tone perfectly up to the length of the wavetable. All what is needed is setting the length of the wavetable longer than the

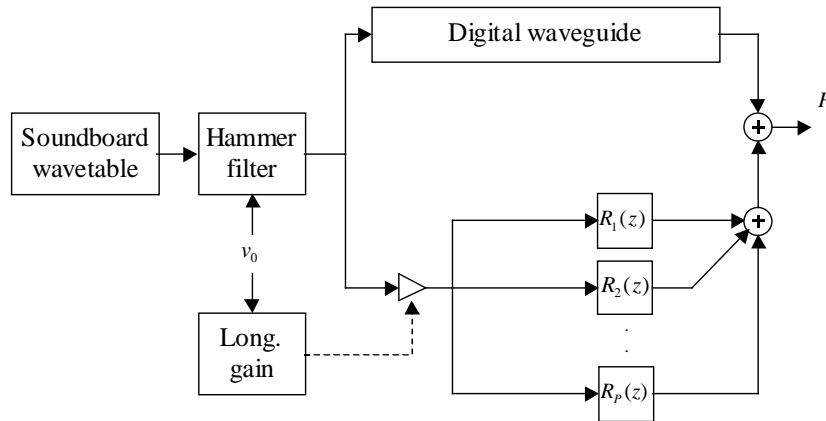


Figure 6.7: Commuted synthesis of longitudinal vibration.

significant portion of the longitudinal component. In the case of the guitar half a second might be appropriate (for the piano, 2–5 seconds would be required). Unfortunately, in this case the ratio of the transverse and longitudinal components is fixed, which varies by dynamic level in real instruments. This can be avoided by either storing different body response tables for the different playing levels, or by factoring out the longitudinal contribution from the inverse filtered signal and storing it in a separate wavetable. This way the amount of longitudinal components can be varied by scaling the output of this second wavetable.

Applications to Sampling

The same ideas can be used to improve the quality of sampling synthesizers. The most common way of implementing the change of the timbre due to the variation of hammer impact speed is to filter the prerecorded sounds by a low-order filter whose parameters depend on the playing level. However, this filter cannot implement the timbre variation due to the appearance of the longitudinal component at higher levels. (Or, if the forte notes are stored, it cannot filter out all the longitudinal components.) The common solution is to store the piano sound at various (3–5) different dynamic levels, and interpolating between these. However, interpolation often causes strange artifacts in the sound. Because of this, some samplers store the sound of the same note at even more different dynamic levels and switch between these without interpolation. Unfortunately, this requires gigabytes of sample memory.

The solution proposed here is to store two wavetables for each note, one containing the transverse component, and the other one containing the longitudinal component. Then, these are filtered through simple lowpass filters implementing the effect of the hammer-string interaction, while the appearance of the longitudinal modes is controlled by the ratio with which the transverse and longitudinal signals are added. This solution leads to a continuous variation of the tone when the dynamic level changes, while requires less memory than the multiple wavetable approaches. The difficulty of the approach is that

the recorded tones have to be decomposed into transverse and longitudinal components.

6.3.8 Comparison

In Section 6.3 many novel techniques have been proposed for the synthesis of the longitudinal component of string vibration. These are summarized in Table 6.1. The column “Physics-based” indicates whether the computation of longitudinal motion is physically meaningful. In practice, this means that the model behaves properly when the string is restruck, damped, or coupled to other strings. The column “Accurate” signs whether the longitudinal bridge force is calculated accurately. If this is not the case, that means that some perception-based simplifications were made, providing different output but similar sound quality. The column “Computational load” shows the approximate computational complexity compared to the full finite-difference model. As noted earlier in Chap. 4, the computational load is approximated by the number of additions and multiplications.

The choice between the different models depends on many factors, for example, whether the task is sound synthesis or scientific investigation (in the latter case “Accurate” models are needed). Moreover, it is influenced by which type of transverse string model should be used because of other reasons. Naturally, it also depends on the available computational power.

	Physics-based	Accurate	Computational load
Finite-difference modeling	×	×	100%
Tension decomposition	×	×	20%
Composite model	×		10%
Resonator-based model	×		5%
Waveguide and resonators	×		2%*, (5%)
Physically informed			1%

Table 6.1: Main features of the different techniques for modeling the longitudinal component of piano tones presented in Sec. 6.3. The computational load is the complete load for computing both the transverse and longitudinal polarizations. The * sign means that the 2% is valid for harmonic string models. In the case of dispersive strings it is about 5 %, as the distributed allpass filters increase the complexity.

6.4 Modeling of Bidirectional Coupling for Sound Synthesis

In this section some of the recent results of bidirectional modeling is presented for completeness. The earlier techniques for modeling the bidirectional coupling are based on finite-difference modeling [Bank and Sujbert 2003; Stopper 2003], leading to large computational complexity. Here a new approach based on tension decomposition is proposed, which produces the same sound quality at lower computational cost.

6.4.1 Finite-difference Modeling

As we have already noted in Sec. 6.3.2, a finite-difference bidirectional string model has been presented in [Bank and Sujbert 2003], which is based on the spatial and temporal discretization of the transverse wave equation Eq. (5.62) and the longitudinal wave equation Eq. (5.27). However, the high sampling frequency (ca. 500 kHz) required because of the higher longitudinal propagation speed leads to large computational complexity. This is a waste of resources as these high frequency components are not present in the vibration of musical instrument strings, and we cannot hear them anyway. On the other hand, computing the longitudinal displacement at each point along the string is not a wastage as it was for the case of unidirectional modeling of Sec. 6.3.2, as now the tension variation has to be known at each point, since it acts back on the transverse vibration of the string. Accordingly, a new method had to be found which still computes the string state at each point along the string but eliminates the need for high sampling rates.

6.4.2 Tension Decomposition

The method is based on the tension decomposition presented in Sec. 5.3.4, but it can also be seen as the extension of the unidirectional tension-decomposition model of Sec. 6.3.3. For computing the transverse displacement, we may use:

$$\mu \frac{\partial^2 y}{\partial t^2} = \frac{\partial \left\{ T(x, t) \frac{\partial y}{\partial x} \right\}}{\partial x} - ES\kappa^2 \frac{\partial^4 y}{\partial x^4} - 2R(\omega)\mu \frac{\partial y}{\partial t} + d_{y,\text{ext}}(x, t), \quad (6.2)$$

which can be obtained from Eqs. (5.62) and (5.63). For modeling losses, the substitution of Eq. (2.48) can be used instead of the term $R(\omega)$. The external excitation force density (such as the effect of the hammer, plucking, or bowing) is taken into account by the term $d_{y,\text{ext}}(x, t)$.

For the first time step, the transverse string motion is computed by $T(x, t) = T_0$. Then, the dynamic responses $\tilde{\xi}_k(t)$ of the first few (5–20) longitudinal modes are calculated as described in Sec. 6.3.3. The first step is the calculation of the excitation-force density $d_\xi(x, t)$ by Eq. (5.28) and the determination of the longitudinal excitation forces $F_{\xi,k}(t)$ by the scalar product of Eq. (5.30b). The instantaneous amplitudes $\tilde{\xi}_k(t)$ are computed by second-order resonators, implementing Eq. (5.48). The string tension is computed as

$$T(x, t) = T_0 + \frac{1}{2} \frac{ES}{L} \int_0^L \left(\frac{\partial y}{\partial x} \right)^2 dx + ES \frac{\pi}{L} \sum_{k=1}^K k \tilde{\xi}_k(t) \cos \left(\frac{k\pi x}{L} \right), \quad (6.3)$$

which is made up the initial tension T_0 , the spatially uniform tension variation computed from the elongation of the string, and the tension coming from the dynamics of longitudinal modes. Then, this tension is used for the finite-difference string model implementing Eq. (6.2), which computes $y(x, t)$ for the next t value, from which the tension $T(x, t)$ can be computed, and so on.

