Proceedings of the 2017 International Symposium on Musical Acoustics

Montreal, QC, Canada

Gary Scavone, Esteban Maestre, Connor Kemp, Song Wang
Schulich School of Music, McGill University

18–22 June 2017

Published by:

The Schulich School of Music
McGill University
555 Sherbrooke Street West
Montreal, QC H3A 1E3
Canada

Credits:
Cover Design: Esteban Maestre
Editors: Gary Scavone, Esteban Maestre, Connor Kemp, Song Wang

Welcome from the ISMA 2017 Organizing Committee

On behalf of the members of the organizing committee, we are pleased to welcome you to the 2017 International Symposium on Musical Acoustics (ISMA) at McGill University in Montreal, Quebec, Canada.

This is the first time an ISMA has been held in Canada. We hope this event helps shine a light on musical acoustics research being conducted from coast to coast to coast in Canada.

It has been a busy year of preparations. We look forward to stimulating you with a workshop, three keynote addresses and more than seventy contributed papers or posters over the next four days. We have tried to organize events to allow time for research discussions with colleagues, both between sessions and in social contexts. We hope you will find our banquet venue, Lion d'Or, to be a most intimate and inviting place to meet and perhaps share some music.

We would like to gratefully acknowledge financial or in-kind support for this event from the Schulich School of Music, the Centre for Interdisciplinary Research in Music Media and Technology, the Social Sciences and Humanities Research Council of Canada, and the Acoustical Society of America. In addition, we must thank those members of the Scientific Committee who participated in the paper review process, as well as the numerous organizing committee members and student volunteers who have spent many hours working out all the details necessary to allow this event to run smoothly. We personally wish to express our deep gratitude for their time and dedication.

We hope you have a wonderful stay in Montreal and highly recommend that you take some time to discover the sites and tastes of the city. And we hope you find this ISMA to be a most interesting, informative, and inspirational event.

Gary Scavone, Conference Chair
Esteban Maestre, Paper Chair

The ISMA 2017 is organized and hosted by the Computational Acoustic Modeling Laboratory in the Schulich School of Music at McGill University, together with the Centre for Interdisciplinary Research in Music Media and Technology, and is also co-sponsored by the Acoustical Society of America. Funding to support the symposium has been provided by the Social Sciences and Humanities Research Council of Canada.

The ISMA 2017 Local Organizing Committee

Gary Scavone	*Conference Chair*
Esteban Maestre	*Paper Chair*
Charalampos Saitis, Connor Kemp, Song Wang	*Paper & Proceedings Support*
Quim Llimona	*Web Support*
Negin Abaeian, Lei Fu, Wen Xiao	*General Support*
Jenine Evans, Simon Benacchio	*Volunteers*

The ISMA 2017 Scientific Committee

Noam Amir
Jose Antunes
Anders Askenfelt
Federico Avanzini
James Beauchamp
Stefan Bilbao
Xavier Boutillon
Jonas Braasch
Murray Campbell
Antoine Chaigne
James Cottingham
Jean-Pierre Dalmont
Benoit Fabre
Claudia Fritz
Francois Gautier
Joel Gilbert
Nicholas Giordano
Mico Hirschberg
Wilfried Kausel
Jean Kergomard
Malte Kob
Thomas Moore
Andrew Morrison
Jean-Francois Petiot
Thomas Rossing
Augusto Sarti
David Sharp
Andrey da Silva
Julius Smith
Tamara Smyth
Christophe Vergez
Maarten van Walstijn
Chris Waltham
Joe Wolfe
James Woodhouse
Shigeru Yoshikawa

Best Student Paper Awards

We are pleased to announce that the two Best Student Paper Awards of the 2017 International Symposium on Musical Acoustics are awarded to:

- "Optical Interferometry for Transient Deformation Analysis" by *Tatiana Statsenko, Vasileios Chatziioannou and Wilfried Kausel*
- "Relationships between Abdominal and Around-Lip Muscle Activities and Acoustic Features when Playing the Trumpet" by *Megumi Satou, Tetsuro Kitahara, Hiroko Terasawa and Masaki Matsubara*

These papers were selected from many outstanding submissions based on comments by the paper reviewers, in consultation with the keynote speakers. The authors of these papers are invited to submit extended versions of these papers for publication in *Acta Acustica united with Acustica*.

Best Poster Presentation Awards

We are pleased to announce that two Best Poster Presentation Awards of the 2017 International Symposium on Musical Acoustics are awarded to:

- "Who Does Control the Attack Transients of the Recorder, the Musician or the Instrument Maker?" by *Augustin Ernoult and Benoît Fabre* (presented by Augustin Ernoult)
- "Study of Magnetic Pickups for Electric Guitars" by *Léo Guadagnin, Bertrand Lihoreau, Pierrick Lotton, Emmanuel Brasseur* (presented by Léo Guadagnin)

These posters were selected by a group of judges in attendance at the conference.

CONFERENCE PROGRAM

Monday, 19 June 2017

Keynote 1

Oral Session 1 — Woodwinds, Reeds (I)

Oral Session 2 — Woodwinds, Reeds (II)

Oral Session 3 — Winds, General (I)

Oral Session 4 — Winds, General (II)

Poster Session 1 — Wind Instruments, Piano, Harpsichord

Tuesday, 20 June 2017

Keynote 2

Oral Session 5 — Percussion, Deformation

Oral Session 6 — Modeling, Synthesis (I)

Oral Session 7 — Plucked Strings

Oral Session 8 — Bowed Strings

Poster Session 2: Percussion, Strings

Wednesday, 21 June 2017

Keynote 3

Oral Session 9 — Modeling, Synthesis (II)

Oral Session 10 — Modeling, Synthesis (III)

Oral Session 11 — Brass Instruments

Oral Session 12 — Contemporary Instruments

Poster Session 3 — Brass, Timbre, Perception, Voice

Thursday, 22 June 2017

Oral Session 13 — Timbre, Perception (I)

Oral Session 14 — Timbre, Perception (II)

KEYNOTE ADDRESS

Harmonic Stepped Waveguides and Their Application to Music

Jean-Pierre Dalmont[†]

Laboratoire d'Acoustique de l'Université du Maine (UMR CNRS 6613), 72085 Le Mans, France
[†]Jean-Pierre.Dalmont@univ-lemans.fr

ABSTRACT

Stepped waveguides made with cylinders of equal length have peculiar properties. Some of them have harmonic eigenfrequencies and can be used as musical instruments. The present talk will explore the large number of solutions and discuss how they can be used as musical instruments.

Reed Chamber Resonances and Reed Tongue Vibrational Modes in Free Reed Excitation

James Cottingham

Physics Dept., Coe College, USA
jcotting@coe.edu

ABSTRACT

Recent studies of the initial transients in Western free reed instruments such as the accordion, harmonica, and harmonium, have concentrated in two areas. The first of these is the significance of higher transverse and torsional modes in initiating reed tongue vibration. The second regards effects of the reed chamber resonances. The motion of the reed tongue in early stages of the initial transient has been studied in some detail using reeds from a reed organ. Oscillation waveforms were obtained using a laser vibrometer system, variable impedance transducer proximity sensors, and high speed video with tracking software. Spectra derived from the vibrational waveforms in each stage have been analyzed, showing that the second transverse mode and the first torsional mode are both observed in the transient, and that the torsional mode appears to be especially significant in the earlier stages of oscillation. Comparison of reed tongues of different design have shown that curving the reed tongue to enhance the role of the torsional mode in the initial excitation may somewhat alleviate the problem of slow speech that typically affects larger, lower-pitched reeds. Although reed chamber resonances can affect tone quality and may have some effect on pitch, the cavity volumes are small and the chamber resonances have high frequencies, so that effects on reed excitation usually tend to be small. However, for higher pitched reeds in the accordion or harmonica, a resonance of the reed chamber can be close to the vibration frequency of the reed tongue. In this case the cavity air vibration can interfere with tongue vibration, inhibiting or possibly preventing the sounding of the reed. Recent experiments have verified that reed chamber resonances can indeed have this effect. Builders typically attempt to alleviate this situation by modification of the reed chamber. In the harmonica, skilled players may be able to overcome this difficulty by appropriate changes in vocal tract resonances.

Objective Comparisons between an Ancient and a Modern Clarinet

Marion Volpe,[1] *Philippe Guillemain,*[1†] *Pierre-André Taillard,*[2] *Jean Kergomard*[1]

[1]Centre Nationnal de la Recherche Scientifique
[2]FHNW, Schola Cantorum Basiliensis
[†]guillemain@lma.cnrs-mrs.fr

ABSTRACT

Impedance measurements and digital simulation of the functioning are used to compare two clarinets whose conception and design are separated by two centuries: a copy of an ancient model by Heinrich Grenser and a modern clarinet (Prestige, from Buffet Crampon). Both instruments were manufactured approximately 30 years ago. The full first register is studied for both instruments. Input impedance measurements have been made in the same conditions and differences between the features of the peaks are studied: amplitudes, quality factors, frequencies, harmonicity and cut off frequency due to the network of open toneholes. These impedances are then used, thanks to the modal estimation of their resonances in a real-time, compatible model of sound production based on a classical physical model of the functioning, including reed dynamics, reed flow and shock between the reed and the table. Mouthpiece pressure and radiated sound are computed with respect to two varying control parameters sampled on a grid: the blowing pressure and a parameter related to the effort exerted by the lip on the reed. From these sounds several bidimensional maps have been computed, describing either the functioning (regime, intonation) or the timbre (attack transient duration, spectral centroid, power). Finally, we try to make a link between the differences observed note by note on the maps and those observed in the input impedances.

Static Stiffness Evaluation and Aging Behaviour of Played Alto Saxophone Reeds

Connor Kemp,[†] *Gary Scavone*

McGill University

[†]connor.kemp@mail.mcgill.ca

ABSTRACT

There are several well known inconsistencies that are frequently observed in woodwind reeds, specifically their variable stiffness (as rated by the manufacturer) and their changing behaviour with time. The present ongoing study is being conducted to describe differences in static stiffness of as-manufactured reeds, each reed with the same geometry, hardness and stiffness rating. The sample set consists of 8 reeds to be measured and tracked. The static measurements are being taken at 6 individual points along the tip and vamp of the reed to increase spatial resolution and capture potential asymmetric effects. Subsequently, the functional life of these reeds will be tracked over the course of regular playing by a professional musician. In this way, initial differences in stiffness are being compared quantitatively using load-deflection testing and qualitatively through stiffness ranking by the musician. These initial values are tracked at regular time intervals over the life of the reed, with the end-of-life determined by the musician. Initial results show that initial static stiffness measurements differ between the reeds, despite an identical rating from the manufacturer. Differences are also observed between spatial positions along the tip. Reeds initially identified as being substantially more stiff than the others also exhibit increased objective stiffness in the load-deflection test. In general, there appears to be a break-in period where the reeds continue to absorb moisture (as measured by reed mass), before stabilizing roughly one month into the experiment. Static stiffness measurements also vary with time, although changing reed masses alone do not account for the difference observed. These findings support the long standing notion of high variability within a box of purchased reeds. The findings of this study will help musicians better understand the reeds they purchase and potentially help manufacturers find ways to better classify and/or pre-condition reeds before packaging.

Strain to Displacement Calibration for Single-Reeds Using a High-Speed Camera

Montserrat Pàmies-Vilà,[1†] *Alex Hofmann,*[1] *Vasileios Chatziioannou*[1]

[1]Department of Music Acoustics (IWK), University of Music and Performing Arts Vienna, Austria
[†]pamies-vila@mdw.ac.at

ABSTRACT

Measurements of reed bending on the clarinet and saxophone have been used to qualitatively analyse the reed behaviour and its interaction with the player's tongue. A similar measurement technique is now aimed at obtaining quantitative information regarding the opening at the reed tip. This paper presents an experimental procedure in which the reed bending and reed-tip opening are simultaneously measured while the reed is artificially excited via a pressure-controlled blowing system. The reed bending is measured in terms of its surface strain and the reed-tip opening is obtained from processed high-speed camera images. For different blowing pressures and lip forces, a comparison of the measured signals indicates a linear strain-to-displacement relationship. Using a linear calibration model, the strain signal has been satisfactorily converted into reed displacement. This procedure opens up the possibility to measure the reed-tip opening under real playing conditions.

1. INTRODUCTION

Sound production in single-reed woodwinds is often characterised by the coupling between an excitation mechanism (reed-mouthpiece) and a linear resonator (bore) [1]. The reed-mouthpiece behaviour in the steady state is commonly described in terms of pressure difference, reed displacement, and volume flow [2]. Regarding the effect of the player's actions upon the excitation mechanism, the vocal tract and the tongue-reed interaction have been the focus of some recent studies [3, 4, 5]. The tongue-reed contact may determine the characteristics of the transient phenomena taking place during articulation of tones.

In order to analyse the reed behaviour during transients, Hofmann *et al*. developed a measurement setup based on the deformation of a strain gauge attached to the reed surface inside the mouthpiece [6]. This strain measurement varies according to the reed bending, leading to qualitative knowledge about the reed motion and its contact with the tongue, which was used to analyse playing techniques [4, 7]. The next step was to determine whether this setup could provide quantitative information regarding the reed-tip opening, a magnitude unavailable to previous physics-based analysis attempts [8, 9, 10]. To correlate strain-gauge signals to reed-tip displacement, static and dynamic tests have been previously performed [6, 11]. The static measurement tested two different lip configurations, whereas the dynamic measurement tested different vibrating frequencies. Both studies showed evidence of a linear relationship between the magnitudes and pointed out the need to corroborate this relationship under more realistic blowing conditions.

This study analyses the use of strain-gauge technology to obtain a calibrated measurement of the reed-tip opening during playing. A uniform displacement across the reed tip is assumed, as the reed is considered to vibrate as a clamped bar that operates below its first resonance frequency, hence the appearance of higher modes is avoided (as verified in [12]). The reed is excited under pressure-controlled artificial blowing conditions and it is simultaneously measured by the strain gauge and tracked by a high-speed camera (HS-camera). These two measurements are compared to establish the strain signal calibration (*i.e*. to transform strain measurements into reed-tip opening). Such a method would allow measurements of the reed-tip displacement in real playing conditions, where the reed vibration takes place inside the musician's mouth, hence no direct visual feedback is available.

2. EQUIPMENT AND METHODS

An artificial blowing setup is used to obtain single-reed vibrations in an artificial mouth (transparent plexiglass, 170 cm^3) allowing independent control of the blowing pressure and the lip force. The pressure in the artificial mouth is controlled by a proportional valve (PVQ33-5G-23 by SMC) in a feedback loop. The artificial lip is made out of soft rubber and its position and force can be adjusted. The experiments are performed on a synthetic clarinet reed (German cut, strength 3, by Légère) fixed on a Bb-clarinet German mouthpiece (W4 by Wurlitzer). The mouthpiece is coupled to a copper cylinder (length: 55 cm, diameter: 18 mm) representing the clarinet bore. The bending measurement is achieved by a strain gauge attached to the reed [6], and the reed motion is optically captured via a high-speed area-scan camera (V12.1 by Phantom). The strain-gauge signal, the blowing pressure, and the lip force are simultaneously recorded at 50 kHz using an acquisition platform (LabView by National Instruments). The camera images are recorded at 20000 frames per second (Camera Control Software by Phantom).

2.1. Strain-gauge reed setup

A strain-gauge sensor (length: 2 mm) is attached on the surface of a synthetic single-reed, inside the mouthpiece. This sensor gives a voltage related to the differences in compressive stress on the reed surface due to changes in the reed bending while it is vibrating. The sensor is placed between the lip position and the reed tip. In this study the sensor is at 11 mm from the reed

tip whereas the artificial lip touches the reed at around 14 mm from the reed tip.

2.2. High-speed camera image processing

A high-speed camera recording followed by digital image processing allows to analyse the vibrations of musical instruments. In this study, the reed and the mouthpiece are placed between a light source and the HS-camera, as schematised in Fig. 1. By doing so, a sharp contrast between the white background and the object –necessary for the image processing– is guaranteed. Magnification and fine focusing are required to assure the detection of single-reed vibrations, which are of the order of one millimetre, hence an optical lens and tube (length: 10 cm) are used. The conversion from pixels to millimetres is obtained by recording a 1-mm grid as a reference. In this study, the image resolution is 63 pixels per mm. The reed motion is then obtained using the following image-processing method.

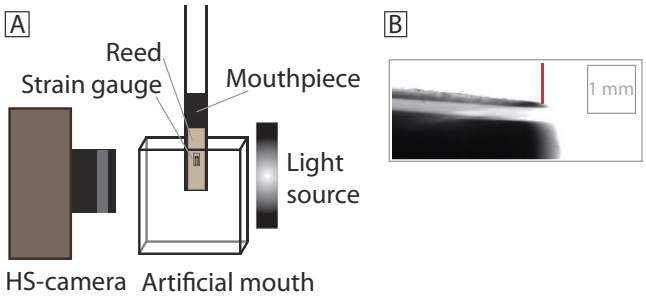


Figure 1. A: Scheme, seen from the top, of the high-speed recording setup. B: One of the recorded images. The red line indicates the vertical line tracked in the image processing.

We aim at measuring the reed-tip opening, *i.e.* the distance between the mouthpiece and the reed tip. The visualisation of the inner surface of the reed in the recorded images is more likely to be affected by shadows than the outer one. Considering that the reed material does not compress in the direction of the tip motion, the vibration can be measured by tracking the vertical motion of the outer side of the reed tip (red line in Fig. 1). An image processing approach has been designed in MATLAB (MathWorks Inc.) to automatically extract the reed tip motion from the recordings, in which the upper limit of the reed tip is detected according to a colour-threshold that discerns the dark objects from the white background. Afterwards, a quantitative measurement of the reed-tip opening is obtained by setting it to zero at the instant of the tip closure.

The setup has been checked to remain constant between image samples. First, the relative position between the camera and the reed-mouthpiece system does not vary, hence there is no need to consider the mouthpiece position. Rotation of the reed-mouthpiece system with respect to the camera view has been proven not to affect the output for angles up to 10°. Finally, the possible error introduced by changes in the shadow boundary during vibration does not present a significant impact on the recorded signals.

2.3. Experimental design

The measurements shown in this paper are performed with a strain-gauge equipped reed clamped to a mouthpiece. The reed-mouthpiece system is excited by controlling the blowing pressure and the lip force in the artificial mouth. The position of both the artificial lip and the reed is kept constant relative to the position of the mouthpiece along the experiment. The variables of the experiment are the blowing pressure and the lip force. For every blowing pressure setting, the lip force (and hence the equilibrium opening) is adjusted so as to obtain oscillations at the fundamental frequency (148 Hz). Four measurements are performed, where the blowing pressure ranges from 4.5 kPa to 7 kPa, and the lip force ranges from 0.5 N to 10 N, as shown in the legend of Fig. 2. An additional strain-gauge measurement is performed under real playing conditions with the same mouthpiece, reed, and cylinder.

3. RESULTS

This section presents a comparison between reed-surface strain and reed-tip opening signals and proposes a strain to tip-opening calibration procedure.

3.1. The linear correlation

Previous studies showed a linear trend between strain signals and reed-tip displacement in both static and dynamic regimes [6, 11]. We aim at validating such a linear correlation by comparing the strain signal to the reed-tip opening obtained from high-speed images in realistic artificial blowing conditions.

Figure 2A shows the original strain $\varepsilon(t)$ and reed-tip opening $h(t)$ signals for one measurement (blowing pressure: 6 kPa, lip force: 1.2 N, label number 3). The strain values are mirrored across the x-axis, so that bigger strain values correspond to reed opening and smaller to reed closing. The two signals are capable of describing the reed-tip motion, which has a closed phase ($h = 0$ mm) and an open phase. These signals are qualitatively comparable to a real playing signal (Fig. 3 bottom; obtained with the same reed and mouthpiece). Then, Fig. 2B plots the tip opening vs. strain for four measurements differing in blowing pressure and lip force. For all measurements, the previously observed linear relationship between strain and reed-tip displacement is confirmed.

A least-squares fit is then formulated as $h = a \cdot \varepsilon + b$, where the parameters a and b express an amplitude factor and an offset factor respectively (straight lines in Fig. 2B). The four linear fits in Fig. 2B present the same slope but a different strain offset, due to the effect of the lip force on the equilibrium position of the reed, and thus on the static bending of the reed. The linear relationship can be corroborated by considering the overlapped signals in Fig. 2C, corresponding to the reed-tip opening obtained from the camera signal and the strain signal after using the linear fit. The fitted signal (blue) accurately describes the reed-tip motion during the whole period. This result suggests that the a and b parameters obtained by performing one reference measurement can then be used to

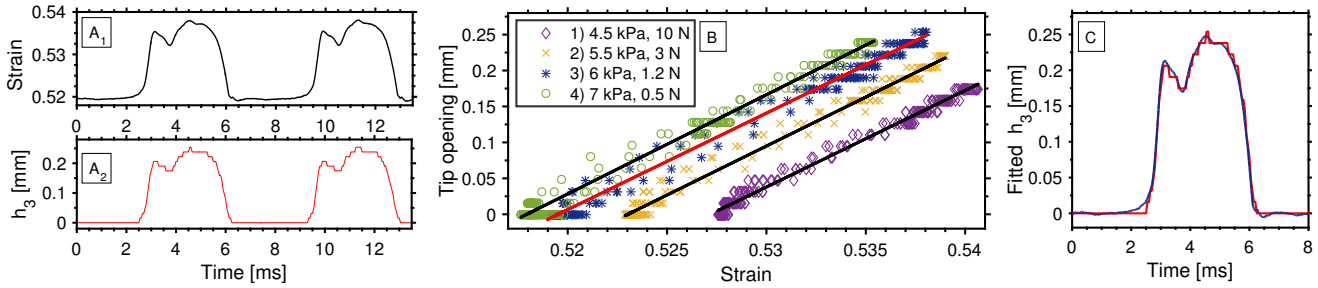


Figure 2. A: Original strain (ε) and its corresponding reed-tip opening (h_3) seen by the HS-camera for the measurement at 6 kPa and 1.2 N (label number 3). B: Reed-tip opening vs. reed strain at different pressures (kPa) and lip forces (N) and their linear fitting model; model 3 (red) is taken as the reference for the calibrated signals in Fig. 3. C: Reed-tip opening from the camera (red) over the fitted strain signal (blue) for measurement number 3.

calibrate other strain signals recorded by the same strain-gauge sensor-reed, as is next analysed.

3.2. The calibration procedure

Using the described setup, the strain at the reed surface and the reed-tip opening have been simultaneously recorded for calibration purposes. The calibration process is realised in four steps: synchronisation of the two signals, extraction of the a and b calibration parameters using a reference signal (a_3 and b_3 for measurement number 3, chosen as the one with the largest displacement amplitude), fitting the desired strain signals (measurements i = 1, 2, 4) by applying the linear fit with the reference parameters $h_i(t) = a_3 \cdot \varepsilon_i(t) + b_3$, and adjustment of the reed-tip opening offset, to attain the minimal opening at $h = 0$ mm.

The plots at the left of Fig. 3 show the three reed-tip opening signals $h_i(t)$ obtained after applying the reference fitting model, plotted over the tip opening seen by the HS-camera. As expected from the constant slope and the varying offset in Fig. 2B, at this stage of the calibration process we obtain an accurate relative displacement but not an absolute one, that is why the offset needs to be adjusted so as to obtain $h = 0$ mm at the position of the reed tip touching the mouthpiece.

The offset adjustment requires a measurement of the strain at the reed closure. When the reed tip closes against the mouthpiece at every period, beating occurs. In this case, the recorded strain signal can be used to determine the reed-tip closure. Hence, for the present study, the reed closure is determined at the instant of minimal strain (*e.g.* for h_1 in Fig. 3 this offset is equal to 0.109 mm). Then, by subtracting this value from the fitted signals, the calibrated signals result in a satisfactory matching to the camera signals, as it is shown at the right of Fig. 3.

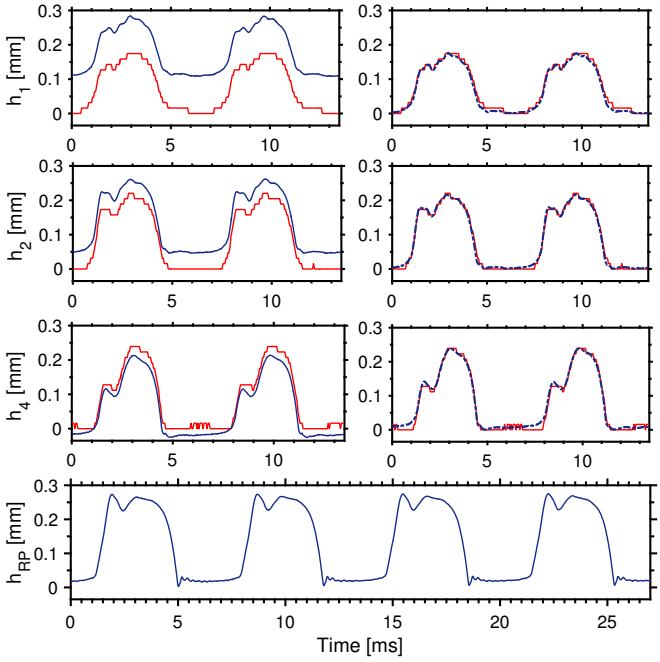


Figure 3. Reed-tip openings obtained after fitting the strain signals (blue) and their corresponding HS-camera signal (red), for measurements 1, 2 and 4, using the fitting model 3 as a reference. At the right, the calibrated signals (blue) are obtained after adjusting the reed-mouthpiece closure ($h = 0$ mm). At the bottom, a calibrated real-playing strain signal, which cannot be compared with a camera signal.

4. DISCUSSION

This paper analyses the use of strain-gauge technology to measure the reed motion and its capacity to provide a calibrated measurement of the reed-tip opening during playing. The use of an artificial blowing setup allows to excite the reed and to perform a high-speed recording of the tip opening. This setup is used to compare the strain at the reed surface with the reed-tip opening, showing a linear relationship between the two magnitudes.

The experimental setup has been improved with respect to [6, 11] by exciting the reed in an artificial mouth and recording its motion using a HS-camera. The artificial blowing conditions allow to reproduce the role of the player's lower lip, which controls the reed equilibrium position, its vibrating length and the amount of damping applied to the reed [1], and to control the pressure in the artificial mouth. By comparing the real playing signal (bottom of Fig. 3) to the artificially

obtained signals, it can be deduced that the artificial blowing machine generates reed oscillations that are qualitatively equivalent to those obtained under real playing conditions and may hence be used to analyse the presented sensing technique.

The experimental design consisting of simultaneously measuring the reed strain and tip opening via a HS-camera results in a satisfactory comparison framework. Although the tip-opening signal obtained from the HS-camera images presents a significant level of quantisation, this quantised signal is only used for calibration and comparison, but will not be used as a measurement tool once the strain gauges are calibrated.

Different blowing pressures and lip forces are tested for one lip position and one tone. The obtained linear fits present the same slope (parameter a) but differ in strain offset, as the reed is preconstrained by the presence of the lip. This is the main drawback of such a setup, implying the need of using the instant of reed closure as a reference strain value. A similar consideration was also pointed out for the optical measurements proposed in [13]. Nevertheless, a minimum in the strain signal might not correspond to the reed closure for different playing configurations. For a real-playing calibrated signal (Fig. 3 bottom) it has been observed that when the reed closes against the mouthpiece (*e.g.* at $t = 5$ ms), the body of the reed experiences a damped vibration at low amplitude. Furthermore, it cannot be guaranteed that the reed will always beat against the mouthpiece lay. Therefore, in the absence of a HS-camera, a complimentary measurement of the reed closure is needed, which can be another instant where beating is forced to occur under the same configuration of the reed-mouthpiece-lip system.

After adjustment of the fitted-strain offset, the calibration procedure results in a satisfactory extraction of the reed-tip opening from the strain signals within the range of this experiment. Future experiments may include an extension of the calibration procedure to the overblown register as well as to transient phenomena. The aim is to calibrate signals in real playing conditions using a reference signal obtained in artificial blowing conditions. Therefore, the fact that under real playing conditions the lip force may vary should also be taken into account. Furthermore, a more accurate determination of the reed closure in both real and artificially obtained signals would allow to better readjust the offset of the fitted signals.

Single reeds equipped with a strain-gauge sensor may be used for obtaining the reed-tip displacement under real playing conditions. As already shown in [7, 12], mounting such strain-gauge sensors on single reeds does not disrupt the performance of woodwind players. This information about the reed displacement is of paramount importance for physics-based synthesis/analysis applications regarding the characterisation of the excitation mechanism in woodwind instruments.

5. ACKNOWLEDGEMENTS

The authors acknowledge Alexander Mayer for technical support. This research is supported by the Austrian Science Fund (FWF): P28655-N32.

REFERENCES

[1] C. J. Nederveen, *Acoustical Aspects of Woodwind Instruments*. Northern Illinois University Press, 1969.

[2] J.-P. Dalmont, J. Gilbert, and S. Ollivier, "Nonlinear characteristics of single-reed instruments: Quasistatic volume flow and reed opening measurements," *JASA*, vol. 114, no. 4, pp. 2253–2262, 2003.

[3] G. P. Scavone, A. Lefebvre, and A. R. da Silva, "Measurement of vocal-tract influence during saxophone performance," *JASA*, vol. 123, no. 4, pp. 2391–2400, 2008.

[4] V. Chatziioannou and A. Hofmann, "Physics-based analysis of articulatory player actions in single-reed woodwind instruments," *Acta Acust united Ac*, vol. 101, no. 2, pp. 292–299, 2015.

[5] W. Li, A. Almeida, J. Smith, and J. Wolfe, "The effect of blowing pressure, lip force and tonguing on transients: A study using a clarinet-playing machine," *JASA*, vol. 140, no. 2, pp. 1089–1100, 2016.

[6] A. Hofmann, V. Chatziioannou, M. Weilguni, W. Goebl, and W. Kausel, "Measurement setup for articulatory transient differences in woodwind performance," in *Meetings on Acoustics*, vol. 19, no. 1, 2013.

[7] A. Hofmann and W. Goebl, "Production and perception of legato, portato, and staccato articulation in saxophone playing," *Frontiers in Psychology*, vol. 5, p. 690, 2014.

[8] G. P. Scavone, "An acoustic analysis of single-reed woodwind instruments with an emphasis on design and performance issues and digital waveguide modeling techniques," Ph.D. dissertation, Stanford University, 1997.

[9] P. Guillemain, C. Vergez, D. Ferrand, and A. Farcy, "An instrumented saxophone mouthpiece and its use to understand how an experienced musician plays," *Acta Acust united Ac*, vol. 96, no. 4, pp. 622–634, 2010.

[10] V. Chatziioannou and M. van Walstijn, "Estimation of clarinet reed parameters by inverse modelling," *Acta Acust united Ac*, vol. 98, no. 4, pp. 629–639, 2012.

[11] M. Pàmies-Vilà, A. Mayer, A. Hofmann, and V. Chatziioannou, "Measurement of dynamic bending and displacement of clarinet reeds," in *7th Congress of Alps-Adria Acoustics Association*, Ljubljana, Slovenia, 2016.

[12] V. Chatziioannou, A. Hofmann, A. Mayer, and T. Statsenko, "Influence of strain-gauge sensors on the vibrational behaviour of single reeds," in *22nd International Congress on Acoustics*, Buenos Aires, Argentina, 2016.

[13] A. Muñoz, B. Gazengel, J.-P. Dalmont, and E. Conan, "Estimation of saxophone reed parameters during playing," *JASA*, vol. 139, no. 5, pp. 2754–2765, 2016.

A Timbral and Musical Performance Analysis of Saxophone Multiphonics Morphings

Martin Proscia,[1] *Pablo E. Riera,*[1,2] *Manuel C. Eguia*[1†]

[1]Laboratorio de Acústica y Percepción Sonora. Escuela Universitaria de Artes. Universidad Nacional de Quilmes.
[2]Laboratorio de Dinámica Sensomotora.. Departamento de Ciencia y Tecnología. Universidad Nacional de Quilmes.
†meguia@unq.edu.ar

ABSTRACT

This work presents a study of the multiphonics of the saxophone from the points of view of timbral analysis and musical performance. A multiphonics categorization into four groups was previously proposed, starting from a musical analysis and further validated by psychophysical experiments and spectral analysis tools. This classification served to identify relevant timbral attributes of the studied multiphonics, such as modulation frequency, spectral center of gravity and roughness. The present work extends the analysis to the spectral morphing of multiphonics, generated by the continuous modification of the playing parameters. These morphings are studied in a comparative view as trajectories in a timbral space, based on the aforementioned attributes, and in terms of their feasibility in the instrument. From this comparative analysis a dynamic characterization of the proposed multiphonic groups is derived.

1. INTRODUCTION

The multiphonic sonority constitutes an idiomatic part of the language of contemporary music. Unlike in the past, when their use had a decorative or accessory character, nowadays this type of sounds has gained a structural space in the musical composition. Many pieces are constructed "from" the multiphonics, starting from their mixed sonority of timbre and harmony as a poetic-musical trigger element. In this context, the saxophone provides a wide range of multiphonics that have been approached from multiple perspectives[1,2,3,4].

In a previous work, we have studied a timbral categorization for the multiphonic tones in the saxophone from a musical, psychophysical and acoustical point of view [5]. This work served to establish a space defined by acoustic parameters that allowed to assign each of the proposed categories a given region in a two-dimensional space. Although this study was performed on static sounds, it was suggested that each multiphonic is rather a dynamical structure capable of traversing different stages. As a consequence, it is possible to think of timbral modulation trajectories, or morphings, between the different sonorities, even for the same multiphonic.

The present work is focused on the study of multiphonics morphings. For this purpose, a set of morphing trajectories, presenting significant changes both at the spectrum level and in other timbral attributes, will be studied. These trajectories, once defined in a two-dimensional space determined by the acoustical parameters, will display the timbral evolution of a multiphonic across the different sonorities. In this way, it is sought to establish a systematization for multiphonic morphing that contemplates the plurality of aspects involved.

2. MULTIPHONICS MORPHINGS

We can say, roughly, that *morphing* is the process during which a sound progresses from one timbre to another. Since timbre is multidimensional in nature, a morphing process does not take place along a single dimension, and entails the modification of many sound parameters in an interdependent way: pitch, dynamic envelope, spectral envelope, degree of roughness, harmony, etc. It was proposed [6] that a morphing process, in order to be perceived as such, must meet three basic requirements: (a) gradualness, which refers to the degree of smoothness of the timbral change, taking into account the gradualness of the timbral differences between the different states, the temporal separation, the duration and the rate of change; (b) directionality, which points out to the relationship between the starting timbre, the arrival timbre and the type of trajectory between both; and (c) time continuity, which reflects how the intermediate states of this modulation take place in time, and whether this occurs continuously or discontinuously.

We will also establish certain guidelines for our study. In the first place, we will be focused on the multiphonics of the alto saxophone as a solo instrument, *i. e.*, we will leave aside the morphing processes that include more than one instrument. In the second place, we will restrict our study to those processes that can be realized in a single gesture (a single *fiato*) discarding those that imply the need to attack the sound more than once[1] . We will therefore undertake the study of multiphonics morphings in the alto saxophone that can be performed by a single instrument, in a single phase and that meet the requirement of being gradual, directed and continuous.

Seven multiphonics of the alto saxophone were selected for the analysis. This choice was made taken into account each particular attribute of the multiphonic, not only based on its sonority but also on the possibilities of tímbrical and dynamical evolution. In order to qualify these attributes, the analysis carried out in the typology previously proposed was considered [5]. This typology consists of four sets of sonorities, or classes, that can be summarized here: (a) *Bichords:* multiphonics with a stable sonority, and a first interval that is around a 3rd, with a velvety quality of surface, and dynamics ranging between *pp* and *mf*, (b) *Complex multiphonics*: have a relatively stable sonority, with emphasis on the higher

[1] We do not discard here the use of circular breathing, as long as we understand it as a single gesture, a continuous sonority that is not interrupted in its evolution by breathing or silence.

harmonics, a high degree of dissonance, a compact and rough surface quality, dynamics between *mf* and *ff* and a first interval that is around a 9th. (c) *Multiharmonics:* are unstable multiphonics from the point of view of their production, they do not present activity in the upper part of the spectrum and their dynamics range between *ppp* and *mp*; also their surface quality is remarkably smooth, with a first interval corresponding to a 7th or a 9th. (d) *Tremolos:* present a homogeneous distribution of the spectral energy, with a surface quality marked by the beat and type of grain typical of a tremolo sound, and a degree of consonance relative to the tuning of its first interval that is close to the octave.

These classes have been proposed for stationary fragments (a three-second cut was made in which the multiphonic was stabilized) from a reduced listening[2] and further validated by a psychophysical study [5]. However, each fingering that produces a multiphonic determines a structure of pitches that the performer must stabilize from the intonation and use of the vocal tract. From modifications in the mouth, air column and intonation, it is possible to obtain different sonorities with the same fingering. In this way, we can say that a multiphonic must be considered a dynamic structure capable of traversing different stages [7]. From a signal analysis point of view, the saxophone multiphonics could be considered as modulated tones, hence having intrinsic property, which is the modulation frequency. This modulation is produced by the superposition of the different pitches and their nonlinear interaction in the reed [8], and is a determinant parameter of the timbre.

A sound morphing is a continuous phenomenon that presents intermediate stages, as well as a visual morphing [9], so that each of the static classes proposed in our typology can be either the starting point, the point of arrival or some of the intermediate points in a multiphonic morphing. In order to perform a multiphonic morphing on the saxophone, the interpreter must control the pitch changes in order to have continuity in each of the intermediate stages.

