KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

SOUND SYNTHESIS BY SIMULATION OF PHYSICAL MODELS OF MUSICAL INSTRUMENTS

Promotor:
Prof. dr. ir. B. De Moor
Prof. dr. ir. R. Lauwereins

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Axel NACKAERTS

Mei 2003

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT TOEGEPASTE WETENSCHAPPEN
DEPARTEMENT ELEKTROTECHNIEK
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee)

SOUND SYNTHESIS BY
SIMULATION OF PHYSICAL MODELS
OF MUSICAL INSTRUMENTS

Jury:
Prof. dr. ir. J. Berlamont, voorzitter
Prof. dr. ir. B. De Moor, promotor
Prof. dr. ir. R. Lauwereins, promotor
Prof. dr. ir. F. Catthoor
Prof. dr. ir. M. Moonen
Prof. dr. ir. Y. Rolain (VUB)
Prof. dr. ir. J. Swevers
Prof. dr. ir. J. Van Impe

Proefschrift voorgedragen tot
het behalen van het doctoraat
in de toegepaste wetenschappen

door

Axel NACKAERTS

Voorwoord

Bij het begin van dit proefschrift zou ik graag een aantal mensen willen bedanken die bijgedragen hebben tot het totstandkomen van deze tekst.

Mijn oprechte dank gaat uit naar Prof. Bart De Moor en Prof. Rudy Lauwereins, de promotoren van dit proefschrift. Toen ik enkele jaren geleden met het idee rondliep om onderzoek te doen over geluidssynthese, was ik aangenaam verrast toen beiden openstonden voor dit niet-conventionele vakgebied. Ik wil hen in het bijzonder bedanken voor de vrijheid die ze me gegeven hebben om het onderzoek naar eigen goeddunken in te richten. Prof. Bart De Moor wil ik bovendien bedanken voor het vertrouwen dat hij toonde in mijn advies omtrent de SISTA computerinfrastructuur.

Ik wens de assessoren Prof. Marc Moonen en Prof. Francky Catthoor te bedanken voor het nalezen van het proefschrift en de nuttige commentaar en aanwijzingen. Hun opmerkingen hebben bijgedragen tot de volledigheid en leesbaarheid van dit werk. Ook de leden van de jury, Prof. Jan Swever, Prof. Jan Van Impe en Prof. Yves Rolain wil ik bedanken voor de interesse. Tenslotte wil ik Prof. Jean Berlamont bedanken voor het waarnemen van het voorzitterschap van de examencomissie.

I would also like to thank Nick Collins for composing the *Prelude* and *Suite* for physically modeled guitar. This collaboration has been a very enriching experience. The road from my initial rendering attempts to the final version has brought many small and large improvements to the model and its implementation.

Dit werk werd gefinancierd door een IWT doctoraatsbeurs, en in de laatste maanden door het FWO. Ik wil hen oprecht bedanken voor hun financiële steun.

Ik wil zeker alle collega's en leden van de onderzoeksgroepen SISTA en ACCA niet vergeten. De vele discussies tijdens en na het werk, de Alma-gesprekken, de SISTA-weekends en de ACCA-activiteiten hebben allemaal bijgedragen tot de heel aangename werksfeer die ik heb mogen ervaren. In het bijzonder heb ik genoten van de vele (on)zinnige discussies met bureaugenoten Frank, Wim, Tom, Julia, en Koen. Stuk voor stuk mensen op wie ik heb kunnen rekenen. Veel mensen hebben zonder het te beseffen bijgedragen tot het welslagen van dit werk. Net als een enkel puzzelstuk niets is zonder alle andere, unieke en onmisbare stukken, zo is ook een mens weinig waard zonder zijn vrienden en familie. Hartelijk dank voor jullie steun en vertrouwen.

Abstract

Over the last decades, several powerful sound sythesis techniques have appeared on the consumer market. This is due to both the increased computation power of processors and the re-emergence of software-based synthesis. At the same time, the need for simple user interfaces to control complex intrument models has become apparent. In this thesis, the digital waveguide technique is used to build models for wind and string instruments. These models allow the simple use of natural haptic interfaces for the measurement of the input parameters, without the need of complex parameter mapping.

This thesis consists of three parts. In the first part, a general introduction is given on musical acoustics and the digital waveguide technique for the discrete-time simulation of musical instruments. In the second part, the focus is on different modeling techniques for wind instruments. Starting from an acoustic model, a new digital waveguide model class with formant filtering is constructed, and several multi-note structures are studied. An implementation using the MPEG-4 SAOL is then discussed. The third part deals with the string instruments. First, a single-string model and its associated parameter estimation routines are described. A new theoretical framework for coupled digital waveguide systems is designed and applied to the guitar. This results in a fully coupled instrument model. Finally, a three-layer SAOL implementation is presented, including a player model, a score interpreter, and a core signal processing routine.

Korte Inhoud

Tijdens de laatste decades zijn een aantal krachtige geluidssynthesetechnieken op de consumentensmarkt verschenen. Dit is te danken aan enderzijds de sterk toegenomen rekenkracht van processoren, en anderzijds aan de her-verschijning van software-gebaseerde synthese. Over dezelfde periode is tevens de noodzaak van een eenvoudige gebruiksinterface om een complex synthetisch instrument te beturen duidelijk geworden. In deze thesis wordt de digitale golfgeleider techniek gebruikt om modellen op te bouwen voor de blaas- en snaarinstrumenten. Deze modellen laten toe om een natuurlijke haptische interface te gebruiken voor de meting van de aansturingsparameters, zonder daarvoor een complexe omrekening nodig te hebben.

Deze thesis bestaat uit drie delen. Het eerste deel omvat een inleiding tot de muzikale akoestiek en de digitale golfgeleidermodellen gebruikt voor de discrete-tijd simulaties van muziekinstrumenten. In het tweede deel ligt de aandacht op de simulatie van de blaasinstrumenten. Vertrekkend van een volledig akoestisch model wordt een nieuwe klasse digitaal golfgeleidermodellen met formant filter opgebouwd, en worden verschillende multinoot modellen bestudeerd. Een implementatie van de modellen in de MPEG-SAOL taal wordt dan besproken. Het derde deel is gericht op de studie van de snaarinstrumenten. Eerst worden een enkel-noot model en de bijhorende parameterschattingsroutines beschreven, waarna een nieuw theoretisch raamwerk wordt afgeleid voor gekoppelde digitale golfgeleidermodellen. Dit leidt uiteindelijk tot een volledig gekoppeld instrumentmodel. Ook dit model wordt geïmplementeerd, ditmaal gebruik makend van een drielaags model dat bestaat uit het spelermodel, de partituur-interpretatiemodule en het kern-signaalverwerkingsalgoritme.

Glossary

Mathematical Notation

$\mathbf{v}$	vector $\mathbf{v}$
$\mathbf{v}(z)$	vector $\mathbf{v}$, function of the z–transform variable
$\mathbf{M}$	matrix $\mathbf{M}$
$\mathbf{M}(z)$	matrix $\mathbf{M}$, function of the z–transform variable
$\underline{\mathbf{v}}$, $\underline{\mathbf{M}}$	frequency–domain equivalents of $\mathbf{v}$ and $\mathbf{M}$
$\mathbf{M}^T$	transpose of matrix $\mathbf{M}$
$\mathbf{M}^*$	complex conjugate of matrix $\mathbf{M}$
$\mathbf{M}^H = (\mathbf{M}^*)^T$	Hermitian transpose of matrix $\mathbf{M}$
$\mathbf{M}^{-1}$	inverse of matrix $\mathbf{M}$
$\mathbf{M}^\dagger$	pseudo–inverse of matrix $\mathbf{M}$
$\det \mathbf{M}$	determinant of matrix $\mathbf{M}$
$\mathbf{v}(m)$	m–th element of vector $\mathbf{v}$
$[\mathbf{v}(z)]_m$	m–th element of vector function $\mathbf{v}(z)$
$\mathbf{M}(m,n)$	element on the m–th row and n–th column of matrix $\mathbf{M}$
$[\mathbf{M}(z)]_{m,n}$	element on the m–th row and n–th column of matrix function $\mathbf{M}(z)$
$x \star y$	convolution of $x[k]$ and $y[k]$
$\mathbb{N}$	the set of natural numbers
$\mathbb{N}_0 = \mathbb{N} \setminus \{0\}$	the set of natural numbers larger than 0
$\mathbb{Z}$	the set of integer numbers
$\mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}$	the set of integer numbers except 0
$\mathbb{Q}$	the set of rational numbers
$\mathbb{R}$	the set of real numbers
$\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$	the set of real numbers except 0
$\mathbb{R}^+$	the set of positive real numbers
$\mathbb{C}$	the set of complex numbers
$\Re\{x\}$	real part of $x \in \mathbb{C}$
$\Im\{x\}$	imaginary part of $x \in \mathbb{C}$
x^*	complex conjugate of x
$\mathrm{conj}(\cdot)$	complex conjugation

$\hat{x}$	estimate of x
$\lfloor x \rfloor$	largest integer smaller or equal to $x \in \mathbb{R}$
$\lceil x \rceil$	smallest integer larger or equal to $x \in \mathbb{R}$
$\mathrm{rnd}(x)$	round $x \in \mathbb{R}$ to the nearest integer
$|\cdot|$	absolute value
$a \approx b$	a is approximately equal to b

Fixed Symbols

c	speed of sound
$\mathcal{C}$	cepstrum
c_{ij}	coupling coefficients
c_{long}	longitudinal wave propagation speed
c_{tors}	torsional wave propagation speed
c_{trans}	transversal wave propagation speed
C_p	specific heat of air
ξ	integer proportionality factor (cylinder, cone)
D	fractional delay
δ_{therm}	thermal layer thickness
δ_{visc}	viscous layer thickness
E	Young modulus
$\mathcal{E}$	spectral envelope
e_k	error sequence
η	viscosity of air
f_0	fundamental frequency
f_s	sampling frequency
G	shear modulus
$h(n)$	filter coefficients
H	Hankel matrix
$\mathcal{H}$	filter transfer fucntion
I	input
I_{eq}	equivalent input
J	Bessel function
k	wave number
k	spring constant
κ	thermal conductivity
K_T	torsional stiffness factor
L	length
L_f	facing length of reed
L_s	spatial sampling distance
m	mass
M	coupling matrix
μ	damping
ω	pulsation

ω_n	resonant frequency (rad/s)
Ω	continuous-time dynamical coupling matrix
p	pressure function (time domain)
p^+	right-traveling pressure function
p^-	left-traveling pressure function
p_Δ	pressure difference
p_{oc}	oral cavity pressure
P	pressure function (Laplace domain)
r	radius, position along radial axis
$\mathcal{R}$	reflectance
ρ	density
S	Area
$\mathcal{S}$	spectrum
t	time
T	temperature
$\mathcal{T}$	transmittance
T_0	string tension at rest
T_s	sampling period
T_{cone}	conical section transmission matrix
T_{cyl}	cylindrical section transmission matrix
u	volume flow
u_r	volume flow in reed channel
u_{rd}	dynamic volume flow in reed channel
u_{rs}	static volume flow in reed channel
v_p	wave velocity
x	position along principal axis
y	displacement
Y	acoustic admittance
Z	acoustic impedance
Z_b	bore acoustic impendance
Z_{in}	acoustic input impedance
Z_L	acoustic load impedance
Z_r	reed channel acoustic impendance

Acronyms and Abbreviations

A/D	Analog–to–Digital converter
ASIC	Application–Specific Integrated Circuit
cf.	`confer` : compare with
CPU	Central Processing Unit
D/A	Digital–to–Analog converter
DFT	Discrete Fourier Transform

DSP	Digital Signal Processor
EDR	Energy Decay Relief
e.g.	`exempli gratia` : for example
Eq.	equation
FFT	Fast Fourier Transform
FIR	Finite Impulse Response filter
HSVD	Hankel Singular Value Decomposition
IDFT	Inverse Discrete Fourier Transform
i.e.	`id est` : that is
IFFT	Inverse Fast Fourier Transform
IIR	Infinite Impulse Response filter
LMS	Least Mean Squares
LS	Least Squares
MFlops	Millions of Floating Point Operations per Second
MIDI	Musical Instrument Digital Interface
MIMO	Multi–Input Multi–Output system
MPEG	Motion Picture Experts Group
PC	Personal Computer
SAOL	Structured Audio Orchestra Language
SASL	Structured Audio Score Language
sHSVD	subband Hankel Singular Value Decomposition
SNR	Signal–to–Noise Ratio
STFT	Short-Time Fourier Transform
SVD	Singular Value Decomposition
vs.	versus

Contents

I   General Acoustic Theory   13

2   Wind Instruments   15

3   String Instruments   23

4   Acoustics   33

Samenvatting

Geluidssynthese door simulatie van fysische modellen van muziekinstrumenten

Hoofdstuk 1 : Inleiding

Geluidssynthese is historisch ontstaan met de wens om een willekeurig geluid te genereren met behulp van wiskundige technieken. Tegenwoordig wordt geluidssynthese alom gebruikt, wat wordt gestaafd door de gezonde economische resultaten van de fabrikanten - bedenk dat in elke PC een geluidskaart zit ! Geluidssynthese door simulatie van een fysisch model van een muziekinstrument onderscheidt zich vooral door de *natuurlijkheid* van de klank en de *responsiviteit* voor subtiele speelvariaties.

In 1948 werd in Parijs door Pierre Schaeffer de eerste elektronische muziekstudio ingericht. Door de combinatie van platenspelers, taperecorders, en microfoons werd een nieuwe kunstvorm geschapen, de *musique concrete*. Dit betekende de start van de elektronische muziek. Dit was een logisch gevolg van de onstopbare evolutie van muziekinstrumenten, van mechanische instrumenten uit de Oudheid, over automatische instrumenten tijdens de Industriële Revolutie, tot de elektronische instrumenten als de Theremin in het begin van de twintigste eeuw. In 1958 genereerde Max Matthews als eerste muziek door middel van een

computer.

Tijdens de halve eeuw durende ontwikkeling van de digitale geluidssynthese hebben de synthesetechnieken de bechikbare rekenkracht gevolgd. In de beginjaren was enkel synthese door middel van de eenvoudigste additieve synthese-algoritmes haalbaar. In diezelfde periode is ook de scheiding ontstaan tussen geluidssynthese in reële tijd, gebruik makend van specifieke apparatuur, en het gebruik van programmeertalen gericht op geluidssynthese. Door de enorme toename in beschikbare rekenkracht zijn deze twee wegen weer stilaan aan het convergeren.

Hoewel de studie van de werking van muziekinstrumenten al millenia oud is, is geluidssynthese door simulatie van een fysisch model van een muziekinstrument is een recente techniek. Deze synthesetechniek tracht de klank van een echt instrument na te bootsen door de volledige mechanisch-akoestische werking ervan te simuleren. Fysische modellering is de generische term voor een scala aan verschillende modelleringstechnieken, die gaan van uiterst gedetailleerde eindige elementen modellen, tot een ruwe benadering van de dominante trillingsmodi. De fysische modellen die gehanteerd worden in dit proefschrift vinden hun oorsprong in de algemene golfvergelijking die de voortplanting van trillingen in het instrument beschrijft.

Geluidssynthese kent heel wat toepassingen. De meest voor de hand liggende toepassing is het gebruik van geluidssynthese als elektronisch of virtueel instrument. Dit laat toe een partituur al dan niet met tussenkomst van een menselijke speler hoorbaar te maken, met de keuze uit een oneindig klankpallet. In sommige gevallen is het gebruik van geluidssynthese de enige mogelijkheid: de partituur van een hedendaagse compositie kan te complex zijn om nog door een menselijke speler gelezen te kunnen worden, de klankvereisten vallen buiten de menselijk voorstellingsvermogen, of de uitvoeringssnelheid overstijgt de menselijke mogelijkheden (*transhumane compositie*). Een tweede toepassingsgebied is het gebruik van geluidssynthese voor parametrische geluidscompressie. Door de klank die door een instrument geproduceerd wordt op symbolische wijze te coderen, kan een extreem hoge compressiegraad gehaald worden. Een natuurgetrouw fysisch model van het instrument in combinatie met een met expressieve meta-informatie uitgebreide partituur, is een zeer compacte beschrijving. Om een voorbeeld te geven: de mono-versie van ⋆ **CD track 30** vereist in ongecomprimeerde vorm 3100kB aan gegevens, de MP3-gecodeerde versie vereist 560kB, en het MPEG-4 model dat een identiek geluid produceert is slechts 7kB groot, wat neer-

komt op een compressieverhouding van 1:5.5 voor MP3 en maar liefst 1:437 voor het fysisch model.

In dit proefschrift wordt een integrale modelleringsmethode aangeboden, van fysico-akoestische beschrijving tot implementatie, voor de blaasinstrumenten en de snaarinstrumenten. De nadruk ligt niet op de akoestische juistheid van de modellen, maar op hun praktische bruikbaarheid en eenvoudige calibreerbaarheid. Dit laatste moet mogelijk zijn zonder gespecialiseerde apparatuur.

Dit proefschrift bestaat uit drie delen. In het eerste deel worden de algemene akoestische en wiskundige concepten uitgewerkt, nodig voor de modellering van de blaas- en snaarinstrumenten. Het tweede deel behandelt de blaasinstrumenten, met als typisch voorbeeld de moderne klarinet, en in het derde deel worden modellen opgebouwd voor de snaarinstrumenten, met de nadruk op de klassieke akoestische gitaar. De belangrijkste bijdragen van dit proefschrift zijn terug te vinden in het tweede en het derde deel.

Deel I : Algemene akoestiek

Het eerste deel vorm een algemene inleiding tot de problemen die behandeld worden in het proefschrift. In de eerste twee hoofdstukken wordt een qualitatieve en muzikale beschrijving gegeven van de struktuur en werking van enkele muziekinstrumenten. De daaropvolgende twee hoofdstukken behandelen de akoestische eigenschappen in continue tijd en discrete tijd. De hier afgeleide eigenschappen dienen als basis voor de hele verdere modellering.

Hoofdstuk 2 : Blaasinstrumenten

Hoofdstuk 2 is een algemene, qualitatieve inleiding tot de akoestiek van blaasinstrumenten. Op enkele uitzonderingen na, hebben de blaasinstrumenten een eeuwenlange ontwikkeling genoten. Ze zijn geëvolueerd tot een relatief klein aantal moderne instrumenten, met uitgekiende mechanische strukturen die stemming en speelbaarheid optimaliseren. Blaasinstrumenten produceren een toon zolang de speler energie blijft toevoeren aan het systeem. Deze oscillatie wordt meestal opgewekt in een luchtkolom, een resonante struktuur bestaande uit cylindrische en kegelvormige secties. Het basisprincipe is hetzelfde als bij een elektronische oscillator:

een signaal wordt gevoed aan een resonerende struktuur, en terugge-
koppeld naar een saturerende versterker. De blaasinstrumenten kunnen
opgedeeld worden volgens het type mechanische versterker. In het geval
van de klarinet en de saxofoon is dit een mondstuk met een enkel riet,
bij de hobo en de fagot, een dubbel riet, bij de koperblazers zijn het de
lippen van de speler, terwijl bij fluiten en sommige orgelpijpen een snelle
luchtstroom gericht op een scheidend oppervlak hiervoor dienst doet.

Het typevoorbeeld dat in dit proefschrift wordt gehanteerd, is de mo-
derne klarinet. Deze bestaat uit een cilindrische buis met eenvoudige
toongaten, een kleine klankbeker, en een mondstuk met enkel riet. De
saxofoon is een heel gelijkaardig instrument, maar gebruikt een kegel-
vormige buis als resonator. De klarinet, dwarsfluit, hobo en sopraan-
saxofoon verschillen vaak in slechts één enkel opzicht: ze hebben een
andere basisvorm als resonator (cilinder of kegel) of gebruiken een an-
der excitatiemechnanisme (enkel riet, dubbel riet, of luchtstroom). De
overeenkomst is zodanig dat een niet-getraind oor slechts moeilijk een
onderscheid kan horen tussen bijvoorbeeld een hobo en een sopraan-
saxofoon. De overeenkomst in fysische struktuur laat toe om de hele
familie van de blaasinstrumenten te modelleren met een beperkt aantal
elementaire modellen.

Hoofdstuk 3 : Snaarinstrumenten

In dit hoofdstuk worden de eigenschappen van de snaarinstrumenten in
het algemeen besproken. We vergelijken hierbij de struktuur en werking
van de akoestische en elektrische gitaren, de viool en de piano. Het
algemene principe waarop alle snaarinstrumenten steunen, is de akoes-
tische of elektrische versterking van de trilling van een snaar. De snaren
worden getokkeld (zoals bij de gitaar, de harp of het clavecimbel), aan-
geslagen (zoals bij de piano) of gestreken (vioolfamilie). De snaren zijn
opgehangen tussen twee punten, waarbij één punt niet als een volkomen
vaste ophanging beschouwd mag worden. Op dit punt is de snaar in
contact met een klankbord (rechtsreeks of via een kam), waarnaar een
deel van de trillingsenergie wordt overgedragen. Dit klankbord is een
deel van een akoestische versterker, die de mechanische trillingen omzet
in kleine luchtdrukverschillen, en tevens de uitstraling ervan bepaalt.
In het geval van een elektrische gitaar wordt de trilling van de snaar
op elektromagnetische wijze gemeten, waarna het signaal elektronisch
wordt versterkt en gevoed aan een luidspreker. Dit basismechanisme
komt bij alle snaarinstrumenten voor, wat toelaat om een gemeenschap-
pelijk model op te stellen dat mits kleine aanpassingen bruikbaar is voor

alle snaarinstrumenten.

Het typevoorbeeld dat behandeld wordt is de klassieke gitaar. Dit instrument, een afstammeling van de zestiende-eeuwse spaanse *vihuela*, kreeg in de negentiende eeuw zijn huidige vorm. De moderne gitaar heeft zes nylon snaren die via een eenvoudige kam verbonden zijn met het bovenblad van de klankkast. Het bovenblad is inwendig versterkt door een symmetrisch of asymmetrisch patroon van houten ribben, die een belangrijke invloed hebben op het uiteindelijke timbre van het instrument. De speler kan de lengte van de snaren bepalen door deze tegen de toets te drukken. In tegenstelling tot bijvoorbeeld de viool, heeft de gitaar metalen frets, wat het aantal verschillende snaarlengtes beperkt. Zowat alle typische eigenschappen van de snaarinstrumenten, zoals de aanwezigheid van twee polarizatierichtingen voor de trillingen, en snaarkoppeling, komen bij de klassieke gitaar voor, met het voordeel dat de fysische omvang van het instrument beperkt is (laag aantal snaren, eenvoudige kam etc.).

Hoofdstuk 4 : Akoestiek

Dit hoofdstuk geeft een inleiding tot de muzikale akoestiek. Vertrekkend van de differentiaalvergelijkingen die de golfvoortplanting door lucht en in snaren beschrijven in **paragraaf 4.2**, worden de algemene eigenschappen van de elementaire bouwblokken van muziekinstrumenten afgeleid in **paragraaf 4.3**. De combinatie van de verschilllende elementen tot een akoestisch systeem wordt dan besproken in **paragraaf 4.4**.

Mechanische, akoestische, en elektrische trillingen zijn de geluidsbronnen bij de muziekinstrumenten. De menselijke perceptie van klank is een gevolg van de versterking van inkomende luchtdrukverschillen door het trommelvlies en het middenoor, en de omzetting ervan in zenuwpulsen door de sensorcellen in de cochlea. De muzikale akoestiek is het onderdeel uit de fysische akoestiek, dat tracht alle mechanismen die hoorbare geluidsgolven produceren, te beschrijven. Dit omvat de generatie, propagatie, en interactie van geluidsgolven in verschillende media.

De basiswet van de muzikale akoestiek is de drie-dimensionale golfvergelijking (4.1). Het oplossen van deze vergelijking met de juiste randvoorwaardes laat toe om de akoestische eigenschappen te bepalen van zowat alle elementen die gebruikt worden bij muziekinstrumenten. Zowel de vlakke en sferische golven in lucht (**paragraaf 4.2.1**), als de transversale, longitudinale en torsionele golven in snaren (**paragraaf 4.2.2**) zijn oplossingen van deze vergelijking. Door het opleggen van randvoor-

waarden zijn de eigenschappen van eenvoudige akoestische strukturen te bepalen (**paragraaf 4.3**). Dit zijn in het bijzonder de cilindrische en konische klankbuis voor de blaasinstrumenten, en de Helmholtz resonator voor de snaarinstrumenten. De verschillende elementaire modellen kunnen door middel van de in **paragraaf 4.4** beschreven methodes, gecombineerd tot volledige instrumenten.

Hoofdstuk 5 : Digitale golfgeleidermodellen

In dit hoofdstuk worden de theoretische akoestische eigenschappen uit hoofdstuk 4 gediscretiseerd in ruimte en tijd. Dit leidt tot de digitale golfgeleidermodellen, afgeleid in **paragraaf 5.1**. Digitale golfgeleidermodellen zijn afgeleid uit de algemene oplossing van de golfvergelijkingen, voorgesteld door links- en rechtslopende golven, en vormen uiterst efficiënte strukturen voor geluidssynthese in reële tijd. Een interessant aspect hierbij is dat een rechtstreekse omrekening van continue tijd naar discrete tijd leidt tot vertragingslijnen die een niet-geheel aantal elementen bevatten. Om dit te kunnen implementeren is een interpollatiemethode noodzakelijk. Naast de ideal laagdoorlaat interpollatie, het gebruik van een allpass filter, en de bepaling van de interpollerende veelterm door Lagrange interpollatie, worden in **paragraaf 5.1.3** ook de kubische Hermite (Catmull-Rom) en B-spline interpollatie besproken. Tabel 5.1 geeft een overzicht van welke methode het meest geschikt is voor de simulatie van een bepaalde instrumentfamilie. In **paragraaf 5.2** worden de discrete beschrijvingen van de akoestische elementen uitgewerkt. In **paragraaf 5.3** worden de methodes beschreven om de discrete elementen samen te voegen tot complete discrete systemen.

Deel II : Blaasinstrumenten

In het tweede deel van het proefschrift worden de modellen voor de blaasinstrumenten uitgewerkt. Eerst wordt een volledige akoestische beschrijving bepaald, die dan wordt vereenvoudigd door gebruik te maken van het zg. formant filter. Zo worden sets van modellen bekomen die aaneengebonden worden in de multinoot modellen.

Hoofdstuk 6 : Volledig akoestisch model

In hoofdstuk 6 bouwen we een volledig akoestisch model op voor een enkel-riet blaasinstrument. Het instrument bestaat uit twee delen: een lineaire resonator, en het niet-lineaire excitatiemechanisme. In **para-**

graaf 6.1 wordt de resonator van de klarinet en de saxofoon bespro-
ken, met verwaarlozing van de toongaten. De resonator bestaat uit een
cilindrische of konische klankbuis, en een kleine klankbeker. Gebruik
makend van de discrete elementaire modellen uit hoofdstuk 5 is de vol-
ledige akoestische beschrijving eenvoudig te bepalen. In **paragraaf 6.2**
bespreken we de trilling van het riet. Hier komen twee aspecten sa-
men: enerzijds is er de beweging en vervorming van het riet door het
drukverschil tussen de mondholte van de speler en de luchtdruk aan de
binnenzijde van het mondstuk, en anderzijds is er de luchtstroom door
de rietopening.

De combinatie van het enkel-riet en het mondstuk gedraagt zich als een
drukgecontroleerde klep, die energie toelaat in het instrument om de
oscillatie aan de gang te houden. We bespreken verschillende modellen
voor de beweging van het riet. Het eenvoudigste model is lineair en
geheugenloos, waarbij de verplaatsing van het riet berekend kan worden
door middel van de wet van Hooke. Indien de massa van het riet in
rekening wordt gebracht, bekomen we een tweede-orde lineair systeem
beschreven door vergelijking (6.10). Hogere-orde modellen beschouwen
een riet als een dunne ingeklemde balk. De oplossingen van de differen-
tiaalvergelijkingen die dit systeem voorstellen zijn de trillingsmodi van
het riet. Het blijkt dat voor alle praktische doeleinden enkel de eerste en
eventueel de tweede trillingsmode belangrijk zijn, aangezien de hogere
modi buiten het bereik van het menselijk gehoor vallen. Als kantte-
kening willen we hierbij opmerken dat de vorm die het riet aanneemt
voor een gegeven drukverschil zeer goed overeenkomt met de (opgeme-
ten) vorm van een mondstuk, wat erop wijst dat mondstukken tot deze
theoretisch optimale vorm zijn geëvolueerd.

Hoofdstuk 7 : Formant gefilterd model

De modellen uit hoofdstuk 6 zijn opgebouwd vanuit akoestisch oogpunt,
gebruik makend van de fysische eigenschappen van de instrumenten. Dit
leidt tot kwalitatief hoogstaande modellen die het gedrag van een instru-
ment goed kunnen voorspellen, maar die ook moeilijk te calibreren zijn,
net door het hoge aantal parameters en de totale complexiteit van deze
modellen. In dit hoofdstuk wordt gezocht naar vereenvoudigde model-
len die toch een realistisch geluid produceren. Door twee verschillende
synthesetechnieken te combineren, bekomen we een nieuwe, efficiënte en
calibreerbare klasse van fysische modellen.

In **paragraaf 7.2** beschrijven we een nieuwe combinatie van bron-filter
synthese en de digitale golfgeleidermodellen. Blaasinstrumenten hebben

een harmonisch rijk spectrum, dat karakteristieke bulten of *formanten* vertoont. Een formant is een bandgelimiteerde concentratie van energie in het spectrum van het signaal. De kern van de bijdrage van dit proefschrift voor de synthese van de blaasinstrumenten ligt in de uitbuiting van deze eigenschap om een synthetisch digital golfgeleidermodel van een instrument te laten klinken als een opname van het echte instrument. Aan de basis hiervan ligt de observatie dat de trillingsmodi van een instrument bepaald worden door de lineair te modelleren resonator. Het niet-lineaire excitatiemechanisme bepaalt welke modes in welke mate geëxciteerd worden. Een *fortissimo* bespeeld instrument kan dus beschouwd worden als een volledig geëxciteerd systeem. Een instrument zachter bespelen betekent dus dat er minder energie naar bepaalde trillingsmodes gaat. Het formant gefilterde golfgeleidermodel beschouwt het model geëxciteerd met maximale ingangswaarden als een complexe oscillator, en tracht de spectra van synthese en opname overeen te doen komen dankzij een dynamisch equalizatiefilter. Het hele model is voorgesteld in figuur 7.8.

Het digitaal golfgeleidermodel met formant filter bestaat uit drie grote delen: de lineaire resonator, het niet-lineaire excitatiemechanisme, en het dynamische formant filter. Het lineaire deel van het model is een voorstelling van de klankbuis, de toongaten, en de klankbeker van het instrument. Aangezien dit model voor slechts een enkele noot geldig is, kunnen alle filteroperaties samengevoegd worden in één reflectiefilter. Het niet-lineaire excitatiemechanisme kan op verschillende manieren geïmplementeerd worden: rechtstreeks als een niet-lineaire functie, als een drukafhankelijke reflectiecoëfficiënt, of door de reflectie te benaderen met een polynomiale functie, of als een combinatie van deze methodes. Het laatste onderdeel is het formant filter. Het formant filter wordt bepaald door vergelijking (7.23) en wordt berekend op basis van de spectrale enveloppes van simulatie en opname [74]. In **paragraaf 7.2.3** worden verschillende beschrijvingen voor de spectrale enveloppe onderzocht. De eenvoudigste beschrijving is een tweede-orde resonante transfertfunctie [68]. Voor een zacht bespeeld instrument zit het grootste deel van de energie in een smalle frequentieband rond de fundamentele frequentie van de gespeelde noot. Bij een luidere toon verschijnen hogere harmonischen. De tweede-orde beschrijving verschaft veel qualitatieve informatie over het gedrag van het instrument, maar is niet krachtig genoeg om het volledige timbre van het instrument te vatten. Hogere-orde modellen kunnen bepaald worden door Lineair Predictieve Codering (LPC), cepstrum-gebaseerde methodes, of Wiener filters te gebruiken. Figuur 7.13 vat de hele procedure samen om een formant

gefilterd digitaal golfgeleidermodel op te bouwen voor een willekeurig blaasinstrument. Dat een natuurgetrouwe synthese mogelijk is, wordt gestaafd door zowel de objectieve meting van de afwijking van de synthetische toon ten opzichte van een opname, als de subjectieve beoordeling. Op de CD die bij het proefschrift hoort, staan enkele geluidsfragmenten die de hier besproken methodes demonstreren. Een volledige omschrijving van de geluidsfragmenten is in **Appendix A** terug te vinden.

Hoofdstuk 8 : Multinoot modellen

In hoofdstuk 8 worden de eigenschappen en de verschillende implementaties van toongaten besproken. De resonator van vele blaasinstrumenten kan door de speler gewijzigd worden door het openen of sluiten van openingen of toongaten, of door het toevoegen van extra buislengte door het indrukken van ventielen. Tijdens de overgang tussen de open en gesloten toestand van een toongat, gaat het instrument door een reeks intermediaire configuraties. Deze transiënten moeten toegevoegd worden aan de formant gefilterde digitale golfgeleidermodellen om een volledig instrumentmodel te verkrijgen [67].

In **paragraaf 8.1.1** geven we de akoestische beschrijving van één toongat en van een sequentie van toongaten. De akoestische modellen zijn niet erg geschikt voor toepassingen in reële tijd, maar geven wel aanwijzingen over mogelijke benaderingen. Daarom werden zes verschillende modellen opgesteld en vergeleken. Eerst wordt een benaderend akoestisch model met een beperkt aantal toongaten opgebouwd dat dient als referentie. Het wordt dan vereenvoudigd tot een model met slechts één toongat, en verschillende seriële en parallelle combinaties worden afgeleid. Figuren 8.5 en 8.6 tonen de verschillende modelstrukturen. De twee modellen met de beste kwaliteit/complexiteitsverhouding zijn een seriëel model met één toongat, en een parallel model met gemeenschappelijke niet-lineairiteit. In **paragraaf 8.2** worden deze vergeleken. Het parallelle model kan verkregen worden uit het seriële model door lineaire superpositie van trillingsmodes in de klankbuis te veronderstellen tijdens de transiënt van de ene noot naar de andere. Er wordt aangetoond dat mits een goede keuze van de modelparameters, een parallel model een bijzonder goede benadering vorm voor een seriëel model, doch met een significant gereduceerde rekencomplexiteit. ⋆ **CD Tracks 18-19** werden berekend door middel van de hier beschreven modellen, en tonen aan dat er zo goed als geen hoorbaar verschil is tussen een seriëel en een parallel model.

Hoofdstuk 9 : Implementatie

Het laatste hoofdstuk uit het tweede deel van het proefschrift behandelt de implementatie van de formant gefilterde digitale golfgeleidermodellen. Er werd geopteerd voor de MPEG-4 Structured Audio programmeertaal om de fysische modellen te implementeren.

MPEG-4 is een initiatief van de MPEG standardisatiecommissie, en beoogt de effeciente codering van multimedia objecten. MPEG-4 Audio is het onderdeel van de standaard dat is toegespitst op de codering van spraak, transformatie-gecodeerd geluid, en geluidssynthese. Voor dit laatste werd een specifieke programmeertaal beschreven die bestaat uit een reeks primitieve signaalverwerkingsroutines en een syntax die toelaat om een virtueel orkest op te bouwen. De implementatie van de formant gefilterde modellen is relatief eenvoudig indien rekening wordt gehouden met de beperkingen van deze taal.

In dit hoofdstuk wordt ook de vergelijking gemaakt tussen de formant gefilterde modellen en andere, bestaande modellen. Er wordt vergeleken met een ander fysisch model, een spectraal model, wavetable synthese, en twee abstracte wiskundige technieken.

Deel III : Snaarinstrumenten

Het derde deel van het proefschrift behandelt de snaarinstrumenten, met als typevoorbeeld de akoestische gitaar. Na de beschrijving van de trilling van een enkele snaar, bouwen we een algemeen model voor meerdere gekoppelde snaren. Samen met de klankkastmodellen vormen deze het uiteindelijke gitaarmodel. De implementatie vereist daarbij nog een eenvoudig spelermodel.

Hoofdstuk 10 : Enkel-snaar modellen

Ondanks de schijnbaar eenvoudige struktuur van de akoestische gitaar, vertoont dit instrument een redelijk complex gedrag. Om het gedrag van het volledige instrument te modelleren is eerst een betere kennis nodig van de eigenschappen van een hypothetische gitaar met slechts één snaar. In hoofdstuk 10 worden algemene methodes afgeleid die als basis dienen voor de beschrijving van de volledige instrumenten.

Paragraaf 10.1 beschrijft een eenvoudig gitaarmodel dat eenvoudig verfijnd kan worden tot een volledig gitaarmodel. Het eenvoudige model is gebaseerd op de mechanische werking van de gitaar en bestaat uit drie grote delen: de snaarexcitatie (de invoer van de speler in het model), de

xxix

snaren zelf, en de klankkast. **Paragraaf 10.2** behandelt de snaarexcita-
tie. Zowel de plaats als de aard van de tokkel hebben een grote invloed
op de klank die het model produceert. In **paragraaf 10.3** worden de
enkel-snaar modellen uitgewerkt. De akoestische eigenschappen van een
ideal snaar werden reeds besproken in hoofdstuk 4. Voor de akoestische
gitaar zijn vooral de transversale golven van belang. Deze kunnen zich
in twee ruimtelijke dimensies (polarizaties) voortplanten: evenwijdig of
loodrecht op het bovenblad van de klankkast. Door de lichte anisotro-
pie van de ophanging verschillen de eigenschappen van de vertikale en
horizontale polarizatie. In een eerste faze wordt slechts één enkele po-
larizatie beschouwd. De evolutie van de transversale golven in tijd en
ruimte kan bepaald worden aan de hand van de algemene golfvergelij-
king, uitgebreid met hogere-orde termen om de invloed van de eindige
snaardikte en de stijfheid van de gebruikte materialen in rekening te
brengen. Dit leidt tot de klankverschillen tussen nylon, bronzen, en sta-
len snaren. De algemene oplossing van de uitgebreide golfvergelijking
kan gediscretiseerd worden, gebruik makend van digitale golfgeleiders.
Figuren 10.8 en 10.9 tonen twee digitale golgeleidermodellen voor een
snaar met één polarizatie, afhankelijk van de gekozen golfveranderlijke
(verplaatsinggolven of krachtgolven). Dit snaarmodel kan vereenvoudigd
worden door de twee vertragingslijnen samen te voegen tot een enkele
lijn. Verdere verfijning van dit lineaire model is mogelijk door verschil-
lende niet-lineaire effecten toe te voegen. Het belangrijkste effect is
de spannings-gemoduleerde golfvoortplantingssnelheid voor transversa-
le golven. Dit is een gevolg van de uitrekking van de snaar door het
tokkelen zelf, en leidt tot een lichte wijziging van de snaarspanning en
dus de golfvoortplantingssnelheid. Het is afhankelijk van het materiaal
waaruit de snaar bestaat, en is meer uitgesproken voor stalen snaren.
Een tweede effect is de beperkte bewegingsruimte van de snaren. Bij te
grote uitwijking slaan de snaren tegen de frets, wat tot gekletter leidt.
Het enkel-snaar, enkel-polarizatie model is de basis waarop alle verdere
modellen worden gebouwd.

In **paragraaf 10.4** wordt de invloed van het bestaan van de twee polari-
zaties besproken. Hoewel de parameters voor de twee polarizatie slechts
weinig verschillen, is de invloed ervan duidelijk hoorbaar. De golven
die zich voortplanten in de twee polarizaties zien een licht verschillende
snaarlengte, wat leidt tot twee fundamentele frequentie met een paar
tiendes van een Hz verschil. Onafhankelijk van elkaar is dit onhoorbaar,
maar wanneer beide polarizaties tegelijk worden geëxciteerd, wat in de
praktijk altijd het geval is, wordt het frequentieverschil hoorbaar als zwe-
vingen. De afwezigheid van zwevingen in de enkel-polarizatie modellen

geeft deze een onnatuurlijk karakter. De twee polarizaties kunnen als onafhankelijke en licht verschillende snaren gemodelleerd worden.

Om een natuurgerouwe synthese te verkrijgen zijn betrouwbare schattingen van de modelparameter nodig. Hiervoor worden in **paragraaf 10.5** paramerestimatieroutines uitgewerkt [71]. De parameters kunnen zowel in het tijdsdomain als in het frequentiedomein bepaald worden. In het tijdsdomein verschaft de amplitude-enveloppe informatie over de demping van het signaal. In het frequentiedomain kan een schatting bekomen worden van de fundamentele frequentie en de maat van inharmoniciteit. De meest doeltreffende techniek die werd ontwikkeld is subband-gebaseerd. Het signaal wordt eerst door middel van de *Short Time Fourier Transform* naar het frequentiedomain getransformeerd. Uit de berekende tijd-frequentieweergave kan een subband geselecteerd worden, waaruit dan weer door toepassing van de Hankel singuliere waarden ontbinding een nauwkeurige en betrouwbare schatting van de frequenties en demping van de twee polarizaties gevonden wordt. De procedure is terug te vinden in paragraaf 10.5.4. Simulatie van het model waarvan de parameters bepaald werden met deze methodes, produceert een heel natuurlijke klank.

Hoofdstuk 11 : Meerdere-snaar modellen

Eén belangrijk aspect dat de natuurlijkheid van de klank bepaalt, is het meetrillen van de andere snaren als een snaar wordt aangeslagen. Dit koppelingsfenomeen leidt tot zwevingen en gewijzigde dempingskarakteristieken. In dit proefschrift werd een algemeen theoretisch raamwerk opgebouwd om verschillende digitale golfgeleidermodellen [73].

Paragraaf 11.1 behandelt de wiskundige achtergrond voor de beschrijving van gekoppelde oscillatoren en gekoppelde digitale golfgeleidersystemen. Snaarkoppeling is een fysisch verschijnsel dat ontstaat wanneer er een energie-overdracht mogelijk is tussen twee of meer verbonden snaren. Indien die overdracht significant is, kunnen de twee snaren meer als onafhankelijk beschouwd worden, maar moeten als één groot systeem beschreven worden. Om dit fenomeen te begrijpen, is het handig om de trilling van een snaar te ontbinden als een reeks oscillatoren van verschillende frequentie, en de koppeling te bepalen tussen de oscillatoren. In paragraaf 11.1.1 wordt dit wiskundig beschreven voor continue tijd. Deze beschrijving laat toe om de trillingsmodi en dempingen te berekenen, en om de stabiliteit van het globale systeem te beoordelen. Door enkele transformaties toe te passen op de systeembeschrijving van de digitale golfgeleidermodellen, vinden we een eenduidig verband tussen

de systemen in discrete tijd en continue tijd.

Dankzij deze theoretische beschrijving kunnen nu in **paragraaf 11.2** verschillende digitale golfgeleidermodellen met gekoppelde snaren geanalyseerd worden. Voor het triviale geval waarbij er geen snaarkoppeling is, leidt dit tot dezelfde vaststellingen als vermeld in het vorige hoofdstuk. Interessant wordt het wanneer de twee polarizaties van dezelfde snaar worden gekoppeld. Voor een gitaar met zes snaren, bekomen we zes onafhankelijke systemen, elk bestaande uit twee gekoppelde snaarmodellen. Van dit systeem kunnen de trillingsmodi berekend worden, evenals de maximale waarde van de koppelingscoëfficiënten. Twee andere in de literatuur verschenen modellen werden vergeleken: het kamgekoppeld model, waarbij de snaarkoppeling wordt bepaald door de admittantie van de kam, en een vereenvoudigd gekoppeld systeem. Beide zijn slechts benaderingen van de werkelijke situatie. Akoestisch bekeken valt een gitaar uiteen in zes zwak gedempte resonatoren (de snaren) en één vrij sterk gedempte resonator (de klankkast), waarbij alle resonatoren verbonden zijn via een eenvoudige kam. Het equivalent fysisch model is voorgesteld in figuur 11.8. Dit model heeft als groot voordeel dat de kam onafhankelijk van de klankkast gewijzigd kan worden, en waarbij de koppelingsmatrix enkel de kam voorstelt. In het specifieke geval van de akoestische gitaar heeft de kam een eenvoudige en vrij stijve struktuur, waardoor de koppelingscoëfficiënten scalaire waarden worden. Voor de akoestische gitaar bevat de koppelingsmatrix in principe 169 coëfficiënten, van elke polarizatie en de klankkast naar elke polarizatie. Dit zijn heel wat parameters, wat de calibratie bemoeilijkt. Om het aantal parameters te reduceren is een beter begrip nodig van de koppelingsmechanismen tussen dezelfde polarizaties van de een andere snaren, tussen de twee polarizaties van dezelfde snaar, tussen de ene polarizatie van een snaar en de andere polarizatie van de andere snaren, en tussen klankkast en snaren [69]. Hiervoor werd een eindige-elementen model van de kam opgesteld, beschreven in **paragraaf 11.3**. Het blijkt dat er een groot verschil is tussen de verschillende soorten koppeling, en dat vele ervan zeer zwak zijn, wat toelaat om een heel aantal coefficienten te verwaarlozen. Op basis van het eindige-elementen model werden heuristische regels afgeleid die een geldige set koppelingscoefficienten oplevert op basis van een beperkt aantal metingen.

Het resultaat van de verschillende stappen kan vergeleken worden op de bijhorende audio-CD. ⋆ **CD Track 20** is de referentie-opname van een Yamaha C70 klassieke gitaar. ⋆ **CD Tracks 21-24** zijn de verschillende deelstappen: enkele-polarizatie, enkele-snaar, dubbele polarizatie,

enkele snaar; meerdere snaren zonder klankkast; en meerdere snaren met klankkast. ⋆ **CD Track 25** is dan de uiteindelijk beste benadering, berekend na calibratie van het volledig gekoppeld model met de heuristische regels.

Hoofdstuk 12 : Klankkastmodellen

De klankkast van de gitaar is een mechanische versterker, die ervoor zorgt dat de trilling van de snaren wordt overgedragen naar de omringende lucht. De klankkast moet ervoor zorgen dat er enerzijds nog voldoende energie in de snaren blijft om een langdurige klank te bekomen, en dat er anderzijds ook voldoende wordt uitgestraald om een hoorbare toon te krijgen. De vorm en de gebruikte materialen hebben een grote invloed op het timbre en het ruimtelijk uitstralingspatroon van de klankkast. De klankkast van een akoestische gitaar is een complexe struktuur, bestaande uit verschillen dunne, min of meer flexibele platen, verstevigingsribben en diverse hout- en lijmverbindingen.

In **paragraaf 12.1** bespreken we de akoestische eigenschappen van de klankkast. De laagste trillingsmode van praktisch belang is de Helmholtz resonantiefrequentie van de ingesloten lucht. Hogere modi zijn toe te schrijven aan het bovenblad, de koppeling tussen ingesloten lucht en het bovenblad, en de trilling van de andere onderdelen. In eerste benadering kan een klankkast als een lineair systeem beschouwd worden. **Paragraaf 12.2.1** behandelt verschillende lineaire modellen. De eenvoudigste techniek maakt gebruik van een FIR filter, dat noodzakelijk duizenden taps groot is. Om de lengte van het filter te verkleinen, kunnen de dominante trilligsmodes gemodelleerd worden met lage-orde IIR filters, en de resterende modes door een korter FIR filter. Dit kan doorgetrokken worden tot alle trillingsmodi door gescheiden tweede-orde IIR filters worden voorgesteld. Deze kunnen parallel werken, of tot de transversale Kautz-struktuur getransformeerd worden.

De klankkast van een gitaar is slechts lineair voor kleine uitwijkingen. Voor grotere amplitudes kan er door de niet-ideale eigenschappen van de gebruikte houtsoort en de struktuur van de verstevigingsribben, verzadiging optreden van de klankkast als mechanische versterker. Dit effect was kwalitatief reeds gekend bij de gitaarbouwers en bepaalt de houtkeuze en de muziekstijl waarvoor een bepaalde gitaar geschikt is. We poneren dat de Wiener-Hammerstein struktuur geschikt is om de klankkast te modelleren [76]. Het Wiener-Hammerstein model bestaat uit een lineaire gedeelte dat de trillingsmodi van de klankkast omvat, en een statische niet-lineariteit die de saturatie van de bovenplaat voorstelt. De

xxxiii

traditionele meetmethodes voor niet-lineaire systemen vereisen het aan-
leggen van specifieke signalen, wat moeilijk haalbaar is bij de meting
van een gitaar. Daarom werd een eenvoudige, benaderende methode
uitgewerkt waarbij een reeks impulsen van stijgende amplitude wordt
opgenomen.

Hoofdstuk 13 : Validatie

In dit hoofdstuk wordt de validatie van de besproken modelleringstech-
nieken en de experimentele bepaling van de modelparameters besproken.
Hiervoor werd een meetopstelling opgebouwd rond een Taylor 514CE
akoestische gitaar met bronzen snaren. Deze gitaar is uitgerust met een
piëzoëlektrische sensor onder de kam, en een elektret microfoon inge-
bouwd in de klankkast. Daarbij werd nog een zeskanaals magnetische
pickup met bijhorende versterker ontworpen om de trilling van de indi-
viduele snaren te kunnen registreren [72]. Deze meetopstelling liet toe
om de koppeling van een snaar naar een andere in detail te bestuderen.
Toepassing van de parameterestimatie algoritmes uit hoofdstuk 10 en
de heuristische regels afgeleid in hoofdstuk 11 zijn voldoende om een
realistisch model af te leiden.

Hoofdstuk 14 : Implementatie

In dit laatste hoofdstuk wordt de implementatie van het volledig ge-
koppeld gitaarmodel. De MPEG-4 SAOL implementatie bestaat uit
drie grote delen: een partituur-vertaler, die de inkomende MIDI no-
tatie vertaalt naar meer gedetailleerde SASL commando's, de eindige-
toestandsmachine, die de SASL-commando's omzet in reeksen fysische
parameters, en de signaalverwerkingsmodule, die zorgt voor de eigenlijke
simulatie [75].

Door de polyfone aard van de gitaar is er een vertaling nodig van parti-
tuur tot de fysische aansturingsparameters van het model. Eén van de
problemen die hierbij opgelost moeten worden, is de toekenning van een
noot aan een bepaalde snaar. Dit is een inherent niet-causaal probleem,
aangezien de snaarkeuze op een bepaald tijdstip invloed heeft op de mo-
gelijke noten die op een later tijdstip gespeeld kunnen worden. Het is dus
noodzakelijk om een eenvoudig spelermodel te gebruiken. Er zijn nog
heel wat andere parameters die semi-automatisch bepaald moeten wor-
den, zoals de aanslagplaats, het type tokkel, en de aanslagsnelheid. De
uiteindelijke implementatie is in staat om een partituur op een redelijk
natuurlijke wijze af te spelen. De mogelijkheden van deze implementatie
zijn te beluisteren op ⋆ **CD tracks 30** en **31**. Het volledig gekoppeld

gitaarmodel wordt dan vergeleken met vier andere fysische modellen, een spectraal model, en wavetable synthese.

Hoofdstuk 15 : Besluit en suggesties voor verder onderzoek

Besluit

Dit proefschrift bestaat uit verschillende delen: **Deel I** handelt over algemene akoestiek, **Deel II** over de blaasinstrumenten, en **Deel III** over de snaarinstrumenten.

Deel I omvat een algemene bespreking van de blaasinstrumenten en de snaarinstrumenten, een inleiding tot de muzikale akoestiek, en de algemene uitwerking van de digitale golfgeleidermodellen.

We begonnen door een kwalitatief overzicht te geven van de blaasinstrumenten in **hoofdstuk 2**. Twee enkel-riet instrumenten (de klarinet en de saxofoon), een luchtstroom-aangedreven instrument (de moderne dwarsfluit), en een lip-aangedreven instrument (de trompet) werden besproken. Er werd besloten dat al deze instrumenten een zelfde basisstruktuur hebben, en met hetzelfde type model kunnen worden beschreven. **Hoofdstuk 3** gaf een kwalitatief overzicht van de akoestische eigenschappen van de snaarinstrumenten. We vergeleken drie verschillende types instrumenten: de akoestische en elektrische gitaren, de vioolfamilie, en de piano. Het blijkt dat de snaarinstrumenten slechts verschillen in de wijze van snaarexcitatie. In **hoofdstuk 4** brachten we een overzicht van de muzikale akoestiek. Vertrekkend van de algemene golfvergelijking, die de golfvoortplanting door lucht en snaren beschrijft, werden de eigenschappen van de basisbouwblokken van de muziekinstrumenten afgeleid. De verschillende elementen werden dan gecombineerd tot akoestische systemen. In **hoofdstuk 5** werden de akoestische elementen gediscretiseerd, gebruik makend van de digitale golfgeleidertechniek. We bespraken het belang van fractionele vertraging, en toonden aan dat Catmull-Rom en B-Spline interpollatie mogelijke alternatieven zijn voor Lagrange interpollatie en allpass filtering. Deze hoofdstukken vormen de basis waarop alle verdere modellen gebouwd werden.

In **Deel II** werden de eigenschappen van de blaasinstrumenten bestudeerd. Hiervoor werden specifieke modellen en parameterestimatietechnieken afgeleid.

In **hoofdstuk 6** werd een volledig akoestisch model voor een enkel-riet blaasinstrument opgebouwd, met verwaarlozing van de toongaten.

De resonator van de klarinet en de saxofoon werd besproken, en er werd dieper ingegaan op de eigenschappen van het excitatiemechanisme. De beweging van het riet werd beschreven door hogere-orde lineaire en niet-lineaire vergelijkingen. De belangrijkste bijdrage voor de blaasinstrumenten is vervat in **hoofdstuk 7**, waarin we een digitaal golfgeleidermodel met formant filter opbouwen. Door de combinatie van bron-gefilterde synthese en de traditionele digitale golfgeleider technieken werd een model bekomen dat zowel performant als eenvoudig te calibreren is. Het model bestaat uit een niet-lineair excitatiemechanisme, een lineaire resonator, en een dynamisch post-equalizatiefilter of formant filter. Verscheidene filter-ontwerptechnieken werden dan vergeleken. Door het toevoegen van toongatmodellen in **hoofdstuk 8** werd een volledig model voor de klarinet bekomen. In **hoofdstuk 9** werd de MPEG-4 SAOL implementatie van het model voorgesteld.

Deel III handelt over de snaarinstrumenten.

In **hoofdstuk 10** werd een enkel-snaar model opgebouwd voor de akoestische gitaar. De eigenschappen van ideale en reële snaren werden besproken, met inbegrip van niet-lineaire effecten als de spanningsgemoduleerde propagatiesnelheid en de beperkte snaarbeweging. Met de dubbele-polarizatie modellen werd rekening gehouden met de beweging van de snaren in twee ruimtelijke dimensies. In dit hoofdstuk werden ook de parameterestimatie methodes uitgewerkt in zowel het tijd- als het frequentiedomein. Er werd aangetoond dat een subband-gebaseerd algoritme, de subband Hankel Singuliere Waarden Ontbinding, in staat is om nauwkeurige schattingen te bepalen van de systeempolen. De enkelsnaar modellen werden in **hoofdstuk 11** uitgebreid tot gekoppeldesnaar modellen. De ontwikkelde methodologie laat toe om alle digitale golfgeleidermodellen voor snaarinstrumenten te vergelijken. Op basis van de theoretische uitwerking en een eindige-elementen analyse werd een nieuw gitaarmodel voorgesteld. Een belangrijk deel van dit model zijn de lineaire en niet-lineaire modellen voor de klankkast die in **hoofdstuk 12** behandelt werden. Voor de uiteindelijke implementatie werd in **hoofdstuk 14** een eenvoudig spelermodel uitgewerkt.

Suggesties voor verder onderzoek

In dit proefschrif ging het grootste deel van de aandacht naar de theoretische uitwerking van de modellen. Zowel model-gerelateerde problemen als verschillende aspecten van de implementatie vereisen verder onderzoek.

Bij de blaasinstrumenten zouden de verschillende strukturen voor het formant filter vergeleken moeten worden vanuit het standpunt van de implementatie. Aangezien dit dynamische filters zijn, moeten transiënte verschijnselen geminimaliseerd worden. Dit vereist het nauwkeurig onderzoek van de filterstruktuur en de effecten van berekeningen met eindige woordlengte. Het niet-lineaire excitatiemechanisme kan leiden tot hoorbare aliasing indien de laagdoorlaatkarakteristiek van het reflectie-filter een the hoge afbreekfrequentie heeft. Overbemonstering en steilere filters kunnen dit probleem oplossen.

Bij de snaarinstrumenten ligt het grootste probleem in de bepaling van de vinger-snaar interactie. Er bestaat zeert weinig literatuur over de fysische processen die plaatsvinden bij het tokkelen, en professionele spelers kunnen geen verklaring geven voor de veranderingen in timbre na jaren oefening. De gekoppelde-snaren modellen kunnen toegepast worden in veel grotere systemen, zoals de harp of de piano. Het kan ook nuttig zijn de digitale golfgeleidermodellen te koppelen aan twee-dimensionale digitale golfgeleider-roosters. In het algemeen kunnen hoger-dimensionale modellen bepaalde problemen effectiever oplossen.

De technieken die in dit proefschrift werden ontwikkeld, kunnen ook toegepast worden om de kwaliteit van circulaire golfgeleiders en golfgeleider-roosters te verbeteren, met als mogelijke toepassingen digitale reverberatie en virtuele akoestiek.

Chapter 1

Introduction

In the first section of this introductory chapter an economic and scientific motivation will be given for the modeling techniques that will be developed in this thesis. In section 1.2, we review the history and taxonomy of computer music and sound synthesis, and indicate the current trends and future possibilities. The core of the thesis is based on the physical modeling sound synthesis techniques developed in the last fifteen years, and is detailed in section 1.2.1. The recent developments of this technique are closely linked to the increase of the available computational power and the appearance of specific sound synthesis programming languages, outlined in section 1.4. Finally, an outline of the thesis and a highlight of the contributions is presented in section 1.5.

1.1 Problem statement

Sound synthesis historically started with the wish to generate any kind of sound using mathematical techniques. As the research evolved, it became possible to generate or transform any sound conceivable. Synthesis techniques are not only used for live music performance, but are also massively used in the entertainment industry, mostly in computer games, and for the design of sound effects and music scores for television broadcasting or feature films. This is illustrated by the healthy economic numbers presented by the industry. For instance, one vendor of electronic musical instruments reported net sales of $2.15 billion, and one major vendor of PC soundcards reported $500 million net sales, with a total of 25 million soundcards installed, for the fiscal year 2001 [25], [56].

Is there a need for more synthesis techniques ? Yes, as the current techniques are capable of the perfect *reproduction* of a sound, but not of the perfect *gener-*

ation of the sound such that the performer can consciously control some aspects that go beyond mere manipulation. This is one of the reasons why although there exist many different synthesis techniques, only a few are actually commercially available on the consumer market. There is a progression toward more interactive and model-based digital instruments. Due to its nature, model-based synthesis gives the player more control over the physical variables that influence the sound production, while non-model-based synthesis uses abstract and meaningless parameters.

1.2   Computer music and sound synthesis

In 1948 Paris, history was made. Pierre Schaeffer, a French radio broadcaster, working for the Radiodiffusion-Television Francaise (RTF), created the first electronic music studio. With a multitude of microphones, phonographs, variable speed tape recorders and sound effect records he created a new art form, *musique concrete*, and with it a world of new music opened up - the world of electronic music.

Electronic music can be divided into three categories: Musique concrete, Synthesizer music, and Computer music. Musique concrete was the first type to be created. It involves using the found sounds in nature, distorted in various ways, to create music. Live, it becomes an exercise in mixing together unexpected sounds into some sort of form while studio musique concrete uses complex tape manipulations to create the effect. Synthesizer music is qualified by the electronic generation of the sounds, and computer music adds compositional aspects.

Automatic generation of sound and music is centuries old: the first mechanical instruments date back to the 2nd century BC, and mechanical, score playing instruments were in use in the late 16th century. During the Industrial Revolution, inventors and composers designed steam-powered and even electrical instruments. The first device to actually synthesize sound was the Dynamophone, also known as the Telharmonium. The instrument weighed over 200 tons and produced music by an alternating current running dynamos. The first truly electronic synthesizer, the Autherophone, was built by the Russian Lev Theremin, and is now known as the "Theremin". The instrument uses 2 vacuum tube oscillators to produce beat notes. Musical sounds were created by "heterodyning" from oscillators that varied pitch. A circuit was altered by changing the distance between 2 elements. The instrument had a radio antenna to control dynamics and a rod sticking out the side that controlled pitch. The performer would move his/her hand along the rod to change pitch, while simultaneously moving his/her other hand in proximity to the antenna. Many composers used this instrument including Edgard Varese (1883-1965).

In 1929, the Hammond organ was launched, and became a very successful electromechanical additive synthesizer. Many other instruments were designed, and used by composers the likes of John Cage (1912-1992), Karlheinz Stockhausen (1928-) or Iannis Xenakis (1922-2002). Synthesizer music and later on, computer music, closely followed the technological advances in electronics. In 1958, Max Matthews generated music by computers.