Note that Eq. (6.3) differs from Eq. (6.1) in the last term because the tension $T(x, t)$ has to be computed all along the string $0 \leq x \leq L$ and not only for $x = L$. This means

that now the contributions of $\tilde{\xi}_k(t)$ has to be projected back to the string by the functions $\cos(k\pi x/L)$, leading to higher computational complexity compared to Eq. (6.1). Accordingly, the computational requirement mostly depends on how many longitudinal modes are implemented. This is because for each new longitudinal mode a scalar product and a “back projection” has to be computed, both requiring N additions and N multiplications, where N is the number of string elements. The accuracy/complexity of the synthesis may be varied by changing the number of longitudinal modes.

As an example, the transverse and longitudinal bridge force computed by the bidirectional tension decomposition model is presented in Fig. 6.8 by a solid line, with $N = 100$ elements in the finite-difference string model and $K = 10$ longitudinal modes. The dotted line shows the output of a full finite-difference model of Sec. 6.4.1 (with $N = 100$) running at ten times higher sampling rate (441 kHz) as a reference. The difference of the outputs is displayed by a dashed line. Note that this difference is mainly because the numerical dispersion of the transverse string model is different at various sampling rates, which also influences the longitudinal vibration due to the coupling of the two polarizations. When also the tension decomposition model is running at 441 kHz, this difference vanishes, as can be seen in Fig. 6.9. The small error signal present in Fig. 6.9 (b) is most probably due to the fact that the longitudinal modes of the full finite-difference model are affected by numerical dispersion, while they are not in the tension decomposition model, since they are computed by a modal-based approach. Naturally, computing the tension decomposition model at a high sampling rate makes no sense in practice, as in that case it would have the same computational complexity as for the finite-difference model of Sec. 6.4.1. Luckily, it is not required, as the difference between the models of Fig. 6.8 (where the tension decomposition model is running at 44.1 kHz) is inaudible. This is because the slight frequency shift of the higher modes due to the different numerical dispersion does not affect sound quality. The perceptual effect is similar to having a less precise inharmonicity modeling, which have been investigated in [Rocchesso and Scalcon 1999].

To sum up, the tension decomposition model is capable of the same sound quality as the full finite-difference model of Sec. 6.4.1 at about 20% of its computational cost. Moreover, it has the advantage that the computational complexity/quality can be varied by changing the number of realized longitudinal modes. Although the model was found to be numerically stable in practice, the stability analysis of the method would be an interesting field of future research. (Note that this can be very difficult, as traditional stability analysis techniques developed for linear models cannot be applied.) Naturally, the idea of tension decomposition could be used for modal-based approaches and for digital waveguide models, too. Modeling the longitudinal modes and its effect on tension variation can also be applied as an add-on to tension modulation string models of Sec. 6.2, where the spatially uniform part of the tension modulation is already modeled.

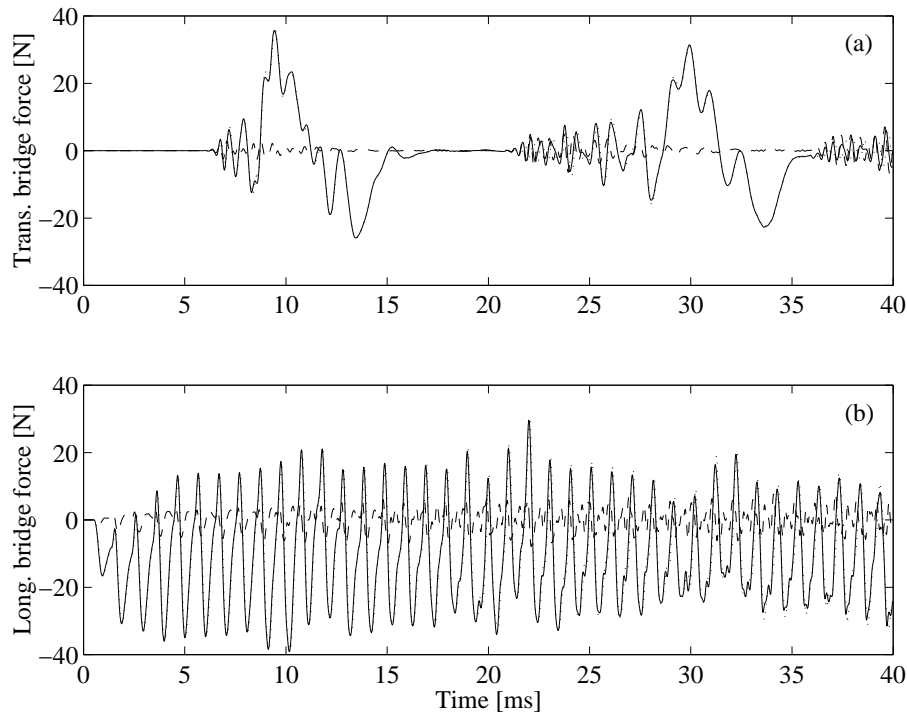


Figure 6.8: The transverse (a) and longitudinal (b) bridge forces of a simulated G_1 piano tone. The solid line displays the output of the model based on tension decomposition running at 44.1 kHz sample rate, while the dotted line presents the output of a finite-difference model running at 441 kHz. The difference of the outputs is displayed by a dashed line.

6.5 Conclusion

In this chapter sound synthesis methods have been presented based on the theory of Chap. 5. Sections 6.1 and 6.2 have overviewed the models that apply the uniform tension approximation.

Section 6.3 proposed many new techniques for modeling the vibration of longitudinal modes. This corresponds to the situation where the tension is spatially nonuniform, but the longitudinal to transverse coupling is negligible. The general idea of the methods is that the motion of longitudinal modes is computed by a nonlinearly excited modal model. This is advantageous because generally a low number of longitudinal modes has to be simulated, and in this case the modal-based approach is the most efficient one. The various models mainly differ in what kind of transverse string model they apply and which simplifications are made in computing the longitudinal response. First, a physically accurate model have been proposed in Sec. 6.3.3 that is based on the decomposition of the tension into a spatially uniform and a space-dependent part, decreasing the computational complexity by an order of magnitude compared to a full finite-difference model. The efficiency have been further increased by the “composite” and “modal” models of Secs. 6.3.4 and 6.3.5 by computing a common excitation signal for the longitudinal modes. These models have been

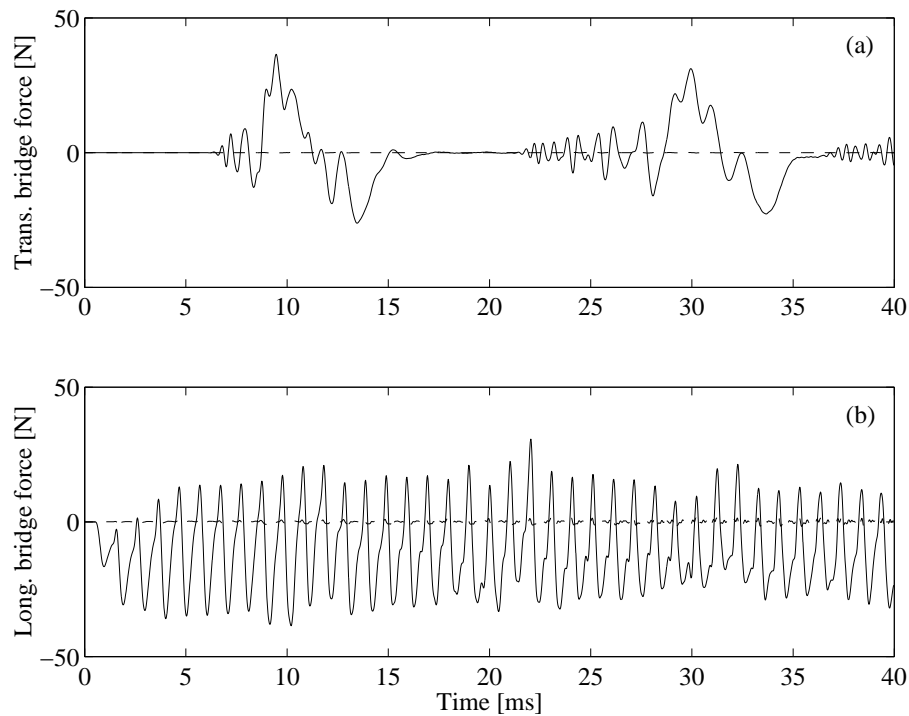


Figure 6.9: The transverse (a) and longitudinal (b) bridge forces of a simulated G_1 piano tone. The line convention and the parameters are the same as for Fig. 6.8. The only difference is that now both models run at the same sample rate (441 kHz).