For the present work we selected representative multiphonics spanning the four proposed classes (see Figure 1). Multiphonics 1 and 4 when played *pp* behave like a Bichord, while as the intensity is increased a third pitch becomes audible and a general timbral change is observed. Multiphonics 5 and 6, belonging to the set of Complex Multiphonics, present a greater degree of roughness when reaching the *mf*, which increases when reaching *ff*. Sounds 2, 3 and 7, although not exclusive to tremolo sonority, have a greater spectral activity, since they include, in contrast to previous cases, a partial sound in the bass record of the instrument. At the same time, as they grow in dynamics they present different evolutions of internal beat and roughness.

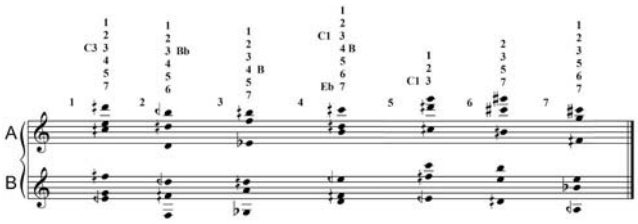


Figure 1. Musical notation for the seven multiphonic alto saxophone under study. Staff A displays the notes transposed to Eb and the digitation that produces that multiphonic. In staff B the pitches are presented in concert pitch.

All the selected multiphonics, starting from *pp* present a sonority close to the Multiharmonic type, while as they evolve spectrally they approach the aforementioned classes. In musical notation (see Figure 2) all the examples start from *pp* and go to *ff*. The phonation of the multiphonic is maintained during the entire fiato, but the partials rise as the intensity of the blow increases. For that reason, we decided to add to the traditional notation a crescendo colored on the score, giving the notion that the number of different pitches increases along with the intensity. The linkage between intensity and the number of different pitches is also apparent when inspecting the spectrogram of the multiphonic morphing, as displayed in Figure 3.

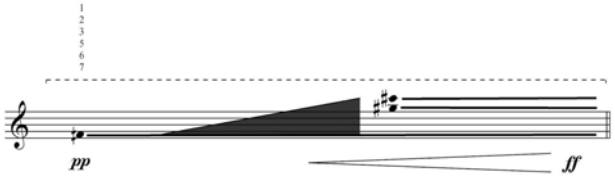


Figure 2. Musical notation of a multiphonics morphing (Multiphonic 7 of Figure 1). As the same fingering is used throughout the entire playing, the structure of pitches of the multiphonic is maintained. Increasing the intensity increases the number of pitches, and a transition between a smooth sonority with little spectral activity (Multiharmonic) and a sonority with a much larger mass and great roughness (Tremolo) take place.

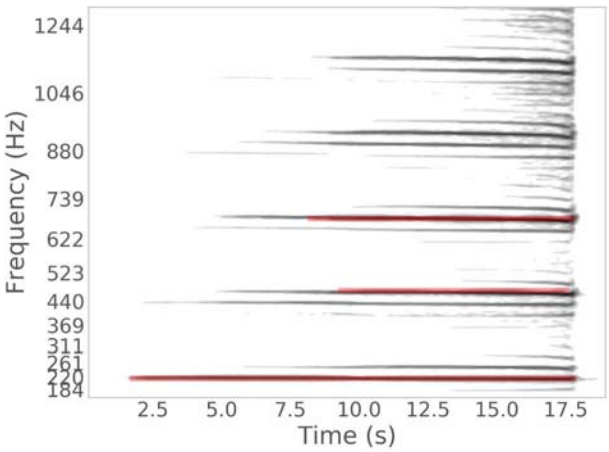


Figure 3. Spectrogram of Multiphonic 7 displayed in Figure 2 in musical notation. The evolving pitches are depicted as shaded lines and were detected by a pitch extraction algorithm [10]

[2] The reduced listening is a concept proposed by Pierre Schaeffer (Schaeffer, 2003) and is a type of isolated listening in which the auditory analysis of a sound object is made independently of the source.

3. RECORDINGS AND ANALYSIS

The multiphonic tones were played with a Selmer Super action 80 Serie III alto saxophone tuned at A = 440 Hz, with a Selmer Serie 80 C* mouthpiece and Vandorem 3½ reeds. The recordings were made using an acoustical measurement microphone (DBX TRA-M) and a Focusrite Sapphire external sound card at a sampling rate of 48 kHz with a resolution of 24 bits. The multiphonics were selected and performed by one of the authors (MP)[3]. The samples were recorded in a room with sound isolation and acoustic treatment (noise floor 19 dBA, reverberation time T60 @ 1 kHz 0.3 s). Recorded samples have an average duration of 19 seconds, minimum of 15 and maximum of 22 seconds.

In order to perform the timbral analysis of the recorded trajectories, first a set of relevant acoustical features must be selected. We will circumscribe this first study to only three parameters: spectral centroid, modulation frequency and roughness. The selected acoustical features are those that showed a higher correlation with the coordinates of a multidimensional analysis performed over a psychophysical study of discrimination of the proposed multiphonic classes [5]. Feature extraction was performed taking frames of 85 ms length, with steps of 42 ms for the roughness and centroid and 1360 ms length and steps of 680 ms for the modulation frequency. For each frame roughness was measured from spectral peaks using the method from [11]. Spectral centroid was computed as the first moment of the spectral power distribution. Modulation frequency was extracted by computing the autocorrelation of the spectrum and identifying the largest and lowest in frequency peak.

For these multiphonics morphings, the intensity, the roughness and the centroid are highly correlated. The main reason of this strong correlation is the following: as the blown intensity is increased more partial components are added to the sound, thus pushing the centroid to higher frequencies and increasing the roughness at the same time. Nonetheless, the spectral centroid is also affected by the pitch structure of each multiphonics setting, where the pitches are higher. In addition, the modulation frequency reflects the average frequency distance between partial components, which also affects the roughness.

In the following figures we display the timbre space trajectories of the multiphonics morphings. Figure 4 depicts the timbral space of roughness vs. modulation frequency. Most morphings, beside Morphing 3, start at low intensity with a low amount of roughness and as they evolve, the roughness is progressively increased (diamond markers indicate the end of the trajectories). Morphing 3 is an example of a sound where the number of partial components remains roughly constant along the trajectory although, at the same time the modulation frequency declines and the roughness displays a slight increment, as the trajectory evolves.

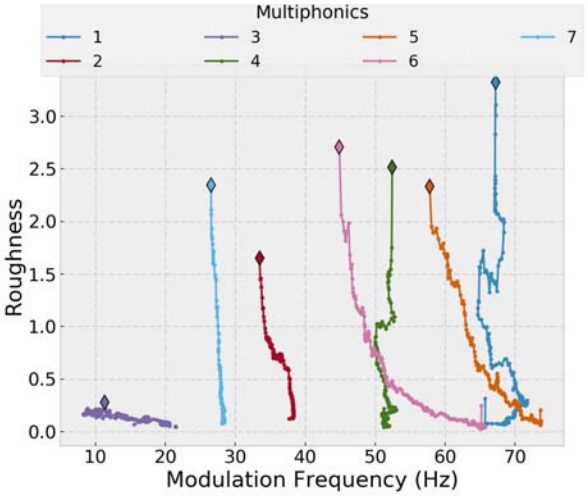


Figure 4. Trajectories of the seven multiphonics morphings in the roughness / modulation frequency space. The diamond marker indicates where the trajectory ends.

In Figure 5 we display the same trajectories unfolded in the timbral space of spectral centroid vs. modulation frequency. The trajectories are somehow similar to those exhibited in the previous figure, but Multiphonics 1 and 5 show different behaviors relative to the maximum values of roughness and spectral centroid. In this Figure we also overlay the timbral space regions that characterize the four multiphonics classes introduced before. These regions were estimated from a set of 118 categorized multiphonics, as described in [5].

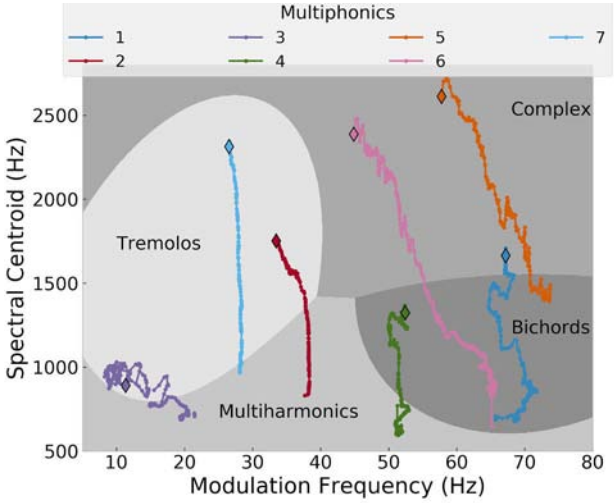


Figure 5. Trajectories of the seven multiphonics morphings in the spectral centroid / modulation frequency space. The diamond marker indicates where the trajectories end. The shaded region indicate the timbric space portion for the multiphonics of the four classes described in the text.

[3] The recordings are available at http://www.lapso.org/multiphonics.html.

We could say that, in general, as the dynamics increases, the spectral centroid rises while the modulation frequency decreases. The first effect is mainly due to the fact the more energy is imprinted to the multiphonic, the higher the stimulation of new partials sounds. As for the modulation frequency, we can conjecture that by increasing the velocity of the air column, in order to keep the sound stable, the player must loosen slightly the embochure, causing a decrease in the beat rate and therefore in the modulation frequency.

4. CONCLUSIONS

In this work, we studied saxophone multiphonics morphings from a dynamical perspective, using a previously proposed classification as anchors in timbral space. These morphings were obtained in most cases by starting from a low blown intensity and ending near a full blown intensity. This technique creates sounds that have clear trajectories in simple timbre spaces made from a small set of acoustic features. As the intensity increases, the sound not only gets louder but also richer in partial components, thus increasing the roughness and its spectral centroid. These changes entail a substantial timbre modification. In addition, the modulation frequency is not static and suffers perturbations produced by variations in the embouchure.

Timbre space trajectories studies often include many different types of acoustic features and dimensionality reduction techniques [12]. In this work we focused on a relative small set of features that capture the essential timbre properties and dimensionality reduction was not necessary.

Sound morphings are a common practice in contemporary and electronic music, however, studies of timbre space trajectories [13,14] have been done mostly with standard instrument techniques and familiar sonorities. The method presented here could be useful for the analysis of musical pieces that make use of these kind of sounds.

REFERENCES

[1] D. Kientzy, *Les sons multiples aux saxophones*. Editions Salabert, 1982.

[2] G. Netti & M. Weiss, *The techniques of saxophone playing,* Bärenreiter, 2010

[3] G.P. Scavone., A. Lefebvre, & A.R. da Silva, *Measurement of vocal-tract influence during saxophone performance.* The Journal of the Acoustical Society of America, 123(4), 2391-2400, 2008.

[4] J.M.Chen, J. Smith, & J. Wolfe. *Saxophonists tune vocal tract resonances in advanced performance techniques.* The Journal of the Acoustical Society of America, 129(1), 415-426, 2011.

[5] P. E. Riera, M. Proscia & M. C. Eguia. *A Comparative Study of Saxophone Multiphonics: Musical, Psychophysical and Spectral Analysis.* Journal of New Music Research, 43(2), 202-213.

[6] C. Mastropietro, *Caracterización de la modulación tímbrica, Aspectos de direccionalidad.* Revista Arte e Investigación, Facultad de Bellas Artes, La plata, 2008, 12(6), pp 28-33, ISSN: 1850-2334.

[7] M. Proscia, *Acercamiento al saxofón multifónico. Una perspectiva de estudio*. Revista del ISM, 1(13), pp 171-194, 2011

[8] N.H. Fletcher, *Nonlinearity, complexity, and the sounds of musical instruments*. The Journal of the Acoustical Society of America 103(5), pp 3062-3062, 1998

[9] M. Caetano & X. Rodet, *Musical instrument sound morphing guided by perceptually motivated features*. IEEE Transactions on Audio, Speech, and Language Processing, 2013, 21(8), pp 1666-1675.

[10] E. Benetos & S. Dixon, *A Shift-Invariant Latent Variable Model for Automatic Music Transcription*, Computer Music Journal, vol. 36, no. 4, pp 81-94, 2012.

[11] W. A Sethares, *Local consonance and the relationship between timbre and scale*. The Journal of the Acoustical Society of America, 94(3), pp 1218-1228. 1993

[12] G. De Poli & P. Prandoni, *Sonological models for timbre characterization.* Journal of New Music Research 26(2) pp 170-197. 1997

[13] M.A. Loureiro, H. Bastos de Paula & H. C. Yehia. "Timbre Classification Of A Single Musical Instrument." In: *Proceedings of the 5th International Conference on Music Information Retrieval*, Barcelona, Spain, October 10-14, 2004

[14] Ch. Hourdin, G. Charbonnea &, T. Moussa, *A multidimensional scaling analysis of musical instruments' time-varying spectra.* Computer Music Journal 21(2) pp 40-55. 1997

Analysis of Nonlinear Characteristics of the Clarinet Exciter Obtained via a New Measurement Method

Pierre-André Taillard,[1†] *Jean-Pierre Dalmont,*[2] *Bruno Gazengel,*[2] *Alberto Muñoz Arancón*[2]

[1]FHNW Schola Cantorum Basiliensis, CP 257, 4009 Basel, Switzerland
[2]Laboratoire d'Acoustique de l'Université du Maine (UMR CNRS 6613), 72085 Le Mans, France
[†]taillard@hispeed.ch

ABSTRACT

A new method for measuring the nonlinear characteristic of the clarinet exciter, which binds the air flow entering into the clarinet with the pressure drop (Δp) across the reed, is described. It uses a clarinet mouthpiece equipped with a reed and an artificial lip whose position ψ is controlled by a micrometer screw. The mouthpiece is connected to a bottle in which a moderate vacuum is generated at the beginning of the experiment. After a short time lapse, the opening of the reed occurs. The thermodynamics of the volume in isochoric conditions enables the calculation of the volume velocity entering the mouthpiece from the pressure measurement. 13 reeds with 10 different embouchures are measured. The measurements enabled the estimation of the equivalent aeraulic section $S(\Delta p, \psi)$. We propose a model of S as a convex function of Δp and ψ, defined as the sum of two 1D stiffening springs plus a porosity constant. The mean standard error of the model is 0.2%.

1. INTRODUCTION

Since the invention of the instrument, about 300 years ago, the clarinettists complain about the difficulty to find musically suitable reeds (reeds with a "good vibration"). Many scientific studies are devoted to this topic, for instance [1, 2, 3, 4]. Some authors measure the nonlinear characteristics of the clarinet exciter (reed+mouthpiece+lip), for instance [5]. This paper describes a new, precise measurement method of these characteristics. We are convinced that the main musical differences between clarinet reeds are related to the quasistatic aeraulic behavior. The "vibration" of the reed is probably not the major point, but the ability of the reed i) not to interfere negatively with the sound production of the clarinet, ii) to enable an efficient sound control by the lip and iii) to allow an operating blowing pressure which is comfortable for the clarinettist.

This paper is organized as follow: in Sec. 2 the setup of the measurement method is described and its thermodynamical behavior is modeled in Sec. 3. Sec. 4 illustrates an application of the method to clarinet reeds and proposes a quasistatic model of the aeraulic section as a function of the pressure drop and the lip pressure.

2. DESCRIPTION OF THE SETUP

The measurement is divided in 2 phases: 1) calibration of the diaphragm, 2) measurement of the nonlinear characteristics of the reeds. The underlaying thermodynamic problem must be solved for each phase.

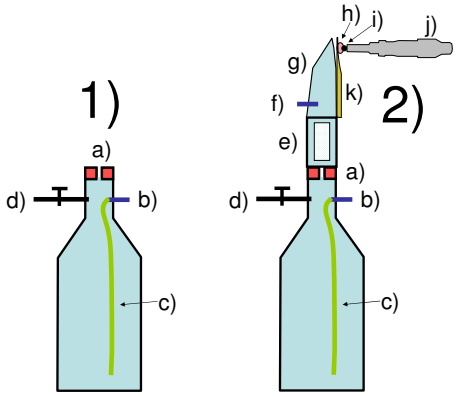


Figure 1. Schematic setup of the problems 1) and 2). a) diaphragm, b) piezoelectric pressure sensor (bottle), c) thin tube (about 1 mm inner diameter), d) cock valve, e) adaptation barrel with absorbing foam (damping the acoustic oscillations), f) piezoelectric pressure sensor (mouthpiece), g) clarinet mouthpiece, h) artificial lip (silicon 10 mm thick), i) steel beam (diameter 3 mm) glued to the artificial lip, j) micrometer screw (controlling the position ψ of the artificial lip), k) clarinet reed

2.1. Problem 1): discharge through a diaphragm

A diaphragm is connected to a hermetically closed, rigid volume (bottle, 3.178 liters, well isolated thermally and opaque to the light). A moderate vacuum (about 15-20 kPa) is generated in the bottle trough a cock valve at the beginning of the experiment (while the diaphragm is closed). After about 1 minute (allowing the observation of the heat exchange with the bottle), the diaphragm is quickly opened and the discharge starts (the air comes back into the bottle). The pressure in the bottle is measured with a piezoelectric sensor (Endevco). See Fig. 1.

The purpose is to determine the effective aeraulic section of the diaphragm S_{dia} from the pressure measurement (via the computation of temperature and flow rate).

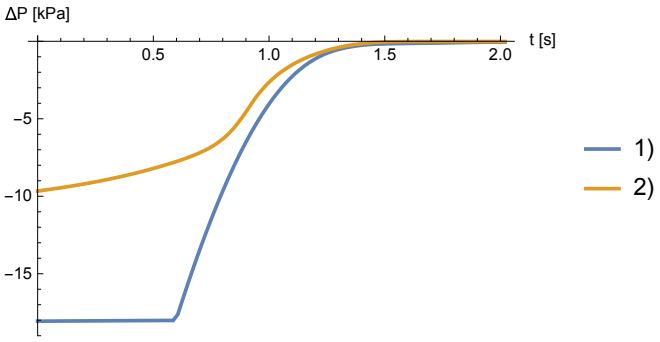


Figure 2. Typical measurements of ΔP during the discharge (zoom) for Problems 1) and 2): 1) diaphragm of 1.5mm diameter. 2) reed #J03 with embouchure $\psi = 7$ [arbitrary units]

2.2. Problem 2): discharge through a varying aeraulic section

The setup of Problem 1) is completed by a clarinet mouthpiece and a reed. An artificial lip (silicon) compresses the reed against the lay of the mouthpiece. Its position (normal to the table of the reed and denoted as ψ) is controlled by a micrometer screw. The reed is not moistened before measurement, in order to avoid a bias in the measurement due to drying.

The purpose is to determine the effective aeraulic section of the channel (slit between reed and mouthpiece) in quasistatic conditions from the pressure measurement inside the mouthpiece p and the aeraulic section of the diaphragm S_{dia} calibrated with Problem 1). The measurement of the pressure in the bottle serves only as a control for the computations. This way the delicate problem of pairing between sensors can be avoided and the precision and the reproducibility of the measurements is increased.

Typical discharge measurements for problems 1) and 2) are depicted on Fig. 2. The total duration of one experiment is about 90 s.

3. THERMODYNAMIC MODEL

3.1. Laws of thermodynamics

This subsection recalls some laws of thermodynamics, using mainly the traditional notations.

3.1.1. Constants and parameters

Ideal gaz constants : $R = 8.314 J/mol/K$
$\gamma = 1.4$ (for diatomic gazes), Density of air : ρ,
Pressure : P Atmospheric pressure : P_0
Pressure drop : $\Delta P = P - P_0$
Volume : V Flow rate : U Aeraulic section : S
$R\times$number of moles of gaz : $N = nR$
Absolute temperature : T Ambient temperature : T_0
Heat capacity at constant volume : $C_v = 1/(\gamma - 1)N$
Internal energy : E Work : W Heat : Q
Sample rate, time step : $f_s = 1/t_s$ (typically $f_s = 5000$ Hz)

Parameters with uppercase generally refer to values inside the bottle. In lowercase the same parameters refers to the value inside the mouthpiece.

3.1.2. Laws

Summary of thermodynamical laws: (1) Ideal gas law, (2) First principle of thermodynamics, (3) Newton's law of cooling (thermostat) and (4) Bernoulli's law:

$$\begin{aligned} P\,V &= N\,T && (1)\\ dE &= dW + dQ && (2)\\ dT_{th}/dt &= -r(T - T_0) && (3)\\ U &= \text{sign}(\Delta P)\,S\,\sqrt{2|\Delta P|/\rho} && (4) \end{aligned}$$

Remarks:
Newton's law of cooling: r is a positive constant, which has to be determined experimentally.
Bernoulli's law: valid for incompressible fluids and large ducts, comparatively to the aeraulic section S, for Reynolds numbers $Re \simeq U/(\nu\sqrt{\pi S}) > 2000$ ($\nu = 15.6 \times 10^{-6}$ for air at $25°C$). For compressible fluids the equation is approximately valid for the conditions at the output of the jet.

3.2. Isochoric model

The heat variation due to the thermostatic effect of the bottle is:

$$dQ = C_v\,dT_{th} \quad (5)$$

The work of small air volume dV leaving the jet outgoing the diaphragm with a temperature T_{jet} and a pressure P is $dW = dV\,P = dN\,T_{jet}$. On the other side, the variation of energy is $dE = 1/(\gamma - 1)((N + dN)(T + dT) - (dN\,T_{jet} + N\,T))$. Applying Eqs. 1, 2, 3 and 5, we obtain the equation of our thermodynamic model in isochoric conditions (*i.e.* at constant volume):

$$dN = \frac{dP\,V + N\,r\,(T - T_0)dt}{\gamma\,T_{jet}} \quad (6)$$

Considering an adiabatic expansion in the jet, the temperature of the air leaving the jet is $T_{jet} = T_0\,(P/P_0)^{\frac{\gamma-1}{\gamma}}$.

3.3. Discrete time scheme for Problem 1

During an experiment, the pressure in the bottle $P[t]$ is measured with time steps t_s : $P_m = P[m\,t_s]$. We have to deduce from the equations above the temperature T_m and the number of moles ($\times R$) of air in the bottle, N_m.

Initialization

At time step $m = 0$ (before generating the vacuum), the air in the bottle is at ambient temperature T_0 and atmospheric pressure P_0. Applying Eq. 1, the initial quantity of air in the bottle is: $N_0 = P_0 V/T_0$.

Iterations for $m > 0$

$$\begin{aligned} dN &= \frac{V(P_m - P_{m-1}) + N_{m-1}\, r\, t_s\, (T_{m-1} - T_0)}{\gamma\, T_{jet}} \\ N_m &= N_{m-1} + dN \\ T_m &= P_m V / N_m \\ U_m &= dN\, T_m / (P_m\, t_s) \end{aligned} \tag{7}$$

3.4. Validation with known diaphragms

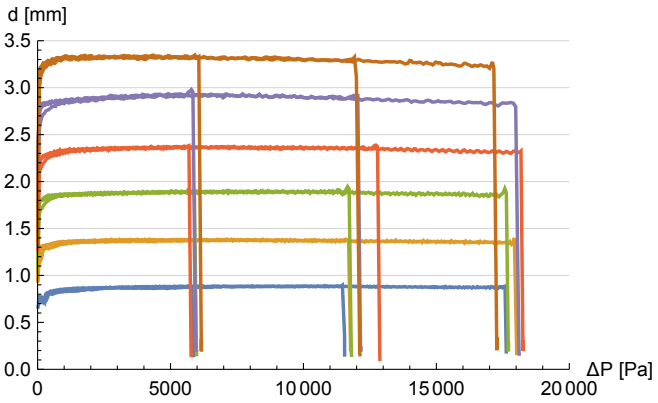


Figure 3. Color lines: aeraulic diameter of the tested diaphragms computed with the isochoric scheme Eq. 7. Each diaphragm is measured 3 times with different initial conditions. Thin dark lines: nominal diameter of the diaphragms.

A validation of the model was performed with a series of chamfered diaphragms of nominal diameters 1, 1.5, 2, 2.5, 3 and 3.5mm. Eq. 4 allows the computation of the equivalent aeraulic section S from U, ρ (in the bottle) and Δp, from which we deduce the equivalent aeraulic diameter of the diaphragms.

Fig. 3 depicts the results. The following values of r were determined by optimization (among 3 tests with different initial depressions for each diaphragm): 0.25, 0.32, 0.39, 0.46, 0.54, 0.61, for the diameters 1 to 3.5mm. This accounts approximately for the greater heat exchange with the bottle when the flow rate through the diaphragm is high. Before the discharge $r = 0.20$ (measured value in static conditions).

The aeraulic section of the diaphragm S_{dia} used in Problem 2) is calibrated with this method (in our case 2.9mm). The adaptation barrel with absorbing foam (damping the acoustic oscillations) belongs formally to the diaphragm (like every aeraulic resistance downstream the mouthpiece sensor).

3.5. Discrete time scheme for Problem 2

This problem is subdivided in 2 subproblems:
A) compute the net mass flow dn entering into the mouthpiece
B) compute the net mass flow dN entering into the bottle

The total mass flow entering trough the channel is $dN_{ch} = dn + dN$, from which $U_{ch} = T_0\,(p/P_0)^{\frac{\gamma-1}{\gamma}}\; dN_{ch}/(p\,dt)$ and its corresponding aeraulic section S_{ch} is deduced with Eq. 4.

Subproblem A) The mouthpiece is treated with Eq. 7 in which the variables of the mouthpiece replace those of the bottle ($P \to p$, $U \to u$, and so on). An adiabatic approximation can be used for this case : $dQ = 0$, thus $r = 0$.

Subproblem B) the combination of Eq. 6 with Bernoulli's law Eq. 4 enables the calculation of the pressure in the bottle P_m (at discrete time m) from the corresponding measurement of the pressure in the mouthpiece, denoted p_m.

Initialization Like Problem 1), additionally: $p_0 = P_0$.

Iterations for $m > 0$

$$\begin{aligned} \delta p &= p_{m-1} - P_{m-1}, \quad \rho = (M\, N_{m-1})/V \\ dN &= \text{sign}(\delta p)\, P_{m-1}\, S_{dia}\, \sqrt{|2\delta p/\rho|}\, t_s / T_{m-1} \\ dP &= (N_{m-1}\, r\, t_s\, (T_0 - T_{m-1}) + dN\, \gamma\, T_{jet})/V \\ N_m &= N_{m-1} + dN \\ P_m &= P_{m-1} + dP \\ T_m &= P_m V / N_m \\ U_m &= dN\, T_m / (P_m\, t_s) \end{aligned} \tag{8}$$

with $M = 0.028965/R = 0.00348388$ for the air.

4. MEASUREMENTS OF CLARINET REEDS

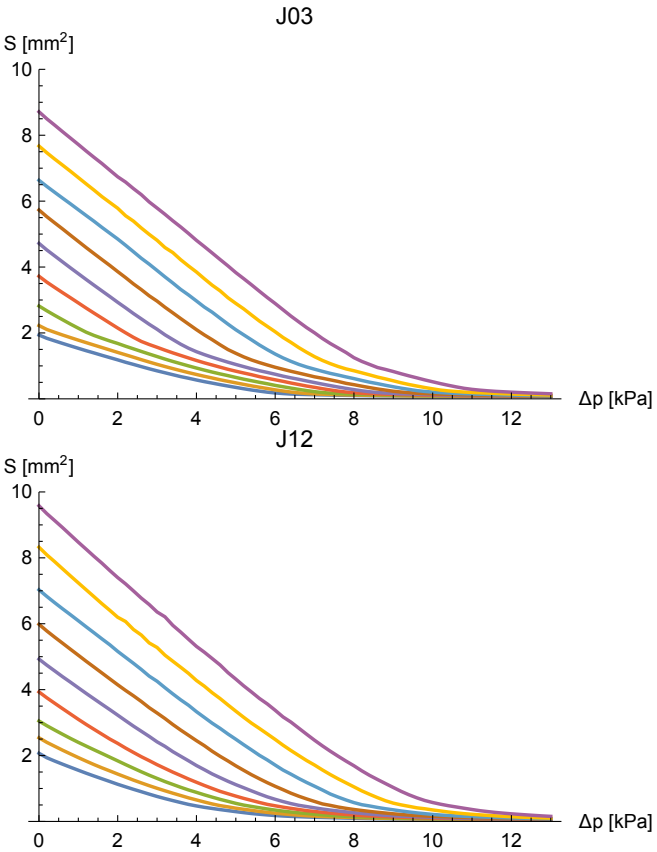


Figure 4. Aeraulic section $S(\Delta p, \psi)$ determined with Eq. 8 for the reeds J03 and J12, $\psi = 1 \ldots 9$ (color lines). Below the transition to turbulent flow, S is approximated from the optic sensors inside the mouthpiece.

A series of 14 clarinet reeds was measured with the described method (Rigotti and Rico, strength 3 and 3 1/2). 10

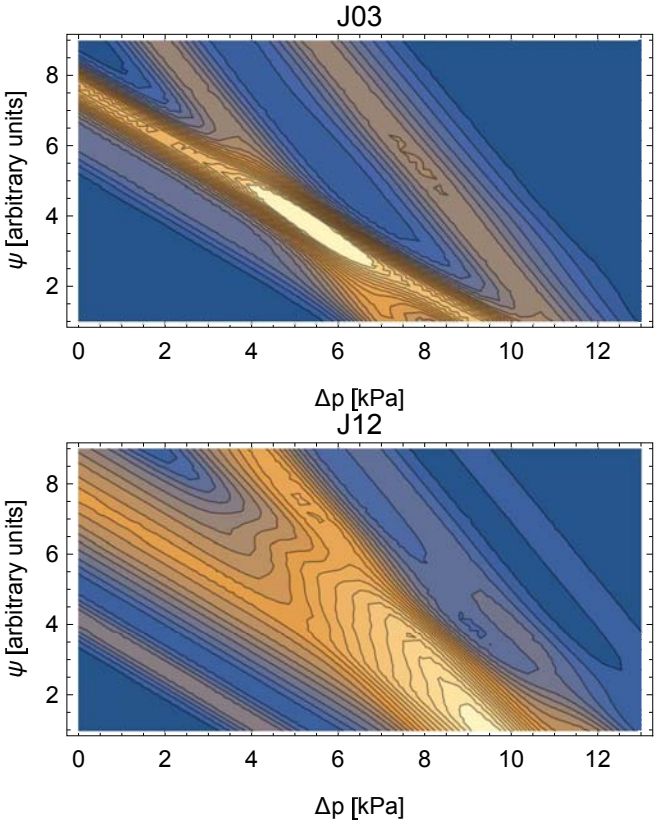


Figure 5. Contour plot of the second derivative of the aeraulic section $\frac{\partial^2 S}{\partial \Delta p^2}$ for the reeds J03 (Rigotti 3) and J12 (Rico 3 1/2), $\psi = 1 \dots 9$, according to the proposed model Eq. 9 (arbitrary color scale)

different positions ψ of the artificial lip were tested (in steps by 0.2 mm). Unfortunately our vacuum cleaner was not powerful enough for creating the vacuum in the bottle necessary to close some strong reeds with a loose embouchure. Complete measurements are available only for 13 reeds and 8 embouchures.

4.1. Measurements of the aeraulic section S

Fig. 4 depicts the results of the measurement of the aeraulic section $S(\Delta p, \psi)$ for 2 reeds. Fig. 5 illustrates $\frac{\partial^2 S}{\partial \Delta p^2}$ for the same reeds. An interesting feature is present for all reeds: we observe 2 different slopes along whom the second derivative is almost constant. The same observation can be done with $\frac{\partial^2 S}{\partial \Delta \psi^2}$. Along these slopes, Δp and ψ are partially decoupled.

The observed behavior can be approximated as a sum of 2 1D stiffening springs and a porosity constant:

$$S(\Delta p, \psi) \simeq S_1(\Delta p + k_1 \psi) + S_2(\Delta p + k_2 \psi) + k_p \quad (9)$$

The following values were determined by optimization for our setup : $k_1 = 1374.2$ and $k_2 = 670.3$ (same values for all reeds). The porosity constant k_p accounts for a residual flow which do not vanish for high values of Δp and must be determined individually for each reed.

This model implies that $S(\Delta p, \psi)$ is a convex function of Δp and of ψ. In other words, the second derivatives of S are non-negative. The stiffening springs can be implemented with non-negative, reproducing kernels (for instance gaussian kernel) as bandlimited functions. We used a kernel allowing a real-time implementation of the measured exciter for synthesis purpose (but this is beyond the scope of this paper).

The mean standard error of the model is around 0.02mm^2. This represents a height of 1.5 μm for a rectangular channel, 13mm wide, or 0.2% of the maximal measured aeraulic section.

5. CONCLUSIONS AND PERSPECTIVES

The precision of the method enables a detailed study of the nonlinear characteristics of clarinet reeds. A great variety of quasistatic behaviors could be measured, confirming the every-day experience of the clarinetists with their reeds and the importance of the non-linear contact between reed and mouthpiece. A comparison with 3D simulations about the dichotomic sensitivity of the reed to the lip pressure shows that k_1 accounts mainly for air coming from the middle of the channel while k_2 concerns mainly the air entering from the sides of the channel.

This short paper could only demonstrate the main features of the method, but could not investigate the musical consequences of the measurements for reed makers and clarinettists, explaining relationships with geometric and flexural measurements, nor demonstrate the ability of the model for a real-time simulation of the measured reeds.

Aknowledgments We thank the engineering school HE-Arc (Neuchâtel, Switzerland) for the facilities granted.

REFERENCES

[1] D. Casadonte, "The Clarinet Reed: An Introduction to its Biology, Chemistry, and Physics," Ph.D. dissertation, Ohio State University, 1995.

[2] M. O. Van Walstijn, "Discrete-time modelling of brass and reed woodwind instruments with application to musical sound synthesis," Ph.D. dissertation, University of Edinburgh, 2002.

[3] T. Guimezanes, "Etude expérimentale et numérique de l'anche de clarinette," Ph.D. dissertation, Université du Maine, Le Mans, France, 2007.

[4] P.-A. Taillard, F. Laloë, M. Gross, J.-P. Dalmont, and J. Kergomard, "Statistical estimation of mechanical parameters of clarinet reeds using experimental and numerical approaches," *Acta Acustica united with Acustica*, vol. 100, no. 3, pp. 555–573, 2014.

[5] J. Dalmont, J. Gilbert, and S. Ollivier, "Nonlinear characteristics of single-reed instruments: Quasistatic volume flow and reed opening measurements," *The Journal of the Acoustical Society of America*, vol. 114, p. 2253, 2003.

Construction and Modeling of a Horn Built of Cylindrical PVC Pipe

Kurt Hoffman

Whitman College
hoffman@whitman.edu

ABSTRACT

The flaring of the bell of brass instruments is crucial in obtaining the desired frequencies of the harmonics of the tube as well as improving the coupling of the standing waves in the tube to the exterior environment. To aid student understanding of the role the flaring tube plays in changing the harmonic frequencies, I built pvc herald trumpets using materials from the local hardware store. I also used Mathematica to make quantitative calculations of the standing waves of the instruments.

By utilizing reducing couplers and a range of tubing sizes, I was able to form a variety of bell profiles that could be used interchangeably to optimize the frequency ratios of the standing waves of the tube. The students were then in a position to explore the affect different lengths and shapes of the bell had on the standing wave frequencies of the horn using their ears as the detector. By playing the horn at different harmonics, they could judge whether the intervals were octaves, fifths, fourths, or some other undesired interval. While the method is trial by error, it does highlight how the shape of the bell can be used to modify the frequencies of different standing waves in the tube.

In considering this problem from a theoretical point of view, the use of a sequence of cylinders to form the bell simplifies the modeling of the tube. The approach to modeling the normal modes of the tube consists of using the impedance of each section of the tube to write the total impedance of the instrument. I then used Mathematica to find the roots of the impedance to identify the normal mode frequencies and to plot the pressure wave of the harmonics. For the case of a brass instrument, the mouthpiece end was treated as a closed end. The appropriate roots in this case are for $1/Z = 0$. This method ignores loss effects and doesn't account for the gradual change in radius that occurs at the reducers in the instrument. Despite these simplifications, the model standing wave frequencies were close to the measured frequencies of the instruments. The calculations also provide a relatively simple way to explore how different lengths and different radii affect the standing wave in the tube.

Printone: Interactive Resonance Simulation for Print-wind Instrument Design

Nobuyuki Umetani,[1†] *Athina Panotopoulou,*[2,1] *Ryan Schmidt,*[3,1] *Emily Whiting*[2]

[1]Autodesk Research, [2]Dartmouth College, [3]Gradientspace

[†]nobuyuki.umetani@autodesk.com

ABSTRACT

We presents an interactive design interface for three-dimensional free-form musical wind instruments [1]. The sound of a wind instrument is governed by the acoustic resonance as a result of complicated interactions of sound waves and internal geometries of the instrument. Thus, creating an original free-form wind instrument by manual methods is a challenging problem. Our interface provides interactive sound simulation feedback as the user edits, allowing exploration of original wind instrument designs. Sound simulation of a three dimensional wind musical instrument is known to be computationally expensive. To overcome this problem, we present a novel fast resonance frequency prediction method based on the boundary element method. Specifically, we formulate the resonance problem as an minimum eigenvalue problem of a nonlinear matrix. Furthermore, we can efficiently compute an approximate resonance frequency using a new technique based on a generalized eigenvalue problem. The designs can be fabricated using a 3D printer, thus we call the results "print-wind instruments" in association with woodwind instruments. We demonstrate our approach with examples of unconventional shapes performing familiar songs.