During the 50-year history of digital sound synthesis, the preferred synthesis methods closely followed the available computing power [102]. In the early days, only extremely simple additive synthesis was possible using a digital computer. At this point, the sound synthesis community started using two different approaches: the real-time, but simple, sound generation, and the non-real-time, complex, and programmable approach using MUSIC-N languages. Real-time synthesis was first dominated by abstract, mathematical synthesis methods, like FM synthesis. By the end of the 1980s, wavetable synthesis became more widespread and in the last decade of the 20th century, we saw the emergence of more powerful synthesis methods like physical modeling and virtual analog modeling. At this point, the hardware became fast enough to run even very complex algorithms in real-time: the "software" synthesizer was born.

1.2.1 Physical Modeling sound synthesis

Physical modeling sound synthesis is a class of synthesis techniques that uses some form of simulation of the physical properties of an object (mainly musical instruments) to generate sound. Physical models go from the extremely detailed finite element analysis to the rough approximations of the dominant vibration modes of the instrument. It is a very interdisciplinary research field, with links to physical acoustics, music, signal processing, system theory and system identification. Figure 1.1 shows how the different aspects are represented in the different chapters of this work.

Several approaches to physical modeling sound synthesis exist. The acoustic approach builds a model based on the physical properties of the different elements of the musical instrument, resulting very complex models with good predictive properties. Due to their high complexity, these models are less useable for real-time synthesis, and difficult to calibrate such that their output match a recording. At the other end, black box models use parametric system identification techniques to model te output of the instrument. This model class requires specialized equipment for the model calibration, and might result in non-physical control parameters. The approach taken in this work is an integral view, assuming that certain properties are common to all instruments of the same family (wind instruments, string instruments), and adding parameters that enable model calibration with only a recording of notes played on the instruments.
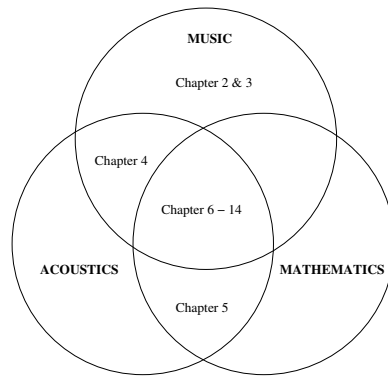
Figure 1.1: Physical modeling sound synthesis is a very interdisciplinary research field. This figure shows the interaction of the different disciplines in the chapters of this thesis.

1.3 Sound synthesis applications

Sound synthesis is widely used in contemporary music, be it as a post-processing step for conventionally generated sound, as an instrument as such, or as a tool to extend the capabilities of a human player.

1.3.1 Score-based synthesis

A score is a symbolic representation of music, and thus a dynamically evolving language. As can be seen on figure 1.2, modern compositions require more complex notation techniques to convey the intentions of the composer. The interpretation of such a score by the performing musicians is not trivial. Some composition can only be played by machines, as their complexity surpasses anything a human is able to absorb.

1.3.2 Transhuman performance

The trend of composers challenging performers seems a constant in musical history [19]. The ever increasing demand in motor skill and concentration will at some time be of reach of even the most trained performers. There are two areas where transhuman performance is necessary: the maximum (micro) repetition rate for a human player seems to be around 30 Hz, implying that all compositions with faster notes are unplayable, and secondly, there is a limit on the (macro) movement that can be made by the human body. For instance, playing three D's of the same pitch, on three different string of a classical guitar
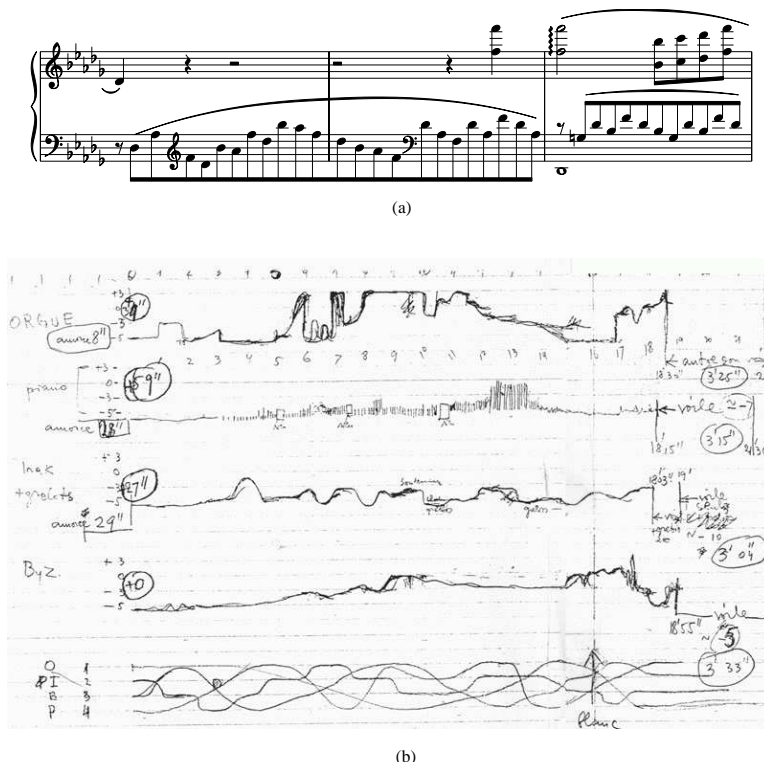
(a)

(b)

Figure 1.2: Two examples of musical scores. (a) conventional musical writing, using bars, clefs, and notes. Excerpt from a composition by Franz Liszt (1811-1888), (b) sketch of the sound contour of *Bohor* by Iannis Xenakis (1962). Evolving composition techniques require an ever increasing set of complex symbolic notation.

at the same time is physically impossible: the human hand cannot be stretched to span 40 cm! In this area, sound synthesis offers unparalleled performance. ⋆ **Track 31** on the CD is an example where exactly this property is used in a composition.

1.3.3   Parametric sound compression

Parametric sound compression is a fairly new field, that tries to compress sound by very efficient coding of the sound. This coding can be done from the point of view of the human perception (what do we hear ?) or the sound source (what is emitting sound?). The most compact representation of a sound is a model of the instrument and the score played on it. The problem of going from sound to

source model is however as yet unsolved. In this work, we present calibration and parameter estimation methods that allow the modeling of a restricted set of instruments.

1.4   Sound synthesis languages

Several types of programming languages are available for the implementation of the algorithms on general-purpose processors. During development, a high-level language like Matlab, dedicated to digital signal processing is well-suited. Final implementations are best programmed using either a generic programming language or a *structured audio* language, specifically designed for sound synthesis.

1.4.1   Development

The initial implementation was programmed in the Matlab environment. Matlab uses a high-level interpreted script language, which allows the fast implementation of changes in the core algorithm, but runs quite slow. Both sample-based processing and block-processing versions were written. A faster version, with compiled core routines, is also created by first identifying the main synthesis loop, eliminating or replacing slow functions using the Matlab profiler, and finally translating to C and compiling the core functions.

1.4.2   Generic programming languages

The use of a generic programming language like `C` or `C++` requires more effort and is only justified when the algorithm development is frozen. None of the generic languages understand the concept of time "out-of-the-box" as needed for sound synthesis, and are thus better suited for offline rendering. Toolkits and libraries exist that are dedicated to sound synthesis, and allow the use of time. The STK toolkit [22], [24] is a set of `C++` classes developed for musical signal processing. The Synthesis ToolKit offers an array of unit generators for filtering, input/output, etc. and uses Tcl/Tk for GUI interaction. The generic programming languages are often used as an intermediate step when compiling a structured audio language.

1.4.3   Structured Audio programming languages

Structured audio representations are semantic and symbolic descriptions that are useful for ultralow-bitrate transmission, flexible synthesis, and perceptually based manipulation and retrieval of sound [112]. This concept has a much

broader application than required for the pure implementation of the physical models developed in this work. Many computer languages for specifying sound-synthesis algorithms have been developed [34]. Some are text-based, like the popular Csound [111] language, the LISP-based Common Lisp Music (CLM) or the object-oriented SuperCollider [59]. Some use a graphical user interface, like `pd` (Pure Data) [87], MAX/MSP [86] or jMAX [27]. Most of these languages are based on unit generators, elementary functional blocks like oscillators, filters, and envelopes that may be connected into networks to describe a signal flow path. Most descend from the MUSIC IV and MUSIC V languages developed in the 1960's by Max Mathews [58], [91]. Recently, a structured audio language was standardized as a part of the MPEG-4 ISO standard. We selected this language for the core development, as its inclusion in a standard ensures widespread implementation and use.

1.5 Outline of the thesis and contributions

In this section, an outline and an overview of the thesis can be found. The main contributions are indicated, and the publications relating the different subjects are referenced.

1.5.1 Objectives and contributions

In this thesis, we build an integral solution, from physics to implementation, for the modeling of both wind and string instruments. The focus is not on the physical accuracy, but on the practical usability and calibration properties of the model. The calibration of the models should be possible without specialized equipment. This has required a survey of the literature and an evaluation of existing models and calibration techniques. The existing modeling and parameter estimation techniques were found not to be satisfactory, as they either require specialized equipment to measure the model parameters, do not allow extensive control by the player, or produce inferior sound quality. New models and techniques have been developed.

Quite an extensive body of literature exists on the subject of musical acoustics, but the digital waveguide techniques do not enjoy this position. The basic elements of digital waveguide modeling are delay lines, fractional delay filters and reflection filters. The traditional fractional delay filters aere evaluated and compared to two other interpolation algorithms.

The calibration of wind instruments models, with their inherent non-linear behavior, has required the design of the formant filtered models with better calibration properties. A more in-depth study has been performed on the properties of the non-linear excitation mechanism. Note transients have been

implemented using a parallel crossfading technique.

The string instruments present a large number of closely spaced system poles due to their physical structure and string coupling mechanisms. First, a mathematical technique for the identification of the system poles has been introduced, and a framework for N coupled strings and N coupled digital waveguides has been developed. Analytic and heuristic parameter estimation rules have been derived, and experimentally validated.

Both the wind and string instrument models are finally implemented using the MPEG-4 Structured Audio language.

1.5.2   Overview of the chapters

The thesis consists of three parts, each of them divided in several chapters. Figure 1.3 shows the schematic overview of the different parts. Note that the introductory chapter and the general conclusions are omitted in this picture. The thesis is accompanied by a CD, containing sound examples. Every part has the same basic structure, with chapters focused on the musico-acoustical description of instruments, continuous time or discrete time modeling techniques, and implementation issues. Figure 1.4 shows which chapters fall into each category. The iconic representations will be used throughout this work to indicate the focus of the chapter.

Part I: General Acoustic Theory

In the first part of the thesis, the basic building blocks for physical modeling synthesis are provided. Both continuous time and discrete time models are discussed for several musically relevant structures.

Chapter 2 provides an introduction to the structure and properties of the clarinet, the saxophone, the modern flute, and the trumpet. It is concluded that they all share the same basic structure, and that a general digital waveguide model can be used to model all the instruments.

Chapter 3 gives an overview of the acoustic properties of string instruments, in general. Three different types of string instruments are compared: the acoustic and electric guitars, the violin family, and the piano.

Chapter 4 provides an introduction to linear musical acoustics. Starting from the general equations describing sound propagation in air and strings, the properties of common acoustic building blocks of musical instruments are derived. The different elements are then combined into acoustic systems.
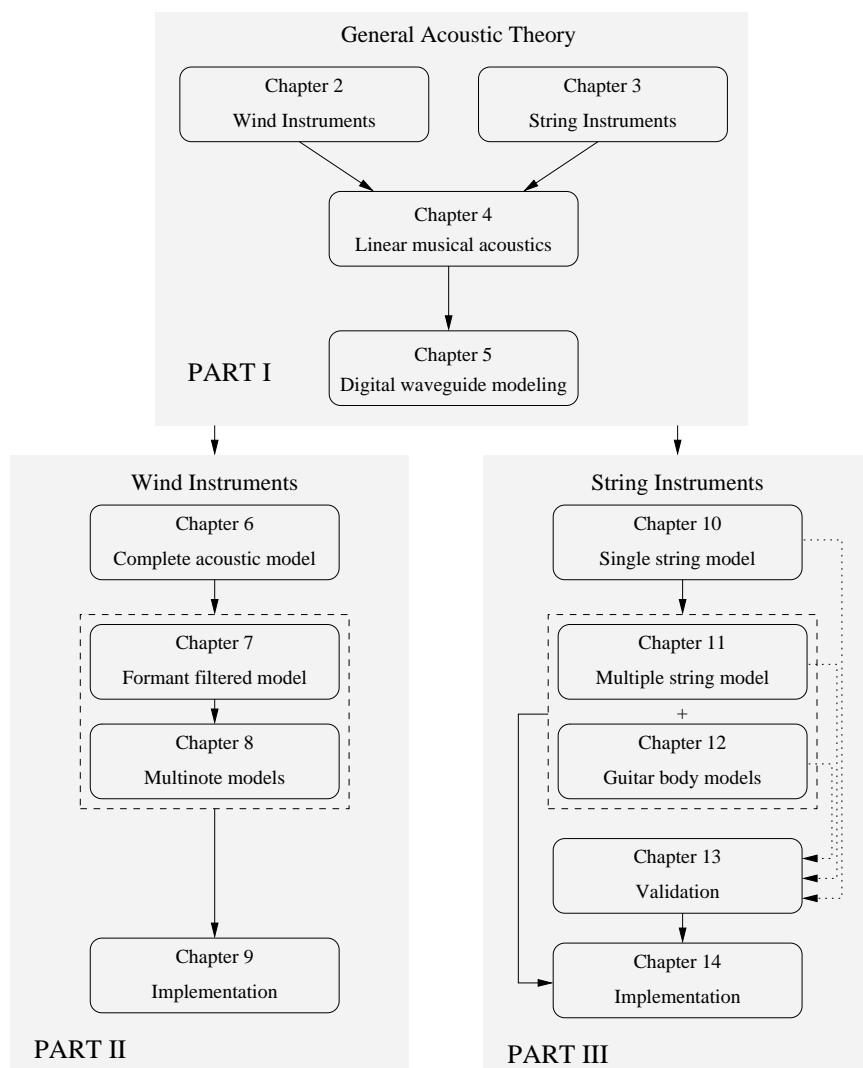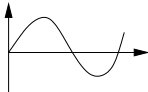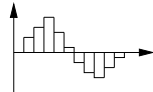
Figure 1.3: Schematic overview of this thesis.

	♫	sine wave	discrete wave	`for i=1:100`
				`{`
				`   ...`
				`}`

Acoustics	Chapter 2 Chapter 3	Chapter 4	Chapter 5	
Wind Instruments		Chapter 6	Chapter 7 Chapter 8	Chapter 9
String Instruments		Chapter 10	Chapter 11 Chapter 12 Chapter 13	Chapter 14

Figure 1.4: This figure shows the classification of the chapters according to the modeling level. From left to right, there is the qualitative musico-acoustical description of instruments, modeling in continuous time, modeling in discrete time, and the implementation issues.

In **chapter 5**, the acoustic elements are discretized using the digital wave-guide modeling technique. The importance of fractional delay is discussed, and it is showed that Catmull-Rom and B-spline based interpolation are possible alternatives to Lagrage interpolation and allpass filtering.

Part II: Wind Instruments

Part II is dedicated to the modeling of the wind instrument family. After an overview of a few common wind instruments, source-filter synthesis and digital waveguide modeling are combined into the formant filtered physical models. The single-note models are linked using a parallel crossfading technique, and a final SAOL implementation was made.

In **chapter 6**, a complete acoustic model is built for a single-reed instrument, disregarding the influence of the toneholes. The resonator of a clarinet and a saxophone is discussed, and we elaborate more in detail on the excitation mechanism. High-order linear and non-linear models are introduced to describe the reed movement.

The core of Part II is contained in **chapter 7**, where a discrete time model of a single-reed instrument is built. Source-filter synthesis is combined with the traditional digital waveguide modeling technique and a new model class is obtained that is both easy to calibrate and flexible from a performance point of view.

The models are then expanded into multinote models in **chapter 8**. Several structures are compared, and we propose a parallel crossfading method to simulate the transition between notes.

Finally, **chapter 9** discusses the actual implementation of the models in Matlab, and using the MPEG-4 SAOL language, and provides a comparison of our model with other synthesis techniques.

The publications relating to this part are [68], [67], [74].

Part III: String Instruments

In Part III, the modeling of the acoustic guitar is discussed. This instrument was chosen because presents many interesting properties of a generic string instrument, while having only a limited number of strings. First, a model and its associated parameter estimation routines for a single string are built, later expanded for multiple strings.

In **chapter 10**, a single-string model for the acoustic guitar is presented. These models are the basis for the more complex coupled string models. The focus of this chapter is on the parameter estimation. Different methods are compared, and it is shown that the subband Hankel Singular Value Decomposition offers both high accuracy and automated estimation.

In **chapter 11**, a novel general framework that allows the performance comparison and design of coupled string structures, is proposed. The framework indicates how to improve the parameter estimation for the structures and how to verify the stability. The properties for N coupled strings and N coupled digital waveguides are derived, and a comparison is made of five coupled string structures. Finally, a fully coupled digital waveguide model is proposed for the acoustic guitar.

The body of an acoustic guitar is the main factor determining the timbre a the instrument. In **chapter 12**, we discuss several linear modeling techniques, and propose a non-linear model for the guitar body.

Chapter 13 elaborates on the experimental measurement of the parameters for a simplified and a fully coupled digital waveguide guitar model. We show that the fully coupled model procudes similar output with better performance possibilities.

The implementation of the fully coupled model is described in **chapter 14**. We show that a player model is necessary, and present a "black box" MIDI/SAOL implementation.

The publications relating to this part are [69], [71], [72], [73], [76], [75].

The main contributions of this thesis are on the subject of the wind instruments, the combination of source-filter synthesis and physical modeling into the formant filtered digital waveguide models, with the associated calibration techniques; and on the subject of the string instruments, the elaboration of a framework for coupled digital waveguides, with the associated parameter estimation methods, illustrated by the fully coupled guitar model. The models proposed in this thesis offer the best currently available balance of model complexity, sound quality, control, and calibration.

Part I

General Acoustic Theory

Chapter 2

Wind Instruments

This chapter provides a short introduction to the acoustic structure of selected wind instruments and gives a qualitative description of their principles of operation.

Wind instruments produce a sustained oscillation as long as the player excites it directly or indirectly (e.g. using an airbag in the case of bagpipes). The oscillation is typically produced in an air column, within a resonating structure, composed of cylindrical and conical sections. Different notes are obtained by modifying the excitation or the resonator, or both. The wind instruments can be classified by the excitation mechanism used: single reed (clarinet, saxophone, etc.), double reed (oboe, bassoon, etc.), air jet (flutes, organ, etc.) or the player's lips (trumpet, trombone, etc.) [12]. The acoustical analysis of the bore shape and the excitation mechanism provides enough information to build a usable digital model.

Woodwinds have slowly evolved over the centuries, and the resulting peculiar features give them their individuality and musicality. Two notable exceptions (in the sence that they were designed) are the saxophone family (designed by the Belgian Adolphe Sax) and the modern flute (designed by Theobald Boehm), both developed at the middle of the nineteenth century. The brass instruments followed a similar development. We selected the clarinet and saxophone (single-reed woodwinds), the modern flute (air jet woodwind), and the trumpet (lip-driven brass) to illustrate the differences and similarities between the instrument families. Of these, only the clarinet and saxophone will actually be modeled in later chapters, but the same basic principles can be applied to
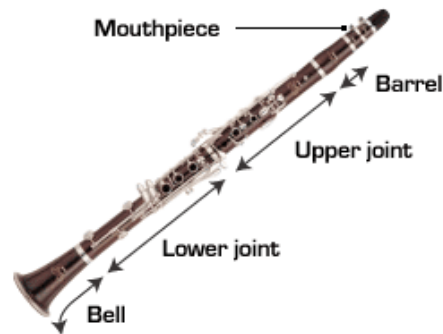
Figure 2.1: The structure of a Bb clarinet. Concert clarinets are traditionally made of ebony and consist of five parts. The lower joint, the upper joint and the barrel together form the resonator of the instrument. The bell acts as the sound emitter, most importantly so for the low notes. The mouthpiece with attached reed is the excitation mechanism. The Boehm Bb clarinet has 17 keys, 6 rings, and is approximately 635mm long. Other fingering systems and tunings also in use.

the others as well.

2.1   The clarinet

The clarinet is essentially an instrument with a cylindrical bore and a single reed. It can be traced back to the late seventeenth century, when it was known as the *chalumeau*. Its relative simplicity makes it an ideal candidate for acoustical study and digital modeling.

The clarinet's resonances are primarily determined as those of a closed-open cylinder, having odd harmonics[1] in the lowest or chalumeau register. Such a bore overblows in the twelfth[2], and the instrument requires many toneholes to be able to play a complete scale. Figure 2.1 shows the structure of a modern Boehm $B\flat$-tuned clarinet. To access the higher or clarion register, register holes were added to the design. The clarion register notes include even and odd harmonics. Acoustically, the clarinet consists of three major parts: the mouthpiece, the resonator (barrel, upper and lower joint), and the bell . The

---

[1]The peaks in the spectrum of the sound produced by an instrument are called *partials* in the general case. *Harmonics* is used when the partials are an integer multiple of the fundamental frequency. The terms "partial" and "harmonic" are often used interchangeably.

[2]Three times the fundamental frequency, spanning 19 semitones. Doubling the fundamental frequency yields the *octave*, or 12 semitones.
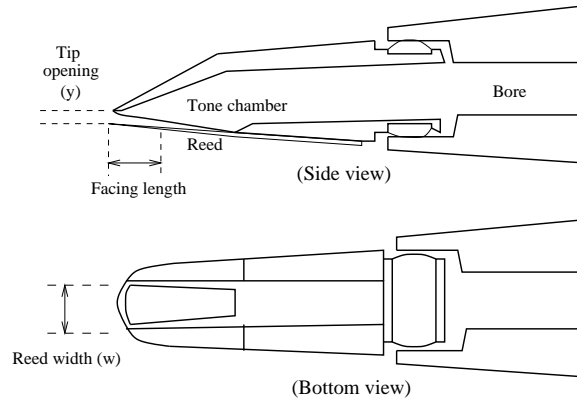
Figure 2.2: The clarinet mouthpiece and reed assembly (side view and bottom view). The reed opening is determined by the pressure difference inside and outside the mouthpiece. The opening at rest and the facing length are chosen by the player by moving the player's lower lip. The reed width is fixed.
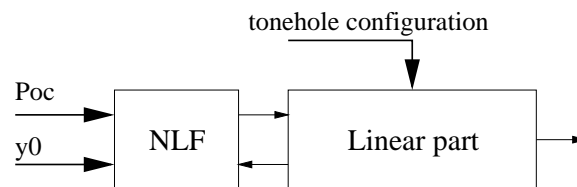

Figure 2.3: The three main parameters controlled by the player during performance are the oral cavity pressure P_{oc}, the tip opening at rest (y_0) and the configuration of the toneholes (the fingering pattern).

Figure 2.4: Comparison of three common woodwind instruments: the modern flute, the clarinet and the oboe. The flute and the clarinet have a cylindrical bore, while a conical bore is used for the oboe. Observe that the three instruments have approximately the same length.

mouthpiece/reed assembly (figure 2.2) is the excitation mechanism. The cane reed combined with the shape of the mouthpiece acts as a pressure-controlled valve. The bore with the toneholes is a resonator. Depending on the configuration of the instrument, certain wave propagation modes are allowed. The bell is the impedance matching device between the surrounding air and the instrument bore, and has a significant effect on the radiation of the lower partials, particularly for low notes. For the higher notes, most radiation comes from the toneholes.

The pressure difference between the oral cavity of the player (p_{oc}) and the pressure inside the mouthpiece induces a movement of the reed. The air flow into the mouthpiece is determined by this pressure difference and the tip opening by aerodynamic laws. This injection creates a pressure wave that propagates into the bore of the instrument. The bore acts as a filter on the traveling wave. When the wave reaches the bell, part of its energy is radiated outwards and part is reflected. The reflected wave travels back to the mouthpiece where it interacts again with the reed. This procedure repeats itself and creates a continuous pressure oscillation. The oscillation exhibits several interesting properties, including period doubling and chaotic modes [93].

The player can select different reeds and mouthpieces with different properties. During performance, the player controls the oral cavity pressure and the displacement of the reed at rest (y_0), as well as the configuration of the toneholes (Figure 2.3). In a lesser extent, the player is able to modify the damping of the reed movement. Figure 2.4 shows a comparison of the bores of a clarinet, a modern flute and an oboe.
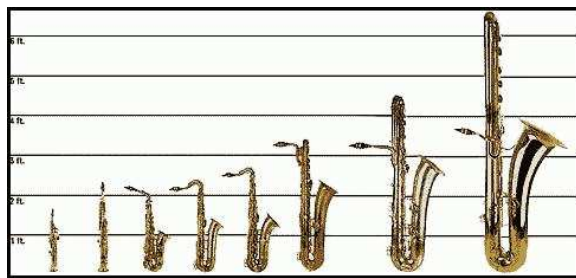
Figure 2.5: A comparison of the complete saxophone family of instruments. From left to right: Eb sopranino, Bb soprano, Eb alto, Bb tenor, Eb baritone, Bb bass and the extremely rare Eb contrabass.

2.2 The saxophone

The family of the saxophones, developed by Adolphe Sax in the mid-nineteenth century, have a wide conical bore, a single clarinet-like reed, and a unified logical system of padded keys covering very large toneholes. The saxophone was designed, and the family has sopranino, soprano, alto, tenor, baritone, bass and contrabass members, shown in figure 2.5. The more exotic sopranino, bass, and contrabass saxophones are rarely used. There exist a few instruments tuned in another key (e.g. the C tenor saxophone). A saxophone has 18-21 toneholes and a corresponding number of keys, actuated by the fingertips or the side of the hands. It has an intricate mechanical construction, that automatically opens or closes combinations of toneholes: the saxophone has more than 300 separate pieces! The saxophones, having a conical bore, overblow in the octave and produce a harmonic-rich spectrum. The structure of an Eb alto saxophone is shown in figure 2.6.

2.3 The modern flute

Instruments of the flute type are of extremely ancient origin, the oldest found dating back to the early Neolithic (8000 BC-5000 BC) [121]. Their principal modern descendants are the orchestral transverse flute and the flue organ pipe. The recorder, panpipes and the shakuhachi are other examples. They all rely on the effect of an air jet striking a hard edge for their acoustical excitation, coupled to a resonating structure. The description of an air jet is non-trivial, and not studied in this work [113], [114]. The essence of the interaction between an air jet and a resonator is shown in figure 2.7. The jet emerges as a plane sheet from a slit or the player's lips, crosses an opening in the resonator

Figure 2.6: The structure of an alto saxophone. The saxophone has 18-21 toneholes, as many keys and has a conventional playing range of about 2 1/2 octaves
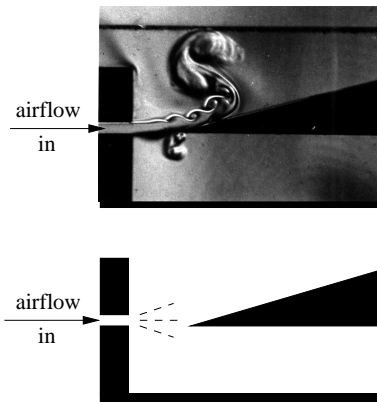

Figure 2.7: Typical air jet used with a recorder-type musical instrument (top) Schlieren photograph of the airflow in a flue organ pipe [114] (bottom) schematic view. The air jet has a laminar flow when entering the cavity. After a short distance, the flow exhibits sinusoidal and later turbulent flow patterns. The sinoidal flow at the sharp interface excites the resonator.

Figure 2.8: The structure of a modern flute. The flute is a closed-open cylinder of about 637mm length, with 18 open or closed keys.

and strikes a sharp edge on the far side. The distance crossed is such that for a given volume flow, the jet is on the edge of becoming turbulent. Depending on some jet parameters, a small disturbance can grow either sinuous or varicose. The resonator of the flute favors one particular frequency of the sinuous disturbances. Since the deflection of the jet is the acoustic driving mechanism and since it is driven by the acoustic flow out of the mouth of the resonator, the system will work best when its flow is maximum or in other words at an impedance minimum. This is in contrast to the single-reed instruments, which operate at the impedance maxima.

Figure 2.8 shows the structure of a modern transverse flute. It consists of a cylindrical bore, stopped at one end, with a blowing hole close to the stopped end and finger holes along its length.

2.4 The trumpet

The trumpet is an example of the lip-driven musical instruments. Historically, these instruments have existed since the dawn of mankind. Early instruments had a more-or-less conical shape (dictated by their origin as seashells or animal horns) or a nearly cylindrical shape (as the Australian aboriginal didgeridoo). In the medieval times, the small conical instruments were provided with side holes, to allow production of notes in between the modes of the complete horn. The other line increased the length of the instrument, and used narrower bores. The length of the bore could be modified by a system of slides or pistons. The modern trumpet is a result of this evolution. For several, mostly practical, reasons, the instruments are built with a cylindrical bore, connected with a more-or-less conical section of comparable length and terminated with a short section of more rapid flare. The main parts of a modern trumpet can be found on figure 2.9. A close approximation to the shape of a real brass instrument is given by the Bessel horn. A main problem with brass bore shapes is that the partials do not form a good harmonic series. Modifications to the bore shape lead to a more-or-less harmonic series $(0.7, 2, 3, \ldots)f_0$. The fundamental mode is very much out of alignment, produces a very weak sound, and is usually
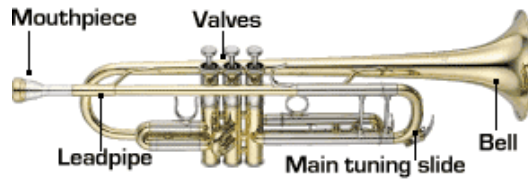
Figure 2.9: The structure of a Bb trumpet.

not used in playing. Combined with the input impedance of the mouthpiece, and the admittance of the player's lips, one can calculate the resonances of the instrument. The lips form, just as with the single-reed instruments, a pressure-controlled valve, but with inverted characteristics: a higher pressure in the oral cavity of the player leads to opening of the valve. The basic shape used in digital waveguide modeling is the cylinder, with a reflection filter fitted such that is closely matches the real modal series of the brass instrument. This is combined with a modified reed model to form the complete trumpet model.

2.5   Conclusion

In this chapter, we have reviewed the operation principles and structure of a few common wind instruments. We conclude from the qualitative descriptions that they all share the same basic physical structure and acoustic properties, and can thus be modeled using the same techniques.

Chapter 3

String Instruments

There is geometry in the ringing of strings. There is harmony in
the spacing of the spheres.

Pythagoras

This chapter gives an overview of the acoustic properties of string instruments,
and compares three different types of string instruments: the acoustic and
electric guitars, the violin family, and the piano.

3.1   Acoustic structure

All string instruments share the same basic structure: they all have one or
more strings which transfer their energy to a soundboard. The soundboard
is an integral part of a resonating cavity, or is placed inside one. The sole
exception is the electric guitar, where the acoustic resonator is replaced by an
electro-acoustical amplifier.

The behavior of waves in strings is discussed in section 4.2.2. Strings are
either plucked with the fingers (lute, guitar, harp), a plectrum (guitar), or jacks
(harpsichord); struck (clavichord, piano); or bowed (violin family). The strings
are too thin to efficiently radiate low frequency sound. The string vibration is
transferred via a structure commonly called a bridge to a soundboard with a

Figure 3.1: Classical guitar shape. This shape has been practically unchanged since its design by Antonio de Torres. The fingerboard usually has 18 frets, with the 12th fret, corresponding to the octave, flush with the body. The distance between two frets is determined such that the pitch difference equals one semitone.

better impedance match with the surrounding air. The bridge input impedance should be high enough such that the major part of the incoming waves are reflected back to the string. This is necessary to obtain a long sustain time, and is especially important for plucked string instruments. On the other hand, enough energy has to be transferred to the soundboard to radiate a high sound volume. This leads to the compromise that string instruments have either a long sustain or a high sound volume. The bridge input impedance has a great influence on the timbre of the instrument. On some instruments, all the strings are attached to the same bridge. This leads to strong string coupling as all the strings have a common movement at the bridge.

3.2   The acoustic guitar

The modern six-string guitar is a descendant of the sixteenth-century Spanish *vihuela*. It was however not until the nineteenth century that the guitar was established as a concert instrument. The shape of the classical guitar as we know it today (see figure 3.1) was largely influenced by the Spanish luthier Antonio de Torres (1817-1892).

The modern guitar has six strings and is about 65cm in length. Normally, the string are tuned to E_2, A_2, D_3, G_3, B_3 and E_4 ($f_0 = 82$, 110, 147, 196, 247, 330Hz), but alternate tunings are widely used. The top is usually cut from spruce (*Picea Sitchensis, Engelmannii*, or *Excelsa*), cedar (*Thuja Plicata*), or redwood (*Sequoia Sempervirens*), planed to about 2.5mm thickness. The back is usually a hardwood (rosewood, maple, mahogany or the more exotic blackwood, koa or walnut). Both top and back plates are braced, and the bracing of the
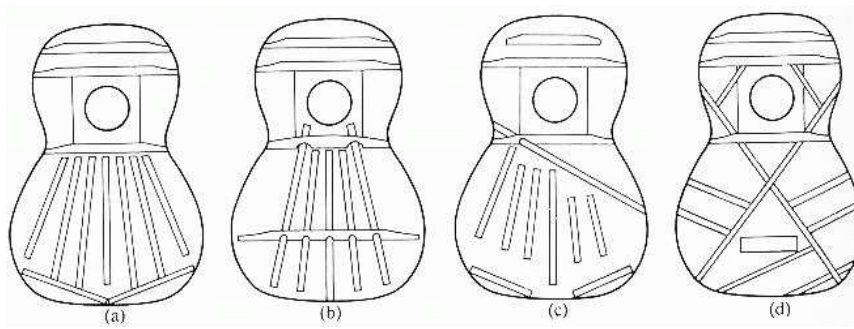
Figure 3.2: Designs for bracing: (a) Torres traditional fan bracing, (b) Bouchet bracing, (c) Ramirez asymmetric bracing, (d) crossed bracing. After [94].

top plate is an important design parameter. Several design for bracing can be seen in figure 3.2.

Acoustic guitars fall in roughly four families of design: classical, flamenco, flat top (folk) and arch top. Classical and flamenco guitar have nylon strings; the flamenco guitar is somewhat thinner, has an extra protective plate on top, and a different head and tuning machine. The flat top and arch top guitars have steel strings. Flat top guitars exist in different sizes (shown in figure 3.3), and are adapted for particular musical styles, and some have a cutaway to facilitate access to the higher frets. Arch-top guitars are based on the design of violins, having an arched top and back and the typical *f*-holes, and were pioneered by Orville Gibson (1856-1918) (see figure 3.4).

The vibrations of the strings are transferred via the saddle and bridge to the top plate of the guitar. The guitar body acts as an acoustic amplifier, as the plates and air cavity of the body provide a better impedance match with the surrounding air than do the strings. The coloration of the sound and the directional pattern of the radiation are mainly determined by the shape and construction of the body. The total string tension for (normal-tension strings) is about 360 N for a classical guitar. Steel string guitars have a higher string tension and require an additional steel bar in the neck for rigidity.

The lowest body resonance of the guitar is typically in the range from 90 to 100Hz (lower for Jumbo-sized flat-tops) and the second one is between 160 to 240 Hz. The lower resonance corresponds to the Helmholtz mode of the air cavity, and the second resonance is the first vibrational mode of the soundboard.

Figure 3.3: Typical shapes for steel string acoustic guitar. From left to right: Taylor "Grand Concert" body with cutaway, "Grand Auditorium", "Dreadnought", and "Jumbo" body shapes. A cutaway enables easier access to the higher frets. The "Jumbo" guitar is characterized by a large and deep body, resulting in a lower resonance frequency, thus enhancing the bass response.

Figure 3.4: Gibson archtop electroacoustic guitar. The arched top plate, with the characteristic f-holes, was originally designed to achieve a louder sound, before the introduction of electric guitar amplification. This is a typical blues and jazz guitar.
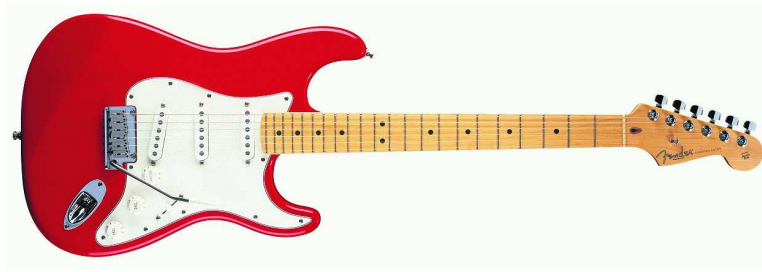
Figure 3.5: Typical solid-body electric guitar (Fender Stratocaster).   The
Fender Stratocaster is a typical rock guitar.

3.3   The electric guitar

Electric guitars get their name from the method of sound transmission from the
strings to the surrounding air. Electric guitars are equipped with electromag-
netic, piezoelectric or, more recently, optical pickups to generate an electrical
analog signal that is fed to an amplification system.

Electric guitars may have either a solid wood body or a hollow body, but both
types emit very little sound and are characterized by a long sustain. The elec-
tromagnetic pickup consists of a coil and a permanent magnet (see figures 3.5
and 3.6). The string movement above the permanent magnet causes a flux fluc-
tuation through the coil, thus inducing an electrical signal. Typically, there is
one permanent magnet for each string, with the coil winding for the individual
coils connected in series. This single-coil setup is very sensitive to EMC noise,
and this lead to the development of "humbucking" pickups, using two coils with
inverted phase to achieve higher common mode rejection. Most electric guitars
have two or three set of pickups at various places along the strings. The pickups
essentially sample the movement of the string directly above it, so the front
pickup (nearest the fingerboard) provides the strongest fundamental, whereas
the rear pickup (near the bridge) is more sensitive to the higher harmonics. The
player can set several combinations of the pickups with a selector switch, and
is usually able to control the output volume and affect tone with a low-pass fil-
ter. Electromagnetic pickups are highly non-linear, and have the most impact
on the actual "clean" sound of the guitar. On the other hand, piezoelectric
pickups are based on ceramics or advanced polymers that generate a voltage
when pressure is applied to it. They are fitted between saddle and string, or
below the saddle, and generate a voltage more or less linearly dependent with
the movement of the string.   Both electromagnetic and piezoelectric pickups
have a high output impedance and require high quality, low noise amplification.
Many amplifiers for electric guitars provide an overdriven stage, where the non-
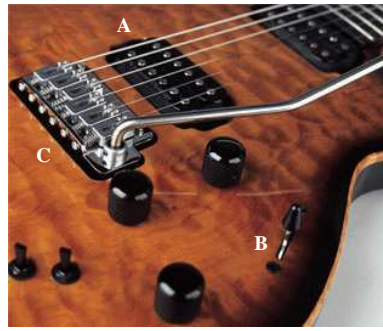linear saturation characteristic of the power amplifier (based on either tubes

Figure 3.6: Detail view of the bridge, pickup and tone control of a solid-body electric guitar. A. Humbucking pickup; B. Pickup selector switch; C. Bridge with individually adjustable saddles and tremolo arm.

or solid-state electronics) is exploited to create musically interesting sounds. There exist several hundred types of different effect boxes to further modify the sound.

From a modeling standpoint, the electric guitar is somewhat easier to model than the acoustic guitar, and only needs an additional pickup model.

3.4   The violin

The violin gradually evolved from the various bowed string instruments used in Europe during the Middle Ages. The instruments of the violin family were developed in Italy during the sixteenth and seventeenth centuries and reached a peak in the eighteenth century in hands of masters such as Antonio Stradivari (1644-1737). The violin is the soprano voice of the violin family, and the viola, violoncello, and contrabass represent the alto, baritone and bass voices. These are all distinctly different instruments. The bow was given its present form by Francois Tourte (1747-1835). The violin has been the object of considerable scientific study, but its acoustical behavior is just beginning to be understood. The violin makers built experimental instruments and acquired a correct understanding of at least some parts of the instrument. Herman von Helmholtz (1821-1894) contributed by his physical and psycho-acoustical experiments. He identified several partials in the tone of violins by listening to resonant cavities (now known as Helmholtz resonators) and the sawtooth waveform or Helmholtz motion of the strings. The standard work on violin acoustics is the book by Cremer [26]. The essential parts of the violin are shown in figures 3.7 and 3.8. The four strings of steel, gut, or nylon (wound with either silver, aluminium, or
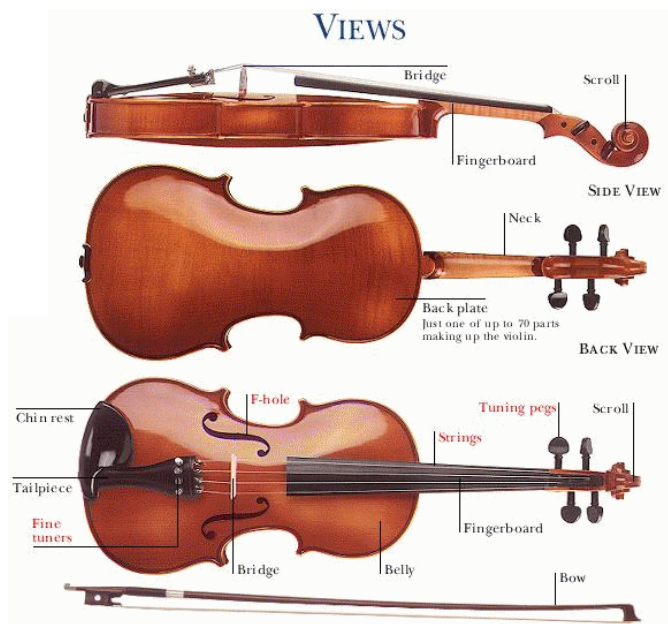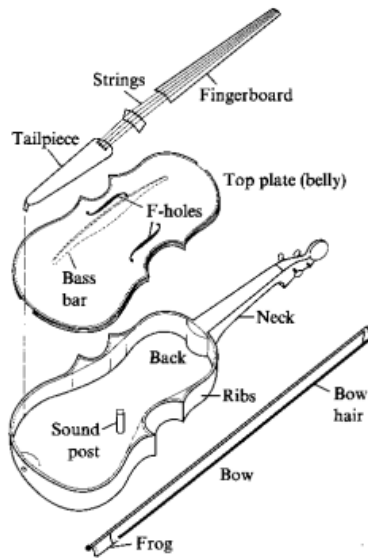
Figure 3.7: Structure of a modern violin.

Figure 3.8: Exploded view of a violin, after [94]

steel) are tuned to G_3, D_4, A_4, and E_5 and transfer their vibrational energy to the top plate through the bridge. The shape of the bridge is a convenient and effective way to alter the frequency response of the instrument. The top plate is generally carved from Norway spruce (*Picea Abies*) or German spruce (*Picea Excelsis*), and the back and the sides from curly maple (*Acer Platanoides*). The fingerboard is made of ebony and the bow stick from pernambuco (*Caesalpinia Echinata*) or composite materials. Running longitudinally under the top is the bass bar, and a sound post is positioned near the treble foot of the bridge.

The total tension of the four strings of a violin is typically about 220 N, and this results in a downward force on the bridge of about 90 N. The strings are excited with a bow. This interaction could be seen as a continuous "stick and slip" mechanism, and excites both the transversal and torsional modes of the strings. When the frictional force on the bow-string interface is larger than the elastic restoring force, the string "sticks" to the bow and induces movement. If the frictional force is too small, the string slips back to its original position. The actual interaction is more complicated and has been extensively studied [90], [13], [26]. One particular effect that affects the members of the violin family are the so-called "wolf tones". These are tones that are particularly difficult to sustain, and are caused by a too-low bridge impedance at that particular frequency.

From the modeling standpoint, it is not difficult to transform our guitar model presented in the next chapters into a violin model. The most difficult part is the bow-string interaction which requires a supplementary non-linear function for the string excitation. All other elements of a violin can be modeled in exactly the same way as in the guitar case.

3.5   The piano

The modern piano is a direct descendant of the harpsichord, and uses hammers instead of jack to excite the strings. Out of its 300-year history, two types of musical instruments have evolved. One is the grand piano, which is built in various sizes, from baby grand to concert grand (up to 3m in length, shown in figure 3.9), and the other is the upright piano, which vertically oriented strings. Most pianos have 88 keys, and concert grand pianos contain 243 strings. For some notes, two or three strings are used in unison, leading to interesting coupling effects. The strings transmit a part of their vibrational energy to the soundboard, which is nearly always made of spruce. To obtain the desired loudness, piano strings are held at a tension that may exceed 1000 N. In order to withstand this total force of the strings and maintain stability of tuning, pianos use a cast iron frame. The piano action is a complex system of levers, dampers, and a hammer, and is designed to transmit the energy from the

Figure 3.9: A concert grand piano (Bösendorfer model 290).

player's fingers to the string. Piano string are damped when not activated by a key, and when the sustain pedal is depressed. This restricts the number of string models needed for the simulation. The hammer-string interaction is quite complex, and depends on the impact speed, the shape, the material properties of the hammer, and the state of the string when hit. The stiffness of the strings also introduce a significant inharmonicity of the higher partials.

The piano could be seen as an extension of the guitar, with more strings and a different excitation. String models and string coupling mechanisms developed for the guitar should also be applicable to the piano.

3.6 Conclusion

In this chapter, a short overview of different types of string instruments was given. All string instruments share a common structure, and differ mostly in the way the strings are excited. The sound of a string instrument is produced mostly using linear vibration modes, but some effects and excitations require a non-linear interaction. String instruments also exhibit weak or strong coupling between the strings through the common saddle, bridge, or soundboard.

Chapter 4

Acoustics
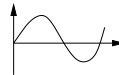
Acoustics, then, or the science of sound, is a very considerable
branch of physics.

Sir J. Herschel.

This chapter provides an introduction to musical acoustics. Starting from the
general equations describing sound propagation in air and the vibration of
strings in section 4.2, we derive the properties of the common acoustical build-
ing blocks of musical instruments in section 4.3. We then discuss the com-
bination of the different acoustic elements into an acoustic system in section
4.4.

4.1   Sound perception

The human ear is an amazing organic mechanical amplifier. It has evolved
to amplify the small air pressure variations we call sound, over a frequency
range of 20Hz-20kHz, and over an amplitude range of 100dB. Before we can
appreciate sound, waves in the air reach the outer ear or auricle (pinna), which
contributes to the localization of the origin of the sound. The sound waves
then reach the tympanic membrane. The membrane vibrates and the vibra-
tion is transmitted to the inner ear by mean of the three small bones, *malleus*

(hammer), *incus* (anvil) and *stapes* (stirrup), visible on figure 4.1. The vibration causes the stirrup to act as a piston that, by displacement of a small and thin membrane on the oval window of the cochlea, displaces the endolymphatic content of the cochlea duct. Displacement causes the portion of the membrane, where hair cells rest, to undulate in conjunction with another membrane on top of the hair cells, making the hair cells transduce the mechanical energy into neural stimulation. From the hair cells of the inner ear, the neural stimulus is transmitted by the afferent cochlear nerve fibers to the brain stem; and from there to the various stations along the brain center up to the cortex where speech and sound are finally decoded [49],[84].

4.1.1   Volume and pitch perception

The perceived volume is a logaritmic function of the amplitude of the sound. It is frequency-dependent, and is subject to psychoacoustical masking phenomena. A quadrupling in amplitude is perceived as a doubling in volume. As for the frequency perception, between about 1,000 Hz and 8,000 Hz, we notice changes between frequencies whose ratios are about 1.002 or 1.003, which is roughly 200 to 350 steps per octave, or something between 1/30 and 1/15 of a musical half step. Outside of this range, discrimination is poorer, but for most of the range of audible frequencies we notice changes in ratios smaller than 1.01, which gives more than 60 steps per octave, or something smaller than 1/5 of a half step. Discrimination of frequencies played in sequence is a bit less - typically about 90 steps per octave or about 1/8 of a half step.

Musicians interested in nonstandard pitches have usually used the *cent*, which is 1/100 of a half step, or the *savart*, which is 1/25 of a half step. In complex sounds, frequency distinctions may be important even though they are less than those preceptible as changes in a simple helical signal. There is also a unit, called the *mel* that is like a pitch measurement, but scaled to the people's judgments that certain pitches are "twice as high" as others in psychological experiments.

Taking frequency resolution between 90 and 360 steps per octave, over a range of 10 octaves, we get 900 to 3600 distinguishable frequencies. But, it seems that we cannot exploit those as independent bits, and the practical information capacity of a single sound is much less.

4.1.2   Music

Music is the art of arranging sounds in time so as to produce a continuous, unified, and evocative composition, as through melody, harmony, rhythm, and timbre. Most western music is based on a 12-tone system that is well suited to form harmonically pleasing chords. Each note has a corresponding fundamental frequency. Tables 4.1 and 4.2 show two such systems. In this work, the
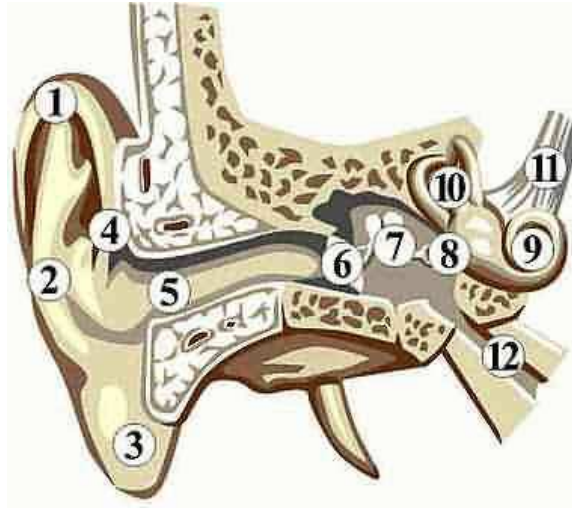
Figure 4.1: The human ear. (1) **Helix** The in-curve rim of the external ear, (2) **Antihelix** A landmark of the outer ear, (3) **Lobule** The very bottom part of the outer ear, (4) **Crest of Helix** A landmark of the outer ear, (5) **External Auditory Meatus** or **External Auditory Canal** The auditory canal is the channel through which the sounds are led from the ear outside to the middle ear, (6) **Eardrum** (tympanic membrane) A thin layer of skin at the end of the external ear canal, (7) **Auditory Ossicles** The three small bones in the middle ear, know as the *malleus*, *incus* and *stapes* which are connected to one another. Together these ossicles are called the ossicular chain. Their purpose is to lead the sound striking the eardrum further into the inner ear, (8) **Oval Window** An opening in the bone between the air filled middle ear cavity and the fluid filled inner ear, and is covered by a thin membrane, (9) **Cochlea** Part of the inner ear that contains part of the hearing organs, (10) **Semicircular Canals** Part of the organ of balance that is part of the inner ear, (11) **Eighth Nerve** Nerve that transmits messages from the inner ear to the brain, and (12) **Eustachian Tube** A tube connecting the middle ear cavity and the pharynx (back of the throat). It can be opened by coughing or swallowing, though it is normally closed. The occasional opening of the Eustachian tube is necessary to equalize the are in the middle ear cavity [84].

semitone	octave						
C#	34.6	69.2	138.5	277.1	554.3	1108.7	2217.4
D	36.7	73.4	146.8	293.6	587.3	1174.6	2349.3
D#	38.8	77.7	155.5	311.1	622.2	1244.5	2489.0
E	41.2	82.4	164.8	329.6	659.2	1318.5	2637.0
F	43.6	87.3	174.6	349.2	698.4	1396.9	2793.8
F#	46.2	92.4	184.9	369.9	739.9	1479.9	2959.9
G	48.9	97.9	195.9	391.9	783.9	1567.9	3135.9
G#	51.9	103.8	207.6	415.3	830.6	1661.2	3322.4
A	55.0	110.0	220.0	**440.0**	880.0	1760.0	3520.0
A#	58.2	116.5	233.0	466.1	932.3	1864.6	3729.3
B	61.7	123.4	246.9	493.8	987.7	1975.5	3951.0
C	65.4	130.8	261.6	523.2	1046.5	2093.0	4186.0

Table 4.1: "Equally tempered" scale. The different *octaves*, or doublings of the fundamental frequency, are subdivided in twelve semitones. The frequency ratio between two semitones is constant and equal to $\sqrt[12]{2}$. The fundamental frequency of the middle A is set to 440 Hz in contemporary music.

equally tempered tuning will be used. Note that a musical instrument is tuned such that it approximates a system. For purely mechanical an constructional reasons, there are very few instruments that can be tuned exactly. Notable exceptions are the violin family, because here, the player is able to exactly determine the fundamental frequency of the note to be played. This is also the reason why there often exist different versions of the same instrument. For instance, flutes exist in contemporary and baroque tuning.

4.2   Acoustic theory

Mechanical, acoustical, or electrical vibrations are the sources of sound in musical instruments. The sensation we call sound is produced primarily by variations in air pressure that are detected by the movement of our ear drums, and the nerve impulses generated by the sensory hair cells inside the cochlea. The field of musical acoustics describes all the mechanisms that result in audible sound waves, including the generation, propagation, and interaction of vibrations in any medium. The field of fundamental physical acoustics goes back to the ancient Greeks, and was developed over the centuries by the likes of Pythagoras (c. 550 BC), Aristotle (384-322 B.C.), the "father of acoustics" Marin Mersenne (1588-1648), Isaac Newton (1642-1727), Euler (1707-1783), Lagrange (1736-1813), and d'Alembert (1717-1783). Today, musical acoustics is but a small part of the research in acoustics, and provides us with mathematical models for almost any type of instrument. For the purpose of this work,

semitone				octave			
C#	34.3	68.7	137.3	274.7	549.4	1098.8	2197.6
D	36.5	73.3	146.7	293.4	586.8	1137.6	2275.2
D#	38.6	77.2	154.5	309.1	618.2	1236.4	2472.8
E	41.2	82.5	165.0	330.1	660.2	1320.4	2640.8
F	43.5	86.9	173.8	347.7	695.4	1390.8	2781.6
F#	45.7	91.5	183.1	366.3	732.6	1465.2	2930.4
G	48.6	97.2	195.5	391.1	782.2	1564.4	3128.8
G#	51.5	103.0	206.0	412.0	824.0	1648.0	3296.0
A	55.0	110.0	220.0	**440.0**	880.0	1760.0	3520.0
A#	57.9	115.8	231.7	463.5	927.0	1854.0	3708.0
B	61.8	123.7	247.5	495.0	990.0	1980.0	3960.0
C	65.1	130.3	260.7	521.5	1043.0	2086.0	4172.0

Table 4.2: "Just" scale. The frequencies of the different notes in this scale are based on the logical integer ratios of the harmonic series. The ratios for the seven notes are 1:1, 9:8, 5:4, 4:3, 3:2, 5:3, 15:8, and finally 2:1. Note that the ratio of the frequencies of two semitones is not a constant.

we restrict the study of acoustics to sound propagation in air and strings.

4.2.1 Sound propagation in air

A sound wave is a pressure wave propagating in a medium, most commonly air. For not-too-high pressure variations, air acts as a linear medium. The properties of sound waves are derived by the analysis of an elementary cubic section of air. The three-dimensional scalar wave equation is found to be:

$$\nabla^2 p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}, \tag{4.1}$$

where p denotes the pressure, and c the speed of sound. Assuming sinusoidal wave components with time dependence $e^{j\omega t}$, equation (4.1) reduces to the scalar Helmholtz equation

$$\nabla^2 p + k^2 p = 0, \tag{4.2}$$

where $k = \omega/c$ is the wavenumber, and the speed of sound is given by

$$c(T) = 332(1 + 0.00166T)\text{m/s}, \tag{4.3}$$

with T the air temperature in degrees Celsius, and at 50% humidity.

We consider two particular cases: plane waves and spherical waves. The true plane wave is the result of the excitation of the air mass by an infinite plane, while the spherical wave is the result of a point source. For real, free-standing

sound sources, the far field is best described as a plane wave, while the spherical description is better suited to the near field. In musical instruments, both types are equally important.

For plane-wave propagation along the x-axis, equation (4.1) reduces to

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}. \tag{4.4}$$

Solutions of this equation describe one-dimensional plane waves. Assuming sinusoidal time dependence, a general frequency-domain solution is

$$P(x,t) = [Ae^{-jkx} + Be^{jkx}]e^{j\omega t}, \tag{4.5}$$

where A and B are the complex amplitudes for wave components going in the right and left directions. The associated air particle velocity is

$$V(x,t) = \frac{1}{\rho c}\left(Ae^{-jkx} - Be^{jkx}\right)e^{j\omega t}. \tag{4.6}$$

It is useful to describe the acoustic properties in a system theoretic context. This allows the definition of the wave impedance or characteristic impedance seen by a wave leaving the source ($B = 0$), at one particular frequency:

$$Z = \frac{P(x,t)}{V(x,t)} = \rho c. \tag{4.7}$$

For plane-wave propagation in air, the wave impedance is real and frequency-independent.

For spherical waves, the Laplacian in equation (4.1) has to be written in spherical coordinates. Assuming one-dimensional propagation, the equation becomes

$$\frac{1}{x^2}\frac{\partial}{\partial x}\left(x^2\frac{\partial p}{\partial x}\right) = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{4.8}$$

where x measures the radial distance from the point source. A general frequency-domain solution to equation (4.8) which assumes sinusoidal time dependence is given by

$$P(x,t) = \left[\frac{A}{x}e^{-jkx} + \frac{B}{x}e^{jkx}\right]e^{j\omega t}, \tag{4.9}$$

and the particle velocity is

$$V(x,t) = \frac{1}{\rho c}\left[\frac{A}{x}\left(1 + \frac{1}{jkx}\right)e^{-jkx} - \frac{B}{x}\left(1 - \frac{1}{jkx}\right)e^{jkx}\right]e^{j\omega t}. \tag{4.10}$$

Again, we determine the impedance for outgoing waves ($B = 0$) as

$$Z(x) = \frac{P(x,t)}{V(x,t)} = \rho c\left(\frac{jkx}{1 + jkx}\right). \tag{4.11}$$

We see that for the far field ($kx \gg 1$), the wave impedance for spherical waves reduces to ρc, the wave impedance for plane waves.

4.2.2 Vibrations in strings

Very similar to sound propagation in air, wave propagation in strings can be
described with a simple differential equation. The transversal motion of an
ideal string is described by

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2},$$ (4.12)

with c the transversal wave propagation speed:

$$c_{\text{trans}} = \sqrt{\frac{T}{\rho}},$$ (4.13)

with T the string tension and ρ the linear density of the string material. For
a finite string, rigidly attached at both ends, one solves equation 4.12 with
boundary conditions $y(0) = 0$ and $y(L) = 0$ and initial conditions $y(x) = y_{\text{init}}(x)$ and $\partial y(x)/\partial t = v_{\text{init}}(x)$. The traveling wave components are inverted
during the reflection. One obtains a sum of normal modes

$$y = \sum_n C_n \sin(\omega_n t + \phi_n) \sin(k_n x).$$ (4.14)

For a string of length L with fixed endpoints, the modal frequencies are

$$f_n = n \frac{c_{\text{trans}}}{2L},$$ (4.15)

with $n = 1, 2, \ldots$. In strings, longitudinal and torsional waves are also pos-
sible. Depending on the instrument, these propagation modes are more or
less important. For instance, torsional waves are very important in bowed
string instruments. Longitudinal waves in a string are much less common than
transversal waves. Unlike transversal waves, their velocity is independent of
string tension and given by

$$c_{\text{long}} = \sqrt{\frac{E}{\rho}}.$$ (4.16)

Torsional waves are also non-dispersive and depend on the shear modulus G
of the material, and can be found as a special case of torsional waves in a rod,
given by

$$c_{\text{tors}} = \sqrt{\frac{GK_T}{\rho I}},$$ (4.17)

where K_T is the torsional stiffness factor that relates a twist to the shearing
strain produced, and I the polar moment of inertia per unit length. For round
strings, $K_T \cong I$ and the velocity is

$$c_{\text{tors}} \approx \sqrt{\frac{G}{\rho}}.$$ (4.18)
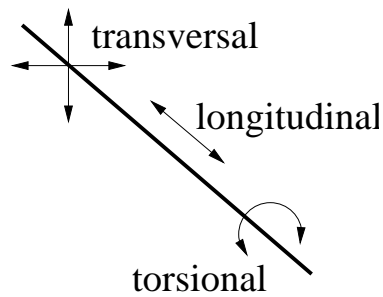
Figure 4.2: There are three distinct wave propagation modes in strings: the horizontal and vertical transversal waves, the longitudinal waves, and the torsional waves.

Unlike wave propagation in air, wave propagation in strings is subject to several non-ideal effects at low amplitude and in the frequency range of interest. Non-ideal string motion will be discussed in chapter 10.

4.3   Acoustic elements

Musical instruments consist of many elementary parts. In this section, we analyze the cylindrical and conical bore, the non-flaring and flaring ends, and the Helmholtz resonator. These structures constitute the base for the models studied later in this work.