published in [Bank and Sujbert 2004, 2005b]. Section 6.3.6 investigated the complications with inharmonic digital waveguides in the context of longitudinal vibration modeling and proposed the solution of distributing the dispersion filter. Physically informed models have been proposed in Sec. 6.3.7, where the connection to physical reality is less tight resulting in two orders of magnitude lower computational cost when compared to the full finite-difference method. A linearized and commuted version of the model has also been proposed. All the methods of Sec. 6.3 increase the sound quality of synthesized piano tones significantly compared to earlier models neglecting the longitudinal vibrations. The choice between them mainly depends on whether the physically meaningful behavior or the low complexity is of highest importance. Although the longitudinal components are the most significant in the case of the piano, the models could be also used for increasing the reality of other synthesis models (guitar, harpsichord, etc.)

A synthesis model implementing the bidirectional coupling of transverse and longitudinal polarizations has been presented in Sec. 6.4 [published in Bank and Sujbert 2005a]. The model is capable of producing a sound indistinguishable from a full finite-difference model at around 20% of its computational cost. The good agreement of the new model with the full finite-difference model confirms the theoretical results of Sec. 5.3, which form the basis of the model. This model can be used for such instruments, where both the effects of tension modulation and longitudinal modes are audible (they are not too many).

It can also be used for producing musically interesting effects, such as exciting a piano string with a much higher amplitude, leading to a pitch glide. Examples of synthesized sounds can be found at [Sound examples n.d.].

As a straightforward extension to the models proposed in Secs. 6.3 and 6.4, two transverse polarizations could be implemented along the lines of the equations presented in Sec. 5.3.6. This requires the addition of a second transverse string model. Moreover, the excitation force density has to be computed as the last term of Eq. (5.57) instead of Eq. (5.28). From that on, the approach is exactly the same as presented in these sections. Another possible development is the implementation of frequency dependent termination impedance. This would be most advantageous for models including two transverse polarizations, as this would reproduce the coupling of the y and z polarizations leading to beating and two-stage decay. We have not dealt with these issues here, as they are orthogonal to the proposed ideas. That is, the longitudinal modeling part does not have to be altered when these extensions are implemented.

Chapter 7

Summary

In this thesis new methods have been presented for the physics-based synthesis of string instruments. Chapter 3 proposed new algorithms for the loss filter design of digital waveguides. Accordingly, the results can be used to improve the parameter estimation of such instrument models, that employ the digital waveguide approach. In Chap. 4 the multi-rate approach have been utilized for increasing the efficiency of the different parts of the instrument model. The multi-rate excitation model of Sec. 4.1 provides a simple method to overcome the numerical stability problems arising from noncomputable loops, which can be used for struck or plucked strings. The parallel resonator bank of Sec. 4.2 decreases the computational complexity of modeling beating and two-stage decay significantly and allows simpler parameter estimation compared to earlier methods. As this phenomenon is prominent in impulsively excited strings, the method is advantageous for plucked and struck instruments. The multi-rate body model presented in Sec. 4.3 provides significant savings in terms of computational complexity and can be used for all instruments, where the sound is generated by the radiation of the body (i.e., all acoustic string instruments including bowed strings).

As we have seen, Chap. 3 and 4 provided either new parameter estimation techniques for existing structures or efficient implementation of well known physical phenomena. On the other hand, for modeling the geometric nonlinearities of strings, no such precise physical models existed that could be used for developing sound synthesis algorithms. Therefore, Chap. 5 provided a new theoretical framework that besides forming the basis of the sound synthesis algorithms of Chap. 6, gives an insight to the phenomenon and explains the measurement results of other authors. Thus, these results are not only useful for the sound synthesis community but for the acousticians and instrument makers, too. Chapter 6 presented the sound synthesis applications of the theoretical model of Chap. 5. For increasing efficiency, simplifications were carried out in a way that the effects of smaller perceptual significance were neglected. The proposed methods can be used to extend various types of existing linear string models making them capable of modeling the effects of geometric nonlinearities. Naturally, modeling the longitudinal vibrations is advantageous for those instruments, where the effect is perceptually significant. Including them in the model greatly improves the quality of synthesized piano sounds, and it is very probable that models of

other string instruments would also benefit from refined modeling.

7.1 Further Research

Measurement methods that could provide more reliable estimates of the instrument body transfer functions could help the accuracy of body modeling. Alternatively, the body impulse response could be computed by a precise physical model. This would mean the combination of the physics-based and post-processing approaches, where the physically computed body impulse response is implemented by an efficient post-processing approach. That would provide the opportunity to the user to modify the geometric and material properties of the instrument body and then play with the new instrument in real-time.

Listening tests should be conducted to assess the importance of the effects of geometric nonlinearities for those instruments where the phenomenon is less prominent than for the piano (guitar, clavichord, harpsichord). For bowed strings, the longitudinal polarization is excited by the bow, too, and not only by the transverse vibration. This should be included in the model, together with the torsional polarization, which has even stronger effect in the case of bowed strings than that of the longitudinal vibration.

It would be beneficial for all the modeling methods described in the thesis to find which are those parameters that give an agreeable sound quality for a given instrument. This problem can be addressed by analyzing the sound of various instruments and trying to find the difference between the ones with better or worse quality. Alternatively, physical models could be used to generate the input for listening tests, where the different features could be varied independently. This set of rules could not only help in parameterizing sound synthesis algorithms but could also be beneficial for instrument makers.

7.2 New Scientific Results

This section summarizes the main results of the thesis in the form of scientific statements.

Statement 1: I have developed new methods for decay time-based design of loss-filters for digital waveguides.

1.1: I have proposed a polynomial regression-based method for designing one-pole loss filters, which is the most common type of loss filters. I have derived a formula for the decay time of a digital waveguide using a one-pole loss filter, and I have established a relationship between the parameters for the one-pole filter and the differential equation of the lossy string.

1.2: I have developed a robust and simple method for designing higher-order loss filters, which minimizes the decay time error through a magnitude-dependent weighting function. The weighting function is derived from the first-order Taylor series approximation for the decay time as a function of filter magnitude.

Statement 2: I have suggested the application of multi-rate techniques for increasing the efficiency of string instrument models.

2.1: I have proposed a new method for maintaining numerical stability within the excitation model. According to the method, the excitation model should operate at a higher (e.g., double) sampling rate than the rest of the instrument model.

2.2: I have shown that the beating and two-stage decay effects can be efficiently modeled by running a few resonators in parallel with the basic string model (e.g., a digital waveguide). The method models the phenomenon only for those partials that are dominated by the effect. The resonators run at a sampling rate lower than that of the string model, which results in considerably lower computational complexity than methods developed earlier.

2.3: I have proposed the multi-rate approach for modeling the force-pressure transfer function of the instrument body. In the lower part of the audible frequency range the body is modeled by a high-order filter running at one fourth or one eighth of the sampling frequency, while for high frequencies a low-order filter approximates the body response. The method requires significantly lower computing power compared to traditional filters, while the degradation in sound quality is marginal.