1. INTRODUCTION

We present *Printone*, a tool to design original functional three dimensional wind instruments [1]. Printone aids the user's instrument creation via interactive simulation of fundamental resonance frequencies. The inputs are a three-dimensional shape that is to be converted into a musical instrument and a set of target frequencies that should be produced through different open/close patterns of the finger holes. To find a design which produces the target frequencies, the user interactively scales the input shape and edits the placement and size of the finger holes. During this process they are guided by interactive feedback from the simulation. We compute the resonance numerically from the free-form geometric shape, hence the user can explore unconventional musical instruments beyond simple cylindrical shapes.

The foremost technical challenge in interactive instrument design is efficient acoustic resonance simulation. Frequency-domain simulation provides an efficient way to simulate three-dimensional acoustics by solving only on the boundary. However, this approach solves the acoustics for a predefined frequency, and thus gives no direct clue about where the resonance frequency occurs. To overcome this problem, we first characterize the resonance as a nonlinear eigenvalue problem.

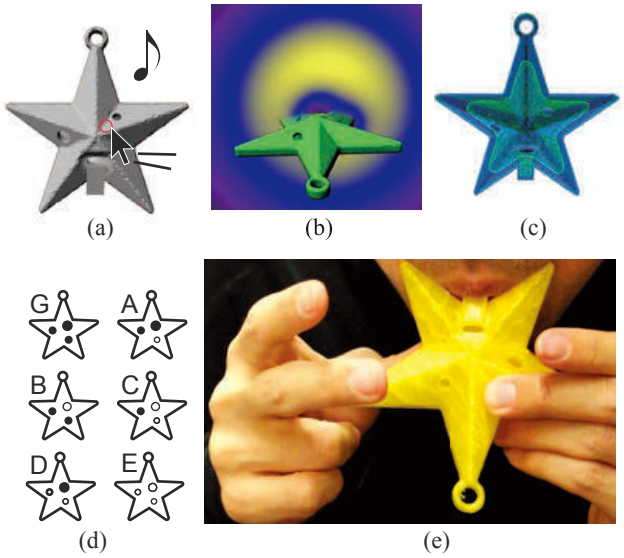


Figure 1. (a) The user can design the shape of wind musical instruments while guided by sound simulation feedback. (b) The acoustic resonance is simulated by solving the wave equation in the frequency domain. (c) The hollow musical instrument has internal structure and finger holes (d) which produce target tones at specific finger configurations. (e) The user can design unconventional functional wind instruments.

Then, we present a technique to efficiently estimate the resonance frequency from a set of sparsely-sampled linear systems in the frequency domain by solving a generalized eigenvalue problem. We further accelerate simulation during geometry editing by introducing sensitivity-based first-order estimation and a local matrix update approach. Our interactive tool also provides an automatic hole size adjustment function, which makes it easy to satisfy multiple frequency targets with different finger configurations.

Our simulation provides interactive response times during design by limiting prediction to a discrete set of fundamental resonance frequencies to different finger configuraitons, avoiding the expense of a full spectrum computation. This means we do not simulate the timbre of the instrument. However we show that predicting resonance frequencies is sufficient for hobby-grade wind musical instruments, where some degree of sound quality can be sacrificed to achieve novel designs. Our mouthpiece is the *fipple*, which is commonly used for recorders or ocarinas because it is easy to fabricate with commonly available FDM printers. The resulting instruments can be fabricated

with commonly available Fused Deposition Modeling (FDM) 3D printers. We validate our model through frequency comparisons between simulated and fabricated instruments, and we demonstrate a wide variety of novel instrument geometries.

2. USER INTERFACE

Fig. 2 shows a screen shot of our system. The 3D user interface consists of geometric editing tools to place finger holes and the fipple. Resonance simulation results are provided as feedback, allowing the user to adjust geometry and finger hole positions that generate desired resonance frequencies.

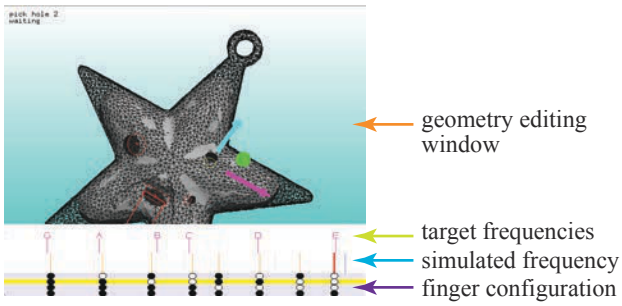


Figure 2. Screen-shot of our wind instrument design tool. The user can manipulate the size and location of finger holes and can open and close these holes while listening to the simulated sound. The fundamental frequency is visualized together with finger configurations and target notes.

Geometric editing Our tool imports existing polygonal meshes, which specifies the exterior shape of the instrument. The tool automatically hollows out the shape to generate an internal cavity. Then, the user positions the fipple and finger holes by clicking on the object's surface. The user can also uniformly scale the entire instrument and adjust the orientation of the fipple. The size of the finger holes can be continuously adjusted to manipulate the frequency of the resonance. The user can also specify whether the finger holes are open or closed to test the sounds produced by their instrument design.

Simulation feedback In the bottom of the window (Fig. 2), the tool shows finger hole open/close patterns along with their simulated resonance frequencies. At the same time, the speaker plays the sound with the simulated resonance frequency in the current finger configuration. During continuous adjustment of the size of an individual hole, the tool approximates the change of the resonance frequencies based on (rough) sensitivity analysis estimation in real-time (>30FPS) and updates the audio and visual feedback. During idle moments, the system computes more accurate frequency estimations for each finger configuration. Each update takes only a few seconds, and thus it does not interfere with the user's interactive editing session.

Interactive finger hole size optimization Our tool supports an optional *automatic* optimization of the finger hole sizing to simplify the design process. The user presses the "AutoTune" button after specifying target frequencies for several finger configurations, and the tool automatically resizes each finger hole such that it best achieves the frequency targets. This is a helpful capability because achieving multiple resonance frequencies requires adjusting multiple finger holes in a coordinated way. In Sec. 3 we explain the forward resonance simulation. For sensitivity analysis and hole-size optimization, please see our original paper [1]

3. ACOUSTIC RESONANCE

Acoustic resonance is a phenomena where the output of an acoustic system becomes very large compared to the magnitude of input at specific frequencies. Our goal is finding such resonance frequencies for a cavity shape, where the enclosed internal hull is connected to the exterior with small holes. Typical methods usually involve specifying locations of a sound-emitting point and a sound-observing point, and then find the ratio between input and output. Instead, we formulate resonance as a minimum eigenvalue problem, which is solely defined by the object's geometry and independent of any source and observation points.

3.1. Background on Acoustic Simulation

We consider a problem where sound waves, which are emitted from external vibration sources, are reflected from an object's surface S. Resonance is a static phenomenon, i.e., the magnitude of a sound wave is maintained via excitation from an external vibration source. In this case it is more convenient to formulate the problem in the frequency-domain, where all the variables are converted to static via temporal Fourier transform. The pressure field is represented as a set of harmonically vibrating fields $p(\mathbf{x})e^{+i\omega t}$, where ω is the angular frequency of the vibration. The wavenumber is given as $k = \omega/c$, where c is the speed of sound in the medium (approximately $340\ \mathrm{m/s}$ in air). The spatial distribution of pressure is denoted $p(\mathbf{x}) \in \mathbb{C}$, and satisfies the Helmholtz equation

$$(-\nabla^2 - k^2)p(\mathbf{x}) = f_{ext}(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{1}$$

where $f_{ext}(\mathbf{x})$ is the source term that represents a forced vibration from an excitation mechanism located inside the medium Ω. For example, a point source at $\mathbf{x}_{src}$ with unit magnitude results in $f_{ext}(\mathbf{x}) = \delta(\mathbf{x}_{src} - \mathbf{x})$, where δ is the Dirac delta function. We assume the object's surface is very hard and reflects the sound completely, i.e., the normal velocities of the air particles are zero at the surface. This can be written as a Neumann boundary condition.

The Helmholtz equation can be solved with many different discretization methods. We use the Boundary Element Method (BEM), which compactly represents the three-dimensional pressure distribution in Ω with the pressure values on the object's surface mesh S. The BEM discretizes the Helmholtz equation into a matrix formulation

$$\mathbf{A}(k)\mathbf{p} = \mathbf{f}_{ext}, \tag{2}$$

where $\mathbf{f}_{ext} \in \mathbb{C}^N$ is a vector storing incident pressure at vertices and $\mathbf{p} \in \mathbb{C}^N$ is the resulting reflection pressure value at the vertices of the object. The coefficient matrix

$\mathbf{A}(k) \in \mathbb{C}^{N\times N}$ is a dense non-Hermitian matrix that varies with respect to the wavenumber k. Note that $\mathbf{A}$ changes nonlinearly with respect to k, since the first derivative of the coefficient matrix is k-dependent. Once the reflection pressure $\mathbf{p}$ is computed by solving (2), the pressure value at the observation point inside the medium $\mathbf{x}_{obs} \in \Omega$ is given by integrating pressure over the surface.

3.2. Resonance as a Small Minimum Eigenvalue

Typically, the resonance is defined as the frequency where pressure at the observation point $p(\mathbf{x}_{obs})$ has a local maximum for a specific source point $\mathbf{x}_{src}$. Such local maximums are typically detected with a *frequency sweep*, i.e., solving for the pressure at $p(\mathbf{x}_{obs})$ at many different wavenumbers. Frequency sweeps are very computationally intensive, we thus first formulate resonance in a different way such that we do not need to compute the pressure at a specific observation point.

The resonance frequency does not depend on the observation position, e.g., the fundamental frequency of a flute does not depend on a listener's position. Furthermore, the resonance frequency typically does not depend on the position of excitation. Hence, instead of comparing the values at these points, we compare the overall magnitude of incident excitation and reflection on the surface. Specifically, we define resonance as the frequency where the magnitude of reflection pressure on surface $|\mathbf{p}|$ takes the maximum value compared to the magnitude of the incident pressure $|\mathbf{f}_{ext}| = |\mathbf{A}(k)\mathbf{p}|$, i.e., the magnitude of the incident wave reaches a minimum relative to the magnitude of the reflection pressure.

$$k_{resonance} = \arg\min_{k} \left(\min_{\mathbf{p}} \frac{|\mathbf{A}(k)\mathbf{p}|}{|\mathbf{p}|} \right), \tag{3}$$

Note that the norm of the reflection pressure $|\mathbf{p}|$ corresponds to the sum of acoustic energy that the vertices emit and the norm of the incident pressure $|\mathbf{f}_{ext}|$ corresponds to the sum of acoustic energies that the vertices receive.

Given a fixed wavenumber k in (3), the problem of finding the minimal reflection pressure $\mathbf{p}$ is equivalent to finding the magnitude of the minimum eigenvalue of matrix $\mathbf{A}(k)$. If this minimum eigenvalue is small, the magnitude of reflection pressure becomes large compared to the magnitude of incident pressure. Since the matrix $\mathbf{A}(k)$ changes nonlinearly with respect to k, finding k that minimizes the smallest eigenvalue of $\mathbf{A}(k)$ by sampling many different wavenumbers is very time consuming. Section 4 describes our novel approximation to solve for this wavenumber efficiently.

4. APPROXIMATING RESONANCE

In the previous section, we defined the resonance frequency as the frequency which produces the smallest minimum eigenvalue of the coefficient matrix $\mathbf{A}(k)$. Computing the matrix $\mathbf{A}$ requires $\mathcal{O}(N^2)$ operations and is very expensive, thus we cannot compute $\mathbf{A}(k)$ for many different wavenumbers if our system is to run at interactive rates. In this section, we present a method to estimate resonance from a few discretely sampled wavenumbers.

4.1. Generalized Eigenvalue Problem Approximation

We sample the wavenumber at k_m $(m = 1, \ldots, M)$. At each discretely sampled wavenumber k_m, we approximate the k-varying coefficient matrix $\mathbf{A}(k)$ with a first-order Taylor expansion as

$$\mathbf{A}(k) \simeq \mathbf{A}_m + (k - k_m)\mathbf{D}_m, \tag{4}$$

where $\mathbf{A}_m = \mathbf{A}(k_m)$ and $\mathbf{D}_m = \partial\mathbf{A}(k)/\partial k$ is the derivative of the coefficient matrix at k_m. The derivative of the operator $\mathcal{A}(k)$ with respect to the wavenumber gives a scaled identity operator $\partial\mathcal{A}(k)/\partial k = -2k\mathcal{I}$. Hence, its discretization $\mathbf{D}_m$ is close to a real-valued diagonally-dominant matrix.

Finding the resonance wavenumber (3) using this first-order approximation (4) is closely related to the generalized non-Hermitian eigenvalue problem:

$$\mathbf{A}'(\lambda)\mathbf{v} = 0, \quad \mathbf{A}'(\lambda) = \mathbf{A}_m - \lambda\mathbf{D}_m, \tag{5}$$

where $\lambda \in \mathbb{C}$ is the eigenvalue and $\mathbf{v} \in \mathbb{C}^N$ is an eigenvector.

Since $\mathbf{A}'$ is just a shifted version of $\mathbf{A}$: $\mathbf{A}(k) = \mathbf{A}'(k_m - k)$, if $(k_m - k)$ is close to λ, the minimum eigenvalue of $\mathbf{A}(k)$ becomes small. In other words, when $(k - k_m)$ is the real part of λ, a vector $\mathbf{p}$ similar to $\mathbf{v}$ makes the $|A(k)\mathbf{p}|/|\mathbf{p}|$ small. Since the first order Taylor expansion (4) is valid only when the wavenumber is close to the sampled one k_m, we are interested in finding a resonance wavenumber nearest to the sampled one k_m. The eigenvalue of smallest magnitude λ^0 gives a resonance wavenumber as

$$k_{resonance} = k_m - \mathrm{Re}(\lambda^0). \tag{6}$$

This is a good approximation because the eigenvalue λ is close to a real number for the cavity resonance. The generalized eigenvalue problem (5) is identical to the conventional eigenvalue problem for matrix $\mathbf{A}_m\mathbf{D}_m^{-1}$. As we discussed in the previous section, $\mathbf{A}_m$ has nearly-real eigenvalues. Since the $\mathbf{D}_m$ is close to a real-valued diagonally-dominant matrix, multiplying by its inverse does not alter the properties of $\mathbf{A}_m$.

Inverse power method The eigenvalue with smallest magnitude λ^0 can be easily computed using the inverse power method. More specifically, by repeatedly multiplying the inverse of the matrix $\mathbf{A}_m\mathbf{D}_m^{-1}$ with an arbitrary vector $\mathbf{v}$ as $\mathbf{v} \leftarrow \mathbf{D}_m\mathbf{A}_m^{-1}\mathbf{v}$, the Rayleigh quotient

$$\theta = \frac{\langle \mathbf{D}_m\mathbf{A}_m^{-1}\mathbf{v}, \mathbf{v}\rangle}{|\mathbf{v}|^2} \tag{7}$$

converges to the inverse of the minimum eigenvalue. Convergence is defined as $|\mathbf{w} - \theta\mathbf{v}|/|\mathbf{v}| < \epsilon$, we use $\epsilon = 10^{-5}$ in this paper. Once the iteration converges, the inverse of the Rayleigh quotient gives the minimum eigenvalue as $\lambda^0 = \bar{\theta}/|\theta|^2$, where $\bar{\theta}$ is the conjugate of θ.

4.2. Wavenumber Sampling

We uniformly sample the wavenumbers such that the range of these frequencies covers the target frequencies. The computation of the matrix derivative $\mathbf{D}_m$ is as expensive as the

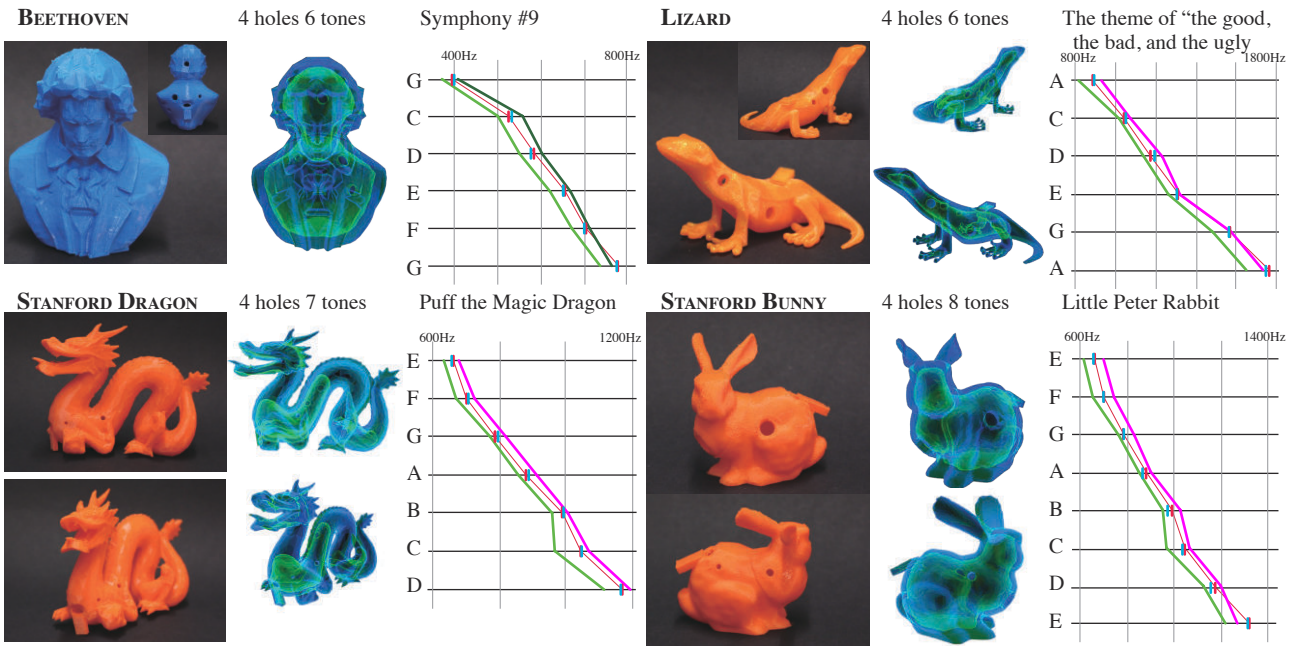


Figure 3. Examples of free-form wind instrument design. The name of the model, number of holes, number of target notes and the target songs are listed together with frequencies of each note. For each instrument, the red lines show the target frequencies, the blue lines show the simulated frequencies, the light green shows the minimum and the dark green line shows the maximum measured frequencies.

computation of the coefficient matrix $\mathbf{A}_m$. Hence we approximate the matrix derivative with a central differential of sampled coefficient matrices

$$\mathbf{D}_m = \frac{\mathbf{A}_{m+1} - \mathbf{A}_{m-1}}{k_{m+1} - k_{m-1}}. \tag{8}$$

For the first and last sampled wavenumber k_1 and k_M, we cannot compute the central difference and we must use forward and backward finite difference schemes.

5. RESULTS

Figure 1 and Figure 3 show five out of sixteen wind instruments created using our tool. Each of these instruments are designed to play a specific song, i.e., notes in the song are the target frequencies. All input models were selected from a public 3D model-sharing website, with choices inspired by the theme of their songs. We encourage readers to watch the supplemental video to hear the wind instruments' sound [1].

Figure 3 compares the target frequencies and actual measured frequencies of the instruments we designed. Red lines are the target frequencies, while the measured frequencies are shown as ranges between the light and dark green lines. This range occurs because the sound is influenced by the speed of blowing (higher speed produces higher frequencies). We gradually increased the blowing speed and recorded the lowest and highest frequencies that produced clear sound. We observe that the target frequencies generally (53 out of 56 target frequencies) fall within measured ranges, i.e., the user can play the exact tones with proper blowing speed. This type of fine-scale tuning by adjusting the blowing speed is very common in wind instruments.

On average our tool runs at 5 frames-per-second and the resonance frequency can be computed in several seconds after the user completes a geometry manipulations (opening/closing, relocating a hole). Our tool is written in C++ and the entire resonance computation is performed by a single CPU thread and we used a MacBook Pro with Core i7 (3.0GHz) CPU. Please see our original paper [1] for more detailed discussion about accuracy and performance.

[1] https://www.youtube.com/watch?v=dWHYLqcCPuU

6. CONCLUSION

We presented an interactive interface to design free-form wind instruments. We simulated only the fundamental frequency, hence simulation of timbre is a future work.

7. ACKNOWLEDGEMENT

We also thank Dr. Alex Barnett for various advice. This project is partially supported by the National Science Foundation under Grant No. 1464267.

REFERENCES

[1] N. Umetani, A. Panotopoulou, R. Schmidt, and E. Whiting, "Printone: Interactive resonance simulation for free-form print-wind instrument design," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 184:1–184:14, Nov. 2016. [Online]. Available: http://doi.acm.org/10.1145/2980179.2980250

Wind Instrument Optimization Made Practical

Burton Patkau,[1†] *Antoine Lefebvre,*[2] *Edward Kort*[3]

[1]Vericom Systems Ltd.
[2]Développement Informatique Antoine Lefebvre
[3]Wood Wind Flutes
[†]bpatkau@vericomsystems.com

ABSTRACT

Mathematical modelling of wind instruments has progressed to the point where it is now feasible to model and optimize instrument designs on a personal computer. WIDesigner, a computer application for wind instrument design optimization, exploits several existing techniques to demonstrate this feasibility. Its current objective is laying out the bore and toneholes of a wind instrument to optimize intonation across a specified scale of notes. To make rapid impedance calculations, WIDesigner uses the Transfer Matrix Method on wind instrument bores, coupled with empirical models of instrument mouthpiece behaviour. The principal optimization algorithms used are BOBYQA for local optimization, and a custom variant of the DIRECT algorithm for global optimization.

1. INTRODUCTION

Historically, the design of wind instruments has depended on iterative improvement through trial-and-error changes to real instruments. Meanwhile, much academic work has gone into developing numerical predictive models of instrument components. This work has progressed to the point that researchers are assembling models to predict the performance of existing designs and create new designs optimized for performance.

This paper discusses key techniques used in one such effort, WIDesigner [1]. Although none of these techniques is novel, taken together they make it feasible to optimize the intonation of wind instruments on a personal computer:

- Transfer Matrix models in the frequency domain to keep computational requirements low.
- Empirical mouthpiece models to encapsulate the influence of the drive mechanism.
- Two optimization algorithms for high-performance local and global optimization.

WIDesigner is a Java application to assist the design of woodwind instruments. In its current manifestation, it calculates instrument dimensions to provide optimum tuning accuracy over the full playing range of an instrument.

2. OPTIMIZATION APPROACH IN WIDESIGNER

WIDesigner works with internal models of instruments and tunings, which can be saved as XML files. The instrument model includes the position and diameter of toneholes, the position of transitions in the instrument bore, and the internal diameter at these bore points. Each class of instrument also records features, specific to the class, of the sound-generation mechanism at the mouthpiece. The tuning model includes, for each note, the pattern for fingering toneholes, and the target frequency for the note.

To evaluate an instrument's tuning, WIDesigner finds a playing frequency close to the target frequency of each note, and calculates the tuning error between the playing frequency and the target frequency. It does not attempt to calculate impedances over broad frequency ranges, predict *a priori* playing frequencies, or even keep track of what register a playing frequency might fall into.

To optimize an instrument, WIDesigner minimizes a real-valued objective function over a constrained parameter space.

$$\text{minimize } F(\vec{x}),\ \text{subject to } \vec{x} \in \mathbb{R}^n,\ l_i \le x_i \le u_i \qquad (1)$$

The parameter vector $\vec{x}$ maps to the instrument geometry. WIDesigner supports over 30 different mappings, from simple mappings such as the vector of tonehole diameters, to compound mappings that cover almost all dimensions of the instrument. Some mappings provide indirect views of the geometry; for example, tonehole spacing rather than absolute tonehole position, or bore diameter ratios rather than absolute bore diameters. These indirect mappings allow the constraints l_i and u_i to represent physical limitations more directly, such as the maximum space between a player's fingers, or the direction of a bore taper. The constraints allow the user to ensure that solutions are physically feasible, within the validity limits of the prediction model, physically usable, and perhaps even aesthetically satisfactory.

The function to be minimized, $F(\vec{x})$, could be based on reactance or reflectance, but we have found the most effective measure to be the sum of squares of tuning errors in cents over all notes in the tuning.

$$F(\vec{x}) = \sum_{i=1}^{m} (g(f_{p,i}(\vec{x}), f_{t,i}))^2 \qquad (2)$$

where $f_{p,i}(\vec{x})$ is the predicted playing frequency for note *i* with geometry $\vec{x}$, $f_{t,i}$ is the target playing frequency for note *i*, and $g(f_1, f_2)$ is the frequency difference in cents:

$$g(f_1, f_2) = 1200 \log_2(f_1/f_2) \qquad (3)$$

For any optimization, the user has full control over the constraint vectors, $[l_1 \quad l_2 \quad \ldots]$ and $[u_1 \quad u_2 \quad \ldots]$. WIDesigner can save constraint vectors as XML files for later re-use.

3. MODELLING THE BORE

WIDesigner uses the Transfer Matrix Method (TMM) to model the instrument's frequency response. A number of published articles discuss the use of transfer matrices [2-5]. In the context of optimization software, which computes the system's response thousands of times for each instrument fingering while the geometry is evolving, it is of primary importance that each evaluation of the system be as efficient as possible. The TMM meets this requirement.

WIDesigner starts with a $[\ P\ U\ ]^T$ state vector for the open termination of the instrument, and multiplies the state vector by the transfer matrix of each bore section and tonehole from the bottom of the instrument to the mouthpiece.

Given the $[\ P\ U\ ]^T$ state vector calculated for the bore, as seen by the mouthpiece, the mouthpiece calculator returns a $[\ P\ U\ ]^T$ state vector representing the state of the entire instrument as seen by the driving source. From this final state vector, WIDesigner calculates a final instrument impedance. Figure 1 illustrates the overall process.

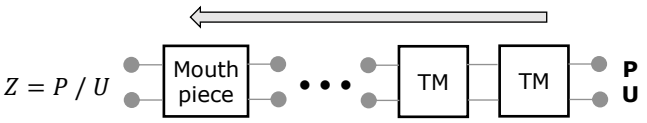


Figure 1. Instrument modelled as a cascade of 2-port networks.

WIDesigner uses the instrument impedance function to predict playing frequencies. Typically, the predicted frequency is the frequency closest to the target frequency at which the imaginary portion of $P\ /\ U$ is zero. Note that the impedance function at this point is not the raw bore impedance; it includes the effect of the drive source and instrument performance characteristics.

The transfer matrix of cylindrical and conical bore segments with boundary layer losses are formulated following Eq. (2) and (5) respectively from [6]. Toneholes use the results given in [7], principally those for unflanged toneholes in a thick pipe. The radiation impedance of the open-end, for the initial $[\ P\ U\ ]^T$ state vector, is calculated following [8] for unflanged tubes and tubes with a finite flange.

4. MODELLING THE MOUTHPIECE

4.1. General Approach

WIDesigner implements a different mouthpiece calculator for each class of instrument. At present, WIDesigner uses empirical mouthpiece models; each model has one or two calibration parameters. A model is calibrated by adjusting its parameters to give the best fit between WIDesigner's tuning predictions and measured tuning from a real instrument. WIDesigner supports calibration functions to assist the process; these are simply optimization functions for which the parameter vector $\vec{x}$ is the mouthpiece model's calibration parameters.

The expectation is that the calibration parameters depend only on the mouthpiece, so for example, calibration done on a whistle before drilling the toneholes would still apply as the toneholes were laid out.

4.2. Native American Flute (NAF) Mouthpiece Model

The NAF mouthpiece model is derived from the FluteCalc program that Dan Gordon developed for transverse flutes [9]. We have adjusted the FluteCalc formulas for use in NAF design, and added two empirical parameters – one static value and one parameter used for calibration.

4.3. Whistle and Flute Mouthpiece Model

The mouthpiece model for whistles and flutes acknowledges that an instrument can produce, for a given note, a range of pitches within one register depending on the blowing pressure. For each note, the model predicts both a maximum frequency, at which the instrument jumps to a higher register, and a minimum frequency, at which the instrument drops to a lower register or fails to sustain the note.

The mouthpiece model calculates an adjusted impedance for the overall instrument. The maximum frequency is that at which the imaginary portion of the impedance is zero. To predict the minimum frequency the mouthpiece model uses a model for loop gain [10] to calculate the frequency at which the gain drops below 1.

In performance, the actual playing frequency will be between the minimum and maximum. While a skilled performer can adjust the blowing pressure from note to note to maintain correct intonation, playing is more comfortable when there is a regular increase in blowing pressure going up the scale. The mouthpiece model predicts actual playing frequencies based on such a regular increase in blowing pressure.

For a transverse flute, the model treats the flute headspace as a closed tube in parallel with main bore, as seen from the embouchure hole.

4.4. Reed Mouthpiece Model

The NAF, whistle, and flute, mouthpiece models return a $[\ P\ U\ ]^T$ state vector, from which WIDesigner calculates an impedance. In contrast, the reed mouthpiece model returns a "normalized admittance" state vector, $[\ ZoU\ P\ ]^T$; the division yields a normalized admittance, ZoY instead of an impedance. This allows the remaining WIDesigner components to treat flute and reed instruments identically. Equation (4) gives the normalized admittance from the reed mouthpiece model, where $[\ P_b\ U_b\ ]^T$ is the state vector of the bore as seen by the mouthpiece.

$$ZoY\ = ZoU_b/P_b\ + j\Delta Y \tag{4}$$

The mouthpiece model calculates the admittance adjustment ΔY as a linear function of frequency. The slope and intercept of this function are the calibration parameters of the model. In general, we expect the adjustment will always be positive for inward-striking reeds (single and double cane reeds), and positive or negative for outward-striking reeds (lip reeds) [11]. In our experience with real instruments, the slope has always been positive for both classes of reeds. At this time, we expect that the calibration values may be different for different players, even with the same reed or mouthpiece.

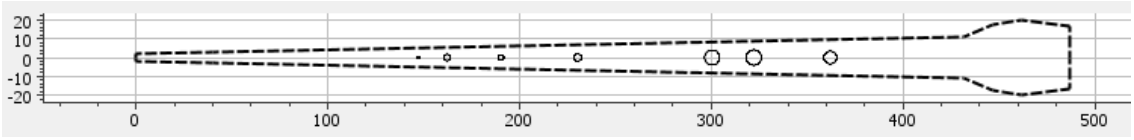


Figure 2. Hypothetical shawm design from WIDesigner

5. OPTIMIZERS

WIDesigner optimization requires optimization algorithms that are multi-variate, constrained, and derivative-free. Of the optimizers available to us, the two we have found most effective are BOBYQA [12] for local optimization, and a custom variant of DIRECT [13, 14] for global optimization.

5.1. BOBYQA

BOBYQA, Bounded Optimization by Quadratic Approximation, performs local optimization using quadratic approximation within a trust region. It serves as a fast algorithm for finding a local optimum, suitable for refining a design that is already reasonable. With BOBYQA, WIDesigner can optimize tonehole sizes and positions in a matter of seconds, testing on the order of 1000 instrument geometries.

5.2. DIRECT

DIRECT, DIviding RECTangles, is a more thorough algorithm for finding a global optimum in the whole search space. It starts with the whole search space between the constraint bounds, and divides it into successively smaller and smaller hyperrectangles. For each division, DIRECT divides a hyperrectangle in three along one of the dimensions, evaluating the function at the centre of each new hyperrectangle. At each iteration, DIRECT chooses hyperrectangles to divide that are large, or have low function values.

At present, WIDesigner uses a modified form of DIRECT:

- The optimizer will not divide hyperrectangles smaller than a specified minimum width. Once the best point found is in a hyperrectangle of the minimum size, the optimizer will return this point if there is no further improvement.
- Except for hypercubes, the optimizer divides a hyperrectangle on only one side in each iteration.
- Under some conditions, the optimizer uses an alternate strategy for selecting which hyperrectangles to divide in an iteration.

DIRECT is particularly useful when the optimization is not starting from a known good design, or when the search space has a number of local minima that can lead BOBYQA astray; this seems a particular problem when bore profile optimization is involved. For global optimization, WIDesigner runs DIRECT first, then uses BOBYQA to refine the solution found by DIRECT. This two-stage approach can optimize full instruments in a matter of minutes, testing tens of thousands of instrument geometries.

Figure 2 shows a sketch of the bore of a hypothetical two-octave shawm in D4, as optimized by WIDesigner. The shawm has 6 finger holes and a thumbhole, and plays 17 diatonic notes, from D4 through D6, plus cross-fingered flattened sevenths C5 and C6. WIDesigner optimized 19 dimensions of the instrument geometry, including the bore diameter at four points on the instrument bell, tuning all notes within 12 cents of Equal Temperament pitch. DIRECT tried about 17000 geometries, BOBYQA tried another 29000, in about 2 minutes on a 2.67GHz laptop CPU.

6. DESIGN AND CONSTRUCTION WITH WIDESIGNER

Design and construction of a new instrument still involves some iteration, even with the introduction of design optimization software such as WIDesigner. The optimization software helps the process converge more directly and reliably to the desired end: a well-tuned instrument. The steps below illustrate a typical process:

- Determine the target tuning, and select a draft design for the new instrument. Set the calibration parameters from an instrument with a similar mouthpiece.
- Produce optimized instrument designs with WIDesigner. Adjust the constraints to get a design that delivers the desired tuning accuracy, satisfies physical limitations of playability and aesthetics, and delivers the desired tone quality according to the designer's experience.
- Begin building the instrument. At intervals in construction, measure the current instrument geometry and tuning. Use these measurements to refine the WIDesigner calibration parameter and geometry values, and re-optimize the instrument design. For example, after crafting the mouthpiece and bore profile, measure the fundamental and one or more overtones, and use these to re-calibrate WIDesigner and re-optimize the tonehole size and location. After drilling undersized toneholes, measure the notes produced to check the calibration and re-optimize the tonehole size.
- Shape the mouthpiece and toneholes to produce the desired tone quality, and bring all the notes into tune. Final shaping proceeds based on measured frequencies rather than frequencies WIDesigner predicts. Tonehole adjustments may include aspects that WIDesigner's geometry model does not capture, such as elliptical toneholes, angled toneholes, or rounded edges.

7. NEXT STEPS

Clearly, the mouthpiece models would benefit from more data and research, to refine the models and ensure they accurately capture properties of real instruments. The current instrument model in WIDesigner is limited in what instruments it can describe: concert woodwinds require a description of keyed toneholes, and a Transfer Matrix model to go with it; brass instruments require description of valves, crooks and slides. Brass instruments would also benefit from enhancement of the

bore Transfer Matrix models to cope with bent crooks and widely flaring bells.

We also hope to investigate other aspects of instrument performance that we can model, such as timbre.

REFERENCES

[1] E. Kort, A. Lefebvre and B. Patkau, "Using WIDesigner," January 2016. [Online]. Available: https://github.com/edwardkort/WWIDesigner/wiki/Using-WIDesigner. [Accessed November 28, 2016].

[2] G. R. Plitnik and W. J. Strong, "Numerical method for calculating input impedances of the oboe.", *J. Acoust. Soc. Am.,* vol. 65, pp. 816-825, 1979.

[3] R. Caussé, J. Kergomard and X. Lurton, "Input impedance of brass musical instruments – Comparaison between experimental and numerical models," *J. Acoust. Soc. Am.*, vol. 75, pp. 241-254, 1984.

[4] D. Keefe, "Woodwind air column models," *J. Acoust. Soc. Am.,* vol. 88, pp. 35-51, 1990.

[5] D. Mapes-Riordan, "Horn modeling with conical and cylindrical transmission-line elements," *J. Audio Eng. Soc.*, vol. 41, pp. 471-482, 1993.

[6] A. Lefebvre, G. Scavone and J. Kergomard, "External tonehole interactions in woodwind instruments." *Acta Acustica united with Acustica*, vol. **99**, no. 6, pp. 975-985, 2013

[7] A. Lefebvre and G. P. Scavone, "Characterization of woodwind instrument toneholes with the finite element method," *J. Acoust. Soc. Am*., vol. 131, no. 4, April 2012.

[8] J.-P. Dalmont, C. J. Nederveen, and N. Joly, "Radiation impedance of tubes with different flanges: Numerical and experimental investigations", *J. Sound Vib.* vol. 244, no. 3, pp. 505–34, 2001.

[9] D. Gordon, "FluteCalc," personal communication, 2006.

[10] R. Auvray, B. Fabre and P.-Y. Lagree, "Regime change and oscillation thresholds in recorder-like instruments," *J. Acoust. Soc. Am.,* vol. 131, no. 2, pp. 1574-1585, February 2012.

[11] N. K. Fletcher, R. K. Silk, L. M. Douglas, "Acoustic admittance of air-driven reed generators," *Acoustica,* vol. 50, pp. 155-159, 1982.

[12] M. J. D. Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives," Cambridge, England: Cambridge University, Department of Applied Mathematics and Theoretical Physics Technical Report NA2009/06, 2009.

[13] D. R. Jones, C. D. Perttunen and B. E. Stuckman, "Lipschitzian optimization without the Lipschitz constant," *J. Optimization Theory and Application,* vol. 79, no. 1, pp. 157-181, October 1993.

[14] S. G. Johnson, "The NLopt nonlinear-optimization package," [Online]. Available: http://ab-initio.mit.edu/wiki/index.php/NLopt_Algorithms#DIRECT_and_DIRECT-L. [Accessed 29 August 2016].

How Do Wood Polishing and Oiling affect Acoustic Dissipation in the Bore of Wind Instruments?