4.3.1   Bores

Bores are most commonly cylindrical or conical cavities, used as resonating chambers for wind instruments. Infinite cylindrical pipes with rigid, perfectly smooth, and thermally insulating walls, have no effect on the wave propagation. Bores have finite dimensions and non-constant diameter, introducing internal reflections at the boundaries and diameter changes. We discuss the properties of cylindrical and conical bores, as a general bore can be approximated by a succession of cylindrical and conical sections.

Cylindrical bore

We derive the properties of cylindrical bores by solving the Helmholtz equation (4.2) in cylindrical coordinates (r, θ, x) with the appropriate boundary conditions (refer to figure 4.3). The wave equation becomes
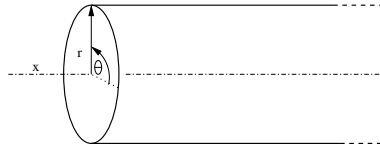
Figure 4.3: The (r, θ, x) circular cylindrical coordinate system used for a cylindrical bore.

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 p}{\partial \theta^2} + \frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}. \tag{4.19}$$

It is separable in circular cylindrical coordinates, and yields

$$\frac{d^2\Theta}{d\theta^2} + m^2\Theta = 0, \tag{4.20}$$

$$r\frac{d}{dr}\left(r\frac{dR}{dr}\right) + (a^2 r^2 - m^2)R = 0, \tag{4.21}$$

$$\frac{d^2 X}{dx^2} + (k^2 - a^2) = 0, \tag{4.22}$$

where a and m are the separation constants. Equation (4.22) describes the motion alongst transverse concentric circles in the pipe, while equation (4.21) describes transverse radial motion. These modes are not very important in this study, as they are quite difficult to excite in conventional musical instruments. The axial wave motion, as described by equation (4.22), is the main wave propagation mechanism. The wave impedance for a cylindrical bore is

$$Z_0(x) = \frac{P(x)}{U(x)} = \frac{\rho c}{S}, \tag{4.23}$$

where $U(x)$ represents the volume flow and S the surface of the opening of the bore.

This is the solution for an infinite cylinder. In finite cylinders, the waves will be partly reflected due to the discontinuities at both ends. Thus, the sinusoidal pressure in the pipe is given by

$$P(x,t) = (Ae^{-jkx} + Be^{jkx})e^{j\omega t} \tag{4.24}$$

where A and B are complex amplitudes. The input impedance for a cylinder terminated with a load impedance Z_L can be calculated as

$$Z_{in} = Z_0\frac{Z_L cos(kL) + jZ_0 sin(kL)}{Z_0 cos(kL) + jZ_L sin(kL)} \tag{4.25}$$

Figure 4.4 shows the input impedance for two particular (ideal) cases: a rigidly terminated cylinder ($Z_L = \infty$) and an open cylinder ($Z_L = 0$). The resonance
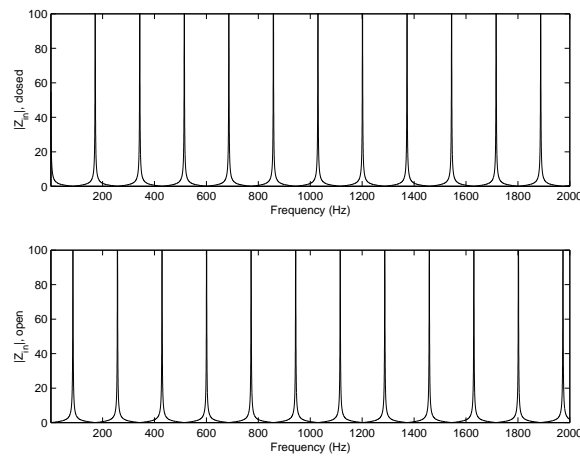
Figure 4.4: This figure shows the acoustic input impedance for (top) a closed cylinder with $Z_L = \inf$, and (bottom) an open cylinder with $Z_L = 0$. The bore has a length of $1m$ and normalized Z_0

frequencies are

$$f = \frac{(2n-1)c}{4L} \tag{4.26}$$

for an open-closed bore, with length L and $n = 1, 2, \ldots$, and

$$f = \frac{nc}{2L} \tag{4.27}$$

for the open-open case.

Example 4.1 The fundamental frequency for a cylindrical bore of length 1m can be found using equation (4.26) for the open-closed case. Filling in the values yields as the fundamental 85.8 Hz. This corresponds to the first zero in the impedance shown in figure 4.4 (top). The higher partials are 257.4 Hz, 429 Hz, ..., or only odd multiples of the fundamental. The open-open cylinder has a fundamental of 171.6 Hz and higher partials at 343.23 Hz, 514.845 Hz, ...   △

Note that these resonance frequencies will occur depending of the excitation. Single-reed instruments are excited at the impedance minima, while air jets excite the bore at the impedance maxima.

Conical bore

Spherical waves of sound can theoretically propagate without reflection or loss away from the apex along the principal axis of an infinite conical bore, assuming

ideal conditions. The wave equation in spherical coordinates is quite complex, but separable, and the resulting differential equations describe sinusoidal wave motion along each of the spherical coordinate axes. The transverse and axial modes are described with Legendre and Bessel functions, respectively. One-dimensional spherical-wave propagation along the central axis has a general solution of the form:

$$P(x) = \frac{C}{x} e^{-jkx}. \tag{4.28}$$

For a cone with finite length, we have

$$P(x,t) = \left(\frac{A}{x} e^{-jkx} + \frac{B}{x} e^{jkx} \right) e^{j\omega t}. \tag{4.29}$$

The wave impedance for spherical waves traveling away from the apex is given by

$$Z_0(x) = \frac{\rho c}{S(x)} \left(\frac{jkx}{1 + jkx} \right) = \frac{1}{\frac{S(x)}{\rho c} + \frac{S(x)}{j\omega \rho x}}. \tag{4.30}$$

Note that this is equivalent to a cylindrical bore in parallel with a lumped inertance of acoustic mass $\rho x / S(x)$. This property has been exploited to "convert" a clarinet (with mostly cylindrical bore) into a sopranino saxophone (with a conical bore), just by adding a stub to the body of the clarinet. The theoretical input impedance seen from the open end, for a complete cone is

$$Z_{in} = \left(\frac{S(L)}{\rho c} \left[\cot(kL) - \frac{1}{kL} \right] \right)^{-1}. \tag{4.31}$$

Figure 4.5 shows the theoretical input impedance and admittance for a complete cone. The resonance frequencies for a complete cone occur where $Z_{in} = 0$ or

$$f = \frac{nc}{2L}, \tag{4.32}$$

the same as in the case of an open-open cylinder.

In reality, a complete cone is not very useful, and a truncated cone with length $L - L_0$ is used instead, where L is the length of the complete cone, and L_0 the length of the truncated part. The resonance frequencies for two open ends are given by

$$f = \frac{nc}{2(L - L_0)}. \tag{4.33}$$

When the truncated end is closed, the resonance frequencies are found as the solutions of

$$\tan(k(L - L_0)) = -kL_0. \tag{4.34}$$

Figure 4.7 shows the fundamental frequency for bores going from a complete cone, over closed-open truncated cones, to the closed-open cylinder (figure 4.6). The structure is morphed by varying the ratio of the tip radius versus the radius
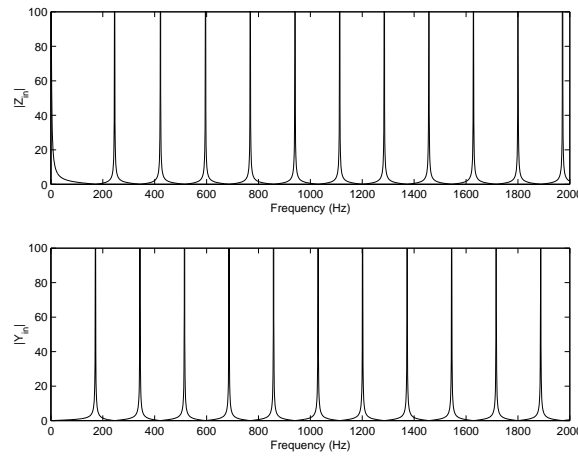
Figure 4.5: Theoretical input impedance Z_{in} and admittance of a complete cone, length $1m$, at the open end.

of the frustrum opening. In the case ratio $= 0$, we have a complete cone, and for ratio $= 1$, we have a cylinder. In the complete cone case, the frequencies of the partials correspond to the zeros of the acoustic impedance shown in figure 4.5, and morph to the impedance zeros of figure 4.4. Note that the frequency ratios of the partials go from nf_0 for the complete cone to $(2n-1)f_0$ for the closed-open cylinder.

Example 4.2 The effect of the truncation of a cone is clearly visible in figure 4.7. For a complete cone with length of 1m, one finds a harmonic series given by equation (4.33), with fundamental frequency of 171 Hz, and all the integer multiples. The partials for a truncated cone do not form a harmonic series (e.g. for a ratio of 0.6, the partials have frequencies of $f_0, 2.75f_0, 4.35f_0, 6f_0, \dots$). $\triangle$

Losses and non-ideal behavior

Wave propagation in real bores is changed by the viscous drag (in region called the viscous layer) and thermal conduction (within the thermal layer) along the bore walls. These phenomena affect the wave propagation at close distance of the walls. As a result, small bores will be affected more than bores with a large opening.

The thickness of the viscous layer is given by [11]

$$\delta_{\text{visc}} = \sqrt{\frac{\eta}{\omega \rho}}, \tag{4.35}$$
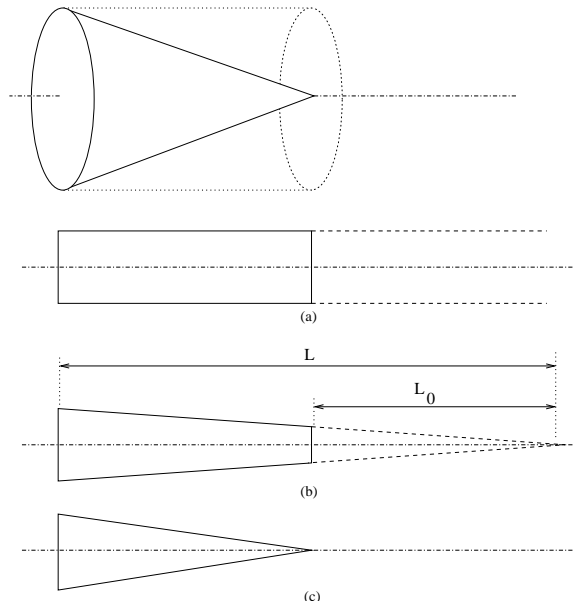
Figure 4.6: (a) Cylinder: length $L_0 = \infty$, and ratio $= 1$, (b) truncated cone, (c) complete cone: length $L_0 = 0$ and ratio $= 0$.
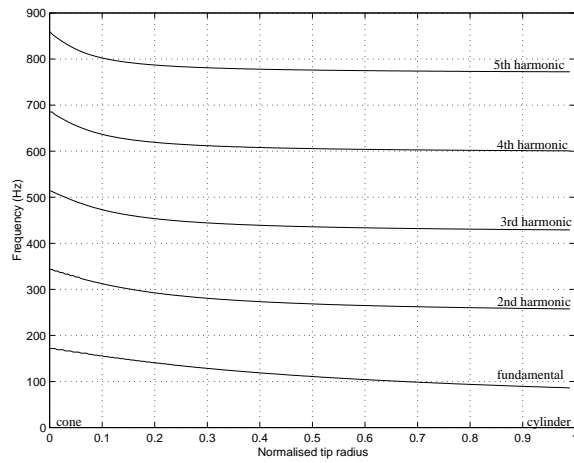
Figure 4.7: The fundamental frequency and the first four harmonics, for a structure morphed from a complete cone (ratio=0) to a closed-open cylinder (ratio=1), with length $1m$.

where η is the viscosity of air, and ρ the density. The thickness of the thermal boundary layer is

$$\delta_{\text{therm}} = \sqrt{\frac{\kappa}{\omega \rho C_p}}, \tag{4.36}$$

with κ the thermal conductivity and C_p the specific heat of air. Friction along the walls, due to the surface roughness of the wall material, is characterized by a parameter r_v, given as the ratio of the (local) bore radius R_b to the viscous boundary layer thickness, or

$$r_v = \sqrt{\frac{\omega \rho}{\eta}} R_b, \tag{4.37}$$

and similarly, the thermal effects are characterized by a parameter r_t:

$$r_t = \sqrt{\frac{\omega \rho C_p}{\eta}} R_b. \tag{4.38}$$

The effects of these losses are incorporated into the wave equation by rewriting the wavenumber k as

$$k = \frac{\omega}{v_p} - j\alpha, \tag{4.39}$$

where v_p is the wave velocity and α the attenuation factor. An expression for v_p and α can be found in (ref Keefe 1984). For the case of air, with $r_c > 2$ and at $300K$, they are well approximated by

$$v_p^{-1} = \frac{1}{c}(1 + 1.045 r_v^{-1}), \tag{4.40}$$

$$\alpha = \frac{\omega}{c}(1.045 r_v^{-1} + 1.08 r_v^{-2} + 0.75 r_v^{-3}). \tag{4.41}$$

The wave amplitude decays as $e^{-\alpha x}$ for plane waves or $e^{-\alpha x}/x$ for spherical waves. Figure 4.8 shows the damping factor α for two common cylindrical bore instruments: the clarinet (typical bore radius $7.3mm$) and the modern flute (typical bore radius $9.5mm$). The clarinet's higher harmonics will be more heavily damped than their flute counterparts, resulting in a more mellow sound.

4.3.2   Ends

The goal of a musical instrument is to emit sound. To achieve this, a part of the energy stored in the resonating cavity has to be radiated outwards, while still keeping enough energy inside to sustain the internal oscillation. Tuning the proportion of reflected versus transmitted energy can be done by altering the impedance mismatch at the end of the bore. Typically, some form of flaring end is used.
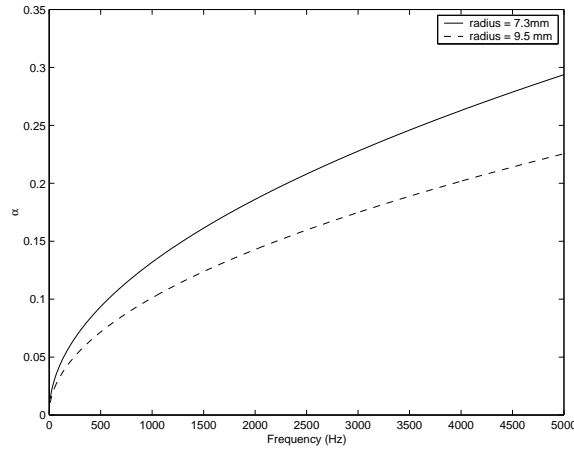
Figure 4.8: The attenuation coefficient α for two bore radii: $7.3mm$ (typical clarinet) and $9.5mm$ (typical modern flute). The damping factor is higher in the case of the clarinet, and the difference increases at higher frequencies. The result is that the flute has a brighter timbre than the clarinet.

Non-flaring ends

Examples of modern instruments without flaring ends are the flutes and organ pipes. For these cases, we have to determine the sound radiation at the end of a cylindrical bore shown on figure 4.9. This is quite similar to e.g. the sound radiation by loudspeaker cones, and has been extensively studied in the past. A flanged end (basically a hole in an infinite wall) has a load impedance given by

$$Z_L = A + jB. \tag{4.42}$$

with

$$A = Z_0 \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(ka)^{2n}}{(n!)^2(n+1)}, \tag{4.43}$$

$$B = \frac{Z_0}{\pi(ka)^2} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(2ka)^{2n+1}}{\left(\frac{(2n-1)!}{2^{n-1}(n-1)!}\right)^2 (2n+1)}, \tag{4.44}$$

with a the mouth radius, and $k = \omega/c$ the wave number. The reflection is then easily calculated with

$$\mathcal{R} = \frac{Z_L - Z_0}{Z_L + Z_0}, \tag{4.45}$$

and the transmission

$$\mathcal{T} = 1 - \mathcal{R}. \tag{4.46}$$

(a)

(b)
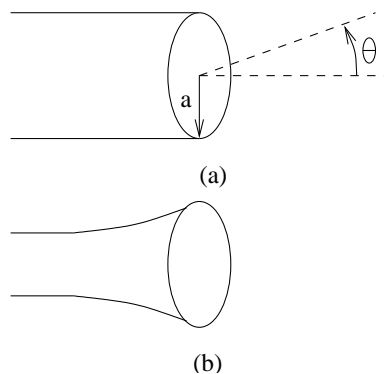
Figure 4.9: Two types of ends. (a) Unflanged, non-flaring end, typically found on flutes, organ pipes, etc. (b) Unflanged, flaring end

The radiation intensity at angle θ away from the axis of the mouth has the form

$$\frac{2J_1(ka\sin\theta)}{ka\sin\theta}, \tag{4.47}$$

with J_1 the first-order Bessel function. At low frequencies, the distribution is nearly independent of the angle θ, but as the frequency increases, the energy is primarily concentrated into a primary lobe centered on the axis. The ends of musical instruments are typically unflanged. An explicit calculation of the sound radiation by unflanged, non-flaring ends [55] yields that the cylindrical bore acts as if it is lengthened by a fraction of the radius. The radiated sound wave is more directional at higher frequencies. There is no explicit expression for radiation from a conical end. One could obtain an approximation by describing a conical end as a sequence of cylindrical sections.

Example 4.1 The radiation intensity is easily calculated with equation (4.47). Figure 4.10 shows the calculated intensity for a cylindrical bore, with a mouth radius of 2 cm, evaluated at 1 kHz, 5 kHz, and 10 kHz. The higher directivity is obvious at the higher frequencies.   △

Flaring ends

The flaring end of a musical instrument constitutes a smooth impedance transition from the bore to the surrounding air. The size and shape of the horn determine its characteristics. The properties of cylindrical-section can be derived by solving the Helmholtz equation. Unfortunately, the Helmholtz equation is only separable in coordinates that are confocal quadratic surfaces or their degenerate forms [66], and only a few of those are likely candidates for
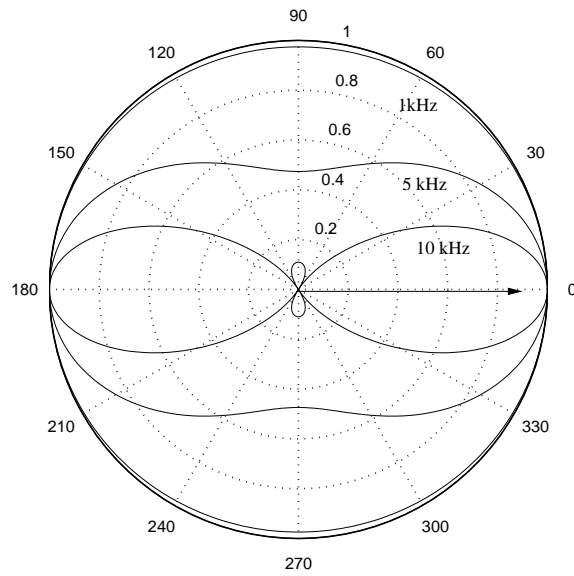
Figure 4.10: Radiation intensity for an unflanged, unflaring end, at three frequencies (1 kHz, 5 kHz, 10 kHz), for a mouth radius of 2 cm. One sees a higher directivity for the higher frequencies.

horn shapes. An approximate solution, for infinite horns with general shape, can be found by solving Webster's equation [118]

$$\frac{1}{S(x)}\frac{\partial}{\partial x}\left(S(x)\frac{\partial p}{\partial x}\right) = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{4.48}$$

where $S(x)$ is the geometrical cross-section. The most important effect of a bell is a lowpass effect on the transmission of the incoming sound waves. The three most widely studied horn shapes are the conical, the exponential and the Bessel horn. Approximate analytical formulas can be derived that adequately describe their properties. Bessel horns in particular, provide a good approximation to horn shapes used in musical instruments [10],[8]. The geometrical cross-section of a Bessel horn complies to

$$S = Bx^{-2\epsilon}, \tag{4.49}$$

where x is the geometrical distance measured from a reference point $x = 0$. If $\epsilon = 0$, the horn is cylindrical, and for $\epsilon = -1$, the horn is a cone. In the most useful case, ϵ is positive, which yields a horn with a rapid flare at the origin. Solving Webster's equation results in an analytical solution for the pressure standing wave $p(x)$

$$p(x) = Ax^{\epsilon+1/2}J_{\epsilon+1/2}(kx), \tag{4.50}$$

where J is a Bessel function and k the wave number. Composite horns consisting of Bessel, exponential, cylindrical, and conical sections can also be described [88]. The propagation in a flaring horn is governed by the value of the horn function F at the point concerned. The horn function in the plane-wave approximation is

$$F = \frac{1}{a}\frac{d^2 a(x)}{dx^2}, \tag{4.51}$$

with $a(x)$ the local radius of the cross-section $S(x) = \pi a(x)^2$. The wave in a horn is propagating if

$$k^2 = \left(\frac{\omega}{c}\right)^2 < F, \tag{4.52}$$

or attenuated if $k^2 > F$. Close to the open mouth, a more complex spherical approximation is needed. The net effect is the existence of a cut-off frequency above which the transmission is attenuated.

4.3.3   Helmholtz resonator

The Helmholtz resonator is a simple one-port structure, consisting of a relatively short open tube, connected to a cavity, as shown in figure 4.11. The acoustic impedance of the tube is given by

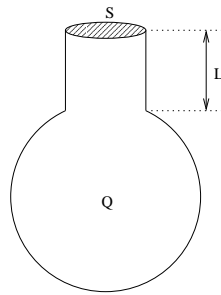$$Z_{\text{tube}}(s) = \left(\frac{\rho L}{S}\right)s, \tag{4.53}$$

Figure 4.11: A Helmholtz resonator, with tube length L, cross section S, and volume Q

where L denotes the length of the tube, and S the cross section. Similarly, the acoustic impedance of the cavity is

$$Z_{\text{cavity}}(s) = \left(\frac{\rho c^2}{Q}\right)\frac{1}{s}, \tag{4.54}$$

with Q the volume of the cavity. Combining these two in series gives the impedance of the Helmholtz resonator:

$$Z_{\text{Helmholtz}}(s) = Z_{\text{tube}}(s) + Z_{\text{cavity}}(s) = \left(\frac{\rho L}{S}\right)s + \left(\frac{\rho c^2}{Q}\right)\frac{1}{s}. \tag{4.55}$$

The Helmholtz resonator is the acoustic equivalent of a mass-spring system. This very simple structure is used as a resonator in many musical instruments. Examples of its use include the ocarina (which is in essence a tunable Helmholtz resonator), the resonator of a sitar, or the body of a guitar.

4.4 Acoustic systems

A real musical instrument is a combination of the acoustical elements discussed in the previous sections. The instrument's bore can be well approximated by a sequence of cylindrical and conical sections. The input impedance of the complete structure can be formulated in terms of transmission matrices, quite similar to the methods used in RF waveguide systems. The input impedance of the different sections can be found with the knowledge of the pressure $P_0(\omega)$ and the volume flow $U_0(\omega)$. These are found as:

$$\begin{bmatrix} P_0(\omega) \\ U_0(\omega) \end{bmatrix} = T \begin{bmatrix} P_L \\ U_L \end{bmatrix}, \tag{4.56}$$

where the transmission matrix T depends on the type of section considered. For cylindrical sections, T is given by

$$T_{\text{cyl}} = \begin{bmatrix} \cos(kL) & jZ_0 \sin(kL) \\ \frac{j}{Z_0} \sin(kL) & \cos(kL) \end{bmatrix}, \tag{4.57}$$

and for a diverging conical section

$$T_{\text{dcone}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{4.58}$$

with

$$a = \frac{R_{L_0}}{R_L} \left[\frac{L_0}{L + L_0} \cos(kL) - \frac{L_0}{k(L + L_0)^2} \sin(kL) \right], \tag{4.59}$$

$$b = \frac{L_0}{L + L_0} jZ_0 \sin(kL), \tag{4.60}$$

$$c = \frac{j}{R_L} \left[\left(\frac{L_0}{L + L_0} + \frac{1}{k^2(L + L_0)^2} \right) \sin(kL) - \frac{L}{k(L + L_0)^2} \cos(kL) \right], \tag{4.61}$$

$$d = \frac{L_0}{L + L_0} \left[\cos(kL) + \frac{1}{kL_0} \sin(kL) \right], \tag{4.62}$$

and for a converging conical section

$$T_{\text{ccone}} = \begin{bmatrix} d & b \\ c & a \end{bmatrix}. \tag{4.63}$$

The transmission matrices can be cascaded to calculate the complete structure of N sections:

$$\begin{bmatrix} P_0 \\ U_0 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = T_1 T_2 T3 \dots T_N \begin{bmatrix} P_L \\ U_L \end{bmatrix}. \tag{4.64}$$

The input impedance is then found as

$$Z_{\text{IN}} = \frac{\beta + \alpha Z_L}{\delta + \gamma Z_L}, \tag{4.65}$$

with Z_L the load impedance after the last section, illustrated in figure 4.12. Using this method, we calculate the theoretical behavior of wind instruments in chapter 6.

The concatenation of different acoustic elements is more important in the case of wind instruments, as these are inherently built of different parts. For string instruments, the acoustical systems mainly consist of a link between the vibrating string, and a resonating cavity. Here, it is important to keep the same wave variables at the boundary of the different element. For instance, the displacement waves on the string are transformed to a force on the bridge, resulting in the bridge velocity and pressure waves in the air inside the body. There is no specific difficulty if one sees the system as parallel and series connections of impedances.
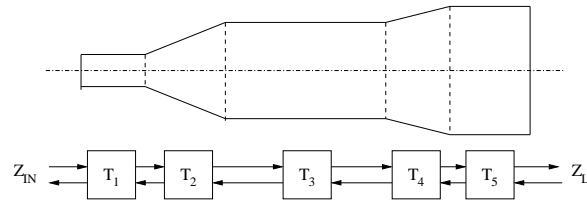
Figure 4.12: A acoustic system consists of several connected elements, each represented by a transmission matrix.

4.5   Conclusion

In this chapter, the general theory of linear acoustics, as applicable to musical instruments, has been reviewed. We have discussed wave propagation in air and in strings, and the most relevant acoustic elements. The continuous-time solutions of this chapter will be translated to discrete-time filter structures in the next chapter.

Chapter 5

Digital Waveguide Models
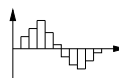
It's quite ironic: We got rid of our analog equipment, replaced it
with digital, then spent the next couple of decades trying to get
the digital to sound like the analog we got rid of.

David Williams

In this chapter, we briefly discuss the basics of digital waveguide systems. Much
work has already been published on the subject of digital waveguides [97],[109].
We first introduce the concept of digital waveguides as a discrete traveling-wave
solution of the wave equation in section 5.1. We then convert the acoustic
elements discussed in section 4.3 to their discrete digital waveguide equivalents
in section 5.2 and build complete digital waveguide systems in section 5.3.

This chapter provides the discrete-time models of the structures described in
chapter 4. These models will be combined in later chapters into efficient models
for complete instruments. The discretization of time and space introduces
artefacts and inaccuracies that must be dealt with, and leaves us with a series
of design options.

5.1   Digital waveguides

Digital waveguide modeling is a simple and efficient way to simulate one-
dimensional wave propagation. The use of digital waveguides was first proposed
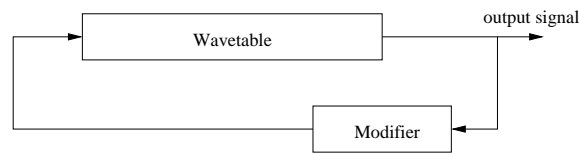
Figure 5.1: Karplus-Strong synthesis. A wavetable, filled with a random signal, is periodically read out and modified. This allows the synthesis of a time-evolving sound, starting with a relatively short waveform. Typically, the modifier consists of a lowpass filter.

by Smith [100] for the simulation of reverberation. It is related to Karplus-Strong synthesis, described in the next section, and enables the practical use of physical modeling for sound synthesis. The basic elements of digital waveguide models are non-integer length delay lines, linear filters, and nonlinear elements. In this section, the focus is on the delay lines and the associated fractional delay filters. After a general formulation of digital waveguides in section 5.1.2, several different fractional delay strategies are compared in section 5.1.3. Finally, we briefly discuss the filters inherently needed to model the wave propagation losses and dispersion in section 5.1.4.

5.1.1   Karplus-Strong Synthesis

Karplus and Strong [44] developed a very simple method for the synthesis of plucked string and drum sounds. It is an extension to the standard wavetable synthesis where a sampled sound signal is periodically read from memory. In the case of Karplus-Strong synthesis, the original wavetable is filled with random values and the samples are modified after each readout, as shown in figure 5.1. Typically, some kind of low-pass filtering is used as the modifier, as many musical instruments have an initial attack with significant high-frequency content, but a steady-state output with lower or decaying high frequencies. In essence, the Karplus-Strong method implements a comb filter, excited with a white noise burst. Karplus-Strong synthesis is not much used in its original formulation.

5.1.2   Digital waveguides for physical modeling

A physical modeling interpretation of Karplus-Strong synthesis was given by Jaffe and Smith [41]. From their point of view, the algorithm consists of a high-order digital filter, which represents the string, and a noise burst, which represents the "pluck". Smith later formally described the two-delay digital waveguide as the implementation of the general solution of the one-dimensional wave equation [100], [101], [104], [103]. Sound propagation in air was discussed

in 4.2 and the scalar one-dimensional wave equation was:

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}. \tag{5.1}$$

A general solution of this equation is

$$p(x,t) = f(t - x/c) + g(t + x/c), \tag{5.2}$$

with $f(.)$ and $g(.)$ continuous functions. This is the traveling-wave solution to the one-dimensional wave equation. Digital waveguide modeling is nothing more than a discretization of this general solution in time and space. The time sampling interval T is $T = 1/f_s$, with f_s the sampling frequency, and the spatial sampling X is equal to the distance traveled by a sound wave during this time interval, or cT meters. With this particular choice, during each temporal interval, the traveling wave components move one spatial interval. Equation 5.2 becomes

$$p(t_n, x_m) = f([n - m]T) + g([n + m]T). \tag{5.3}$$

The notation can be simplified by defining

$$p^+(n, m) = f([n - m]T), \tag{5.4}$$

$$p^-(n, m) = g([n + m]T), \tag{5.5}$$

such that

$$p(t_n, x_m) = p^+(n, m) + p^-(n, m). \tag{5.6}$$

Figure 5.2 shows a digital waveguide section that implements the traveling-wave solution with two digital delay lines. The upper delay line simulates the right-traveling wave, and the lower delay line the left-traveling wave. Direct calculation is possible at each spatial sampling interval. In a real application, it is not likely that the desired length is an integer multiple of cT, and an interpolation scheme is required. If we are only interested in some values, and not the complete solution, the different unit-delay elements can be lumped into two long delay lines. Furthermore, as it is a linear system, all linear operations can be commuted to the ends of the delay lines. This provides us with a very efficient system.

5.1.3 Fractional Delay

The use of a fixed sampling frequency leads to an interesting problem. The spatial sampling distance cT, at a sampling frequency of 44.1 kHz, is approximately $7.9mm$. Any length different from an integer multiple of this distance has to be derived be interpolation. Figure 5.3 shows such an non-integer-length delay line. Several techniques were developed to achieve this [50], [109]. Linear interpolation methods are widely used, and are approximations of the ideal
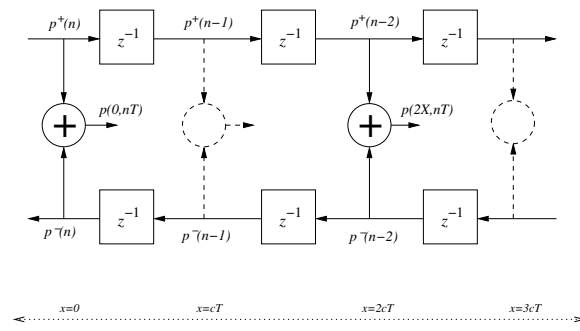
Figure 5.2: Digital waveguide implementation of the traveling-wave solution, using two delay lines. The upper delay line simulates the right-traveling wave, and the bottom delay line, the left-traveling wave. The sum of the outputs of every unit delay of the top and the bottom delay lines is a valid solution of the wave equation.
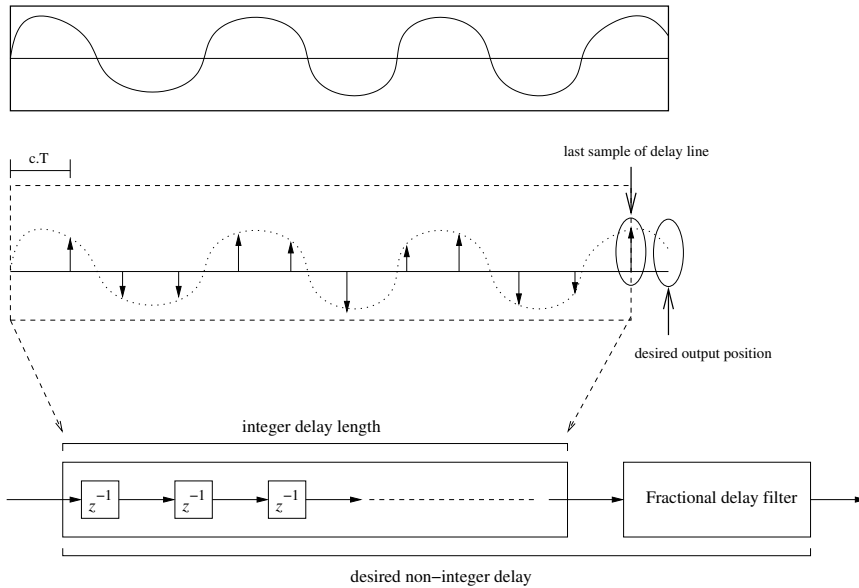

Figure 5.3: Schematic representation of a delay line with non-integer length. (top) The continuous-time, continuous-length tube is spatially discretized with sampling length cT_s (middle). As the length does not translate into an integer number of samples, the discrete version is either too short or too long. Interpolation is needed to simulate the desired output point. (bottom) The true delay is split in a conventional integer-length delay line and a fractional delay filter.

sync interpolator. The most widely used interpolation methods are either all-pass filtering or Lagrange interpolation. We also investigated the use of the cubic Hermite and B-spline interpolation methods.

Lagrange interpolation is polynomial interpolation, and yields an interpolating FIR filter with coefficients given by

$$h(n) = \prod_{k=0, k \neq n}^{N} \frac{D - k}{n - k},$$ (5.7)

with $n = 0, 1, 2, \ldots, N$, D the fractional delay, and N the order of the filter. The fractional delay D is found as

$$D = \frac{L}{cT_s} - \lfloor \frac{L}{cT_s} \rfloor.$$ (5.8)

The coefficients found with this equation have to be normalized to preserve unity gain at DC. The fractional delay and filter order should fall in the range $(N-1)/2 \leq D \leq (N+1)/2$ for odd N and $N/2 - 1 \leq D \leq N/2 + 1$ for even N [109]. The first-order Lagrange interpolator is simply linear interpolation between two samples. Higher-order Lagrange interpolators implicitly use the derivatives of the signal to achieve higher precision, but have the disadvantage that the fractional delay D is larger than one (depending on the order), and thus require the shortening of the "main" delay line. Figure 5.4 compares first- and third-order Lagrange interpolation.

The second method is IIR allpass interpolation. The coefficients of an N-th order allpass filter are given by

$$a_k = (-1)^k \binom{N}{k} \prod_{n=0}^{N} \frac{D - N + n}{D - N + k + n}$$ (5.9)

where

$$\binom{N}{k} = \frac{N!}{k!(N-k)!},$$ (5.10)

The coefficient of a first-order allpass is thus

$$a = -\frac{D-1}{D+1}.$$ (5.11)

The filter is given by

$$\mathcal{H}_{\text{allpass}}(z) = \frac{a + z^{-1}}{1 + az^{-1}}.$$ (5.12)

Allpass interpolation presents no attenuation, but has a somewhat higher error in its phase characteristic, and the time-domain behavior results in artefacts. Figure 5.5 compares the delay of the first-order Lagrange interpolators and the
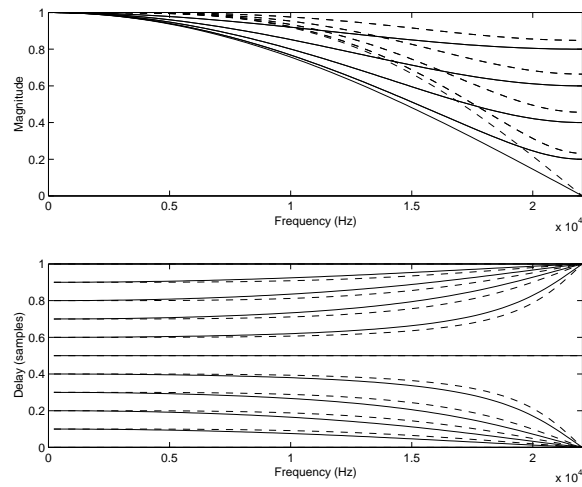
Figure 5.4: Frequency response and delay in samples of (solid lines) a first order Lagrange interpolator, for delay values $D = 0, 0.1, 0.2, \ldots, 1$, and (dashed line) a third order Lagrange interpolator for delay values $D = 1, 1.1, 1.2, \ldots, 2$. $F_s = 44.1kHz$. To make the two responses comparable, the delay of the 3rd order filter was decreased by one. The third order Lagrange filter has a significantly lower damping, and a flatter delay up to 5 kHz.
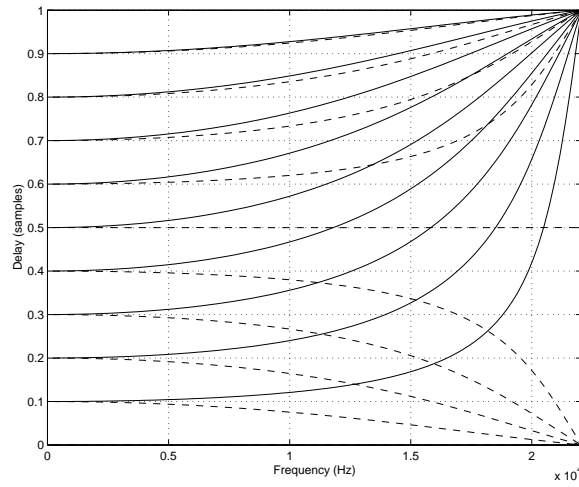
Figure 5.5: Comparison of the delay of (solid lines) first-order allpass filters, and (dashed lines) first-order Lagrange interpolators, for $D = 1, 1.1, 1.2, \ldots, 2$ and $F_s = 44.1 kHz$. Both filters are quite equivalent up to 3 kHz. Allpass interpolation degrades very fast after 5 kHz, making it unsuitable for higher pitched notes.

first-order allpass filters. For frequencies lower than 2-3 kHz, both are equally fit for use.

Cubic Hermite interpolation, and the specific form of the Catmull-Rom splines, use the derivatives of the signal to determine a cubic spline that interpolates two points. Hermite interpolation is piecewise polynomial interpolation. Each cubic Hermite spline is a linear combination of four basis functions. The Hermite basis functions are

$$
\begin{align}
f_1 &= 2D^3 - 3D^2 + 1, \tag{5.13}\\
f_2 &= -2D^3 + 3D^2, \tag{5.14}\\
f_3 &= D^3 - 2D^2 + D, \tag{5.15}\\
f_4 &= D^3 - D^2. \tag{5.16}
\end{align}
$$

Hermite interpolation uses two control points (P_1 and P_2), and two tangents (T_1 and T_2). In matrix form, we get the interpolated value y as a function of the fractional delay D

$$
y = \begin{bmatrix} D^3 & D^2 & D & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 & 0 \\ -2 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ T_1 \\ T_2 \end{bmatrix}. \tag{5.17}
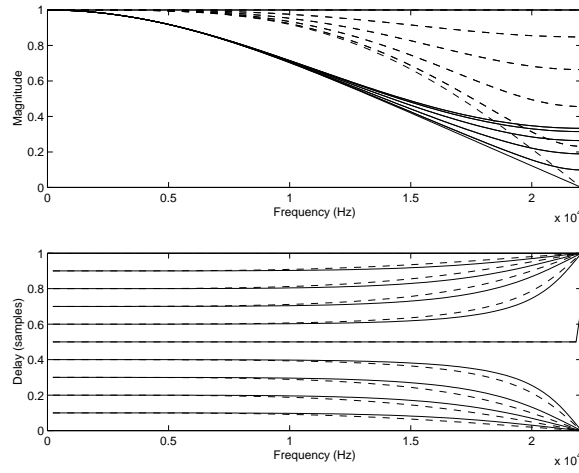$$
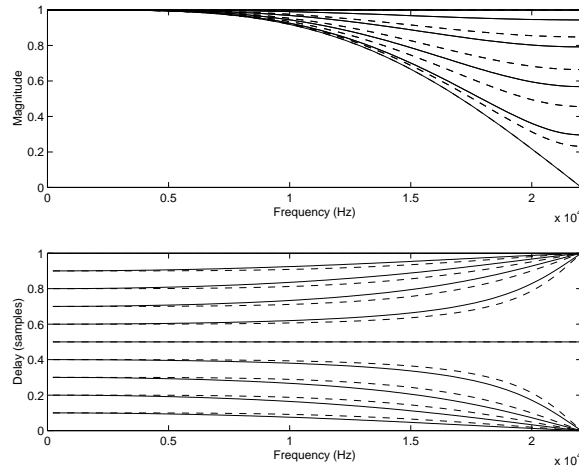
Figure 5.6: Comparison of the magnitude and delay of (solid line) Catmull-Rom cubic interpolation, and (dashed line) 3rd order Lagrange interpolation. Catmull-Rom cubic interpolation exhibits larger dispersion than LI3, but with lower damping.

If there is no extra tangent information, it should be derived from extra data points. With the so-called cardinal splines, the tangent is calculated as

$$T_i = a(P_{i+1} - P_{i-1}), \tag{5.18}$$

where the factor a controls the "tightness" of the spline. Catmull-Rom splines are a special case of cardinal splines, where

$$T_i = \frac{P_{i+1} - P_{i-1}}{2}. \tag{5.19}$$

In matrix form, the interpolated value is found as

$$y = \begin{bmatrix} D^3 & D^2 & D & 1 \end{bmatrix} \begin{bmatrix} -0.5 & 1.5 & -1.5 & 0.5 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix}. \tag{5.20}$$

This is again a FIR filter, for a fixed fractional delay D. Figure 5.6 compares the magnitude and delay characteristics of Catmull-Rom interpolation (CRI) and third order Lagrange interpolation (LI3). CRI has the same delay characteristics as LI1, but significantly better magnitude characteristics than LI3.

The last method studied here are the B-splines. These are not true interpolating polynomials, but approximate the curves, and have better continuity than

Figure 5.7: Comparison of the magnitude and delay of (solid line) B-Spline approximation, and (dashed line) 3rd order Lagrange interpolation

Hermite polynomials. The B-Spline cubic approximator is defined by:

$$y = \begin{bmatrix} D^3 & D^2 & D & 1 \end{bmatrix} \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix}. \tag{5.21}$$

Figure 5.7 compares the magnitude and delay characteristics of B-Splines and LI3. B-Spline approximation has a severe low-pass effect, but a more linear phase characteristic than LI3.

The different methods all represent compromises between delay accuracy, damping and filter order. In table 5.1, we compare the features of the techniques and provide directions as to which method to use for a specific situation.

5.1.4   Losses

In section 4.3.1, we have provided an overview of the losses caused by thermal conductivity and viscous drag. The easiest way to account for losses in the digital waveguide implementation, is to lump all losses to the ends of the delay lines. The filter that represents lossy wave propagation over the distance cT is

$$H(z) = e^{-\alpha cT} z^{-c/v_p} \tag{5.22}$$

with α and v_p as defined in section 4.3.1. For a total delay of m samples (bore of length mcT), we can represent the commuted losses, without the main delay

Type	Order	Damping	Range	Situation	Example
LI1	1st	high	5 kHz	medium-high damp.	clarinet
Allpass	1st	zero	3 kHz	low damp., low freq.	piano (low)
LI3	3rd	medium	10 kHz	medium damp.	guitar
CR	3rd	low	7 kHz	high freq.	piano (high)
B-Spline	3rd	very high	12 kHz	high damp., high freq.	violin

Table 5.1: Comparison of the different methods: first order Lagrange (LI1), 3rd order lagrange (LI3), Allpass, Catmull-Rom (CR), and B-Spline interpolation. The range value is the highest frequency with acceptable deviation of the intended delay. Partials above this value experience a different delay, resulting in dispersion. Low order filters are computionally cheaper, but have the lowest useable frequency range. First order Lagrange interpolation can be used in situations with medium-high damping and frequencies up to 5 kHz - this covers most wind instruments.

of z^{-m}, as

$$H(z) = e^{-\alpha mcT} z^{m(1-c/v_p)}. \tag{5.23}$$

For the case of air and at $300K$, this is well approximated by

$$H(z) = e^{-\alpha mcT} z^{m1.045/r_v} \tag{5.24}$$

The loss filter can be commuted or combined with the fractional delay and reflection filters.

5.2 Discrete Acoustic Elements

5.2.1 Cylindrical bores

In section 4.3.1, we discussed the wave propagation in cylindrical bores. A cylindrical bore, terminated with a reflectance $\mathcal{R}_L(z)$ can be modeled as shown in figure 5.8. The length of the delay line depends on the sampling rate used. The spatial sampling distance corresponds to the distance traveled by the sound wave in air during one sampling interval, or

$$L_s = \frac{c}{F_s} \tag{5.25}$$

where F_s is the sampling rate and c the speed of sound in air. The digital waveguide equivalent length of a cylindrical bore with length ℓ is

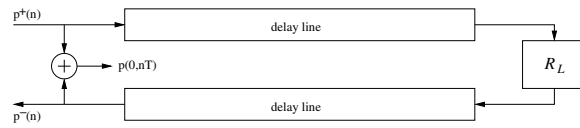$$L = \frac{\ell}{L_s} = \frac{\ell c}{F_s}. \tag{5.26}$$

Figure 5.8: Two-delay-line digital waveguide implementation of ideal plane-wave propagation in a cylindrical bore, with reflection filter R_L. The right-traveling wave is represented by the top delay line, consisting of a chain of unit delays and a fractional delay filter. The signal is reflected back into the bottom waveguide, depeding on the load impedance.
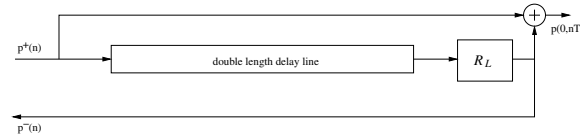

Figure 5.9: Single-delay-line version of figure 5.8. The load impedance of figure 5.8 has been commuted with the bottom delay line. This results in one delay line of double length. Note that only one fractional delay filter is needed in this case.

The output of the right-traveling waveguide is filtered by the reflection filter R_L and reflected back into the left-traveling waveguide.

Example 5.1 The spatial sampling distance at 44.1 kHz can be found with equation (5.25), and is 7.78 mm. A bore with a length of $50cm$ results in a delay line of length 64.26 samples. This can be implemented as two delay lines of 64 samples, and two additional fractional delay filters with $D = 0.26$. The response of a cylindrical bore of length 25cm and 50cm can be found on ⋆ **CD Tracks 1 and 2**. △

An even simpler model uses only one delay line (of double length), as shown in figure 5.9.

5.2.2   Conical bores

The acoustic properties of the conical bore were discussed in section 4.3.1. The cone is defined by the total cone length, the length of the truncated part, and the ratio of the smallest and largest radii, consistent with the definitions in section 4.3.1. The continuous-time traveling-wave solution for a truncated cone, shown in figure 5.11, is given by

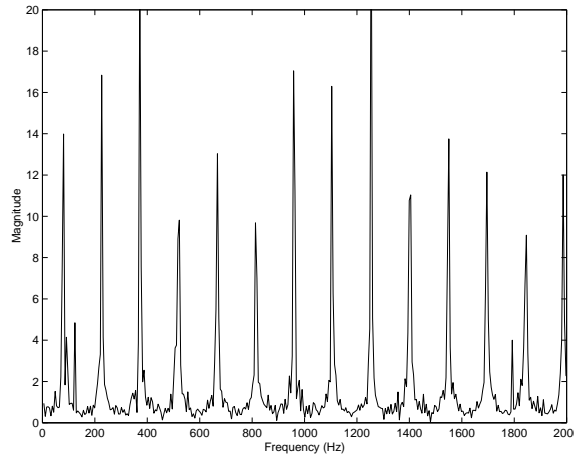$$p(x,t) = \frac{f(t - x/c)}{x} + \frac{g(t - x/c)}{x}. \tag{5.27}$$

Figure 5.10: The magnitude of the spectrum of a two-delay-line digital wave-guide, exited with white noise. The two delay lines are 150 samples long. The reflection filter is a scalar, $R_L = 0.99$, and the sampling frequency F_s=44.1 kHz. This is a simulation of an open-open cylindrical bore, with fundamental frequency of 73.5 Hz, and odd harmonics.

This expression is similar to that for plane waves with the exception that the traveling-wave components are inversely proportional to their distance form the cone apex. This solution can be discretized as [103]

$$p(t_n, x_m) = \frac{f([n-m]T)}{x} + \frac{g([n+m]T)}{x}. \qquad (5.28)$$

Assuming as a boundary condition that the pressure is equal to zeros at the open end of the cone, we can construct the digital waveguide implementation shown in figure 5.12. Aside from the $1/x_0$ scaling, the digital waveguide implementations of the cylindrical and conical bore appear to be identical. The difference occurs at discontinuities: waves traveling in a truncated cone experience a complex acoustic impedance, causing a phase difference between the pressure and the volume flow.   The pressure wave reflectance seen from the entrance of a conical bore terminated with a load impedance Z_L is given by

$$\mathcal{R}(x) = e^{-2jkL} \frac{Z_L Z_0^*(L) - Z_0(0) Z_0^*(L)}{Z_L Z_0(L) + Z_0(L) Z_0^*(L)}, \qquad (5.29)$$

with k the wave numer, and with $Z_0(0)$ and $Z_0(L)$ the complex impedance of the cone at its beginning or end, respectively, determined with equation (4.30). Figure 5.12 corresponds to the case where $Z_L = 0$. The boundary conditions for a closed apex impose that the volume components reflect with an inversion, and the pressure component reflects with a phase angle $2\angle Z_0(x)$. In the case
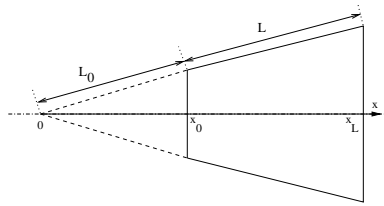
Figure 5.11: Cross-section of a truncated cone. The cone has a total length of $L - L_0$ and is defined by the ratio of the radii at x_0 and x_L.
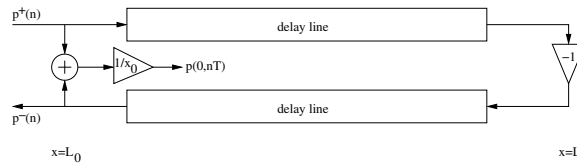

Figure 5.12: Two-delay-line digital waveguide implementation of ideal spherical wave propagation in a conical bore, assuming zero pressure at the conical end.

of a complete cone, this is a 180° phase shift, but for the truncated cone, the situation is more complicated. The pressure reflectance at L_0 is found to be

$$R_0(s) = \frac{p^+}{p^-} = \frac{L_0 s - c}{L_0 s + c}. \tag{5.30}$$

This expression can be discretized using the usual methods, and represents an allpass structure. The filter is then incorporated into the model shown in figure 5.13.

Example 5.2 The output of a complete cone, excited at the apex with white noise, can be found on $\star$ CD Track 3. Compare this track ti the output of the cylindrical bore.   △
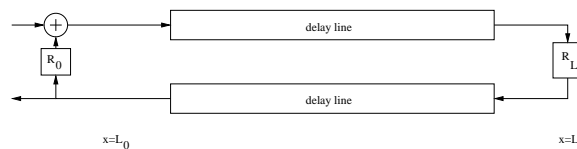
Figure 5.13: Two-delay-line digital waveguide implementation of a truncated conical bore with reflection R_0 at the truncated start and R_L at the open end.
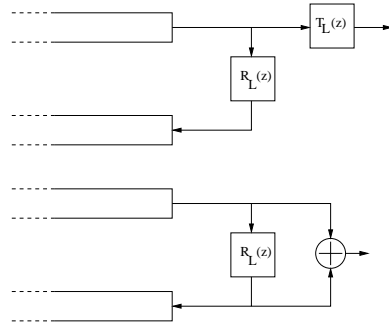
Figure 5.14: Integration of ends into digital waveguide models: (top) complementary transmittance filter, (bottom) residual sound output. The right-traveling wave (in the top delay line) is reflected back according to the reflection filter R_L. The transmittance is either explicitely defined (top), or calculated as the residual signal (bottom).

5.2.3 Ends

The acoustic properties of unflanged and flanged ends were discussed in section 4.3.2. Flaring ends (exponential or Bessel horns) inherently require multi-dimensional wave propagation, and are not efficiently modeled by digital waveguides. It is often enough to model the end as a flanged, non-flaring end, represented by a load impedance Z_L

$$Z_L(\omega) = A(\omega) + jB(\omega), \tag{5.31}$$

with

$$A(\omega) \;=\; Z_0 \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(ka)^{2n}}{(n!)^2(n+1)}, \tag{5.32}$$

$$B(\omega) \;=\; \frac{Z_0}{\pi(ka)^2} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(2ka)^{2n+1}}{\left(\frac{(2n-1)!}{2^{n-1}(n-1)!}\right)^2 (2n+1)}, \tag{5.33}$$

with a the mouth radius, and $k = \omega/c$ the wave number. The reflection is then easily calculated with

$$R_L = \frac{Z_L - Z_0}{Z_L + Z_0}, \tag{5.34}$$

and the transmission

$$T_L = 1 - \mathcal{R}. \tag{5.35}$$

These are quite cumbersome expressions, and the transmission is often approximated by a simple linear first order low-pass filter, with cut-off frequency

corresponding to the wave propagation cut-off given by

$$\omega_{3\mathrm{dB}} = \min_{x} c \sqrt{\frac{1}{a(x)} \frac{d^2 a(x)}{dx^2}}, \tag{5.36}$$

with $a(x)$ the local radius of the cross-section of the end. The reflection and transmission can be explicitly included as filters, or the transmittance can be calculated as the residual of the reflection, as shown in 5.14.

5.2.4 Strings

For transversal waves, strings are almost equivalent to cylindrical bores. The equation governing the transversal wave was given in section 4.2.2. The length of the delay line equivalent to a string with length ℓ is

$$L = \frac{\ell F_s}{c_{\mathrm{trans}}}, \tag{5.37}$$

with c_{trans} the propagation speed of transversal waves, or, when starting from a desired fundamental frequency f_0 and a string with fixed ends,

$$L = \frac{F_s}{2 f_0}. \tag{5.38}$$

Strings also exhibit nonlinear effects due to their stiffness and length change during motion. Digital waveguide systems for the simulation of non-ideal strings are discussed in chapter 10.

5.3 Discrete acoustic systems

To model complete instruments, we need to concatenate a sequence of different cylindrical and conical bores. At the boundary of two discontinuous sections, there is a change in acoustic impedance, resulting in partial reflection and transmission. The reflectance is calculated by

$$\mathcal{R} = \frac{Z_2 - Z_1}{Z_2 + Z_1}. \tag{5.39}$$

If we consider a sequence of two cylindrical sections, this can be simplified to

$$\mathcal{R} = \frac{\mathcal{S}_1 - \mathcal{S}_2}{\mathcal{S}_1 + \mathcal{S}_2}, \tag{5.40}$$

where $\mathcal{S}_{1,2}$ represents the cross-sectional area of the two cylinders. A possible digital waveguide implementation of such a scattering junction is shown in figure 5.15. The junction is passive if $-1 \leq \mathcal{R} \leq 1$. It is clear from equation (5.40)
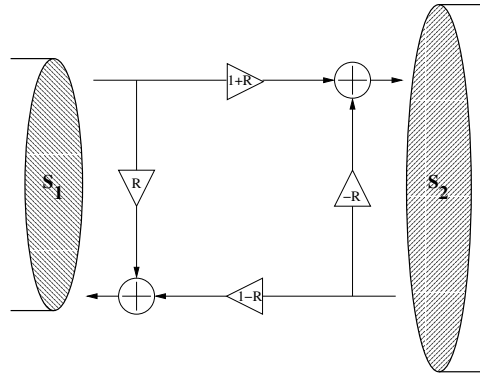
Figure 5.15: An implementation of a scattering junction at the boundary of two cylindrical section with different radius. The reflection and transmission are simple scalars.

that this is always the case. Conical bores yield slightly more complicated expressions, as the scattering coefficients become frequency-dependent. Filling in the expressions for the input and output impedance of truncated cones in equation (5.39) yields the reflection and transmission filters:

$$\mathcal{R}^- = \frac{B-1}{B+1} - \frac{2B\gamma}{(B+1)(j\omega + \gamma)}, \tag{5.41}$$

$$\mathcal{R}^+ = -\frac{B-1}{B+1} - \frac{2\gamma}{(B+1)(j\omega + \gamma)}, \tag{5.42}$$

$$\mathcal{T}^- = 1 + \mathcal{R}^+, \tag{5.43}$$

$$\mathcal{T}^+ = 1 + \mathcal{R}^-, \tag{5.44}$$

where B is the ratio of the surface area of the two cones at the junction $\mathcal{S}_1/\mathcal{S}_2$ and γ is given by

$$\gamma = -\frac{c}{\mathcal{S}_1 + \mathcal{S}_2}\left(\frac{S_1}{x_1} - \frac{S_2}{x_2}\right). \tag{5.45}$$

Note that in this formulation, the scattering junction filters are unstable for negative γ. This makes a sequence of expanding and contracting conical sections especially difficult to describe - fortunately, this only occurs with very few instruments (the English horn and *oboe d'amore* being notable exceptions).

It is in general better to calculate the total input impedance of the system, as discussed in section 4.4, and to convert the continuous-time variables to the digital waveguide variables. The conversion is given by:

$$\begin{bmatrix} P \\ U \end{bmatrix} = \begin{bmatrix} P^+ + P^- \\ \frac{P^+}{Z_0} - \frac{P^-}{Z_0^*} \end{bmatrix}. \tag{5.46}$$
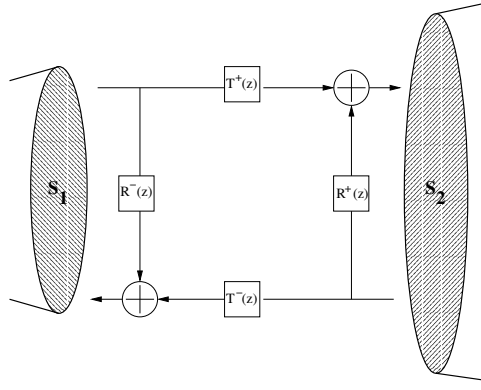
Figure 5.16: An implementation of a scattering junction at the boundary of two conical sections with different radius. Contrary to the cylindrical case, we now have non-scalar reflection and transmission filters.

Applying this to the two bores at the junction, the continuous-time transmission matrix of which is given by

$$\begin{bmatrix} P_1 \\ U_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P_2 \\ U_2 \end{bmatrix}, \tag{5.47}$$

where the transmission matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ depends on the bore type (cfr. section 4.4), one obtains

$$\begin{bmatrix} P_1^- \\ P_2^+ \end{bmatrix} = \begin{bmatrix} \mathcal{R}^+ & \mathcal{T}^- \\ \mathcal{T}^- & \mathcal{R}^+ \end{bmatrix}, \tag{5.48}$$

for the waveguide system, where

$$\mathcal{R}^- = \frac{Z_1^*}{\kappa}(BZ_2^* - DZ_1Z_2^* + AZ_2Z_2 - CZ_1Z_2Z_2^*), \tag{5.49}$$

$$\mathcal{T}^- = \frac{Z_1^*}{\kappa}(AD - BC)(Z_1Z_2 + Z_1Z_2^*), \tag{5.50}$$

$$\mathcal{R}^+ = \frac{Z_2}{\kappa}(BZ_1 + DZ_1Z_1^* - AZ_1Z_2^* - CZ_1Z_1^*Z_2^*), \tag{5.51}$$

$$\mathcal{T}^+ = \frac{Z_2}{\kappa}(Z_1^*Z_2^* + Z_1Z_2^*), \tag{5.52}$$

with

$$\kappa = Z_1Z_2(B + Z_1^* + AZ_2 + CZ_1^*Z_2). \tag{5.53}$$

The resulting coefficients can then be approximated by simpler IIR filters. Systems containing strings are built by linking compatible wave variables, essentially in the same way as for digital waveguide systems for wind instruments.