Statement 3: I have developed a comprehensive model for the nonlinear vibration of metal strings that can be efficiently used for sound synthesis. The model takes into consideration the coupling of transverse and longitudinal polarizations.

3.1: I have introduced a classification for the nonlinear behavior of strings, which estimates from the physical parameters of the string and from the amplitude and frequency content of the transverse vibration which phenomena dominate the vibration. This “nonlinearity map” clearly shows the similarities and differences between the various cases.

3.2: I have determined the closed solution for the nonlinear differential equation of the string for the case where the tension on the string is spatially nonuniform, but the variation of tension has a negligible effect on the transverse vibration. This approximation is valid for highly stretched metal strings used in most string instruments. The proposed modal model describes the free vibration of longitudinal modes and the generation of phantom partials jointly.

3.3: I have derived a relationship between the modal model and the uniform tension approximation by decomposing the tension into a spatially uniform and a space-dependent part. I have shown that the spatially uniform tension approximation, which is the most widely used approach, is a special case of the proposed modal model.

Statement 4: I have extended the most common types of physical string models by making them capable of modeling the longitudinal vibrations, too.

4.1: I have proposed various composite string models where the response for longitudinal modes is calculated by nonlinearly excited second-order resonators. For computing the transverse vibration, both the finite-difference and modal-based approaches are appropriate. The method is capable of modeling the bidirectional coupling of transverse and longitudinal polarizations. The proposed models require significantly lower computational cost than the techniques that compute both polarizations through a finite-difference scheme.

4.2: I have shown that dispersive digital waveguides cannot be used for computing the excitation force of the longitudinal polarization in their original form. The problem can be avoided through distributing the dispersion filter, at the expense of increased computational complexity.

4.3: I have proposed "physically informed" techniques for modeling the longitudinal vibration, which are even more efficient than the above proposed ones. These use a physics-based transverse string model extended by a signal model whose parameters are computed from the physical parameters of the string.

Bibliography

- Adrien, J.-M. (1991). The missing link: Modal synthesis, *in* G. De Poli, A. Piccialli and C. Roads (eds), *Representations of Musical Signals*, The MIT Press, Cambridge, Massachusetts, USA, pp. 269–297.
- Anand, G. V. (1969). Large-amplitude damped free vibration of a stretched string, *J. Acoust. Soc. Am.* **45**(5): 1089–1096.
- Aramaki, M., Bensa, J., Daudet, L., Guillemain, P. and Kronland-Martinet, R. (2002). Resynthesis of coupled piano string vibrations based on physical modeling, *Journal of New Music Research* **30**(3): 213–226.
- Arfib, D. (1979). Digital synthesis of complex spectra by means of multiplication of non-linear distorted sine waves, *J. Audio Eng. Soc.* **27**(10): 757–768.
- Bader, R. (2003). Physical model of a complete classical guitar body, *Proc. Stockholm Music Acoust. Conf.*, Stockholm, Sweden, pp. 121–124.
- Bank, B. (2000a). Nonlinear interaction in the digital waveguide with the application to piano sound synthesis, *Proc. Int. Computer Music Conf.*, Berlin, Germany, pp. 54–58.
- Bank, B. (2000b). *Physics-based sound synthesis of the piano*, Master’s thesis, Budapest University of Technology and Economics, Hungary. Published as Report 54 of HUT Laboratory of Acoustics and Audio Signal Processing, URL: <http://www.mit.bme.hu/~bank/thesis>.
- Bank, B. (2001). Accurate and efficient modeling of beating and two-stage decay for string instrument synthesis, *Proc. MOSART Workshop on Curr. Res. Dir. in Computer Music*, Barcelona, Spain, pp. 134–137.
- Bank, B. and Sujbert, L. (2003). Modeling the longitudinal vibration of piano strings, *Proc. Stockholm Music Acoust. Conf.*, Stockholm, Sweden, pp. 143–146.
- Bank, B. and Sujbert, L. (2004). A piano model including longitudinal string vibrations, *Proc. Conf. on Digital Audio Effects*, Naples, Italy, pp. 89–94.
- Bank, B. and Sujbert, L. (2005a). Efficient modeling strategies for the geometric nonlinearities of musical instrument strings, *Proc. Forum Acusticum 2005*, Budapest, Hungary.

- Bank, B. and Sujbert, L. (2005b). Generation of longitudinal vibrations in piano strings: From physics to sound synthesis, *J. Acoust. Soc. Am.* **117**(4): 2268–2278.
- Bank, B. and Välimäki, V. (2003). Robust loss filter design for digital waveguide synthesis of string tones, *IEEE Sign. Proc. Letters* **10**(1): 18–20.
- Bank, B., Avanzini, F., Borin, G., De Poli, G., Fontana, F. and Rocchesso, D. (2003). Physically informed signal-processing methods for piano sound synthesis: a research overview, *EURASIP J. on Appl. Sign. Proc.* **2003**(10): 941–952.
- Bank, B., De Poli, G. and Sujbert, L. (2002). A multi-rate approach to instrument body modeling for real-time sound synthesis applications, *Proc. 112th AES Conv., Preprint No. 5526*, Munich, Germany.
- Bank, B., Välimäki, V., Sujbert, L. and Karjalainen, M. (2000). Efficient physics-based sound synthesis of the piano using DSP methods, *Proc. 10th Eur. Sign. Proc. Conf.*, Tampere, Finland, pp. 2225–2228.
- Bensa, J. (2003). *Analysis and synthesis of piano sounds using physical and signal models*, PhD thesis, Université de la Méditerranée, Marseille, France. URL: <http://www.lma.cnrs-mrs.fr/~bensa>.
- Bensa, J. and Daudet, L. (2004). Efficient modeling of "phantom" partials in piano tones, *Int. Symp. on Musical Acoust.*, Nara, Japan, pp. 207–210.
- Bensa, J., Bilbao, S., Kronland-Martinet, R. and Smith, J. O. (2003). The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides, *J. Acoust. Soc. Am.* **114**(2): 1095–1107.
- Bensa, J., Jensen, K. and Kronland-Martinet, R. (2004). A hybrid resynthesis model for hammer–string interaction of piano tones, *EURASIP J. on Appl. Sign. Proc.* **2004**(7): 1021–1035.
- Bilbao, S. (2004a). Energy-conserving finite difference schemes for tension-modulated strings, *Proc. IEEE Int. Conf. Acoust., Speech, and Sign. Proc.*, Montreal, Canada, pp. 285–288.
- Bilbao, S. (2004b). Modal-type synthesis techniques for nonlinear strings with an energy conserving property, *Proc. Conf. on Digital Audio Effects*, Naples, Italy, pp. 119–124.
- Borin, G. (2001). Personal communication.
- Borin, G. and De Poli, G. (1996). A hysteretic hammer–string interaction model for physical model synthesis, *Proc. Nordic Acoust. Meeting*, Helsinki, Finland, pp. 399–406.
- Borin, G., De Poli, G. and Rocchesso, D. (2000). Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems, *IEEE Trans. Acoust., Speech, and Sign. Proc.* **8**(5): 597–606.