Henri Boutin,[1†] *Sandie Le Conte,*[1] *Jean-Loïc Le Carrou,*[2] *Benoît Fabre*[2]

[1]CRC, USR 3224-ECR Musée de la Musique, Cité de la Musique, Philharmonie de Paris
[2]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7190, Institut Jean Le Rond d'Alembert
[†]boutin@lam.jussieu.fr

ABSTRACT

In the laboratory of Musée de la Musique, Paris, a part of research activities aims to provide acoustical information on collections. Concerning wind instruments, various wood species are used in the making procedure, having undergone different treatment techniques and showing various wearing and aging conditions. Such differences involve large variations of roughness and porosity, affecting the acoustic dissipation in the resonator, with a potential impact on playability and sound.

The influence of polishing and oiling on the acoustic dissipation is experimentally investigated with a corpus of five cylindrical pipes, manufactured by a flute maker. Their species are typically used in wind instrument making: maple (Acer Pseudoplatanus), pear wood (Pyrus Communis L.), boxwood (Buxus Sempervirens) and African Blackwood (Dalbergia Melanoxylon). The pipes are parallel to the grain, except one whose axis forms an angle of 60 degrees relative to the fiber direction.

A simple method is suggested to estimate the pipe attenuation factors. It relies on two impedance measurements with the far-end sealed and two lengths of air column, one being twice as long as the other. They are adjusted with an airtight silicon cylinder inside the bore. The attenuation factors are measured before and after polishing, and then after oiling. They are compared to a theoretical value calculated using a model of cylinder with rigid, smooth and non-porous walls.

Before oiling, the estimated attenuation factors are larger than the theoretical ones. Polishing reduces these factors but does not modify their order: maple, boxwood, pear wood and African Blackwood from largest to smallest. Their differences reveal variations of visco-thermal losses and wall impedance among the pipe directions and species considered. After oiling, the estimated attenuation factors in the pipes parallel to the fiber direction are similar to the theoretical one.

As a result, in addition to increasing solidity and durability of wind instruments, as reported by makers involved in this study, selecting the wood species and the direction of the wood pieces, as well as polishing and oiling the resonator tend to reduce the acoustic dissipation inside the bore. A theoretical model of propagation in cylinders is suggested, involving finite impedance on the inner wall. It allows us to estimate this parameter for the species considered before and after polishing, and after oiling. Further studies will aim at investigating the mechanisms of oil impregnation and the influence of periodic oil applications in wood species commonly used in wind instrument making.

Parameter Optimisation of a Viscothermal Time-Domain Model for Wind Instruments

Sebastian Schmutzhard,[1†] *Vasileios Chatziioannou,*[1] *Alex Hofmann*[1]

[1]Department of Music Acoustics (IWK), University of Music and Performing Arts, Vienna, Austria
[†]schmutzhard@mdw.ac.at

ABSTRACT

Viscothermal losses are vital for realistic numerical models of wind instruments. Numerical methods for modelling acoustic tubes that include viscothermal losses are well understood in the frequency domain for many years now. In this paper, motivated by the aim to numerically investigate note transitions on wind instruments, we study time-domain models for acoustic tubes. Time-domain modelling is better suited to such nonlinear, as well as time-varying systems. Recent methods use rational function approximation for including viscothermal losses in the time domain. We pursue the same approach and present a semi-analytic technique for the numerical integration of losses. Finally, in an attempt to accurately reproduce experimental findings, a time-domain bore optimisation routine is presented that includes the effect of viscothermal and radiation losses.

1. INTRODUCTION

Boundary layer losses need to be considered when modelling wave propagation in acoustic ducts. Physical modelling approaches including viscothermal losses are mostly performed in the frequency domain (see, e.g. [1, 2] or [3, Ch. 3.2.3]). The results of such frequency-domain approximations have been shown to be in good agreement with experimental measurements [4]. Nevertheless, it is sometimes required to model acoustic tubes in the time domain. This is particularly the case when systems with time-varying properties, as well as nonlinear systems need to be modelled [5].

In an attempt to simulate experimental arrangements for the investigation of transient processes in woodwind instruments (e.g. note transitions), it is therefore desirable to obtain accurate time-domain wave propagation models that include the effect of viscothermal losses. Recently, the use of rational function approximation has been presented to this end [6, 7, 5]. In this work, following the same approach, we present a semi-analytic technique for incorporating time-domain losses. The resulting wave propagation model can be coupled to a nonlinear excitation model that captures the interaction between the player's embouchure and the reed-mouthpiece system [8].

In order to accurately simulate experimental setups, a bore optimisation process is presented that, given either the input impedance or the impulse response of an acoustic tube, can arrive at an effective bore shape that represents the tube used in the experiment. Performing the optimisation in the time-domain ensures that the time-domain formulation of losses is consistent with subsequent attempts to model transient phenomena. Section 3 describes how both viscothermal losses as well as losses that occur at the open end of the tube due to sound radiation may be included in such models. Eventually optimisation may be performed in order to estimate model parameters (including bore geometry) based on time-domain pressure signals. Hence physics-based analysis attempts (see, e.g. [9]) may focus on characterising dynamic player-instrument interactions that occur at the mouthpiece during articulation.

2. MODELLING WAVE PROPAGATION IN TUBES

2.1. Frequency-domain model

We are considering a tube of length L and variable cross-sectional area $S = S(x), 0 \leq x \leq L$. The propagation of plane waves in tubular ducts, taking viscothermal losses into account, can be modelled in the frequency domain by [5, 10]

$$\partial_x P + ZV = 0, \tag{1a}$$

$$\partial_x (SV) + YSP = 0, \tag{1b}$$

where $P = P(\omega, x)$ is the acoustic pressure and $V = V(\omega, x)$ the particle velocity and ω denotes the frequency. The functions $Z = Z(\omega, x)$ and $Y = Y(\omega, x)$ are called the specific series impedance and the shunt addmitance, respectively. Specific formulas for Z and Y are given in [2].

To make the model complete, one has to impose boundary conditions at $x = 0$ and $x = L$. The left boundary, $x = 0$, is governed by a prescribed particle velocity $V(\omega, 0) = V_{\text{in}}(\omega)$ and the right, $x = L$, by a stipulated radiation impedance

$$Z_r(\omega) = \frac{P(\omega, L)}{S(L)V(\omega, L)}. \tag{2}$$

2.2. Time-domain model

In order to transfer the model (1) to the time domain, Z and Y can be split as [5]

$$Z = i\omega\rho + Z_v, \quad Y = \frac{i\omega}{\rho c^2} + Y_\theta, \tag{3}$$

where i denotes the imaginary unit, ρ the density of air and c the speed of sound in air. Since the cross-sectional area S does not depend on the time t, (1) transfers into

$$\partial_x p + \rho \partial_t v + z_v * v = 0, \tag{4a}$$

$$\partial_x (Sv) + \frac{S}{\rho c^2} \partial_t p + S y_\theta * p = 0, \tag{4b}$$

where $*$ denotes convolution with respect to the time variable t, and lower case function names refer to the time-domain versions of the capitalised functions. The boundary conditions in the time domain are given by $v(t,0) = v_{\text{in}}(t)$ at $x = 0$ and (2) transfers for $x = L$ to

$$p(t, L) = S(L)z_r * v(t, L). \tag{5}$$

In addition to the boundary conditions, we assume that the pressure p and the velocity v vanish for $t < 0$.

3. NUMERICAL SOLUTION OF THE TIME-DOMAIN MODEL

3.1. Semi-analytic computation of the losses

In order to deal with the convolutions $z_v * v$ and $y_\theta * p$ we pursue a similar approach as in [7] and approximate the functions Z_v and Y_θ, given in (3), by rational functions Z_v^K and Y_θ^K,

$$Z_v^K = R_0 + \sum_{k=1}^{K} \frac{R_k i\omega}{L_k + i\omega},\ Y_\theta^K = \sum_{k=1}^{K} \frac{G_k i\omega}{C_k + i\omega}. \tag{6}$$

One can show that for $\omega = 0$, $Y_\theta(0, x) = 0$, so a constant term in the approximation of Y_θ is not needed. R_k, L_k, G_k and C_k can be found by minimizing the differences $Z_v - Z_v^K$ and $Y_\theta - Y_\theta^K$ on a fixed set of frequencies by Newton's method [7]. Behold that the coefficients depend on the tube radius and hence on the spatial variable x, but in order to simplify notation, the x-dependency is not denoted. Hence, we approximate

$$z_v * v \approx R_0 v + \sum_{k=1}^{K} w_k \text{ and } y_\theta * p \approx \sum_{k=1}^{K} q_k \tag{7}$$

where the frequency-domain versions of w_k and q_k are given by the relations

$$W_k = \frac{R_k i\omega}{L_k + i\omega} V \text{ and } Q_k = \frac{G_k i\omega}{C_k + i\omega} P. \tag{8}$$

For $\lambda > 0$ and $H(t)$ denoting the Heaviside function, the Fourier transform of $H(t)e^{-\lambda t}$ is $\frac{1}{\lambda + i\omega}$, therefore

$$w_k(t) = R_k \int_0^t e^{-L_k(t-\tau)} \partial_t v(\tau) d\tau, \tag{9a}$$

$$q_k(t) = G_k \int_0^t e^{-C_k(t-\tau)} \partial_t p(\tau) d\tau. \tag{9b}$$

3.2. Computation of the boundary conditions

The boundary condition at $x = 0$ is $v(t, 0) = v_0(t)$. At the radiating end, we are using an approximation of Z_r of the form [3, Eq. (3.29b)]

$$Z_r(\omega) = \frac{i\omega R_r}{L_r + i\omega}. \tag{10}$$

The coefficients R_r and L_r depend on the radius and moreover on the flange of the tube. For example, the boundary condition discussed in [11, Eq. (9.9)] is of this form with $R_r = \frac{\rho}{S(L)\alpha_1}$ and $L_r = \frac{\alpha_2}{\alpha_1}$. In our study, we are considering flanged tube ends, hence the approximation (10) is accurate. When dealing with experimental results, more refined versions [12] should be considered, which will be optimised based on measured signals. With this approximation, the boundary condition at $x = L$ is, because of (5), given by the differential equation

$$S(L)R_r \partial_t v(t, L) = L_r p(t, L) + \partial_t p(t, L), \tag{11}$$

which can be discretised by finite differences.

3.3. Finite Difference scheme

We substitute (7) into (4) in order to arrive at

$$\partial_x p + \rho \partial_t v + R_0 v + \sum_{k=1}^{K} w_k = 0, \tag{12a}$$

$$\partial_x (Sv) + \frac{S}{\rho c^2} \partial_t p + S \sum_{k=1}^{K} q_k = 0. \tag{12b}$$

We compute approximations p_m^n and v_m^n to the solutions p and v of (12), respectively, at discrete points (t_n, x_m) in time and space, where $t_n = n\Delta_t$, $n = 0, 1, 2 \ldots$ and $x_m = m\Delta_x$ for $m = 0, \ldots, M$ and $L = M\Delta_x$, for fixed Δ_t and Δ_x.

One iteratively computes p_m^{n+1} and v_m^{n+1} from results obtained at previous time steps. In order to derive a discrete version of (12), we need the following discretisations of (9),

$$w_{k,m}^{n+1} \approx\ e^{-L_{k,m}\Delta_t} w_{k,m}^n + R_{k,m}(v_m^{n+1} - v_m^n) e^{-L_{k,m}\frac{\Delta_t}{2}}, \tag{13}$$

and

$$q_{k,m}^{n+1} \approx\ e^{-C_{k,m}\Delta_t} q_{k,m}^n + G_{k,m}(p_m^{n+1} - p_m^n) e^{-C_{k,m}\frac{\Delta_t}{2}}. \tag{14}$$

The boundary condition on the left gives for v_0^{n+1}

$$v_0^{n+1} = v_{\text{in}}^{n+1}. \tag{15}$$

Equation (12a) is discretised by finite differences and, using (13), one can compute $v_m^{n+1}, m = 1, \ldots, M$ from

$$\frac{p_m^n - p_{m-1}^n}{\Delta_x} + \rho \frac{v_m^{n+1} - v_m^n}{\Delta_t} + R_{0,m} v_m^{n+1} + \sum_{k=1}^{K} \left[e^{-L_{k,m}\Delta_t} w_{k,m}^n + R_{k,m}(v_m^{n+1} - v_m^n) e^{-L_{k,m}\frac{\Delta_t}{2}} \right] = 0. \tag{16}$$

Discretising (12b), and using (14), one computes p_m^{n+1}, $m = 0, \ldots, M-1$ from

$$\frac{S_{m+1} v_{m+1}^{n+1} - S_m v_m^{n+1}}{\Delta_x} + \frac{S_m}{\rho c^2} \frac{p_m^{n+1} - p_m^n}{\Delta_t} + S_m \sum_{k=1}^{K} \left[e^{-C_{k,m}\Delta_t} q_{k,m}^n + G_{k,m}(p_m^{n+1} - p_m^n) e^{-C_{k,m}\frac{\Delta_t}{2}} \right] = 0. \tag{17}$$

Taking finite differences in (11) yields for $m = M$

$$S_M R_r \frac{v_M^{n+1} - v_M^n}{\Delta_t} = L_r p_M^{n+1} + \frac{p_M^{n+1} - p_M^n}{\Delta_t}, \qquad (18)$$

from which p_M^{n+1} can be computed. Finally for $k = 1, \ldots, K$ and $m = 1, \ldots, M$, $w_{k,m}^{n+1}$ and $q_{k,m}^{n+1}$ are updated by (13) and (14), respectively.

4. BORE RECONSTRUCTION

In this section we use the time-domain model and its numerical implementation to estimate shape parameters of a tubular duct using the input impedance,

$$Z_{\text{in}}(\omega) = \frac{P(\omega, 0)}{S(0)V(\omega, 0)}. \qquad (19)$$

Having solved the numerical scheme (15) - (18) for timesteps $t_n = n\Delta_t$, $n = 0, \ldots, N-1$, we compute an approximation of the input impedance at frequencies $\omega_k = \frac{k}{N\Delta_t}$, $k = 0, \ldots, N-1$ with the fast Fourier transform (fft) by

$$Z_{\text{in,num}}(\omega_k) = \frac{\{\text{fft}\,\mathbf{p_0}\}_k}{S_0\{\text{fft}\,\mathbf{v_0}\}_k}, \qquad (20)$$

where $\mathbf{p_0} = \{p_0^n\}_{n=0}^{N-1}$ and $\mathbf{v_0} = \{v_0^n\}_{n=0}^{N-1}$. We assume that the cross-sectional area $S(x)$ depends on a set of parameters $\Pi = \{\pi_1, \ldots, \pi_d\}$, hence $S(x) = S(x, \Pi)$. Since the function S appears in the models (1) and (4) as well as in the numerical scheme (15) - (18), also the functions P, V, p and v and the numerical approximations p_n^m and v_n^m depend on the shape parameter set Π. Moreover, all functions further depend on the length L of the tube. Of course, this also holds for the computed input impedance, hence $Z_{\text{in,num}}(\omega_k) = Z_{\text{in,num}}(\omega_k, \Pi, L)$. Given a reference impedance $Z_{\text{in,ref}}$ of a given bore, we want to estimate the parameter set Π and the length L such that $Z_{\text{in,num}}$ is close to the reference value at $\omega_k, k = 0, \ldots N^* - 1$. The numerically computed input impedance is an approximation to the input impedance only for $k < \frac{N}{2}$ and more accurate for low frequencies, hence we only take $\omega_k, k < N^*$ into account for the comparison, where $N^* < \frac{N}{2}$. To this end, we are minimizing the function

$$f(\Pi, L) = \sum_{k=0}^{N^*-1} |Z_{\text{in,ref}}(\omega_k) - Z_{\text{in,num}}(\omega_k, \Pi, L)|^2 \qquad (21)$$

with respect to the variables $\pi_1, \ldots, \pi_d$ and L. Starting from initial values Π^0 and L^0, the Levenberg-Marquardt method [13, Chapter 10] iteratively computes Π^j and L^j, $j = 1, 2, \ldots$, which are approximations to a local minimum of f. The algorithm stops as soon as the difference between two consecutive approximations is smaller than a given tolerance. Convergence to a local minimum requires a good starting point for the iteration. In this case any cylindrical tube with a plausible wind instrument radius proved sufficient for convergence to the desired minimum[1].

[1] Alternatively the Rosenbrock optimisation method [9] can be used that may identify the global minimum of a search space.

5. NUMERICAL EXPERIMENTS

In our numerical experiments, we model an alto saxophone mouthpiece adjuncted to a neck as a concatenation of a cylinder and a frustum of a cone. The mouthpiece itself is modeled as a cylinder attached to a frustum, the neck as a single frustum with the same aperture. The radial bore profile of a concatenation of a cylinder and a cone is described by

$$r(x, \Pi) = \pi_1 \qquad \text{for } 0 \le x \le \pi_4 \qquad (22a)$$

$$r(x, \Pi) = (x - \pi_4)\pi_3 + \pi_2 \qquad \text{for } \pi_4 \le x. \qquad (22b)$$

The cross-sectional area is given by $S(x) = r(x)^2\pi$. The reference input impedance is simulated with ARTool [14], using a frequency-domain plane wave model that includes viscothermal losses [4]. The corresponding parameters of the model (22), as well as the length of the tube, are given in the last column of Table 1.

We are using the profile model (22) in a first experiment, hence we want to estimate the parameters $\pi_1, \ldots, \pi_4$ as well as the length L. We start from the initial shape of a cylinder of radius 5mm and length 230mm. After 61 steps, the Levenberg-Marquardt method stops. Figure 1 shows the initial input impedance and the initial bore profile, as well as the input impedances and bore profiles of the reference and the result after the parameter optimization. The results of the Levenberg-Marquardt method are summarized in Table 1.

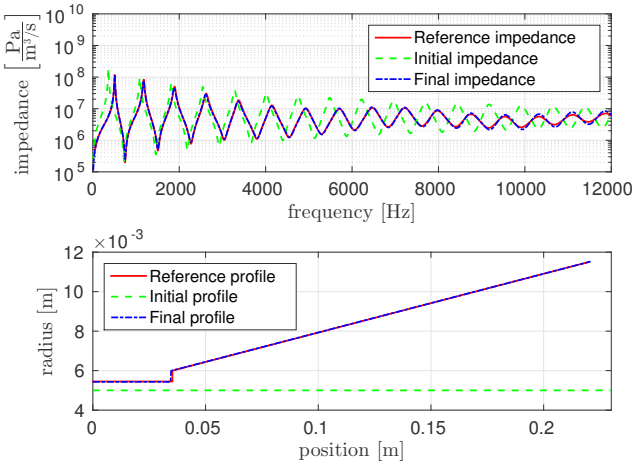


Figure 1. Reconstruction of instrument length, radii, and aperture of the cylinder-cone model (22).

Without prior information on the bore, one can model the profile by a linear interpolation through the points (x_l, π_l), $l = 1, \ldots, d$ for fixed $0 = x_1 < x_2 < \ldots < x_d$. Therefore

$$r(x, \Pi) = \frac{x - x_{l-1}}{x_l - x_{l-1}}(\pi_l - \pi_{l-1}) + \pi_{l-1} \text{ for } x_{l-1} \le x < x_l \qquad (23a)$$

$$r(x, \Pi) = \frac{x - x_{d-1}}{x_d - x_{d-1}}(\pi_d - \pi_{d-1}) + \pi_{d-1} \text{ for } x_d < x. \qquad (23b)$$

In a second experiment we take 20 uniformely distributed nodes, $x_l = \frac{l-1}{19}220 \times 10^{-3}, l = 1, \ldots, 20$. Again, as starting

	Initial (step 0)	Final (step 61)	Reference
L	230 mm	220.84 mm	220 mm
π_1	5 mm	5.4329 mm	5.45 mm
π_2	5 mm	5.9861 mm	6 mm
π_3	0	0.029739	0.029778
π_4	40 mm	34.566 mm	35.3 mm

Table 1. Summary of the reconstruction algorithm.

point we choose a cylinder of radius 5mm and length 230mm, i.e. $\pi_l^0 = 5 \times 10^{-3}$ for $l = 1, \ldots, 20$ and $L^0 = 230 \times 10^{-3}$. After 41 steps, the stopping criterion is satisfied. The initial, the reference and the final input impedances as well as the corresponding bore profiles are shown in Figure 2.

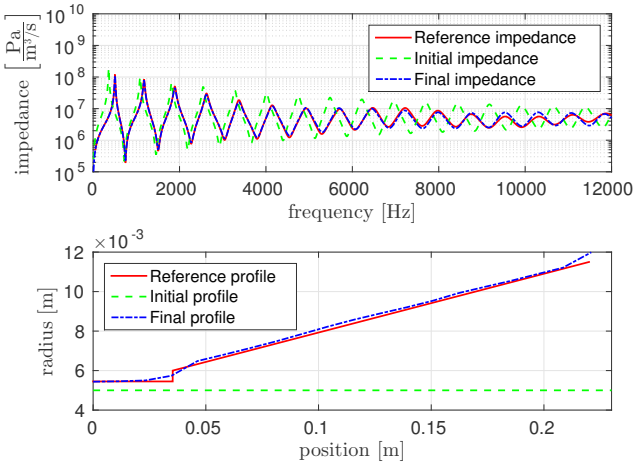


Figure 2. Reconstruction of instrument length and radii of linear interpolation model (23).

6. CONCLUSIONS

In this paper we study time-domain models for acoustic tubes, introducing a semi-analytic variant for the computation of viscothermal losses. It is demonstrated that impedances computed by time-domain models are in good agreement with reference impedances computed by plane wave frequency-domain models. Further, we describe an algorithm for the estimation of bore shape parameters using a time-domain model. Fine tuning the bore radius also updates the relevant parameters in the terms that model viscothermal and radiation losses. Current research focuses on performing the optimisation based on time-domain pressure measurements in order to estimate both the geometry of the instrument, as well as the parameters related to the player-instrument interaction.

7. ACKNOWLEDGEMENTS

This research is supported by the Austrian Science Fund (FWF): P28655-N32.

REFERENCES

[1] A. H. Benade, "On the propagation of sound waves in a cylindrical conduit," *JASA*, vol. 44, no. 2, pp. 616–623, 1968.

[2] D. H. Keefe, "Acoustical wave propagation in cylindrical ducts: Transmission line parameter approximaions for isothermal and nonisothermal boundary conditions," *JASA*, vol. 75, no. 1, pp. 58–62, 1984.

[3] L. R. Rabiner and R. W. Schafer, *Digital processing of speech signals*, ser. Prentice-Hall signal processing series. Englewood Cliffs, N.J. Prentice-Hall, 1978.

[4] J. A. Kemp, "Theoretical and experimental study of wave propagation in brass musical instruments," Ph.D. dissertation, University of Edinburgh, 2002.

[5] S. Bilbao and R. Harrison, "Passive time-domain numerical models of viscothermal wave propagation in acoustic tubes of variable cross section," *JASA*, vol. 140, no. 1, pp. 728–740, 2016.

[6] S. C. Thompson, T. B. Gabrielson, and D. M. Warren, "Analog model for thermoviscous propagation in a cylindrical tube," *JASA*, vol. 135, no. 2, pp. 585–590, 2014.

[7] S. Bilbao, R. Harrison, J. Kergomard, B. Lomard, and C. Vergez, "Passive models of viscothermal wave propagation in acoustic tubes," *JASA*, vol. 138, no. 555, 8 2015.

[8] V. Chatziioannou and A. Hofmann, "Physics-based analysis of articulatory player actions in single-reed woodwind instruments," *Acta Acust united Ac*, vol. 101, no. 2, pp. 292–299, 2015.

[9] V. Chatziioannou and M. van Walstijn, "Estimation of clarinet reed parameters by inverse modelling," *Acta Acust united Ac*, vol. 98, no. 4, pp. 629–639, 2012.

[10] R. Caussé, J. Kergomard, and X. Lurton, "Input impedance of brass musical instrumentscomparison between experiment and numerical models," *JASA*, vol. 75, no. 1, pp. 241–254, 1984.

[11] S. Bilbao, *Numerical Sound Synthesis*. Chichester: Wiley, 2009.

[12] F. Silva, P. Guillemain, J. Kergomard, B. Mallaroni, and A. Norris, "Approximation formulae for the acoustic radiation impedance of a cylindrical pipe," *J Sound Vib*, vol. 322, no. 12, pp. 255 – 263, 2009.

[13] J. Nocedal and S. J. Wright, *Numerical optimization*, ser. Springer Series in Operations Research and Financial Engineering. Berlin: Springer, 2006.

[14] A. Braden, D. Chadefaux, V. Chatziioannou, S. Siddiq, C. Geyer, S. Balasubramanian, and W. Kausel, "Acoustic research tool (ARTool)." [Online]. Available: http://artool.sourceforge.net/

Numerical Simulations of Transverse Modes in the Initial Transient of an Organ Pipe

Jost Fischer,[†] *Rolf Bader*

Institute of Systematic Musicology, University of Hamburg
[†]jost.leonhardt.fischer@uni-hamburg.de

ABSTRACT

A new approach to investigate the role of transverse modes in the initial transient of an organ pipe is presented. By numerical simulations solving the compressible Navier-Stokes equations under suitable boundary and initial conditions, it is possible to retrace initial transient process of an organ pipe including the generation and attenuation of transverse modes in the resonator as well as their radiation into free space. Furthermore, their contribution to the forming process of the fundamental frequency of the organ pipe is discussed. Utilizing methods of coarse-graining and of spectral analysis, a wide range of aspects of the dynamics of transverse modes in the initial transient are considered. In particular the nonlinear damping processes in the air column and in the boundary layers close to the inner walls of the organ pipe's resonator are spotlighted. The presented numerical approach which is verified by measurements at real instruments allows to study the initial transient of organ pipes in extraordinary detail. It enhances the understanding of the underlying first principles of initial sound generation of organ pipes and other wind instruments.

On the Second Register's Playability of the Clarinet: Towards a Multicriteria Approach

Alexis Guilloteau,[1†] *Philippe Guillemain,*[1] *Jean Kergomard,*[1] *Michael Jousserand*[2]

[1]Laboratoire de Mécanique et d'Acoustique, CNRS
[2]Buffet-Crampon
[†]guilloteau@protonmail.com

ABSTRACT

During the centuries of the clarinet evolution process, woodwind makers had to empirically understand how to choose a design that allows a good control of its musical attributes. The most important ones can be defined as, intonation, ease of playing and tone color. A preliminary study focuses on the efficiency of a gradient based optimization algorithm to consider the intonation criterion. Promising results were obtained by using the first two resonance frequencies of the input impedance resulting from the geometry being optimized. It also shows the need of introducing a second criterion dedicated to the ease of playing for the second register, as the reactive attenuation of the first impedance peak can't be controlled by only the targeted resonance frequencies. Indeed, in this preliminary study, the first impedance peak remained higher than the second one which prevents register change. This article presentation will propose different objective definitions of the required ratios between first and second impedance peaks. Then a multicriteria optimization algorithms (goal attainment method) is applied to take advantage of each, intonation and second register playability criteria. Two prototypes were realized and results were validated by musicians and impedance measurements. They confirm the diversity of a register hole functionalities. A dual propensity to balance inter-register inharmonicity and second register playability can be counterproductive and demonstrate the real benefit of increasing and dedicating register holes, which is not the case on the current Boehm clarinet.

An Investigation of the Clarinet Reed's Vibrating Surface Area

Whitney Coyle,[†] *Carter Richard*

Rollins College
[†]wcoyle@rollins.edu

ABSTRACT

A hope for many musicians is that they will find an instrument that will play "in-tune" without significant effort on their part. Thus, in order to design a "playable" clarinet the tuning tendencies and tuning homogeneity throughout the range of the instrument must be known. The playing frequencies of the instrument depend on several control parameters including the blowing pressure, reed opening and the instrument's input impedance. Using analytical formulas we can now rapidly predict these frequencies based on the different control parameters and resonance frequencies. The analytical formulas take into account four effects known to influence the playing frequency. The effects are each included separately: reed-induced flow, the reed dynamics, the inharmonicity of the resonator and the temperature gradient within the clarinet. The effect of reed-induced flow has been found to be the cause of the largest frequency shift and it is therefore necessary to ensure the accuracy of the input parameters for this effect. One user input constant for this effect is the vibrating surface area of the reed, Sr. Due to the nature of playing this reed instrument, this value is difficult to determine experimentally. Past studies give a range of values for this parameter, Sr, but none agree what this value should be. The analytical formulas also assume a uniform, rectangular region that is vibrating as a cantilever beam. This may not be an accurate enough description of the reed vibrations and the model could be improved through better understanding of this parameter. This paper will present a brief parameter study showing the sensitivity of the current model to the parameter choice for reed-induced flow and detail the techniques used to investigate more closely the actual value of Sr including the use of vibration mode analysis of the reed using ESPI (electronic speckle pattern interferometry) as well as high-speed camera analysis of the reed while it is being driven by an artificial blowing machine.

Beneficial Aspects of Toneholes Undercutting Applied to Clarinet Design

Alexis Guilloteau,[1†] *Philippe Guillemain,*[1] *Michael Jousserand*[2]

[1]Laboratoire de Mécanique et d'Acoustique, CNRS
[2]Buffet-Crampon
[†]guilloteau@protonmail.com

ABSTRACT

Traditionally woodwind makers use holes with undercuts for many reasons. First of all, when the instrument is crafted, it is easier to round the hole inner edge than change its place to correct the tuning. In the case of the development of a numerical optimization heuristic, known analytical model don't take into account undercutting in holes, so that they often simply consider a dual cylinder connection. For this specific geometry a turbulent flow layer effect has already a prominent dissipative action, even for a low excitation pressure equivalent to a piano musical dynamics. This study has experienced those aspects. We first make a prototype of logical clarinet, with numerical optimization tools, without taking into account any undercutting aspects for tone and register holes. Then experienced clarinetists who played this prototype, had difficulties to play some notes they were also lower in tune. We realized that those specific notes were also corresponding to holes with particularly tight diameters (around 4 mm), and enhance the thesis of a modified reactive effect (or equivalent length) due to the vena contracta phenomenon when turbulent flow appears in the hole. To fix this problem, and because it's better to not make models more costly, we've decided to replace every sharp edge of holes with specific shapes of undercuts depending on the hole radius due to making constraints. They had already proven their efficiency for significantly reduce nonlinear effects, so that we supposed to only take linear contribution in the acoustic models. A specific theoretical discussion is proposed and some analytical formulas are provided corresponding to the chosen undercutting shapes. Then, after the making process of a second optimized prototype which include those previous specifications (not discussed here), we expose some measurements on artificial mouth. Indeed, the inter-thresholds distance (oscillation and extinction) was significantly improved, for a corresponding hole diameter and similar excitation parameters. This inter-thresholds distance should also be an interesting objective measurement tool to evaluate the clarinet ease of playing, but deserves to be completed with a psychoacoustic study.

Who Does Control the Attack Transients of the Recorder, the Musician or the Instrument Maker?

Augustin Ernoult,[†] *Benoît Fabre*

LAM, Institut Jean le Rond d'Alembert, Université Pierre et Marie Curie
[†]ernoult@lam.jussieu.fr

ABSTRACT

Physical models of the recorder allow predicting their quasi-stationary behaviors. The evolution of sound qualities is predictable and well understood when control parameters, such as mouth pressure, evolve slowly compared to the response time of the instrument. However, the musician continuously varies his control to articulate the music. For example, during the attack transients, an abrupt increase of the supply pressure induces the progressive development of acoustic oscillations. The musician varies the characteristics of attack transients as a function of the musical context. The control of these attacks is therefore an essential aspect of learning to play the recorder and of making this instrument. They are few studies on recorder attack and they are generally only qualitative and/or focus on one specific aspect observed on few data. The first aim of this study is to identify and quantify the characteristics of the attack transients of recorder-like instruments in musical context in the temporal and spectral domains. The second one is to understand how the musicians and the instrument makers can control these characteristics.

Attack transients produced by one novice and two experienced players on four hand-made recorders (bass, alto, soprano and sopranino) are compared. This analysis leads to dividing recorder attack transients in two phases: the onset of the acoustic oscillation and its growth. The duration and the spectral content of the onset varies from one musician to another. A model is proposed to interpret the origin of the input of energy necessary to initiate the oscillation of the instrument. During its formation, the jet can been interpreted as a flow source coupled with the resonator inducing acoustic oscillations. This model enhances the relative importance of control parameters and instrument design on the onset characteristics. The growth duration, similar for the three players, is imposed by the instrument properties which are themselves controlled by the instrument makers. The growth is also characterized by the simultaneous evolution of different spectral components. A specific empirical study allows identifying the stability conditions for each components. This characterization allows interpreting the spectral evolutions observed during experimental attack transients with energetic considerations. These stability conditions being imposed by the manufacture, the musician can or not encourage the development of high components by controlling the temporal evolution of the mouth pressure.

Music or Mechanics? Understanding the Role of a Bagpiper's Arm

Cassandre Balosso-Bardin,[1,2] *Augustin Ernoult,*[3] *Patricio de La Cuadra,*[4] *Ilya Franciosi,*[3] *Benoît Fabre*[3]

[1]University of Lincoln
[2]Chaire Geste-Acoustique Musique, Sorbonne Universités
[3]Institut Jean le Rond d'Alembert, Université Pierre et Marie Curie
[4]Pontificia Universidad Católica de Chile
[†]cassandre.balbar@gmail.com

ABSTRACT

Despite their many organological and esthetical differences, bagpipes are all played thanks to the movement of the arm on a bag, creating enough pressure to activate the reeds and produce sound. Repertoires, scales and registers vary according to the instruments and their musical cultures, going from a fully chromatic scale over two octaves (such as the uilleann pipes from Ireland) to a diatonic scale within a range of a sixth (such as the Greek tsampouna or the Tunisian mizwid). Unlike fingerings and melodic ornamentation, the musician's arm technique is rarely discussed in bagpipe literature, nor is it particularly verbalized during a piper's tuition. According to Simon McKerrell, 'each player learns it individually and develops their own technique' [1]. Despite this lack of verbalization, bagpipe experts seem to agree that the breathing technique and the bag are essential elements of their playing [1],[2].

In this research, carried out during the Geste-Acoustique-Musique program (Sorbonne Universités, Paris), we endeavor to understand how the bagpiper exerts control on his/her bag. Understanding this may enhance our comprehension of the importance of the arm in a musical context. Our main questions are: what role does the arm have in the control of the instrument? Is the bag controlled with musical intention? Leading from this, further questions can be asked such as how does this influence the instrument's repertoire and the musician's performance?

To answer these questions, we will present data collected during three experiments in different cultural contexts and with musicians of different levels. Using acoustic equipment, we were able to measure the insufflated airflow, the pressure in the bag, the angle of the arm as well as make videos and record the sound. In order to complement our scientific data, we carried out an online questionnaire, which allowed us to gather information on the perception of musicians and their subjective impressions on the control of their instrument. With acoustic measurements, qualitative data and an ethnomusicological framework, this research offers a multidimensional and interdisciplinary study of the control of the bagpipe's bag.

REFERENCES

[1] S. McKerrell, "Sound performing: sound aesthetics among competitive pipers," in International review of the aesthetcs and sociology of music, 42/1. Croatian Musicological Society, 2011.

[2] T. Rice, "Evaluating artistry on the Bulgarian bagpipes," in Ethnomusicological encounters with music and musicians. Farnham, Ashgate Publishing Limited, 2011.

Reconstructing an 1832 Boehm Conical Flute: an Interdisciplinary Collaboration

Cassandre Balosso-Bardin,[1†] *Patricio de La Cuadra,*[2] *Camille Vauthrin,*[3] *Benoît Fabre*[4]

[1]University of Lincoln
[2]Pontificia Universidad Católica de Chile
[3]Chaire Geste-Acoustique-Musique, Sorbonne Universités
[4]Université Pierre et Marie Curie Paris 6
[†]cassandre.balbar@gmail.com

ABSTRACT

In 1832, flute-maker Theobald Boehm launches a brand new flute after conducting a series of acoustical experiments. The result was a flute with a conical bore, and a revolutionary key-system borrowing the best elements from different contemporary flute-makers to accommodate a complete modification of the instrument's tone holes. However, he did not patent his 1832 model and "left it free for use and imitation" [1]. This means that despite its short life-span – by 1848 Boehm had already invented the modern flute with a cylindrical bore – and much rivalry from other manufacturers, the 1832 flute was built by many renowned makers until the first decades of the 20th century. Thus, although this instrument – that we named transition flute – was based on a specific model developed by Boehm, there was space for a wide variety of models. The lack of standardization means that when flautists play on transition flutes today, they are confronted to a wide variety of choice as each instrument is different.

When commissioned by musicians to build a transition flute based on the 1832 conical model, the Flutes Roosen workshop was faced with a dilemma as the flutes all varied significantly. The following article shows how the geometrical and acoustic study of four transition flutes allowed us to propose a computer-generated model that simultaneously respected Boehm's inventions and took into account the variety of models available. Rather than produce an exact copy of one instrument, we opted for recreating a new flute based on acoustic and geometrical measurements of different transition flutes and on Boehm's desired improvements, detailed in his 1847 Essay on the construction of flutes [1].

REFERENCES

[1] T. Boehm, An essay on the construction of flutes. London: Rudall, Carte & co, 1832.

Finite Element Modeling of Sound Radiation at the Open End of a Conic Frustum

Song Wang,[†] *Gary Scavone*

Computational Acoustic Modeling Laboratory (CAML), McGill University
[†]song.wang5@mail.mcgill.ca

ABSTRACT

The linear acoustic response of wind music instruments can be modeled by approximating their air column geometries with piece-wise combinations of cylindrical and conical sections. For sections that couple to a radiating domain (toneholes and open ends), theoretical load impedance or reflection coefficient characterizations must be used. While theoretical approximations for the open end of flanged and unflanged cylinders are well documented, none exist for flanged or unflanged cones. In this study, we make use of the finite element method to evaluate the reflection coefficient at the open end of a flanged and unflanged conic frustum of varying taper angle and radius, with the goal of providing approximate formulas for the corresponding reflection coefficient magnitudes and length corrections.