5.4   Conclusion

In this chapter, we have introduced the digital waveguide modeling technique used for the computationally efficient modeling of musical instruments in discrete time. After the general description of the traveling-wave solution, we have compared several fractional delay strategies. Besides the established techniques of Lagrange and allpass interpolation, we also studied the cubic Hermite and B-spline interpolation, and indicated in which situations a particular method can be used. The acoustic elements of chapter 4 were converted to discrete time models, and the method used to concatenate the elements into digital waveguide system was explained.

Part II

Wind Instruments

Chapter 6

Complete Acoustic Model

The greater you understand the structure of something, the more amazed you'll be at the tiniest movement within it. In that sense the possibilities are limitless.

Brian Eno

In this chapter, a complete acoustic model for a single-reed instrument is built. The instrument consists of two parts: the resonator, needed to sustain and acoustically amplify the partials of the sound, and the excitation mechanism. In section 6.1, the resonator of the modern clarinet and the alto saxophone is discussed, without the toneholes. The toneholes are the subject of chapter 8. In section 6.2, the movement of the reed is studied. We start by modeling the reed as a memoryless spring system, and continue with second- and higher order linear and nonlinear models. This is combined with the aerodynamical equations for the airflow to obtain a single-reed model. Finally, the resonator model is combined with the single-reed model to form a complete instrument. This chapter provides a continuous-time, single-note reference model that will be approximated in the next chapters.
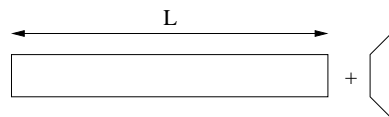
Figure 6.1: The simplest acoustic model of a clarinet consists of a long cylindrical bore and a conical bell.

6.1 The resonator

The simplest approximation of the clarinet body is a cylindrical tube combined with a small conical flare, as shown in figure 6.1. This model ignores the effect of the toneholes. The acoustic properties of a cylindrical bore were discussed in section 4.3.1. When the reed channel is closed, we obtain a closed-open cylinder, with resonant frequencies of

$$f = \frac{(2n-1)c}{4L}. \tag{6.1}$$

The bell adds dispersion at the higher frequencies. The input impedance of the resonator can be calculated using the transmission matrices determined in section 4.4. Figure 6.2 shows the input impedance of such a resonator, for three different bells. At low frequencies, the bell has very little influence (cfr. section 4.3.2). Adding the viscothermal losses described in sections 4.3.1 and 5.1.4 results in the damping of the high frequencies (cfr. equation (5.24)).

The resonator is then discretized using the digital waveguide methods described in chapter 5. At the boundary between the cylindrical and conical parts of the resonator, a scattering junction is needed as described in section 5.3. In this particular case, there is no diameter discontinuity, and the scattering filters reduce to simple lowpass and highpass filters. Filling in $B = 1$ in equations (5.41) to (5.44) yields

$$\mathcal{R}^-(s) = \frac{-\gamma}{(s+\gamma)}, \tag{6.2}$$

$$\mathcal{R}^+(s) = \frac{-\gamma}{(s+\gamma)}, \tag{6.3}$$

$$\mathcal{T}^-(s) = \frac{s}{(s+\gamma)}, \tag{6.4}$$

$$\mathcal{T}^+(s) = \frac{s}{(s+\gamma)}, \tag{6.5}$$

with $\gamma = c/(2x)$ (x is the length of the truncated part of the conical section). After discretization of the filters, we obtain a structure shown in figure 6.3. The addition of toneholes represents a significant increase in model complexity.
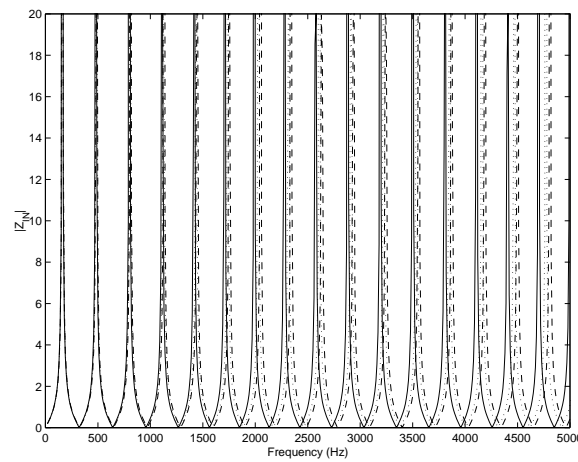
Figure 6.2: A comparison of the input impedance of the simple resonator of figure 6.1, with $L = 0.5m$, $L_{\text{bell}} = 5cm$, for three different cone opening angle (solid line) 30°, (dotted line) 45°, (dashed line) 60°, and ignoring losses. The input impedances were determined using the methodology of section 4.4.

Several methods will be discussed in chapter 8. The resonator of a saxophone is completely similar, except that one uses a conical section instead of a cylindrical one.

6.2 The excitation mechanism

The single reed combined with the mouthpiece structure, shown on figure 6.4, acts as a pressure-controlled valve, allowing energy into the instrument for the initialization and maintenance of oscillations in the resonator [46],[97]. It can be seen as a mechanical oscillator of the blown-shut type (see figure 6.5). The operation of the single-reed excitation mechanism can be broken down in two parts: the mechanical movement of the reed, and the aerodynamical airflow through the tip opening. Of course, the two are interrelated, but a first approximation can be obtained by describing both separately.

6.2.1 Reed movement

The reed is a sheet of cut cane, thin at the tip and gradually thicker towards the heel. The reed movement is controlled by the pressure difference between the player's oral cavity and the pressure in the reed channel or inside the mouthpiece. As a first approximation, the pressure in the oral cavity is considered
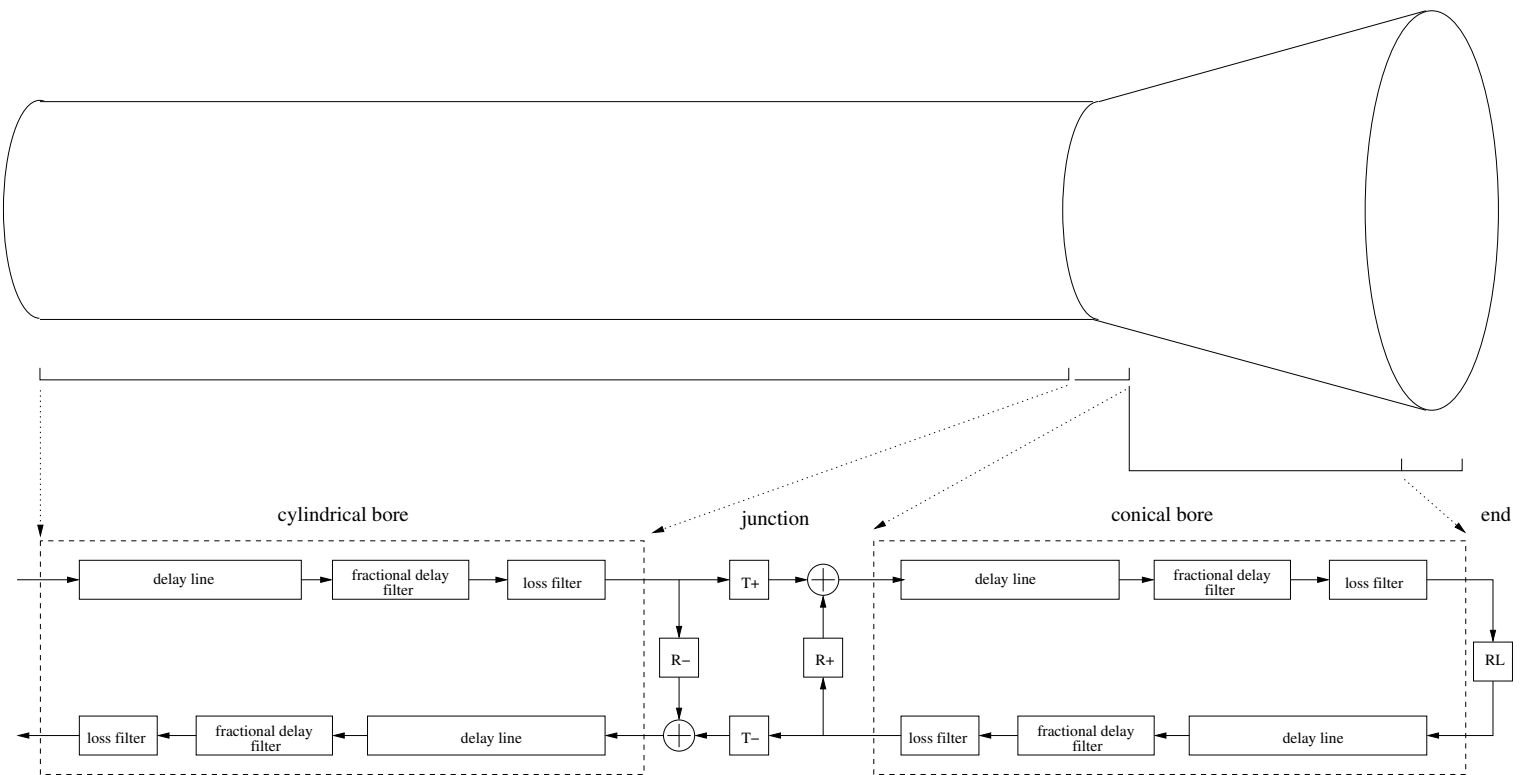
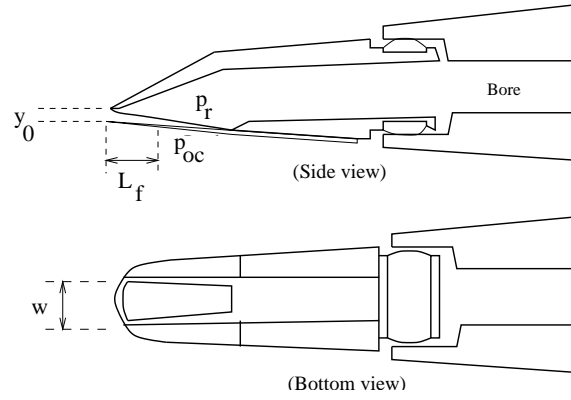Figure 6.3: A complete waveguide model for the resonator of a clarinet.

Figure 6.4: Side view and bottom view of the reed-mouthpiece structure. We define the rest opening y_0, the reed facing length L_f and width w. The pressure inside the mouthpiece, directly above the reed is p_r, and the pressure in the oral cavity of the player is p_{oc}.
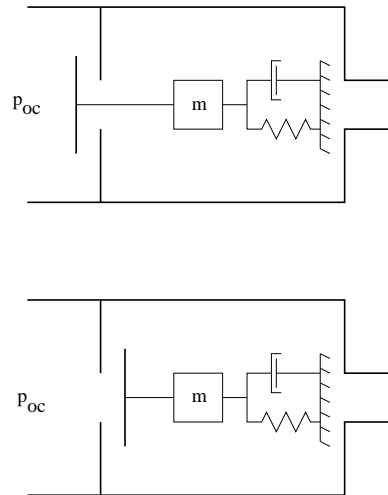

Figure 6.5: Schematic representation of (top) blown-shut (single-reed), and (bottom) blown-open type mechanical oscillator

constant and controlled by the player.

Simple model

A common simplification for woodwind instruments has been to neglect the effect of the mass of the reed. This is equivalent to selecting a reed with infinite resonance frequency. The force F_r on the reed is

$$F_r = wL_f p_\Delta \tag{6.6}$$

where w is the width of the reed, L_f the facing curve length, and p_Δ the pressure difference between the pressure in oral cavity of the player p_{oc} and the pressure in the mouthpiece p_r. The displacement of the reed from its rest position y_0, which is also the tip opening of the mouthpiece quoted by the manufacturers, is simply determined by Hooke's law

$$\Delta y = \frac{F_r}{k}, \tag{6.7}$$

with k the spring constant of the reed. The actual tip opening is

$$y = y_0 - \Delta y. \tag{6.8}$$

Example 6.1 For an alto saxophone mouthpiece and reed, a typical rest opening is in the range of 0.5-1 mm, and values for k range from 1000-2000 N/m. The reed width is w=16 mm. For a middle-of-the-road reed/mouthpiece combination with L_f=10mm, y_0=0.7 mm, k=1500 N/m, the force required to close the reed opening is 1.05 N, or a pressure difference of p_Δ=6562 Pa.   △

Second order model

A more precise approach is to describe the reed as a complete damped mass-spring system. The reed movement for a given applied force is described by the differential equation

$$m\frac{d^2y}{dt^2} + \mu\frac{dy}{dt} + ky = F_r(t). \tag{6.9}$$

The transfer function is easily found to be

$$\mathcal{H}_{\text{reed}}(s) = \frac{1}{ms^2 + \mu s + k} = \frac{1/m}{s^2 + 2\xi\omega_n s + \omega_n^2}. \tag{6.10}$$

This is implemented as an IIR filter. The parameters m, μ and k can be measured by several methods.

Reed	$k(Nm^{-1})$	$\mu(s^{-1})$
Vandoren Classic 2	1020	150
Vandoren Classic $2\frac{1}{2}$	1170	220
Vandoren Classic 3	1660	280
Vandoren Classic $3\frac{1}{2}$	2150	380

Table 6.1: Measured parameters of several alto saxophone reeds, averaged over several reeds of the same strength. For this particular manufacturer, half a "strength" more means a reed with a 1.3 times larger spring constant and damping. This is expected, as reeds of the same type (here "Classic") differ only in thickness.

Higher order model

A reed is in essence a thin bar (cantilever), clamped at one end, with a force applied on its surface. The partial differential equation that describes this situation is

$$EI\frac{\partial^4 y(x,t)}{\partial x^4} = -wp_\Delta(t) - \rho\frac{\partial^2 y(x,t)}{\partial t^2}, \tag{6.11}$$

with E Young's modulus, and I the moment of inertia, or

$$\rho\frac{\partial^2 y(x,t)}{\partial t^2} + EI\frac{\partial^4 y(x,t)}{\partial x^4} = -wp_\Delta(t). \tag{6.12}$$

We first consider the homogenous equation

$$\rho\frac{\partial^2 y(x,t)}{\partial t^2} + EI\frac{\partial^4 y(x,t)}{\partial x^4} = 0. \tag{6.13}$$

We write the reed displacement $y(x,t)$ as

$$y(x,t) = X(x)T(t), \tag{6.14}$$

such that equation 6.13 becomes

$$\frac{T(t)''}{T(t)} + \frac{EI}{\rho}\frac{X(x)''''}{X(x)} = 0, \tag{6.15}$$

or

$$\frac{T(t)''}{T(t)} = -\frac{EI}{\rho}\frac{X(x)''''}{X(x)} = -\omega^2. \tag{6.16}$$

This separates in two equations

$$T(t)'' + \omega^2 T(t) = 0, \tag{6.17}$$

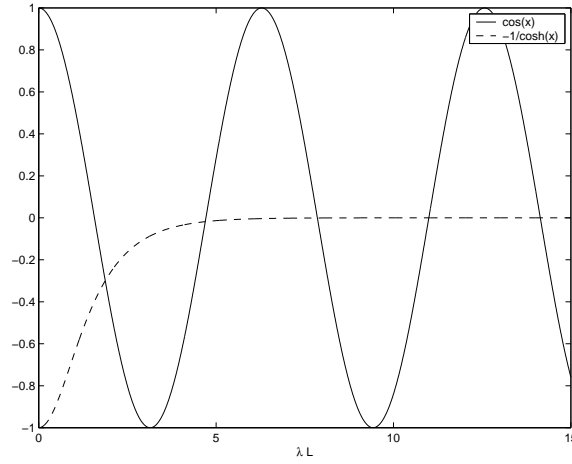$$X(x)'''' - \frac{\rho}{EI}\omega^2 X(x) = 0. \tag{6.18}$$

Figure 6.6: Graphical representation of the condition for the non-trivial solutions. The crossings are the roots of the nonlinear equation.

The solutions for these equations are

$$
\begin{aligned}
T(t) &= C_1 \sin(\omega t) + C_2 \cos(\omega t) && \text{(6.19)} \\
X(x) &= C_3 \sinh(\lambda x) + C_4 \cosh(\lambda x) + C_5 \sin(\lambda x) + C_6 \cos(\lambda x) && \text{(6.20)}
\end{aligned}
$$

where $\lambda = \left(\frac{\omega^2 \rho}{EI}\right)^{1/4}$. The boundary conditions for a cantilever of length L, fixed at one end, translate to

$$
\begin{aligned}
y(0,t) &= X(0) = 0, && \text{(6.21)} \\
y'(0,t) &= X'(0) = 0, && \text{(6.22)}
\end{aligned}
$$

for the displacement and speed of the fixed end, and

$$
\begin{aligned}
y''(L,t) &= X''(L) = 0, && \text{(6.23)} \\
y'''(L,t) &= X'''(L) = 0, && \text{(6.24)}
\end{aligned}
$$

for the free end. Applying these boundary conditions on equation (6.20), and for the non-trivial solution ($\omega \neq 0$), one gets the condition

$$
\cos(\lambda L)\cosh(\lambda L) + 1 = 0. \tag{6.25}
$$

Figure 6.6 shows the graphical solution of this equation.

A good numerical approximation can be found for both small and large values of λL. For small λL, we determine the Taylor expansion around $\lambda L = \pi/2$ and

λL (solved)	λL (approx.)	ω
1.875	1.862	1370 Hz
4.694	4.712	8568 Hz
7.859	7.853	24 kHz
10.996	10.996	47 kHz
14.137	14.137	77 kHz

Table 6.2: Roots of equation 6.25 with corresponding natural resonance frequencies, calculated for a facing length of $8mm$, reed width of $17mm$ and thickness $0.4mm$. The left column shows the result of the numerical solution, the middle column the result obtained with the approximations.

get

$$-1/\cosh(\lambda L) \approx -\mathrm{sech}(\frac{\pi}{2}) + \mathrm{sech}(\frac{\pi}{2})\tanh(\frac{\pi}{2})\left(\lambda L - \frac{\pi}{2}\right) + \mathcal{O}\left\{(\lambda L - \frac{\pi}{2})^3\right\},$$
(6.26)

$$\cos(\lambda L) \approx -\left(\lambda L - \frac{\pi}{2}\right) + \mathcal{O}\left\{(\lambda L - \frac{\pi}{2})^3\right\}.$$
(6.27)

Solving equation with these approximations yields

$$\lambda L = \frac{-\pi - 2\mathrm{sech}(\frac{\pi}{2}) - \pi\mathrm{sech}(\frac{\pi}{2})\tanh(\frac{\pi}{2})}{-2 - 2\mathrm{sech}(\frac{\pi}{2})\tanh(\frac{\pi}{2})} = 1.862.$$
(6.28)

For larger values of λL, $-1/\cosh(\lambda L) \approx 0$, and thus the solutions are

$$\lambda L = \pi/2 + k\pi,$$
(6.29)

with $k = 1, 2, \ldots$. The values for the numerical solution and approximations of λL and resonance frequencies obtained are shown in table 6.2. Only the lower two modes are important, as the higher modes fall outside the human hearing range, which implies that it is possible to model the cantilever approximation of the reed movement at the tip of the reed with two second-order models, with resonances at the first two natural resonances of the cantilever.

We now examine the static solution of the differential equation. At the cut-off pressure (static pressure at which the tip opening is closed), the reed should lay smoothly on the mouthpiece face. If we consider that the pressure exerts a constant force $p_\Delta w$ per surface area of the reed, the reed shape is

$$\Delta y(x) = \frac{p_\Delta w x^2 (6L_f^2 - 4L_f x + x^2)}{24D},$$
(6.30)

where x is the position along the reed, L_f is the facing length of the mouthpiece, and D is the flexural rigidity of the reed given by
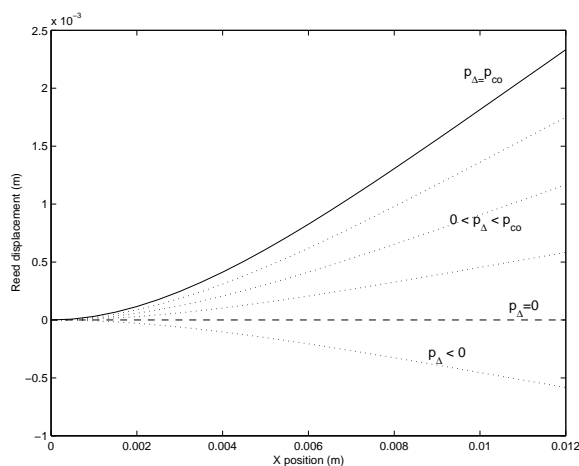
$$D = \frac{Eh^3}{12(1 - \nu^2)},$$
(6.31)

Figure 6.7: Reed displacement for different values of p_Δ for a tip opening at rest of $2.8mm$ and a facing length of $12mm$. Note that the displacement near the tip of the reed is lineary proportional to the pressure difference. This justifies the use of the low-order models for the simulation of the airflow through the tip opening.

with h the thickness, E the Young modulus, and ν the Poisson ratio (ratio of transverse contraction strain to longitudinal extension strain in the direction of stretching force) of the reed [51]. This shape ensures the smoothest possible closing. The rigidity can be determined for a given cut-off pressure p_{co}: the maximum displacement

$$\max(\Delta y) = \frac{p_{co} w L_f^4}{8D} \qquad (6.32)$$

is then equal to the tip rest opening y_0, or

$$D = \frac{p_{co} w L_f^4}{8y_0}. \qquad (6.33)$$

This rigidity value corresponds to the "ideal" reed for the given facing length and tip opening. Harder or softer reeds assume a different shape. The shape of the ideal reed for different values of p_Δ is shown in figure 6.7. If the mouthpiece does not follow this shape, the reed will touch it, limiting the movement of that part of the reed. This changes the effective facing length, and thus the spring constant and resonance frequency. The same holds if a softer or harder reed is used. Figure 6.8 shows the measured shape of the curve of two alto saxophone mouthpieces with different tip opening, and figure 6.9 compares the shape of the ideal reed for a facing length of $22mm$ and a tip opening of $2.413mm$, at cut-off pressure, with the measured shape of a Berg Larsen 95/1 SMS stainless steel
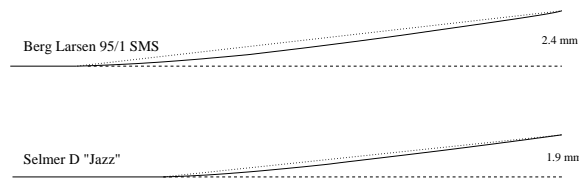
Figure 6.8: Measured shape of the curve of a (top) Berg Larsen 95/1 SMS, and (bottom) Selmer D "Jazz" metal alto saxophone mouthpieces.
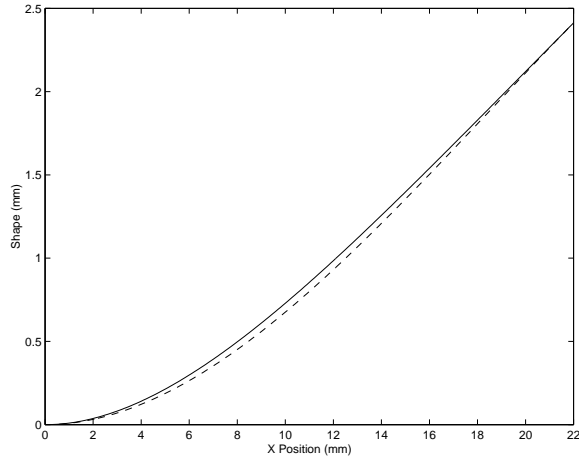

Figure 6.9: Comparison of (solid line) the theoretical shape of the reed at p_{co}, and (dashed line) the measured shape of a Berg Larsen 95/1 SMS alto saxophone mouthpiece. The shapes are the same to within the measurement error, suggesting that manufacturers found this shape to be optimal.

alto saxophone mouthpiece. The shapes match up to the measurement error on the shape of the mouthpiece. This suggests that mouthpiece manufacturers, knowingly or not, determined that this theoretical shape is the optimal one for a given tip opening and facing length. In a practical playing situation, the musician will partly close the tip opening by pressing the reed with his/her lips, such that only a part of the curve is actually used, as shown in figure 6.10. This however strongly depends on the personal playing style of the musician. For a given facing length and tip opening, it is now possible to calculate the "strength" of the reed that will provide the player with the largest dynamic range.
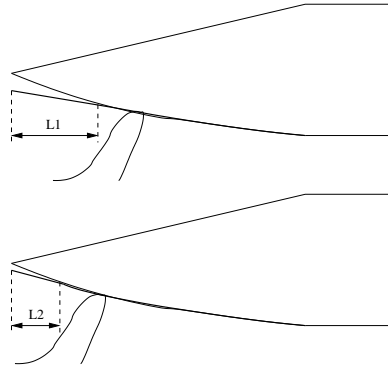
Non-linear model

Figure 6.10: The player usually only uses a part of the facing length by pressing closer to the mouthpiece at different places.

The previous systems assume that the properties of the reed stay constant over the movement range of the reed. If the reed is completely linear, one expects to measure an impulse spectrum comprised of damped sinusoids with constant damping. It was experimentally determined that this is not the case, and that the damping is amplitude-dependent [29]. At larger amplitudes, the damping is significantly (a factor 2) larger than at small amplitudes. A good fit for the observed data is of the form

$$\mu(A) = (\alpha A + \beta)\mu_0, \tag{6.34}$$

where A denotes the amplitude, α and β are the parameters, and μ_0 the low-amplitude damping. Typical values for $\alpha = 1.25$ and $\beta = 0.75$ were found for reeds of different brands and hardness. Note that this is not an acoustically correct description, but an engineering model that produces appropriate output.

6.2.2   Airflow

The airflow over the reed into or out of the mouthpiece chamber is primarily determined be the surface of the opening S_o and the pressure difference p_Δ. The volume flow is determined from the Bernoulli equation and is described by [6],[120]

$$u_r = S_o \text{sgn}\,(p_\Delta)\sqrt{\frac{2p_\Delta}{\rho}}. \tag{6.35}$$

A key assumption behind the derivation of this equation is that the air jet separates from the wall at the end of the reed channel, forming a free jet into the mouthpiece. As a result, the pressure in the reed channel p_r is equal to the
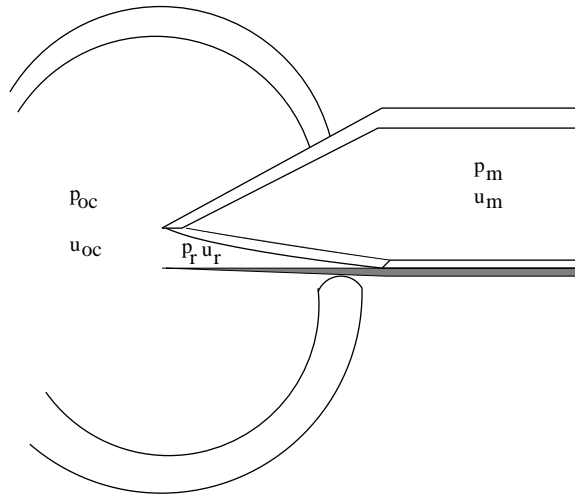
Figure 6.11: The different pressures (p) and volume flows (u) considered in the airflow model. We differentiate between the oral cavity area (oc), the reed channel (r), and mouthpiece (m).

pressure in the mouthpiece p_m. This is justified by the large abrupt transition in cross-sectional area from reed channel to mouthpiece.

The influence of blowing pressure on the playing frequency of the clarinet has been systematically studied in [7]. Bak *et al.* observed an increase of frequency with increasing blowing pressure. This is easily incorporated by slightly modifying the rest opening for higher blowing pressures.

6.2.3 Single-reed models

We now have enough information to build an approximate model for the complete single-reed structure. If we use the simple motion model, and combine equations (6.7) and (6.35), we get that

$$u_r = w(y_0 - \frac{wL_f p_\Delta}{k})\text{sgn}\,(p_\Delta)\sqrt{\frac{2p_\Delta}{\rho}}, \qquad (6.36)$$

or with cut-off pressure $p_{co} = \frac{ky_0}{S_r}$:

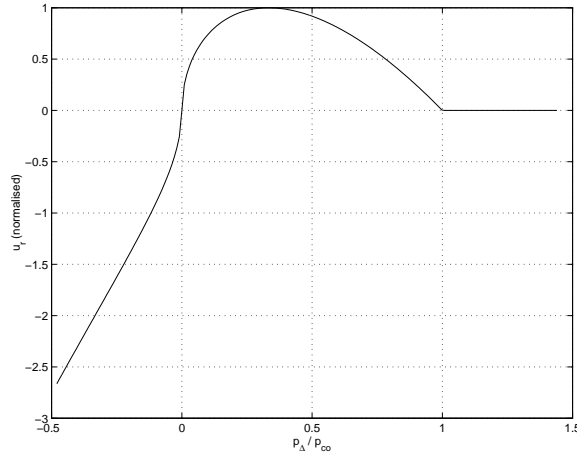$$u_r = wy_0(1 - \frac{p_\Delta}{p_{co}})\text{sgn}\,(p_\Delta)\sqrt{\frac{2p_\Delta}{\rho}}. \qquad (6.37)$$

Figure 6.12: Normalized total volume flow through a single-reed structure, with p_Δ relative to the cut-off pressure p_{co}. For $p_\Delta/p_{co} < 0$, the flow increases as the pressure difference increases, as there is no limit to the displacement of the reed. When $p_\Delta = 0$, there is no airflow at all. For $0 < p_\Delta < p_{co}$, the airflow first increases due to the increase pressure difference, and then decreases because the surface of the tip opening is decreasing. At $p_\Delta/p_{co} = 1$, the reed channel is closed and no air can flow.

The total airflow through the reed channel is converted into a pressure by the acoustic impedance of the bore

$$p = u_r Z_{\text{bore}}. \tag{6.38}$$

Figure 6.12 shows the total flow. For the steady-state (DC) flow, the input impedance of the bore can be considered zero, such that p_Δ can approximated by p_{oc}. The static flow is now given by

$$u_{rs} = wy_0(1 - \frac{p_{oc}}{p_{co}})\sqrt{\frac{2p_{oc}}{\rho}}, \tag{6.39}$$

and is shown in figure 6.13. The dynamic component is

$$u_{rd} = \begin{cases} u_{rs}\left(1 + \frac{p_r}{p_{co}-p_{oc}}\right)\sqrt{1 - \frac{p_r}{p_{oc}}} & \text{if } p_{oc} > p_r > p_{oc} - p_{co} \\ 0 & \text{if } p_r \leq p_{oc} - p_{co} \\ -u_{rs}\left(1 + \frac{p_r}{p_{co}-p_{oc}}\right)\sqrt{\frac{p_r}{p_{oc}} - 1} & \text{if } p_{oc} > p_r > p_{oc} - p_{co} \end{cases} \tag{6.40}$$

Several different implementations exist for single-reed models. These will be discussed in section 7.2.2.
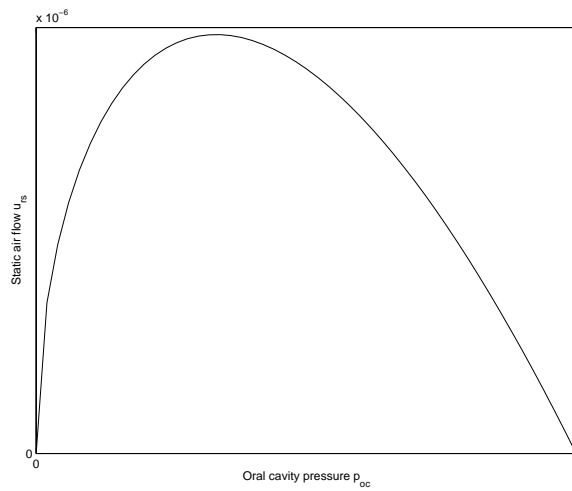
Figure 6.13: Static volume flow. The total airflow through the bore of the instrument consists of this static (DC) airflow (carrying no audio information), and the dynamic flow given by equation (6.40). This figure essentially shows the operating point chosen by the player for a particular oral cavity pressure.

6.3 The complete instrument

To obtain a complete acoustic model of the instrument, the excitation mechanism has to be connected to the resonator. The input of the instrument model are the oral cavity pressure p_{oc} and the reed opening at rest y_0. These two variables allow the calculation of the airflow through the reed channel. This airflow is converted by the bore input impedance into a pressure. This pressure determines the pressure difference between mouthpiece and oral cavity, and propagates along the bore. This presents no particular difficulties if a valid implementation of the airflow equations is chosen. The output of the mouthpiece/reed model is connected to the inputs of the digital waveguide model of the bore, and vice versa. This will be done in more detail in the next chapter.

Example 6.2 $\star$ CD Track 4 is the output of a complete instrument model of the clarinet. It consists of the single-reed excitation mechanism coupled to a cylindrical bore. $\star$ CD Track 5 is generated with the same single-reed model, but coupled to a conical bore, resulting in an approximation of a saxophone. $\triangle$

6.4   Conclusion

In this chapter, a complete acoustic model for a single-reed instrument was built. We first discussed a simple resonator model, composed of a cylindrical bore and a conical bell. It was shown that the fundamental frequency is largely unaffected by the presence of the bell. However, we notice a slight shift for the higher partials, resulting in an inharmonic spectrum. We studied both low- and high order linear and non-linear models for the reed movement. The higher-order linear models showed that the movement of the tip under static pressure is proportional to the applied pressure, and that only two vibration modes are within audible range. This allows the use of one or two simple second-order models to model the reed behavior. Finally, we described the airflow through the reed tip opening. The airflow, and thus outgoing pressure, depends on the pressure difference between the oral cavity of the player, and the pressure inside the mouthpiece, and is strongly non-linear. In this chapter, the resonator and excitation mechanism were discussed as essentially independent elements. In the next chapter, a holistic solution will be investigated.

Chapter 7

Formant Filtered Model

Jazz is the only music in which the same note can be played night
after night but differently each time.

Ornette Coleman

In the previous chapter, complete acoustic models for the clarinet and the saxophone were discussed. Here, we present practical implementations of simplified models, and take a closer look at the calibration issues. We combine the traditional physical modeling view with source-filter synthesis by adding a dynamic formant filter.

The models from the previous chapter were built from the acoustics perspective, using the physical properties of the different elements of the instrument to obtain a model. In this chapter, we approach the same problem from the synthesis standpoint, and see how both approaches can be combined. For synthesis, simplicity is the key. Unfortunately, the simplest models have a very poor degree of realism, and thus more elements derived in the previous chapter have to be introduced. By exploiting the formant-like spectrum of wind instruments and the efficient digital waveguide synthesis, we obtain a new model class that offers a high sound quality, while circumventing the problems of a pure acoustic solution. This is achieved by adding a dynamic equalization filter, or Formant filter, to a digital waveguide model.

From the synthesis standpoint, the simplest existing models for wind instrument sound synthesis are the McIntyre-Schumacher-Woodhouse models.
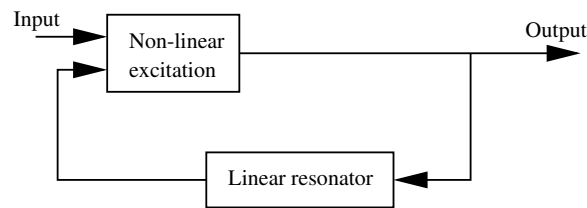
Figure 7.1: The McIntyre, Schumacher, and Woodhouse model simplifies the full acoustic model by dividing the instrument in a linear and a non-linear part.

7.1   McIntyre-Schumacher-Woodhouse model

A simple and generic physical model has been studied by McIntyre, Schumacher, and Woodhouse (MSW) [60]. In MSW, tone production can be divided into two main parts: a nonlinear excitation and a linear resonator (see figure 7.1). For a given instrument, MSW synthesis models the objects and actions as a compact set of equations. The most complicated and instrument-specific equations describe the excitation, while the linear resonator primarily consists of a delay line and a linear filter. To obtain realistic results with this approach, a different model is needed for each note. The effective bore length (and thus the time delay) changes for different notes, as does the tonehole configuration and thus the internal reflection function. Notes played in a different register by using register holes or a different embouchure on the real instrument can also be modeled with a different bore length.

The structure of these models is very much like the full acoustic model structure derived in the previous chapter, but instead of having one model for all notes, we have one model per note. In the next section, we determine the different parts of this model, and see how the calibration of the non-linear excitation can be simplified by the addition of a formant filter.

7.2   The Formant filtered model

By combining source-filter synthesis with waveguide synthesis, we obtain a model that is both accurate and easy to calibrate.

In source-filter synthesis, or subtractive synthesis, the sound is obtained by filtering an excitation signal with a time-varying filter. It was mostly used to produce speech, but also has musical applications [63],[64],[91]. The classical source-filter synthesis is depicted in figure 7.2. The idea is to have a broadband or harmonically rich excitation signal which is filtered to obtain the desired
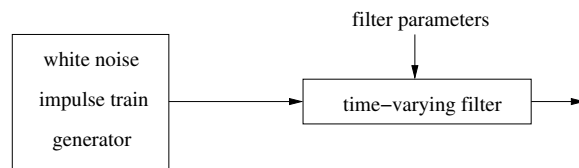
Figure 7.2: Source-filter synthesis. A time-varying filter is excited with white noise or an impulse train. The control parameters are the filter coefficients and the pitch of the impulse train.
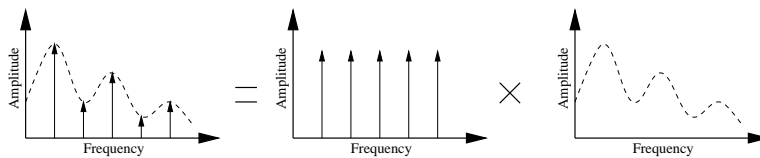

Figure 7.3: The formant-like structure of the spectrum of a wind instrument can be synthesized by filtering the pre-whitened output of a digital waveguide model.

signal. Typically, a white noise or impulse train generator is used.

Wind instruments have a spectrum that is harmonically rich, and the spectral envelope of the partials is very formant-like. A formant is a (band-limited) concentration of energy in the spectrum of the signal. Formants are well-known features in speech recognition and synthesis, and specific formant wave-functions were developed for musical application, mostly for the synthesis of singing voices [92].

By exploiting the formant-like spectrum of wind instruments and the efficient digital waveguide synthesis, we obtain a model that has better calibration properties than plain digital waveguide models. We achieve this by introducing a formant filter. This formant filter is a post-equalization filter that corrects the amplitudes of the partials such that they match a recording. This is quite similar to source-filter synthesis, but the excitation for the filter is a waveguide model of a wind instrument. The principle is illustrated in figure 7.3 This method has the advantage that the transients are mostly affected by the waveguide model, and the spectral envelope of the steady-state sound is controlled by the formant filter.

7.2.1 Linear part

The linear part of the model represents the instrument bore, the toneholes, and the bell. As we are only modeling one note, it is possible to lump all filtering effects together in one filter, and all time-delay into one delay line.

The time delay originates mainly in the length of the bore. Its length can be calculated with

$$L = \frac{F_s}{\chi f_0}, \tag{7.1}$$

where F_s the sampling frequency, χ the integer proportionality factor depending on the bore type (conical or cylindrical) and f_0 the fundamental frequency of the note being modeled. It is clear that L seldomly has an integer value. The different fractional delay implementation were explained in section 5.1.3. We used a first order Lagrange (linear) interpolator because of its simplicity (refer to table 5.1). The coefficients of the general Lagrange interpolator are given by

$$h(n) = \prod_{k=0, k \neq n}^{N} \frac{D - k}{n - k}, \tag{7.2}$$

where D is the fractional delay and N the order of the filter, $n = 0, 1, \ldots N$. The first order FIR filter $(N = 1)$ is then calculated as

$$h(n) = \begin{bmatrix} h(0) & h(1) \end{bmatrix} = \begin{bmatrix} -(D - 1) & D \end{bmatrix}. \tag{7.3}$$

Due to the physical structure of the bore, the toneholes, the bell, dispersion, and losses, the frequencies of the partials might be inharmonic. A very similar situation exists when modeling string instruments: the stiffness of the string introduces inharmonic behavior for the partials. We can easily correct these small frequency differences by using an allpass filter [41]. The allpass filter introduces a frequency-dependent phase delay, leading to slightly different frequencies for the higher partials. The bore length yielding an n-th partial of frequency ν_r is

$$L = \chi n \frac{c}{\nu_r}, \tag{7.4}$$

where c is the speed of sound in air. The frequency dependent bore length difference between the recording (with frequencies ν_r) and the synthetic sound (ν_s) can now be calculated as

$$\Delta L = \chi n \left(\frac{c}{\nu_r} - \frac{c}{\nu_s} \right). \tag{7.5}$$

The phase $\theta(\nu)$ of the all-pass filter now becomes

$$\theta(\nu) = 2\pi \frac{\Delta L}{c} \nu_s. \tag{7.6}$$

Several techniques exist for the design of all-pass filters [78]. In this model, we used a simple first order all-pass filter of the form

$$\mathcal{H}_{ap}(z) = \frac{a + z^{-1}}{1 + az^{-1}}, \tag{7.7}$$

where a is a real coefficient, fitted such that the phase of $\mathcal{H}_{ap}(z)$ follows equation (7.6). Note that this filter is not able to correct all the partials individually. The slight inharmonicity of the partials is almost inaudible.

7.2.2   Non-linear part

Several non-linear models for the single-reed mechanism were discussed in section 6.2. Here, we choose one specific implementation for the excitation mechanism. A good choice is the second order reed motion model with constant parameters. There are several possible implementations:

- direct implementation of the non-linear function;

- using a pressure-dependent reflection coefficient;

- using a reed-reflection polynomial.

Non-linear function

The linear and non-linear models discussed in section 6.2 can be directly implemented. It is however not trivial to obtain a good implementation. All the models are based on the pressure difference $p_\Delta = p_{oc} - p_r$. The pressure in the mouthpiece is determined by the incoming volume flow u_r,

$$u_r = wy \cdot \mathrm{sgn}\,(p_\Delta) \sqrt{\frac{2|p_\Delta|}{\rho}}, \tag{7.8}$$

the displacement flow u_{rs} caused by the reed motion,

$$u_{rs} = S_r \frac{dy}{dt}, \tag{7.9}$$

and the incoming pressure from the bore. Combining these gives the outgoing pressure wave

$$p^+(t) = Z_{0b}\left(u_r(t) - u_{rs}(t)\right) + p^-(t) \tag{7.10}$$

Solving this equation to p^+ yields several solutions (depending on the sign of p_Δ)

$$p^+ \quad = \quad \frac{wyZ_{0b}\left(wyZ_{0b} \pm \sqrt{(wyZ_{0b})^2 - 4\rho p^- - 2\rho p_{oc} + 2\rho Z_{0b} u_{rs}}\right)}{\rho}$$
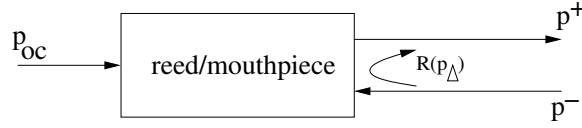
Figure 7.4: The reed/mouthpiece structure as a black box. The output of the reed/mouthpiece structure can be analytically determined, or approximated by a reflection function.

$$-Z_{0b}u_{rs} + p^-, \tag{7.11}$$

$$p^+ = \frac{wyZ_{0b}\left(-wyZ_{0b} \pm \sqrt{(wyZ_{0b})^2 - 4\rho p^- + 2\rho p_{oc} + 2\rho Z_{0b}u_{rs}}\right)}{\rho}$$

$$-Z_{0b}u_{rs} + p^-. \tag{7.12}$$

If we disregard the displacement flow, which is two orders of magnitude smaller than the incoming volume flow, this simplifies to

$$p^+ = \frac{wyZ_{0b}\left(wyZ_{0b} \pm \sqrt{(wyZ_{0b})^2 - 4\rho p^- - 2\rho p_{oc}}\right)}{\rho} + p^-, \tag{7.13}$$

$$p^+ = \frac{wyZ_{0b}\left(-wyZ_{0b} \pm \sqrt{(wyZ_{0b})^2 - 4\rho p^- + 2\rho p_{oc}}\right)}{\rho} + p^-. \tag{7.14}$$

The analytical calculation of the reflection coefficient $\mathcal{R}(p_\Delta)$ of figure 7.4 is non-trivial when using the complete non-linear function.

Pressure-dependent reflection coefficient

An efficient alternative was proposed in [101], where the reed/bore junction is modeled with a memory-less reflection coefficient that varies in response to the pressure changes. The reed system is represented by a lumped acoustic impedance. Assuming continuity of the volume flow, it is easy to see that

$$p^+ = \frac{p_{oc}}{2} + \mathcal{R}(p_\Delta)\left(p^- - \frac{p_{oc}}{2}\right), \tag{7.15}$$

where the pressure-dependent reflection coefficient is calculated as

$$\mathcal{R}(p_\Delta) = \frac{Z_r(p_\Delta) - Z_{0b}}{Z_r(p_\Delta) + Z_{0b}}, \tag{7.16}$$

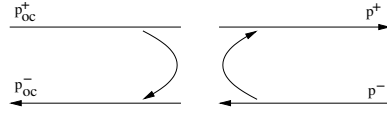with Z_r the acoustic impedance of the reed channel and Z_{0b} the characteristic impedance of the bore.

Figure 7.5: The scattering junction used to determine the reed-reflection polynomial model. The pressure difference $\tilde{p}_\Delta$ is assumed to be independent of the outgoing pressure p^+.

Reed-reflection polynomial

Another technique approximates the excitation mechanism by a polynomial [20],[21]. The reflection and transmission properties of the junction are assumed to depend on $\tilde{p}_\Delta = p_{oc}^+/2 - p^-$. Figure 7.5 shows the reed-bore junction used considered to determine this model. When the reed channel is, perfect non-inverting reflection occur at the junction. The reflection is approximated by a polynomial such that

$$p^+ = p_{oc}^+ - \left(c_1\tilde{p}_\Delta + c_2\tilde{p}_\Delta^2 + c_3\tilde{p}_\Delta^3\right). \tag{7.17}$$

Combination

The Pressure-Dependent Reflection Coefficient method in combination with a damped mass-spring system (see figure 7.6) allows the use of natural parameters (reed stiffness, resonant frequency, reed tip/mouthpiece distance) and is inherently passive, thus improving the numerical stability. The reflection coefficient is calculated as

$$\mathcal{R} = \frac{Z_r - Z_b}{Z_r + Z_b}, \tag{7.18}$$

where Z_r is the acoustic impedance of the reed/mouthpiece structure and Z_b represents the acoustic impedance of the instrument bore. Z_r is determined by

$$Z_r = \frac{p_\Delta}{u_r}, \tag{7.19}$$

where the pressure difference $p_\Delta = p_{in} - p^-$ between the oral cavity and the tone chamber results from the feedback and input. The model is free of delayless loops as there is no implicit feedback with this definition of p_Δ, so no special mathematical problems occur [14]. Note that in reality, the pressure difference is given by $p_\Delta = p_{in} - (p^+ + p^-)$. Implementing this requires either the introduction of a delay, the analytical solution given by equations 7.13 and 7.14, or an iterative numerical solution. Both choices for p_Δ generate quite similar outputs.
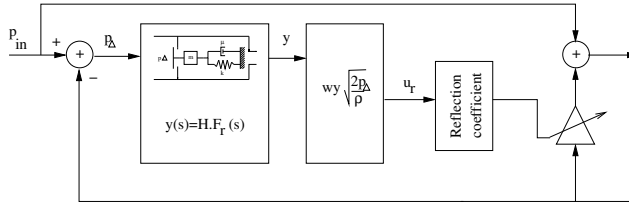
Figure 7.6: The non-linear excitation mechanism. The input and reflected wave determine the pressure difference p_Δ which drives a second-order damped mass-spring system, yielding the new reed opening y. With this opening, the volume flow u_r is calculated. The acoustic impedance of the reed/mouthpiece Z_r can now be calculated with u_r and p_Δ. Knowledge of Z_r leads to the reflection coefficient $\mathcal{R}$.

The volume flow u_r or the airflow through the tip opening is calculated as

$$u_r = w(y_0 - y)\sqrt{\frac{2p_\Delta}{\rho}}, \tag{7.20}$$

(w is the width of the reed, y the tip opening, y_0 the tip opening at rest and ρ the density of air). Further refinements of this formula were proposed in [6] and were discussed in section 6.2. The tip opening was found as the displacement of the reed modeled as a damped mass-spring system with transfer function

$$\mathcal{H} = \frac{A}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \tag{7.21}$$

in which ζ is the damping coefficient and ω_n the natural resonance frequency of the reed. The factor A is chosen such that the displacement y is equal to the rest opening y_0 when the DC input pressure equals the cut-off pressure. The displacement is then

$$y(s) = \mathcal{H}F_r(s), \tag{7.22}$$

where $F_r = S_r p_\Delta$ represents the force on the reed with surface S_r. The actual implementation uses a discretization $H(z)$ obtained by applying the backward Euler transform on (7.21). Figure 7.6 represents the complete non-linear part. The more accurate model described in section 6.2, including displacement-dependent damping as a parameter has also been implemented [29]. Other models can be found in [48], [14].

In the MSW model, we can interpret the linear part as the part that determines the allowed frequencies, and the non-linear part as the distributor of the incoming and reflected energy at these frequencies. Of course, there is some interaction between the two parts through the feedback (see figure 7.1). In the normal playing range of the instrument, the reed behavior can be approximated

by a memoryless function [6],[77],[60]. By doing this, the whole non-linearity
becomes memoryless and will not introduce any phase change in the loop. As
the frequencies of the partials are defined by the total time delay of the loop
(true delay and phase changes), they are independent of the non-linearity itself.
In this light, a fortissimo-played (but not overblown) instrument can be seen
as a fully excited system, in which the amplitude of every allowed frequency
has reached its maximum. It also means that when played on a softer level the
amplitudes of some of the allowed frequencies will be lower, but the frequencies
themselves do not change. Figure 7.7 shows the spectra of a clarinet at three
playing levels. We see that for low-amplitude input, only the fundamental fre-
quency and the first harmonic are present. Louder playing levels introduce
more harmonics. Physically, this is caused by the non-linearity of the reed at
higher displacements. For (theoretical) amplitudes larger than the opening at
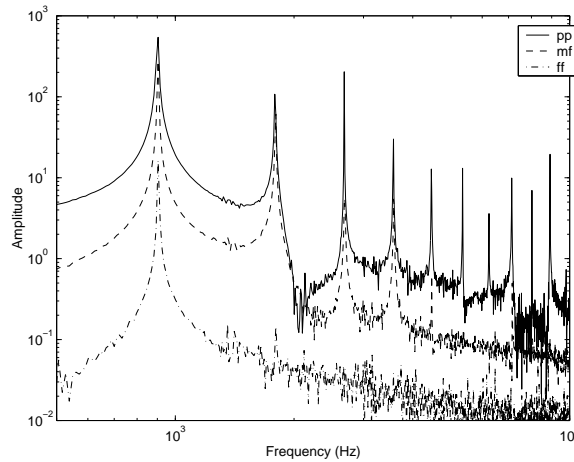rest of the reed, hard clipping (*beating*) occurs.

Figure 7.7: The measured amplitude spectrum of a clarinet played at three
input levels (*pianissimo*, *mezzoforte*, and *fortissimo*). We see that for low-
amplitude input, only the fundamental frequency and the first harmonic are
present. Louder playing levels introduce more harmonics. Physically, this is
caused by the non-linearity of the reed at higher displacements. For (theo-
retical) amplitudes larger than the opening at rest of the reed, hard clipping
(*beating*) occurs.

7.2.3   Formant Filter

The idea is now to split the non-linearity into two parts: a non-linear function,
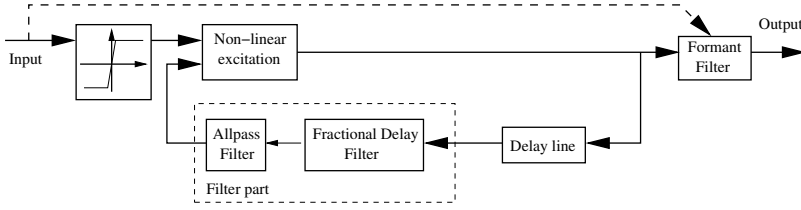with maximum p_{oc} input (but still a variable y_0) and a linear filter, dependent

Figure 7.8: The MSW model has been combined with a formant filter. The physical model now acts as an oscillator with maximum input. The linear part has been split in a delay line and a filter part including a fractional delay filter and the allpass filter. The output goes through the input-dependent formant Filter.

on p_{oc}. The linear filter will adjust the amplitudes of the partials for all possible input levels. By doing this, the player loses direct control over the blowing pressure in the reed model (the player now controls the filter) but still controls the reed opening.

To achieve this division, a formant filter will be used. The first step in the formant filter design is to determine the spectral envelopes of the maximum-input model ($\mathcal{H}_{\text{maximum-input envelope}}$) and of a recording ($\mathcal{H}_{\text{desired envelope}}$). The formant filter is given by:

$$\mathcal{H}_{\text{filter}} = \frac{\mathcal{H}_{\text{desired envelope}}}{\mathcal{H}_{\text{maximum-input envelope}}}, \tag{7.23}$$

Note that the poles of this filter might introduce unacceptable gains, so some precautions might be necessary: adding a high-pass filter to decrease the DC gain, adding a lowpass filter to decrease the high-frequency gain, and damping the resonant peaks. The order of this filter and the methodology used to calculate the envelopes will determine the final precision of the filter. Ideally, this filter should be determined at all possible inputs. In practice, the filter will only be determined for a few input pressures. By using interpolation, one is able to calculate intermediate approximations.

Second order filters

The second-order spectral envelope can be described by

$$\mathcal{H}_{\text{env}} = \frac{A_{\text{env}}\omega_{\text{env}}^2}{s^2 + 2\xi_{\text{env}}\omega_{\text{env}}s + \omega_{\text{env}}^2}, \tag{7.24}$$

in which A_{env} denotes the amplitude, ξ_{env} the damping coefficient ($0 < \xi_{\text{env}} \leq 1$) and ω_{env} the resonant frequency in the absence of damping [68]. The lower limit for the parameter ξ_{env} ($\xi_{\text{env}} > 0$) ensures that a stable filter is obtained.
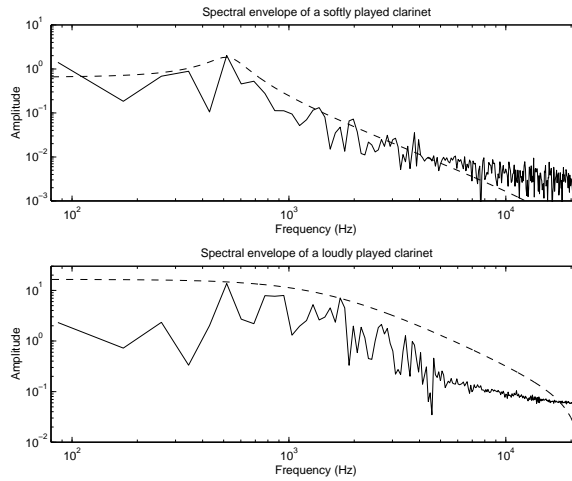
Figure 7.9: The second-order envelope spectrum of equation (7.24) (dashed line) fitted to the spectrum of a softly and loudly played clarinet (full line).

Without the upper limit $\xi_{\text{env}} \leq 1$, the poles of system would split to two different, real poles. The coefficients of this filter can be found by minimizing the weighted least squares problem (WLS) with cost function

$$F = \sum W|S|(|S| - |\mathcal{H}_{\text{env}}|)^2. \tag{7.25}$$

Here, S is the spectrum of the signal as determined by the Short-Time Fourier Transform. W is a scaling factor, chosen large when $|S| > |\mathcal{H}_{\text{env}}|$ and small otherwise. The goal of this cost function is to minimize $(|S| - |\mathcal{H}_{\text{env}}|)^2$, but weighted such that the dominant spectral peaks are accurately followed. This is done by an extra multiplication with $|\mathcal{S}|$: strong peaks have more impact that small peaks or the noise floor. The extra weighting with W is needed to ensure that the spectral envelope is "dressed" on top of those peaks. Standard LS fits the envelope through the peaks in the spectrum: giving more weight to the spectrum above the fitted envelope pulls the fit upwards. The WLS problem is solved with a Gauss-Newton algorithm. Figure 7.9 shows two such spectral envelopes. Plotting the extracted coefficients (see figure 7.10) as a function of time for a soft-to-loud sequence reveals that the spectral envelopes go from highly-resonant (with resonance peak around the fundamental frequency), to fully damped (with -3dB frequency around three times the fundamental frequency). The resulting formant filter according to eq. (7.23) is

$$\mathcal{H}_{\text{filter}} = \frac{A_r \omega_r^2}{A_s \omega_s^2} \frac{s^2 + 2\xi_s \omega_s s + \omega_s^2}{s^2 + 2\xi_r \omega_r s + \omega_r^2}, \tag{7.26}$$
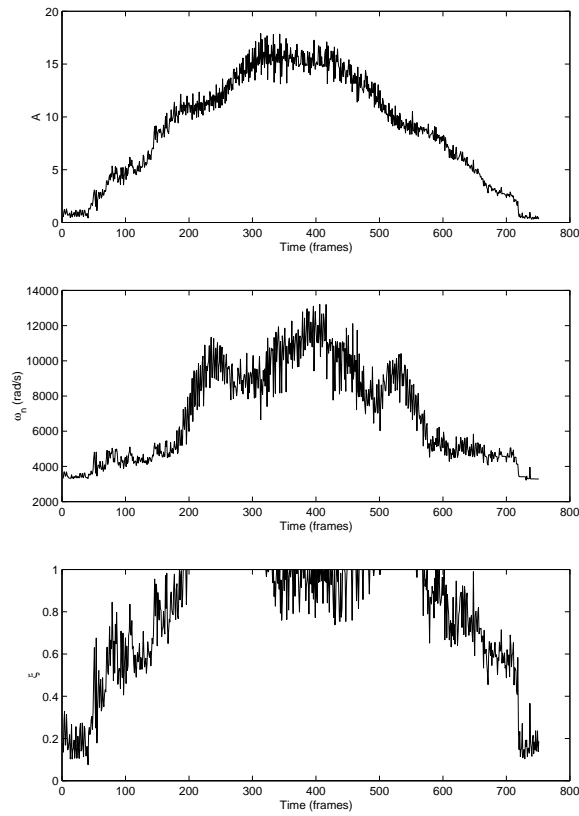
Figure 7.10: Evolution of the envelope parameters for a *pp* to *ff* to *pp* recording. Shown are the extracted amplitude (A_{env}), resonant frequency ($\omega_{n,\text{env}}$) and damping coefficient (ξ_{env}). Note that ξ_{env} clips to its maximal value (1) for the highest amplitudes. The frames are 1024 samples or 23.2 ms long (44.1kHz sampling rate), the complete signal has 17.4 s duration.

in which the subscripts r and s denote the parameters estimated from respectively the recording and the synthetic output.

Linear Predictive Coding

Extension to N-th order is possible using a technique like Linear Predictive Coding. LPC fits an allpole model with transfer function

$$\mathcal{H}_{lpc}(z) = \frac{G}{1 + \sum\limits_{k=1}^{p} a_k z^{-k}}, \tag{7.27}$$

where G is a gain factor. LPC does not exactly yield the spectral envelope, but a best fit according to some cost function. The standard LP method uses least squares as the error measure. Linear prediction has two important spectral matching properties [57]: globally, the matching process should perform uniformly over the whole frequency range, irrespective of the general shaping of the spectrum and locally, the minimization of the error measure results in a model spectrum $\hat{\mathcal{S}}$ that is a good estimate of the spectral envelope of the signal spectrum $\mathcal{S}$. Using the standard LP error measure has the disadvantage of the cancellation of errors: the contributions to the error when $\mathcal{S} > \hat{\mathcal{S}}$ cancel those when $\mathcal{S} < \hat{\mathcal{S}}$. The cost function (7.25) solves this, but has to be minimized using a non-linear optimization routine. When allowing enough poles, the resulting filter $\mathcal{H}_{\mathrm{lpc}}$ still follows the envelope closely when using standard LPC. Note that if too many poles are selected, the result will be the spectrum itself and not the envelope. If the LP order is kept far below the number of partials, this is not an issue. We choose to use this filter as the description of the envelope ($\mathcal{E}_{\mathrm{lpc}}$). An example of an envelope determined with the LPC method can be seen in figure 7.11.