- 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, *Proc. Int. Computer Music Conf.*, Thessaloniki, Greece, pp. 350–353.
- Boutillon, X. (1988). Model for piano hammers: Experimental determination and digital simulation, *J. Acoust. Soc. Am.* **83**(2): 746–754.
- Caramaschi, E. (2004). *Un modello fisico del pianoforte per la sintesi del suono*, Master’s thesis, University of Verona, Italy.
- Chaigne, A. and Askenfelt, A. (1994). Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods, *J. Acoust. Soc. Am.* **95**(2): 1112–1118.
- Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation, *J. Audio Eng. Soc.* **21**(7): 526–534.
- Conklin, H. A. (1996). Design and tone in the mechanoacoustic piano. Part III. Piano strings and scale design, *J. Acoust. Soc. Am.* **100**(3): 1286–1298.
- Conklin, H. A. (1999). Generation of partials due to nonlinear mixing in a stringed instrument, *J. Acoust. Soc. Am.* **105**(1): 536–545.
- Cuzzucoli, G. and Lombardo, V. (1999). A physical model of the classical guitar, including the player’s touch, *Computer Music J.* **23**(2): 52–69.
- Erkut, C. (2001). Loop filter design techniques for virtual string instruments, *Int. Symp. on Musical Acoust.*, Perugia, pp. 259–262.
- Erkut, C., Karjalainen, M., Huang, P. and Välimäki, V. (2002). Acoustical analysis and model-based sound synthesis of the kantele, *J. Acoust. Soc. Am.* **112**(4): 1681–1691.
- Erkut, C., Välimäki, V., Karjalainen, M. and Laurson, M. (2000). Extraction of physical and expressive parameters for model-based sound synthesis of the classical guitar, *Proc. 108th AES conv., Preprint No. 5114*, Paris, France.
- Fletcher, N. H. and Rossing, T. D. (1998). *The Physics of Musical Instruments*, Springer-Verlag, New York, USA. 2nd ed. (1st ed. 1991).
- Garnett, G. E. (1987). Modeling piano sound using digital waveguide filtering techniques, *Proc. Int. Computer Music Conf.*, Urbana, Illinois, USA, pp. 89–95.
- Giordano, N. and Jiang, M. (2004). Physical modeling of the piano, *EURASIP J. on Appl. Sign. Proc.* **2004**(7): 926–933.

- Giordano, N. and Korty, A. J. (1996). Motion of a piano string: Longitudinal vibrations and the role of the bridge, *J. Acoust. Soc. Am.* **100**(6): 3899–3908.
- Giordano, N. and Winans II, J. P. (1999). Plucked strings and the harpsichord, *J. Sound and Vib.* **224**(3): 455–473.
- Giordano, N. and Winans II, J. P. (2000). Piano hammers and their compression characteristics: Does a power law make sense?, *J. Acoust. Soc. Am.* **107**(4): 2248–2255.
- Gough, C. (1984). The nonlinear free vibration of a damped elastic string, *J. Acoust. Soc. Am.* **75**(6): 1770–1776.
- Graff, K. F. (1975). *Wave Motion in Elastic Solids*, Oxford University Press, London, England.
- Hanson, R. J., Anderson, J. M. and Macomber, H. K. (1994). Measurements of nonlinear effects in a driven vibrating wire, *J. Acoust. Soc. Am.* **96**(3): 1549–1556.
- Hiller, L. and Ruiz, P. (1971a). Synthesizing musical sounds by solving the wave equation for vibrating objects: Part 1, *J. Audio Eng. Soc.* **19**(6): 462–470.
- Hiller, L. and Ruiz, P. (1971b). Synthesizing musical sounds by solving the wave equation for vibrating objects: Part 2, *J. Audio Eng. Soc.* **19**(7): 542–550.
- Jaffe, D. A. and Smith, J. O. (1983). Extensions of the Karplus-Strong plucked-string algorithm, *Computer Music J.* **7**(2): 56–69.
- Järveläinen, H. and Tolonen, T. (2001). Perceptual tolerances for decay parameters in plucked string synthesis, *J. Audio Eng. Soc.* **49**(11): 1049–1059.
- Järveläinen, H., Välimäki, V. and Karjalainen, M. (2001). Audibility of the timbral effects of inharmonicity in stringed instrument tones, *Acoustic Research Letters Online* **2**(3): 79–84. available at <http://ojps.aip.org/ARLO>.
- Karjalainen, M. and Erkut, C. (2004). Digital waveguides versus finite difference structures: Equivalence and mixed modeling, *EURASIP J. on Appl. Sign. Proc.* **2004**(7): 978–989.
- Karjalainen, M. and Smith, J. O. (1996). Body modeling techniques for string instrument synthesis, *Proc. Int. Computer Music Conf.*, Hong Kong, pp. 232–239.
- Karjalainen, M. and Välimäki, V. (1993). Model-based analysis/synthesis of the acoustic guitar, *Proc. Stockholm Music Acoustics Conf.*, Stockholm, Sweden, pp. 443–447.
- Karjalainen, M., Backman, J. and Pölkki, J. (1993). Analysis, modeling, and real-time sound synthesis of the kantele, a traditional finnish string instrument, *Proc. IEEE Int. Conf. Acoust., Speech, and Sign. Proc.*, Vol. 1, Minneapolis, , MN, USA, pp. 229–232.
- Karjalainen, M., Esquef, P. A. A., Ansalo, P., Mäkivirta, A. and Välimäki, V. (2002). Frequency-zooming ARMA modeling of resonant and reverberant systems, *J. Audio Eng. Soc.* **50**(12): 1012–1029.

- Karjalainen, M., Välimäki, V. and Tolonen, T. (1998). Plucked-string models: from Karplus-Strong algorithm to digital waveguides and beyond, *Computer Music J.* **22**(3): 17–32.
- Karjalainen, M., Välimäki, V., Räsänen, H. and Saastamoinen, H. (1999). DSP equalization of electret film pickup for acoustic guitar, *Proc. 106th AES Conv., Preprint No. 4907*, Munich, Germany.
- Karplus, K. and Strong, A. (1983). Digital synthesis of plucked-string and drum timbres, *Computer Music J.* **7**(2): 43–55.
- Kurmyshev, E. V. (2003). Transverse and longitudinal mode coupling in a free vibrating soft string, *Phys. Lett. A* **310**(2–3): 148–160.
- Laakso, T. I., Välimäki, V., Karjalainen, M. and Laine, U. K. (1996). Splitting the unit delay – tools for fractional delay filter design, *IEEE Sign. Proc. Mag.* **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 Trans. Circ. and Syst.–II: Analog and Digital Sign. Proc.* **41**(1): 40–48.
- Laroche, J. and Jot, J.-M. (1992). Analysis/synthesis of quasi-harmonic sounds by use of the Karplus-Strong algorithm, *Proc. 2nd French Congr. on Acoust.*, France.
- Le Brun, M. (1979). Digital waveshaping synthesis, *J. Audio Eng. Soc.* **27**(4): 250–266.
- Leamy, M. J. and Gottlieb, O. (2000). Internal resonances in whirling strings involving longitudinal dynamics and material non-linearities, *J. Sound and Vib.* **236**(4): 683–703.
- Legge, K. A. and Fletcher, N. H. (1984). Nonlinear generation of missing modes on a vibrating string, *J. Acoust. Soc. Am.* **76**(1): 5–12.
- Lehtonen, H.-M., Rauhala, J. and Välimäki, V. (2005). Sparse multi-stage loss filter design for waveguide piano synthesis, *Proc. IEEE Workshop Appl. of Sign. Proc. to Audio and Acoust.*, New Paltz, NY, USA, pp. 331–334.
- Leissa, A. W. and Saad, A. M. (1994). Large amplitude vibrations of strings, *Journal of Applied Mechanics* **61**: 296–301.
- Mathworks (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, *J. Acoust. Soc. Am.* **74**(5): 1325–1345.
- Morse, P. M. (1948). *Vibration and Sound*, McGraw-Hill. Reprint, (1st ed. 1936).
- Morse, P. M. and Ingard, K. U. (1968). *Theoretical Acoustics*, McGraw-Hill.
- Nakamura, I. and Naganuma, D. (1993). Characteristics of piano sound spectra, *Proc. Stockholm Music Acoust. Conf.*, pp. 325–330.