Radiation Characteristics of Grand Piano Soundboards in Different Stages of Production

Niko Plath,[†] *Plath Pfeifle, Christian Koehn, Rolf Bader*

University of Hamburg
[†]niko.plath@uni-hamburg.de

ABSTRACT

In an on-going project, a series of radiation measurements is taken on two concert grand piano soundboards. The pianos are accompanied during the entire production process. Measurements are performed at seven discrete stages of the instruments' construction. A noninvasive microphone array method is utilized for the present work. The array consists of 105 microphones successively placed parallel to the soundboard, resulting in a total number of 1289 microphones covering the complete surface. The soundboard is excited using an acoustic vibrator at 15 positions associated with string termination points on the bass and main bridge. Sensors at the driving points measure input force and acceleration. Impulse responses are obtained using a sweeped sine technique. The measured sound pressure is back-propagated to the radiating soundboard surface using a minimum energy method. For the final production stage, i.e. the instrument being completely assembled and concert tuned, a different measurement setup is used: utilizing a binaural dummy head within an anechoic chamber, single notes are recorded. The analysis is divided into three parts: First, radiation characteristics are described as radiated power and radiation efficiency (after Suzuki 1986), as well as the ratio of sound pressure and velocity of the soundboard (after Giordano 1998), to be compared to obtained deflection shapes. Impact of the attachment of components, such as bridge and ribs, on the critical frequency range where the coincidence effect occurs can be given. Second, measured sound pressure is forward-propagated into space surrounding the piano and far field radiation patterns are presented for the first resonances, most important for aural impression. Third, binaural measures (interaural level difference, interaural phase difference and binaural quality index) are utilized to connect the obtained physical data to the sensation of directivity and apparent source width as perceived by a player. The empirical findings will contribute to a software tool, based on a real time physical model, to help piano makers estimate the impact of design changes on the generated sound.

Experimental Study of the Plectrum Shape and the Key Velocity on the Harpsichord's Sound

Jean-Loïc Le Carrou,[†] *Arthur Paté, Arthur Givois, Alexandre Roy*

Sorbonne Universités, UPMC Univ Paris 06, CNRS, LAM / Institut ∂'Alembert
[†]jean-loic.le_carrou@upmc.fr

ABSTRACT

The sound of the harpsichord originates from the acoustic radiation of a soundboard excited by a vibrating string through coupling. The string is put into motion by a complex interaction with a plectrum. This small piece of feather or plastic catches the harpsichord makers' and players' greatest attention: The voicing process (working of the shape, dimensions and material or the plectrum) is a crucial stage of the making and adjustment of the instrument. Different geometrical and mechanical features alter indeed the initial conditions of the vibrating string, hence presumably the sound. The study presented here aims at investigating the influence of the musician's gesture and the plectrum shape on the sound of the instrument. Different speeds of the finger on the key (0.05 m/s to 0.3 m/s), and four different plectrum shapes (manually carved, exaggerating features found in actual plectra) are investigated. A complete set of measurements is carried out for each speed / shape configuration during the plectrum / string interaction phase: microphones for the radiated sound, optical contactless sensors for the displacement and velocity of the string vibration, displacement sensor for the key depression, a high-speed camera for the plectrum deflection obtained by image processing and a robotic finger for pressing the key with different velocities. Whereas the speed at which the finger presses the key does not seem to have a strong influence, the results indicate a clear influence of the plectrum shape on its deflection, as well as on the initial conditions of the string (deduced from the string's motion), and on the loudness of the sound.

Controlling Piano Tone by Varying the "Weight" Applied on the Key

Caroline Traube,[1†] *Manon Moulin,*[2] *Felipe Verdugo*[1]

[1]Université de Montréal
[2]Université de Mons
[†]caroline.traube@umontreal.ca

ABSTRACT

To control the tone of their instrument, piano teachers at University of Montreal recommend to act on the double escapement action by modifying the "weight" applied on the key. When the pianist uses more weight, the key is pressed to the bottom of the keyboard and the pianists feel a bump and a sensation of friction when they pass the escapement. When using less "weight", they play more at the surface of the keyboard. The present study aims to verify if this variation of weight applied on the keys has an impact on the double escapement action and on the piano tone. Two series of recordings were analysed. In a first series, pianists played a short musical phrase varying several control parameters (with/without weight, with/without pedal) and at several intensity levels. In a second series of recordings, isolated notes were played in the same conditions. Simultaneously to the recording of the piano tones, a video image of double escapement grand piano action was captured with a camera placed inside the piano. The video monitoring shows that the piano action behaves differently depending on the amount of weight. More weight on the keystroke provokes friction between the jack and the hammer knuckle. The tone is different with or without this friction. In fact, the analysis of the audio data shows that the piano tones produced with and without weight differ along several timbre-related acoustical descriptors (temporal and spectral features). The main parameters which are modified are related to the quality of the attack.

The Production of Phantom Partials in the Sound of the Modern Piano

Camille Adkison, Eric Rokni, Lauren Nelder, Thomas Moore[†]

Dept. of Physics, Rollins College
[†]tmoore@rollins.edu

ABSTRACT

The spectra of the sound produced by the modern piano includes frequency components that are not equal to any of the resonance frequencies of the transverse string motion. These frequency components, termed phantom partials, have been attributed to forced longitudinal vibrations in the string that occur at the sums and differences of the frequencies of the string resonances. The frequencies of the phantom partials do not coincide with the frequency components produced by the transverse motion of the string because the stiffness of the string causes the frequencies of the string overtones to differ from an exact harmonic relationship. This inharmonicity increases with the mode number, so that the frequency of the phantom partial produced by summing the frequencies of the mth and nth transverse mode of the string falls below the resonance frequency of the (m+n) mode.

Although the frequencies of the phantom partials appear to correspond to a sum or difference of the resonant frequencies of the piano string, it is not obvious that the nonlinear coupling that produces these frequency components occurs only in the string. It is possible that the motion of the soundboard, bridges, case, and frame all contribute to the production of phantom partials, although to our knowledge the origins of phantom partials have never been systematically investigated.

We have experimentally investigated the production of phantom partials that occurs in a modern grand piano and found that while the string plays an important role in the production of phantom partials, other parts of the piano also contribute. Mechanical excitation of the bridge and soundboard at two discrete frequencies can produce phantom partials even when the driving frequencies do not correspond to resonance frequencies of the string and/or all string motion is damped. A comparison of the power in a phantom partial when a string is free to vibrate versus when the string is damped provides an estimate of the importance of the piano structure to the production of phantom partials.

KEYNOTE ADDRESS

Acoustics of Pianos: an Update of Recent Results

Antoine Chaigne[†]

Department of Music Acoustics (IWK), University of Music and Performing Arts, Vienna, Austria
[†]antchaigne@gmail.com

ABSTRACT

In the first part of this lecture, recently published results related to the acoutics of pianos will be presented. The topics include the piano action, the stringing scheme, the vibroacoustics of the soundboard and the radiation of the instrument. The second part is devoted to the simulations of pianos, including historic instruments of the 19th century. These instruments show a large variety of building parameters, in terms of geometry, materials and making process. Therefore, the objective of establishing links between their physical properties and their tonal quality is challenging. Examples will be given related to the influence of string's parameters (amplitude, tension, coupling points), soundboard design and listening environment. In addition, some aspects on the analysis and perception of piano tones will be presented and discussed.

Physics-Based High-Efficiency Analysis of Membranophones Using a Spectral Method

Yozo Araki,[1†] *Toshiya Samejima*[1]

[1]Faculty of Design, Kyushu University
[†]araki.yozo.051@m.kyushu-u.ac.jp

ABSTRACT

We propose a high-efficiency analysis method of membranophones such as drums or timpani. The main strategies to make the computation more efficient are as follows. Since most of membranophones have axisymmetric shapes, they can be analyzed theoretically in the circumferential direction by using Fourier series expansions. This strategy can reduce the spatial dimensions to be discretized. Then, the rotating section is analyzed by using spectral methods, which are high-accuracy numerical methods. Fourier series expansions and spectral methods in the analyses of membranophones are formulated for each component of the membranophones. Although spectral methods have restrictions concerned with applicable shapes and boundary conditions, applying several techniques enables the methods to deal with complex shapes and boundary conditions. Components of a membranophone are coupled through coupling conditions of pressure and velocity. It is shown that the proposed method can analyze membranophones accurately with a small number of degrees of freedom and low memory requirement.

1. INTRODUCTION

Sound analysis and synthesis of musical instruments are useful for their acoustic design and playing virtual instruments. In particular, physics-based synthesis has been studied and developed actively in recent years. It can reflect changes of physical parameters such as properties of materials or excitation conditions to timbres[1]. That is a difficult task with synthesis based on signal processing. However, physics-based synthesis requires much computational cost for spatial discretization in general numerical calculations. Thus, most of studies concerned with it focused on implementation of computational methods on hardware such as GPU or FPGA rather than the methods themselves[2, 3].

This study deals with membranophones such as drums and timpani, and proposes an efficient computational method of them. In general, most of membranophones have axisymmetric shapes. Therefore, their vibration displacements and acoustic velocity potentials can be expanded in Fourier series in the circumferential direction. It means that it is unnecessary to discretize the circumferential direction, thereby, the spatial dimension to be discretized can be reduced.

The rotating section, which generates the total axisymmetric shape, is analyzed by using spectral methods, which are high-accuracy numerical methods. Global interpolation functions used in spectral methods enable the methods to offer greater accuracy than other numerical methods such as finite difference methods (FDMs) or finite element methods (FEMs), which use local interpolation functions. In other words, when only lower accuracy is required, spectral methods demand less computational cost than the others. We have formulated spectral methods for vibration analyses of a stretched circular plate and cylindrical shell, and shown that they could analyze the vibration fields more accurately with a small number of degrees of freedom and less memory than FEMs in our previous studies[4, 5].

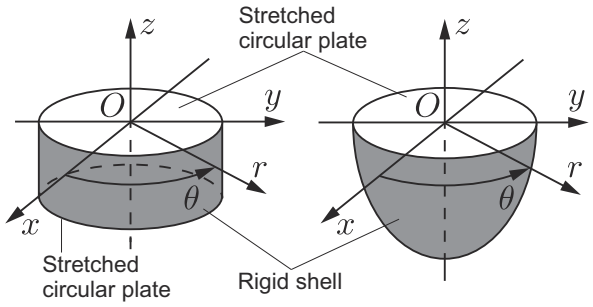


Figure 1. Analysis models of a drum (left) and timpani (right).

Although spectral methods have restrictions concerned with applicable shapes and boundary conditions, applying several techniques enables the methods to deal with complex shapes and boundary conditions. After applying Fourier series and spectral methods to each component of membranophones, obtained matrix equations are coupled through coupling conditions of velocity and pressure.

Finally, calculated and measured frequency response functions are compared and it is shown that the proposed method is effective as an analysis method of membranophones.

2. PHYSICAL MODEL

Analysis models are shown in Fig. 1. Although the shells slightly vibrate in a real situation, they are considered rigid for ease in this study.

2.1. Stretched Circular Plate

A head of a membranophone is modeled as a stretched circular plate since it has both tension and bending stiffness. The

vibration equation of the stretched plate is expressed as

$$(B\nabla^4 - T\nabla^2)(\zeta + \beta\frac{\partial\zeta}{\partial t}) + \rho_{\mathrm{p}}\frac{\partial^2\zeta}{\partial t^2} = f + p, \qquad (1)$$

where ζ, B, T, ρ_{p}, f and p are vibration displacement, bending stiffness, tension, surface density, mechanical excitation pressure and sound pressure, respectively. β is a coefficient determining damping effect proportional to stiffness.

2.2. Sound Field

The wave equation governing sound fields inside and outside a membranophone cavity is expressed as

$$\nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = 0, \qquad (2)$$

where ϕ and c are acoustic velocity potential and sound velocity, respectively. Sound pressure p is given by $p = \rho_{\mathrm{a}}(\partial\phi/\partial t)$, where ρ_{a} is the density of air.

3. ANALYSIS METHOD

3.1. Fourier Series Expansion

The physical models shown in Fig. 1 have axisymmetric shapes. Therefore vibration displacement and velocity potential can be expanded in the circumferential direction as follows

$$\zeta(r,\theta,t) = \sum_{n=-\infty}^{\infty}\zeta_n(r,t)e^{jn\theta}, \qquad (3)$$

$$\phi(r,\theta,z,t) = \sum_{n=-\infty}^{\infty}\phi_n(r,z,t)e^{jn\theta}. \qquad (4)$$

In practice, the above infinite series has to be truncated at a finite order. Substituting the displacement and velocity potential into the differential equations in the previous section and using the orthogonality of trigonometric functions, we obtain decoupled differential equations with respect to only r, z and t; the equations do not include differentiation with respect to θ.

3.2. Spectral Method

The differential equations obtained in the previous section are analyzed numerically by using spectral methods with only discretizing the rz-plane. Spectral methods are some versions of weighted residual methods and classified according to the type of interpolation function and weighting function. In this study, spectral nodal Galerkin methods[6], which use Lagrange basis functions spreading over the whole analyzed domain as both interpolation functions and weighting functions, are adopted. Although spectral methods can offer greater accuracy than FDMs or FEMs, they have drawbacks concerned with applicable shapes and boundary conditions. Some special treatments are introduced to overcome the drawbacks.

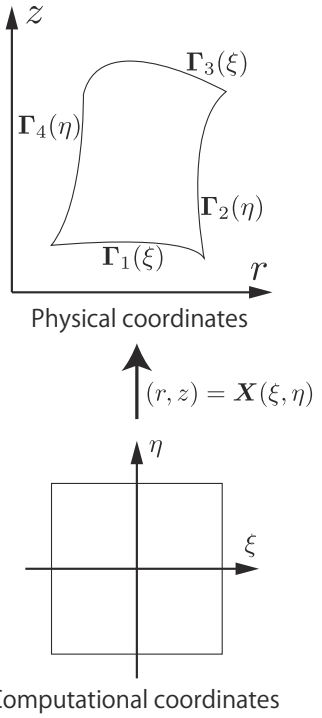


Figure 2. A map between the physical space coordinates (r, z) and computational space coordinates (ξ, η).

3.2.1. Plate vibration field

In a typical spectral method with Lagrange type of differentiation matrices, only one equation per one unknown is obtained. However, for the plate vibration field, two boundary conditions have to be imposed at a boundary node because the vibration equation of the stretched plate is a fourth-order differential equation.

Use of Hermite interpolation type of differentiation matrices can overcome the problem[5]. The differentiation matrices are derived by using Hermite interpolation that interpolates with nodal derivatives in addition to nodal values instead of Lagrange interpolation. This treatment enables two boundary conditions to be imposed at the boundary node.

Discretizing the differential equation of the plate vibration field according to spectral nodal Galerkin formulation, we obtain the same form of matrix equation as one in an FEM,

$$\boldsymbol{M}^{\mathrm{p}}\ddot{\boldsymbol{\zeta}}_n + \boldsymbol{C}_n^{\mathrm{p}}\dot{\boldsymbol{\zeta}}_n + \boldsymbol{K}_n^{\mathrm{p}}\boldsymbol{\zeta}_n = \boldsymbol{f}_n + \boldsymbol{p}_n, \qquad (5)$$

where $\boldsymbol{M}^{\mathrm{p}}$, $\boldsymbol{C}_n^{\mathrm{p}}$, $\boldsymbol{K}_n^{\mathrm{p}}$, $\boldsymbol{\zeta}_n$, $\boldsymbol{f}_n$ and $\boldsymbol{p}_n$ are the mass matrix, damping matrix, stiffness matrix, vibration displacement vector, mechanical excitation force vector and sound pressure vector, respectively. The subscript n represents an order of the Fourier series. The entries of the damping and stiffness matrices depend on the order n while those of the mass matrix are independent of it.

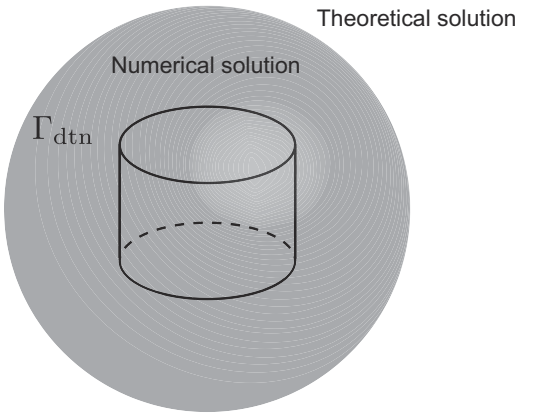


Figure 3. A spherical artificial boundary around an obstacle for a DtN method.

3.2.2. Sound field

Spectral methods can ordinarily analyze only regular-shaped domains due to their global interpolation functions. However, the shapes of the sound fields inside and outside a membranophone cavity are not always regular. Thus, we introduce generalized curvilinear coordinates (ξ, η) in order to analyze such an irregular-shaped domain.

Transform of Eq. 2 with a map $\boldsymbol{X}(\xi, \eta)$ between the physical coordinates (r, z) and computational coordinates (ξ, η) shown in Fig. 2 yields the wave equation in the generalized curvilinear coordinates. Discretizing the wave equation as in the case of the plate vibration field, we obtain

$$\boldsymbol{M}^{\mathrm{a}}\ddot{\boldsymbol{\phi}}_n + \boldsymbol{K}_n^{\mathrm{a}}\boldsymbol{\phi}_n = \boldsymbol{v}_n, \tag{6}$$

where $\boldsymbol{M}^{\mathrm{a}}$, $\boldsymbol{K}_n^{\mathrm{a}}$, $\boldsymbol{\phi}_n$ and $\boldsymbol{v}_n$ are the mass matrix, stiffness matrix, velocity potential vector and acoustical particle velocity vector, respectively.

3.2.3. Boundary and coupling conditions

We discuss boundary and coupling conditions here. First, the shell or kettle of a membranophone is considered acoustically rigid. Then acoustical particle velocity has to be zero on it.

Next, although the sound field outside a membranophone cavity is an infinite domain, it has to be truncated at a finite domain and a proper boundary condition has to be imposed at the artificial boundary when it is analyzed with spectral methods. In this study, a Dirichlet-to-Neumann (DtN) method[7] is adopted. It is the method that introduces a spherical artificial boundary Γ_{dtn} around an obstacle and couples a numerical solution with a theoretical solution at the boundary (Fig. 3). Since it expresses an exact non-reflecting boundary condition, one can make a domain which is numerically analyzed smaller than other approximated non-reflecting boundary conditions.

Finally, the plate vibration field and sound fields inside and outside the cavity are coupled through coupling conditions of pressure and velocity. The sound pressure vector to the plate $\boldsymbol{p}_n$ is given by the difference between the interior and exterior sound pressure $\rho_{\mathrm{a}}\dot{\boldsymbol{\phi}}_n$. The acoustical particle velocity vectors to the sound fields $\boldsymbol{v}_n$ are given by the vibration velocity vector $\dot{\boldsymbol{\zeta}}_n$.The coupling conditions yields a coupled matrix equation. Solving the equation and substituting ζ_n and ϕ_n into Eqs. (3) and (4), we obtain vibration displacement and acoustic velocity potential at arbitrary points.

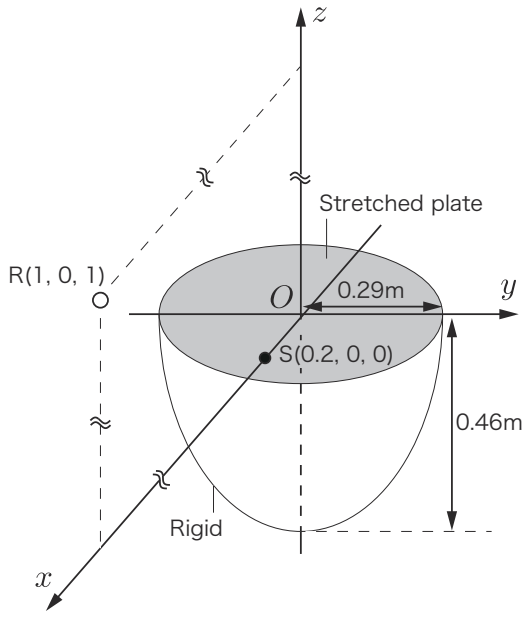


Figure 4. The size of the measured and calculated timpani drum and positions of excitation point and receiving point.

3.2.4. Analysis of temporal differentiation

Equations (5) and (6) include temporal differentiation and they can be solved by using time integration schemes. However, they need verification concerned with the stability. Therefore, at a stage prior to the analysis in the time domain, we let the time factor be $e^{j\omega t}$ and solve the equations in the frequency domain in this study.

4. NUMERICAL RESULTS

In order to validate the proposed method, frequency response functions of a timpani drum were measured and calculated.

4.1. Measurement and Calculation Condition

Figure 4 shows the size of a measured and calculated timpani drum and the positions of a excitation point S and receiving point R. The parameter values were the following: $T = 5800\ \mathrm{N/m}$, $B = 1.83 \times 10^{-4}\ \mathrm{Pa \cdot m^4}$, $\rho_{\mathrm{p}} = 0.25\ \mathrm{kg/m^2}$, $\beta = 1.0 \times 10^{-6}\ \mathrm{s}$, $c = 344\ \mathrm{m/s}$ and $\rho_{\mathrm{a}} = 1.21\ \mathrm{kg/m^3}$.

In measurement, the timpani head was excited impulsively by using a drum stick and a frequency response function of radiated sound was measured with an FFT analyzer.

In calculation, the artificial boundary to which a DtN map was applied was set on a sphere with a radius of 0.5 m and

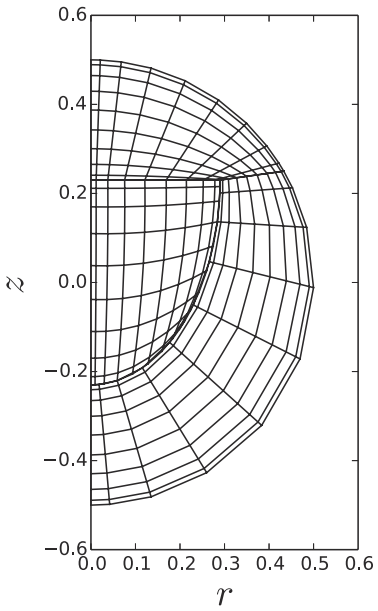


Figure 5. The nodes of the calculated timpani drum.

the rz-plane was discretized as shown in Fig. 5. The degree of freedom of the coupled matrix equation to be solved was 310 and the Fourier series in the circumferential direction was truncated at the 10th order.

4.2. Comparison Between Measurement and Calculation

Calculated and measured frequency response functions are shown in Fig. 6. The blue line and red one represent the calculated frequency response function and measured one, respectively.

The level differences of the first and fourth peaks, which correspond to (0, 1) and (0, 2) mode, are seen. It is possible that the differences are attributed to disagreement of parameter values or excitation force. However, the other peaks show near levels and both the results show good agreement in terms of eigenfrequencies. Therefore, the result suggests that the proposed method be valid as an analysis method of a membranophone.

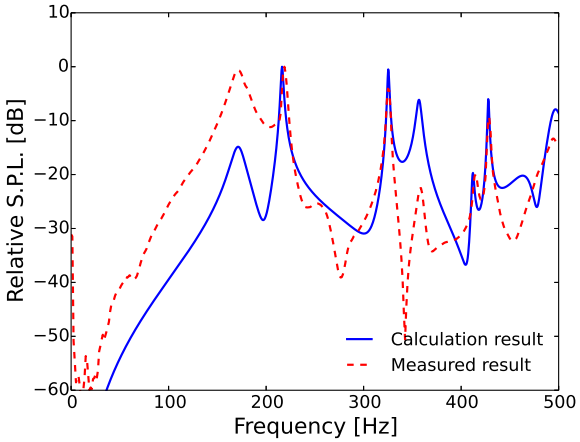


Figure 6. Comparison of calculation and measured results.

5. CONCLUSIONS

A high-efficiency analysis method of membranophones was proposed. It was shown that a frequency response function of a timpani drum calculated with the proposed method showed good agreement with measured ones. The proposed method is based on Fourier series expansions and spectral methods as well as proposed methods in our previous studies[4, 5]. From results in the studies, it is probable that the proposed method in this study could also significantly reduce the degree of freedom necessary for computation and analyze membranophones more efficiently than FEMs.

Although matrix equations were solved in the frequency domain in this study, time integration schemes such as a central difference method or Newmark-β method also could be applied. In particular, since mass matrices in spectral methods are diagonal matrices, explicit schemes, which have a problem of stability, would be effective and could reduce computational cost. Analysis in the time domain is our future work.

REFERENCES

[1] S. Bilbao, *Numerical sound synthesis: finite difference schemes and simulation in musical acoustics*. John Wiley & Sons, 2009.

[2] S. Bilbao and C. J. Webb, "Physical modeling of timpani drums in 3d on gpgpus," *Journal of the Audio Engineering Society*, vol. 61, no. 10, pp. 737–748, 2013.

[3] F. Pfeifle and R. Bader, "Real-time finite-difference method physical modeling of musical instruments using field-programmable gate array hardware," *Journal of the Audio Engineering Society*, vol. 63, no. 12, pp. 1001–1016, 2016.

[4] Y. Araki and T. Samejima, "Spectral method for vibration analysis of circular stretched plates," *The Journal of the Acoustical Society of Japan*, vol. 72, no. 4, pp. 173–181, 2016 (in Japanese).

[5] Y. Araki and T. Samejima, "Fourier series expansion type of spectral collocation method for vibration analysis of cylindrical shells," *Acoustical Science and Technology*, vol. 37, no. 5, pp. 211–220, 2016.

[6] D. Kopriva, *Implementing spectral methods for partial differential equations: Algorithms for scientists and engineers*. Springer Science & Business Media, 2009.

[7] D. Givoli, *Numerical methods for problems in infinite domains*. Elsevier, 1992, vol. 33.

A Review of Bell Design and Tuning Techniques From Magic to Science

Miguel Carvalho,[1†] *Vincent Debut,*[1] *Jose Antunes*[2]

[1]INET-md, Universidade Nova de Lisboa
[2]C2TN, Instituto Superior Tecnico
[†]miguel.carvalho@ctn.tecnico.ulisboa.pt

ABSTRACT

The sound of a bell presents numerous partials resulting from its vibrational modes. Since the 17th century, it is commonly accepted among bellfounders, that to sound pure and clear, the first five partials (at least) of a bell must fall into specific musical frequency ratios, established at 0.5:1:1.2:1.5:2. Beyond this internal tuning, for carillons, bellfounders must also correctly tune their bells in relation to each other, in order to form a musical scale and provide a musical instrument able to play melodies and chords. Other peculiarity of bells sounds stems from the bell near-axial symmetry. Actually, during casting, the presence of slight defects in symmetry is unavoidable, which breaks the degeneracy of the normal modes and leads to audible beats in the radiated sound, a phenomenon which is usually referred to as warble by campanologists. On top of all this, the perceived pitch of a bell is a delicate phenomenon, related to psychoacoustic effects, which may correspond to a modal frequency or arise as a virtual pitch. Even today, the nature of the so-called bell strike-note is still a topic of debate.

One therefore easily understands the complexity of designing and tuning carillon bells. The evolution carillons face until they become musical instruments results from more than ten centuries of gradual transformation and development. It is based on the founders experience, acquired by trial and error, and has benefited from the increasing scientific interest for the phenomenon of bell partials. The secrecy under which such knowledge was kept is also another reason why there has been only a few bell founders capable of successfully casting and tuning carillons throughout the history. Nowadays, following the advances in structural dynamics modelling and computational techniques, bells can be designed on a scientific basis, the casting process can be fully monitored, and tuning can be accurately achieved by coupling physical modelling and optimization algorithms.

Based on an extensive bibliographical research, we provide in this paper an overview of bell design and tuning techniques, since its early stages of empiricism to the most recent scientific advances. Our work sums up and complements the scarce existing literature on bell tuning, and presents innovative strategies for assessing and improving the tuning of bells, using 3D-scan imaging technology, physical modelling and computational methods.

Optical Interferometry for Transient Deformation Analysis

Tatiana Statsenko†*, Vasileios Chatziioannou, Wilfried Kausel*

Department of Music Acoustics, University of Music and Performing Arts Vienna
†statsenko@mdw.ac.at

ABSTRACT

Electronic speckle pattern interferometry is performed for temporal analysis of the surface deformation of vibrating structures. This requires specially designed image processing methods in order to reconstruct the dynamic deformation of an object. The reconstruction procedure which is based on a numerical optimization algorithm is applied on a set of interferometric images recorded by a high speed camera. While designing a phase-recovery algorithm one should also consider the ill-posedness of the problem due to the ambiguity in the recorded interferometric patterns, speckle noise, and electronic noise of the camera. The presented method is applied to reconstruct temporal surface deformation of a violin body under real playing conditions. The results show the feasibility of the algorithm and the correspondence of the calculated intensity images with the experimental data.

1. INTRODUCTION

Nondestructive optical methods for inspection and analysis of surface deformations and radiated sound fields become prevalent in the musical acoustics field [1]. Various measurements of musical instruments can be performed optically, such as string motion, modal analysis, acoustic wave propagation, etc [2, 3, 4]. Most common measurement techniques involve laser Doppler vibrometry, Schlieren imaging, particle image velocimetry, holographic interferometry and electronic speckle pattern interferometry (ESPI) [1]. Each method is appropriate for performing a certain task. In case of surface deformation analysis speckle interferometry has several advantages such as full-field spatial recovery and high temporal resolution.

Temporal ESPI (TRESPI) is a general concept for speckle interferometry methods applied to record and reconstruct non-harmonic surface deformation. Initially, pulsed lasers were used for recording of transient deformations [5, 6, 7]. However, pulsed ESPI is only applied for repeatable events and thus cannot succeed in case of single transient deformation measurement because conventional cameras are not fast enough. In this case, high-speed cameras are utilized, and pulsed laser is substituted with a continuous wave laser [8, 9]. The quality of the recorded data depends directly on the camera resolution and minimum attainable exposure time.

The aim of this work is to recover out-of-plane deformation of the violin body during a single string bowing. The expected output is a dynamic deforming structure representing the relative motion of the surface in time. Section 2 describes the methodology of the measurement procedure for the TRESPI arrangement, relevant formulas and image processing strategy. Section 3 presents the results of the deformation recovery. Section 4 discusses the issues and possible extensions of the method.

2. METHODOLOGY

Using a high-speed camera allows capturing fast deformations in the microsecond range. Small exposure time reduces the impact of the environmental noise. However, electronic noise may be an obstacle to obtain correct values of the deformation since it reinforces the intensity graininess in addition to the random speckle noise caused by the summation of the optical waves reflected by the rough surface of the object [10].

Implementation of mathematical optimization into speckle interferometry methods has shown success for the case of harmonic vibrating surfaces [11]. Instead of solving directly the inverse problem, the optimization algorithm iteratively updates vibration amplitude values to minimize the mismatch between filtered experimental data and calculated intensity. Optimization algorithm utilizes the continuousness of the surface and thus reduces the errors that may appear during pixel-by-pixel direct calculations. Applying boundary conditions and adding constraints on the possible surface profile improves the resulting quality.

2.1. Experimental setup

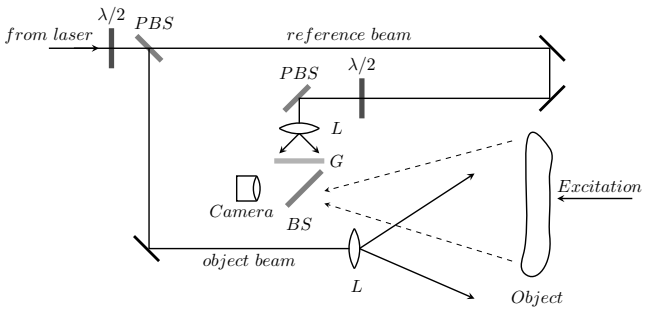


Figure 1: Experimental arrangement for a TRESPI measurement. The relative intensity of the beams and the polarization are controlled by half-wave plates ($\lambda/2$) and polarizing beamsplitters (PBS). Microscopic objective lenses (L) expand the beams. Ground glass (G) scatters the reference beam. A high-speed camera records the interference between the beam reflected from the object and the reference beam, combined by a beamsplitter (BS).

The TRESPI experimental arrangement is shown in Figure 1. The illumination is provided by a continuous wave laser (Verdi V5, Coherent Inc.) with a wavelength λ of 532 nm. The high-speed camera (Phantom v12.1, Vision Research) records interferometric frames at 21000 frames per second at a resolution of 512 by 512 pixels. The bit depth of the camera is 12 bit. Since the exposure time of the camera is extremely small (47 µs), the irradiance should be sufficient for the pixels to get enough light, so the output power of the laser is about 2 W. The violin was fixed on the optical table and excited by a bow manually. The image capturing was done using camera PCC software (Vision Research) and images were processed in MATLAB (MATLAB and Image Processing Toolbox R2014a, The MathWorks, Inc., Natick, Massachusetts, USA). The ESPI arrangements are generally sensitive to a single direction of deformation. In the current setup, the out-of-plane deformation is measured, where the plane of the object is parallel to the plane of the camera sensor.

Subtraction of the frames taken at different times produces the correlation frames which reflect the deformation profile in the form of bright and dark fringes. These fringes may be seen as lines of equal displacement of $\lambda/2$. Thereby, the qualitative assessment of the deformation that happened between two time moments during bowing is performed by the subtraction of the pair of corresponding frames. Figure 2 shows such a subtraction intensity of the violin deformation. The shape of the violin is complicated for automated analysis because the processing algorithm cannot distinguish parts such as bridge and strings. Therefore, only part of the violin is considered (256 by 256 pixels). The square region of interest that covers the quarter of the image is highlighted in Figure 2. The subtraction intensity frame shows that the tailpiece is moving differently from the violin body, which creates discontinuity of the interference pattern.

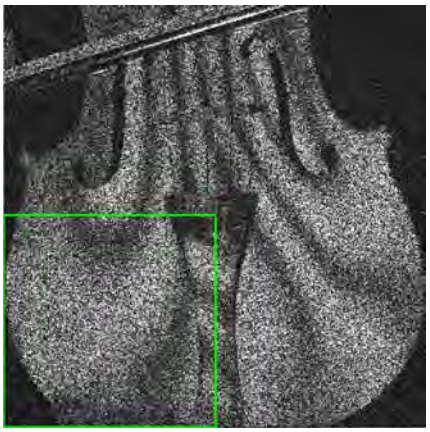


Figure 2: Qualitative ESPI image of the violin. Subtraction intensity of the interference patterns corresponding to a time difference of 4.8 ms (100 frames). The green box features the region of interest.

2.2. Theory

The intensity value recorded from the interference of two coherent beams for every pixel of the camera at time t is described as:

$$I(t) = A + B \cdot \cos(\phi(t)), \tag{1}$$

where A is the sum of the intensities of both beams, B is the square root of their multiplication, referred to as contrast, ϕ is a global phase. The global phase is constituted of the random speckle phase and the phase offset between the two beams that changes due to the surface motion.

The coefficients A and B can be considered constant during the short image acquisition time and detected for every pixel as:

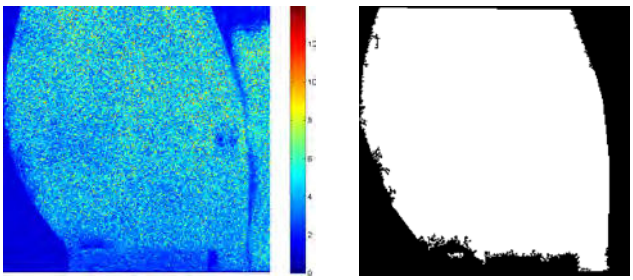
$$A = \left(\frac{\max(I(t)) + \min(I(t))}{2} \right), \tag{2}$$

$$B = \left(\frac{\max(I(t)) - \min(I(t))}{2} \right). \tag{3}$$

The coefficient B is the pixel modulation that shows the range of the intensity difference. If the value of B is smaller than the temporal noise, the pixel is regarded invalid since the intensity deviation is lower than noise, so it has to be removed from further processing. Therefore, a measure of pixel quality Q is introduced:

$$Q = B/N, \tag{4}$$

where N is the noise value which is calculated as the standard deviation of the temporal intensity spatially averaged over the background region, taken from a corner of the image. Pixel quality matrix is shown in Figure 3a. Based on the Q values it is possible to automatically extract background pixels, non-moving parts and badly illuminated regions. The tailpiece part was defined and subtracted manually. The output mask of relevant pixels is shown in Figure 3b, calculated with a signal-to-noise ratio of 3 as a threshold for pixel validity.



(a) Pixel quality matrix (b) Mask

Figure 3: (a) Matrix coefficient Q and (b) mask applied to preserve valid pixels during processing.