In the low-order case, it is fairly easy to parameterize the filter for the complete input range. For the Nth order LPC case, the spectrum envelope was extracted for only three playing levels ($\mathcal{E}_{pp}$, $\mathcal{E}_{mf}$, $\mathcal{E}_{ff}$) and for the model output ($\mathcal{E}_{synth}$). The formant filter frequency response ($\mathcal{H}_{ff,mf,pp}$) was then calculated as

$$\mathcal{H}_{ff,mf,pp}(z) = \frac{\mathcal{E}_{ff,mf,pp}(z)}{\mathcal{E}_{synth}(z)}. \tag{7.28}$$

These can be found in figure 7.12. The filters representing all other levels were obtained by interpolating between these three levels depending on the input level. This has been validated by comparing the interpolated filters with the filters extracted of a soft-to-loud sequence. Only a slight difference could be heard when the *mf* point was well chosen. The formant filters are thus

$$\mathcal{H} = a_1 \mathcal{H}_{pp} + a_2 \mathcal{H}_{mf} + a_3 \mathcal{H}_{ff}, \tag{7.29}$$
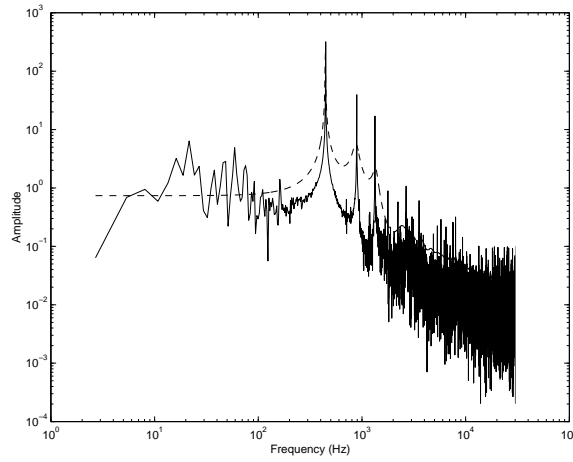
Figure 7.11: The spectrum envelope (dashed line) as determined with the LPC method for a *pp* played clarinet (full line). The sampling frequency is 44.1 kHz.

where a_1, a_2 and a_3 determine the interpolation point. The interpolation parameters are simply determined by the local amplitude A_{signal} and the amplitude of the *pp*, *mf* and *ff* signals (A_{pp}, A_{mf}, A_{ff}) as

$$a_1 = 1 - q \quad a_2 = q \quad a_3 = 0, \tag{7.30}$$

where $q = (A_{\text{signal}} - A_{pp})/(A_{mf} - A_{pp})$ in the case $A_{\text{signal}} \leq A_{mf}$, and

$$a_1 = 0 \quad a_2 = 1 - q \quad a_3 = q, \tag{7.31}$$

where $q = (A_{\text{signal}} - A_{mf})/(A_{ff} - A_{mf})$ when $A_{\text{signal}} > A_{mf}$. Note that the formant filters obtained by LPC are IIR type filters.

Cepstral techniques

When using a formant filter, the output of the model can be seen as

$$y(t) = w(t) * h(t), \tag{7.32}$$

with $w(t)$ the output of an oscillator and $h(t)$ an amplitude envelope. This is equivalent to

$$Y(\omega) = W(\omega)\mathcal{H}(\omega), \tag{7.33}$$

in the frequency domain. The output of the digital waveguide model has a harmonic spectrum, with lots of sharp partials. This is a high-frequency envelope. We are mainly interested in the low-frequency envelope. The power cepstrum of a signal is defined as

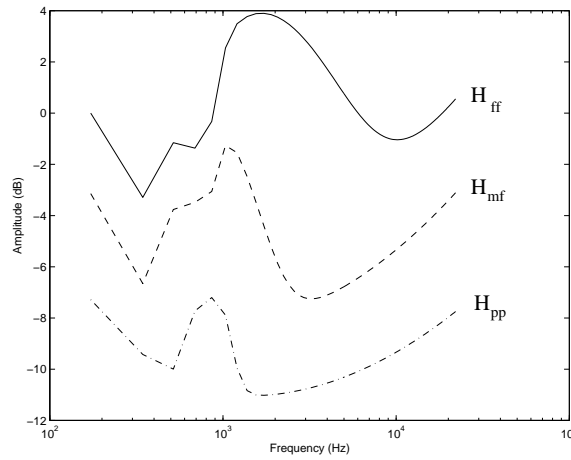$$\mathcal{C} = \mathcal{F}^{-1} \left\{ \log(|\mathcal{F}\{y(t)\}|) \right\}. \tag{7.34}$$

Figure 7.12: Three extracted Formant Filters for three amplitudes determined with the LPC method.

It is clear that by calculating the cepstrum of the global model output, we get the sum of the cepstra of the formant filter and the digital waveguide output, or

$$C = \mathcal{F}^{-1}\left\{\log(|W(\omega)|)\right\} + \mathcal{F}^{-1}\left\{\log(|H(\omega)|)\right\}. \qquad (7.35)$$

The low-quefrency[1] cepstral coefficients correspond to the envelope, while the output of the digital waveguide mostly contributes to the high-quefrency coefficients. By keeping only p low-quefrency bins, and applying the inverse tranformation, we find the magnitude of the envelope.

An improvement on this technique is the discrete cepstrum spectral envelope, developed in [33]. This method does not consider the complete spectrum, but uses only the peaks of the partials. The spectrum can be described as

$$Y = \sum_{k=1}^{N} a_k \delta(\omega_k). \qquad (7.36)$$

We consider this as the convolution of the waveguide output and the envelope. The waveguide output is

$$W = \sum_{k=1}^{N} w_k \delta(\omega_k). \qquad (7.37)$$

The description of the envelope is chosen such that the cepstral coeffients be-

---

[1]Cepstral equivalent of the frequency in the spectral domain.

come easy to calculate:

$$\mathcal{H} = \prod_{k=0}^{p} \exp(h_k cos(k\omega)). \tag{7.38}$$

The envelope coefficients can be found by finding the minimum of a cost function defined by

$$E = \sum_{k=1}^{N} (\log{(w_k \mathcal{H}(\omega_k))} - log(a_k))^2. \tag{7.39}$$

The minimum is found by setting the gradient to zero and using expression (7.38) to get a linear matrix equation

$$Ah = b, \tag{7.40}$$

where A has order $p + 1$ with elements given by

$$a_{ij} = \sum_{k=1}^{N} \cos(\omega_k i) \cos(\omega_k j), \tag{7.41}$$

h is the vector of filter parameters h_k and b is a column vector given by

$$b_i = \sum_{k=1}^{N} \cos(\omega_k i) \log(\frac{a_k}{w_k}). \tag{7.42}$$

It is now quite easy to construct a linear phase FIR filter with the magnitude spectrum of the formant filter, calculated with the low-frequency cepstral coefficients.

Wiener filter

It is also possible to directly calculate a FIR implementation of the formant filter, by calculating the Wiener filter that transforms the synthetic output to the desired, recorded output.

7.2.4   Calibration procedure

The calibration of the physical model is quite important, as inappropriate parameters might make the model totally unplayable. Following the design flow of figure 7.13, we will calibrate the formant filtered model. The first part of the calibration process is the estimation of the parameters of the reed model. It is known that the clarinet, like most single-reed instruments, operates in the stiffness-dominated region of the mass-spring system. Normal playing frequencies for a clarinet are under 1kHz so we expect the resonant frequency

Single / Double / Lip / Jet reed ?
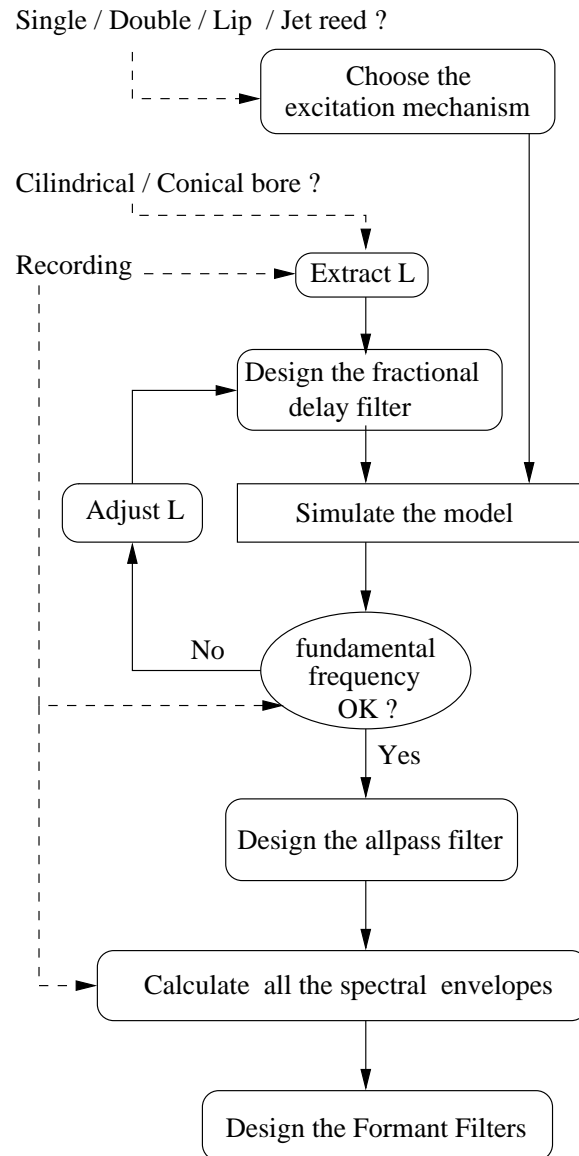
Cilindrical / Conical bore ?

Recording

Figure 7.13: The design flow for calibrating the formant filtered physical models. One starts by building a general model using an appropriate non-linear characteristic. The procedure then designs the linear part such that the simulation matches the recording.

to be in the range 2-3kHz. The damping coefficient ranges from 0.05 to 0.35 with little effect on the frequency domain behavior. With a lower damping coefficient for the reed, the model tends to "lock in" and oscillate at the reed's resonant frequency. Detailed measurements of different reeds have confirmed these parameter ranges [29]. These parameters do influence the time-domain (transient) response but have little effect on the steady-state tone. In the normal playing range, they affect the amplitude of the higher harmonics. The main contribution of these parameters is in the player interaction: they determine the oral cavity pressure and rest opening needed to generate a tone of a certain amplitude. This effect is captured in the parameterization of the formant filter.

Example 7.1 ⋆ `CD Tracks 6-17` are examples of the use of an LPC-dermined formant filtered clarinet model. We have a sequence of a low note, at low amplitude and high amplitudes, and a high note, again at low and high amplitudes. The first track of each series is the reference note as recorded. We then compare a second-order filter with a twentieth-order filter. For the high notes, with widely spaced partials, the twentieth-order filter yields a realistic sound. For the low note, a higher order is needed.   △

7.2.5 Performance

To determine the relative performance of the different techniques, the models were first implemented using MatLab. The calibration data are 16-bit 44.1kHz recordings of a clarinet, made in a dry recording room, using an AKG C4000B microphone. The musician played several chromatic scales with different loudness, as well as soft-to-loud sequences. These recordings were segmented per note by hand-selecting fragments with steady loudness. The segments were used by the calibration algorithm.

The derived models have a very good spectral match with the recordings. Informal listening tests indicated that it is difficult to discern real and synthetic sound for steady state tones. Dynamic changes also sound natural if the input is derived from real data. Listen to ⋆ `CD Tracks 3-18` for different examples of low and high notes played and simulated, with various formant filters types (refer to Appendix A for the details).

As an objective test, we calculated the weighted relative power spectral error (RPSE) defined as

$$\text{RPSE} = \frac{\sum_{\text{bins}} W |\mathcal{S}^2_{\text{reference}} - \mathcal{S}^2_{\text{synthetic}}|}{\sum_{\text{bins}} \mathcal{S}^2_{\text{reference}}}, \tag{7.43}$$

in which $\mathcal{S}$ is the amplitude of the spectrum and W a weighting factor. The weighting factor W was chosen flat up to $8kHz$ and declining for higher frequencies. One could use a perceptual auditory model for a more accurate
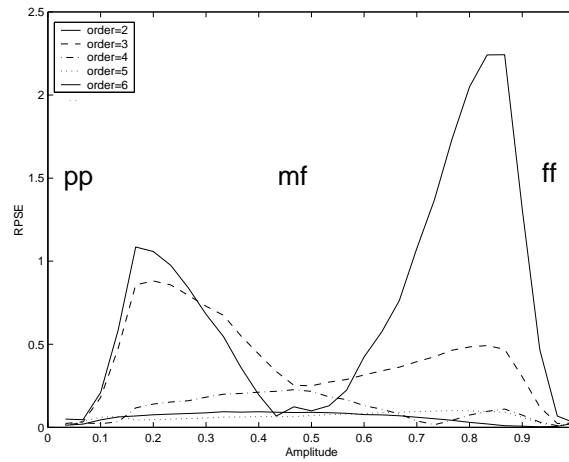
Figure 7.14: This figure shows the Relative Power Spectral Error for the formant filtered model. The formant filters are IIR filters calculated using LPC. The three calibration points are indicated. Higher order models have a significantly lower overall error.

representation of the subjective sound quality [81]. Figures 7.14 and 7.15 show the RPSE for increasing filter order, using the LPC-derived IIR formant filters and the Wiener FIR filters. As can be seen on the figures, it is clear that higher-order filters provide a lower RPSE. The amplitudes at which $\mathcal{H}_{pp}$, $\mathcal{H}_{mp}$ and $\mathcal{H}_{ff}$ were determined are indicated.

The transient behavior at onset of a note mostly depends on the chosen non-linear excitation mechanism. Small adjustments of the parameters of the non-linear part can alter the attack time and overblowing sensitivity. Figure 7.16 shows the spectrogram of a tone going from *pp* to *ff*.

7.3 Conclusion

In this chapter, we have looked at the practical implementation of a single-reed digital waveguide model. Several implementation techniques for the non-linear excitation mechanism were compared. By combining source-filter synthesis and the digital waveguide modeling techniques, we obtain a formant filtered model that is both accurate and easier to calibrate. We discussed several methods to determine the formant filters, and compared the performance of LPC-based formant filters and high-order Wiener filters. The per-note formant filtered models proposed in this chapter will be combined into multi-note models.
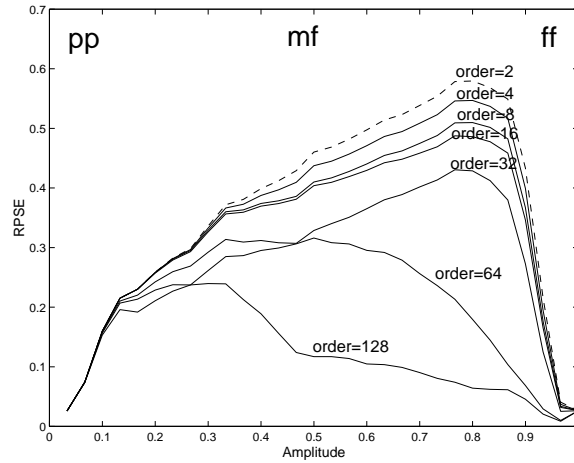
Figure 7.15: This figure shows the Relative Power Spectral Error for the formant filtered model. The formant filters are FIR Wiener filters with _order_ taps. Note that the scale of the Y axis differs from the scale of figure 7.14.
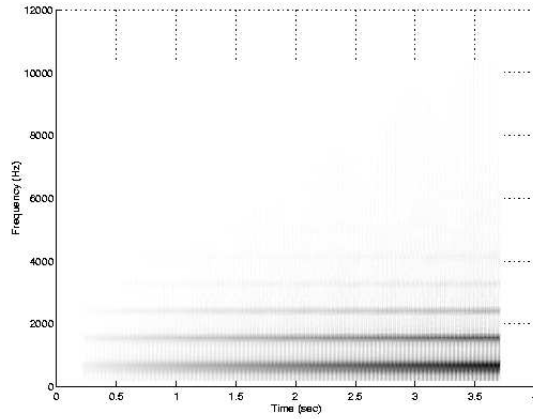

Figure 7.16: Spectrogram of the formant filtered physical model of a clarinet, from _pp_ to ff. A simple second order filter was used. At low playing amplitude, only the first three partials have significant amplitude, while at high playing amplitude, six partials are important.

Chapter 8

Multinote models

In this chapter, the single-note digital waveguide models are expanded to multinote models. Section 8.1 discusses different tonehole models and the possible digital waveguide implementations. Two specific digital waveguide models are then compared in section 8.2.

8.1   Multinote models

When playing several notes in a sequence on a real wind instrument, one effectively changes the acoustic configuration of the instrument. The start and the end note can be modeled by separate systems. During the transition of one note to another, the real instrument goes through a whole range of intermediate states. The output of the physical model of the instrument should match the output of the real instrument during these states as closely as possible. Changing the configuration of the toneholes results in a change in the digital waveguide resonator. The next sections describe several solutions to the note transition problems when using waveguide models.

8.1.1   Tonehole model

The most complete solution is to make a full acoustic model of the instrument including all toneholes and to model the effect of opening and closing a tonehole (see figure 8.5,(a)) [97]. This requires an accurate tonehole model and specific
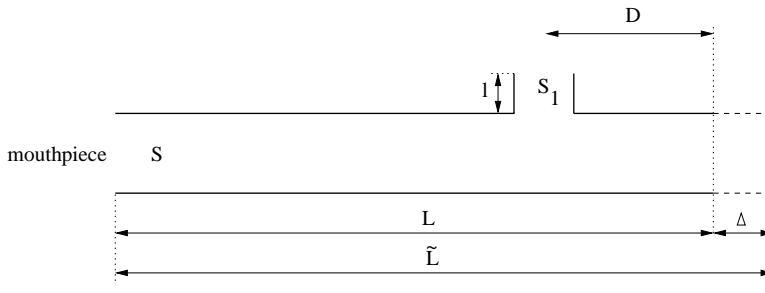
Figure 8.1: Representation of a cylindrical bore with one tonehole. The bore length is L is extended to $\tilde{L}$ to account for the radiation impedance.

control information when simulating the model (like the intermediate key or finger position above the tonehole) [96]. The resulting model is quite complex and difficult to calibrate but yields the best result. This model is used as the ultimate benchmark for the other solutions.

First, we look at the influence of one tonehole in a cylindrical bore, show in figure 8.1. When the hole is closed, the input impedance at the mouthpiece is approximated by

$$Z = jZ_0 \tan k\tilde{L}, \tag{8.1}$$

with $\tilde{L} = L + \Delta$ the length of the bore including the length correction for radiation, and $Z_0 = \rho c / S$ the characteristic impedance of the bore. The hole has a cross section S_1 and acoustic length l. At the position of hole, the admittance presented towards the mouthpiece is

$$
\begin{aligned}
Y &= Y_{\text{hole}} + Y_{\text{rest of bore}}, \\
&= -j\frac{S_1}{\rho c} \cot kl - j \cot k(D + \Delta),
\end{aligned}
\tag{8.2}
$$

and if both D and l are small compared to the wavelength,

$$Y \approx -\frac{j}{\rho\omega}\left(\frac{S_1}{l} + \frac{S}{D + \Delta}\right). \tag{8.3}$$

This impedance approximately corresponds to a reduction of the length of the original tube such that

$$\tilde{D} = D - \frac{S_1(D + \Delta)^2}{S_1(D + \Delta) + Sl}. \tag{8.4}$$

From these results, it can be concluded that a cylindrical bore with an open tonehole is not equivalent to a cylindrical bore truncated at the tonehole.

Benade presented an approximation to a bore with toneholes by assuming that the tonehole lattice of a woodwind instrument consists of a sequence of equidistant, equal-size holes as shown in figure 8.2 [9]. For the infinite case, the acous-
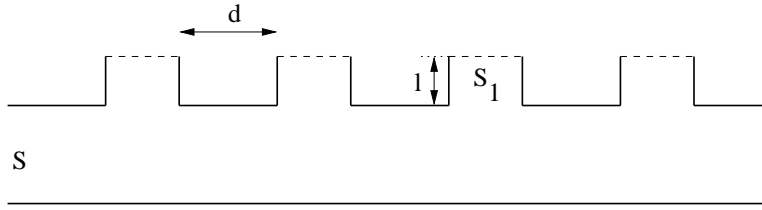
Figure 8.2: The tonehole lattice. There is an infinite sequence of identical toneholes, at regular intervals.

tic impedance is given by

$$Z_c = Z_0 \sqrt{\frac{1 + j(Z_0/2Z_b)\tan ks}{1 - j(Z_0/2Z_b)\cot(ks)}}, \tag{8.5}$$

where Z_b is the input impedance of the branches, and Z_0 the wave impedance of the main air column. These equations are derived by assuming continuity of volume flow and conservation of mass at the junctions of the main bore and branches. The branch impedances are approximated by the impedances of open-open or open-closed cylindrical sections. For a lattice of closed toneholes, it was shown that in the low-frequency limit, the impedance reduces to

$$Z_c^{\text{closed}} = Z_0\sqrt{1 + Q^{\text{closed}}}, \tag{8.6}$$

$$Z_c^{\text{open}} = jZ_o kd \sqrt{\frac{1 + Q^{\text{open}}}{Q^{\text{open}}}}, \tag{8.7}$$

where $2d$ is the spacing between the branches, and

$$Q^{\text{closed}} = \frac{1}{2}\frac{S_1}{S}\frac{l}{d}, \tag{8.8}$$

$$Q^{\text{open}} = \frac{1}{2}\frac{S_1}{S}\frac{d}{l}. \tag{8.9}$$

The phase velocity for closed toneholes is

$$v_p^{\text{closed}} = \frac{c}{\sqrt{1 + Q^{\text{closed}}}}. \tag{8.10}$$

The wavenumber for a lattice with open toneholes

$$2\Gamma d = 2\sqrt{Q^{\text{open}}\left(1 - \frac{S}{S_1}2k^2 dl(1 + \frac{2}{3}Q^{\text{open}})\right)}. \tag{8.11}$$

It follows from this equation that there is no wave-like propagation in the open-hole section of the bore, but an exponentially damped pressure variation. This
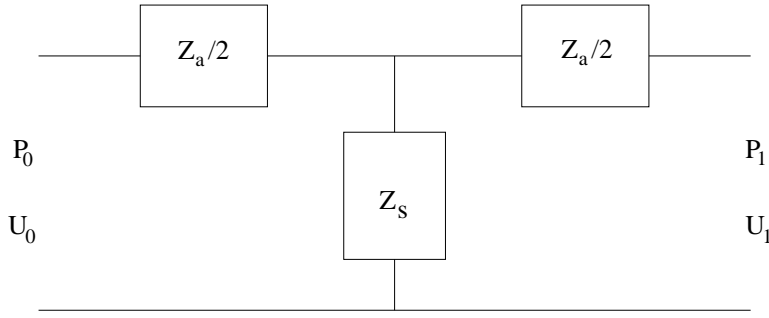
Figure 8.3: Symmetric transmission line scattering junction representing one tonehole.

basically means that only the first few open toneholes are important, allowing us to discard further toneholes.

A far more accurate representation is found if all the toneholes are modeled separately. Keefe [47] made a detailed acoustical analysis of a single tonehole. He obtained the symmetric scattering junction representation shown in figure 8.3.

The transmission matrix of the tonehole is determined by

$$
\begin{bmatrix} P_0 \\ U_0 \end{bmatrix} = \begin{bmatrix} 1 + \frac{Z_a}{2Z_s} & Z_a\left(1 + \frac{Z_a}{4Z_s}\right) \\ Z_s^{-1} & 1 + \frac{Z_a}{2Z_s} \end{bmatrix} \begin{bmatrix} P_1 \\ U_1 \end{bmatrix}, \tag{8.12}
$$

where

$$
\begin{align}
Z_s^{(open)} &= Z_a(a/b)^2(jkt_e + \xi_e), \tag{8.13} \\
Z_s^{(closed)} &= -jZ_0(a/b)^2\cot(kt), \tag{8.14} \\
Z_a^{(open)} &= -jZ_0(a/b)^2 kt_a^{open}, \tag{8.15} \\
Z_a^{(closed)} &= -jZ_0(a/b)^2 kt_a^{closed}. \tag{8.16}
\end{align}
$$

The height t of the tonehole is not the physical height, but is chosen such that the cylinder with the cross-sectional area of the tonehole encloses the same geometrical volume as the tonehole. For a typical tonehole structure shown in figure 8.4, this (closed tonehole) height is approximated by

$$
t = t_w + \frac{1}{8}b(b/a)[1 + 0.172(b/a)^2], \tag{8.17}
$$

where a is the radius of the main bore, b the radius of the tonehole, and t_w the shortest distance from the top of the tonehole to the main bore. The open-tonehole effective length depends on the presence of a pad above it. If no pad
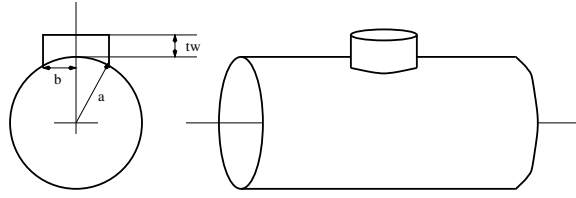
Figure 8.4: The structure of a tonehole. A tonehole is a small cylindrical section through the wall of the bore.

is present, t_e is given by

$$t_{e1} = \frac{k^{-1}\tan(kt) + b[1.40 - 0.58(b/a)^2]}{1 - 0.61kb\tan(kt)}, \qquad (8.18)$$

and if a pad of diameter $2R$ is placed on an average height of h above the tonehole, then

$$t_{e2} = \frac{k^{-1}\tan(kt) + b\left\{0.61(R/b)^{0.18}(b/h)^{0.39} + (\pi/4)[1 - 0.74(b/a)^2]\right\}}{1 - 0.61(R/b)^{0.18}(b/h)^{0.39}kb\tan(kt)}.$$
$$(8.19)$$

The specific resistance of an open tonehole is given by

$$\xi_e = 0.25 \cdot (kb)^2 + \alpha t + 0.25 \cdot k\delta_v \ln(2b/r_e), \qquad (8.20)$$

where r_e is the effective radius of curvature of the external end of the tonehole wall, $\delta_v = \sqrt{2\eta/(\rho\omega)}$ is the viscous boundary layer thickness and α the attenuation coefficient (see 4.3.1). The open and closed tonehole series equivalent lengths are

$$t_a^{open} = \frac{0.47b(b/a)^4}{\tanh(1.84t/b) + 0.62(b/a)^2 + 0.64(b/a)}, \qquad (8.21)$$

$$t_a^{closed} = \frac{0.47b(b/a)^4}{\coth(1.84t/b) + 0.62(b/a)^2 + 0.64(b/a)}. \qquad (8.22)$$

To use the full acoustic description of the tonehole in a digital waveguide system, the transmission matrix has to be transformed to the digital waveguide variables to obtain a two-port scattering junction with

$$\begin{bmatrix} P_1^- \\ P_2^+ \end{bmatrix} = \frac{1}{Z_a Z_s + 2Z_0 Z_s + Z_0^2} \begin{bmatrix} Z_a Z_s - Z_0^2 & 2Z_0(Z_s - Z_a) \\ 2Z_0 Z_s & Z_a Z_s - Z_0^2 \end{bmatrix} \begin{bmatrix} P_1^+ P_2^- \end{bmatrix}. \qquad (8.23)$$
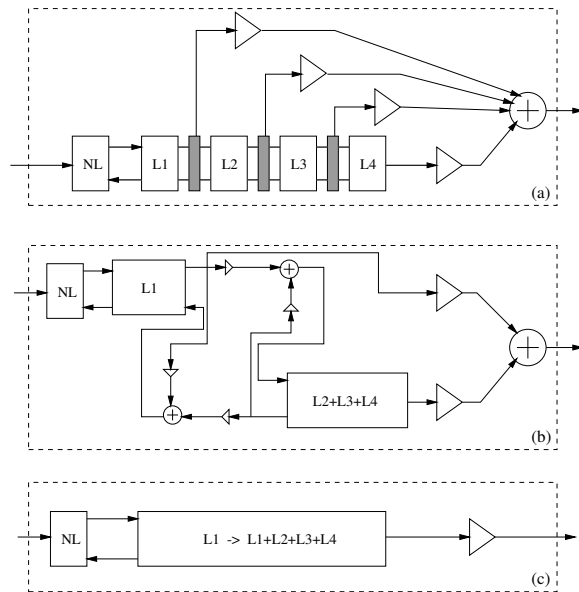
Figure 8.5: (a) The tonehole model. The different resonators (lengths $L1$ to $L4$) are linked through tonehole scattering junctions (gray boxes). (b) Serial crossfading: the two resonators are linked through a simple scattering junction. (c) Model morphing: the model parameters gradually change. (NL=non-linear function)

8.1.2   Serial Crossfading

A simpler solution is obtained by approximating the full acoustic model with a one-tonehole model (figure 8.5,(b)), effectively "plugging in"a tonehole when needed. Further, we could model the tonehole itself as a two-port, with simple scalars as reflection and transmission coefficients. This technique keeps a large part of the quality of the full acoustic model, while being a lot simpler. Ideally, the two first open toneholes should be modeled.

8.1.3   Model morphing

The previous system can be approximated once more by gradually changing or morphing the model for the first note to that of the second note (figure 8.5,(c)) during a note transition [50]. When using a waveguide model, this means a gradual change of the length of the delay line and changes of most or all of the filter coefficients. This, however, can introduce audible artifacts. This solution is useable for small changes in delay length and filters but becomes increasingly difficult and less natural-sounding for larger changes.

8.1.4   Model switching

At the other end of the spectrum, one could simply use a note-specific model and switch to another model when needed (figure 8.6,(a)). The outputs of the two separate models existing during the note transition can be crossfaded to obtain a legato effect. This very simple solution needs careful tuning of the mixing function used to avoid an audible "gap" between the two notes. It is very easy to implement.

8.1.5   Linked Model Switching

Instead of using two completely separate models, one could use two models with the output of the first model coupled to the input of the second model and vice versa. During the transition, both models are linked. This creates more intermediate states and makes a smoother sound. As can be seen on figure 8.6(b), the resulting system is still quite complex and requires two non-linear functions.

8.1.6   Joined Parallel Crossfading

A further possibility is by using two separate models for the resonator part and only one for the non-linear function (figure 8.6,(c)). It is a quite simple solution that gives output similar to the serial crossfading solution, while being a lot easier to implement. This solution is related to the banded waveguide
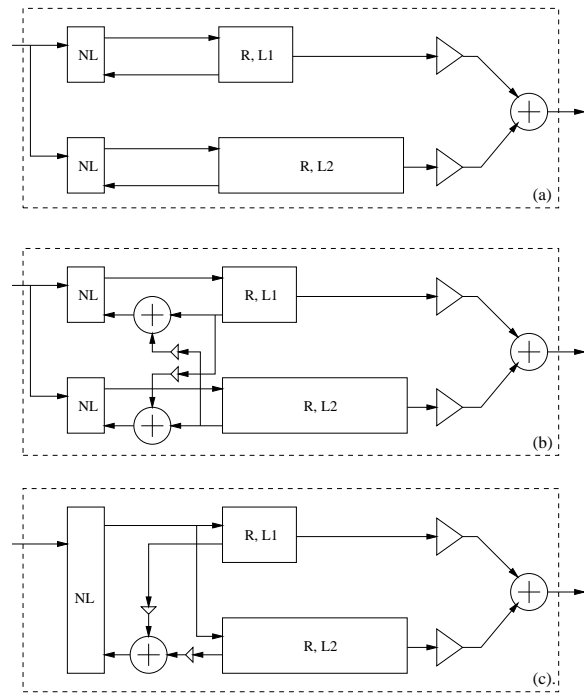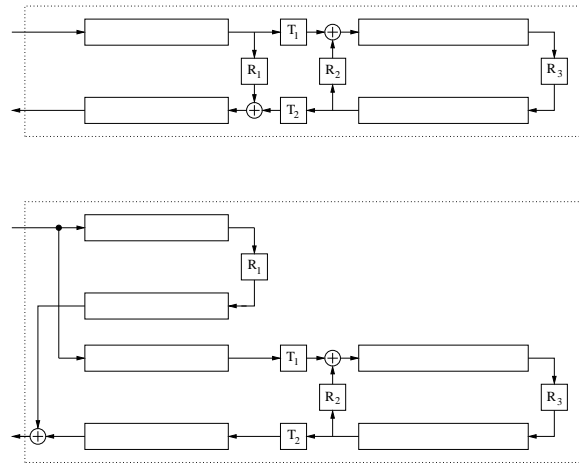
Figure 8.7: Explosion of one serial crossfading model into two. Exploiting the linearity of the global model, the scattering junction can be used to separate the top system. The top and bottom systems have identical input impedances.

models [23], without bandpass filtering. Before and after the transition, we have a one-note model while during the transition, we need an extra resonator. This model could be seen as a simplification of the equivalent systems shown in figure 8.7

8.2   Model comparison

The models are now compared for two possible applications: as true acoustic soundhole models, and in the context of formant filtered models.

8.2.1   Acoustic models

The models were detailed for the case with only one soundhole. As a real instrument has multiple soundholes, it is necessary to include all the soundholes to capture the instrument's properties. The obvious candidates are the full tonehole model, and the serial and parallel crossfading methods. We start by comparing the one-tonehole cases, and provide a qualitative comparison of multiple-tonehole models.

Figure 8.8 shows a general comparison of the models in terms of acoustic correctness and serial or parallel linking of the resonators. What we are looking for is a model that is fairly simple to implement and that is capable of a real-
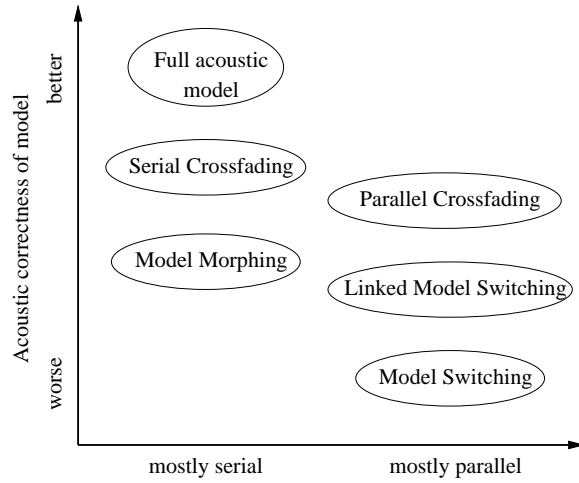
Figure 8.8: Comparison of the different methods in terms of the acoustic correctness and the type of resonators used. Parallel resonators are easier to implement.

istic note transition. Of the models presented in section 8.1, the serial and the parallel crossfading methods have similar complexity. To determine which one of these two is best regarding computational cost vs. performance, we need an objective comparison.

For a start, assume that we want to play two notes sequentially, and the first note has a higher pitch than the second one. This leads to a waveguide structure with the first note necessitating a delay length $L1$ and the second note a delay length $L1 + L2$. The two models can be found in figure 8.9. We choose the output of the non-linear function ($X1$) as the output of the system. The resonator part of the models can be seen as a black box with transfer function M, connected to the non-linear function.

As a first test, we calculate $\mathcal{M}$ for the two cases. The input of the black box is $X1_{\mathrm{ser, par}}$, the output is $Y1_{\mathrm{ser, par}}$. If the transfer function is similar in the two cases, this test gives a strong indication that the output of the two cases will also be similar. The transfer function is determined by

$$\mathcal{M}_{\mathrm{ser, par}} = \frac{Y1_{\mathrm{ser, par}}}{X1_{\mathrm{ser,par}}}. \tag{8.24}$$

For the serial case, we find the following input-output relation:

$$Y1_{\mathrm{ser}} = \left[a_{11}\mathcal{H}_1^2 + \frac{a_{12}a_{21}\mathcal{H}_1^2\mathcal{H}_2^2 R}{1 - a_{22}R\mathcal{H}_2^2} \right] X1_{\mathrm{ser}}, \tag{8.25}$$
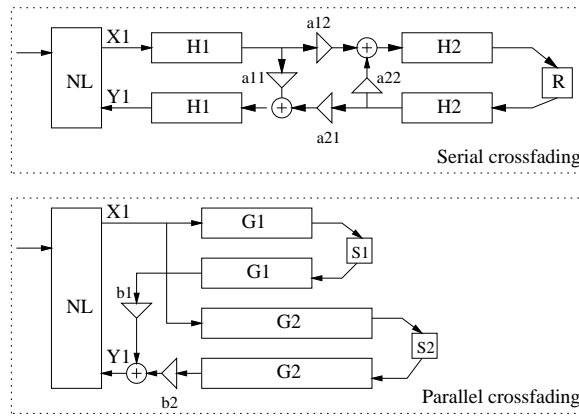
Figure 8.9: The two compared techniques. (top) serial crossfading and (bottom) parallel crossfading. $\mathcal{H}1$, $\mathcal{H}2$, $\mathcal{G}1$ and $\mathcal{G}2$ are the transfer functions of single delay lines with fractional delay, dispersion damping filters. R, $S1$ and $S2$ are reflection functions.

and for the parallel case:

$$Y1_{\text{par}} = \left[b_1 S_1 \mathcal{G}_1^2 + b_2 S_2 \mathcal{G}_2^2 \right] EX1_{\text{par}}. \tag{8.26}$$

Note that the denominator of equation 8.25 only has influence during the transition if we assume that a closed tonehole effectively "disappears" from the system. As detailed in the analysis of the true acoustic models in section 8.1.1, this is not exactly the case, but the approximation is satisfactory. To make the two cases comparable, we now determine the values for the parameters b, S and $\mathcal{G}$ of equation 8.25 such that $\mathcal{M}_{\text{par}}$ matches $\mathcal{M}_{\text{ser}}$. This translates to:

- the two systems must have the same output for the steady state case

- the outputs should be as close as possible during the note transition.

Steady state behavior

When the model is in steady state, we can make a few assumptions about the parameters for both sequentially played notes. During the first note, we assume that the system behaves as if the part of the instrument behind the open tonehole is not there (full reflection at the junction). During the second note, it is as if there is no tonehole at all (full transmission at the junction). Compared to the full acoustic model, this is a quite rough approximation of reality. We can now translate these conditions to the two cases. For the serial

case these assumptions become (cfr. figure 8.9, top):

$$\text{Open} \quad \begin{cases} a_{11} = -1, \\ a_{12} = a_{21} = 0, \end{cases} \tag{8.27}$$

$$\text{Closed} \quad \begin{cases} a_{11} = a_{22} = 0, \\ a_{12} = a_{21} = 1, \end{cases} \tag{8.28}$$

while for the parallel case (figure 8.9,bottom):

$$\text{Open} \quad \begin{cases} b_1 = 1, \\ b_2 = 0, \end{cases} \tag{8.29}$$

$$\text{Closed} \quad \begin{cases} b_1 = 0, \\ b_2 = 1. \end{cases} \tag{8.30}$$

This yields the following relations between the two models:

$$S_1 = a_{11}, \tag{8.31}$$
$$S_2 = R, \tag{8.32}$$
$$\mathcal{G}_1 = \mathcal{H}_1, \tag{8.33}$$
$$\mathcal{G}_2^2 = \mathcal{H}_1^2 \mathcal{H}_2^2. \tag{8.34}$$

Transition behavior

Using these relations in equations 8.25 and 8.26 yields as a possible solution:

$$b_1 = a_{11}, \tag{8.35}$$
$$b_2 = \frac{a_{12} a_{21}}{1 - a_{22} R \mathcal{H}_2^2}. \tag{8.36}$$

We now approximate this by putting the denominator equal to 1. This corresponds to removing the resonance in the second part of the bore (behind the tonehole). This will have a minor effect as the resonance is damped. The parameter b_2 now becomes frequency-independent:

$$b_2 = a_{12} a_{21}. \tag{8.37}$$

The spectrogram of the time domain simulation for three multinote models are shown in figure 8.10.

Example 8.1 ⋆ CD Track 18-19 are examples of serial and parallel crossfading. There is very little audible difference between the two methods.
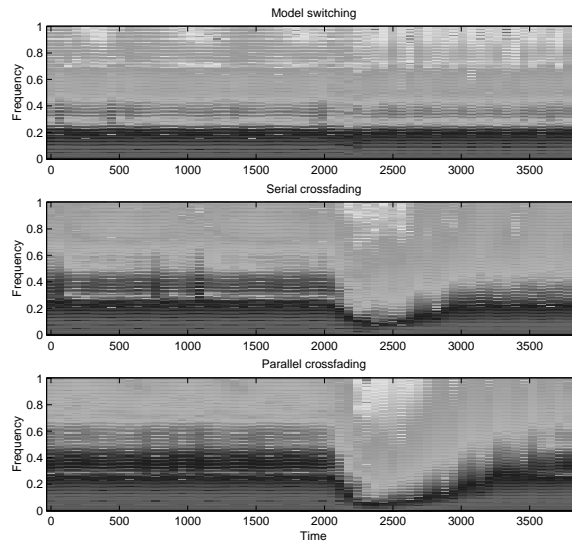△

Figure 8.10: This figure shows the transition between two notes, for three different multinote models. The tonehole is closed at time=2100. (top) Model switching, (middle) Serial crossfading, (bottom) Parallel crossfading. Both serial and parallel crossfading show the characteristic drop in amplitude during the transition.

8.2.2   Formant filtered models

In the context of formant filtered models, we have a calibrated model for each note. Every single-note model implicitely incorporates the effects of the tone-hole configuration for that particular note. This means that the only thing we actually have to model is the transition, or the opening and closing of one tonehole. In this case, the previous one-tonehole analysis is perfectly adequate.

Again, this is not as accurate as the full acoustic model, but we have to take into consideration that during normal playing the transition time between two notes is rather short, making it difficult for us to actually hear the difference. We conclude from this theoretical analysis that serial and parallel crossfading produce very similar output. Simulation of the two cases confirms this analysis. Subjectively, there is very little audible difference between the two crossfading methods.

From the implementation standpoint, the parallel crossfading technique is the superior choice. A drawback of the serial crossfading method is that one must be able to "tap"into the delay line or to attach a new delay line when needed. This is not the case with parallel crossfading. Another factor common to all scattering junction-based methods is that potentially unnecessary calculations are done in the part of the model behind the tonehole. With the parallel crossfading method, after the transition is complete, the unused resonator can be cleared and replaced as needed, making it more flexible. One has to choose between one large model running all the time or two smaller models existing in parallel during the transition.

8.3   Conclusion

In this chapter, the acoustic properties of a tonehole and a tonehole lattice were discussed. Six possible digital waveguide implementations of toneholes were qualitatively compared, and a more detailed comparison of the serial and parallel crossfading techniques was given.

Chapter 9

Implementation

```
           for i=1:100
           {
    ◀━━━━━━▮   ...
           }
```

In this chapter, we discuss the actual implementation of the physical models using the MPEG-4 Structured Audio Language. First, we provide an overview of the properties of MPEG-4 and the Structured Audio Orchestra Language in section 9.1. In section 9.2, we discuss the implementation of the different elements of digital waveguide models and the issues specific to the wind instruments. Finally, in section 9.4, this implementation is compared to other synthesis techniques.

9.1   MPEG-4

MPEG-4 is an initiative of the MPEG standardization committee, and is a standard for the coding of multimedia objects. MPEG-4 Audio, the part of MPEG-4 dedicated to audio objects, integrates the worlds of speech, high quality audio coding, sound synthesis and natural audio representation. The sound synthesis part is comprised of tools for the realization of symbolically defined music and speech [39].

Synthetic audio is described by first defining a set of "instrument" modules that can create and process audio signals under the control of a script or score file. An instrument is a small network of signal processing primitives that can emulate the effects of a natural acoustic instruments. MPEG-4 does not standardize a particular set of synthesis methods, but a method for describing synthesis methods, although a more careful reading of the MPEG-4

standard shows that MPEG-4 Structured Audio (MPEG-4 SA) is best suited for wavetable synthesis [40].

9.1.1 Major elements of MPEG-4 SA

MPEG-4 Structured Audio is comprised of several parts:

Structured Audio Orchestra Language SAOL is a digital signal processing language which allows for the description of arbitrary synthesis and control algorithms as part of the content bitstream.

Structured Audio Score Language SASL is a simple score and control language which is used to describe the manner in which sound-generation algorithms described in SAOL are used to produce sound.

Structured Audio Sample Bank Format The format used for the transmission of banks of audio samples to be used in wavetable synthesis and the description of simple processing algorithms to use with them.

Scheduler A supervisory run-time element of the Structured Audio decoding process . It maps structural sound control, specified in SASL or MIDI, to real-time events dispatched using the normative sound-generation algorithms.

MIDI reference The MIDI standard for communication between digital musical instruments [4]. Although more restricted, MIDI can be used as an alternative to SASL.

For the practical implementation of the models presented in the previous chapters, only SAOL, SASL and MIDI are used directly. The scheduler in inherently used during synthesis, but as it is integrated into the compiler/renderer, it is not controllable by the sound designer.

9.1.2 SAOL

SAOL is a declarative unit-generator-based structured audio language. SAOL was designed to be highly readable, modular, expressive and functional. It allows the declaration of *instruments* as a set of digital signal processing algorithms that produce sound, and an *orchestra* as a set of instruments. One significant feature of SAOL is the distinction between sample-rate and control-rate variables. Sample-rate variables are variables updated every sampling period (typically at 44.1kHz), while control-rate variables are only updated at the usually much slower control rate (typically 100Hz). This allows some sort of block processing, although SAOL is designed as a sample-processing language

for the sample-rate variables. Unfortunately, no rate-free variables exist, which makes it quite hard to implement new non-core opcodes for complex signal processing tasks. SAOL is also heavily dependent on stored-function tables or wavetables [98]. This allows the efficient processing of periodic signals but is fairly useless in the case of physical modeling synthesis.

SAOL allows the user to define new opcodes. Using the same core functionality as instruments, user-defined opcodes perform a certain calculation and then return their results. This way, one could build a library of special opcodes that can easily be re-used in several instruments. SAOL also uses bus-routing to connect the output of one or more instruments to other instruments or effects, but the busses are only defined for sample-rate dataflows. The routing of control-rate data has to be done through global variables, which was a restriction encountered during the implementation of the physical models.

9.1.3 SASL and MIDI

SAOL instruments may be controlled using either the MPEG-4 SA specific Structured Audio Score Language SASL or by using the MIDI standard. For the cases where MIDI is used, a set of standard names allows the programmer access to the standard MIDI control, pitch-bend, and aftertouch. Channels and preset mappings are also supported.

9.1.4 MPEG-4 Software

At the time of writing, only a restricted set of software solutions exist for the compilation of MPEG-4 Structured Audio.

`saolc` A reference implementation of the Structured Audio decoder and renderer was added to the ISO standard. `saolc` is a SAOL-to-C translator. The resulting C program can be compiled and run to generate audio.

`sfront` `sfront` is an efficient SAOL-to-C compiler [53]. We used several versions of `sfront` for the final SAOL implementation.

We are however confident that more solutions will become available as the acceptance of MPEG-4 as a multimedia encoding standard increases.

Both MPEG-4 solutions translate the native SAOL code to intermediate C, that must be compiled. `Saolc` writes a direct translation of the SAOL code, but unfortunately does not support all the possibilities of the SAOL language. After compilation, the performance is mediocre at best. `sfront` applies some optimizations to the code, but as it does not apply any advanced code transformations, the end result is a very long C program with a somewhat better

performance after compilation. It seems that applying some higher-level code transformations might result in much higher performance after compilation. This has however not been studied in this work.

9.2   Digital waveguide elements

The three major elements of the digital waveguide system are the delay lines, the linear filters, and the non-linear functions. In the case of a wind instrument, we have one large delay line, one non-linear function for the excitation mechanism, a small reflection filter, and a more complex, high-order formant filter.

9.2.1   Delay lines

There are two possible implementations for the delay lines used in the digital waveguide models. The first one is to use a FIFO buffer, with all elements moved one space every sample period. The second possibility is to use moving pointers in a circular buffer. The FIFO implementation is well-suited for specific hardware solutions that include a hardware FIFO, but require too many memory operations in all practical other cases. The moving pointers present a very efficient solution and are the choice for this implentation.

9.2.2   Filters

Implementations of filters require the solution of several issues. The two major types of filters, IIR and FIR, have specific problems. The SAOL language has two opcodes specific for the calculation of FIR and IIR filters. However, one is still free to define his own filter opcode, but this results in extremely inefficient code, as the user opcode runs within the limitations of the SAOL language, while the predefined opcodes can be hardware-optimized.

Time domain calculation

Both FIR and IIR filters can be calculated in the time domain. FIR filters are based on time-domain convolution and are inherently slow ($\mathcal{O}\{N^2\}$). Some DSPs however have an instruction set specifically adapted for this type of calculation (MAC instruction, hardware modulo-addressing, parallel data and coefficient memory access). Even then, it remains a fact that filters of length 1024 and more cannot be calculated in real-time. FIR filters present no special numerical problems. IIR filters are best calculated in the time domain, as they have low order and include feedback loop. IIR filters are sensitive to numerical problems resulting from the fixed-point calculation, the finite precision, and

the particular filter structure and coefficient encoding method used. Potential problems are unstable filters and limit cycles.

Frequency domain calculation

The effect of a linear filter can also be calculated in the frequency domain by using block-processing. This is an inherently fast method, but introduces a calculation delay equal to the block size. Several fast-convolution algorithms exist, and are perfectly suited for non-real-time calculations, or when a short delay is not an issue.

9.2.3 Non-linear functions

The implementation of the non-linear function depends on the complexity of the function and the available time for the calculation. The highest precision is achieved when the function is calculated explicitely. This requires multiple multiplications, divisions and the evaluation of a square root. Local linearizations are possible, but introduce an overhead as the algorithm first has to decide in which region the approximation is valid. The use of a look-up table is a third possibility, but as the non-linear functions for the single-reed excitation mechanism depends on two variables, this requires a lot of memory and relatively unstructured memory access. For this implementation, we prefer the direct calculation of the non-linear function. Care must be taken to use the complete solution of the equations, as hidden loops and thus an unstable model might occur otherwise.

9.3 Clarinet model implementation

Single-note clarinet models are quite straightforward to implement. Multinote models require the addition of a virtual instrument layer when MIDI input is used, as the MPEG-4 interpretation of MIDI results in a new, independent instance of the instrument, thus forbidding any interaction between the previous note played and the next note. The use of MIDI also restricts the user to fixed and hard-coded key opening and closing times, as no suitable controller exists in the MIDI protocol to transfer this data. The use of pure SASL with a dedicated controller for each key is a better choice. A very simple control parameter translation layer is needed to transform the encoded breath sensor and pitch bend data to the right model parameters ($p_o x$ and initial reed displacement) and formant filters.

	Description
Formant [74]	Physical model
	Formant filtered
	Joint parallel crossfading
PhyMo [97]	Physical model
	Full acoustic model
Add [64]	Spectral model
	Per note
Sampling	Processed recording
	Per note
Shaping [3]	Abstract algorithm
	Per note, per amplitude
FM [17]	Abstract algorithm
	Per note, per amplitude

Table 9.1: Description of the different techniques available for the synthesis of wind instrument tones.

9.4   Comparison with other synthesis techniques

In this section, the properties of the formant filtered physical models are compared to other currently existing synthesis techniques.

There exist many different sound synthesis techniques, and quite a few can be used for the synthesis of wind instruments. We selected the most successful and widespread types, and compare them to the formant filtered model. It is a qualitative comparison that looks at different aspects of the synthesis techniques. We compare the formant filtered model (Formant) with the full acoustic model developed in [97] (PhyMo), additive synthesis (Add), wavetable synthesis (Sampling), waveshaping (Shaping), and FM synthesis (FM). Table 9.1 summarizes the different techniques.

The most advanced model is the PhyMo technique, a full acoustic digital waveguide model with multiple toneholes. It is the most accurate acoustic model. The calibration of the model requires the careful measurement of the physical dimensions of all the different parts of instrument, and some assumptions on the parameters of the non-linear excitation mechanism. There is no complete calibration solution, except for extremely simple instruments, and the calibration requires specialized equipment. Using acoustic reflectometry, one is able to determine the internal shape and reflection characteristics of the real instrument. The large number of parameters makes the model calibration quite hard. Using the model requires complex controllers, as all the physical control variables need to be measured. The (un-calibrated) sound quality is very high, as is the model's response. Compared to Formant, the model has better pre-

dictive properties. As the basic properties of both `PhyMo` and `Formant` is the same, both techniques are very close in sound quality and control when used without calibration. The `Formant` model is a paramtrization of the physical model that has better calibration properties and a lower computational load.

All other techniques are fundamentally different, as they are based on the perception of the sound. The `Add` model tries to reconstruct the spectrum of the sound by adding the outputs of multiple oscillators [64]. This is a very powerful technique that can only be used after calibration. It very easily calibrated, by analyzing the frequency, amplitude, and time-frequency envelope of all the partials of the signal. Automated tools exist for the calibration and design of the additive model. The main drawback is that this techniques requires calibration, as is cannot deduct properties of other notes without meta-information about the sound production. For steady-state tones, the sound quality is equal or better than either `Formant` or `PhyMo`. Transients are very difficult to achieve, and control is restricted to previously analyzed ranges.

Wavetable synthesis, or sampling, is extremely simple. In its early implementations, tape recorders were used to record a sound and play it back at a different pitch. By looping a part of the recording, a longer tone can be obtained. Pitch shifting is achieved by changing the sample playback speed, by resampling the signal, or interpolated table reads. `Sampling` is not very well-suited for wind instrument tones. The best quality is achieved when every note is recorded at different amplitudes. The player has very little direct control over the timbre, and note transients are not modeled at all. Compared to `Formant`, wavetable synthesis requires much more storage space, offers little player control, and an in general inferior sound quality.

Waveshaping, also called, non-linear distortion, is a simple synthesis technique that uses a non-linear mapping from input to output [3]. Typically, a memory-less function is used. The choice of the non-linear function has a large impact on the calibration properties. When using sum of Chebychev polynomials, the amplitude of the different harmonics can be tuned by changing the gain of the corresponding polynomial. Waveshaping presents the disadvantage that noisy signals cannot be generated. By combining Chebychev-based waveshaping and a noise source, we built a fairly convincing model for the clarinet. The steady-state performance is very good, however, no transients can be generated. Accounting for inharmonic partials is more complex, and care must be taken not to introduce significant amounts of aliasing. Compared to `Formant`, this technique is cheap, easy to calibrate, but only usable for the notes and amplitudes it was calibrated for. It is possible use formant filters with a waveshaper calibrated for full excitation as the source. This yields a model that has about the same computational load as `Formant`, but with less control properties and no transients.

FM synthesis uses the modulation of the frequency of an oscillator by another

oscillator. In its simplest form, the spectrum of the sound generated is a sum of Bessel functions. Certain combinations of parameters produce formant-shaped spectra, resembling to output of wind instruments. The calibration of FM synthesis is not easy, but can be achieved by using general non-linear optimization techniques with a perceptual/spectral cost function [38]. In general, the sound quality is poor, but the computational load is very low.

9.5   Conclusion

The implementation of the wind instrument models present no important problems if care is taken for the choice of the filter structures, and if the non-linear function is programmed without hidden loops. The MIDI input data only needs a simple translation to the internal model parameters, as the sensor measurement of a typical wind controller have a good match with the required control parameters for the model. Compared to other synthesis techniques, the formant filtered physical models have better calibration properties and player control for a given sound quality level.

Part III

String Instruments

Chapter 10

Single-String Models
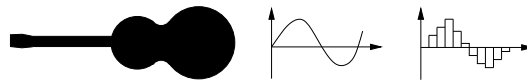
I learned a long time ago that one note can go a long way if it's
the right one, and it will probably whip the guy with twenty notes.

Les Paul

In this chapter, a single-string model for the acoustic guitar is presented. The
methods and properties are completely general and can easily be applied to
other string instruments with minor modifications. This single-string model
will be extended to a fully coupled, multiple-string model in the next chapter.

In section 10.1 we start by building a simple acoustic model for the acoustic
guitar, and discuss the excitation of the strings in section 10.2. A simple
single-polarization single-string model is then described in section 10.3 and
dual-polarization models in section 10.4. Finally, section 10.5 deals with several
parameter estimation algorithms.

10.1   Simple guitar models

In section 3.2, we described the structure and the acoustic operation of the
acoustic guitar. A basic structure reflecting the acoustic operation of the guitar
is shown in figure 10.2. This structure can be refined to include string coupling,
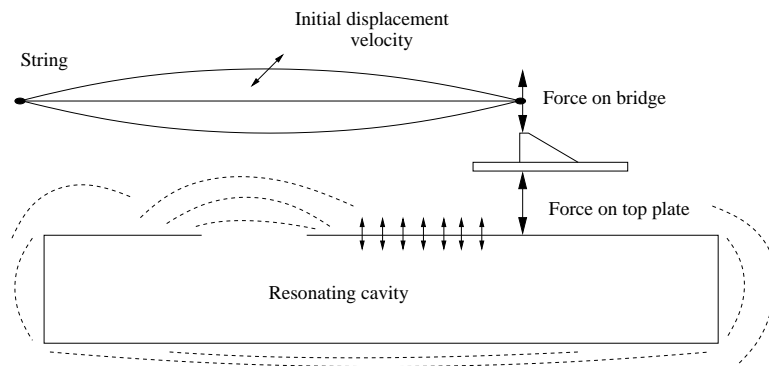
Figure 10.1: The working principle of an acoustic guitar. The player gives an initial displacement and velociy to the string. The string vibrates, and a part of the energy is transferred to the other strings and the top plate. The top plate is a part of the guitar body (the resonating cavity).
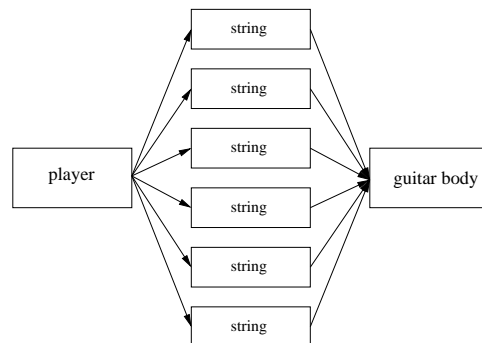

Figure 10.2: The basic system-level structure of a guitar model.

but it is adequate as a starting point. The player defines the initial excitation of the strings. The excitation consists of the initial displacement and velocity, and includes a plucking position. It is then fed to the string models. The string models are the core of the guitar model, as they define the vibration frequencies and the damping over time. The string vibration is observed at the bridge and used as the input of a guitar body model. The player also controls the string length, damping, and, to a lesser extent, the string tension.

10.1.1 Excitation

The excitation of the strings is the single most important parameter determining the tone of a given guitar. It takes a human player years of practice to achieve a consistent high-quality tone. In the case of the acoustic guitar, the strings are plucked with the finger, or using a sharp, rigid or flexible device (a plectrum, thumbpick, or fingerpick). In section 10.2, the different excitation functions will be discussed.

10.1.2 Strings

Wave propagation in the strings is possible in several spatial directions. The most important modes for the guitar are the transversal waves, parallel and perpendicular to the soundboard. The longitudinal modes become important when modeling the piano, and the torsional modes are imperative for a high-quality violin model. The simplest string models only describe the string movement in one dimension (section 10.3). By combining several single-polarization models, one obtains a complete string model (section 10.4). Finally, several complete string models can be combined using a bridge model to accurately model string coupling (chapter 11).

10.1.3 Body

The body of the guitar acts as an acoustic amplifier, coupling the vibrational modes of the string to sound waves in air. The body is in essence a Helmholtz resonator, with added top and back plate vibrations. The lowest resonance mode is the Helmholtz resonance, whose frequency is determined by the cavity volume and the area of the soundhole. The higher modes are determined by the principal modes of the top and back plate, and the higher air cavity modes, resembling the standing waves in a rectangular box. The lower vibrational modes of the top and back plates resemble those of the rectangular plate. The guitar body models are the subject of chapter 12.
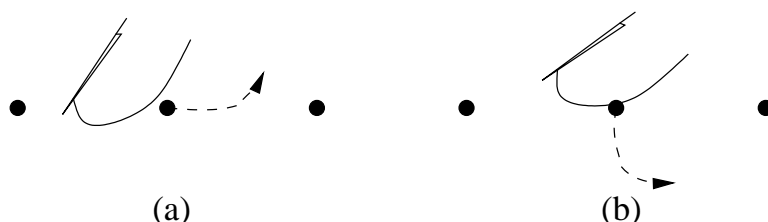
Figure 10.3: Initial string displacement for (a) the *tirando* stroke, and (b) the *apoyando* stroke

10.2   Excitation

In the classical and flamenco style of playing, two main types of plucking are commonly used. The strokers are the *apoyando* and *tirando*. In the *apoyando* or rest stroke, the plucked string is pulled down towards the soundboard, where it slips over the fingernail. The finger now rests on the adjacent string. This stroke is mostly used to emphasize the main melody. With *tirando* or free stroke, the string is first pulled away parallel to the soundboard and released slightly upwards. The end result is in essence the same string motion as with the *apoyando* stroke, but with a somewhat better control of the amount of vertical and horizontal movement. These plucking methods are widely used to accentuate certain notes of the musical phrase, or to achieve a different timbre. Figure 10.3 shows the difference between the strokes.

When the player plucks a string that is already vibrating, the finger-string interactions first dampens all or part of the motion of the string. The *apoyando* stroke results in the damping of the adjacent string. A more strumming excitation is also widely used in flamenco and contemporary guitar playing. The *rasgueado* stroke consists of the rapid up- or downstroke of one or more right-hand fingers. Several studies have been published on string-finger interaction [82], [31].