- Narasimha, R. (1968). Nonlinear vibration of an elastic string, *J. Sound and Vib.* **8**(1): 134–146.
- Oplinger, D. W. (1960). Frequency response of a nonlinear stretched string, *J. Acoust. Soc. Am.* **32**(12): 1529–1538.
- Oppenheim, A. V. and Schaffer, R. W. (1975). *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, USA.
- Pakarinen, J., Karjalainen, M. and Välimäki, V. (2003). Modeling and real-time synthesis of the kantele using distributed tension modulation, *Proc. Stockholm Music Acoust. Conf.*, Stockholm, Sweden, pp. 409–412.
- Pakarinen, J., Karjalainen, M., Välimäki, V. and Bilbao, S. (2005a). Energy behavior in time-varying fractional delay filters for physical modeling of musical instruments, *Proc. IEEE Int. Conf. Acoust., Speech, and Sign. Proc.*, Philadelphia, PA, USA, pp. 1–4.
- Pakarinen, J., Välimäki, V. and Karjalainen, M. (2005b). Physics-based methods for modeling nonlinear vibrating strings, *Acta Acust. – Acust.* **91**(2): 312–325.
- Rauhala, J., Lehtonen, H.-M. and Välimäki, V. (2005). Multi-ripple loss filter for waveguide piano synthesis, *Proc. Int. Computer Music Conf.*, Barcelona, Spain, pp. 729–732.
- Roads, C. (1995). *The Computer Music Tutorial*, The MIT Press, Cambridge, Massachusetts, USA.
- Rocchesso, D. and Scalcon, F. (1996). Accurate dispersion simulation for piano strings, *Proc. Nordic Acoust. Meeting*, Helsinki, Finland, pp. 407–414.
- Rocchesso, D. and Scalcon, F. (1999). Bandwidth of perceived inharmonicity for physical modeling of dispersive strings, *IEEE Trans. Speech Audio Proc.* **7**(5): 597–601.
- Rowland, D. R. and Pask, C. (1999). The missing wave momentum mystery, *American Journal of Physics* **67**(5): 378–388.
- Schafer, R. W. and Rabiner, L. R. (1973). A digital signal processing approach to interpolation, *Proceedings of the IEEE* **61**(6): 692–702.
- Scheier, E. D. (1999). Structured Audio and Effects Processing in the MPEG-4 Multimedia Standard, *Multimedia Systems* **7**: 11–22.
- Serafin, S. (2004). *The Sound of Friction: Real-time Models, Playability and Musical Applications*, PhD thesis, Stanford University, California, USA. URL: <http://www.imi.aau.dk/~sts/serafinthesis.pdf>.
- Serra, X. and Smith, J. O. (1990). Spectral modeling synthesis: a sound analysis/synthesis system based on deterministic plus stochastic decomposition, *Computer Music J.* **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.
- Smith, J. O. (1986). Efficient simulation of the reed-bore and bow-string mechanisms, *Proc. Int. Computer Music Conf.*, The Hague, the Netherlands, pp. 275–280.
- Smith, J. O. (1991). Viewpoints on the history of digital synthesis, *Proc. Int. Computer Music Conf.*, Montreal, Canada, pp. 1–10.
- Smith, J. O. (1992). Physical modeling using digital waveguides, *Computer Music J.* **16**(4): 74–91. URL: <http://ccrma.stanford.edu/~jos/wg.html>.
- Smith, J. O. (1993). Efficient synthesis of stringed musical instruments, *Proc. Int. Computer Music Conf.*, Tokyo, Japan, pp. 64–71.
- Smith, J. O. (1997). Nonlinear commuted synthesis of bowed strings, *Proc. Int. Computer Music Conf.*, Thessaloniki, Greece, pp. 264–267.
- Smith, J. O. (2005). *Physical Audio Signal Processing for Virtual Musical Instruments and Audio Effects*, Center for Computer Research in Music and Acoustics, Stanford University, W3K Publishing, USA. Draft, URL: <http://ccrma.stanford.edu/~jos/pasp>.
- Sound examples (n.d.). Recorded and synthetic sound examples are available at <http://www.mit.bme.hu/~bank/phd> and on the attached CD-ROM.
- Stopper, B. (2003). Minimens 1.0 audio piano string simulator. URL: <http://www.piano-stopper.de/homepe.htm>.
- Stulov, A. (1995). Hysteretic model of the grand piano felt, *J. Acoust. Soc. Am.* **97**(4): 2577–2585.
- Tolonen, T. (1998). *Model-based analysis and resynthesis of acoustic guitar tones*, Master's thesis, Helsinki University of Technology, Espoo, Finland. Report 46, HUT Laboratory of Acoustics and Audio Signal Processing, URL: <http://www.acoustics.hut.fi/~ttolonen/thesis.html>.
- Tolonen, T. and Järveläinen, H. (2000). Perceptual study of decay parameters in plucked string synthesis, *Proc. 109th AES Conv., Preprint No. 5205*, Los Angeles, USA.
- 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. URL: http://www.acoustics.hut.fi/~ttolonen/sound_synth_report.html.
- Tolonen, T., Välimäki, V. and Karjalainen, M. (2000). Modeling of tension modulation nonlinearity in plucked strings, *IEEE Trans. Speech Audio Proc.* **8**(3): 300–310.

- Trautmann, L. and Rabenstein, R. (1999). Digital sound synthesis based on transfer function models, *Proc. IEEE Workshop Appl. of Sign. Proc. to Audio and Acoust.*, New Paltz, NY, USA, pp. 83–86.
- Trautmann, L. and Rabenstein, R. (2000). Sound synthesis with tension modulated nonlinearities based on functional transformations, *Proc. Acoust. and Music: Theory and Applications*, Montego Bay, Jamaica, pp. 444–449.
- Trautmann, L. and Rabenstein, R. (2003). *Digital Sound Synthesis by Physical Modeling using the Functional Transformation Method*, Kluwer Academic/Plenum Publishers, New York, USA.
- Van Duyne, S. A. and Smith, J. O. (1994). A simplified approach to modeling dispersion caused by stiffness in strings and plates, *Proc. Int. Computer Music Conf.*, Århus, Denmark, pp. 407–410.
- 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, *Proc. Int. Computer Music Conf.*, Århus, Denmark, pp. 411–418.
- Välimäki, V. (1995). *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*, PhD thesis, Helsinki University of Technology, Espoo, Finland. URL: http://www.acoustics.hut.fi/~vpv/publications/vesa_phd.html.
- Välimäki, V. and Tolonen, T. (1998). Development and calibration of a guitar synthesizer, *J. Audio Eng. Soc.* **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, *J. Audio Eng. Soc.* **44**(5): 331–353.
- Välimäki, V., Pakarinen, J., Erkut, C. and Karjalainen, M. (2006). Discrete-time modelling of musical instruments, *Reports on Progress in Physics* **69**(1): 1–78.
- Välimäki, V., Penttinen, H., Knif, J., Laurson, M. and Erkut, C. (2004). Sound synthesis of the harpsichord using a computationally efficient physical model, *EURASIP J. on Appl. Sign. Proc.* **2004**(7): 934–948.
- Watzky, A. (1992). Non-linear three-dimensional large-amplitude damped free vibration of a stiff elastic stretched string, *J. Sound and Vib.* **153**(1): 125–142.
- Weinreich, G. (1977). Coupled piano strings, *J. Acoust. Soc. Am.* **62**(6): 1474–1484.
- Woodhouse, J. (2004a). On the synthesis of guitar plucks, *Acta Acust. – Acust.* **90**(5): 928–944.
- Woodhouse, J. (2004b). Plucked guitar transients: Comparison of measurements and synthesis, *Acta Acust. – Acust.* **90**(5): 945–965.

Appendix

A.1 Parameter Dependence of Nonlinear Components in String Motion

Here we outline the derivations of Eqs. (5.6)–(5.8) included in Sec. 5.1. We assume that only the transverse polarization is excited by the excitation, while the longitudinal polarization gains energy from nonlinear coupling. Therefore, the amplitude dependence of the nonlinear components is written as a function of transverse slope $\partial y/\partial x$.