After the intensity normalization (during which coefficients A and B are removed), the intensity at a discrete time moment t_n can be written as:

$$I(t_n) = \cos(\phi(t_n)) = \cos(\phi(t_{n-1}) + \Delta\phi(t_{n-1}, t_n)), \tag{5}$$

where $\Delta\phi(t_{n-1}, t_n)$ is the phase difference that occurs during the motion of the surface between time moments t_{n-1} and t_n. The relative deformation $\Delta D(t_{n-1}, t_n)$ linearly depends on $\Delta\phi(t_{n-1}, t_n)$ and λ:

$$\Delta D(t_{n-1}, t_n) = \frac{\lambda \Delta\phi(t_{n-1}, t_n)}{4\pi}. \quad (6)$$

The direct solution of Eq. 5 by applying an arccos function may suffer from the periodicity of the cosine function and sign ambiguity. General phase extracting methods produce a phase map with so-called "wrapped phase" that is further processed on a pixel basis in order to "unwrap" it. The algorithm based on the Rosenbrock optimization is applied to outperform pixel phase unwrapping methods. The method minimizes the target objective function while adjusting the surface displacement which is generated by bilinear interpolation on the grid points. The density of the grid points is gradually adapted to the scale of the deformation during the optimization.

The target error function F is the pixel sum of the mismatch between the normalized intensity and the intensity calculated from the predicted deformation $\Delta D(t_{n-1}, t_n)$ for every pixel with coordinates x, y:

$$F = \sum_{x,y} \left(I(t_n) - cos\left(\phi(t_{n-1}) + \frac{4\pi \Delta D(t_{n-1}, t_n)}{\lambda} \right) \right)^2. \quad (7)$$

First, the starting frame is assigned and the initial phase is calculated. Initially, the deformation is assumed positive, so that the initial phase $\phi^+(t_1)$ is calculated as:

$$\phi^+(t_1) = \arccos(I(t_1)) \cdot sign(I(t_2) - I(t_1)). \quad (8)$$

Then, the relative deformation is calculated with the respect to the first frame and $\phi^+(t_1)$. The initial guess for the deformation is a flat surface with 4 operating grid points at the edges of the image. During the optimization the surface deforms and the number of grid points increases, until the target error function is minimized or maximum value of iteration steps is attained. The optimization results in the absolute value of the deformation $|\Delta\phi(t_1, t_2)|$. The sign can be corrected through the detection of the nodal lines and differentiation between the regions that move in-phase and out-of-phase to provide the mechanical stability of the system, as proposed in [11]. After the sign correction procedure, the real phase value of $\phi(t_1)$ is obtained through pixel-by-pixel multiplication of $\phi^+(t_1)$ by the sign of the resulting phase difference:

$$\phi(t_1) = \phi^+(t_1) \cdot \text{sign}(\Delta\phi(t_1, t_2)). \quad (9)$$

This three-step algorithm that involves calculation of ϕ^+, optimization of $\Delta\phi$ and the sign correction for $\Delta\phi$ can be applied to find a phase difference between any pair of frames when the deformation does not exceed $\lambda/2$ which corresponds to 2π phase difference. However, direct summation of the obtained phase differences to obtain temporal deformation of the frame series results in a high cumulative error because of the noisy random speckle phase and the camera noise. To reduce the error, for every next recovery step $\Delta D(t_n, t_{n+1})$ the initial phase $\phi(t_n)$ is presented as the sum of the phase and the recovered deformation calculated in the previous step:

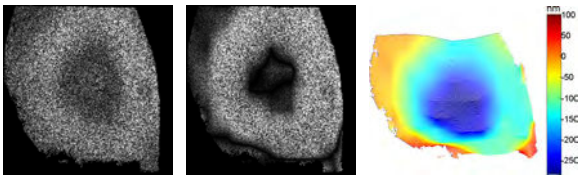
$$\phi(t_n) = \phi(t_{n-1}) + \Delta D(t_{n-1}, t_n). \quad (10)$$

The total deformation D between the first and the n-th frame is:

$$D(t_n) = \Delta D(t_1, t_2) + \Delta D(t_2, t_3) + ... + \Delta D(t_{n-1}, t_n). \quad (11)$$

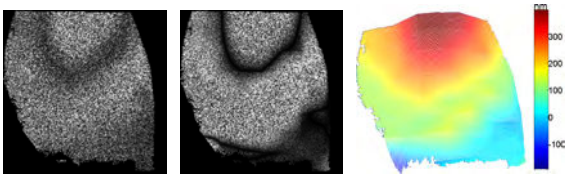
3. RESULTS

During the experiment, the third string of the violin was bowed. The deformation recovery results, for various time moments for the chosen region of interest, are shown in Figure 4. Calculated intensity profiles are in good correspondence with the experimentally obtained subtraction frames. However, the effects of the cumulative error are present at the edges of the mask. This happens because the fixed boundary condition has not been applied due to the complicated shape of the mask and free boundary conditions on the top edge of the region of interest should remain. The results show the feasibility of the method in case of positive and negative deformation values, as well as in case of deformation differences higher than $\lambda/2$.



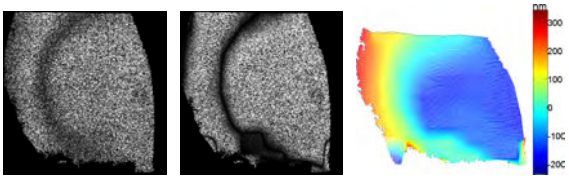
Experiment Calculation Deformation

(a) 0.52 ms



Experiment Calculation Deformation

(b) 1.1 ms



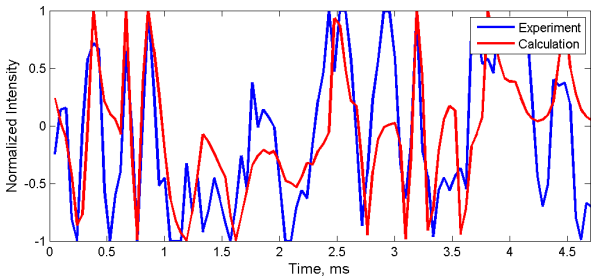
Experiment Calculation Deformation

(c) 2.9 ms

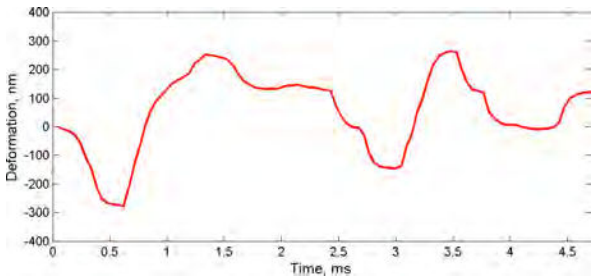
Figure 4: Results of the relative deformation reconstruction for different time intervals.

Figure 5a shows the experimentally obtained normalized intensity profile and the calculated intensity profile for a single pixel taken from the central part of the region of interest. Both contours are heavily affected by noise. However, the calculated

intensity follows the experimental data to a certain extend. This comparison proves the continuity of the deformation and robustness of the method for time range of 10 ms. Figure 5b shows a single pixel deformation profile taken from the central part of the region of interest.



(a) Temporal intensity



(b) Temporal deformation

Figure 5: (a) Normalized intensity and calculated intensity of a single pixel taken from the central region of the image and (b) recovered temporal deformation.

4. CONCLUSIONS

Transient deformations of the violin body during bowing have been investigated by means of high-speed electronic speckle pattern interferometry. The presented image processing algorithm produces a dynamic out-of-plane deformation profile with high spatial and temporal resolution. This optical method can be applied in musical acoustics for transient deformation analysis of surfaces and is valuable for studying non-linear vibrations of musical instruments.

Future work will involve stability analysis, application of boundary conditions for complicated object shapes and reduction of the phase cumulative error.

5. ACKNOWLEDGMENTS

The authors would like to thank Alexander Mayer for technical assistance. This research work has been funded by the European Commission within the ITN Marie Curie Action project BATWOMAN under the 7th Framework Programme (EC grant agreement no. 605601).

REFERENCES

[1] N. E. Molin and L. Zipser, "Optical methods of today for visualizing sound fields in musical acoustics," *Acta Acustica united with Acustica*, vol. 90, no. 4, pp. 618–628, 2004.

[2] T. R. Moore and S. A. Zietlow, "Interferometric studies of a piano soundboard," *The Journal of the Acoustical Society of America*, vol. 119, no. 3, pp. 1783–1793, 2006.

[3] G. Bissinger and D. Oliver, "3-D laser vibrometry on legendary old Italian violins," *Sound And Vibration*, vol. 41, no. 7, pp. 10–14, 2007.

[4] D. Russell, D. E. Parker, and R. S. Hughes, "Analysis of standing sound waves using holographic interferometry," *American Journal of Physics*, vol. 77, no. 8, pp. 678–682, 2009.

[5] A. Fernández, A. J. Moore, C. Pérez-López, A. F. Doval, and J. Blanco-García, "Study of transient deformations with pulsed TV holography: application to crack detection," *Applied Optics*, vol. 36, no. 10, pp. 2058–2065, 1997.

[6] C. Zang, G. Chen, D. Ewins, and I. A. Sever, "Review of current status of full-field measurement equipment hardware and software, test planning and integration tools, suggestions for future development," Tech. Rep., 2005.

[7] A. R. Ganesan, "Holographic and laser speckle methods in non-destructive testing," in *Proceedings of the National Seminar & Exhibition on Non-Destructive Evaluation*, 2009, pp. 126–130.

[8] A. J. Moore, D. P. Hand, J. S. Barton, and J. D. Jones, "Transient deformation measurement with electronic speckle pattern interferometry and a high-speed camera." *Applied Optics*, vol. 38, no. 7, pp. 1159–1162, 1999.

[9] C. Pérez-López, M. H. De la Torre-Ibarra, and F. Mendoza Santoyo, "Very high speed cw digital holographic interferometry," *Optics Express*, vol. 14, no. 21, pp. 9709–9715, 2006.

[10] M. Servin, J. C. Estrada, J. A. Quiroga, J. F. Mosiño, and M. Cywiak, "Noise in phase shifting interferometry," *Optics Express*, vol. 17, no. 11, pp. 8789–8794, 2009.

[11] T. Statsenko, V. Chatziioannou, T. Moore, and W. Kausel, "Deformation reconstruction by means of surface optimization. Part I: Time-averaged electronic speckle pattern interferometry," *Appl. Opt.*, vol. 56, no. 3, pp. 654–661, 2017.

Finite-Difference Model (FDM) of Mode Shape Changes of the Myanmar Pat Wain Drum Circle with Tuning Paste on Drum Head

Rolf Bader

Institute of Systematic Musicology, University of Hamburg
R_Bader@t-online.de

ABSTRACT

The tuning of the Myanmar pat wain, a unique drum circle consisting of 20 double-headed drums of a hsain wain orchestra is investigated. The drums have clear pitches over three and a half octaves, unusual for a percussion instrument, and are tuned by adding a paste pa sa on the drum head consisting of rice and ashes. It was investigated by the author during a field trip in northern Myanmar during its tuning process. The drum head is modeled using a time-dependent Finite-Difference Method (FDM), solving the differential equation of the two membranes on a 63x63 grid each, with a sample rate of 96kHz, using a Gauss-shaped impulse as driver. The internal air is modeled as a delay line between the two drums. The aim was to estimate pitch and mode changes on the beating membrane due to adding of the paste. To estimate the influence of viscoelasticity on the pitches, the membrane differential equation was enlarged to a viscoelastic model using a complex, frequency dependent Young's modulus transferred into the time domain using a Laplace transform. The calculation was performed on a NVIDIA GTX960 GPU, which is highly parallel, where calculating one second of a drum sound needed about 10 seconds. The sound was calculated as integral of the radiating of the beating membrane points with respect to a place above this membrane. The material properties on the FDM grid were implemented according to the mass distribution of the leather membrane and the paste. It appears that the strong pitch impression of the drum is caused by several factors. The fundamental shows a long decay due to the interaction of the front with the back membrane. The second and third harmonics are enhanced by the inclosed air. Both are tuned near harmonic relations to the fundamental by the distribution of the paste, which changes the harmonic relations, a result previously also found for Indian tabla drums. The higher harmonics are damped due to strong internal damping of the membrane. The viscoelastic FDM for the upper drum head showed strong that the viscoelastic internal damping due to the paste does not change the pitches considerably. The results are compared to the field recordings, which show considerable damping and pitch glides with low-pitched drums. Also the tuning of the drum circle is compared to different tuning systems showing similarities to a tuning pattern of Cambodian roneat deik metalophone.

Toward a Real-Time Waveguide Mesh Implementation

Tamara Smyth,[1†] *Jennifer Hsu,*[2] *Ryan Done*[1]

[1]Department of Music, UC San Diego, La Jolla, California

[†]trsmyth@ucsd.edu

ABSTRACT

This work presents an analytic solution to a 2-D waveguide mesh, made square with input and output at the center to produce a symmetry that reduces redundancy and increases computational savings. The result is a frequency-domain representation of the mesh, a parametric transfer function describing the ratio of output to input, symbolically computed to retain parameters. In this form, the mesh, known for being computationally prohibitive, can be further made to perform in real-time, with (symbolic) parameters allowing for interaction by the user. This work is motivated by a desire to explore real-time interactive/parametric percussion synthesis.

1. INTRODUCTION

It is well known that one-dimensional (1-D) wave propagation can be *efficiently* modeled using a 1-D digital waveguide, a bi-directional delay line that models acoustic propagation delay associated with waves travelling to the right and left along one-dimensional systems such as strings and cylindrical (and conical) bores found in musical instruments. This efficiency is significant for virtual musical instrument design because it allows these systems to be modeled *parametrically* for use in *real time*, allowing the user to change parameter values while hearing the (perceptually) immediate response.

The efficiency of the digital waveguide is lost when brought to higher spatial dimensions. Though 2-D and 3-D waveguides may be used to model wave propagation on membranes, plates, and cavities [1], typical (time-domain) implementations are computationally prohibitive for real-time performance. A 2001 paper [2] reports running a 6×6 2-D mesh in real time on a Pentium III 500 MHz (and 10×10 on a Athlon 1.3 GHz), with more current tests (on an Intel core i5 2.2 GHz) supporting (but struggling at) up to 45×45. Many musically interesting sounds are produced with meshes far above 100×100 and [2] point out that further research is clearly required to optimise the mesh algorithm if it is to be more widely used. As a result, their use in virtual musical instruments has been largely limited to *offline* applications: analysis, non-interactive sound synthesis, or modeling of components not expected to change in real time (e.g. resonators such as soundboards or instrument bodies). The lack of real-time low-latency interaction has made musical exploration of these systems more difficult than their 1-D counterpart. Here, an analytic solution of a simplified (made to have computation-saving symmetry) 2-D waveguide mesh (so called because of the interlaced nature of its time-domain signal flow) is presented, providing an alternate, yet equal, symbolic frequency-domain representation (a transfer function that retains boundary loss parameters), that is intended for parametric percussion synthesis.

2. THE WAVEGUIDE MESH

The 2-D mesh can be viewed as a matrix of junctions with each junction $J_{i,j}$ having 4 (north, south, east, and west) input-output (I/O) ports, as indicated by the '$+$' and '$-$' superscripts, respectively (see Figure 1). In this work, the mesh is constrained to be square ($M \times M$).

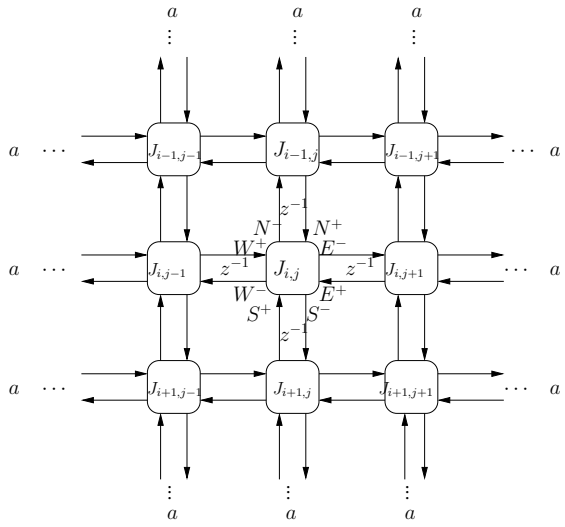


Figure 1. Mesh junctions showing input/outputs on the north N, south S, west W and east E ports and a delay of one sample (z^{-1}) between port inputs and neighbouring port outputs.

The port inputs on junction $J_{m,n}$ may be expressed as a function of neighbouring port outputs with a delay of one sample, with the exception of a mesh boundary where the round-trip from output to input on a port introduces a 2-sample delay and a multiply with boundary scalar reflection loss a:

$$
\begin{aligned}
N^+_{i,j}(n) &= \begin{cases} aN^-_{i,j}(n-2), & \text{if } i = 1 \text{ (north bndry);} \\ S^-_{i-1,j}(n-1) & \text{otherwise.} \end{cases} \\
S^+_{i,j}(n) &= \begin{cases} aS^-_{i,j}(n-2), & \text{if } i = N \text{ (south bndry);} \\ N^-_{i+1,j}(n-1) & \text{otherwise.} \end{cases} \\
E^+_{i,j}(n) &= \begin{cases} aE^-_{i,j}(n-2), & \text{if } j = M \text{ (east bndry);} \\ W^-_{i,j+1}(n-1), & \text{otherwise.} \end{cases} \\
W^+_{i,j}(n) &= \begin{cases} aW^-_{i,j}(n-2), & \text{for } j = 1 \text{ (west bndry);} \\ E^-_{i,j-1}(n-1), & \text{otherwise.} \end{cases}
\end{aligned} \tag{1}
$$

The output on each port at time sample n is given by the junction's total velocity (the wave variable used here),

$$vJ_{i,j}(n) = \frac{N^+_{i,j}(n) + S^+_{i,j}(n) + E^+_{i,j}(n) + W^+_{i,j}(n)}{2}, \quad (2)$$

minus the port input which, when applying (2), yields:

$$\begin{aligned}
N^-_{i,j}(n) &= vJ_{i,j}(n) - N^+_{i,j}(n) \\
&= \frac{-N^+_{i,j}(n) + S^+_{i,j}(n) + E^+_{i,j}(n) + W^+_{i,j}(n)}{2} \\
S^-_{i,j}(n) &= vJ_{i,j}(n) - S^+_{i,j}(n) \\
&= \frac{N^+_{i,j}(n) - S^+_{i,j}(n) + E^+_{i,j}(n) + W^+_{i,j}(n)}{2} \\
E^-_{i,j}(n) &= vJ_{i,j}(n) - E^+_{i,j}(n) \\
&= \frac{N^+_{i,j}(n) + S^+_{i,j}(n) - E^+_{i,j}(n) + W^+_{i,j}(n)}{2} \\
W^-_{i,j}(n) &= vJ_{i,j}(n) - W^+_{i,j}(n) \\
&= \frac{N^+_{i,j}(n) + S^+_{i,j}(n) + E^+_{i,j}(n) - W^+_{i,j}(n)}{2}.
\end{aligned} \quad (3)$$

Substituting (1) into (3) yields final expressions (given here in matrix form) for port outputs (column vector[1] $\mathbf{J}_{i,j}$) as a function of neighbouring port outputs (column vector $\mathbf{b}_{i,j}$), with the now dispensable '$-$' superscript henceforth omitted:

$$\mathbf{J}_{\mathbf{n}i,j} = \mathbf{A}\mathbf{b}_{\mathbf{n}i,j} \quad (4)$$

where

$$\mathbf{J}_{\mathbf{n}i,j} = (N_{i,j}(n), S_{i,j}(n), E_{i,j}(n), W_{i,j}(n)), \quad (5)$$

$$\mathbf{A} = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \quad (6)$$

and where the neighbouring port dependencies is given by

$$\mathbf{b}_{\mathbf{n}i,j} = \begin{bmatrix} S_{i-1,j}(n-1) \\ N_{i+1,j}(n-1) \\ W_{i,j+1}(n-1) \\ E_{i,j-1}(n-1)) \end{bmatrix} \quad (7)$$

with exceptions as per boundary conditions in (1):

$$\begin{aligned}
\mathbf{b}_{\mathbf{n}1,j}(1) &= aN_{1,j}(n-2) \quad \text{(north boundary)} \\
\mathbf{b}_{\mathbf{n}M,j}(2) &= aS_{M,j}(n-2) \quad \text{(south boundary)} \\
\mathbf{b}_{\mathbf{n}i,M}(3) &= aE_{i,M}(n-2) \quad \text{(east boundary)} \\
\mathbf{b}_{\mathbf{n}i,1}(4) &= aW_{i,1}(n-2) \quad \text{(west boundary).}
\end{aligned} \quad (8)$$

2.1. Mesh Input and Output

The mesh may be "excited" at a location corresponding to junction $J_{i,j}$ using an input signal $x(n)$ symmetrically distributed about the junction's port inputs (or equivalently, the neighbouring junction's port outputs), such that the output signal $y(n)$ is the junction's total velocity:

$$\begin{aligned}
y(n) &= vJ_{i,j}(n) \\
&= \frac{1}{2}\sum_{u=1}^{4} \mathbf{b}_{\mathbf{n}i,j}(u) + x(n) \\
&= \frac{1}{2}\sum_{u=1}^{4} \left(\mathbf{b}_{\mathbf{n}}(u) + \frac{x(n)}{2}\right),
\end{aligned} \quad (9)$$

with boundaries according to (8). Equation (9) can be further reduced if the I/O location is at the center,

$$\begin{aligned}
y(n) &= vJ_{\frac{M+1}{2},\frac{M+1}{2}}(n) \quad \text{for odd } M, \\
&= \frac{1}{2}\left(4S_{\frac{M-1}{2},\frac{M+1}{2}}(n-1) + 4\frac{x(n)}{2}\right) + ... \\
&= 2S_{\frac{M-1}{2},\frac{M+1}{2}}(n-1) + x(n),
\end{aligned} \quad (10)$$

where, due to symmetry, the input on all ports is equal.

3. REDUCTION BY MESH SYMMETRY

Here, the input/output location is chosen to be the center of a square $M \times M$ mesh, for odd M, creating a symmetry and redundancy that reduces computation (not taken advantage of in a typical mesh implementation which performs the same computation regardless of the I/O location). Of course, this same symmetry that reduces computation also results in cancelled modal resonances—the most reduced spectrum for a given dimension.

When the mesh is excited at the center junction $J_{\frac{M+1}{2},\frac{M+1}{2}}$, this junction's port outputs are equal in all directions,

$$N_{\frac{M+1}{2},\frac{M+1}{2}} = S_{\frac{M+1}{2},\frac{M+1}{2}} = E_{\frac{M+1}{2},\frac{M+1}{2}} = W_{\frac{M+1}{2},\frac{M+1}{2}}, \quad (11)$$

(with n being suppressed for compactness), creating a symmetry that is propagated outward to the mesh edges, and a redundancy requiring computation of only half a quadrant of the entire mesh (as indicated by shading in (12)),

$$\begin{bmatrix}
J_{1,1} & J_{1,2} & \dots & J_{1,\frac{M+1}{2}} & \dots & J_{1,M} \\
J_{2,1} & J_{2,2} & \dots & J_{2,\frac{M+1}{2}} & \dots & J_{2,M} \\
J_{3,1} & J_{3,2} & \dots & J_{3,\frac{M+1}{2}} & \dots & J_{3,M} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
J_{\frac{M+1}{2},1} & J_{\frac{M+1}{2},2} & \dots & J_{\frac{M+1}{2},\frac{M+1}{2}} & \dots & J_{\frac{M+1}{2},M} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
J_{M,1} & J_{M,2} & \dots & J_{M,\frac{M+1}{2}} & \dots & J_{M,M}
\end{bmatrix} \quad (12)$$

a reduction from M^2 junctions to

$$Nj = \frac{\frac{M+1}{2}\left(\frac{M+1}{2}+1\right)}{2}, \quad \text{for odd } M. \quad (13)$$

[1] Notation whereby column vectors are notated with parentheses and commas is adopted when compactness is necessary

3.1. New Boundary Conditions

Reducing the mesh to half a quadrant introduces new boundaries (points of symmetry) along the mesh's first half of the

- **center column (col)** ($J_{1,\frac{M+1}{2}}, J_{2,\frac{M+1}{2}}, \ldots, J_{\frac{M+1}{2},\frac{M+1}{2}}$):

$$W_{i,j+1}(n-1) = E_{i,j-1}(n-1), \text{ for } j = \frac{M+1}{2}. \tag{14}$$

- **quadrant diagonal (diag)** ($J_{1,1}, J_{2,2}, \ldots, J_{\frac{M+1}{2},\frac{M+1}{2}}$):

$$\left.\begin{aligned} N_{i+1,j}(n-1) &= W_{i,j+1}(n-1) \\ E_{i,j-1}(n-1) &= S_{i-1,j}(n-1) \end{aligned}\right\} \quad \text{for } i = j. \tag{15}$$

Employing these symmetries (14-15) alone in a direct implementation of (4-8), (10)), and can significantly reduce mesh computation times (see Table 1, column 3 vs. column 2).

4. OBTAINING THE TRANSFER FUNCTION

Taking the z-transform of (4) and rearranging elements (for convenience in algorithm) yields

$$\mathbf{J}_{\mathbf{z}i,j} = \mathbf{H}_{\mathbf{z}i,j}\mathbf{b}'_{\mathbf{z}i,j}, \tag{16}$$

where for non boundaries

$$\mathbf{J}_{\mathbf{n}i,j} = (W_{i,j}(z)N_{i,j}(z), E_{i,j}(z), S_{i,j}(z)), \tag{17}$$

$$\mathbf{b}_{\mathbf{z}i,j} = (E_{i,j-1}(z), S_{i-1,j}(z), W_{i,j+1}(z), N_{i+1,j}(z)), \tag{18}$$

and

$$\mathbf{H}_{\mathbf{z}i,j} = \frac{z^{-1}}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \tag{19}$$

and where boundary conditions are later applied following (8) (see Section 4.1).

At this point, the mesh is represented as a system of equations for each junction (expressed as a function of neighbouring junctions) which, when excited and tapped in the center, yields an output as a function of the southern port output of the center junction's northern neighbour and the input (9), the z-transform of which is given by

$$Y(z) = 2S_{\frac{M-1}{2},\frac{M+1}{2}}(z)z^{-1} + X(z). \tag{20}$$

What is needed to produce a transfer function however, is an output solely as a function of the input,

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + \cdots + b_{N_b} z^{N_b}}{1 + a_1 z^{-1} + \cdots + a_{N_a} z^{N_a}}, \tag{21}$$

or alternatively (and perhaps preferably) in factored form. Thus, beginning with (16) and an initialization step, dependencies of junctions on their neighbors is sequentially eliminated, starting at the top row, and then working down the (triangular) mesh.

4.1. Initialization

Elements in (16) can be initialized to values that, in the case of boundaries, lead to reduced sizes for $\mathbf{J}_\mathbf{z}$, $\mathbf{b}_\mathbf{z}$ and $\mathbf{H}_\mathbf{z}$. Applying boundaries (8, 14, 15), yields initial conditions

$$\mathbf{J}_{i,j} = \begin{cases} E_{1,1}(z), & \text{top left corner} \\ N_{\frac{M+1}{2},\frac{M+1}{2}}(z), & \text{bottom apex} \\ \Big(W(z), \quad S(z)\Big)_{1,\frac{M+1}{2}}, & \text{top right corner} \\ \Big(N(z), \quad E(z)\Big)_{i,j}, & \text{diagonal } (i=j) \\ \Big(W(z), \quad E(z), \quad S(z)\Big)_{1,j}, & \text{top row} \\ \Big(W(z), \quad N(z), \quad S(z)\Big)_{i,\frac{M+1}{2}}, & \text{right column} \end{cases} \tag{22}$$

and the corresponding $\mathbf{b}_\mathbf{z}$ vector with cases given respectively (but with case statements removed for brevity):

$$\mathbf{b}_{\mathbf{z}i,j} = \begin{cases} W_{1,2}(z), \\ \Big(S_{\frac{M-1}{2},\frac{M+1}{2}}(z), \quad X(z)\Big), \\ \Big(E_{1,\frac{M-1}{2}}(z), \quad N_{2,\frac{M+1}{2}}(z)\Big), \\ \Big(S_{i-1,j}(z), \quad W_{i,j+1}(z)\Big), \\ \Big(E_{1,j-1}(z), \quad W_{1,j+1}(z), \quad N_{2,j}(z)\Big), \\ \Big(E_{i,\frac{M-1}{2}}(z), \quad S_{i-1,\frac{M+1}{2}}(z), \quad N_{i+1,\frac{M+1}{2}j}(z)\Big), \end{cases} \tag{23}$$

and corresponding $\mathbf{H}_\mathbf{z}$ matrix:

$$\mathbf{H}_{\mathbf{z}i,j} = \begin{cases} az^{-3}, & \text{top right corner} \\ \begin{bmatrix} z^{-1} & z^{-1} \end{bmatrix} & \text{bottom apex} \\ \begin{bmatrix} H_r & H_b \\ 2H_b & H_a \end{bmatrix}, & \text{top right corner} \\ \begin{bmatrix} 0 & z^{-1} \\ z^{-1} & 0 \end{bmatrix}, & \text{diagonal} \\ \begin{bmatrix} H_a & H_b & H_b \\ H_b & H_a & H_b \\ H_b & H_b & H_a \end{bmatrix}, & \text{top row} \\ \begin{bmatrix} 0 & \frac{z^{-1}}{2} & \frac{z^{-1}}{2} \\ z^{-1} & -\frac{z^{-1}}{2} & \frac{z^{-1}}{2} \\ z^{-1} & \frac{z^{-1}}{2} & -\frac{z^{-1}}{2} \end{bmatrix}, & \text{right column} \end{cases} \tag{24}$$

where

$$H_r = \frac{az^{-3}}{2 + az^{-2}}, \quad H_a = \frac{-z^{-1}}{2 + az^{-2}}, \quad H_b = \frac{az^{-3} + z^{-1}}{2 + az^{-2}}, \tag{25}$$

4.2. Dependency Elimination Algorithm

Eliminations are done by updating and operating on the dependency matrix $\mathbf{H}_\mathbf{z}$. Vector $\mathbf{b}_\mathbf{z}$ is shown for reference only. First steps are shown for illustration, but later supressed for brevity.

4.2.1. Traverse first row from left to right:

eliminate $E_{1,1}$ in $\mathbf{b}_{\mathbf{z}1,2}$:

$$\mathbf{J}_{\mathbf{z}1,2} = \mathbf{H}_{\mathbf{z}1,2}\left(\cancel{E_{1,1},}\overset{W_{1,2}}{\nearrow} W_{1,3}, \quad N_{2,2}\right)$$

by multiplying the first column of $\mathbf{H}_{\mathbf{z}1,2}$ by the first row of $\mathbf{H}_{\mathbf{z}1,1}$ and updating $\mathbf{H}_{\mathbf{z}1,2}$:

$$\mathbf{H}_{\mathbf{z}1,2} \rightarrow \left[\begin{bmatrix} H_{1,1} \\ H_{2,1} \\ H_{3,1} \end{bmatrix} \times \begin{bmatrix} H_{1,1} \end{bmatrix}_{1,1} \quad \begin{matrix} H_{1,2} & H_{1,3} \\ H_{2,2} & H_{2,3} \\ H_{3,2} & H_{3,3} \end{matrix}\right]_{1,2}$$

eliminate $W_{1,2}$ in $\mathbf{b}_{\mathbf{z}1,2}$ (recursive)

$$\mathbf{J}_{1,2} = \mathbf{H}_{\mathbf{z}1,2}\left(\cancel{W_{1,2},} \quad W_{1,3}, \quad N_{2,2}\right),$$

$$\begin{aligned}\mathbf{H}_{\mathbf{z}1,2} \quad \rightarrow \quad & \left[\begin{bmatrix} 1 \\ H_{2,1} \\ H_{3,1} \end{bmatrix} \times \begin{bmatrix} H_{1,2} & H_{1,3} \end{bmatrix}\right]_{1,2} / (1 - \begin{bmatrix} H_{1,1} \end{bmatrix}_{1,2}) \\ & + \begin{bmatrix} 0 & 0 \\ H_{2,2} & H_{2,3} \\ H_{3,2} & H_{3,3} \end{bmatrix}_{1,2} \end{aligned} \tag{26}$$

eliminate $E_{1,2}$ from $\mathbf{b}_{\mathbf{z}1,3}$

$$\mathbf{J}_{1,3} = \mathbf{H}_{\mathbf{z}1,3}\left(\cancel{E_{1,2},}\overset{W_{1,3}}{\nearrow} W_{1,4}, \quad N_{2,3}\right),$$

$$\mathbf{H}_{\mathbf{z}1,3} \rightarrow \left[\begin{bmatrix} H_{1,1} \\ H_{2,1} \\ H_{3,1} \end{bmatrix} \times \begin{bmatrix} H_{2,1} & H_{2,2} \end{bmatrix}_{1,2} \quad \begin{matrix} H_{1,2} & H_{1,3} \\ H_{2,2} & H_{2,3} \\ H_{3,2} & H_{3,3} \end{matrix}\right]_{1,3}$$

eliminate $W_{1,3}$ in $\mathbf{b}_{\mathbf{z}1,3}$

$$\mathbf{J}_{1,3} = \mathbf{H}_{\mathbf{z}1,3}\left(\cancel{W_{1,3},}\overset{N_{2,2}}{\nearrow} W_{1,4}, \quad N_{2,3}\right), \tag{27}$$

$$\begin{aligned}\mathbf{H}_{\mathbf{z}1,3} \rightarrow & \left[\begin{bmatrix} 1 \\ H_{2,1} \\ H_{3,1} \end{bmatrix} \times \begin{bmatrix} H_{1,2} & H_{1,3} & H_{1,4} \end{bmatrix}\right]_{1,3} \\ & / \left(1 - \begin{bmatrix} H_{1,1} \end{bmatrix}_{1,3}\right) + \begin{bmatrix} 0 & 0 & 0 \\ H_{2,2} & H_{2,3} & H_{2,4} \\ H_{3,2} & H_{3,3} & H_{3,4} \end{bmatrix}_{1,3} \end{aligned}$$

eliminate $E_{1,j}$ in $J_{1,j+1}$ to row end;
eliminate $W_{1,j+1}$ in $J_{1,j+1}$ to row end;

4.2.2. Traverse first row from right to left:

eliminate $W_{1,j}$ from $J_{1,j-1}$, for $3 <= j <= \frac{M+1}{2}$.

4.2.3. Traverse from first to second row

eliminate $S_{1,j}$ in $J_{2,j}$, for $2 <= j <= \frac{M+1}{2}$.

4.2.4. Traverse second row from left *to* right

eliminate $E_{2,j}$ in $J_{2,j+1}$ to row end;
eliminate $W_{2,j+1}$ in $J_{2,j+1}$ to row end;

4.2.5. Traverse second row from right *to* left

eliminate $W_{2,j}$ from $J_{2,j-1}$ to row end;

4.2.6. Traverse second row from left *to* right

eliminate $N_{2,j}$ from $J_{2,j}$, $J_{2,j+1}$, for $2 <= j <= \frac{M+1}{2}$

4.2.7. Traverse to next (and subsequent) rows and repeat

Once the bottom junction is reached, the transfer function is given by the final expression in $\mathbf{H}_{\mathbf{z}\,\frac{M+1}{2},\frac{M+1}{2}}$ for an unit-impulse input.

MXM (order)	full mesh	symmetric mesh	DF II
3×3 (8)	2 ms	0.7 ms	0.6 ms
5×5 (16)	6 ms	1.1 ms	1.0 ms
7×7 (26)	13 ms	1.2 ms	1.5 ms
9×9 (38)	22 ms	4.1 ms	2.0 ms
11×11 (52)	34 ms	6.5 ms	2.9 ms

Table 1. Mesh computation times for the square $M \times M$ mesh, mesh with applied symmetry, and direct form II implementation of transfer function with given order (and where comparable timings are expected for a second-order section cascade) for one second of 44100-Hz-audio).

5. CONCLUSION

Because of the computational limitations of the 2-D waveguide mesh, it is not often used for real-time applications. This work presents an initial step toward creating a frequency-domain representation, with the eventual aim of developing a real-time implementation. Though the current algorithm yields a transfer function, which for low-orders can be implemented in canonical form (direct form II), higher orders experience coefficient quantization that perturbs the poles, often outside the unit circle, and leading to instabilities. Furthermore, direct form II would not be an ideal real-time implementation due to known instabilities that arise when coefficients are made time varying. It is thus desirable to obtaining the transfer function in factored form so that alternate implements, such as a back of second order sections, or a parametric impulse response, can be used. This work is being developed in Mathematica, a symbolic programming language, allowing for analysis of symbolic (rather than numeric) results, less prone to precision error. Table 1 shows current timing results.

REFERENCES

[1] S. V. Duyne and J. O. Smith, "The 2-d digital waveguide mesh," in *Proceedings of the 1993 IEEE Workshop of Applied Signal Processing to Audio and Acoustics*. New Paltz, New York: IEEE Press, October 1993.

[2] D. T. Murphy, C. J. C. Newton, and D. M. Howard, "Digital waveguide mesh modelling of room acoustics: Surround-sound, boundaries and plugin implementation," in *Proceedings of the Conference on Digital Audio Effects (DAFX-01)*, Limerick, Ireland, December 2001.