The main difference between plectrum and finger plucks is the width of the interaction region. A finite-width pluck will excite all harmonics, irrespective of the plucking position. This is shown in figure 10.4. The pluck position determines which harmonics will be excited. Integer ratios of the string length lead to the absence of the harmonics corresponding to the ratio, e.g. plucking at 1/8 of the string length will not excite the 8th, 16th, 24th, etc. harmonics. Figure 10.5 shows the harmonics excited by plucking at the three tradional positions: on the fingerboard (*sul tasto*), the normal position, and near the bridge (*sul ponticello*). Note that playing closer to the bridge generates more higher harmonics.
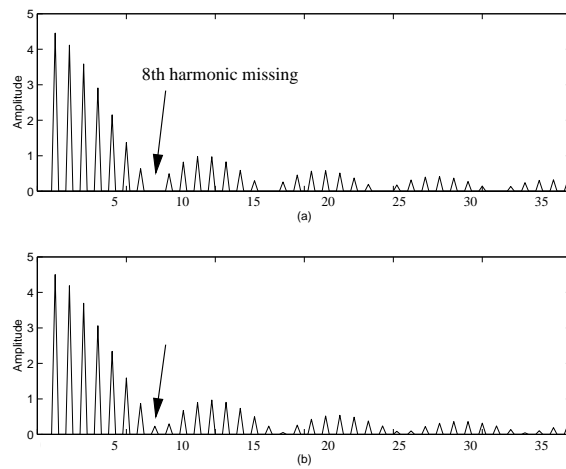
Figure 10.4: This figure illustrates the difference between (a) a zero-width pluck, and (b) a finite-width pluck. The string is plucked at 1/8 of the string length, causing the absence of the 8th, 16th, etc. harmonic.
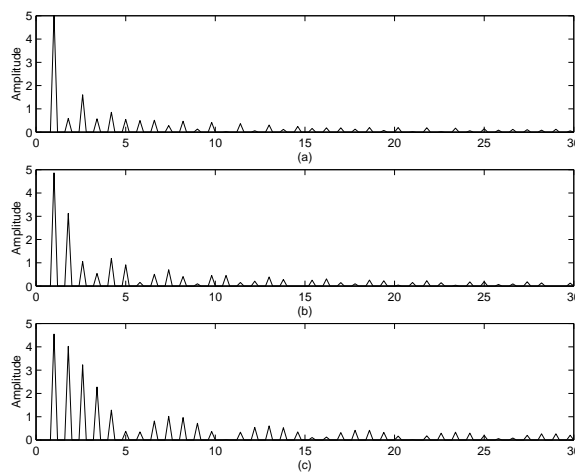

Figure 10.5: The figure shows the spectrum of the pluck at three different plucking positions: (a) *sul tasto* (near the fingerboard, ratio $\beta = 0.46$), (b) normal pluck point (ratio $\beta = 0.28$), and (c) *sul ponticello* (near the bridge, ratio $\beta = 0.15$). Plucking near the bridge excited more high vibration modes, resulting in a bright and somewhat harsh tone.
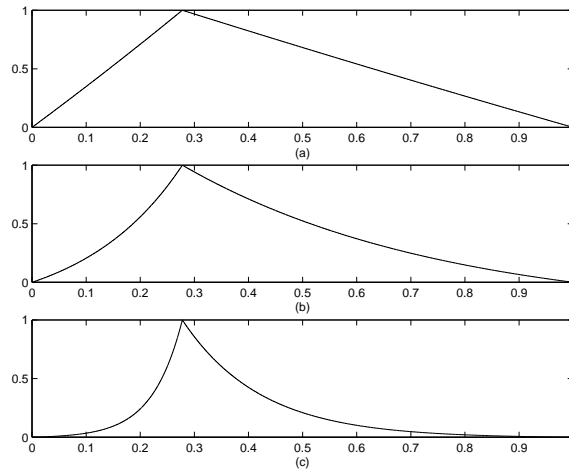
Figure 10.6: Three plucks with different sharpness: the slopes increase and decrease exponentially accordig to equation (10.2) with the sharpness parameters (a) $s = 0.01$, (b) $s = 1.5$, (c) $s = 5$.

Fourier analysis of the transverse force on the bridge gives the following expression for the amplitude of the nth harmonic:

$$A_n = \frac{2dT_0}{n\pi L_0} \frac{1}{\beta(1-\beta)} \sin \beta n \pi, \tag{10.1}$$

where T_0 and L_0 are the string tension and length at rest, and d the initial displacement [30].

For synthesis purposes, we also define pluck sharpness, which enables us to excite more higher harmonics as the excitation becomes more impulse-like. Figures 10.6 and 10.7 show three different plucks with their corresponding spectra. The plucks consist of exponential ramps given by

$$P = e^{sx}, \tag{10.2}$$

White or colored noise can also be added to the excitation, resulting in snare-drum like effects.

10.3   Single-polarization string models

The transversal motion of an ideal string is described by

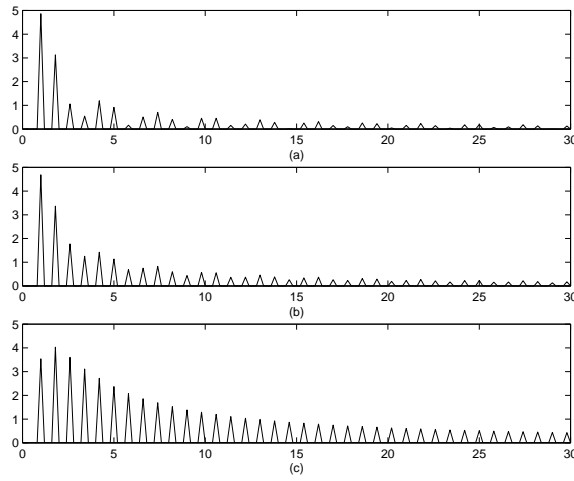$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \tag{10.3}$$

Figure 10.7: The spectra corresponding to the pluck shapes of figure 10.6. The sharper the pluck, the more vibration modes are excited. The limit, for infinite sharpness, is an impulse, exciting all modes evenly.

with c the transversal wave propagation speed:

$$c = \sqrt{\frac{T}{\rho}}, \tag{10.4}$$

with T the string tension and ρ the linear density of the string material. These equations assume infinitely thin strings, which is a good approximation for treble strings. Real strings have a finite thickness and stiffness. Including string stiffness, and frequency-dependent losses yields a more complex equation:

$$T\frac{\partial^2 y}{\partial t^2} = \rho\frac{\partial^2 y}{\partial x^2} - \kappa\frac{\partial^4 y}{\partial x^4} - 2b_1\frac{\partial y}{\partial t} + 2b_3\frac{\partial^3 y}{\partial t^3}, \tag{10.5}$$

where b_1 and b_3 are the loss parameters, and the parameter κ is given as

$$\kappa = \frac{E\pi r^4}{4}, \tag{10.6}$$

with E the Young modulus and r the radius of the string. For the lossless case and small string stiffness, the modal frequencies are

$$f_n = f_0\sqrt{1 + Bn^2}, \tag{10.7}$$

with

$$B = \frac{\pi^3 E r^4}{512 T L^2}. \tag{10.8}$$

	Nylon	Bronze	Stainless steel
E4	1.08e-5	0.89e-5	2.25e-5
B3	2.5e-5	2.7e-5	6.74e-5
G3	5.8e-5	13.4e-5	30.4e-5
D3	12e-5	35.8e-5	104e-5
A2	26e-5	119e-5	318e-5
E2	66e-5	314e-5	820e-5

Table 10.1: Comparison of the values of the inharmonicity coefficient B for three string types: classical nylon strings (D'Addario EJ45) $E = 2.5GPa$, bronze 80/20 folk strings (D'Addario EJ40) $E = 110GPa$ and stainless steel electric guitar strings (D'Addario XLS600 Jazz medium) $E = 200GPa$. String diameter and normal tension as published by the manufacturer.

The inharmonicity coefficient is quite different for the strings commonly used on guitars. Table 10.1 compares the values for B for three string types: classical nylon strings (D'Addario EJ45), bronze 80/20 folk strings (D'Addario EJ40) and stainless steel electric guitar strings (D'Addario XLS600 Jazz medium). The nylon strings are the least affected by string stiffness.

10.3.1 Double delay line model

The digital waveguide implementation of string motion is rather straightforward. Figure 10.8 shows a two-delay line model designed for acceleration waves. The delay lines and the reflection filters R_b and R_f form a loop in which the waveform circulates, as was discussed in chapter 5. The two reflection filters simulate the reflection of the waveform at the termination points (the bridge and the fret or nut). To comply to the boundary conditions for a string with fixed-fixed ends, the reflection filters must be phase-inversive. The delay lines are of non-integer length and consist of an integer-length delay line and a fractional delay filter. Losses and inharmonicity are cared for by an allpass and a lowpass filter. Note that the fractional delay filter itself introduces a lowpass effect, and the lowpass filter should be designed taking into account these existing losses. The input signal is injected into the delay lines at the plucking point. It has been shown in [103] that the ideal pluck of the string can be approximated by a unit impulse if acceleration waves are used. We considered both combinations of acceleration/unit impulse and displacement/initial displacement. The equivalent models produce identical output, but the displacement input is easier to interpret and modify by the user, as it reflects what the player does (that is, moving the string).

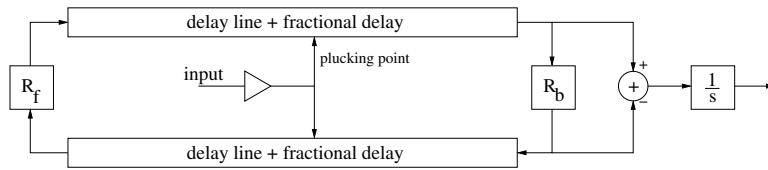## 10.3.2 Single delay line model

Figure 10.8: Double delay line model with acceleration as the wave variable. The final output is a velocity, obtained by integration of the acceleration waves. Both reflection coefficients R_f and R_b represent fixed ends, and are thus phase inversive.
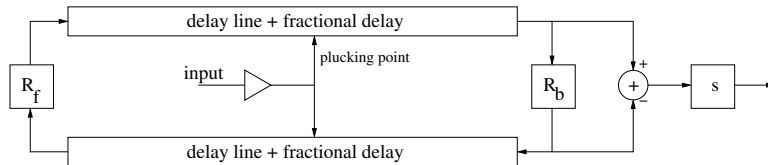

Figure 10.9: Double delay line model with displacement as the wave variable. The output is again the bridge velocity, obtained by differentiating the displacement waves.

In certain cases, the initial double delay line digital waveguide model can be reduced to a loop containing only a single delay line [43]. In this case, only the observation of the force signal output at the bridge is possible. To obtain the same output as the two-delay line model, one has to apply an "equivalent" input to the single delay line model. Figure 10.10 shows the two-delay line model of which we now derive the single-delay line equivalent model. The signal X_2 entering the right-going lossy and dispersive delay line can be written as

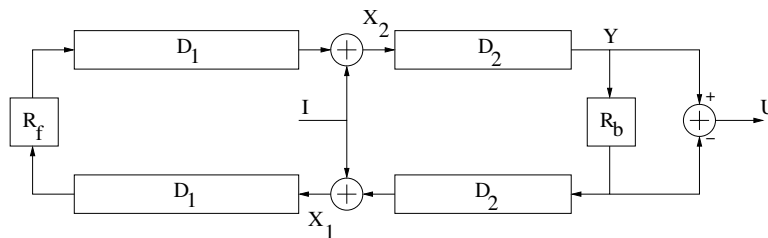$$X_2 = I + R_f D_1^2 (I + R_b D_2^2 X_2), \tag{10.9}$$

Figure 10.10: Reference two-delay line model. This model will be simplified such that it only contains one delay line.
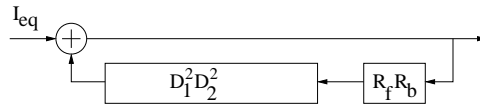
Figure 10.11: The single delay line string loop used in conjunction with I_{eq} has lower computational complexity than the double delay line digital waveguide model it is derived from.

.

or solved to X_2

$$X_2 = I(1 + R_f D_1^2)\frac{1}{1 - R_f R_b D_1^2 D_2^2}. \tag{10.10}$$

The first factor provides us with an equivalent input I_{eq}

$$I_{eq} = I(1 + R_f D_1^2). \tag{10.11}$$

Similarly, the output U can be written as

$$
\begin{aligned}
U &= Y - R_b Y \\
&= D_2 X_2 (1 - R_b) \\
&= I_{eq}\frac{1}{1 - R_f R_b D_1^2 D_2^2} D_2 (1 - R_b).
\end{aligned} \tag{10.12}
$$

The rightmost factor consists of a pure time delay and the filtering effect of the bridge reflection. Of this, only the filtering effect is relevant for the model. The middle factor is the single delay line loop shown in figure 10.11. The complete single delay line model that produces the same output as the two delay line model of figure 10.10 is shown in figure 10.12.

Example 10.1 ⋆ **Track 21** is an example of a low A note, generated using a single delay line, single polarization model, excited with a triangular pluck. △

10.3.3   Non-linear effects

Several non-linear effects occur in typical string instruments. For the acoustic guitar, the two main non-linear effects are tension modulation (inherently caused by the vibration of a real string), and fret beating caused by the physical dimensions of the instrument.

Tension modulation non-linearity

As the shape of the string changes during its motion, its length and tension changes as well. Perceptually, two phenomena occur: a pitch variation and the
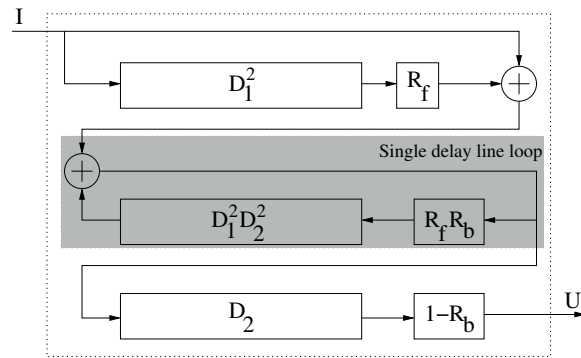
Figure 10.12: The complete single delay line string model equivalent to the two delay line model of figure 10.10
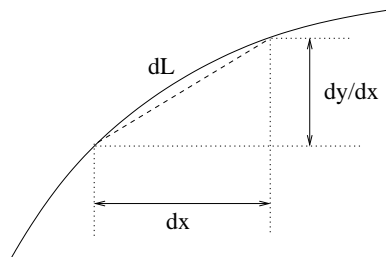

Figure 10.13: The length of an infenitesimal piece of string due to a displacement in the y direction is found to be $dL = \sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2}$.

generation of harmonics. The non-linear properties of string have been studied both theoretically and experimentally [16], [54], [65]. The tension of a vibrating string is found as

$$T = T_0 + \Delta T = T_0 + \frac{ES\Delta L}{L_0},\qquad(10.13)$$

with E the Young modulus, S the cross-sectional area of the string, and T_0 and L_0 the string tension and length at rest. The length change ΔL can simply be calculated as the sum of all the length changes along the string (refer to figure 10.13) :

$$\Delta L = \int_0^{L_0} \sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2}\, dx - L_0.\qquad(10.14)$$

For displacements small compared to the string length, this can be simplified

to

$$\Delta L \approx \frac{1}{2} \int_0^{L_0} \frac{\partial y}{\partial x}^2 \, dx. \tag{10.15}$$

For the exponentially damped fundamental waveform of the form

$$f_0(x,t) = A \sin(\omega_0 t) sin(\frac{\pi x}{L}) e^{-a_0 t}, \tag{10.16}$$

with A the initial amplitude, ω_0 the fundamental frequency, and a_0 the damping, the length change is well approximated by

$$\Delta L \approx \frac{A^2 \pi^2}{4L} \sin^2(\omega t) e^{-2a_0 t}. \tag{10.17}$$

We see that ΔL decays exponentially at twice the rate of the sinusoid. More complex expression for the string displacement introduce sum and difference frequencies into ΔL, causing the generation of missing harmonics. This analysis is valid if the tension is immediately spread over the whole string by longitudinal wave propagation. The wave propagation speed for longitudinal waves in a string is constant and given by

$$c_L = \sqrt{\frac{E}{\rho}}, \tag{10.18}$$

or quite a lot faster than for transversal wave propagation, so it can be safely assumed that this condition is satisfied. The tension modulation affects the transversal speed c as

$$c = \sqrt{\frac{T}{\rho}}, \tag{10.19}$$

and the fundamental frequency is linearly related to the average wave propagation speed. This equation can be combined with equation 10.13:

$$c = \sqrt{\frac{T_0 + \frac{ES\Delta L}{L_0}}{\frac{\rho L}{L+\Delta L}}}, \tag{10.20}$$

or

$$c = c_0 \sqrt{1 + \left(1 + \frac{ES}{T_0}\right) \frac{\Delta L}{L_0} + \frac{ES}{T_0} \left(\frac{\Delta L}{L_0}\right)^2}. \tag{10.21}$$

Considering only the fundamental, we get that

$$c \approx c_0 \left(1 + \frac{A^2 \pi^2}{8L^2}(1 + \frac{ES}{T_0}) e^{-2a_0 t} \sin(\omega_0 t)\right). \tag{10.22}$$

It follows from these equations that steel strings are more affected by tension modulation because of their much larger Young modulus, despite of their generally smaller radius. The deviation from the nominal wave propagation speed is

$$\Delta c = c - c_0. \tag{10.23}$$

For use in the digital waveguide framework, a discrete formulation is necessary. The discrete formulation estimates the length of the string by directly integrating equation (10.14), and adjusts the length of the delay line and the parameters of the fractional delay filter dynamically to match the desired length [108], [105]. In this case, only the fractional delay filter has to be adjusted. The discrete length change for a two-delay line digital waveguide model is approximated by

$$\Delta L(n) = \sum_{k=0}^{L_0-1} \sqrt{1 + [s_r(n,k) + s_l(n,k)]^2} - L_0, \qquad (10.24)$$

where $s_r(n,k)$ and $s_l(n,k)$ denote the right and left going slope waves, respectively. If displacement waves are used, a first differentiation is needed. If we assume that $[s_r(n,k) + s_l(n,k)]^2 \ll 1$, equation (10.24) can be approximated by

$$\begin{aligned} \Delta L &\approx \sum_{k=0}^{L_0-1} [1 + \frac{[s_r(n,k) + s_l(n,k)]^2}{2}] - L_0 \\ &= \frac{1}{2} \sum_{k=0}^{L_0-1} [s_r(n,k) + s_l(n,k)]^2. \end{aligned} \qquad (10.25)$$

The deviation from the nominal wave propagation speed leads to a deviation in total delay time

$$\begin{aligned} d(n) &= -T \sum_{l=n-1-L_0}^{n-1} \Delta c(l) \\ &= -Tc_0 \sum_{l=n-1-L_0}^{n-1} \sqrt{1 + \left(1 + \frac{ES}{T_0}\right)\frac{\Delta L}{L_0} + \frac{ES}{T_0}\left(\frac{\Delta L}{L_0}\right)^2} \\ &\approx -\frac{1}{2} \sum_{l=n-1-L_0}^{n-1} (1 + \frac{ES}{T_0}\frac{\Delta L(l)}{L_0}), \end{aligned} \qquad (10.26)$$

and $Tc_0 = 1$ since T is the sampling interval and c_0 equals the sampling frequency in a digital waveguide system. If one is only interested in a crude approximation of the tension modulation effects, it is also possible to estimate the delay deviation $d(n)$ on the behavior of the fundamental as given by equation (10.22). If the measured fundamental frequency at the start of the signal is $\tilde{f}_0$, the initial length difference is

$$\Delta L = \frac{F_s}{f_0} - \frac{F_s}{\tilde{f}_0}. \qquad (10.27)$$

The decay rate $e^{-a_0 t}$ of the string is known in advance, and thus the delay deviation $d(n)$ is

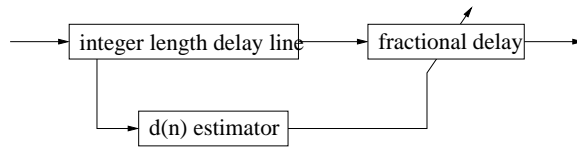$$d(n) = \Delta L e^{-2at}. \qquad (10.28)$$

Figure 10.14: A tension modulated digital waveguide. The deviation of the delay $d(n)$ is estimated and used as a parameter for the fractional delay filter.

This can easily be calculated recursively.

There are several practical implementations of the time-varying fractional delay filters [107]. A schematic representation of a digital waveguide with tension modulation is shown in figure 10.14 A switching nonlinear filter is well-suited for this purpose [85].

Example 10.2 ⋆ `CD Track 26-27` show the difference between a model of a nylon string and steel string. The only difference included here are the properties of the string material, all other (damping, fundamental) are the same. There is a subtle difference in timbre, especially in the first second second. △

Fret beating

The vertical transversal displacement of a guitar string is bounded by the frets and the fingerboard. When the excitation is large enough for the strings to touch the frets, a non-linear distortion of the signal occurs. This can be considered as either a hard, saturating non-linearity (the fret limits the movement), or more accurately as both a hard saturation of the displacement and an extra velocity input (elastic bouncing on the frets). It is easy to include this effect in the model, by comparing the output of the delay lines to a desired maximum output and by saturating and adding the velocity input. When including fret beating in a digital waveguide model, it is more efficient to use the two-delay line model, as the fret beat has to be added in the two propagation directions.

10.4   Dual-polarization string models

Transversal waves can be decomposed in waves traveling in a plane parallel to the soundboard (y_H), and traveling in a plane perpendicular to the soundboard (y_V). Due to the structure of the bridge and the nut, the terminations are not isotropic, causing a difference in the properties of the wave propagation
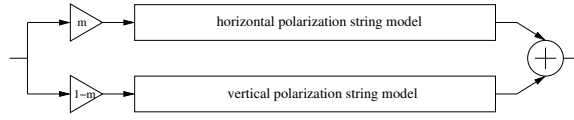
Figure 10.15: Schematic representation of a dual-polarization model. Two independent string models are use for the horizontal and vertical transversal wave propagation modes.

on the string. The differences are large enough such that the string exhibits slightly different resonance frequencies and a different damping for y_H and y_V. The slight difference in resonance frequencies is audible as beating in most of the partials. In this section, we consider the two polarizations as uncoupled, independent, single-polarization strings. The coupled case will be analyzed in chapter 11. Figure 10.15 shows a schematic reprentation of a dual-polarization model. The total output of the dual-polarization string can be written as

$$y(t) = \sum_{k=0}^{N} A_{k,1} e^{-a_{k,1}t} \sin(\omega_{k,1}t + \phi_{k,1}) + A_{k,2} e^{-a_{k,2}t} \sin(\omega_{k,2}t + \phi_{k,2}). \quad (10.29)$$

This equation can be reformulated such that it allows for easier parameter estimation, by setting

$$A_k = \min(A_{k,1} e^{-a_1 t}, A_{k,2} e^{-a_2 t}), \quad (10.30)$$

$$\Delta A_k = |A_{k,1} e^{-a_1 t} - A_{k,2} e^{-a_2 t}|, \quad (10.31)$$

$$\omega_k = \frac{\omega_{k,1} + \omega_{k,2}}{2}, \quad (10.32)$$

$$\Delta \omega_k = \frac{\omega_{k,2} - \omega_{k,1}}{2}, \quad (10.33)$$

$$\phi_k = \frac{\phi_{k,1} + \phi_{k,2}}{2}, \quad (10.34)$$

$$\Delta \phi_k = \frac{\phi_{k,2} - \phi_{k,1}}{2}, \quad (10.35)$$

it can be written as

$$y(t) = \sum_{k=0}^{N} A_k \cos(\Delta \omega_k t + \Delta \phi_k) \sin(\omega_k t + \phi_k) + \sum_{k=0}^{N} \Delta A_k \sin(\omega_1 t + \phi_1), \quad (10.36)$$

assuming that $A_{k,1} e^{-a_1 t} > A_{k,2} e^{-a_2 t}$. The amplitude envelope for each partial is then simply

$$\mathcal{E}_k = A_k \cos(\Delta \omega t + \Delta \phi) + \Delta A_k. \quad (10.37)$$

The local maxima of the amplitude envelope correspond to

$$\text{maxima} = \max(A_{k,1} e^{-a_1 t}, A_{k,2} e^{-a_2 t}) + A_k, \quad (10.38)$$
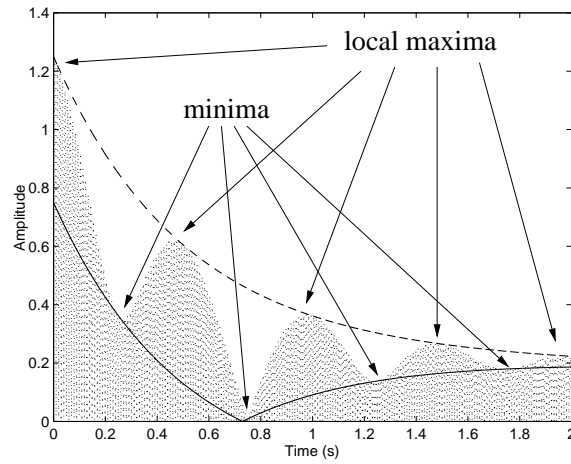
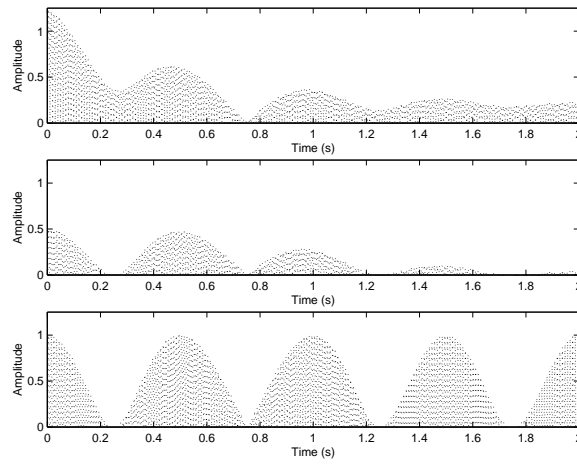

Figure 10.17: (top) The signal, (middle) Detrended signal, (bottom) Amplified signal

Transform, look at the Energy Decay Relief method and propose the subband Hankel Singular Value Decomposition. The different methods are then compared. For simplicity, we first analyze the methods for the theoretical single-polarization case, and elaborate further on the more accurate dual-polarization case.

We assume that the signals can be written as

$$y(t) = \sum_{i=0}^{N} A_i e^{-a_i t} sin(\omega_i t + \phi_i) + n(t), \qquad (10.40)$$

where $n(t)$ is additive white noise.

The determination of the frequency of a partial is a well-studied subject and an essential part of additive synthesis and sinusoidal modeling. The frequency can be determined in either the time or the frequency domain [91].

10.5.1   Time domain estimation

Time-domain methods include several forms of zero-crossing and autocorrelation. One method commonly used for the determination of the fundamental frequency, is based on the autocorrelation function [89]. The short-time autocorrelation function operating on a short segment of signal $x(n)$ is defined

as

$$\phi_k(m) = \frac{1}{N} \sum_{i=0}^{N-1} x(i+k)x(i+k+m)w(i)w(i+m), \qquad (10.41)$$

with $0 \le m \le M_C - 1$, M_C is the number of autocorrelation points to be computed, N the number of samples, and $w(n)$ the window function. The estimated fundamental frequency is obtained by searching for the maximum value of $\phi_k(m)$ for each k in a meaningful range of m. Quadratic interpolation is then used to have sub-sample accuracy for the maximum. It appears that the type of windows used has an important effect on the results obtained. Rectangular windows yield poor results, while Hamming, Hanning, or Blackman windows have a more acceptable quality. This method can be extended to determine all the partials, by first filtering the signal with a bandpass filter centered around the partial. However, it is not suited for the detection of two closely spaced partials, and it is much more efficient to use the frequency domain methods.

10.5.2   Short-Time Fourier Transform

The easiest method is based on the Short Time Fourier Transform (STFT) of a recording. The amplitude-time relationship is determined for each band containing a partial, and the parameters are fitted such that the model output matches the recording. This is a two-step process. First, the frequencies of the partials are determined, and secondly, the damping ad amplitudes are computed.

In the frequency domain, the way to determine the frequency of a partial is to use a long FFT, with fine frequency resolution. There are several disadvantages to this method, the most important one being the long size of the signal required to achieve sub-Hertz resolution. It is quite hard to produce and record a guitar note of more than ten seconds, especially for the higher strings. In the single-polarization case, we know that there is only one sine in each bin for relatively small FFT sizes. For the lowest note on a guitar with normal tuning, the FFT size of 1024 is already sufficient. A sinewave is represented by a sinc-like waveform centered around its frequency in the frequency domain. Using quadratic interpolation, we are able to determine the frequency with sub-bin accuracy. The peak with frequency f and amplitude A in bin k_p of spectrum X is found as the top of a parabola $y(x) = a(x-p)^2 + b$ with

$$p = \frac{1}{2}\frac{\alpha - \gamma}{\alpha - 2\beta + \gamma}, \qquad (10.42)$$

$$f = k_p + p, \qquad (10.43)$$

$$A = \beta - \frac{1}{4}(\alpha - \gamma)p, \qquad (10.44)$$

where

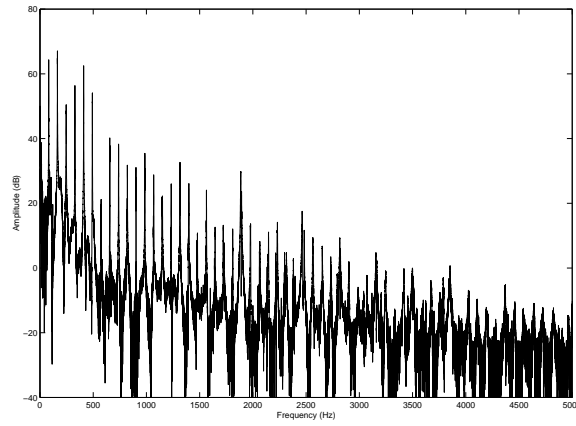$$\alpha = y(-1) = 20 \log_{10} |X(K_p - 1)|, \qquad (10.45)$$

Figure 10.18: The spectrum of a recording of low E played on a Yamaha C70 classical guitar. The spectrum was calculated using a 262.144-point FFT, with bin bandwidth of 0.17 Hz.

$$\beta = y(0) = 20 \log_{10} |X(k_p)|, \tag{10.46}$$

$$\gamma = y(1) = 20 \log_{10} |X(k_p + 1)|. \tag{10.47}$$

The situation is more complicated in the dual-polarization case, where each peak is actually composed of two very closely spaced partials. To resolve two sinusoids separated by Δ Hz, it is required that the main lobe bandwidth $B_f < \Delta$. This can be achieved by using a long window length, but is neither practical nor accurate. Figure 10.18 shows a part of the spectrum of a recording of an acoustic guitar, calculated using a $2^{1}8$-point FFT, with bin bandwidth of 0.17 Hz. Figure 10.19 shows a detailed view of the frequency bins around the third harmonic, where the splitting of the peak in two partials is visible.

The damping of the system poles can be determined in the pure frequency domain by evaluating the 3dB bandwidth of the peak, but due to the fact that the signal is only slightly damped, the peaks are very sharp and this method is therefore extremely inaccurate. To determine the damping and amplitudes accurately, and the frequency difference for the dual-polarization case, one has to switch to time-frequency representations. The most widely used one is the Short-Time Fourier Transform. An excellent overview of the Short Time Fourier Transform and the use of windowing can be found in [2], [36]. Other methods include the Bark Transform and the various wavelet transformations. All these methods have a time resolution-frequency resolution trade-off.

The STFT is composed as a series of DFTs computed on windowed segments of the signal. The segments are separated in time by the hop size parameter, padded with zeros if necessary, and multiplied by a window function. The same
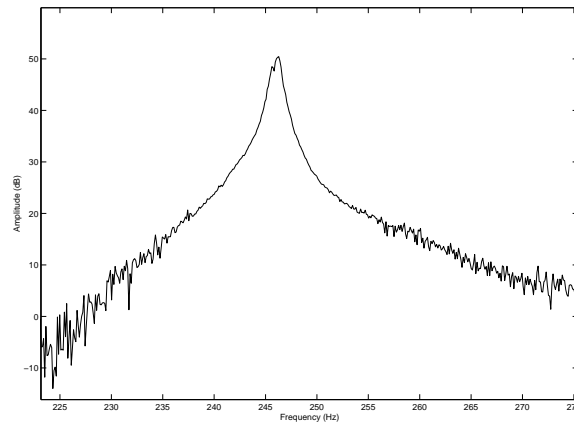
Figure 10.19: Detailed view of figure 10.18. The double peak due to the dual-polarization nature of the guitar is visible.

frequency bin in the sequence of FFTs represents one subband of the STFT. Each subband of the STFT is critically sampled with sampling frequency F_s/N, where F_s is the sampling frequency of the original signal and N the window length.

The STFT subbands containing the partials of the signal provide the most useful information. The amplitude of the partial over time allows the determination of both the beating frequency and the damping as discussed in section 10.4. The accuracy of this method is limited by the presence of noise, and the compromise between frequency resolution and time resolution. The size of the sequence of FFTs has to be chosen such that each subband only contains one partial. Taking into account the effects of windowing (leakage), this defines a minimum window size. Healivy damped partials tend to damp in a few windows length, and low amplitude partials are drowned in noise, making the procedure more difficult. By using non-linear optimization, we however achieved acceptable results.

10.5.3 Energy Decay Relief

A more accurate method was developed for the determination of room reverberation parameters. This Energy Decay Curve method is based on the backwards integration of an impulse response, and allows the determination of the average of the decay curves with a single measurement. The Energy Decay Relief (EDR) is a frequency-domain generalization, computing the backward integration on the STFT [42]. The EDR of amplitude envelope $A_k(n)$ is calculated as

[42]

$$A_{k,\text{EDR}}(n) = \sum_{i=n}^{\infty} A_k(i), \tag{10.48}$$

where k denotes the partial or subband of the STFT. An efficient way of computing this sum is to reverse the time axis.

For the single-polarization case, we expect a exponentially decaying amplitude for each partial. A linear model is fitted to the log-EDR, weighted such that the noise floor is not taken into account. The linear model $y_k = a_k + b_k n$ is found by minimizing

$$E_k(a_k, b_k) = \sum_{i=0}^{N} W_k(i)[A_{k,\text{EDR}}(i) - y_k]^2. \tag{10.49}$$

The slope b_k corresponds to exponential decay rate.

This method assumes that there is only one partial in the frequency band of interest - in other words, it assumes a single-polarization string. As we have seen, this situation is very unrealistic due to the construction of most string instruments. This method is thus unsuitable for the estimation of guitar strings parameters.

10.5.4 Subband Hankel Singular Value Decomposition

The previous methods are either only suited to find some of the parameters, or lack the precision needed for the accurate determination of all the model parameter. A new subband based method was developed with better properties. It is a subband extension of the Hankel Singular Value Decomposition. We start with the mathematical description of the method and evaluate the results in section 10.5.6.

The goal of the algorithm is to determine the system poles in a signal of length Q. We use a dual-polarization string model, with one delay line for each string, with lengths L_1 and L_2. If the lengths L_1 and L_2 are kept constant, the output signal has the form

$$y(t) = \sum_{i=1}^{N} A_i e^{-\alpha_i t} \sin(\omega_i t + \phi_i), \tag{10.50}$$

where t represents time, N the number of partials, α_i the damping and A_i the initial amplitude of each partial. We now consider a sampled version of this signal (sampling frequency F_s), written as

$$s[n] = \sum_{i=1}^{N} c_i V_i^n, \tag{10.51}$$

in which $c_i \in \mathbb{C}$ represents the initial phase, $V_i = e^{-\alpha_i + j\omega_i} \in \mathbb{C}$ are the poles and $n \in \mathbb{N}$ is the discrete time. One technique commonly used for the identification of exponentially damped sinusoids is based on the Hankel Singular Value Decomposition (HSVD) [110]. This method is analytically exact if the signal conforms to equation (10.50). We start by building a Hankel matrix from the signal of length Q

$$ H = \begin{pmatrix} s(1) & s(2) & \dots & s(Q/2) \\ s(2) & s(3) & & s(Q/2+1) \\ \vdots & & & \vdots \\ s(Q/2+1) & s(Q/2+2) & \dots & s(Q) \end{pmatrix}. \tag{10.52} $$

We then calculate the Singular Value Decomposition (SVD) of this matrix:

$$ H_{(Q/2+1)\times Q/2} = U_{(Q/2+1)\times N} \cdot \Sigma_{N\times N} \cdot T^{\dagger}_{N\times Q/2}, \tag{10.53} $$

for a signal with $N \leq Q/2$ partials. Note that when noise is present, the matrix Σ will be of size $Q/2 \times Q/2$. If we only want $M \leq N$ poles, we select the M largest singular values and corresponding singular vectors. This is equivalent to calculating a rank-M approximation of the signal. We now construct a matrix Z':

$$ Z' = (U_{\downarrow})^{-1} U_{\uparrow}, \tag{10.54} $$

where $U_{\uparrow}$ denotes the upper $Q/2$ rows of U and $U_{\downarrow}$ the lower $Q/2$ rows. It is well known that the eigenvalues of Z' give estimates for the poles of the system [110].

$$ \mathrm{eig}(Z') = (\hat{V}_1, \hat{V}_2, \dots, \hat{V}_N). \tag{10.55} $$

The estimates for ω_i and α_i of the sampled signal s follow:

$$ \hat{\omega}_i = \Im(\log(\hat{V}_i)), \tag{10.56} $$
$$ \hat{\alpha}_i = -\Re(\log(\hat{V}_i)). \tag{10.57} $$

To accurately determine the poles in the presence of noise, a long signal ($Q > 500$) is needed. Due to the size of H this may lead to computational problems in the calculation of the SVD. These problems can be overcome by using a subband scheme that analyzes several critically sampled signals (with a reduced number of partials) instead of the original signal, as this replaces one large H by several smaller ones [35].

Applying a short-time Fourier Transform (STFT) of length L to signal s yields us an L-band representation

$$ S_m[k] = \sum_{i=1}^{N} \sum_{n=0}^{L-1} e^{-j2\pi mn/L} s[n+kL], \tag{10.58} $$
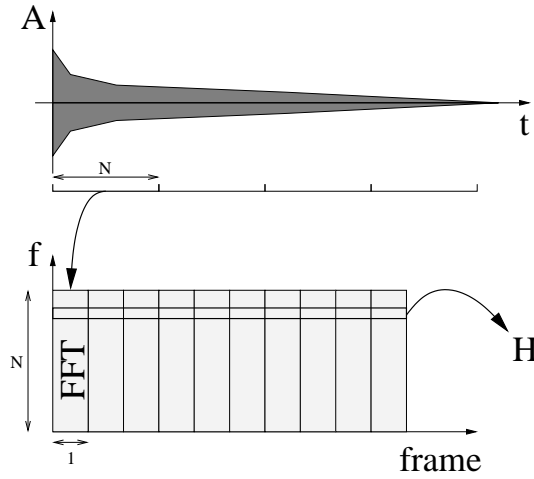
Figure 10.20: The complexity of the calculation is reduced by building an N-band representation and calculating the HSVD on each subband separately. As the number of partials is smaller than the number of bands, we only calculate the poles for the relevant bands.

where $m = 0, 1, \ldots, L - 1$ and k the number of the time frame. Each subband can thus be written as a sum of exponentially damped sinusoids, and therefore conforms to equation (10.51). This implies that we can perform the HSVD algorithm on each subband of the STFT (one sub-band is the sequence of complex samples formed by the selection of one bin of the STFT over all the time frames). This is illustrated by figure 10.20. For each sub-band, the sampling frequency is $F_s^{(m)} = F_s/L$. We obtain an estimate for the poles $\hat{V}_i^{(m)}$ in that sub-band. The damping and frequency for the original pole V_i of signal y relates to $\hat{V}_i'$ as

$$\omega_i = \hat{\omega}_i F_s^{(m)} + \text{bin base frequency}, \tag{10.59}$$
$$\alpha_i = \hat{\alpha}_i F_s^{(m)}. \tag{10.60}$$

Using Least Squares, we can finally determine the phase $\hat{\phi}_i$ and initial amplitudes $\hat{A}_i$. This algorithm is called sub-band HSVD (sHSVD).

For the estimation of the parameters of the dual-polarization model of figure 10.15, the window length of the STFT is chosen large enough such that only one pair of partials (one partial for each mode) can be found in one frequency bin. For each bin that contains partials, the entire time-frame sequence is used as the input for the sHSVD algorithm but calculate only two poles, one for each polarization mode. The estimates for ω_i and α_i are thus obtained. The length of the delay lines L_1 and L_2 can easily be found given the knowledge of

ω_i. Once these lengths are known, the reflection coefficient R is found as

$$R_i = e^{-\hat{\alpha}_i \frac{L_i}{F_s}}. \tag{10.61}$$

In the end, the reflection coefficients for both polarizations are known, at every partial. This complete set of reflection coefficients enables the fitting of a linear filter that approximates the frequency-dependence. Result for a test case and a recorded signal are shown in section 10.5.6.

10.5.5   Other methods

Several other methods have been proposed for the determination of the string parameters, usually applied to piano strings. These methods include the general system identification methods, Prony's method, or matrix pencil techniques. Due to the structure of the piano, one is able to obtain longer and more controled signals than is the case with the acoustic guitar, and the parameters can be determined with greater accuracy.

10.5.6   Comparison

A first test consisted of estimating the parameters of a known system. To synthesize the test signal, a dual-polarization waveguide model was used, in which one polarization had a frequency-independent reflection coefficient and the other polarization had a lowpass reflection characteristic. White noise was added to simulate recording imperfections. The results can be seen in figure 10.21. For each partial, we get a pair of frequencies (f_1, f_2). These are then assigned to one polarization mode (horizontal or vertical) by grouping all the highest and all the lowest frequencies. This way, we get two series of frequencies that almost satisfy $f_1, 2f_1, 3f_1, \ldots$ and $f_2, 2f_2, 3f_2, \ldots$. The deviation of the measured series compared to the theoretic series is an indication of the inharmonic behavior. It is clear that the sHSVD algorithm is capable of accurately separating the characteristics of the two polarizations. Table 10.2 gives a comparison of the result obtained using both the STFT and the sHSVD methods. The sHSVD is more accurate, which is especially important at the lower frequencies, as these results are used to determine the length of the delay lines of the digital waveguide models.

The second test was done on a recording of a nylon string guitar. The goal is to determine the parameters of the model shown in figure 10.22. Plotting the frequency difference of the partials, relative to the number of the partial, gives us information on the two polarization modes (figure 10.23). The fundamental frequencies have a 0.082 Hz difference, which is confirmed by the audible beating in the higher partials. This leads to the delay line lengths L_1 and L_2, and the fractional delays. The inharmonicity was determined by calculating the difference between the ideal harmonic series and the measured frequencies,
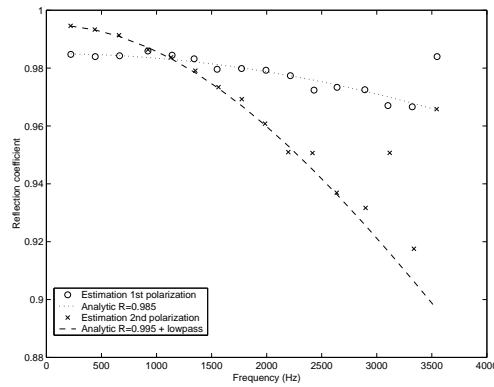
Figure 10.21: Estimated and analytic reflection coefficients for a dual-polarization waveguide model. The estimation is the result of the sHSVD analysis of a test signal where the first polarization had a reflection coefficient R=0.985 (the slight decrease at higher frequencies is caused by the first order Lagrange interpolator used, cfr. chapter 5) and the second polarization had a reflection coefficient R=0.995 combined with a first-order lowpass filter (-3dB at 7 kHz).

as can be seen on figure 10.24. The reflection coefficients for all the partials are shown in figure 10.25. We now determined the reflection filters $\mathcal{H}_1$ and $\mathcal{H}_2$ of the model. The lowpass and allpass parts of the reflection filter were determined such that a good fit was obtained for both the required reflection coefficient (results from the amplitude of $\mathcal{H}_1$ and $\mathcal{H}_2$) and inharmonicity (results from the phase). The model was excited with an impulse (comb-filtered to approximate the plucking position) and its output was filtered using the measured impulse response of the guitar body at the bridge. Figure 10.26 compares the time evolution of the fundamental of the recorded signal with the output of the dual-polarization model. Figure 10.27 are waterfall plots of the STFT of both signals. The differences at higher frequencies are probably due to differences in the excitations and the absence of coupling in the model.

Example 10.3 ⋆ CD Track 22 was generated using a dual-polarization string model, without string coupling. The two polarizations have slightly different decay rates and fundamental frequency. Compare this to the single-polarization case of track 21. △

10.6 Conclusion

In this chapter, a single-string model for the acoustic guitar has been built. The properties of ideal and real strings have been discussed, including ten-
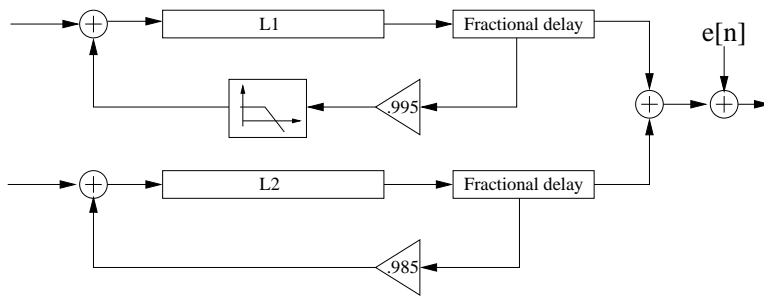
Figure 10.22: The dual-polarization digital waveguide system used to test the parameter estimation routines. One polarization has a reflection coefficient of 0.995, and its output is filtered with a first-order lowpass filter with cut-off at 7 kHz. The second polarization has a reflection coefficient of 0.985. First-order Lagrange interpolation was used as the fractional delay method. The system was excited with a triangular initial displacemnt. The estimation error increases for higher frequencies, due to the lower amplitude of these partials, and thus to the lower signal-to-noise ratio.

Partial	Analytic Exact f	Exact R	STFT f	R	sHSVD F	R
1	220.5	.995	218.1	.995	220.5	.995
	220.0	.989	217.6	.989	220.0	.989
3	661.5	.995	653.3	.995	661.5	.995
	660.0	.983	651.9	.983	660.0	.982
5	1102.5	.995	1136.9	.995	1102.6	.995
	1100.0	.969	1139.3	.974	1100.3	.972
7	1543.5	.995	1557.3	.995	1547.1	.995
	1540.0	.940	1560.2	.984	1544.6	.961

Table 10.2: Numerical results for the estimation using either the STFT method or subband HSVD. The accuracy at the lower frequencies is the most important factor in favor of the sHSVD method.
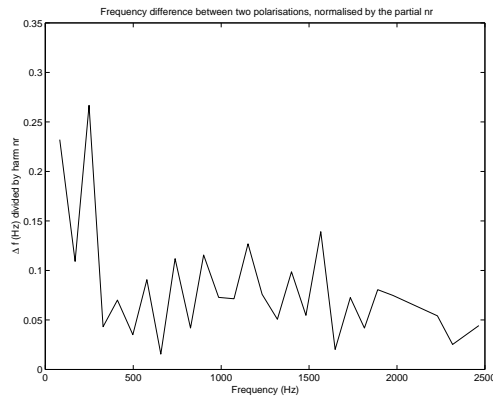
Figure 10.23: The frequency difference between the two polarizations divided by the number of the partial in the harmonic series. From this graph, we can conclude that the two polarizations have fundamental frequencies with a mean difference of 0.082 Hz
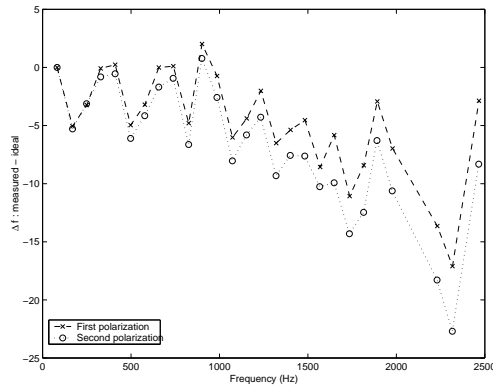

Figure 10.24: The frequency difference between the ideal harmonic series and the extracted frequencies. This is due to the string stiffness and can be simulated with an allpass filter.

sion modulation non-linearity and fret beating. The dual-polarization models account for vibrations in two spatial planes. In section 10.5, several time-domain, frequency-domain and subband parameter estimation algorithms have been proposed and compared. Very accurate single-string models can be determined using the techniques from this chapter.
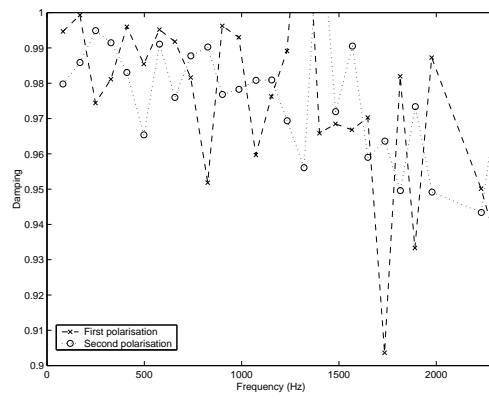
Figure 10.25: Estimated reflection coefficients for a dual-polarization waveguide model. The estimation is based on a recording of a low E on a Yamaha C70 nylon string guitar with AKG C4000B microphones ($F_s = 44.1$kHz, 16 bit).

Figure 10.26: Amplitude of the fundamental of the recorded guitar signal and simulation of the model shown in figure 10.22 with the parameters obtained with sHSVD. The model was excited with an impulse and its output was filtered using the impulse response of the guitar body.

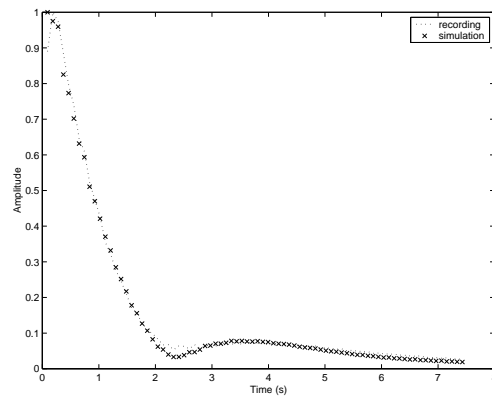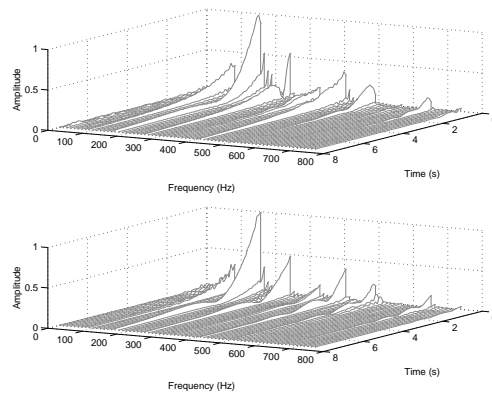Figure 10.27: Short-time Fourier spectra of (top) the recorded signal and (bottom) the output of the calibrated dual-polarization model.

Chapter 11

Multiple String Models

Approach your guitar intelligently, and if there are limits, don't
deny them. Work within your restrictions. Some things you can
do better than others, some things you can't do as well. So
accentuate the positive.

Chet Atkins

One important aspect that determines the natural quality of the sound of a
string instrument is the sympathetic coupling between the strings [31]. As one
string is struck, plucked or bowed, other strings of the instrument start to vi-
brate and change the decay or add beating to some partials of the sound. This
is an important aspect as a sound without some amount of beating is almost
immediately recognized as synthetic. Several structures to simulate the sym-
pathetic coupling between strings have been proposed. In this chapter, a novel
general framework is proposed that allows the comparison of the performance
of these structures and their effect on the sound. The framework also gives
clues as how to improve the parameter estimation for the structures and how
to verify the stability.

The first section of this chapter describes the mathematical background of the
behavior of N coupled oscillators and N coupled digital waveguide systems. In
section 11.2, several string coupling models are compared within the proposed
framework. Finally, we propose a simplified, fully coupled instrument model
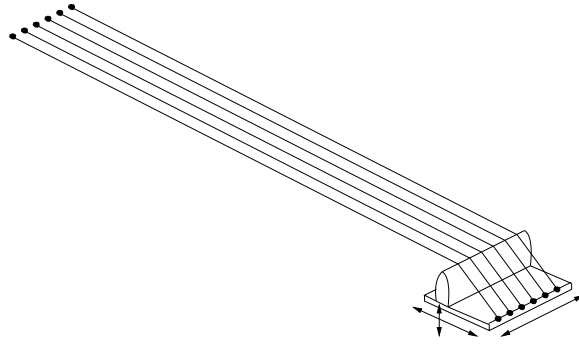and its associated parameter estimation routines in section 11.3.

Figure 11.1: All strings interact at the bridge level. The common movement of the bridge and the top plate is an extra input to the strings.

11.1   String coupling

Coupling is a physical phenomenon where energy flows between two or more connected systems. If the energy flow is large enough, the systems cannot be considered as separate entities, but as one large, interacting system, as shown in figure 11.1. There exist several studies on the coupling and synchronization of systems, but these studies are concerned with the forward problem (given the coupling, what is the effect) and not as much with the estimation problem (given the effect, how strong is the coupling).

11.1.1   Non-conservatively coupled oscillators

The differential equation describing a harmonic oscillator is simply

$$\ddot{y} + \omega_0^2 y = 0. \tag{11.1}$$

Several independent oscillators can be coupled by adding a coupling term f_k. This term denotes the driving of the oscillators by the other oscillators of the system. The equation becomes

$$(\frac{d}{dt} - j\omega_{0k})(\frac{d}{dt} + j\omega_{0k})y_k = f_k, \tag{11.2}$$

for $k = 1, 2, \ldots, N$ coupled oscillators. After some simplifications, we have

$$(\frac{d}{dt} - j\omega_{0k})y_k = \sum_{k'} \Omega_{kk'} y_{k'}, \tag{11.3}$$

or in vector form

$$\frac{d\mathrm{Y}}{dt} = \Omega \mathrm{Y}, \tag{11.4}$$

in which the dynamical matrix Ω is made up of the elements $\Omega_{kk'} + j\omega_0$ [119]. These equations describe only one partial. A signal containing P partials on an instrument with N strings has a set of $N \times P$ equations associated with it. The horizontal and vertical polarizations are taken care of by introducing two independent (but coupled) strings; one for each polarization.

The typical situation for the guitar is that we have two polarizations for each string and possibly a coupling with a matching mode on another string. This gives as a minimum size for Ω a two-by-two and a maximum of twelve-by-twelve matrix, assuming that the system is linear and that the string mode frequencies are well-separated. When we look at only one string with two polarizations, the dynamical matrix is

$$\Omega = \begin{bmatrix} j\omega_1 + \xi_{11} & \xi_{21} \\ \xi_{12} & j\omega_2 + \xi_{22} \end{bmatrix}. \tag{11.5}$$

where ω_k is the string/mode frequency and ξ_{kk} a measure for the coupling. An acceptable situation is that $\xi_{11} \neq \xi_{22}$ and $\xi_{12} = \xi_{21}$. The eigenvalues λ_k of this matrix are the poles of the system. We have

$$\Omega = W\Lambda W^{-1}, \tag{11.6}$$

where W contains the eigenvectors and Λ the complex eigenvalues. The generic solution of equation 11.2 is a sum of damped sinusoids:

$$y_k(t) = \sum_n A_{kn} \exp(j\lambda_n t), \tag{11.7}$$

where A_{kn} denotes the amplitude of the sinusoid. The vector $\mathbf{p}$ of the amplitudes is determined by calculating

$$\mathbf{p} = W^{-1}\mathbf{s}, \tag{11.8}$$

where $\mathbf{s}$ contains the desired start amplitudes. In the two-oscillator case, we have

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \tag{11.9}$$

$$W = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}, \tag{11.10}$$

$$\vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}. \tag{11.11}$$

The solution of equation 11.3 is

$$y_1 = p_1 w_{11} e^{\lambda_1 t} + p_2 w_{12} e^{\lambda_2 t}, \tag{11.12}$$

$$y_2 = p_1 w_{21} e^{\lambda_1 t} + p_2 w_{22} e^{\lambda_2 t}, \tag{11.13}$$

$$y = y_1 + y_2. \tag{11.14}$$

The signal y is the theoretical approximation of the output of a system of coupled oscillators.

11.1.2 Coupled waveguides

When we look at single delay-line (SDL) waveguide implementations of a guitar with coupled strings, we have a structure as shown in figure 11.2 [103], [43], with $N = 12$ fractional delay lines. In the most general case, the delay line input $\tilde{y}$ is calculated as the sum of the system input I and the product of the vectorized output y of the delay lines and a coupling matrix M

$$\tilde{y} = My + I, \tag{11.15}$$

The coupling matrix is

$$M = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1N} \\ c_{21} & c_{22} & c_{23} & & \\ c_{31} & c_{32} & c_{33} & & \vdots \\ \vdots & & & \ddots & \\ c_{N1} & & \cdots & & c_{NN} \end{bmatrix}, \tag{11.16}$$

where the elements c_{km} $(k \neq m)$ represent the (lumped) transmission from delay line m to delay line k, and the elements c_{kk} represent the reflection for delay line k. Note that, in general, the c_{km} elements are complex and frequency-dependent and should be implemented as linear filters. To study the system, it is easier to consider a delay-length independent coupling matrix $\tilde{M}$

$$\tilde{M} = \begin{bmatrix} \tilde{c}_{11} & \tilde{c}_{12} & \tilde{c}_{13} & \cdots & \tilde{c}_{1N} \\ \tilde{c}_{21} & \tilde{c}_{22} & \tilde{c}_{23} & & \\ \tilde{c}_{31} & \tilde{c}_{32} & \tilde{c}_{33} & & \vdots \\ \vdots & & & \ddots & \\ \tilde{c}_{N1} & & \cdots & & \tilde{c}_{NN} \end{bmatrix}, \tag{11.17}$$

where

$$\tilde{c}_{kk} = c_{kk}^{1/L_k}, \tag{11.18}$$

$$\tilde{c}_{km} = c_{km}/L_k., \tag{11.19}$$

and L_k is the length of the delay line in samples. This discrete-time coupling matrix $\tilde{M}$ could be seen as an equivalent unit-delay coupling matrix to the matrix M and is related to the continuous-time Ω matrix by a z-to-s plane transformation. For instance, using an inverse backward Euler transform, in the two-oscillator case, we calculate Ω as

$$\Omega = \begin{bmatrix} j(\omega_0) + (\tilde{c}_{11} - 1)F_s & \tilde{c}_{21}F_s \\ \tilde{c}_{12}Fs & j\omega_1 + (\tilde{c}_{22} - 1)F_s \end{bmatrix}, \tag{11.20}$$

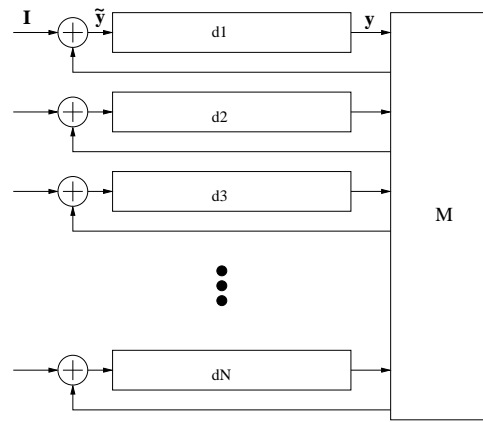where F_s is the sampling frequency.

Figure 11.2: A general waveguide structure with string coupling through a coupling matrix M. Elements of M are complex and frequency dependent.

One digital waveguide can be described as a group of oscillators, one for each partial. The complete dynamical matrix is quite large if we include every partial/oscillator. Two interesting approximations of the complete matrix are analyzed here. The first one includes only the "fundamental" oscillator (at the fundamental frequency for each string). The ensuing Ω_f matrix gives an indication of the overall amplitude envelope for the dominant frequencies. The second variation discards the imaginary part: this is equivalent to the worst case possible, where all oscillators have the same frequency and are easily excited when coupled. This is the case for the lowest common multiple of the string fundamentals. To ensure the global stability of the system, we have to ensure that the real part of the eigenvalues of this worst-case matrix Ω_w are negative (we need a damped system).

11.2 Coupled string models

Several different coupled string models can now be compared within this framework. Starting from the un-coupled situation, we proceed to the more advanced bridge coupled and simplified coupled models found in literature. Finally, we propose a more complete fully coupled model.

11.2.1 No coupling

The simplest case is when there is no coupling at all (the digital waveguides are independent systems). The matrix M is simply given by

$$
M = \begin{bmatrix} c_{11} & 0 & \cdots & 0 \\ 0 & c_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & c_{NN} \end{bmatrix}, ,
\tag{11.21}
$$

and the corresponding Ω_f matrix is

$$
\Omega_f = \begin{bmatrix} j\omega_1 + (\tilde{c}_{11} - 1)F_s & & 0 \\ & \ddots & \\ 0 & & j\omega_N + (\tilde{c}_{NN} - 1)F_s \end{bmatrix}.
\tag{11.22}
$$

The eigenvalues of Ω_f are the diagonal elements. Stability is ensured if $\mathrm{eig}(\Omega_w) = \mathrm{eig}(\Re\{\Omega_f\}) = \Re\{\tilde{c}_{kk} - 1\} < 0$ or $\tilde{c}_{kk} < 1$. This is physically sound as no amplification occurs.