From Eq. (5.5) it follows that the linear transverse component (the component which would arise if the string was ideal) of the bridge force $F_{t,\text{lin}}$ has the magnitude

$$\|F_{t,\text{lin}}\| = T_0 \left\| \frac{\partial y}{\partial x} \right\|, \quad (\text{A.1})$$

where the Euclidean norm (root mean square value) of the transverse slope at the termination ($x = L$) is referred as $\|\partial y/\partial x\|$.

Equation (5.2) shows that the magnitude of longitudinal vibration depends on the transverse slope according to a square law. The force at the bridge (see Eq. 5.4) has the same behavior, as it is the sum of a term linearly depending on $\partial\xi/\partial x$ (which is a second order function of $\partial y/\partial x$) and a term having square law dependence on $\partial y/\partial x$. The magnitude of longitudinal vibration does not depend on ES , as both the nonlinear excitation force and the restoring force are proportional to ES in Eq. (5.2). However, the appearance of the longitudinal string motion in the bridge spectra has a linear dependence on ES , as can be seen in Eq. (5.4). Thus, the magnitude of longitudinal force at the bridge $\|F_1\|$ is approximately described by

$$\|F_1\| \approx C_1 ES \left\| \frac{\partial y}{\partial x} \right\|^2, \quad (\text{A.2})$$

where C_1 is a constant in the order of unity, which depends on the type of string excitation. Note that we have assumed that the longitudinal motion is so small that it has no influence on the transverse vibration.

In Eq. (5.3) the first term on the right-hand side is the dominant restoring force, leading to the well known string motion. The next term is a nonlinear forcing term, which adds some new components to the transverse vibration. The magnitude of this force is a third order function of the transverse slope, as the longitudinal slope $\partial\xi/\partial x$ is the second-order

function of $\partial y/\partial x$. As for the force on the bridge, these excited nonlinear components appear through the first term $-T_0\partial y/\partial x$ (already filtered by the string dynamics) in Eq. (5.5) and directly through the next term, which is again a third order function of $\partial y/\partial x$. The significance of the nonlinear forcing term increases as a linear function of ES in Eq. (5.3). This nonlinear component appears as a linear function of ES in the transverse bridge force, too. As a result, the magnitude of the nonlinear transverse component can be approximated as a third order function of transverse slope:

$$\|F_{t,\text{nonlin}}\| \approx C_t ES \left\| \frac{\partial y}{\partial x} \right\|^3, \quad (\text{A.3})$$

where C_t is a constant in the order of unity. Again, we have assumed that the nonlinear transverse component is so small that it can be neglected in comparison with the standard linear transverse component.

Figure A.1 shows the Euclidean norm of bridge forces for the first 100 ms of simulated struck piano strings, computed by the nonlinear string model of Sec. 6.4. Figure A.1 (a) displays a string with the physical parameters μ , T_0 , E , S , and L corresponding to a G_1 piano string. Losses and dispersion are also included in the simulation. The dotted lines show the approximate curves computed by Eqs. (A.2) and (A.3) with $C_1 = 0.25$ and $C_t = 10$. These C_1 and C_t values have been found to be also acceptable approximations for other kind of excitations, such as plucking. The thick solid line shows the Euclidean norm of the linear transverse bridge force. The magnitude of nonlinear transverse component (thin solid line) is computed by subtracting the output of a linear string model from the output of the nonlinear model. This component does not necessarily mean “new” components in the spectrum, as it might correspond to the amplitude and frequency change of the already present transverse modes. A good example for this is when the string tension is increased due to the geometric nonlinearity, the transverse modal frequencies are also increased. Finally, the dashed line displays the Euclidean norm of the longitudinal bridge force.

Figure A.1 (b) shows a simulation with the same parameters, except that the Young’s modulus is increased by a factor of 100, corresponding to a loosely stretched string. It can be seen that now the magnitude of the longitudinal component and that of the nonlinear transverse component reach the level of the transverse component at a much lower transverse slope compared to Fig. A.1 (a).

Figure A.1 (a) and (b) demonstrate that the approximate curves follow the simulated ones until the nonlinear transverse component reaches the level of the linear transverse component. The reason for this is that the generation of the longitudinal motion draws energy from the transverse vibration (which is not included in Eqs. (A.2) and (A.3)), and this energy leakage from the transverse motion starts to be significant only above a certain level.

In the next simulation, all the parameters of the modeled piano string were fixed, but the Young’s modulus E was changed continuously. This corresponds to comparing strings with the same fundamental frequency but being of different material. The value 10^0 of the x axis corresponds to the E value of a G_1 piano string with an (ES/T_0) ratio of 400. The hammer excitation was set in a way that it gave $\|\partial y/\partial x\| = 10^{-3}$ in all the cases. The

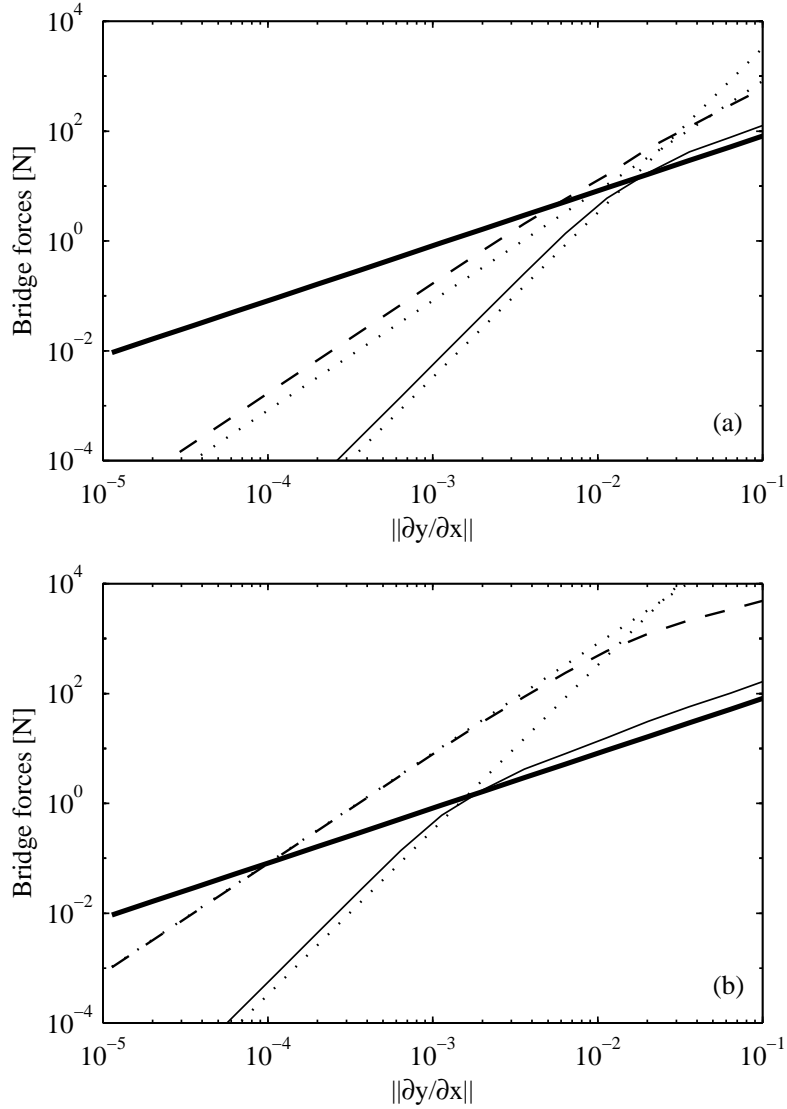


Figure A.1: Euclidean norm of simulated bridge forces as a function of the Euclidean norm of the transverse slope $\|\partial y/\partial x\|$ at the bridge: linear transverse force $\|F_{t,lin}\|$ (thick solid line), longitudinal force $\|F_1\|$ (dashed line), and nonlinear transverse force $\|F_{t,nonlin}\|$ (thin solid line). The approximate values computed by Eqs. (A.2) and (A.3) are displayed by dotted lines. Fig. 1 (a) has the parameters of a G_1 piano string, while Fig. 1 (b) has a 100 times higher E value.