Real-Time Physical Model of a Cymbal Using the Full von Kármán Plate Equations

Florian Pfeifle

University of Hamburg
Florian.Pfeifle@uni-hamburg.de

ABSTRACT

Metallic idiophones like cymbals or gongs have been shown to exhibit nonlinear acoustic effects such as mode coupling, pitch glides or transitions from deterministic to chaotic regimes for a wide range of excitation patterns. Capturing these effects in a physical model is essential when aiming at realistic simulation and synthesis results for research or musical applications. The von Kármán (VK) plate equations have been successfully applied to model nonlinear effects typically found in cymbals, gongs or stiff membranes. But, capturing these complexities in a structural vibration model comes with the drawback of higher computational cost when compared to more simple plate theories as the Kirchhoff or Berger plate models. Indeed, when written in full form, the VK plate equation is composed of a system of three coupled differential equations, one for the transverse motion and two for the in-plane (longitudinal) motion of the modeled plate. This leads to the predicament that the VK equations are not easily implementable in real-time on standard hardware for reasonably sized simulation domains. Utilising specialised hardware platforms such as general purpose graphics processing units (GPGPUs) or field programmable gate array (FPGA) hardware, has been shown to be a viable alternative for accelerating physical models of musical instruments enabling real-time simulation and auralisation of complex large scale geometries. In this work, a real-time version of the full VK equation system implemented on FPGA hardware is presented. The model is computed on a development board consisting of two FPGA chips that can be parametrized and controlled from a personal computer. The VK model is discretized using finite difference approximations for the spatial domain and is iterated in time using a locally symplectic iteration scheme leading to a fully explicit formulation of the problem. The excitation is modeled as a simplified stick-cymbal interaction that can be parametrized from a personal computer. Different stroke patterns are used to excite the model and resulting vibrations are compared to measurements taken on a real cymbal showing the variability of the simulated sound and giving realistic physical bounds for the proposed implementation.

String / Fretboard Interaction in Electric Basses: Simulations and Experiments

Clara Issanchou,[1†] *Jean-Loïc Le Carrou,*[1] *Cyril Touzé,*[2] *Benoît Fabre,*[1] *Olivier Doaré*[2]

[1]Sorbonne Universités, UPMC Univ Paris 06, CNRS, LAM / Institut ∂'Alembert
[2]IMSIA, ENSTA ParisTech-CNRS-EDF-CEA, Université Paris Saclay
[†]i17.clara@gmail.com

ABSTRACT

The non-linear behaviour of string instruments, either intended or not, is decisive for instrument making and playing techniques. In particular, collisions between a string and an obstacle arise in numerous musical instruments, among which tanpuras, sitars, electric basses, etc. In the case of electric basses, contact occurs for specific playing techniques such as pop or slap, for which the musician plucks or hits a string strongly enough to make a string / fretboard interaction appear. In the present study, this non-linear vibratory behaviour is both simulated and measured for an electric bass. The numerical model is based on a modal description together with a regularised contact force, and was recently developed and applied to the contact between a vibrating string and a point obstacle [1]. The scheme is conservative and includes a weak coupling between the string and the structure. Besides, highly controlled experiments are led on an electric bass in order to compare numerical and experimental data. A good agreement between numerical simulations and experiments is found. In particular, specificities of the resulting sounds are faithfully reproduced. Finally, the influence of numerical and physical parameters, possibly related to tuning issues, is exhibited through a numerical parametric study.

[1] C. Issanchou, S. Bilbao, J.-L. Le Carrou, C. Touzé, and O. Doaré, A modal-based approach for the non linear vibration of strings against a unilateral obstacle: simulations and experiments in the pointless case, Journal of Sound and Vibration, 393, pp. 229-251 (2017).

Calibration for Impact Measurements of Qin Sound Hole Velocities

Chris Waltham

Department of Physics & Astronomy, University of British Columbia, Vancouver BC, Canada V6T 1Z1
cew@phas.ubc.ca

ABSTRACT

The culture of making and playing the ancient Chinese zither, the qin, indicates that the radiation from the instrument's sound holes is crucial for the quality of the sound. However, the contribution of the air modes is much less clear than that of the air modes of Western instruments like the guitar and violin. Presented here is a method of measuring air velocities in sound holes with a small microphone array, using impact excitation. The results are verified in a three-step calibration process involving measuring the sound hole velocities for a guitar, which are known independently using knowledge of the character of low frequency modes and measurement of the monopole radiation component with a larger 4π microphone array. The technique is then applied to the case of the qin.

1. INTRODUCTION

The sound boxes of nearly all string instruments consist of a light wooden shell with one or more "sound holes". The acoustic radiation from the sound box is a result of vibrations of both the wooden shell and the air in the mouth of the sound hole. Assessing the wood contribution from measured operating deflection shapes (hereafter referred to as "mode shapes") is straightforward. Measuring the air contribution, which may be substantial, and even dominant, is more involved. Matched pairs of microphones are used for acoustic velocities in other applications but for musical instruments, the wavelength is often much larger than the whole instrument, making it hard to see a phase difference between two microphone in the small space of a sound hole mouth. The common approach is to use a small planar microphone array to image the acoustic velocity in the sound hole.

Data from microphone arrays are usually analyzed within the framework proposed by Williams et al.[1], in the early 1980s. In this formulation, radiation is emitted by a flat plane surface vibrating in a complex manner, within an infinite baffle, sampled by an array of sensors in the near field (NF). The measured acoustic field is then transformed into k-space, reversed back to source, and then transformed back into (x, y) space to reveal a velocity map of the surface of the radiator.

Real measurements deviate in two significant ways from that assumed in the analysis. First, the instrument is generally not mounted in a heavy, immobile baffle, and the edge of any instrument is never more than a fraction of a wavelength away from the edge of the array. Second, even small microphones cannot be considered of negligible size compared to the separation of array and radiator (i.e. a few cm). Hence the microphones and mountings can affect the sound field in a manner that is hard to estimate.

Here NF-array measurements are used to deduce sound hole velocities which are compared with values known from other means. Two guitar family instruments with simple low frequency mode structures [2] have been used. Their mode shapes and further-field ("FF", the size of the anechoic chamber precluded a true "far-field" array) radiation patterns where measured, and the monopole components extracted. In all three cases - mode map, FF- and NF-arrays, the instruments were excited by a small impact hammer.

The method used is similar to that reported by Bissinger et al.[3] for the violin, but differing in the following respects. First, the simplicity of the low frequency guitar modes was employed to make a reliable estimate of the monopole flow using only the mode maps, and no mathematical model was required except to assume the relative phase of wood and air for the A0 and T1 modes (i.e. π out of phase and in-phase, respectively). The size of the guitar relative to the violin, is also an advantage, as is the flat soundboard. Second, the velocity in the sound holes was assumed to be uniform and fitted, rather than imaged, so no back-propagation was needed. Thus, as a verification procedure for NF-array analysis, it is fairly simple. The output is a normalized sound hole velocity in m/s/N. This technique will now be used in a qin construction project which will attempt to identify the contributions to the qin's radiation, which are much more subtle, even at low frequency, than those of the guitar. In particular, the effect of wooden "absorbers" placed in the sound hole openings will be examined; these are considered to be very important elements of the qin sound[4].

Section 2 describes the principle and sequence of the calibration procedure. Section 3 gives details of the measurements themselves. Section 4 outlines how the calibration results were analyzed and Section 5 gives the results and how the overall uncertainties were established. Section 6 shows some preliminary results of applying this technique to the question of what sound holes contribute to the radiation of a qin. Section 7 indicates the next steps with regard to the qin construction project.

2. CALIBRATION PRINCIPLE

Guitars have simple mode structures at low frequency. The lowest mode with substantial radiation is invariably one labelled A0 with the air and front plate each moving largely as a piece and in antiphase with each other. There follows a higher

mode, often labelled T1, which is similar except air and wood are now in phase with each other. The far-field radiation at low frequency is in large measure isotropic, i.e. the monopole component is much larger than any higher multipole. Hence if one knows, from modal analysis, the wood contribution to the monopole radiation, one can deduce the air contribution from both the absolute and relative sizes of the A0 and T1 radiation peaks. Thus direct measurements of sound hole velocities using the NF-array can be compared with those deduced from the FF-array and mode maps.

3. CALIBRATION MEASUREMENTS

The NF-array was constructed from 30 Panasonic WM-61D quarter-inch electrets mounted in aluminum tubes arranged in a 6×5 matrix of 3 cm pitch. The 30 cm long tubes were held in place by an plywood sheet, with as many holes as structural stability would allow, to minimize reflection. Each microphone was fitted with a Linkwitz preamplifier to improve linearity, which was a concern given the large dynamic range of impact measurements. The microphones were calibrated by a piston calibrator, at 250 Hz and 1 kHz, 94 dB and 110 dB. Each instrument was suspended, in turn, 1-2 cm away from the NF-array and excited by an instrumented impact hammer, striking one side of the bridge. The mode maps were measured with the same hammer, and a 0.6 g accelerometer placed on a 5 cm grid around the instrument (including the ribs, to check that their contribution to the radiation is small).

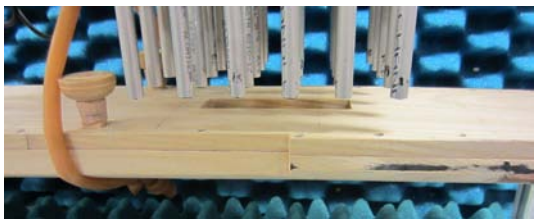


Figure 1. Measurement of an unlacquered qin sound hole using the NF-array.

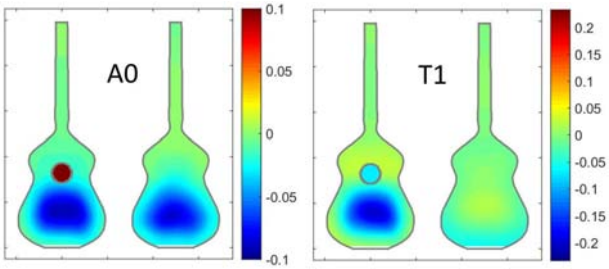


Figure 2. Maps of the A0 and T1 modes of the calibration guitar. The contours represent admittances in m/s per newton force at the treble end of the bridge. The hole values are those deduced from the complete process described in this paper but reduced by a factor 20 so the wood shapes can be seen.

A similar impact method was used to measure each instrument's radiation patterns by placing them inside a 4π 30-microphone FF-array of radius 92 cm, rotatable to give 240 microphone positions[5]. Pressure data, Fourier-transformed and normalized to hammer force, were recorded using the FF-array in a total of 240 microphone positions (8 rotations of the array). The complex data from each microphone position were then simply added together, weighted for the solid angle fraction that each position represented. The result was the monopole component of the radiation, as higher multipoles would sum to zero in this procedure.

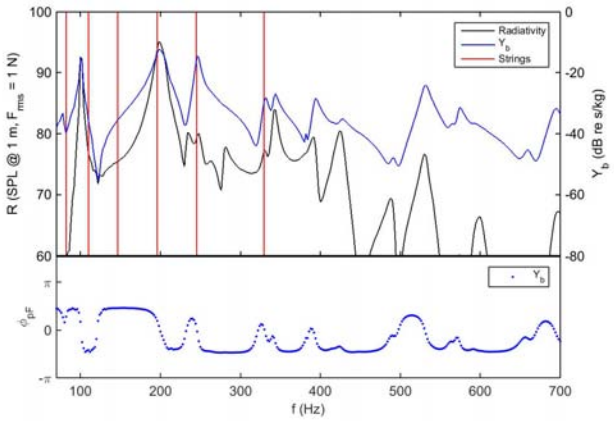


Figure 3. The monopole radiation component from the calibration guitar, measured with the FF-array, is shown in black; the bridge admittance is shown in blue; the string fundamental frequencies are shown in the red. The lower plot shows the phase of the bridge admittance.

The NF and FF measurements were made in an anechoic chamber. However, it is noted that the NF measurements made in a non-anechoic room, while commissioning the NF-array, turned out to be essentially the same as those subsequently made in the anechoic chamber.

4. ANALYSIS

Conversion of the monopole pressure p_0 to a volume flow Q was effected with a monopole model whose output depended weakly on an assumed monopole radius a, and the microphone distance r from the centre of the monopole (Eqn. 1). The monopole radius was estimated from angular distributions of the radiation to be 15 cm. The uncertainties introduced into the volume flow (in m^3/s/N) by varying the radius over reasonable values (10-20 cm) were small compared to experimental effects.

$$p_0 = \frac{\rho c Q}{2\pi a r \sqrt{1 + \frac{1}{k^2 a^2}}} \tag{1}$$

The contribution of the wood to the volume flow (in m^3/s/N) was calculated using mode map data. Knowing the area of the sound hole, it was then possible to estimate the sound hole velocity required to produce the measured volume flow, given the

phase relationship assumed from the mode type identification, A0 or T1.

Starting with a trial value for the acoustic velocity in the sound hole, and the mode map based on the accelerometer measurements, a 2-D velocity map ($u_s(x,y)$) was created for the whole top surface of the guitar for each frequency of interest corresponding to a propagating wavelength λ. The result was Fourier-transformed into $\hat{u_s}(k_x,k_y)$, and converted to a surface pressure map in k-space $\hat{p_s}(k_x,k_y)$ via the complex wavenumber K, given by Eqns. 2 and 3.

$$\begin{aligned} K &= ((\frac{2\pi}{\lambda})^2 - k_x^2 - k_y^2)^{1/2}, \text{ if } k < \frac{2\pi}{\lambda} \\ K &= j(k_x^2 + k_y^2 - (\frac{2\pi}{\lambda})^2)^{1/2}, \text{ if } k > \frac{2\pi}{\lambda} \end{aligned} \tag{2}$$

$$\hat{p_s}(k_x,k_y) = \frac{-\omega\rho}{K}\hat{u_s}(k_x,k_y) \tag{3}$$

The surface pressure can be propagated into a pressure field $p_{NF}(x,y)$ at the plane of the NF-array, a distance d from the surface of the guitar, using the transform $\hat{G}$ (Eqn. 4) and inverse fourier transform F^{-1} (Eqn. 5).

$$\hat{G}(k_x,k_y,d) = e^{jKd} \tag{4}$$

$$p_{NF}(x,y) = F^{-1}(\hat{p_s}(k_x,k_y)\hat{G}(k_x,k_y,d)) \tag{5}$$

The pressure field $p_{NF}(x,y)$ could then be sampled at the positions of the physical microphones for comparison with measured values.

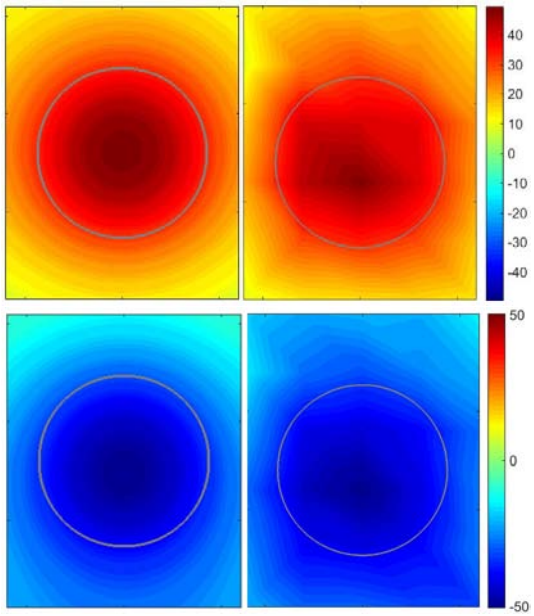


Figure 4. Maps of acoustic pressure 2 cm above the calibration guitar sound hole for the 100 Hz A0 mode (upper) and 200 Hz T1 mode (lower): comparison of expected NF-array pressures, simulated using surface velocities from mode maps and monopole radiation data (left), and raw NF-array data (right). Uniform sound hole velocities of 2.8 and -1.3 m/s/N were used for the simulation of A0 and T1 modes respectively. The contours of all maps show normalized pressures in Pa/N.

5. CALIBRATION RESULTS

The results for the A0 and T1 modes are shown in Fig. 4. The magnitude and shape of the data and simulated peaks are essentially the same. Apart from some non-uniformity in the microphone signals, there is no evidence of any complex structure in the signal, which supports the original assumption that the sound hole velocity is uniform across the surface of the hole. The quality of the agreement remains good up to about 400 Hz; above that the mode shape become more complex.

A summary of the numerical results for the A0 and T1 resonances is shown in Table 1. After analyzing the discrepancies one can assess the overall systematic uncertainties associated with the method to be $\pm 20\%$ (twice the value given in Ref. [3]). There are contributions from several sources, particularly the difficult-to-assess deviation from an ideal baffled radiator, and also the monopole radius assumption, microphone calibration between arrays, and seasonal variability of the instruments due to environmental changes.

6. QIN CAVITY MODES

The qin has a complex shape, formed of two halves, upper and lower, with two sound holes on the underside. Into the top are carved two absorbers that partially obscure the sound holes. The basic radiation mechanisms have recently been reported[6]. The instrument has a long narrow soundbox with two sound holes partly obscured by "absorbers". According to ancient lore, the holes and absorbers are crucial to the tone of the qin. However, it is by no means obvious why or how this is so. The sound holes contribute in a subtle manner to the overall radiation, unlike those in Western instruments (whose purpose is to ensure efficient radiation at low frequencies). The effect of the absorbers is hard to assess without cutting up the instrument (they are carved with the top plate out of one piece).

The technique for measuring sound hole velocities tested on the guitar have been applied to a qin ("qin A" of Ref. [6]). The velocities in the holes are small (a few % of those seen in the guitar) and not greatly larger than that of the surrounding wood for most modes, hence there is often some difficulty extracting the "air"-signal. The relatively small area of the holes compared to that of the wood reduces their effect on the overall radiation still further. A unusually simple example of an air-mode is shown in Fig. 5. Here the second cavity mode, previously identified by placing a string of small microphones in the cavity, can be seen to be the cause of the velocities appearing at the sound holes. The variation of velocity along the length of the sound holes is so large that they were divided into ten strips for fitting. In the case of the larger hole, most of the activity is observed in one end of the hole; this is a due to the nature of the cavity mode, with the nodal points occurring somewhere in the opening. There is very little wood motion at this frequency; the wood-air coupling is small.

Freq. (Hz)	u_{hole} (model) (m/s/N)	Q_{wood} (meas.) (m^3/s/N)	Q_{total} (thy.) (m^3/s/N)	Q_{total} (meas.) (m^3/s/N)	p_{NF} (thy.) (Pa/N)	p_{NF} (meas.) (Pa/N)
102 (A0)	2.8	-0.0046	0.0122	0.0147	48.0	48.3
201 (T1)	-1.3	-0.0056	-0.0135	-0.110	-49.7	-48.5

Table 1. Results of the sound hole velocity analysis for the A0 and T1 modes of the calibration guitar. The model soundhole velocities u_{hole} were used to calculate the hole volume flow which, added to the wood contribution to the monopole flow Q_{wood} (meas.) from mode map measurements, yielded a "theoretical" total flow Q_{total} (thy.). This value is compared to Q_{total} (meas.) taken from FF-array measurements. The "theoretical" maximum pressure p_{NF} (thy.) 2 cm from the sound hole is compared to that predicted, p_{NF} (thy.), from the model velocity and wood contribution, using the propagation method.

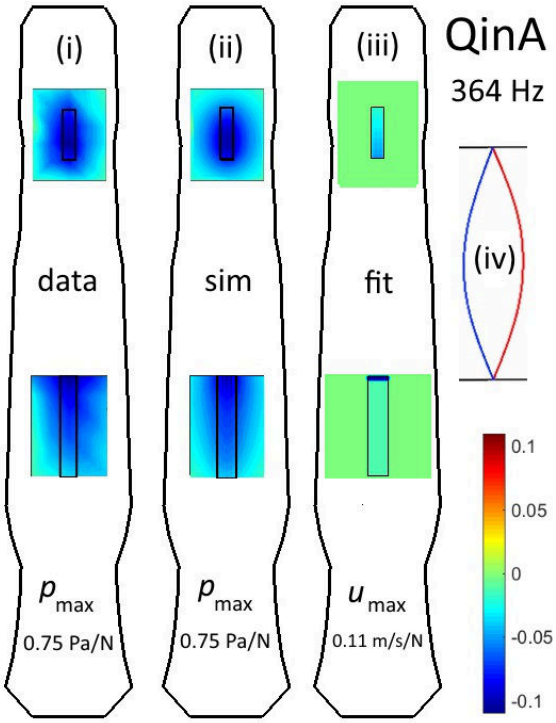


Figure 5. Cavity mode 2 for qin A. Left to right: (i) NF-array contours maps of acoustic pressure 2 cm above sound holes; (ii) simulation of (i); (iii) deduced velocity profiles in the sound holes used in simulation (ii); (iv) deduced internal cavity mode. Note all the flow in the larger hole is concentrated at the top end.

7. NEXT STEPS

A valuable way to understand how the vibrational properties of a soundbox arise from its components, in the absence of a credible physical model, is to characterize each component separately, either as the instrument is being built, or by dismantling an existing instrument. Both of these procedures have been used extensively in the case of the violin family and are now being applied to the qin. To investigate fully the acoustical nature of the qin, a new instrument has been made from existing parts such that a complete qin can be easily assembled and disassembled. The instrument is made of a paulownia upper piece and a catalpa lower piece. The whole will covered in traditional fashion by a thick layer of lacquer, mixed with deer horn. The vibroacoustic effect of this combination has not yet been measured. The absorbers in the sound holes of the qin have been made to be removable, and thus using the technique described in this paper, their function may possibly be discerned.

Acknowledgements

The author thanks Alan Thrasher, and the UBC Chinese Orchestra for the loan of instruments; Murray Hodgson and Glenn Jolly of the UBC Mechanical Engineering Department and Gar Fisher of the UBC Physics and Astronomy Department for technical assistance.

REFERENCES

[1] E. G. Williams, J. D. Maynard, and E. Skudrzyk, "Sound source reconstruction using a microphone array," *J. Acoust. Soc. Am.*, vol. 68, pp. 340–344, 1980.

[2] O. Christensen and B. Vistisen, "Simple model for low-frequency guitar function," *J. Acoust. Soc. Am.*, vol. 68, pp. 758–766, 1980.

[3] G. Bissinger, E. G. Williams, and N. Valdivia, "Violin f-hole contribution to far-field radiation via patch near-field acoustical holography," *J. Acoust. Soc. Am.*, vol. 121, pp. 3899–3906, 2007.

[4] J. Binkley, *Abiding with Antiquity: Translations from the Yu-ku-chai-ch'in'pu*. Lulu Press, 2006. [Online]. Available: http://web.cecs.pdx.edu/~jrb/chin/index.html

[5] C. E. Waltham, E. Koster, N. Smit-Anseeuw, and A. Zimmer, "Acoustic imaging of string instrument soundboxes," *Proc. Meet. Acoust.*, p. 035004, 2013.

[6] C. E. Waltham, E. Koster, and Y. Lan, "Acoustics of the qin," *J. Acoust. Soc. Am.*, vol. 139, pp. 1592–1600, 2016.

Non-Occidental Lutes: Relationship between Making Process and Functionality

Henri Boutin, Sandie Le Conte,[†] *Philippe Bruguiere*

Cité de la musique Philharmonie de Paris
[†]sleconte@cite-musique.fr

ABSTRACT

The Musée du Quai Branly and the Musée de la musique both in Paris keep in their collections more than 300 non-occidental lutes coming from Africa (Ud, Ginbri, Kwitra can be given as examples), Asia (Setar, Sarod, Ektara, Biwa, Yueqin etc.) and South Africa (Chiapas, Charango etc.). From the organological classification [Grove] these instruments present some common characteristics (without taking into account the sound): chordophones in which a string bearer and a resonator are united and cannot be separated, and in which the strings runs parallel with the soundboard. However do they have acoustical common characteristics? Trying to answer this question a large study has been led on more than one hundred instruments. In most cases, because these instruments are kept in museums they are not played anymore. Moreover they are not always in a good state (cracks on the soundboard and soundbox, missing parts), so a methodology has to be applied on every instrument independently on their conservation state and not too much time consuming, regarding the number of instruments to be analysed. The methodology consists in measuring systematically the geometry of some elements of this wide corpus (soundboard and soundbox dimensions, neck dimensions, vibrating length, rosacea dimensions when it exists). In addition to these measurements, vibrational parameters are acquired using an impact testing method. The accelerometer, measuring the vertical acceleration, is fixed close to the treble foot of the bridge, while impacts are applied on 9 different points on the soundboard. Then the accelerometer is fixed on the neck and the neck is impacted in 2 more points. For each instrument, the dimensions of the soundbox are compared to the maximum string vibrating length. Generally, in string instruments, some body resonances, resulting from the coupling between structure and air vibration in the soundbox, amplify the radiated sound. Thus, in order to investigate the acoustics of the lutes, the lowest resonance, called Helmholtz frequency, is measured, thanks to a microphone positioned very close to the hole and recording the hammer impact. So, from the vibrating measurements the frequencies of the two largest inertance peaks match those of the sound spectrum. It is then possible to show some correlation and some common characteristics between geometrical and vibrating measurements.

Analysis of Sound Post Adjustment using a Hybrid Bowed String Model

Frédéric Ablitzer,[†] *François Gautier*

Laboratoire d'Acoustique de l'Université du Maine
[†]frederic.ablitzer@univ-lemans.fr

ABSTRACT

The position of the sound post on a violin is considered by makers as an important adjustment parameter, which affects both tonal and playing qualities of the instrument. The aim of this study is to investigate the influence of sound post adjustment on the response of the instrument when played. Mobility measurements at the bridge are made to evaluate the consequences of typical variations in sound post position on the dynamic behavior of the body as seen by the string. Then, modal identification is performed for each configuration using the high resolution method ESPRIT. The modal parameters of the body are used to feed a hybrid physical model of bowed string. In this model, the dynamic behavior of each string is described using a theoretical modal basis, while that of the body is represented by an experimental modal basis. The set of modal equations thus obtained is combined with a friction model to simulate the bowed string dynamics. Time-domain simulations are made to obtain the response of the instrument to the same gesture, for different adjustments of the sound post. A set of indicators is proposed to highlight the effects of these adjustments on both transient and steady-state regimes.

Influence of the Bridge, Island Area, Bass Bar and Soundpost on the Acoustic Modes of the Violin.

Colin Gough

School of Physics and Astronomy, University of Birmingham, B15 2TT, UK
profgough@googlemail.com

ABSTRACT

The influence of the bridge, f-holes, island area, soundpost and bass bar on the vibro-acoustic modes of the violin has been investigated over a wide range of frequencies using COMSOL shell structure finite element (FEA) computations. The body shell modes other than in the island area have been strongly damped to reveal the smoothly varying frequency spectrum of the bridge-island area input filter controlling the energy transfer from the bowed string to the radiating shell modes. Of special relevance is the so-called *BH*-feature between around 2-3 kHz, often correlated with the quality of fine sounding violins.

1. INTRODUCTION

In a pair of related publications [1, 2], the violin was modeled as a thin-walled, guitar-shaped box, with arched top and back plates coupled around their edges by shallow ribs. COMSOL 3.5 finite element shell structure software was used as an aid to understanding the acoustically important signature modes below around 1 kHz and their relationship to the freely supported top and back plates before assembly. This paper extends the investigation to address the influence of the bridge, f-holes, the island area between them and the bass bar and offset soundpost on the vibro-acoustic modes of the violin up to 10 kHz. These components introduce an acoustic filter between the transverse bowed string forces at the top of the bridge and the major radiating surfaces of the body shell, hence quality of sound of an instrument.

Of particular interest is a rather broad peak in amplitudes of the excited plate resonances often observed in the range 2-3 kHz, referred to as the *BH* or bridge hill feature. This was originally associated with the in-plane mode of the top section of the bridge rocking about its waist originally identified by Reinicke [3] - typically at around 3 kHz, when mounted on a rigid base.

Dunnwald [4] and Jansson[5] both noted the appearance of such a feature in several classic Cremonese Italian violins. However, Bissinger [6], observed little evidence for any such correlation in measurements on 12 violins of widely varying quality.

Jansson and Niewczyk[7] also observed the *BH* hill feature on some instruments with a rigid bridge, with no rocking mode resonance in the frequency range of interest. They therefore concluded that it was not necessarily a property of the bridge alone. Subsequently, Durup and Jansson [8], using simplified inverted L-shaped "f-holes" cut into a rectangular flat plate, identified a strong correlation of the *BH* feature with a *flapping* resonance of the upper inner-corner "wings" of the island area. Jansson and co-researchers [9] have recently identified a similar mode from finite element computations for simply modeled f-holes cut into both flat and arched rectangular plates.

At high frequencies, there is a high density and large overlap of the damped resonances of the body shell. In such cases a statistical approach can be used to describe the frequency-averaged response of the plate modes, as described by Cremer [10], Chpt.11 and Woodhouse[11].

Woodhouse used a SEA (Statistical Energy Analysis) approach to derive "skeleton" curves describing the smoothly varying, mode-averaged, input filter response of a simply modeled bridge mounted on a rectangular plate. In this model, the bridge vibrations were strongly damped by the resistive, thin plate, point driving-impedance of the supporting plates, $Z_o = \frac{\sqrt{3}}{4} c_{gm} \rho h^2$, where $c_{gm}\rho$ is the geometric mean of the along- and cross-grain characteristic impedances of longitudinal waves and h is the plate thickness. The resulting losses significantly broaden the rocking mode resonance. In addition, the mutual plate deformations around the two feet of the bridge result in a filter response dependent on their separation.

As recognized by Woodhouse, the original model ignored the influence of the nearby free edges of the f-holes forming an island area with its own localised modes between them, though they were included for a parallel slit geometry in a later hybrid SEA + FEA MSc project by Smith[12].

Such complications are automatically taken into account in the present paper, which describes the use of finite element computations to investigate and understand the influence of the bridge and plate properties on the *BI*, bridge-island, input filter and the additional role played by the offset soundpost and bass bar. Hopefully, such understanding may provide a helpful guide to luthiers in making and setting up any instrument of the violin family to optimize its sound.

2. THE FINITE ELEMENT MODEL

The body shell geometry of the violin used in this paper is identical to that described earlier papers [1, 2], except that an island area is now defined between the upper and lower bout areas of the top plate, as illustrated in Fig.1. The bridge and island area are only weakly damped (η = 0.01), while the remaining areas of the top and back plates are strongly

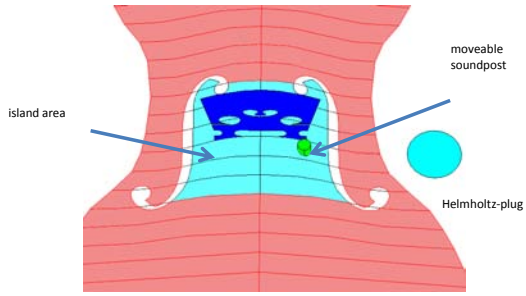


Figure 1. The division of the top plate into an island area separating the strongly damped lower and upper bout areas.

damped (η =1). The filter response of the bridge and island area can then be computed as a smoothly varying function of frequency uncomplicated by the multi-resonant response of the body shell vibrations, as illustrated in figure 2.

This is similar to Woodhouse's analytic "skeleton" curve model [11], but with the coupling to the vibrating plate modes now moved to the upper and lower boundaries of the island area, rather than under the two bridge feet. As the energy radiated from the ends of the island area rapidly decays to a small value at the plate edges, we have simply pinned the top plate edges. Coupling to the back plate modes is then only via the soundpost, when present.

Because the influence of the along- to cross-grain elastic anisotropy, $A = E_{along}/E_{cross}$, is important for the localised island area modes, the lengths of the top and back plates have been decreased by $A^{1/8}$ and widths increased by the same factor maintaining the same area and geometric mean $E_{gm} = \sqrt{E_{along}E_{cross}}$. At high frequencies, when arching is no longer important, the bending wave solutions of the transformed geometry then reproduce those of the unscaled anisotropic thin plates (Cremer [10]§11.2). For a spruce top plate with typical anisotropy of 20, a scaling factor of 1.45 has been used, and 1.22 for the maple back plate with typical anisotropy of around 5.

A bridge of mass 2.2 g has been used, with its thickness tapered from 4 mm at the bridge feet to 1.5 mm at the top. A density of 660 kgm^{-3} has been assumed with an isotropic Young's constant of 4e9 GPa giving a rocking frequency resonance of the top half of the bridge about the waist of 3.08 kHz, when mounted on a rigid base. To distinguish between the uncoupled bridge modes and those of the island area, additional computations were made with the rigidity of the bridge increased by a factor of 10^4, with its mass unchanged. This raises the rocking mode frequency far beyond any frequency of interest.

Initially, an unscaled, simply-shaped, bass bar was used of thickness 5 mm, maximum height 16 mm at its center and length 12.5 cm running lengthwise under the bass foot of the bridge. Exact dimensions are unimportant, as both the bridge and bass bar properties have been varied over several orders of magnitude to identify and understand their influence on the bridge-island filter action.

The body shell is excited by a 1N sinusoidal force at the top of the bridge in line with the bridge plane, either parallel $F_\parallel$ or perpendicular $F_\perp$ to the supporting ribs. The resulting amplitude and phase of the in-line velocity gives the input admittance A. For brevity, only $F_\parallel$ will be considered here with plots of the admittance in phase with the exciting force. This determines the rate of energy transfer from the vibrating strings to the acoustically radiating body shell modes, apart from small losses in the bridge and island area.

3. ISLAND AREA AND F-HOLES

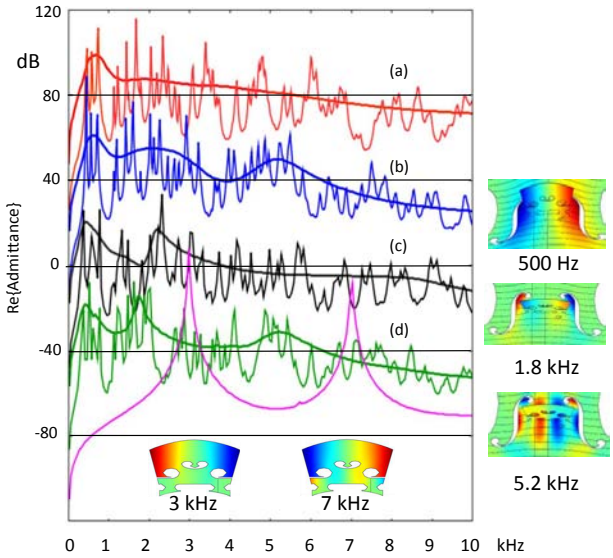


Figure 2. The computed in-line, in-phase, $F_\parallel$ admittance, for a rigid bridge mounted on the empty body shell, initially without bass bar or f-holes, with both lightly ($\eta = 0.01$) and heavily damped ($\eta = 1$) plate modes: (a) for a bridge with rigidity increased by 10^4 and (b) with normal rigidity. Then, after opening the f-holes, (c) with bridge rigidity increased by 10^4 and (d) for normal rigidity. The curves have been shifted by 20 dB to identify the separate characteristics. The mode shapes to the right of the plot are the dominant resonant island area modes, while those below are the in-plane resonances of the bridge resting on a rigid base including the rocking mode at 3 kHz.

Figure 2 illustrates how heavily damping the body shell modes reveals the underlying frequency dependent filtering of the bridge-island area. The measurements shown are for an isotropic top plate before the plates were scaled to account for anisotropy. The upper two sets of curves illustrate the admittances of the empty body shell before the f-holes are cut into the top plate. First in (a), for a rigid bridge with the relatively featureless admittance determined by the energy radiated from under the two feet towards the outer edges of the top plate and in (b), with a normal bridge resulting in the superposition of two rather broad resonances at a significantly lower frequency than those of the bridge alone mounted on a rigid base, which are also illustrated. The rigidly mounted

bridge modes are the previously described rocking mode at $\sim$ 3 kHz and a higher frequency mode at $\sim$ 7 kHz involving bending of the lower areas of the bridge rotating in anti-phase with the top of the bridge.

The lower two curves illustrate the frequency dependence of the admittance at the top of the bridge with f-holes cut into the top plate. Figure (c) illustrates the admittance for a rigid bridge, with two strong broad resonances centered on $\sim$ 500 and 1.8 kHz. These arise from the resonant excitation of the first two intrinsic island area modes illustrated to the right of the plot. The lowest frequency mode is a transverse standing wave across the width of the island area confined approximately to its length. This mode will couple to many anti-symmetric modes of the body shell at both low and high frequencies. The resonance at around 1.8 kHz is equivalent to that of the "flapping" inner-wing mode originally identified in highly simplified f-hole structures by Durup and Jansson[8].

Unsurprisingly, there is no evidence in any of these plots of a strongly peaked resonance at the "rocking" frequency of the rigidly supported bridge. This is because the resonance is spread out over a very wide range of frequencies by its damping and coupling to the plate modes. It is therefore important to understand the influence on the admittance as the coupling between the bridge and island area is varied.

4. BRIDGE-ISLAND COUPLING

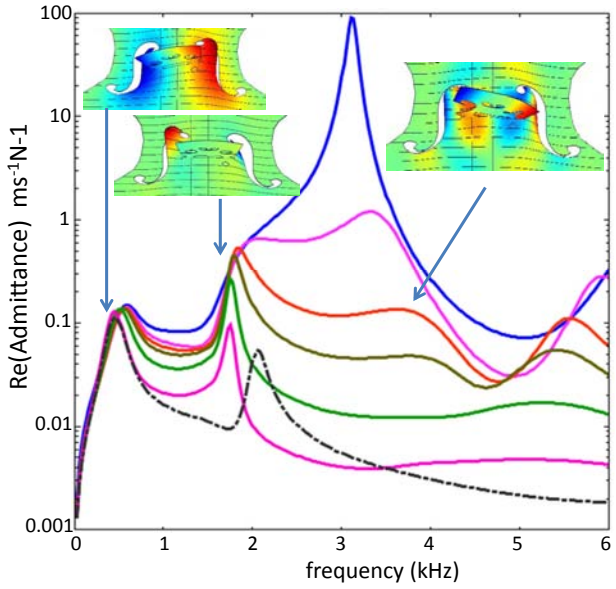


Figure 3. Dependence of the in-line, in-phase, $F_{\parallel}$ admittance of a tapered bridge of mass 2.2 g on the top plate island area of the body shell without bass bar or soundpost, as its mass and density are scaled by factors 0.01, 0.1, 0.3, 0.5, 1 and 2 from their unscaled value leaving the 3 kHz rigidly supported bridge rocking frequency unchanged. The bridge-island mode shapes are illustrated for frequencies just below their resonant peaks. The dash-dotted curve is for a rigid unscaled bridge mass with in-plane rocking frequency 100 times higher.