It is now trivial to calibrate the matrix M and the delay lengths with a recording. One has to measure the frequency and damping of the partials by applying the sub-band Hankel Singular Value Decomposition. For the estimation of the parameters using sHSVD, we choose the window length of the STFT large enough such that only one pair of partials (one partial for each mode) can be found in one frequency bin. For each bin that contains partials, we use the entire time-frame sequence as the input for the sHSVD algorithm but calculate only two poles, one for each polarization mode. We obtain the estimates for ω_i and $\alpha_i = (\tilde{c}_{ii} - 1)F_s$. The length of the delay lines L_1 and L_2 can easily be found given the knowledge of ω_i. Based on these results, one calculates the c_{kk} elements. For every partial, a different set of coefficients is found. One could fit a simple IIR filter (typically, a lowpass filter) and use it to calculate the reflection.

11.2.2   Dual polarization coupling

In this model, the two polarizations are coupled but there is no interstring coupling. A schematic view of a dual-polarization string pair is shown in figure 11.3. As a result, the Ω_f matrix breaks down in $N/2$ blocks of two oscillators.

$$
\Omega_f = \begin{bmatrix} \Omega_{f,1} & & \cdots & 0 \\ 0 & \Omega_{f,2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Omega_{f,6} \end{bmatrix},
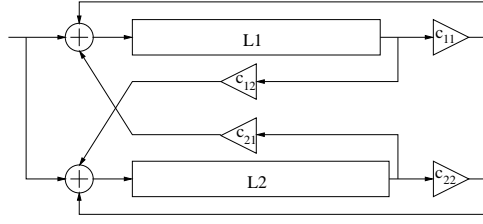\tag{11.23}
$$

Figure 11.3: A Dual-polarization string pair. The two delay lines represent the two polarization modes. There is no interstring coupling. The coefficients c_{11}, c_{12}, c_{21} and c_{22} are positive real.

with

$$\Omega_{f,k} = \begin{bmatrix} j\omega_1 + (\tilde{c}_{11} - 1)F_s & \tilde{c}_{21}F_s \\ \tilde{c}_{12}Fs & j\omega_2 + (\tilde{c}_{22} - 1)F_s \end{bmatrix} \tag{11.24}$$

The global system is stable if all the $\Omega_{f,k}$ represent stable systems. The calibration of the matrix M is somewhat more difficult than in the previous case. Using the sHSVD, we measure for each partial the two poles $\lambda_1 = \alpha_{1m} + j \times \omega_{1m}$ and $\lambda_2 = \alpha_{2m} + j\omega_{2m}$. These measured poles are the eigenvalues of the coupled system $\Omega_{f,k}$. We now analytically calculate the eigenvalues of equation (11.24). When solving this expression to find the unknown coefficients $\omega_1, \omega_2, \tilde{c}_{11}, \tilde{c}_{12}, \tilde{c}_{21}$ and $\tilde{c}_{22}$ with known λ_1 and λ_2, there are still two degrees of freedom left. If we add the extra constraints that the sum of the outputs is equal or less than the input, we obtain a set of physically meaningful solutions. One interesting case is when the loss at the reflection point is zero or

$$\tilde{c}_{11} + \tilde{c}_{12} = 1, \tag{11.25}$$
$$\tilde{c}_{21} + \tilde{c}_{22} = 1. \tag{11.26}$$

This gives a unique solution: the maximal values for the coefficients. Note that these constraints also guarantee the stability if $0 < \tilde{c}_{jk} < 1$. The solution is given by

$$\omega_0 = (\omega_{1m} + \omega_{2m})/2, \tag{11.27}$$
$$\epsilon = \sqrt{(\omega_{1m} - \omega_{2m})^2 + 4\alpha_{1m}\alpha_{2m}}, \tag{11.28}$$

and

$$\omega_1 = \omega_0 + \epsilon/2, \tag{11.29}$$
$$\omega_2 = \omega_0 - \epsilon/2. \tag{11.30}$$

The coefficients are

$$\tilde{c}_{21} = \frac{\alpha_{1m}\omega_{2m} + \alpha_{2m}\omega_{1m} - (\alpha_{1m} + \alpha_{2m})\omega_2}{-F_s\epsilon}, \tag{11.31}$$
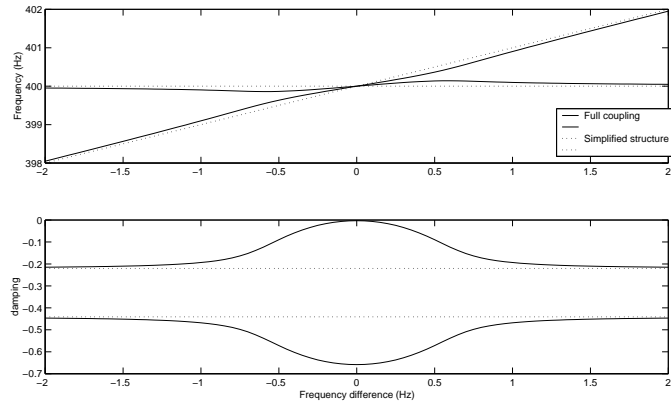
Figure 11.4: This figure shows the comparison of the real (damping) and imaginary (frequency) parts of the system poles, for a resistively coupled model and the un-coupled simplified case, as a function of the mistuning.

$$\tilde{c}_{12} = \frac{\alpha_{1m} + \alpha_{2m}}{-F_s} - \tilde{c}_{21}, \tag{11.32}$$

$$\tilde{c}_{11} = 1 - \tilde{c}_{12}, \tag{11.33}$$

$$\tilde{c}_{22} = 1 - \tilde{c}_{21}, \tag{11.34}$$

where F_s is the sampling frequency. This approach is valid when the movement of the two polarizations are in phase, which is the case for the lower frequency partials. At higher frequencies it is possible that the coefficients $\tilde{c}_{jk}$ become complex-valued. The length of the delay lines can be determined from ω_1 and ω_2, and the reflection coefficients c_{jk} can now simply be determined with equations (11.18) and (11.19). In figure 11.4, we show the fundamental frequencies and corresponding damping (the real part of the pole) for true resistively coupled oscillators and for the un-coupled case, as a function of the mistuning of the two strings. For large mistuning, the two models have the same output. At small mistuning, there is a clear change in decay properties in the coupled case.

Example 11.1 $\star$ `CD Track 23` uses the same dual-polarization model as track 22, but now including string coupling. Notice the different decay rate and beating. $\triangle$

11.2.3   Bridge coupling

In the bridge coupled model, all the outputs of the digital waveguides are tied to the same node. The outputs are summed and input to a linear filter
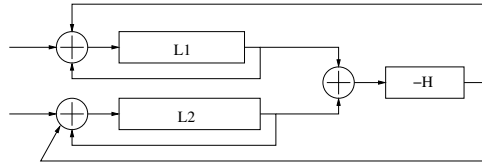
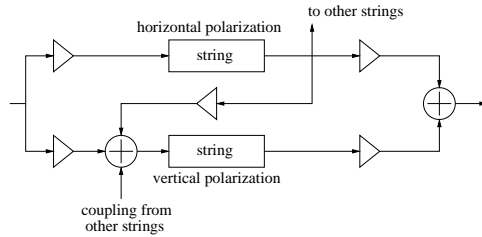Figure 11.5: The bridge coupled model for the case with two delay lines.

Figure 11.6: The simplified model including sympathetic coupling from other strings.

representing the bridge admittance, as can be found on figure 11.5. For the two-string case, the coupling matrix is

$$M = \begin{bmatrix} 1 - H & -H \\ -H & 1 - H \end{bmatrix}.$$ (11.35)

This is system is stable if $|2H(e^{j\omega t}) - 1| \leq 1$. We see that for $H \equiv 0$ we have two independent systems, and $H \equiv 1$ we have one long delay line. It is quite cumbersome to analytically calculate the eigenvalues, as $H(e^{j\omega t}) \in \mathbb{C}$. H can be estimated by numerically optimizing its value such that the eigenvalues of Ω_f match the measured poles.

The bridge coupled model is interesting because it implicitly integrates the resonating body of the instrument into the bridge admittance. The bridge cannot be seen independently from the body as its admittance is measured when it is attached to the body.

11.2.4   Simplified coupling

One simple structure including sympathetic coupling was proposed in [43] and is shown in figure 11.6. In this structure, a coupling matrix C determines the ratio of the output signal of the horizontal polarization to the other strings. We can easily calculate the coupling matrix M. After some manipulations of M and the calculation of the Ω matrix, we obtain the matrix structure shown
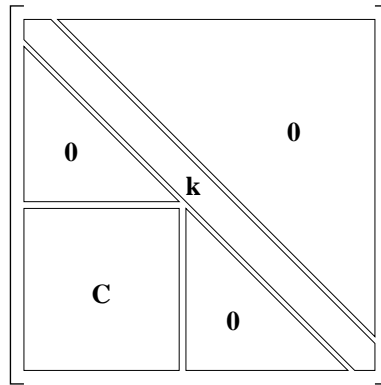
Figure 11.7: The structure of the Ω matrix after grouping the horizontal and vertical polarizations.

in figure 11.7. The eigenvalues are the same as for the un-coupled case. This means that the system poles are unaffected by the matrix C. Strictly speaking, this simplified coupling is not a coupled string model, but an excited resonator bank driving another resonator bank. To reproduce the two-stage decay accurately, one has to use different reflection coefficients for each individual mode and try to satisfy all combinations of strings. This is possible when only the two polarizations are used, but may not be possible when coupling all the strings. For instance, strings 1 and 2 have two-stage decay rates d_1 and d_2 resulting in reflection coefficients c_{11} and c_{22}; strings 5 and 6 have c_{55} and c_{66}. The reflection coefficients needed for accurate coupling of strings 1 and 6 are $c_{11}^{\dagger}$ and $c_{66}^{\dagger}$. It is easy to find a situation where $c_{11}^{\dagger} \neq c_{11}$ and $c_{66}^{\dagger} \neq c_{66}$. It is thus not possible to accurately model the coupling between all the strings at the same time. This model includes sympathetic vibration of the strings, but no coupling effects. An advantage is the guaranteed stability of the complete model if the individual string models are stable.

11.2.5   Full Instrument coupling

If we look at the general structure, we basically have a coupling node and a series of resonators. We can add yet another resonator to the structure: the body itself, as can be seen on figure 11.8. This is fairly similar to the N-dimensional loaded waveguide junction [28], but with feedback to the same junction. It is now possible to give an independent input to the body (e.g. a slap as in *flamenco* playing) that will cause all the strings to vibrate.
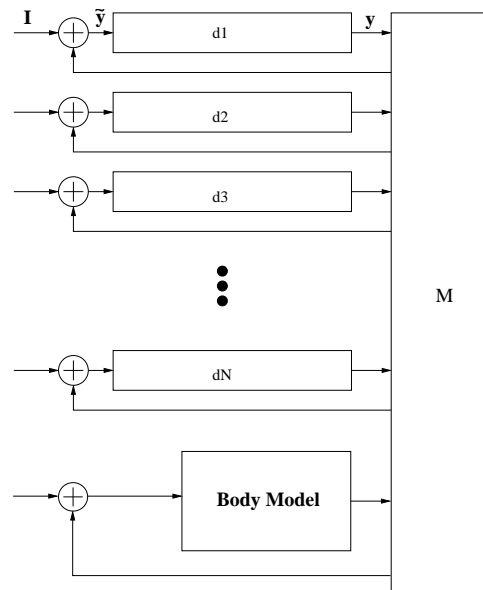
Coupling matrix

Figure 11.8: Adding the body itself as a resonator with full coupling with the strings.
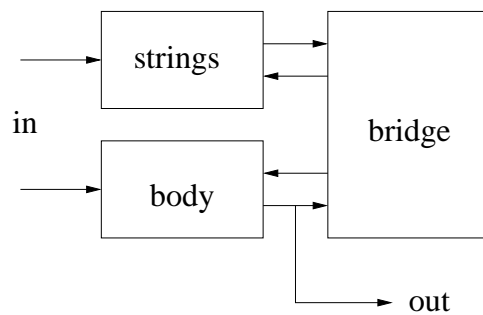

Figure 11.9: The physical structure equivalent to the Full Instrument Coupling model
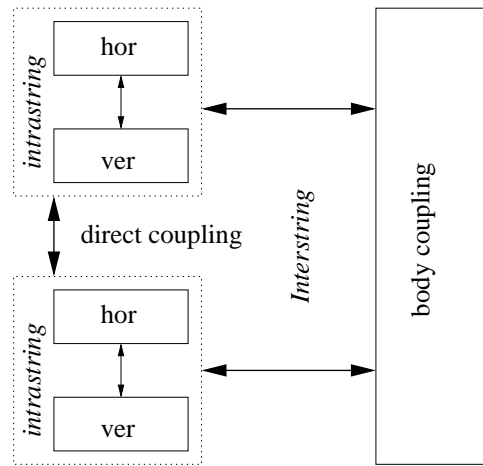
Figure 11.10: The different classes of coupling possible.

It is clear from the bridge coupled model that the coupling matrix does not represent the bridge as such, but the combination of body and bridge (the admittance of the whole system measured at the bridge). If we include the body as an independent resonator, we effectively simplify the coupling matrix as we represent only the bridge. In the acoustic guitar case, the saddle/bridge is a fairly rigid construction, with rather high-frequency vibration modes compared to the body: this will lead to near-frequency independent coefficients. We have an added degree of freedom to describe the coupling: it is possible to couple strings through direct bridge coupling, without influence of the body (see figure 11.10). Grouping the polarizations gives the global structure shown in figure 11.11 The body itself could be implemented as a long FIR filter or as several IIR filters (e.g. Kautz filters, [83]) or a combination of both. We can now build a new waveguide model, by using the complete matrix and by approximating it without sacrificing too much quality.

11.3   Simplified full instrument coupling

A valid approximation to the coupling matrix can be found if we have a better understanding of the coupling mechanism itself [70]. As before, we will model an acoustic nylon string guitar. To analyze the coupling between the strings, we constructed a finite element model of the saddle of the guitar, and calculated the displacement and internal stress for the static load and for unit force in X and Y directions. The Finite Element model was built and calculated using the ALADDIN package [5]. Figures 11.12 and 11.13 show schematic views of
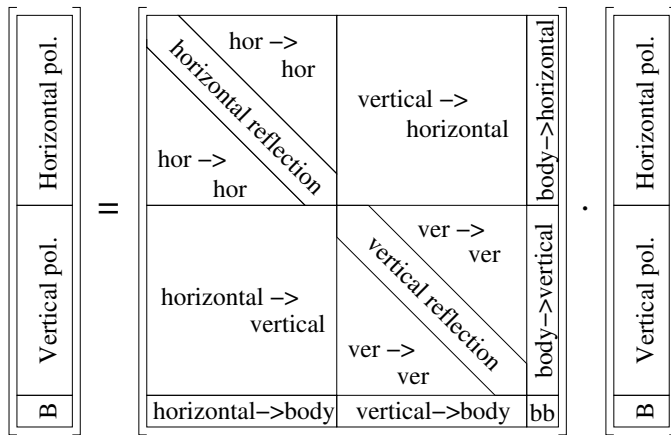
Figure 11.11: The coupling matrix M after inclusion of the body. The direct bridge coupling is separated from the body coupling.
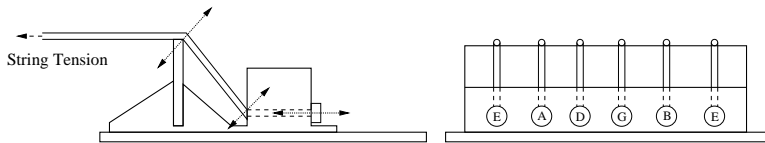
Figure 11.12: Mechanical structure of the bridge/saddle construction of a classical acoustic guitar. The string tension results in steady-state forces in the bridge and the top plate. The vibration of the strings adds a dynamic load.
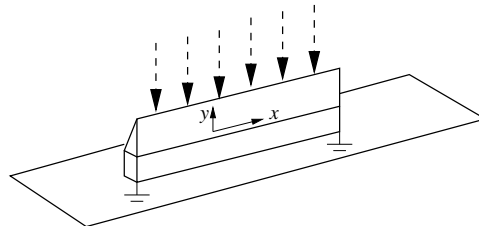
Figure 11.13: Schematic view of the saddle. The two lower corners of the finite element model of the saddle were fixed nodes (allowed no movement in X or Y direction)

Figure 11.14: The static stress for forces in the Y direction. The stress was calculated using a finite element element model with a spatial resolution of 0.1 mm. Darker shades indicate a higher stress.

the saddle/bridge structure. The lower two corners of the saddle model were fixed in X and Y direction (this is needed for a valid FE simulation). This is valid because the top plate is flexible compared to the saddle/bridge (in Y direction) and as rigid as the saddle/bridge in the X direction. We first determined the stresses at rest with six string attached. The forces exerted by the strin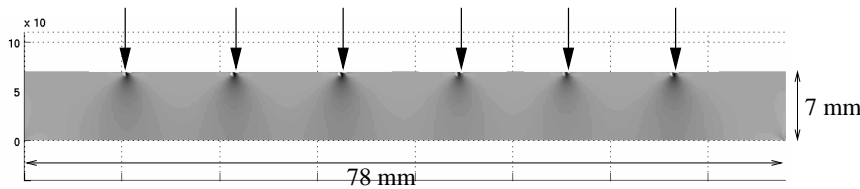gs on the saddle were calculated for d'Addario Pro Arte EJ45 normal tension strings. Figures 11.14 and 11.15 show the calculated internal stresses. The experiment was repeated for a unit force in the X direction (figure 11.17) and in the Y direction (figure 11.16). The results for a load in Y direction (string moving in vertical polarization) show that this load results in stress in both the X and Y direction and that the effect of the load is fairly localized around the load position (load point $x = 33$mm, substantial stress between 26mm $< x <$ 40mm). This means that almost all of the force applied by the string is transmitted to the top plate, without *directly* influencing the other strings. Of course, the displacement of the top plate due to the load will influence *all* the strings. We conclude that interstring vertical-vertical coupling occurs through body coupling. A load in the X direction (string moving in horizontal polarization) gives a different picture. Again, there is stress in both X and Y direction, but the effect is not localized around the load point. The whole bridge sees a stress (and the related displacement). The X component of the stress results in a *direct* coupling in the X direction for all the strings. As the Y component in this case is rather small compared to the Y component when loading in the Y direction, it results in less transmission to the top plate and thus a lower damping. We conclude that interstring horizontal-horizontal coupling occurs through direct coupling and interstring horizontal-vertical coupling through a combination of direct and body coupling.

Intrastring coupling is somewhat more difficult to determine. A load in Y direction results in a symmetric stress pattern in the X direction. A load in X direction creates an asymmetric stress in the Y direction. We believe that this leads to stronger horizontal-vertical intrastring coupling than vertical-horizontal instrastring coupling. These calculations are also valid for the dynamic case with
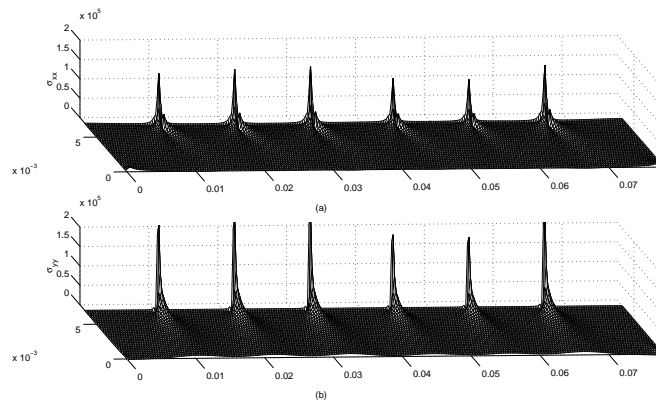
Figure 11.15: The static stresses inside the saddle were obtained by finite element analysis. The forces on the saddle were calculated for the saddle/bridge of a Yamaha C70 acoustic guitar using d'Addario Pro Arte 45 strings under normal tension. The saddle material was bone ($E = 1.7\text{GN/m}^2$). The saddle dimensions are 78mm × 7mm × 2.5mm. (a) σ_{xx} stress in X direction, (b) σ_{yy} stress in Y direction
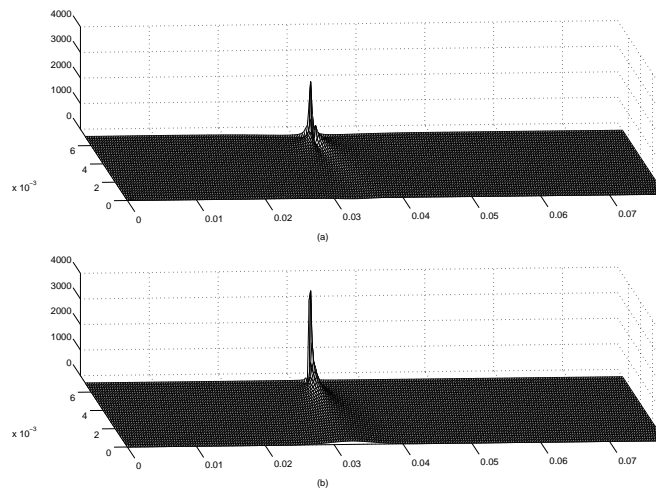
Figure 11.16: The effect of unit force in the Y direction at the third string. The internal stresses are symmetric in the X and Y direction. The influence at $y = 0$ (on top board) is small in X direction and larger in Y direction.
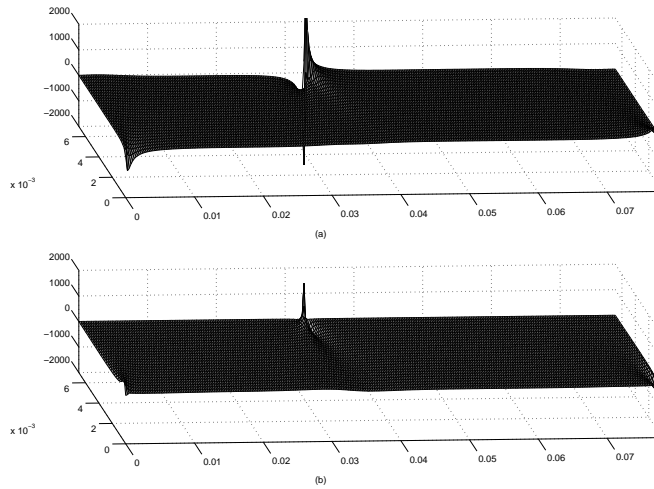
Figure 11.17: The effect of unit force in the X direction at the third string. The internal stresses are asymmetric in the X and Y direction. The influence at $y = 0$ (on top board) is small in X direction and larger in Y direction.

changing load. The time constant to achieve stress equilibrium in the saddle is very small compared to the frequency of the load change (this is valid up to 10 kHz). This is due to the high speed of sound in bone, combined with the small dimensions of the saddle. Including these considerations in the coupling matrix gives the result shown in figure 11.18 or

$$H_{\text{in}} \quad = \quad M1_{6\times6} H_{\text{out}} + I_{\text{hor}}, \tag{11.36}$$

$$V_{\text{in}} \quad = \quad M2_{13\times6} \begin{bmatrix} H_{\text{out},} \\ V_{\text{out}} \\ B_{\text{out}} \end{bmatrix} + I_{\text{ver}}, \tag{11.37}$$

$$B_{\text{in}} \quad = \quad M3_{13\times1} \begin{bmatrix} H_{\text{out},} \\ V_{\text{out}} \\ B_{\text{out}} \end{bmatrix} + I_{\text{body}}, \tag{11.38}$$

where H, V and B are the inputs and outputs of the horizontal polarization, the vertical polarization and the body respectively. The bridge displacements are in phase with the string movement, so $0 < c_{jk} < 1$.

Example 11.2 $\star$ CD Track 24 is the output of a coupled dual polarization model including the guitar body. $\triangle$

11.3.1   Calibration

Figure 11.18: Approximation for the coupling matrix based on the finite element analysis.

We now have to determine the numerical values for the coupling coefficients. The calibration process breaks down into four steps

1. Determine the decay rate for the fundamental of each string, for the two polarizations, with all other strings damped;

2. Determine the body-string coupling by analyzing a recording of an impulse to the body;

3. Determine string-string coupling by exciting and damping one string;
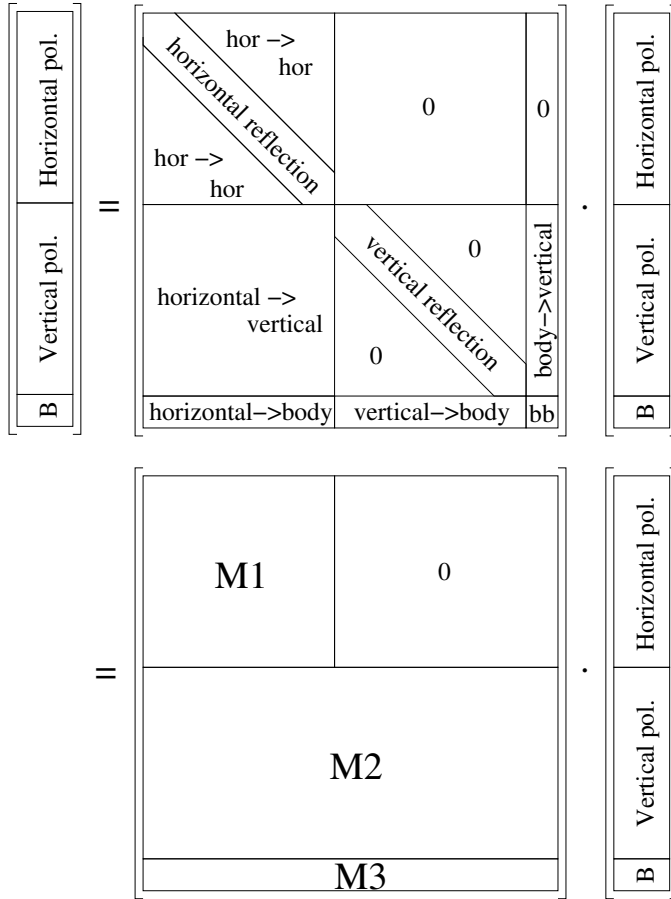
4. Adjust the values found in step 1.

The first step is easy and results in a useful dual-polarization model. The second step requires a recording of an impulse applied to the body while leaving one string undamped and repeating this for each string. The resulting vibration of the string is a measure for the body-string coupling. The third step is very similar, except that we excite and damp one string while leaving one other string undamped. This results in coupling through the bridge and through the body. One way to determine the coefficients is to use an optimization algorithm with a cost function including the predicted output (eigenvalues of the theoretical Ω matrices) and the measured output. Here, the finite element analysis provides a starting point for the values of the coefficients.

The easiest coefficients to determine are the vertical reflection (c_{Vr}) and vertical-to-body ($c_{V \to B}$) coefficients (see figure 11.18). The first step has already determined the value for c_{Vr}. A starting value for $c_{V \to B}$ is

$$\tilde{c}_{V \to B} = 1 - \tilde{c}_{Vr}. \tag{11.39}$$

The horizontal reflection (c_{Hr}) coefficients can be determined with the same procedure. Based on the results for the stress in X direction, we conclude that the horizontal-horizontal ($c_{H \to H}$) coupling depends on the distance of the "source" string to the "target" string. We model this by taking the coefficient (linearly) proportional to the distance. The same holds for $c_{H \to V}$. According to the finite element analysis, the displacement in the horizontal dimension is about three times larger than in the vertical dimension, which leads to the ratio

$$c_{H \to H} = 3 c_{H \to V}. \tag{11.40}$$

The transfer from the horizontal polarization to the instrument body is more difficult to determine. It has been found that

$$c_{H \to B} = 0.1(1 - c_{H \to H}), \tag{11.41}$$

gives acceptable results. We can now scale the $c_{H \to V}$ and $c_{H \to H}$ values such that

$$\sum c_{H \to V} + \sum c_{H \to H} = 0.9(1 - c_{Hr}). \tag{11.42}$$
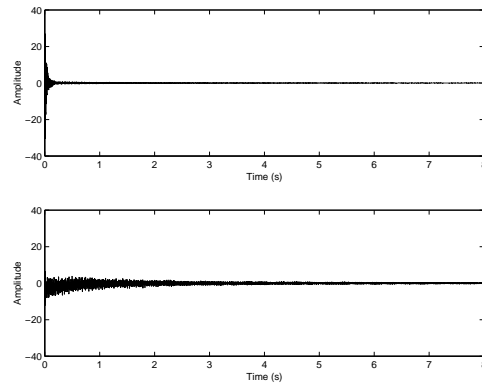
Figure 11.19: Exciting the body part of full model results in the impulse response of the guitar body (top) and the sympathetic vibration of the strings (bottom).

The last coefficients to be determined are $c_{B \to V}$.

Figure 11.19 shows the model output when the body is excited with an impulse.

The last step is to adjust the coupling coefficients such that

$$\sum_k c_{jk} = 1, \tag{11.43}$$

or the column sums of the coupling matrix are one.

Another possible method consists of exciting the body or the strings while leaving all the strings undamped. As coupling with more than one string is now possible, we have to determine the number of strings a specific partial is coupled to. Assuming a linear system, the number of coupled strings depends on the fundamental frequency of each string. Figure 11.20 shows the spectra of two different strings. Only the overlapping parts (the common multiples of the fundamentals) will be modified by the coupling if the system is linear. This property enables us to calculate the number of oscillators in Ω to include for each partial. For the calibration of the instrastring coefficients, we select un-coupled partials. Using the same technique, we can select partials that are coupled to only one other string to calibrate the coefficient for that specific interstring coupling.

Example 11.3 $\star$ CD Track 25 is an example of the complete model, calibrated to sound like the reference note of track 20. Note that the attack is slightly different due to the different excitation. $\triangle$
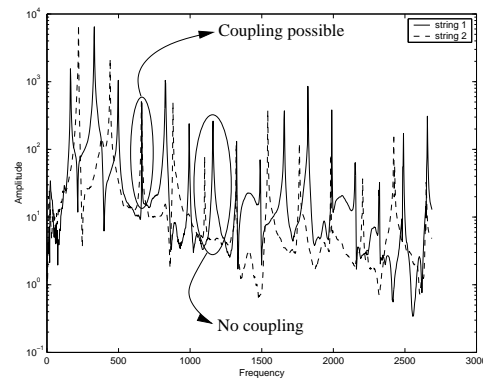
Figure 11.20: The spectra of two different strings show the overlap at the common multiples of the fundamentals. Only the partials will be influenced by the interstring coupling

11.4   Conclusion

In this chapter, the single-string models were extended to multiple-string models, including the very important string coupling mechanism. A mathematical description for the behavior of coupled oscillators was transformed for use with digital waveguide systems. This allowed the comparison of several different coupled string models. Finally, we proposed a more accurate fully coupled instrument model, with its associated parameter estimation routines.

Chapter 12

Guitar Body Models

The guitar body acts as an acoustic amplifier for the vibrations of the strings. It ensures that enough energy is reflected to the strings to sustain the note and that a note of significant volume is audible to the listeners. The shape and materials used are determining factors of the timbre and the spatial sound radiation pattern of the instrument. A guitar body is a complex structure consisting of thin, more or less flexible plates, bracings, and an enclosed air cavity. We first discuss the acoustic properties of a guitar in section 12.1. Several possible linear modeling techniques are then reviewed in section 12.2.1. A non-linear Wiener-Hammerstein model of the guitar body is finally developed in section 12.3.

12.1   Acoustic properties

The frequency response of a guitar body (figure 12.1) shows many resonances and anti-resonances. When mechanically driven at the bridge, the lowest resonance is usually a bar bending mode, but at too low a frequency to be excited by the vibration of real strings [31]. Most guitars have three string resonances in the 100-200 Hz range due to coupling between the fundamental modes of the top and back plates, and the Helmholtz mode of the air cavity. At the lowest of the three resonances, the top and back plates move in opposite directions. The guitar top vibrates in many modes; those of low frequency bear considerable resemblance to those of a rectangular plate when no bracings are present.

The bracings are necessary to add mechanical rigidity to the top plate, and have a strong impact on the vibration modes of the top plate. The higher air cavity modes resemble standing waves in a rectangular box. Disregarding the ribs and the back plate, the top-plate and air cavity can be modeled as two coupled oscillators. Adding the back plate yields a three-oscillator model. In most guitars, the addition of the back results in a downward shift of the two primary resonances [18], [61], [95]. The higher two resonances usually occur around 200 Hz, depending upon the stiffness of the the top and back plates. The motion of the air and the top plate is in the same direction, thus resulting in strong sound radiation. The resonances of the top plate, back plate, and air cavity generally combine to give at least one strong resonance around 300 Hz in a classical guitar, but closer to 400 Hz in a cross-braced folk guitar. Above 400 Hz, the coupling between top and back plate appears to be relatively weak, and the observed resonances are due to resonances in one of the plates. There is a definite link between the subjective quality of a guitar and its frequency response [62], and the knowledge of the acoustical properties of the guitar body leads to improved guitar designs.

12.2   Linear models

For sufficiently small input signals, the body of a guitar is a linear system, and can be modeled using all the conventional methods. The impulse response of a guitar body excited at the bridge shows a large number of poles. It has been shown that more than 300 poles are needed to obtain a model that is indistinguishable from the original by the listener [83]. Figure 12.3 shows the singular values of the Hankel matrix built from the impulse response data shown in figure 12.2. The order needed to accurately model the body is approximately 300, which incidentally means that the human ear is quite sensitive.

12.2.1   FIR

The easiest way of modeling the guitar body is by using the complete impulse response of the body as an FIR filter. The basis functions used in FIR filters are the orthonormal functions z^{-k}. A FIR model of a system $G(z) \in H_2$ (H_2 denoting the space of stable, strictly-proper transfer functions) consists of a finite number of expansion terms and takes on the form

$$\hat{G}(z) = \sum_{k=1}^{N} g(k) z^{-k}, \tag{12.1}$$

where $g(k)$ are impulse response coefficients. If the decay rate of the signal is low compared to the sampling rate, very high order FIR filters are needed.
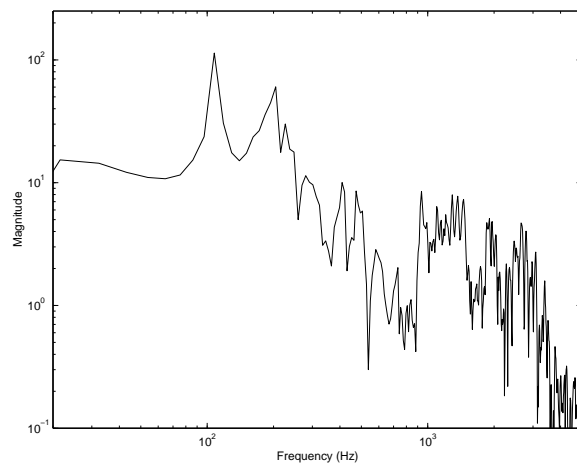
Figure 12.1: Spectrum of the impulse response of a Yamaha C70 classical guitar.
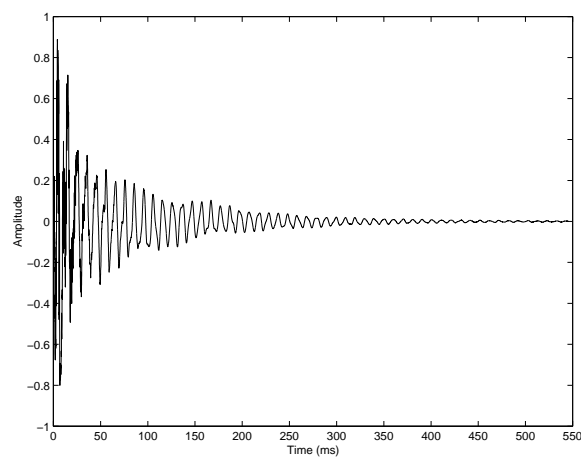
Figure 12.2: Impulse response of a Yamaha C70 classical guitar. The impulse was applied at the bridge. The response was recorded in front of the soundhole, at a distance of 30cm, using an AKG C4000B microphone. The signal was fed to a Tascam DAP-1 DAT recorder and sampled at 16 bit, 44.1 kHz.
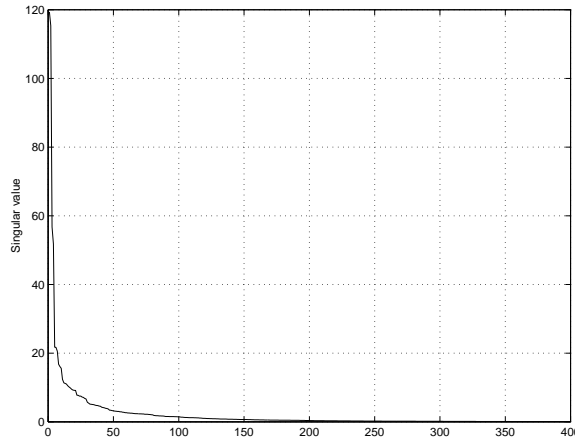
Figure 12.3: The singular values of the Hankel matrix built from the impulse response data. There are about 350 non-zero singular values. This is a strong indication that the model order is less or equal to 350.

The computational cost is quite high, especially when using a sample-oriented method. When block-processing is used, several fast-convolution frequency domain techniques can be used to speed up the computation. The long FIR size of several tens of thousands of samples is the main bottleneck of using this method in the fully coupled model, as this model does not easily allow block processing.

12.2.2   IIR+FIR

The size of the FIR filter can be significantly reduced by using IIR resonators to model the first few body resonances, and using the FIR filter only for the residual signal. The first resonance is the Helmholtz resonance of the air cavity. The impedance of a Helmholtz resonator has been discussed in section 4.3.3 and is given by

$$Z_{\text{Helmholtz}}(s) = \left(\frac{\rho L}{S}\right) s + \left(\frac{\rho c^2}{Q}\right) \frac{1}{s}. \tag{12.2}$$

The resonance peaks occur where the acoustic impedance is zero. This gives the following transfer function for the Helmholtz resonance:

$$\mathcal{H}_{\text{Helmholtz}}(s) = \frac{S}{\rho L} \frac{s}{s^2 + c^2 S/LQ}. \tag{12.3}$$

This is a standard second-order tranfert function, and can be described in discrete time by using several transforms. For frequencies much lower than

Nyquist, the backwards Euler transform is usable. This transform has the disadvantage that it tends lower the sharpness of resonant peaks. Therefore, the bilinear transform is better suited. A parametric representation for a second-order peak filter is given as

$$\mathcal{H}_{\text{peak}} = (1-b)\frac{1-z^{-2}}{1-2b\cos(\omega_0)z^{-1}+(2b-1)z^{-2}}, \tag{12.4}$$

where

$$\omega_0 = \frac{2\pi f_0}{f_s}, \tag{12.5}$$

$$\Delta\omega = \frac{2\pi \Delta f}{f_s}, \tag{12.6}$$

$$b = \frac{1}{1+\tan(\Delta\omega/2)}, \tag{12.7}$$

with Δf the 3dB bandwidth of the peak. The peak frequency is best found in the frequency domain, by quadratic interpolation of the peak in the spectrum, as was discussed in section 10.5. In this case, the damping is quite high, so the peaks are wider, making the manual determination of the 3dB bandwidth a possibility. Methods using the STFT are less interesting, exactly because of the higher damping, making the signals too short in time for an accurate measurement of the exponential slope. A better approach in this case is to use an optimization technique to minimize the cost function

$$\mathcal{E} = \sum W(\mathcal{S} - \mathcal{H}_{\text{peak}})^2, \tag{12.8}$$

where $\mathcal{S}$ is the spectrum of the body impulse response, and W a windowing function around the peak. The residual signal is calculated by either a time-domain subtraction of the output of the IIR resonator, or by filtering the body impulse response with a notch filter. The procedure is then repeated for the other modes, until the residual signal is short enough to be used as a FIR filter.

Example 12.1 The lowest resonance for the body with the spectrum shown in figure 12.1 occurs at 94.99 Hz, with a 3dB bandwidth of 4.41 Hz. This was determined using cost function (12.8), and a Gaussian window W around the first resonance. $\triangle$

In general, the guitar body is a linear time invariant model, and could be described using the standard LTI model structures. The complete model structure is given by

$$A(z)y_k = \frac{B(z)}{F(z)}u_k + \frac{C(z)}{D(z)}e_k, \tag{12.9}$$

where e_k is Gaussian white noise. We assume that the recording of the body impulse only adds white noise to the output and obtain the output error (OE)

model structure

$$y_k = \frac{B(z)}{F(z)} u_k + e_k. \tag{12.10}$$

As was determined before, the order of $B(z)/F(z)$ is approximately 350. The model parameters can be found with either prediction error methods or correlation methods.

Prediction error methods are based on the minimization of the prediction error sequence $\hat{e}_k$

$$\hat{e}_k = y_k - G(z,\theta) u_k, \tag{12.11}$$

with $G(z,\theta) = B(z)/F(z)$. This is achieved by optimizing the parameter vectorθ with as a cost function

$$V(\theta) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{2} (\hat{e}_k^f(\theta))^2, \tag{12.12}$$

where $\hat{e}_k^f(\theta)$ is the error sequence, filtered to stress a certain part of the spectrum. The frequency domain interpretation of the prediction error method shows that this is just a generalization of the cost function (12.8) or

$$V(\theta) = \int_{-\pi}^{\pi} \frac{1}{2} |G_0(e^{j\omega}) - G(e^{j\omega},\theta)|^2 |L(e^{j\omega})|^2 |U(\omega)|^2 d\omega, \tag{12.13}$$

with $G_0(z)$ the real transfer function, $G(z,\theta)$ the model, $U(\omega)$ the spectrum of the input, and $L(z)$ the filter used to filter the prediction error. For our purposes, instead of determining one high-order model, we calculate a set of low-order models. Each low order model is obtained by band-pass filtering the prediction error around the previously selected peaks in the spectrum. If the input is a true impulse, this method simplifies to equation (12.8). When calculating a second order model for each subband, we obtain

$$y_k = \sum_{i=1}^{N} \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{a_0 + a_1 z^{-1} + a_2 z^{-2}} u_k + e_k, \tag{12.14}$$

with N the total number of bands considered. This procedure results in a parallel set of IIR filters.

The practical determination using these conventional system identification methods require long, persistent excitation signals. Therefore, a synthetic output is generated by filtering white noise with the complete impulse response data as FIR coefficients. The resulting longer input-output signals are used during identification.

Alternatively, one could use the sHSVD algorithm to determine the poles directly, and to design a filterbank with the same poles. This is quite similar to the determination of a state space model.

12.2.3 Kautz filters

Kautz filters [45], [116], [115], [37] are fixed-pole IIR filters structurally organized to produce orthonormal tap-output impulse responses. The transversal Kautz filter can be seen as a generalization of FIR and Laguerre filter structures, providing IIR-like parametric modeling combined with the superior numerical properties of FIR filters [80], [83].

Taking into account the high FIR filter order needed for accurate modeling, one could use a different orthonormal basis of H_2. Let the functions $F_k(z)$ with $k \in \mathbb{N}$ denote the basis elements of such a general basis. Then the transfer function $G(z) \in H_2$ can be expanded as

$$G(z) = \sum_{k=1}^{\infty} c_k F_k(z), \tag{12.15}$$

The aim is to choose the basis $\{F_k(z)\}_{k \in \mathbb{Z}}$ such that the expansion coefficients c_k rapidly converge to zero. A straightforward approach to this problem is to ortho-normalize the set of functions

$$f_{i,j}(z) = \frac{1}{(z - a_i)^j}, \tag{12.16}$$

where the poles a_i can in principle be any complex number with modulus smaller than one, and $i \in \mathbb{N}, 1 \le j \le m_i$. Applying a Gram-Schmidt procedure to the sequence of functions $f_{i,j}(z)$ yields the orthonormal functions

$$\phi_k(z) = \frac{\sqrt{1 - |\xi_k|^2}}{z - \xi_k} \prod_{i=1}^{k-1} \frac{1 - \bar{\xi}_i z}{z - \xi_i}, \tag{12.17}$$

with $k \in (N)$ [117]. The set of orthonormal functions $\{\phi_k(z)\}_{k \in \mathbb{Z}}$ is called the Takenaka-Malmquist construction. Kautz proposed a method of rational orthonormal basis function construction that is very close to the Takenaka-Malmquist construction, but where the functions are constrained to be real-rational functions [45]. The general discrete-time Kautz functions are defined as

$$\phi_{2k-1}(z) = \frac{C_{k,1}(1 - \alpha_{k,1} z)}{(z - \beta_k)(z - \bar{\beta}_k)} \prod_{i=1}^{k-1} \frac{(1 - \bar{\beta}_i z)(1 - \beta_i z)}{(z - \beta_i)(z - \bar{\beta}_i)}, \tag{12.18}$$

$$\phi_{2k}(z) = \frac{C_{k,2}(1 - \alpha_{k,2} z)}{(z - \beta_k)(z - \bar{\beta}_k)} \prod_{i=1}^{k-1} \frac{(1 - \bar{\beta}_i z)(1 - \beta_i z)}{(z - \beta_i)(z - \bar{\beta}_i)}, \tag{12.19}$$

with $k \in \mathbb{N}$. The parameter $\beta_k \in \mathbb{C}$ are to be chosen such that $|\beta_k| < 1$ and the parameters $\alpha_{k,1}, \alpha_{k,2} \in \mathbb{R}$ must satisfy

$$(1 + \alpha_{k,1}\alpha_{k,2})\left(1 + |\beta_k|^2\right) - (\alpha_{k,1} + \alpha_{k,2})(\beta_k + \bar{\beta}_k) = 0. \tag{12.20}$$

The constants $C_{k,1}$ and $C_{k,2}$ are given by

$$C_{k,1} = \sqrt{\frac{(1 - \beta_k^2)(1 - \hat{\beta}_k^2)(1 - \beta_k \hat{\beta}_k)}{(1 + \alpha_{k,1}^2)(1 + |\beta_k|^2) - 2\alpha_{k,1}(\beta_k + \bar{\beta}_k)}}, \qquad (12.21)$$

$$C_{k,2} = \sqrt{\frac{(1 - \beta_k^2)(1 - \hat{\beta}_k^2)(1 - \beta_k \hat{\beta}_k)}{(1 + \alpha_{k,2}^2)(1 + |\beta_k|^2) - 2\alpha_{k,2}(\beta_k + \bar{\beta}_k)}}. \qquad (12.22)$$

By the nature of the Kautz construction, poles can only occur in complex conjugate pairs. In general the complexity of the computations and the analysis involved in using rational bases increases as the number of different poles in the basis increases. The best-known simplified bases that are used for filtering purposes are the Laguerre and the two-parameter Kautz basis. The Laguerre basis is obtained from the Takekana-Malmquist construction, using only one real-valued pole in the construction, i.e. $\xi_k = \xi$ with $-1 < \xi < 1$. This makes the Laguerre basis less suitable for modeling systems that have complex conjugate pole pairs in their transfer functions, as it is the case with the guitar body. The two-parameter Kautz constructions is a simplification of the general Kautz construction where $\beta_k = \beta$ and $\alpha_{k,1} = \alpha_1$, $\alpha_{k,2} = \alpha_2$. The set is given by

$$\phi_{2k-1} = \frac{\sqrt{1 - c^2}(z - b)}{z^2 + b(c - 1)z - c} \prod_{i=1}^{k-1} \frac{-cz^2 + b(c - 1)z + 1}{z^2 + b(c - 1)z - c}, \qquad (12.23)$$

$$\phi_{2k} = \frac{\sqrt{(1 - c^2)(1 - b^2)}}{z^2 + b(c - 1)z - c} \prod_{i=1}^{k-1} \frac{-cz^2 + b(c - 1)z + 1}{z^2 + b(c - 1)z - c}, \qquad (12.24)$$

with the parameters α_1 and α_2 are chosen to be

$$\alpha_1 = \frac{1 + |\beta|^2}{\beta + \bar{\beta}}, \qquad (12.25)$$

$$\alpha_2 = 0, \qquad (12.26)$$

and the parameters b and c are given by

$$b = \frac{\beta + \bar{\beta}}{1 + |\beta|^2}, \qquad (12.27)$$

$$c = -|\beta|^2. \qquad (12.28)$$

In the case of a guitar body model, we know the dominant poles, and can use the full Kautz functions. A structure based on the Kautz basis functions that is well-suited for implementation as a transversal-like filter is

$$\phi_k(z) = a_k \frac{z\sqrt{1 - |\beta_k|^2}}{1 - \beta_i z} \prod_{i=0}^{k} \frac{1 - \bar{\beta}_i z}{z - \beta_i}. \qquad (12.29)$$
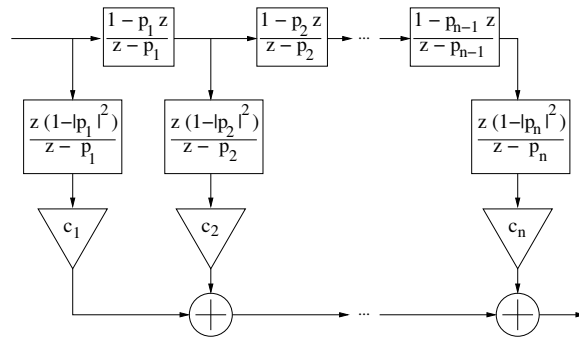
Figure 12.4: The general Kautz filter with complex poles p_i consists of an allpass backbone and all-pole tap-output filters. Both the FIR and Laguerre filters can be implemented with this structure. In the Laquerre case, $p_i \in \mathbb{R}$, $-1 < p_i < 1$, and for the FIR case $p_i = 0$.

This set of basis functions leads to the complex Kautz filter structure shown in figure 12.4 [79]. The real Kautz filter structure that is derived from equations (12.18) and (12.19) is shown in figure 12.5. The tap output coefficients c_k can easily be found by projection on the basis functions or

$$c_k = \langle G(z), \phi_k(z, \beta) \rangle = \frac{1}{2\pi j} \oint_\Gamma G(z) \phi_k(z^{-}1, \beta) z^{-1} dz, \qquad (12.30)$$

with Γ a circle of radius higher than 1 and lower than $1/|\beta|$. In practice, the coefficients are found by feeding the time-inverted impulse response of the guitar body $h(-n)$ to the Kautz filter and reading the tap outputs $x_i(n) = G_i[h(-n)]$ at $n = 0$ such that $c_i = x_i(0)$.

For an efficient model, a good set of poles should be selected. The determination of the poles is however not a trivial problem, and several design procedures can be found in the literature [15]. The lowest-order models are obtained when the basis function poles correspond to the system poles, which can be determined by the methods described in section 12.2.2. This procedure results in a transversal set of IIR filters.

12.3   Non-linear models

The body of a guitar is only linear for small excitations. For larger amplitudes, the wooden top plate tends to saturate the output due to its limited flexibility and bracings. The effect was measured using series of impulses with increasing amplitude [76].
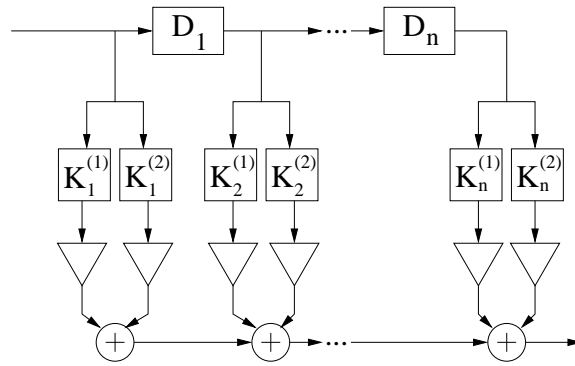
Figure 12.5: The real Kautz filter consists of cascaded delay blocks and dual-tap outputs according to equations (12.18) and (12.19). The tap output weights are computed by projection on the basis functions.

Wiener-Hammerstein systems consist of two linear blocks and a static non-linearity, show in figure 12.6. We assume that in the case of a guitar body, the input is saturated, yielding a Wiener system. The output of the Wiener-Hammerstein is

$$
\begin{align}
x_2(t) &= g(x_1(t)), \tag{12.31}\\
U(s) &= H_1(s)X_2(s), \tag{12.32}
\end{align}
$$

The static non-linearity is described by a polynomial function

$$
g(x) = \sum_{k=1}^{N} b_k x^k. \tag{12.33}
$$

The experimental determination of the coefficients is possible by recording a series of impulses with increasing amplitude. If the system is linear, amplification of a low-amplitude impulse response will approximately yield the high-amplitude impulse response. Figure 12.7 shows a polynomial approximation of the non-linearity. The dots on the figure indicate the input-output relation between the amplified low-amplitude impulse response, and the high-amplitude
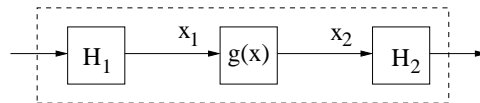
Figure 12.6: A complete Wiener-Hammerstein system with two linear blocks and a static non-linearity.
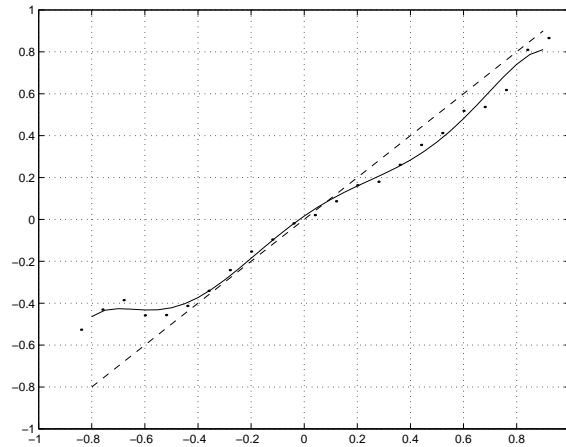
Figure 12.7: This figure shows the static non-linearity added to the body model. The dashed line shows a linear input-output relation. The dots are the measure input-output relations for a given amplitude, and the solid line is a polynomial fit. The compression effect is visible for the negative input values. This means that above a specific playing level, the output amplitude will only slightly increase, and the sound will be distorted. This effect depends on the type of wood used and is well known by guitar players.

impulse response. First, two sequences of impulses of the same amplitude were recorded and the expected responses were calculated. The expected values are then placed in amplitude bins, and the input-output relation is determined. Each dot on the figure represents thousands of samples for the lower amplitudes to about hundred samples at high amplitude. We observe a saturation characteristic at higher amplitudes. Further experiments are needed to determine the cause of this behavior.

Example 12.2 ⋆ CD Track 29 explicitely uses the guitar body. The fully coupled model, with the guitar body simulated as an IIR+FIR model, is excited by an impulse on the body model. The body impulse is audible, as are the sympathetic vibrations of the strings. △

12.4 Conclusion

In this chapter, several different guitar body modeling techniques have been discussed. All the linear techniques are equivalent with respect to their output, but they differ in computational complexity and sensitivity to errors. Long

FIR models perfectly capture the properties of the body, but due to their long length, are not well suited for synthesis. The FIR filters can be simplified by combining IIR models for the dominant poles with a residual FIR filter, yielding a set of parallel filters. A different approach is to use the Kautz filter structure, where an efficient, low-order FIR-like structure is obtained. Finally, it has been determined that a guitar body is better modeled with the addition of a static non-linearity that simulates the saturation of this mechanical amplifier. The body models are an important part of the complete calibrated models used in the next chapter.

Chapter 13

Validation

In this chapter, we describe the experimental determination of the coupling parameters for two coupled string waveguide models. First, the measurement setup is described in section 13.1, and all the string parameters are estimated in section 13.2. The coupling parameters are determined for the simplified coupled string structure and the fully coupled structures of the previous chapter [69], [72].

13.1 Measurement setup

All measurements were done on a Taylor 514CE steel-string acoustic guitar. This guitar has a built-in Fishman Stereo Blender with a Fishman Acoustic Matrix undersaddle pickup and an internal electret microphone. We added a modified Roland GK-2A hexaphonic magnetic pickup at 1 cm of the saddle, not attached to the guitar but independently mounted on a rigid construction. The 8-channel sensor outputs were recorded on a PC equipped with 4 synchronised LynxOne soundcards, at 24-bit resolution and a 44.1 kHz sampling rate. The recording setup is schematically shown in figure 13.1 and figure 13.2 is a photograph of the actual setup. The recorded signals were imported in Matlab (Mathworks Inc.) for further signal processing.

Figure 13.3 shows the outputs of the magnetic pickup, undersaddle pickup and built-in electret microphone for the high E string. The envelopes are quite
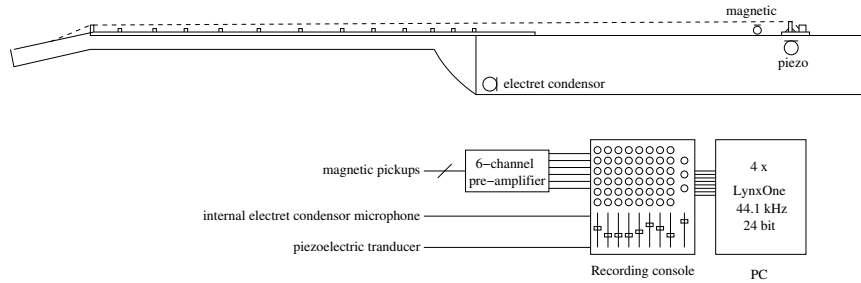
Figure 13.1: Schematic overview of the measurement setup. The guitar is equipped with a Fsihman Acoustic Matrix undersaddle pickup and an internal electret microphone. These signals are preamplified by a Fishman Stereo Blender built into the guitar. The movement of each string is separately measured with a modified Roland GK-2A hexaphonic magnetic pickup, amplified through a custom-built six-channel guitar preamplifier. All signal go through a recording console for level adjustment. The signals are sampled using four synchronized LynxOne soundcards, at 24 bit, 44.1 kHz.

Figure 13.2: Photograph of the measurement setup.

Figure 13.3: Amplitude vs time plots of the recording for the high E string (a) magnetic pickup, (b) undersaddle pickup, (c) electret microphone

similar. When looking at at the spectra of these signals (figure 13.4), the difference is quite clear. The output of the magnetic pickup clearly shows the comb-filter effect of the plucking point, while the internal microphone signal shows the influence of the instrument's body.

The magnetic pickups are inherently non-linear. Their principle of operation is the modification of the magnetic flux through a coil by moving a ferromagnetic string in the field of a permanent magnet. A rough calculation yields that the output of the pickup $V \sim y^{-3}$, with y the pickup-string distance. This gives the trade-off (for the same magnet): larger distance, smaller distortion (relative to signal strength) but weaker signal (lower signal-to noise ratio). At the bridge level, the string displacement is relatively small, so we disregard this non-linear distortion. The undersaddle pickup is an electret-film transducer, which could be seen as a capacitor with variable interplate distance. Its output is proportional to the differential movement of the saddle and the top plate of the guitar. Figure 13.5 shows the magnitude response of the magnetic pickup signal and the undersaddle pickup over a wider frequency range. The response of the undersaddle pickup shows the characteristic double peak of the body resonances.