results are displayed in Fig. A.2, using the same line convention as in Fig. A.1. Figure A.2 shows that the nonlinear components are larger where E is larger at a fixed T_0 , i.e., the nonlinearity is more prominent at higher ES/T_0 values. The qualitative explanation for this is when the initial stretching of the string is small (T_0 is negligible in comparison with ES), the change of the string length during vibration is no more negligible in comparison with this initial stretching, leading to a significant variation in tension. On the other hand,

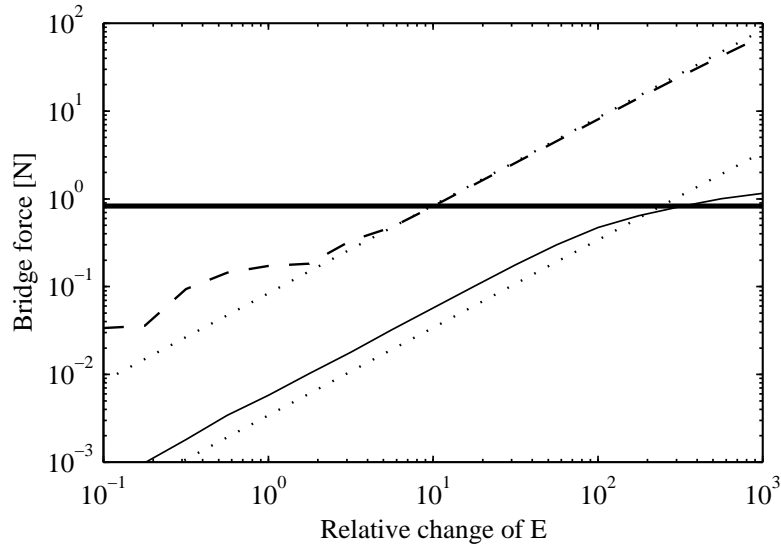


Figure A.2: Euclidean norm of simulated bridge forces as a function of the Young’s modulus E : linear transverse force $\|F_{t,\text{lin}}\|$ (thick solid line), longitudinal force $\|F_l\|$ (dashed line), and nonlinear transverse force $\|F_{t,\text{nonlin}}\|$ (thin solid line). The approximate values computed by Eqs. (A.2) and (A.3) are displayed by dotted lines. The value 10^0 of the x axis corresponds to the E value of a G_1 piano string. The Euclidean norm of the string slope at the bridge is fixed at $\|\partial y/\partial x\| = 10^{-3}$.

if the string is stretched significantly when it is mounted on the instrument (which is typical for nylon strings), the length change during vibration will be much smaller in comparison to the initial stretch.

It can be seen that the approximations of Eqs. (A.2) and (A.3) are quite accurate, but for lower E values the amplitude of the longitudinal component (dashed line) departs from the theoretically predicted one significantly. This is because here the longitudinal modes are excited at their resonance resulting in a larger longitudinal motion (see Sec. 5.3), which is not covered in Eqs. (A.2) and (A.3).

A.2 The Effect of Nonrigid String Terminations

Here we estimate whether the equations of Sec. 5.3 developed for rigid string terminations are applicable for real musical instruments with finite bridge impedance (this Appendix belongs to Sec. 5.3.6). Let us take a look at a special case when all the transverse modes have the same length $L + \delta L$, and the longitudinal modes have the length L . This corresponds to a bridge whose admittance does not depend on frequency and can only move in the transverse direction. In this case L has to be substituted by $L + \delta L$ in Eq. (5.35). The “sum terms” $(m + n)$ in Eq. (5.35) will be of the form $C_p \sin[p\pi x/(L + \delta L)]$, where $p = m + n$. Note that these terms excite only longitudinal modes $k = p = m + n$ in the case of perfectly rigid termination ($\delta L = 0$). For $\delta L \neq 0$, the excitation force originating

from the term p is calculated as

$$\begin{aligned} F_{\xi,k,p}(t)^+ &= \int_{x=0}^L C_p \sin\left(\frac{p\pi x}{L+\delta L}\right) \sin\left(\frac{k\pi x}{L}\right) dx = \\ &= C_p \frac{1}{2} \left[\frac{L}{(p+pd+k)\pi} \sin[(p+pd+k)\pi] - \frac{L}{(p+pd-k)\pi} \sin[(p+pd-k)\pi] \right]_0^L, \end{aligned} \quad (\text{A.4})$$

where the notation $d = L/(L + \delta L) - 1$ has been chosen to simplify the derivations. For $d \approx 0$, the sinusoidal functions can be approximated by their first order Taylor series (i.e., $\sin x \approx x$):

$$F_{\xi,k,p}(t)^+ \approx \pm C_p \frac{L}{2} \left[\frac{pd}{p+pd+k} - \frac{pd}{p+pd-k} \right], \quad (\text{A.5})$$

where the \pm sign is positive if $p+k$ is even, and negative if $p+k$ is odd. Note that the sign has no real importance, as we are interested in the amplitude of the nonlinear components. For $d = 0$ (which holds for a perfectly rigid termination $\delta L = 0$), this expression is zero for $p \neq k$ and equals to $\pm C_p L/2$ for $k = p$.

For $d \neq 0$ and $p = k$, we obtain

$$F_{\xi,k,p=k}(t)^+ \approx \pm C_p \frac{L}{2} \left[\frac{d}{2} - 1 \right], \quad (\text{A.6})$$

which, for realistic $d \approx \delta L/L$ values ($d = 0.01..0.001$), almost equals with the one computed by assuming rigid terminations. This means that the amplitude of the dominant phantom partial (which would be present for rigid terminations) is not changed significantly if the termination is nonrigid.

For $d \neq 0$ and $p \neq k$, i.e., for the secondary phantom partials, which would not be present for rigid terminations, we get

$$F_{\xi,k,p \neq k}(t)^+ \approx \pm C_p \frac{L}{2} \frac{2d}{\frac{p}{k} - \frac{k}{p}}, \quad (\text{A.7})$$

which is maximal for neighboring k and p values. For $p = k + 1$,

$$F_{\xi,k,p=k+1}(t)^+ \approx \pm C_p \frac{L}{2} dk \approx C_p \pm \frac{L}{2} \frac{\delta L}{L} k. \quad (\text{A.8})$$

This means that those terms which would excite mode k only for rigid terminations, excite the neighboring modes $k+1$ and $k-1$ at a relative level $k\delta L/L$ compared to exciting mode k . More distant longitudinal modes $k+2$, $k-2$, etc., are excited even less.

In practice, $\delta L/L = 0.01..0.001$ for musical instruments. For the first longitudinal mode, this means that the terms which are not included in Eq. (5.36a) are 40..60 dB lower compared to the ones present in Eq. (5.36a). For $k = 10$, this difference reduces to 20..40 dB, but we can still state that the dominant terms are those included in Eq. (5.36a), which was computed by assuming infinitely rigid terminations. For longitudinal modes with mode numbers $k \approx 100$, these nonstandard components would have the same magnitude as the theoretical ones in Eq. (5.36a), meaning that Eq. (5.36) is not valid anymore. However, these very high longitudinal modes (around 100 kHz) are not excited effectively due to

the band-limited excitation from the transverse vibration. Note that even if they were effectively excited, they would be well above the audible frequency range. The same derivations can be performed for the “difference terms” of Eq. (5.35), having the form $C_p \sin[p\pi x/(L + \delta L)]$, where $p = |m - n|$, which give exactly the same results.