Finally, figure 13.6 shows the amplitude spectrum of the estimated transfer functions from magnetic pickup to undersaddle pickup to microphone. The
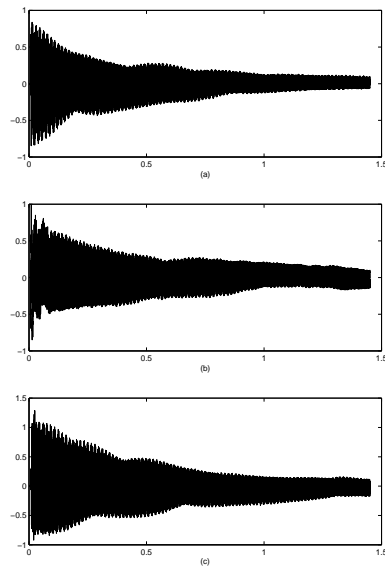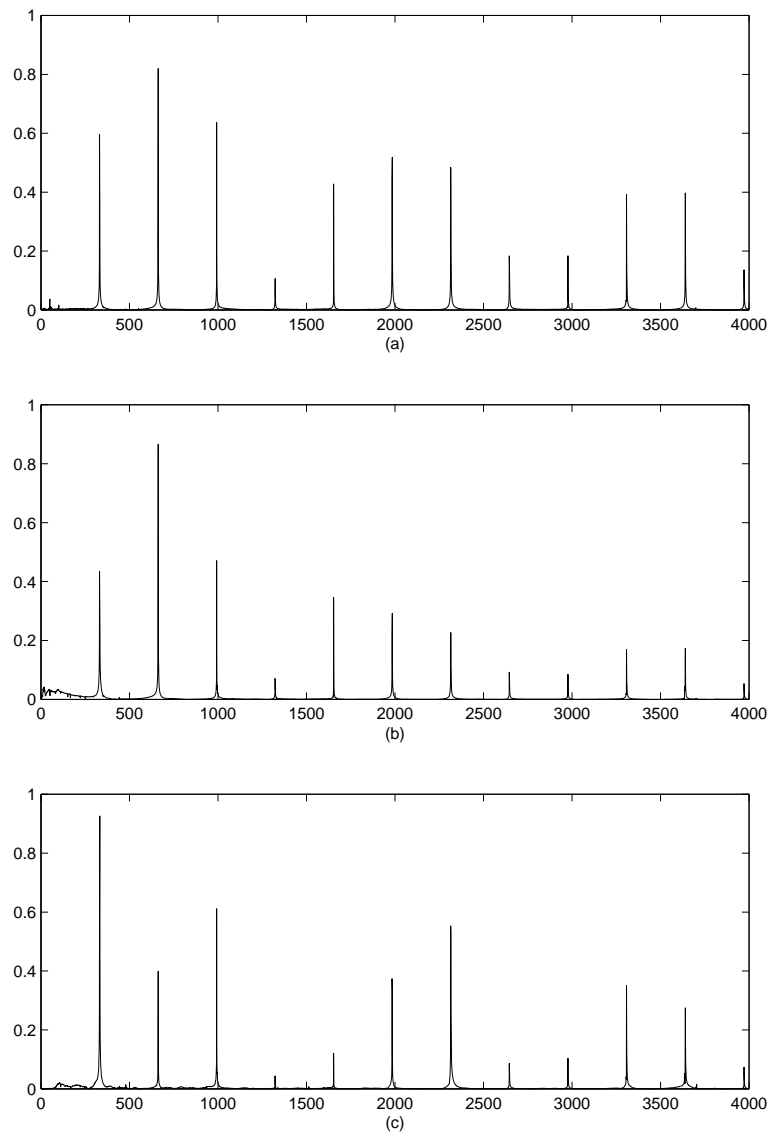
Figure 13.4: Spectrum of the recording for the high E string (a) magnetic
pickup, (b) undersaddle pickup, (c) electret microphone
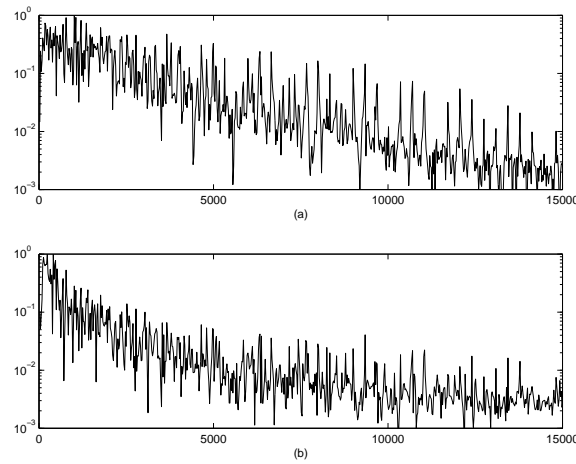
Figure 13.5: Amplitude vs Frequency plot of (a) the magnetic pickup signal and (b) the undersaddle pickup

transfer function from undersaddle pickup to microphone is essentially flat, while the transfer function from magnetic pickup to undersaddle pickup closely resembles the body impulse response. The transfer functions were calculated using Welch's averaged periodogram method.

13.2   Parameter estimation

Figure 13.7 shows a part of the recorded signals. The figure shows the outputs of the hexaphonic pickup, when the high E string is plucked. We see light coupling to the D string and stronger coupling to the A and low E string. Examination of the spectrum of the signals shows that coupling only occurs for matching modes. This confirms the near-linearity of the guitar body and string coupling mechanism discussed in chapter 11 . In table 13.1, we indicate for each string which harmonic corresponds to a harmonic of the high E string. Based on this table, we expect strong coupling between the high E and the A and low E strings, and only slight coupling to the other strings. This is exactly as seen in figure 13.7.

To simplify things, we assume that only two strings interact. The procedure is however applicable to N coupled strings. First, we determine the parameters of the model shown in figure 13.9 and discussed in section 11.2.4. This is the simplified coupling structure with two strings and two polarizations per string.
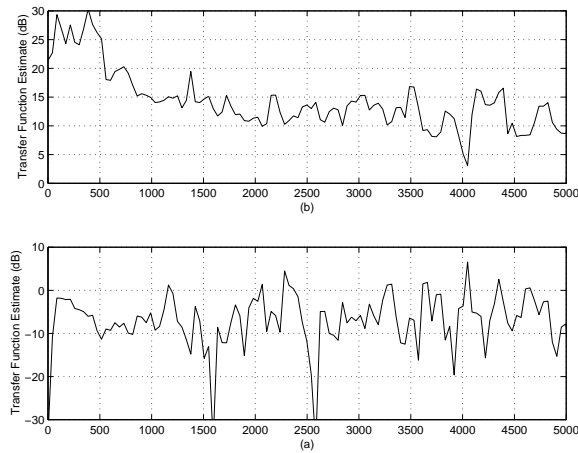
Figure 13.6: Amplitude response of the estimated transfer function from (a) the magnetic pickup to undersaddle pickup and (b) the undersaddle pickup to electret microphone

high E	329.6 Hz	fund	2nd	3rd	4th
B	246.9 Hz			4th	
G	195.9 Hz			5th	
D	146.8 Hz				9th
A	110.0 Hz	3rd	4th	9th	12th
low E	82.4 Hz	4th	8th	12th	16th

Table 13.1: This table indicates which harmonics of the non-excited strings correspond to within a few Hz to the harmonics of the high E strings. Based on this table, we expect strong coupling to the A and low E strings, and weak coupling to the other strings
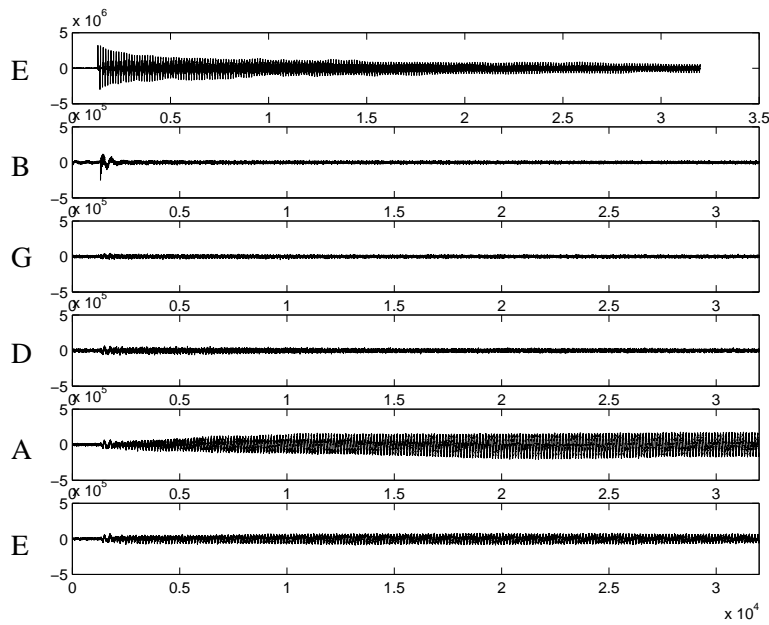
Figure 13.7: This figure shows the output of the hexaphonic magnetic pickup. From top to bottom: high E, B, G, D, A, low E. The high E string was plucked. Note that the scale for the high E string is ten times larger.

The input of the digital waveguides is characterized by

$$X = I + \begin{bmatrix} R_{11} & 0 & 0 & 0 \\ 0 & R_{12} & c_{21} & 0 \\ 0 & 0 & R_{21} & 0 \\ c_{12} & 0 & 0 & R_{22} \end{bmatrix} Y, \qquad (13.1)$$

where Y denotes the output of the waveguides and I the excitation input, R_{ij} are the reflection coefficients and c_{ij} the coupling parameters. This system consists of a pair of digital waveguides driving a second pair of waveguides. It is always stable if all $R_{ij} < 1$. Several ways are possible to determine the model parameters, as described in chapter 10 [99], [52], [71]. One could use the analytical solution for two coupled oscillators, and determine the pair-wise coupling parameters, or use a non-linear optimization algorithm [73]. We need two sets of measurements: one with the high E string excited and the low E string at rest, and one were the low E string is excited. The coupling parameters in both directions should be concurrently determined. Using non-linear optimisation, we obtain the set of parameters shown in table 13.2 (for delay line lengths of 533.23 for the low E and 133.22 for the high E string respectively). We obtain one lightly and one strongly damped polarization,
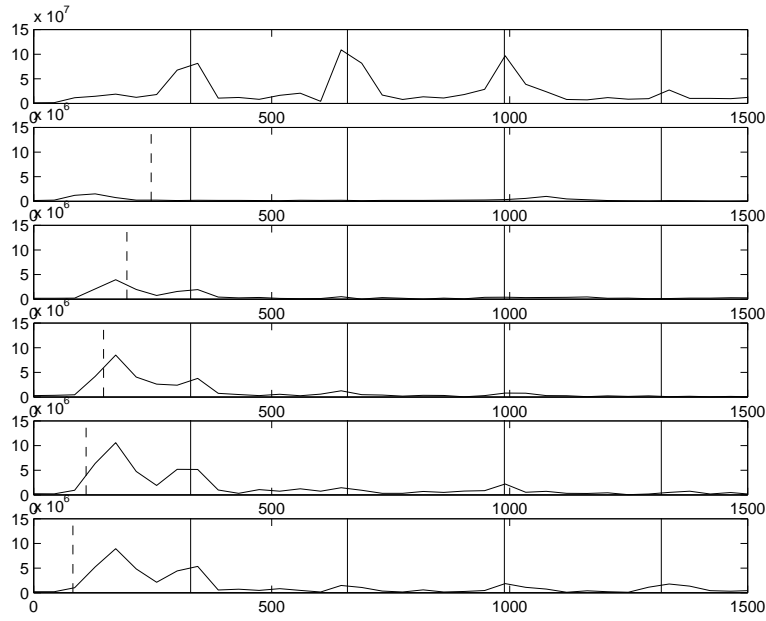
Figure 13.8: This figure shows a part of the spectrum of the signals shown in Figure 13.7. The full lines are the fundamental and the first three harmonics of the high E string. The dashed lines indicate the fundamentals of the other strings. We see coupling to matching modes only. Note that the scale for the high E string is different

and quite similar coupling coefficients c_{12} and c_{21} between high E and low E strings. Note that $R_{11} + c_{12} \approx R_{21} + c_{21} \approx 1$. Figure 13.11 shows the recorded and simulated amplitude envelopes for a common harmonic of the high and low

	Simplified Coupling	Full Coupling
R_{11}	0.9969	0.9969
R_{12}	0.9142	0.9642
R_{21}	0.9896	0.9896
R_{22}	0.9423	0.9823
c_{12}	0.0188	
c_{21}	0.0204	

Table 13.2: The parameters determined for the two models. For the simplified coupling case, we obtain one strongly and one lightly damped polarization, while in the full coupling case, both polarizations are lightly damped.
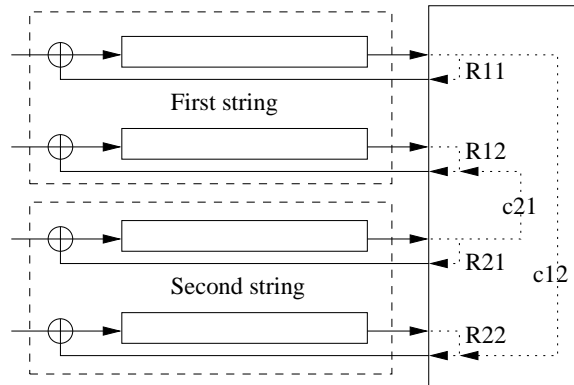
Figure 13.9: This simplified coupled string structure has two polarizations per string. The first polarization of the first string drives the second polarization of the second string, with no feedback.
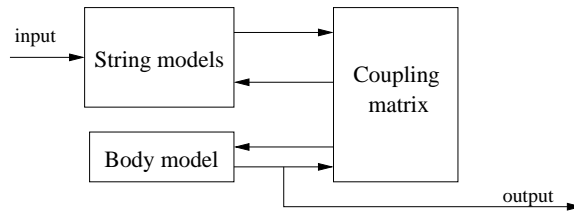
Figure 13.10: The fully coupled string model has twelve separate string models (two for each string) and a body model. All the outputs are mixed using a 13×13 coupling matrix and used as inputs.
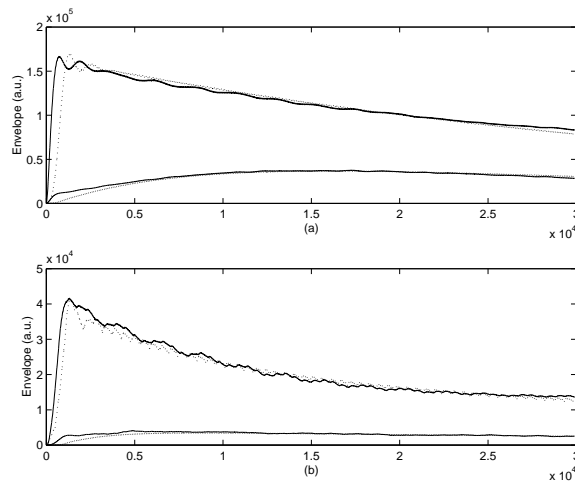
Figure 13.11: This figure shows the amplitude envelope of (full line) the recording of magnetic pickup, and (dotted line) the actual model output. Shown are (a) high E string excited (top pair), coupling with low E string (bottom pair); and (b) low E excited, coupling with high E

E strings. The fully coupled model is more difficult to calibrate. In this model, we have twelve string models and one body model, all coupled using a 13×13 coupling matrix. As there is feedback, care must be taken to ensure that the global model is stable. The heuristic rules as described in section 11.3 give a good starting point for further optimisation [73]. When using these rules, one only determines the decay rate of the fundamentals for the two polarizations of each string. All other model parameters follow from basic principles. Exactly calibrating the parameters for this model requires non-linear optimisation, as the system stability has to be ensured. Slightly altering the feedback from the body to the strings yields the output shown in figure 13.12. This is a close match to the recording. Note that the two-stage decay now results from true coupling, as the two polarizations have comparable reflection coefficients.

Example 13.1 $\star$ `CD Track 28` is the output of a fully coupled guitar model, where the coupling parameters were set such that the system was at the limit of stability. An arpegiated chord was played on the six strings. $\triangle$
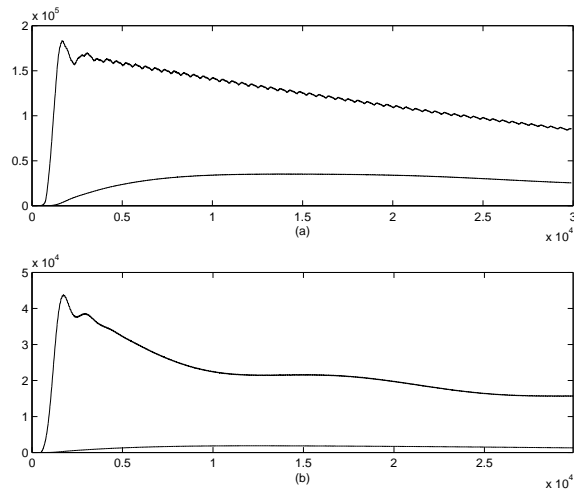
Figure 13.12: Amplitude envelopes for the same two cases as figure 13.11, when using the fully coupled model and after having slightly altered the parameters determined using the heuristic rules.

13.3 Conclusion

In this chapter, a practical measurement setup was described that allows the determination of all the digital waveguide model parameters. We determined the coupling parameters for the simplified and the fully coupled structures. Both models are capable of closely simulating the behavior of the real instrument, where the fully coupled model holds the advantage that two-stage decay and beating are included in the model.

Chapter 14

Implementation

```
for i=1:100
{
   ...
}
```

In this chapter, the implementation issues specific to string instruments are reviewed. After introducing the global structure of the implementation in section 14.1, we focus on an MPEG-4 SAOL implementation. Section 14.2 discusses the core DSP routines, and in section 14.3, the score interpreter and the player model are introduced.

14.1   Structure

The implementation of the guitar model consists of multiple layers, show in figure 14.1. The two main layers are the core signal processing routine `stringcore` and the score interpreter. The score interpreter can be subdivided in two subparts, depending on the score language used: one that calculates the correct parameters to control the digital waveguide model, and one that models the behavior of a player. This multi-layered implementation structure is independent of the actual programming language used.

14.2   `stringcore` implementation

The core of the implementation is based on the digital waveguide models proposed in chapter 11. It consists of the twelve digital waveguides, the body
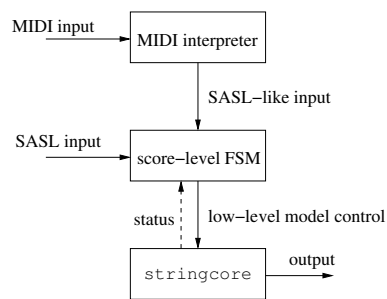
Figure 14.1: System overview of the complete guitar model. The single channel of multichannel MIDI input is interpreted by the MIDI interpreter, which generates SASL-level score information. The score-level Finite State Machine translates the requests to play or end a note into a series of low-level control commands for the `stringcore` datapath-like opcode.

model and the coupling matrix, as shown in figure 14.2. The `stringcore` routine is the lowest computation level of the model. It represents the physical behavior of a real instrument and cannot be modified, only controlled through a restricted number of parameters. `stringcore` provides the higher-level routines with status information.

14.2.1   Sample based processing

The simplest implementation is sample-based, like the implementations of digital waveguide models for string instruments discussed in chapter 9. In the case of string instruments, the actual implementation of the string loop is quite straightforward, as the model is purely linear, and no danger of aliasing exists if the model input is band-limited. When using SAOL, this is the default mode of operation, as the SAOL language is sample-oriented. The actual computation might be done otherwise, but this is invisible for the programmer's viewpoint, and depends on the SAOL compiler. Sample based processing is not known for its high performance, but adequate speeds can nonetheless be achieved [53]. The Matlab implementation is straightforward, while the SAOL implementation requires the definition of a few user opcodes for fractional delay and matrix-vector products.

14.2.2   Block processing

Block processing is more complex, as the outputs of all the individual string models have to be combined by the coupling matrix. The block size is thus limited to the shortest delay line used in the model. Just as with the wind
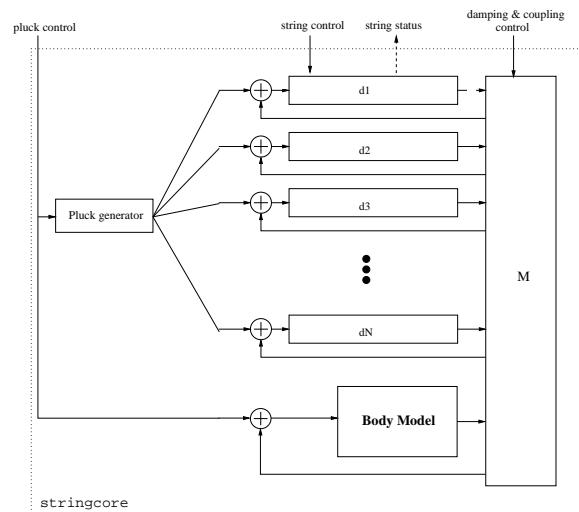
Figure 14.2: `stringcore` is the core signal processing routine of the implementation. It is implemented as a SAOL user opcode, and consists of the twelve digital waveguides, the coupling matrix, the body model, and a pluck generator.

instrument implementations, it is easier to define a fixed processing block size, by using the shortest block size. In the case of the acoustic guitar, we could set the highest note playable as the 16th fret of the high E string. This corresponds to a frequency of 830.61 in the equally tempered tuning, or a delay line length of about 53 samples. This implementation type is not possible using SAOL.

14.3   Score interpreter implementation

The score interpreter is the part of the implementation that transforms score-level directives (note pitch, attack velocity, duration, etc.) into digital waveguide control parameters (delay line length, excitation, etc.). The score interpreter largely determines the naturalness of the synthesis of a sequence of notes, while `stringcore` is the main responsible for the naturalness of a single note. The score interpreter has to make decision on the note-to-string allocation, and on the type and place of the string excitation. As will become clear, score interpretation for a polyphonic multiple-fret instrument is non-causal and thus not possible in low-latency, real-time conditions. We only discuss the SAOL implementation, as this more clearly shows the different design decisions. We start by looking at the use of the MPEG-4 native SASL control language, and then extend this with single-channel and multichannel MIDI.

Action	Commands
damp a string	alter coupling
pluck a string	accept input
select fret	**if** string active
	begin
	set coupling (damp)
	wait 100ms
	end
	rescale delay line
	set coupling for new delay length

Table 14.1: Breakdown of the most common SASL-level events into `stringcore` low-level control.

14.3.1   SASL-MPEG interpreter

Using SASL allows the programmer to define all control parameters directly. The control structure of figure 14.1 is used. The guitar model is not accessed directly by the SASL input, but by a score-level Finite State Machine. The FSM translates the physical control parameters like string number, fret number, pluck type, position etc. into a sequence of elementary commands for the digital waveguide model `stringcore`. The `stringcore` routine has a very low-level functionality. The algorithm recognizes commands to rescale the delay lines, alter the coupling coefficients, or to accept input. Some actions, like a fret change, require a sequence of commands. The most commonly used low-level commands are summarized in table 14.1.

14.3.2   MIDI

In the MPEG-4 interpretation of the MIDI standard, every MIDI note received by de orchestra results in a new instance of the instrument requested. The note is then sent to that particular instance, and when the `Note Off` message is received, that instance dies. This is obviously not what we want, as all the notes should be played on the same virtual instrument. To solve this, the MIDI interpreter dynamically creates a unique instance of the instrument, and sets a flag through a global variable to direct all other MIDI notes to that instance. This is schematically represented in figure 14.4. We consider two different types of MIDI input: single-channel MIDI, and multichannel MIDI, where one MIDI channel is used for every string.

Single channel MIDI-MPEG interpreter
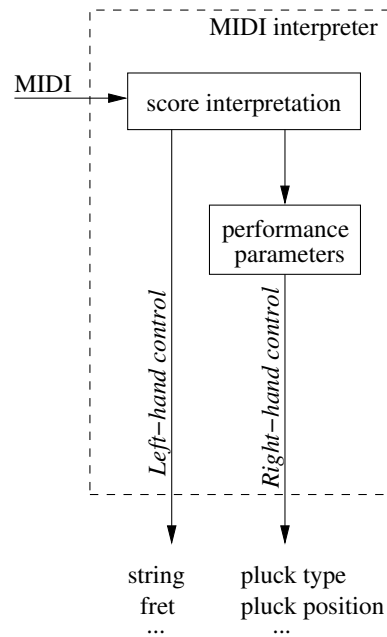
Figure 14.3: The MIDI interpreter generates SASL-like output for both the left hand control parameters (string number, fret), and the right hand control parameters (pluck type, position, angle, amplitude, velocity, etc.).
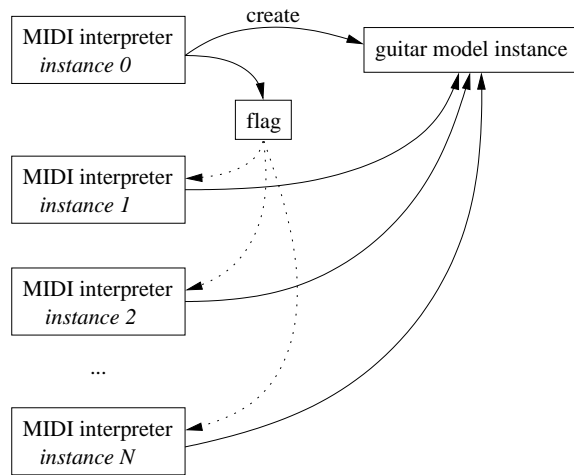
Figure 14.4: The first instance of the MIDI interpreter creates a dynamic instance of the guitar model. All subsequent instances of the interpreter direct their output to that particular instrument instance.

The guitar inherently requires a more complex player model than any wind instrument, or even many string instruments. The problem with the guitar is that different notes can be played on more than one string simultaneously, and the same note can be played on more than one string. The player has to decide on which string a particular note has to be played. A real player takes into account

- the note to be played

- the duration of the note

- the next notes to be played

- the specific indications written on the score.

This requires practice of the piece.

In the context of streaming MIDI, the virtual player model can only decide the "best" string and fret combination to play a particular note on the past notes. The algorithm selected to achieve this is fairly simple and tries to use the string with minimum fret number (that is, use the string with the lowest possible playing position if available)

- calculate `min(fret)` over all strings

- select the string with minimum fret number

- if the string corresponding to `min(fret)` is busy, select another string with minimal fret

- if no other valid string/fret combination exists, damp the vibration on the string corresponding to `min(fret)`, and select that string

- resize the delay line and excite the model

This system of string selection guarantees that every note of the MIDI stream is played, but does not guarantee that every note is played during the required time. The composer is also somewhat restricted to the preprogrammed abilities of the player model.

Multiple channel MIDI-MPEG interpreter

The most versatile implementation for streaming MIDI input, assigns a MIDI channel to each string. The composer can now decide on which string every note is to be played.
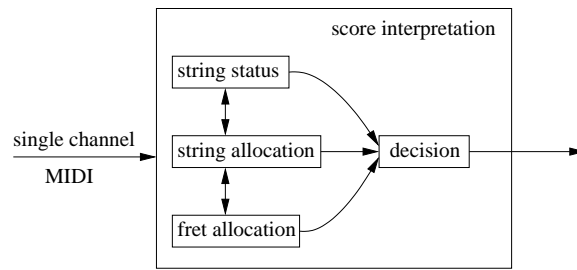
Figure 14.5: The score interpretation part of the MIDI interpreter decides on which string and which fret the incoming MIDI note will be played.

Example 14.1 ⋆ `CD Track 30` is a straightforward rendering of the first sixteen bars of *Leyenda* (Asturias) by Isaac Albeniz. It is the unedited output of the SAOL implementation with a very simple score and almost no dynamic information. Every pluck is somewhat different by the addition of a random pluck angle and amplitude variation to mimic a human player. ⋆ `CD Track 31` has been synthesized using the multiple MIDI channel SAOL implementation. It is the fifth movement of a piece written for physically modeled guitar by Nick Collins. It shows the power of the physical model in rendering transhuman performance while keeping a natural sound. In this case, the pluck position was determined automatically depending on the string the notes were played, and the pluck type depends on the dynamic indication in the score. A seven-track output was generated (one track for each string, and one for the body). Finally, the different track were mixed by the composer using Protools. Some reverberation was also added. It is the result of several rendering-listening iterations, needed to improve the musical interpretation of the score by the rendering engine. △

14.4 Comparison with other synthesis techniques

In this section, the properties of the Fully coupled digital waveguide model is compared to other techniques commonly used for the synthesis of string instrument tones.

Table 14.2 gives an overview of the techniques that are compared here. We selected the most successful and widespread types, and compare them to the fully coupled model presented in this work. It is a qualitative comparison that looks at different aspects of the synthesis techniques. We compare the fully coupled model with a commuted digital waveguide model with the simplified coupling mechanism described in section 11.2.4 `SimCo` [43], a model based on

	Description
`FullCo` [69]	Physical model
	Digital waveguide
	Simplified Full Coupling
`SimCo` [43]	Physical model
	Commuted digital waveguide
	Simplified coupling
`NumSol` [106]	Physical model
	Multi-dimensional Transfer Function Model
`Modal` [1]	Physical model
	Vibrating substructures
`CORDIS` [32]	Physical model
	Mass-spring networks
`Add` [64]	Spectral model
	Additive synthesis
	Per note
`Sampling`	Processed recording
	Per note

Table 14.2: Description of the different techniques available for the synthesis of string instrument tones.

the functional transformation of the partial differential equation describing the string movement `NumSol` [106], with a modal synthesis model `Modal` [1], a model based on the simulation of mass-spring networks `CORDIS` [32], a spectral additive synthesis model `Add` [64], and with wavetable synthesis `Sampling`.

Most of the existing models are physics-based, starting from the basic mechanical vibration properties of the instrument. The models differ in the way the solution to the equations describing the system is determined.

The `SimCo` digital waveguide model for the acoustic guitar is comparable to the `FullCo` model. It is a single-delay-line, dual polarization model with the simplified coupling mechanism described in section 11.2.4. The main difference is in the use of an excitation signal database, obtained by inverse-filtering recorded tones with the calibrated model, and the absence of a body model in the coupling mechanism. The `FullCo` model has more possibilities due to the way the body model is included, and a more accurate coupling mechanism. The `SimCo` model produces tones with a more natural attack because of the excitation database. The parameter estimation techniques that we developed in this work can also be used for the calibration of the `SimCo` model.

The `NumSol` method transforms a continuous model for the vibrating body, given a partial differential equation, into a multi-dimensional transfer function

model, taking into account the initial and boundary conditions, and the excitation function. The functional transformation yields a model with the same accuracy as a finite element model, but with a much lower computation load. These models use the physical parameters of the strings (length, cross section, density, etc.) which makes calibration quite hard. The sound quality is very high. The method is more accurate than `FullCo`, but requires a considerable to build each individual model, and does not account for the resonator.

Modal synthesis is based on the simulation of vibrating structures. The user defines a set of coupled substructures, defined by modal data. The method can be applied to structures of arbitrary complexity, calibration is possible, but the model requires an enormous computational effort, and is not suited for the real-time synthesis of practical systems. It is used for the theoretical analysis of complex coupled vibrating systems. It has the same synthesis capabilities as `FullCo`, but no calibration, a much higher computational load, and more difficult model design.

The `CORDIS` system attempts to model the acoustical system using simple ideal mechanical elements. By interconnecting a large number of these elementary systems, a model of a musical instrument can be obtained. As with all numerical acoustics synthesis methods, the sound quality can be very high, but with high computational cost.

The most widely used technique is `Sampling`, or wavetable synthesis. In this synthesis method, a prerecorded tone is played back at the desired pitch. Contrary to the wind instrument case, this method is quite acceptable for string instrument tones, as these have clearly defined attack and decay parts. In some cases, no looping is needed at all. By sampling every single note of the instrument with different attacks, very convincing tones can be produced. The player has then to select the appropriate sample for playback, according to the pluck amplitude, pluck type and plucking point. The main disadvantage is the very high memory requirement for good-quality synthesis. The best sample-based digital piano models have sample databases containing more than a gigabyte of data. Compared to `FullCo`, it is much more difficult to have realistic sympathetic vibrations.

14.5   Conclusion

The control of a physical model is both easy and complicated. The main advantage for a human player is that sensor inputs are directly mapped to the parameters controlling the model. If the sound has to be synthesized based on a score, the situation is more complex, and artificial input has to be generated. Part of this task can be delegated to the composer of the piece when either SASL or multichannel MIDI are used. When the usual single-channel MIDI is used,

a string/fret allocation algorithm is needed, which is in essence a player model. The actual input of the model could be either selected from an excitation signal database or could be generated during synthesis by the use of a parametric input model. The most practical and compact solution is the combination of a single-channel MIDI input interpreter and a parametric pluck generator. The best sound quality and musical expression is however achieved with a multichannel MIDI model, excited with samples out of a recomputed database.

Chapter 15

Conclusions and Further Research

In the first section of this chapter, the global conclusions are formulated. In section 15.2 some suggestions are given for further research.

15.1   Conclusions

This thesis was divided in three major parts: **Part I** on acoustic theory, **Part II** on wind instruments, and **Part III** on string instruments.

In **part I**, an introduction has been provided on musical acoustics and the digital waveguide modeling technique.

We started by providing a qualitative overview of the wind instrument family in **chapter 2**. We discussed two single-reed instruments (clarinet and saxophone), an air-jet driven instrument (the modern flute) and a lip-driven instrument (the trumpet). It was concluded that they all share the same basic structure, and that a general digital waveguide model could be used to model all the instruments. **Chapter 3** gave an qualitative overview of the acoustic properties of string instruments. Three different types of string instruments were compared: the acoustic and electric guitars, the violin family, and the piano. It is concluded that the members of string instrument family mostly differ in the string excitation mechanism, while all other properties are comparable. **Chapter 4** provided an introduction to linear musical acoustics. Starting from the general equations describing sound propagation in air and strings, the properties of common acoustic building blocks of musical instruments were derived. The

219

different elements were then combined into acoustic systems. In **chapter 5**, the acoustic elements were discretized using the digital waveguide modeling technique as the general traveling-wave solution to the wave equation. We discussed the importance of fractional delay, and showed that Catmull-Rom and B-spline based interpolation are possible alternatives to Lagrage interpolation and allpass filtering. These two chapters were then used as a base for the development of models for specific musical instruments.

In **Part II**, the properties of wind instruments have been studied and specific modeling and parameter estimation techniques were developed.

In **chapter 6**, a complete acoustic model has been built for a single-reed instrument, disregarding the influence of the toneholes. We discussed the resonator of a clarinet and a saxophone, and elaborated on the excitation mechanism. We introduced high-order linear and non-linear models to describe the reed movement. The core of Part II was contained in **chapter 7**, where we built a discrete time model of a clarinet, as an example of the single-reed instruments. We combined source-filter synthesis with the digital waveguide modeling technique and obtained a novel model class that is both easy to calibrate and flexible from a performance point of view [74]. The model consists of a non-linear excitation mechanism, a linear resonator, and an input-dependent dynamic filter. This post-equalization filter, or Formant filter, can be obtained using several filter design techniques. We compared low-order direct estimation [68], Linear Predictive Coding, cepstrum based methods, and simple Wiener filter designs. For sufficiently high filter orders, a very good match with a recorded signal was obtained. The models were then expanded into multinote models in **chapter 8** [67]. We compared several structures, and propose a parallel crossfading method to simulate the transition between notes. Finally, an MPEG-4 SAOL implementation was presented in **chapter 9**.

Part III addressed string instruments.

In **chapter 10**, we presented a single-string model for the acoustic guitar. The properties of ideal and real strings were discussed, including tension modulation non-linearity and fret beating. The dual-polarization models accounted for vibrations in two spatial planes. The focus of this chapter was on the parameter estimation: we compared different methods in the time and frequency domains, and showed that the subband Hankel Singular Value Decomposition offer both high accuracy and automated estimation [71]. These single-string models were then extended in **chapter 11**, where we proposed a new general framework that allows the performance comparison of coupled string structures [69]. The framework indicated how to improve the parameter estimation for the structures and how to verify the stability. We derive the properties for N coupled strings and N coupled digital waveguides, and compared five coupled string structures. Finally, we proposed a fully coupled digital waveguide model for the acoustic guitar [73]. This model allowed expressive musical effects not

achieveable with previous models. The body of an acoustic guitar is the main factor determining the timbre a the instrument, and is an important part of the fully coupled digital waveguide model In **chapter 12**, we discussed several linear modeling techniques for the guitar body, and proposed a non-linear extension to catch the subtle differences between the different tone woods used in guitar building [76]. The model has been validated in **chapter 13** [72]. Finally, **chapter 14** discussed the implementation of the guitar model. It was shown that a primitive player model is needed to obtain natural-sounding output. A multi-layered MPEG-4 SAOL model was presented [75]. The complete model has an equivalent or better sound quality than the best of the other published techniques, and with more and different sonic abilities and easier control by the player.

15.2 Further research

This thesis put more emphasis on the theoretical aspects of the modeling. There are some implementation issues that were not addressed, and some model-related problems as well.

For the wind instruments in particular, the numerical properties of the different types of formant filters should be compared from the implementation point of view. As these are dynamic filters, transient noises must be minimized. This requires a careful examination of the filter structure and its finite-precision properties. The non-linear excitation mechanism could result in audible aliasing if the cut-off frequency of the reflection filter is set quite high. Oversampling and steeper filters can alleviate this.

For the string instruments, the main problem remains in the determination of a suitable parametrization of the finger-string interaction. There is very little quantitative literature to be found on what exactly happens when a note is plucked. Professional guitar players are not able to exactly describe what they do, the tonal character of the notes played merely results from years of practice and subtle improvements of the playing technique. Further, the coupled string models should be applied to larger systems, like the orchestral harp or the piano. It might be beneficial to link the coupled digital waveguide representations of the strings with a waveguide mesh implementation of the resonator, to obtain a more accurate interaction. In general, higher dimensional digital waveguide structures can be used to solve some sub-problems encountered. One of the issues is the way a one-dimensional waveguide is linked to a two- or three-dimensional waveguide mesh.

Finally, some of the techniques used in this thesis can be used to improve the output of the circular waveguides and the waveguide meshes, with possible applications in reverb and virtual acoustics.

Appendices

Appendix A

CD track list

The first three tracks relate to chapter 5, and are the simulated response of cylindrical and conical bores excited with white noise. Note the difference in timbre between the first and third tracks.

Track 1. Cylindrical bore, closed-open, excited with white noise, length=25 cm

Track 2. Cylindrical bore, closed-open, excited with white noise, length=50 cm

Track 3. Conical bore, complete cone, excited with white noise, length=50 cm

The next 16 tracks illustrate some issue in chapters 7 and 8. First, the same single-reed model is used in conjunction with a cylindrical and a conical bore. The formant filter examples show that for the higher notes, the model output is very close to the reference when higher-order filters are used. For the lower note, a higher filter order is needed. Finally, tracks 18 and 19 are examples of serial crossfading and the much simpler joint parallel crossfading techniques.

Track 4. Cylindrical bore with reed model (clarinet-like)

Track 5. Conical bore with reed model (saxophone-like)

Track 6. Clarinet, low note, *pp*, reference

Track 7. Clarinet, low note, *pp*, 2nd order LPC formant filter

Track 8. Clarinet, low note, *pp*, 20th order LPC formant filter

Track 9. Clarinet, low note, *ff*, reference

Track 10. Clarinet, low note, *ff*, 2nd order LPC formant filter

Track 11. Clarinet, low note, *ff*, 20th order LPC formant filter

Track 12. Clarinet, high note, *pp*, reference

Track 13. Clarinet, high note, *pp*, 2nd order LPC formant filter

Track 14. Clarinet, high note, *pp*, 20th order LPC formant filter

Track 15. Clarinet, high note, *ff*, reference

Track 16. Clarinet, high note, *ff*, 2nd order LPC formant filter

Track 17. Clarinet, high note, *ff*, 20th order LPC formant filter

Track 18. Clarinet, multinote, serial crossfading

Track 19. Clarinet, multinote, joint parallel crossfading

The last series of tracks relate to the string instruments models. Track 20 is the reference recording of a low A played on a classical guitar. The next tracks are a sequence of model improvements, culminating in the fully coupled model, with parameters extracted from the reference recording. Tracks 26 and 27 are examples of tension modulation non-linearity, with the parameters of nylon strings and steel strings - there is a quite subtle difference in timbre during the attack. Track 28 is a fully coupled guitar model where the coupling parameters are maximized to the limit of system stability. Track 29 is a fully coupled guitar model, excited by an impulse to the body. Note the sympathetic vibration of the strings. Track 30 is the output of the MPEG-4 SAOL implementation of the fully coupled model. Finally, Track 31 is the last movement of the *Suite for Physically Modeled Guitar*, and shows the many compositional possibilities offered by a physical model. The piece cannot be play by a human.

Track 20. Guitar, reference

Track 21. Guitar, simulation, single string, single polarization

Track 22. Guitar, simulation, single string, dual polarization without coupling

Track 23. Guitar, simulation, dual polarization with coupling

Track 24. Guitar, simulation, dual polarization with coupling and guitar body

Track 25. Guitar, simulation, matched fully coupled model

Track 26. Guitar, simulation, tension modulation non-linearity, nylon string

Track 27. Guitar, simulation, tension modulation non-linearity, steel string

Track 28. Guitar, simulation, maximum coupling, arpeggio

Track 29. Guitar, simulation, body impulse with undamped strings

Track 30. Guitar, simulation, normal playing, excerpt from *Leyenda* by Albéniz

Track 31. Guitar, simulation, transhuman performance, *Suite for Physically Modeled Guitar, mvt 5* by Nick Collins

Bibliography

[1] J.M. Adrien. Dynamic modeling of vibrating structures for sound synthesis, modal synthesis. In *Proceedings of the AES 7th International Conference*, pages 291–300, Toronto, Canada, 1989. Audio Engineering Society.

[2] Jont B. Allen and Lawrence R. Rabiner. A unified approach to short-time fourier analysis and synthesis. *Proceedings of the IEEE*, 65(11):1558–1564, November 1977.

[3] D. Arfib. Digital synthesis of complex spectra by means of multiplication of nonlinear distorted sine waves. *Journal of the Audio Engineering Society*, 27(10):757–768, 1979.

[4] MIDI Manufacturers Association. The complete midi 1.0 detailed specification v.96.2, 1996.

[5] M.A. Austin, X.G. Chen, and W.J. Lin. Aladdin: A computational toolkit for interactive engineering matrix and finite element analysis. Technical Report ISR Technical Research ReportTR95-74, University of Maryland, 1995.

[6] J. Backus. Small vibration theory of a clarinet. *J. Acoust. Soc. Am.*, 35(3):591–599, 1963.

[7] N. Bak and P. Domler. The relation between blowing pressure and blowing frequency in clarinet playing. *Acustica*, 63:238, 1987.

[8] A.H. Benade and E.V. Jansson. On the plane and spherical waves in horns with nonuniform flare. *Acustica*, 31:80–98, 1974.

[9] Arthur A. Benade. On the mathematical theory of woodwind finger holes. *Journal of the Acoustical Society of America*, 32(12):1591–1608, 1960.

[10] Arthur H. Benade. On woodwind instrument bores. *Journal of the Acoustical Society of America*, 31(2):137–146, 1959.

[11] Arthur H. Benade. On the propagation of sound waves in a cylindrical conduit. *Journal of the Acoustical Society of America*, 44(2):616–623, 1968.

[12] Arthur H. Benade. *Fundamentals of musical acoustics, 2nd edition.* Dover Publications, New York, 1990.

[13] L.L. Beranek. *Acoustics*, chapter 3. McGraw-Hill, 1954. Reprinted 1986, Ac. Soc. Am.

[14] Gianpaolo Borin, Giovanni De Poli, and Davide Rocchesso. Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems. *IEEE Transactions on Speech and Audio Processing*, 8(5), September 2000.

[15] Albertus C. Den Brinker, Frank P.A. Benders, and Tomas A.M. Oliveira e Silva. Optimality conditions for truncated kautz series. *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, 43(2):117–121, Februari 1996.

[16] G. F. Carrier. On the non-linear vibration problem of the elastic string. *Questions on Applied Mathematics*, 3:157–165, 1945.

[17] J.M. Chowning. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, 21(7):526–534, 1973.

[18] O. Christensen. Qualitative models for low frequency guitar function. *Journal of Guitar Acoustics*, (6):10–25, 1982.

[19] Nick Collins. Relating superhuman virtuosity to human performance. In *Proc. of MAXIS 2002*, 2002.

[20] Perry R. Cook. Non-linear periodic prediction for on-line identification of oscillator characteristics in woodwind instruments. In *Proc. of the 1991 International Computer Music Conference*, pages 157–160, 1991.

[21] Perry R. Cook. A meta-wind-instrument physical model, and a meta-controller for real-time performance control. In *Proceedings of the International Computer Music Conference*, pages 273–276, 1992.

[22] Perry R. Cook. Synthesis toolkit in c++, version 1.0. In *Course #17 & #18, Creating and Manipulating Sound to Enhance Computer Graphics, SIGGRAPH 1996*, May 1996.

[23] Perry R. Cook. Towards physically-informed parametric synthesis of sound effects. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, 1999.

[24] Perry R. Cook and Gary P. Scavone. The synthesis toolkit (stk). In *Proc. of the 1999 International Computer Music Conference*, Beijing, China, 1999.

[25] Yamaha Corporation. Annual report, 2002.

[26] L. Cremer. *The Physics of the Violin.* MIT Press, Cambridge, Massachusetts, 1984.

[27] Francois Déchelle, Riccardo Borghesi, Maurizio De Cecco, Enzo Maggi, Butch Rovan, and Norbert Schnell. jmax: a new java-based editing and control system for real-time musical applications. In *Proceedings of the International Computer Music Conference*, 1998.

[28] Scott A. Van Duyne and Julius O. Smith. Physical modeling with the 2-D digital waveguide mesh. In *Proc. of the International Computer Music Conference*, pages 40–47, Tokyo, Japan, 1993.

[29] Wim Ermens. Single-reed wind instruments: The non-linear oscillator. Master's thesis, Dept. of Electrical Engineering, Katholieke Universiteit Leuven, 1999.

[30] Neville H. Fletcher. Plucked strings - a review. *Catgut Acoust. Soc. Newsletter*, (26):13–17, 1976.

[31] Neville H. Fletcher and Thomas D. Rossing. *The Physics of Musical Instruments.* Springer-Verlag Inc., New York, 1991.

[32] J.-L. Florens and C. Cadoz. The physical model: modeling and simulating the instrumental universe. In G.D. Poli, A. Piccialli, and C. Roads, editors, *Representations of Musical Signals*, pages 227–268. The MIT Press, Cambridge, Massachusetts, USA, 1991.

[33] Thierry Galas and Xavier Rodet. An improved cepstral method for deconvolution of source-filter systems with discrete spectra: application to musical sound signals. In *Proceedings of the 1990 Internation Computer Music Conference(ICMC)*, Glasgow, September 1990.

[34] G. E. Garnett. Music, signals, and representations: A survey. In G. De Poli, A. Picialli, and C. Roads, editors, *Representations of Musical Signals*, pages 325–370. MIT Press, Cambridge, MA, 1991.

[35] B. Halder and T. Kailath. Efficient estimation of closely spaced sinusoidal frequencies using subspace-based methods. *IEEE Signal Processing Letters*, 4(2):49–51, Feb 1997.

[36] Fredric J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–84, January 1978.

[37] Thomas Jan De Hoog. *Rational orthonormal bases and related transforms in linear system modeling.* PhD thesis, Technische Universiteit Delft, 2001.

[38] Andrew Horner. Nested modulator and feedback FM matching of instrument tones. *IEEE Transactions on Speech and Audio Processing,* 6(4):398–409, July 1998.

[39] International Standards Organization. MPEG-4 Audio, ISO 14496 (MPEG-4), part 3 (audio), 1998. Geneva, CH:ISO.

[40] International Standards Organization. International Standard ISO 14496 (MPEG-4), part 3 (audio), subpart 5 (Structured Audio), 1999. Geneva, CH:ISO.

[41] D.A. Jaffe and J.O. Smith. Extensions of the Karplus-Strong plucked string algorithm. *Computer Music Journal,* 7(2):56–69, 1983.

[42] Jean-Marie Jot. An analysis/synthesis approach to real-time artificial reverberation. In *Proceedings of the IEEE International Conferentce on Acoustics, Speech , and Signal Processing (ICASSP),* pages 221–224, San Fransisco, CA, USA, 1992.

[43] Matti Karjalainen, Vesa Välimäki, and Tero Tolonen. Plucked-string models: From the Karplus-Strong algorithm to digital waveguides and beyond. *Computer Music Journal,* 22(3):17–32, Fall 1998.

[44] Kevin Karplus and Alex Strong. Digital synthesis of plucked-string and drum timbres. *Computer Music Journal,* 7(2), 1983.

[45] W. H. Kautz. Transient synthesis in the time domain. *IRE Transactions on Circuit Theory CT-1,* pages 29–39, 1954.

[46] D. H. Keefe. Wind-instrument reflection function measurement in the time domain. *Acoust. Spc. Am.,* 99(4):2370–2381, 1996.

[47] Douglas H. Keefe. *Woodwind Tone-hole Acoustics and the Spectrum Transformation Function.* PhD thesis, Case Western Reserve University, 1981.

[48] Douglas H. Keefe. Physical modeling of wind instruments. *Computer Music Journal,* 16(4):57–73, Winter 1992.

[49] J.P. Kelly. *Principles of Neural Science,* chapter Auditory system, pages 258–268. Elsevier, 1981.

[50] T. L. Laakso, V. Välimäki, M. Karjalainen, and U. Laine. Real-time implementation techniques for a continiously variable digital delay in modeling musical instruments. In *ICMC,* pages 14–18, 1992.

[51] L.D. landau and E.M. Lifschitz. *Theory of Elasticity.* Pergamon Press, Oxford, England, 1986.

[52] J. Laroche. A new analysis/synthesis system of musical signals using prony's method. application to heavily damped percussive sounds. In *Proc. IEEE ICASSP-89*, pages 2053–2056, Glasgow, May 1989.

[53] J. Lazzaro and J. Wawrzynek. Compiling MPEG-4 Structured Audio into C. In *Proceedings of the Second IEEE MPEG-4 Workshop and Exhibition (WEMP)*, San Jose, CA, USA, June 18-20 2001.

[54] K. A. Legge and N. H. Fletcher. Nonlinear generation of missing modes on a vibrating string. *Journal of the Acoustic Society of America*, 76(1):5–12, 1984.

[55] H. Levine and J. Schwinger. On the radiation of sound from an unflanged pipe. *Physical Review*, 43:383–406, 1948.

[56] Creative Technology Ltd. Annual report, 2001.

[57] J. Makhoul. Linear prediction: A tutorial review. *Proc. IEEE*, 63:561–580, 1975.

[58] Max Mathews. *The Technology of Computer Music.* MIT Press, 1969.

[59] James McCartney. Supercollider: a new real time synthesis language. In *Proceedings of the International Computer Music Conference*, Hong Kong, 1996.

[60] M. E. McIntyre, R. T. Schumacher, and J. Woodhouse. On the oscillations of musical instruments. *Journal of the Acoustical Society of America*, 74(5), Nov. 1983.

[61] J. Meyer. Die abstimmung der grundresonanzen von guitarren. *Das Musikinstrument*, (23):176–186, 1974.

[62] J. Meyer. Quality aspects of the guitar tone. In E. V. Jansson, editor, *Function, Construction, and Quality of the guitar.* Royal Swedish Academy of Music, Stockholm, 1983.

[63] J.A. Moorer. The use of linear prediction of speech in computer music. *Journal of the Audio Engineering Society*, 27(3):134–140, 1979.

[64] J.A. Moorer. Signal processing aspecs of computer music: a survey. In J. Strawn, editor, *Digital Audio Signal Processing: An Anthology*, chapter 5, pages 149–220. William Kauffmann, Inc., 1985.

[65] P.M. Morse. *Vibration and Sound.* McGraw-Hill, 1948.

[66] P.M. Morse and H. Feshbach. *Methods of Mathematical Physics*, pages 494–523, 655–666. McGraw-Hill, New York, 1953.

[67] Axel Nackaerts, Bart De Moor, and R. Lauwereins. Note transitions in physical models of a wind instrument. In *Proc. of Acoustics and Music: Theory and Applications 2000 (AMTA 2000, CD-ROM)*, pages 381–385, Montego Bay, Jamaica, December 2000.

[68] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. Dynamic timbre matching for sound synthesis using time-varying parametric filters. In *Proc. of the 43rd IEEE Midwest Symposium on Circuits and Systems (MWSCAS 2000, CD-ROM)*, Lansing, Michigan, August 2000.

[69] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. Coupled string guitar models. In C.E. D'Attellis, V.V. Kluev, and N.E. Mastorakis, editors, *Mathematics and Simulation with Biological, Economical and Musicoacoustical Applications*, pages 289–294. WSES Press, 2001.

[70] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. Coupled string guitar models. In *Proceedings of Acoustics and Music : Theory and Applications (AMTA 2001, CDROM)*. WSES, 2001.

[71] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. Parameter estimation for dual-polarization plucked string models. In *Proc. of the International Computer Music Conference (ICMC 2001)*, pages 203–206, 2001. La Habana, Cuba.

[72] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. Measurement of guitar string coupling. In *Proc. of the International Computer Music Conference (ICMC 2002)*, pages 321–324, Göteborg, Sweden, September 2002.

[73] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. Parameter estimation for coupled string models. Technical Report TR 02-01, Dept. Elektrotechniek, ESAT-SCD, Katholieke Universiteit Leuven, 2002. ftp://ftp.esat.kuleuven.ac.be/pub/sista-/nackaerts/reports/report_JAES02.ps.gz.

[74] Axel Nackaerts, Bart De Moor, and Rudy Lauwereins. A formant filtered physical model for wind instruments. *IEEE Transactions on Speech and Audio Processing*, 11(1):36–44, January 2003.

[75] Axel Nackaerts, Rudy Lauwereins, and Bart De Moor. Player modeling and control of a physical model of a string instrument. In *Proc. of the 1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002)*, Leuven, Belgium, November 2002.

[76] Axel Nackaerts, Bert Schiettecatte, and Bart De Moor. Non-linear guitar body models. Technical Report TR 03-35, ESAT-SISTA, Katholieke Universiteit Leuven, 2003.

[77] C. J. Nederveen. *Acoustical aspects of woodwind instruments.* Frits Knuf, Amsterdam, The Netherlands, 1969.

[78] T. Q. Nguyen, T. I. Laakso, and R. D. Koilpillai. Eigenfilter approach for the design of allpass filters approximating a given phase response. *IEEE Transactions on Signal Processing*, 42(9):2257–2263, September 1994.

[79] T. Paatero. Generalized linear-in-parameter models. In *Finnish Signal Processing Symposium (FINSIG'99)*, pages 217 – 221, Oulu, Finland, 1999.

[80] T. Paatero and M. Karjalainen. Kautz filters and generalized frequency resolution: theory and audio applications. In *110th AES Convention*, Amsterdam, The Netherlands, 2001.

[81] B. Paillard, P. Mabilleau, S. Morisette, and J. Soumagne. Perceval: Perceptual evaluation of the quality of audio signals. *J. Audio eng. Soc.*, 40:21–31, 1992.

[82] M. Pavlidou and B. E. Richardson. The string-finger interaction in classical guitar: Theoretical model and experiments. In *Proceedings of the Institute of Acoustics*, volume 19, pages 55–60, Edinburgh, United Kingdom, 1997.

[83] H. Penttinen, M. Karjalainen, T. Paatero, and M. Järveläinen. New techniques to model reverberant instrument body responses. In *Proc. of the Interntional Computer Music Conference (ICMC 2001)*, pages 182–185, Havana, Cuba, September 2001.

[84] J.O. Pickles. *An Introduction to the Physiology of Hearing.* Academic Press, London, 1982.

[85] John R. Pierce and Scott A. Van Duyne. A passive nonlinear digital filter design which facilitates physics-based sound synthesis of highly nonlinear musical instruments. *Journal of the Acoustic Society of America*, 101(2):1120–1126, 1997.

[86] Miller Puckette. Combining event and signal processing in the max graphical programming environment. *Computer Music Journal*, pages 68–77, 1991.

[87] Miller Puckette. Pure data. In *Proceedings of the International Computer Music Conference*, pages 269–272, Hong Kong, 1996.

[88] R.W. Pyle. Effective length of horns. *Journal of the Acoustical Society of America*, (57):1309–1317, 1975.

[89] L. R. Rabiner. On the use of autocorrelation analysis for pitch detection. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 25(1):24–33, 1977.

[90] C.V. Raman. On the mechanical theory of the vibrations of bowed strings and of musical instruments of the violin family, with experimental verification of the results. *Bull. 15, The Indian Association for the Cultivation of Science*, 1918.

[91] Curtis Roads. *The Computer Music Tutorial*. The MIT Press, Cambridge, Massachusetts, USA, 1985.

[92] Xavier Rodet. Time-domain formant-wave-function synthesis. *Computer Music Journal*, 8(3):9–14, 1980.

[93] Xavier Rodet. Stability/instability of periodic solutions and chaos in physical models of musical instruments. In *Proc. of the 1994 International Computer Music Conference*, Aarhus, Danemark, 1994.

[94] T.D. Rossing. *The Science of Sound*, chapter Chapter 10. Addison-Wesley, Reading, Massachusetts, 1982.

[95] T.D. Rossing, J. Popp, and D. Polstein. Acoustical response of guitars. In *Proc. of SMAC 83*, Stockholm, 1985. Royal Swedish Academy of Music.

[96] Gary P. Scavone and Perry R. Cook. Real-time computer modeling of woodwind instruments. In *Proc. of the 1998 International Symposium on Musical Acoustics*, Leavenworth, WA, 1998.

[97] Gary Paul Scavone. *An acoustic analysis of single-reed woodwind instruments with an emphasis on design and performance issues and digital waveguide techniques*. PhD thesis, CCRMA at Stanford University, March 1997. URL: http://www-ccrma.stanford.edu/~gary/.

[98] Eric D. Scheirer and Barry L. Vercoe. SAOL: The MPEG-4 Structured Audio Orchestra Language. *Computer Music Journal*, 23(2):31–51, 1999.

[99] M. R. Schroeder. New method for measuring reverberation times. *J. Acoust. Soc. Amer.*, 37:232–235, 1965.

[100] Julius O. Smith. A new approach to digital reverberation using closed waveguide networks. In *Proc. 1985 International Computer Music Conference*, pages 44–53, Vancouver, Canada, 1985. Computer Music Association.

[101] Julius O. Smith. Efficient simulation of the reed-bore and bow-string mechanisms. In *Proc. 1986 International Computer Music Conference*, pages 275–280, Den Haag, Netherlands, 1986. Computer Music Association.

[102] Julius O. Smith. Viewpoints on the history of digital synthesis. In *Proceedings of the ICMC*, pages 1–10, Montreal, Oct. 1991.

[103] Julius O. Smith. Physical modeling using digital waveguides. *Computer Music Journal*, 16(4):74–91, 1992.

[104] Julius O. Smith. Discrete-time modeling of acoustic systems with applications to sound synthesis of musical instruments. online publication, Apr. 1996.

[105] Tero Tolonen, Vesa Vïimäki, and Matti Karjalainen. Modling of tension modulation nonlinearity in plucked strings. *IEEE Transactions on Speech and Audio Processing*, 8(3):300–310, May 2000.

[106] Lutz Trautmann and Rudolf Rabenstein. Sound synthesis with tension modulation nonlinearities based on functional transformations. In Nikos Mastorakis, editor, *Mathematics and Computers in Modern Science: Acoustics and Music, Biology and Chemistry, Business and Economics*, Mathematics and Computers in Science and Engineering, pages 69–74. WSES Press, 2000.

[107] V. Välimäki, T. Tolonen, and M. Karjalainen. Signal-dependent nonlinearities for physical models using time-varying fractional delay filters. In *Proceedings of the International Computer Music Conference (ICMC 1998)*, pages 264–267, Ann Harbor, Michigan, USA, 1998.

[108] V. Välimäki, T. Tolonen, and M. Karjalainen. Plucked-string synthesis algorithms with tension modulation nonlinearity. In *Proceedings of the IEEE International Conference in Acoustics, Speech, and Signal Processing (ICASSP 99)*, volume 2, pages 977–980, March 15-19 1999. Phoenix, Arizona, `http://www.acoustics.hut.fi/∼vpv/publications/icassp99-tm.htm`.

[109] Vesa Välimäki. *Discrete-time modeling of acoustic tubes using fractional delay filters*. PhD thesis, Helsinki University of Technology, Faculty of Electrical Engineering, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland, 1995. Report no. 37.

[110] S. Van Huffel, H. Chen, C. Decanniere, and P. Van Hecke. Algorithm for time-domain NMR data fitting based on Total Least Squares. *J. Magn. Reson.*, A110:228–237, 1994.

[111] B. Vercoe. *Csound: a manual for the audio processing system.*
      Massachusetts Institute of Technology Media Lab, Cambridge, MA,
      1996.

[112] Barry L. Vercoe, William G. Gardner, and Eric D. Scheirer. Structured
      audio: Creation, transmission, and rendering of parametric sound
      representations. *Proceedings of the IEEE*, 86(5):992–940, May 1998.

[113] M.P. Verge, B. Fabre, W.E.A. Mahu, A. Hirschberg, R.R. van Hassel,
      A.P.J Wijnands, J.J. de Vries, and C.J. Hogendoorn. Jet formation and
      jet velocity fluctuations in a flue organ pipe. *J. Acoust. Soc. Am.*,
      95:1119–1132, 1994.

[114] M.P. Verge, A. Hirschberg, and Causs. Sound production in recorderlike
      instruments. ii a simulation model. *J. Acoust. Soc. Am.*, 101:2925–2939,
      1997.

[115] Bo Wahlberg. System identification using kautz models. *IEEE Trans.
      on Automatic Control*, 39(6):1276–1281, 1991.

[116] Bo Wahlberg. System identification using laguerre models. *IEEE Trans.
      on Automatic Control*, 36(5):551–562, 1991.

[117] J.J. Walsh. *Interpolation and approximation by rational functions in the
      complex domain.* American Mathematical Society, Rhode Island, 1956.
      2nd edition.

[118] A.G. Webster. Acoustical impedance, and the theory of horns and of
      the phonograph. In *Proc. of the US National Academy of Sciences*,
      volume 5, pages 275–282, 1919.

[119] Gabriel Weinreich. Coupled piano strings. *Journal of the Ac. Soc. of
      Am.*, 62(6):1474–1484, December 1977.

[120] T.A. Wilson and G.S. Beavers. Self-sustained oscillations of the
      clarinet: An integral equation approach. *Journal of the Acoustical
      Society of America*, 56:653, 1974.

[121] Juzhong Zhang, Garman Harbottle, Changsui Wang, and Zhaochen
      Kong. Oldest playable musical instruments found at jiahu early
      neolithic site in china. *Nature*, 401:366–368, 1999.