Consider first the in-line point admittance A_{input} at the top of the bridge, for a central driving force $F_{\perp}e^{i\omega t}$ exciting the *bouncing* mode of the bridge coupled to the localised symmetric modes of the island area. This can be written as

$$A_{input} = \frac{A_{island} + i\omega/m\omega_o^2}{[(1-\omega^2/\omega_o^2) + i\omega m A_{island}]}. \qquad (1)$$

A_{island} is the effective admittance of the local island area modes under the two feet of the bridge, which are strongly damped by energy radiated into the upper and lower bouts of the top plate, while m is the effective mass of the *bouncing* mode of the rigidly supported bridge with $\omega_o = 2\pi f_{bouncing}$. An equivalent expression describes the admittance for $F_{\parallel}$, with appropriate A_{input}, m and *rocking* frequency.

The computed admittance is illustrated in Figure 3 for $F_{\parallel}$ exciting the heavily damped, anti-symmetric, body shell modes (without offset soundpost or bass bar), as the Young's modulus and density of the bridge are simultaneously scaled by the same wide-ranging factor, $bs = 0.01$ to 2, maintaining $f_{rocking}$ constant. The coupling of the bridge to the island area results in markedly different admittance curves from those for the bridge alone [11].

For the very light and flexible bridge (bs=0.01), the admittance at the top of the bridge is dominated by that of the bridge alone, with a strongly peaked resonance just above the rigidly supported bridge rocking frequency at 3.0 kHz. But as the admittance of the rocking bridge approaches that of the island area on which is stands, their modes of vibrations are strongly coupled together resulting in the illustrated veering and splitting of the resonant frequencies and damping of the *normal* modes describing their coupled vibrations.

The dash-dotted curve is for a rigid bridge of unscaled mass, but rocking frequency increased by a factor of 100. This illustrates the unperturbed resonances of the localized island area modes and the decrease in amplitude at higher frequencies from the inertial mass of the rocking bridge.

The lowest frequency peak in the admittance at $\sim$ 400 Hz involves the initially rigid-body twisting vibrations of the island area coupled to the vibrations of the upper and lower bouts of the top plate. This transforms with increasing frequency into a strongly localized transverse bending mode of the island area. The strong, but relatively weakly damped resonance, just below 2 kHz involves the rigid body and rocking vibrations of the bridge vibrating in anti-phase with the "flapping" resonances inner wings of the island area, identified earlier by Durup and Jansson [8] for highly simplified f-hole shapes on a flat rectangular plate. The strongly damped resonance above the weakly coupled "rocking" 3 kHz frequency of the bridge rises to around 4 kHz, with the bridge now coupled to the next highest frequency transverse bending mode of the island area, with the top of the bridge now rocking in anti-phase with the top plate under its feet.

5. BASS BAR AND OFFSET SOUNDPOST

The addition of the bass bar and offset sound post has a very strong influence on the modes of the island area, hence cou-

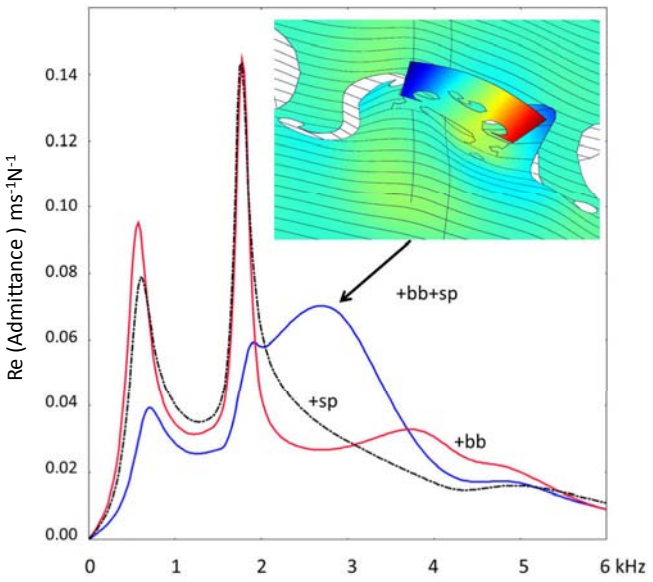


Figure 4. Influence of the bass bar and offset sound post on the BI admittance input filter (+bb+sp), the sound post alone (+sp) with bass bar removed and bass bar alone (+sb) with soundpost removed.

pling to the plate modes and sound of an instrument. Our model enables the contributions of the bass bar and offset soundpost to the overall filtering action of the bridge-island area to be investigated individually, as their coupling strengths are varied independently from zero to typical normal values. Likewise, the influence of arching, anisotropic Young's moduli, plate and bridge masses and rigidity can all be varied in a similar way - to be described in a more detailed account of the present investigation.

Here, we simply consider the influence of the bass bar and soundpost on the *BI* input filter, starting first with them both in place at full coupling strength and then removing them separately, to leave just the sound post or bass bar in place, as illustrated in figure 4. With both bass bar and soundpost in place, the computations reveal a pronounced *BI*-hill feature, with a very broad resonant response similar to the observed *BH* feature when present. The illustrated mode shape is clearly that of a very strongly damped normal mode, with the bridge rocking in anti-phase with the transverse vibrations of the island area. This clearly involves the inner-wings flapping resonance. However, the mode is slightly asymmetric, which is not surprising, as the bass bar and soundpost introduce asymmetry in opposite directions, which will only compensate for each other by accident.

In contrast, when only the bass bar or soundpost is present, the strong mid-frequency resonance disappears, with almost all the plate activity now concentrated in the region of the flapping inner-wing island area regions, with a top plate thickness dependence to be published later just below 2 kHz. Although the resonant frequencies are very similar, the modes are strongly asymmetric, with the vibrational amplitudes on the opposite side of the island always larger than close to the soundpost or bass bar, which tend to create nodal areas around or along them.

Figure 4 suggests that the input filter involving the localized vibrations of both the bridge and island area will strongly influence both the intensity and tonal balance of sound for all members of the violin family. This in turn will be strongly dependent on the offset soundpost and bass bar, "tuned" to balance the asymmetry induced in opposite directions. This explains why bass bar adjustments can play such a crucial role in the quality of sound of an instrument - as widely recognized by luthiers when setting up an instrument to optimize its sound.

Measurements are now required to test the predictions of the above model.

REFERENCES

[1] C. Gough, "Violin plate modes," *J. Acoust. Soc. Am.*, vol. 137, pp. 139–153, 2015.

[2] C. Gough, "A violin shell model:vibrational modes and acoustics," *J. Acoust. Soc. Am.*, vol. 137, pp. 1210–1225, 2015.

[3] W. Reinicke, "Die ubertrangungseigenschaften des streichinstrumentenstegs," 1973.

[4] E. Dunnwald, "Ein erweitertes vefahren zur objectiven bestimmung der klangqualitat von violinen," *Acustica*, vol. 71, pp. 269–276, 1990.

[5] E. Jansson, "Admittance measurements on 25 high quality violins," *Acta Acustica United with Acustica*, vol. 83, pp. 337–341, 1997.

[6] G. Bissinger, "The violin bridge as filter," *J. Acoust. Soc. Am.*, vol. 120, pp. 482–491, 2006.

[7] E. Jansson and B. Niewczyk, "On the acoustics of the volin: bridge or body hill?" *J. Catgut Acout. Soc. Series 2*, vol. 3, pp. 23–27, 1999.

[8] F. Durup and E. Jansson, "The quest of the violin bridge-hill," *Acta Acustica United with Acustica*, vol. 91, pp. 206–213, 2005.

[9] E. Jansson, R. Barczewski, and A. Kabala, "On the violin bridge hill- comparison of experimental testing and fem," *Vibrations in Physical Systems*, vol. 27, pp. 151–160, 2016.

[10] L. Cremer, "The physics of the violin," 1984.

[11] J. Woodhouse, "On the "bridge hill" of the violin," *Acta Acustica United with Acustica*, vol. 91, pp. 155–165, 2005.

[12] M. Smith, "Hybrid modelling of violin acoustics," Master's thesis, Engineering, Cambridge University, 2008.

The Return of the Messiah: Modal Analysis and Bridge Admittance Modeling

Chris Chafe,[1†] *Esteban Maestre,*[2] *Augusto Sarti,*[3] *Antonio Canclini,*[3] *Gary Scavone,*[2] *Julius Smith,*[1] *Fabio Antonacci*[3]

[1]CCRMA / Stanford University
[2]CIRMMT / McGill University
[3]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano
[†]cc@ccrma.stanford.edu

ABSTRACT

Phase one of a larger project to fully model the Messiah violin (1716) involves measurement of bridge admittance and radiativity profiles. The Messiah is believed to be the only Stradivarius in existence in as new state. Currently, the instrument is part of the collection of the Ashmolean Museum in Oxford, England and has been seldom played, in part because of an untreated internal crack in the top plate. In the context of the 300th anniversary of its construction, the violin was made available for one hour to perform a reduced set of minimally-invasive measurements from which to study its vibrational and acoustical properties. We present our first results of computing admittance and radiativity transfer functions from impact hammer, laser doppler vibrometer, and microphone measurements. The overall project goals are two-fold: to create a playable physical model and to "repair" the crack in simulation. Phase two and beyond incorporate a bow model with simplified thermal friction and hair dynamics and performance rendering from gesture capture. The present state of the instrument prohibits bringing its strings up to pitch and consequently measurements were required to be made at reduced tension. A second instrument, the Cremonese Stradivarius (1715), was subjected to the same measurement protocol twice: in a similarly detuned state and at pitch. The Cremonese was made by Antonio Stradivari just one year before the Messiah and its measurements provide a reference for compensation of string tension intended to regain an intact Messiah (by inference).

The paper presents details of admittance models of different resolutions combining modes estimated from admittance and from reference radiativity measurements. Bridge velocity and radiated sound pressure measurements were obtained from orthogonal impact hammer strikes in directions corresponding to the horizontal and vertical directions of the transverse motion of the strings (at the corner of the bass side of the bridge). Modal analysis from admittance measurements uses a technique which explores different resolutions (model orders). Modal analysis of the high-frequency region uses reference radiativity transfer functions. We compare the three sets of admittance models obtained: Messiah (detuned), Cremonese (detuned) and Cremonese (at pitch). A discussion of the next phase describes the 16 hemispherically-distributed microphone measurements also obtained and how these will be used to construct a spherical radiativity model.

On the Role of the Second Membrane of the Tabla

Sylvain Maugeais

Laboratoire Manceau de Mathematiques, Université du Maine
Sylvain.Maugeais@univ-lemans.fr

ABSTRACT

The tabla is an Indian percussion instrument that is remarkable in many ways, in particular because its first partials are harmonic thanks to the black patch at its center. This harmonic quality makes it desirable for the sound to have a very long decay time as it produces a nice sustain which is particularly striking: contrary to the drum set, for which drummers have a tendency to damp their instrument, the tabla player will always look for a good sustain.

When watching a tabla, the second characteristic that catches the eye after the black patch is the presence of an annular shaped membrane breaded at the outer edge of the principal one. The two membranes are not attached to each other except at the rim, and they can easily be separated from each other. In fact, it is customary for tabla players to insert a thread between the two to have a "brighter sound and a better sustain". This assertion suggests that this second membrane, and in particular the disconnection between the two during the vibrations, may be the reason why the tabla have such a long decay time.

In this article, we first present measurements made on a tabla showing the separation of the membranes during the vibration. We then propose a model for the drum head of the tabla consisting of the two membranes together with the kettle, coupled by a nonlinear collision force. A numerical scheme approaching this model is then introduced and calibrated so that it replicates the data obtained by the experiments, thus giving an explanation for the increased sustain.

Finite Element Design and Manufacturing of a Nylon-String Guitar Soundboard from Sandwich-Structured Composites

Negin Abaeian,[1†] *Ulrich Blass,*[2] *Gary Scavone*[1] *Larry Lessard*[2]

[1]Computational Acoustic Modeling Laboratory, Schulich School of Music, McGill University
[2]Structures and Composite Materials Laboratory, Dept. of Mech. Engineering, McGill University
[†]negin.abaeian@mail.mcgill.ca

ABSTRACT

Composite materials have many potential advantages over wood in the construction of string instruments, including less sensitivity to humidity and temperature, better durability, as well as uniformity of mechanical properties that can lead to more predictable vibrational behaviour. While composite instruments have been commercially available for several decades, the acoustic properties of these alternative materials are still relatively unknown to many luthiers. The aim of this project is to use Finite Element Methods (FEM) to design and manufacture the soundboard of a nylon-string guitar from sandwich-structured composites, and compare its vibrational behaviour with that of a reference wooden one. In order to minimize the effects of other factors in this replacement, the soundboard is treated independently from the rest of the body, and the pattern of the braces behind the soundboard remains unchanged. First, following an experimental modal analysis on the reference wooden soundboard, the wooden soundboard is modelled in Abaqus, where the properties of the model are adjusted to match the real soundboard. The design process then starts by designing the composite top plate to have the same bending stiffness and density as the reference wooden top plate. The two top plates are then simulated using FEM under free boundary conditions, where the thickness and material properties of the core and the composite layers are varied so that the designed top plate matches the simulated wooden one in mode-frequencies of up to 1000 Hz and in the first few mode-shapes. The same procedure is followed after adding the braces to both models, until the mode-frequencies and the mode-shapes of the simulated composite soundboard are reasonably similar to those of the simulated wooden soundboard. Finally, to choose between the most promising designs, the composite soundboard designs are compared with the simulated wooden soundboard under hinged-boundary conditions, i.e. fixed-translation and free-rotation. The designed soundboard is then manufactured, followed by an experimental modal analysis on the composite soundboard, from which the mode-shapes, mode-frequencies and the damping values are estimated and compared with those of the wooden soundboard.

An Exploration of Guitar Neck Admittance Measurements Taken at Different String Stopping Locations

Mark Rau,[1†] *Esteban Maestre,*[2] *Julius O. Smith,*[1] *Gary Scavone*[2]

[1]Center for Computer Research in Music and Acoustics (CCRMA), Stanford University
[2]Computational Acoustic Modeling Laboratory (CAML), McGill University
[†]mrau@ccrma.stanford.edu

ABSTRACT

String instrument synthesis models have typically ignored the influence of the string stopping position when modeling the string boundary conditions at the neck. This simplification neglects the differences between frequency-dependent wave reflection phenomena taking place at the string boundary when it is stopped at different positions along the neck. Driving-point admittance measurements were taken at the bridge of an electric guitar, and also at various stopping positions along the neck: open string, 2nd, 4th, 6th, 8th, and 10th fret. We explore the acquired data and compare boundary reflectance functions as computed for the bridge and the various stopping positions. As a test case, we compare measured and theoretical string decay times for one particular note.

1. INTRODUCTION

Stringed instruments, such as guitars, involve the transfer of vibrational energy from a set of strings into a body, which serves as a more efficient radiator of acoustic energy. The energy of vibrating guitar strings is transferred to the instrument at the string termination points, the bridge and the stopping location along the neck. The mechanical "input admittance" or "driving point mobility" is defined in the frequency domain as the velocity of the structure divided by the input force driving the structure at a given location and in a given direction. The admittance shows the ability of the guitar body to be displaced as a function of frequency, providing insight about the vibrational modes and string decay time [1, 2]. Input admittance is often measured at the bridge on a hollow instrument as this will give a good estimate of the instrument's vibrational characteristics which will result in sound [2]. Admittance is the inverse of impedance which can be used to calculate the string reflection coefficient

$$r = \frac{Z_b - Z_c}{Z_b + Z_c} \tag{1}$$

where Z_c and Z_b respectively represent the characteristic string and bridge impedance functions at the boundary. This amplitude reflection coefficient is the ratio of an incoming wave amplitude to the reflected wave amplitude. The string termination along the rigid neck will not result in a significant amount of sound compared to the termination at the mobile bridge, so this termination is often ignored when considering the instrument's radiated sound [3]. However, this termination will affect the stopping location string reflectance as there are vibrational modes of the neck, effecting the string decay rates [3].

The purpose of this work is to use input admittance measurements to study the boundary conditions at the bridge and stopping location along the neck of an electric guitar. We perform a first exploration of how much energy is lost at the string ends by computing the corresponding reflectances, with the aim of foreseeing the importance of including a model for the neck losses to attain accurate string decay times in a future synthesis model.

The rest of the paper is organized as follows. Section 2 gives details on the experimental setup used to measure neck and bridge admittances as well as string decay times. In Section 3 we present the admittance measurements as well as an example of computed reflectance functions. A simplified method to predict string decay times is presented and compared to an example measured string decay. Section 4 discusses the measurements, missing factors in the decay time prediction, and proposes future steps.

2. MEASUREMENTS

Admittance measurements were made on a guitar to observe the body and neck vibrations. String decay measurements were then made to explore the relationship between the neck vibrations and the string decay. A hollow body electric guitar with a floating bridge and two electromagnetic pickups was measured. The measurements were taken in a semi-anechoic room with the guitar hung from the ceiling by the tuning pegs, and lightly resting against foam for stability. The setup for the admittance measurements is shown in Figure 1. The guitar was strung with flatwound strings which were tuned to pitch.

2.1. Admittance

To measure the admittance of the guitar, the instrument was struck with a force sensing hammer (B&K 8203) while a laser Doppler vibrometer (Polytec PDV 100) was used to measure the velocity. The hammer strike and laser measurement locations are made as close as possible to obtain a driving point admittance. The signals were pre-amplified to the appropriate level and recorded through a National Instruments data acquisition card with a sample rate of 44.4k Hz and 16-bit

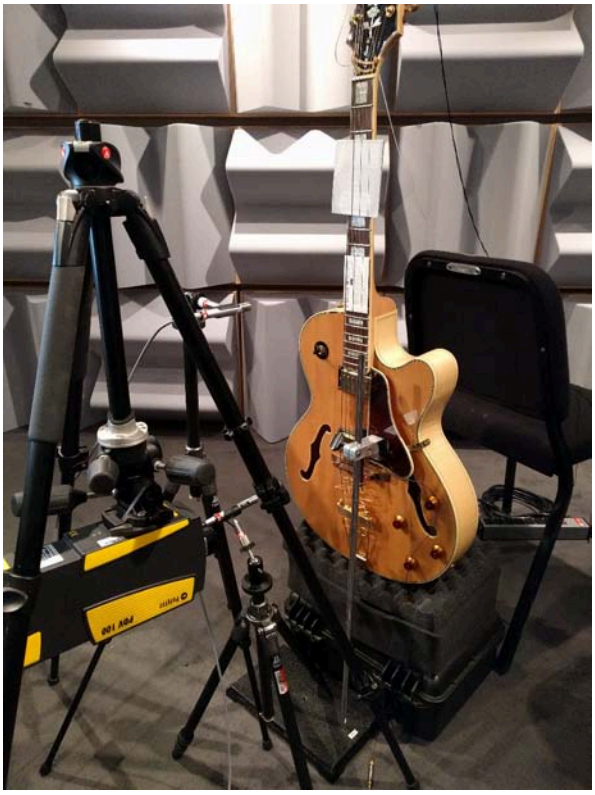


Figure 1: Experimental setup for admittance measurements.

sample resolution. All strings were damped using a combination of thick card stock and elastic bands. Each measurement was triggered as soon as the impact hammer came in contact with the guitar, and lasted for 2 seconds. For each location, 5 consecutive measurements were made and averaged, observing the coherence between the measurements and discarding any erroneous measurements. The bridge admittance was measured in the vertical (normal to the guitar top plate) and horizontal (along the direction of the bridge) directions. The impacts were made near the low E string; this position was chosen because the low E is the thickest string, thus providing strong coupling to the body. Admittance measurements were also made on the neck at the open position, 2nd, 4th, 6th, 8th, and 10th frets. The open position admittance was measured with a driving point on the nut in the vertical and horizontal directions. The admittance was measured at each fret position on the fret for the vertical direction, and near the fret on the side of the neck for the horizontal direction. All admittance measurements were also recorded with calibrated microphones to later calculate the instruments radiativity for synthesis purposes.

2.2. String Decay

In order to justify the stopping location admittance method and evaluate later synthesis, string decay measurements were made. A "wire break" method was used as it is more reproducible than a human pluck. A copper wire is looped around and pulled across the string such that it abruptly snaps at a repeatable level of stress, imparting an approximate step function (being more flat at low frequencies and rolling off at high frequencies above 10k Hz) in a known direction. It is difficult to know the exact force the wire exerts when it snaps but it is consistent between tests and the measurements are only concerned with the relative amplitude of vibrations [4]. Copper wire of gauge 40 on the American wire gauge scale (0.0799 mm) was chosen as it produced a string displacement similar to that of an average guitar plectrum. The signals were pre-amplified to the appropriate level and recorded through a National Instruments data acquisition card with a sample rate of 44.4k Hz and 16-bit sample resolution. Each measurement was triggered as soon as the wire broke, and lasted for 10 seconds. The output from the electromagnetic pickup nearest to the neck was recorded with the guitar's tone and volume controls at a maximum value, imparting no additional filtering. The vibrational axis of the strings will change with time [5], but guitar pickups are most sensitive in the direction perpendicular to the top plate of the guitar [6], so the measurements will approximate the vertical transverse string motion. The string plucks were also recorded with two calibrated microphones to later compare the radiated sound to that of the strings alone, as measured by the electromagnetic pickups. The strings were plucked at a constant distance away from the bridge at an angle of 45° between the vertical and horizontal axes. In order to measure the fretted notes, a "capo" (a device that clamps all of the strings to a particular fret) was placed on the corresponding fret. All strings that were not being measured were damped using heavy card stock and elastic bands.

Once recorded, the string decays measured through the electromagnetic pickup were analyzed. A short-time Fourier transform was used to determine the amplitudes of the first five harmonics at each time frame. The peaks of the first five harmonics were determined in a general range based on their theoretical frequency. Since the string frequencies will vary slightly over the duration of the decay, the frequency and amplitudes were calculated more carefully by using parabolic interpolation. The harmonics were then analyzed to determine the time each took to decay by 15dB assuming an exponential decay.

3. PRELIMINARY OBSERVATIONS

Bridge admittance measurements in the vertical and horizontal directions are shown in the right columns of Figures 2 and 3. The left columns of Figures 2 and 3 shows the vertical and horizontal admittance measurements made at the stopping locations corresponding to the open position, 4th fret, and 8th fret. When observing Figures 2 and 3, it can be seen that there are three modes of high amplitude at 85, 175, and 332 Hz which are likely caused by the neck bending modes. The amplitudes of the modes at 175 and 332 Hz decrease at positions along the neck approaching the bridge, suggesting that they are near the nodes of the bending modes.

Figure 4 shows the magnitude responses of the bridge and open position admittance in the vertical and horizontal directions as well as string decay plots which show the time for the first 5 harmonics to decay by 15dB in the vertical direction

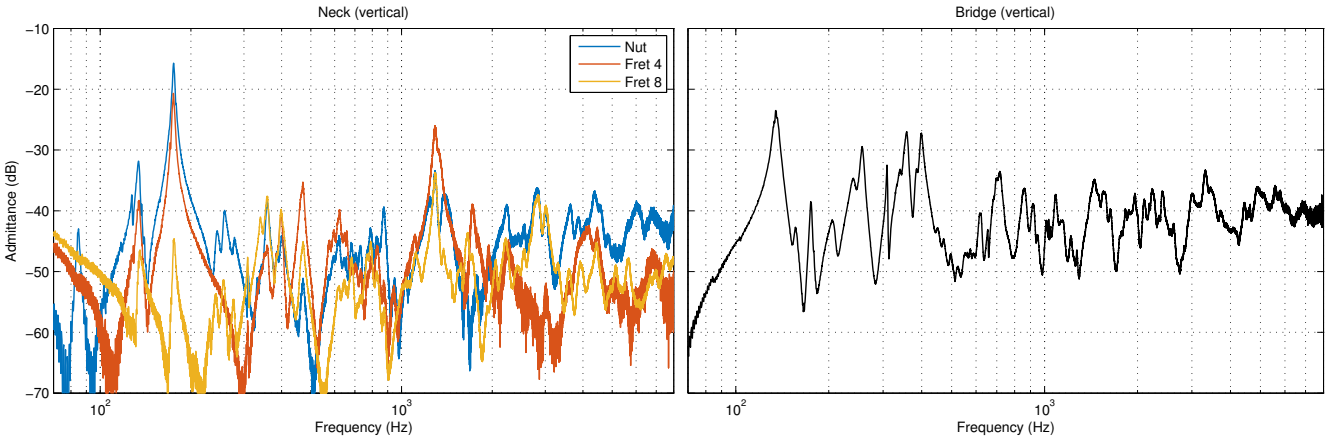


Figure 2: Magnitude response of the vertical admittance measurements made at the bridge, open position, 4th fret, and 8th fret.

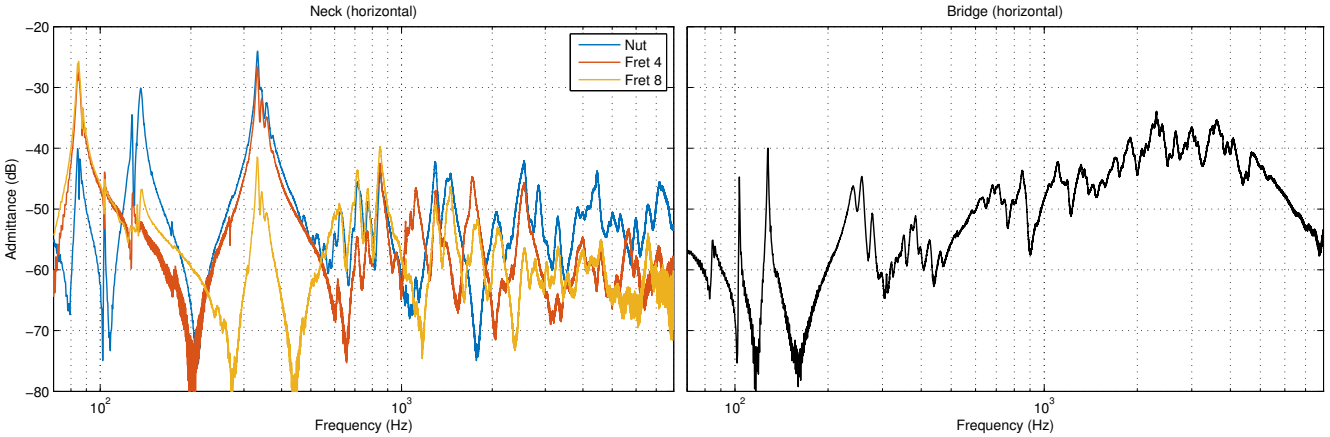


Figure 3: Magnitude response of the horizontal admittance measurements made at the bridge, open position, 4th fret, and 8th fret.

as measured through the electromagnetic pickup. Each circle colour corresponds to the first five harmonics of an individual string. This serves as a way to observe the overall decay rate of the strings with respect to the measured admittance.

The reflection coefficient functions for the bridge, open position, 4th fret, and 8th fret of the low E string were calculated as described in equation 1, and the amplitude is shown in Figure 5. The string characteristic impedance was calculated from manufacturer data to be to 0.8702 N m s^{-1} [7]. As it can be observed, losses at the neck should not be neglected: for some frequencies they are comparable to those at the bridge.

Assuming that energy is only lost through the bridge admittance, neck admittance, and propagation losses, the change in amplitude ratio G over one period can be calculated as

$$G = \frac{A}{A_0} = r_B \cdot r_N \cdot r_S \tag{2}$$

where A_0 is the original amplitude, A is the final amplitude after one period, r_B is the bridge reflection coefficient, r_N is the stopping location reflection coefficient, r_S is the propagation loss coefficient for one period of vibration. The reflection coefficients r_B and r_N represent the losses through one reflection from each end respectively. This can be arranged to calculate the decay time t_L for the transverse string motion to decay by a certain amplitude as follows

$$t_L = \frac{\gamma_L}{f_0 20 \log_{10}(r_B r_N r_S)} \tag{3}$$

where f_0 is the fundamental frequency in Hertz, and γ_L is the decay amplitude ratio in dB. This method can be used to predict the vertical and horizontal transverse string decay by using the respective vertical and horizontal reflection coefficients.

Table 1 shows the measured decays as well as the computed vertical transverse decays of the first 5 harmonics of a pluck of the open low E string. The amount of decay, γ_L was set to -15 dB and the string air propagation loss r_S is assumed to be equal to 1 at all frequencies as an initial simplification. As expected from the imposed simplifications, we observe significant differences.

Harmonic	1	2	3	4	5
Measured Decay (s)	5.45	4.49	0.86	0.87	1.12
Predicted Decay (s)	13.29	0.68	0.24	0.33	0.33

Table 1: Measured and predicted decay times for the first 5 harmonics of the open low E string to decay by 15dB.

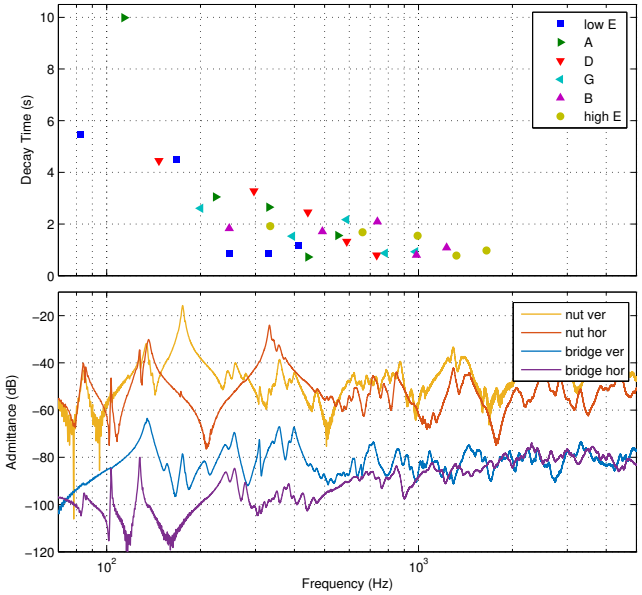


Figure 4: Magnitude responses (lower) of the bridge and open position admittance in the vertical and horizontal directions. The bridge admittance measurements are scaled by -40dB for clarity. String decay times (upper plot) for the first 5 harmonics to decay by 15dB.

4. DISCUSSION AND FUTURE WORK

Observing the admittance measurements, it is clear that the neck modes absorb a significant amount of energy and will affect the string decays. The stopping location reflection coefficient will change as the note is stopped at different positions along the string, suggesting that multiple stopping location transmission coefficients should be used for accurate string decay synthesis across the instrument.

A practical method for measuring the propagation losses in the string from simple measurements is not known, and it would be useful if these could be learned from string decay and boundary admittance measurements. We hoped that we would be able to approximate these frequency-dependent losses from the above analysis, but some of the predicted decay times are shorter than the measured decays. This suggests that the boundary conditions are more complicated than stateless reflectances, and that horizontal-vertical transverse motion coupling, and string-body and string-neck coupling should not be ignored as was done in this preliminary simplification. Future work will look into predicting the amount of energy transferred back into the string from the motion of the body and neck terminations, if modeled as resonant systems.

Along these lines, further steps will be to perform mode fitting on the admittance measurements made at the bridge as well as the neck stopping locations. This will result in efficient digital filters which are different at each end of the string termination, providing more accurate string decay rates for a digital waveguide model using the technique described in [8]. Multiple stopping location filters will be used with a different filter corresponding to each fretted note position. Synthesis examples will be computed, both including the stopping location filters, and without them. The synthesis will be compared to the measured decay rates to further validate the effectiveness of this method at correctly synthesizing the string decay.

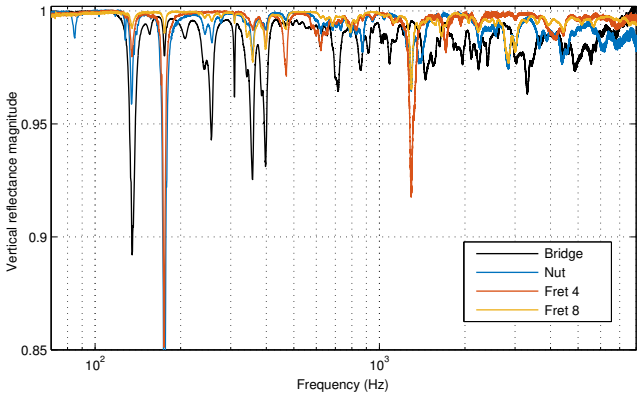


Figure 5: Low-E string reflectance magnitudes for the vertical direction, corresponding to the bridge and different stopping locations.

REFERENCES

[1] H. Mansour, V. Fréour, C. Saitis, and G. P. Scavone, "Post-Classification of Nominally Identical Steel String Guitars Using Bridge Admittances," *Acta Acustica united with Acustica*, vol. 101, pp. 1–14, 2015.

[2] J. Woodhouse and R. S. Langley, "Interpreting the input admittance of violins and guitars," *Acta Acustica united with Acustica*, vol. 98, no. 4, pp. 611–628, 2012.

[3] H. Fleischer, "Mechanical Vibrations of Electric Guitars," *Acta Acustica united with Acustica*, vol. 84, no. January, pp. 758–765, 1998.

[4] J. Woodhouse, "Plucked guitar transients: Comparison of measurements and synthesis," *Acta Acustica united with Acustica*, vol. 90, no. 5, pp. 945–965, 2004.

[5] T. D. Rossing, *The Science of String Instruments*. New York: Springer, 2010.

[6] A. Paté, "Lutherie de la Guitare Électrique Solid Body: Aspects Méchaniques et Perceptifs," Ph.D. dissertation, L'Universite Pierre Et Marie Curie, 2014.

[7] D'Addario, "Catalog Supplement/String Tension Specifications: A complete technical reference for fretted instrument string tensions," 1999.

[8] E. Maestre, G. P. Scavone, and J. O. Smith, "Digital modeling of string instrument bridge reflectance and body radiativity for sound synthesis by digital waveguides," *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA 2015*, 2015.

Effects on Temporal Change of Vibrational Properties on Guitar String by Stringing and Plucking

Kota Otsuka,† *Koichi Mizutani, Naoto Wakatsuki, Tadashi Ebihara*

University of Tsukuba
†otsuka@aclab.esys.tsukuba.ac.jp

ABSTRACT

It is well known that the tone of string instruments gradually changes if the string is used for a long time with stringing and plucking, however, the factors and the mechanisms causing this have not been enough elucidated. The purpose of this study is to provide an insight into effects on temporal change of vibrational properties by stringing and plucking. String vibration was measured by the optical probe that consists of LEDs (Light-Emitting Diodes) and photodiode arrays to investigate effects of stringing and plucking on vibrational properties. Using the optical method, we can measure not only sound but also string vibration without effects of body resonance. We measured temporal changes on vibrational properties and tensions in two condition: (1) just remain strung without plucking and (2) continue being plucked. The string is tuned as 110 Hz, which is a same value as an actual guitar, for several times on both conditions. As a result, the tension required to make frequency of fundamental mode 110 Hz gradually descend for 24 hours on the condition that string remains strung. It is considered that cause of this phenomenon is decrease of line density of the string. Line density ranges as time progresses by relaxation, creep, and rust. However, tension is adjusted for several times to keep fundamental frequency constant. This is a relaxation and stretch process which is similar to actual playing condition. Moreover, experimental period is too short to rust string. Thus, relaxation could be a dominant factor of decrease of line density. In all experiments, a 5th string of EFT16(D'Addario) is used. It is made of plain steel with a phosphor bronze wound. Length of the string is set as 600 mm in the measurement system. On this condition, it seemed that line density become 99.34% by calculation from the value of tension decrease. It means the string got 4 mm longer in this system. On the other hand, descend of tension finished in an hour and then tension keeps a constant value on the condition that string continues being plucked. It is considered that it takes shorter time to make string a stable condition if a string continues being plucking. As touching vibrational properties, both fundamental and the second modes attenuate most rapidly immediately after stringing. Attenuation of them become gradual as time progresses. It is considered that rapid attenuation property is unique in a brand-new string.

Study of Magnetic Pickups for Electric Guitars

Léo Guadagnin,[†] *Bertrand Lihoreau, Pierrick Lotton, Emmanuel Brasseur*

Laboratoire d'Acoustique de l'Université du Maine
[†]leo.guadagnin@univ-lemans.fr

ABSTRACT

Electric guitars use sensors to convert the string vibration into an electrical signal. The most common type of sensors is the magnetic pickup based on the principle of magnetic induction. These pickups are generally modeled using an equivalent Thevenin's generator whose electrical impedance is measured but whose voltage generator is not discussed. The aim of this work is to study Thevenin's generator of a single coil guitar pickup.

An experimental bench is carried out allowing to measure the response of the pickup to a harmonic motion of the guitar string for a single polarization. Several magnetic pickups are built in order to easily control each geometric and magnetic parameters. A parametric study is led and highlights the influence of several parameters on the induced voltage (frequency, material and section of the string, cinematic of the string, etc.).

A modeling of the output voltage is first proposed where the magnetic permanent field created by the magnet is supposed to be uniform. The approximation is rough but required to derive an analytical expression of the induced voltage where the influence of all parameters (coil, magnet, string) are explicit. Predictions of the modeling are compared to the experimental results: although the model succeeds to accurately predict the influence of some parameters, the assumption of uniform magnet field prevents the model to be fully satisfactory.

Consequently, in order to investigate the limitations due to the assumption of uniform magnet field , the model is compared to the numerical model proposed by Horton and Moore which takes into account the exact topology of the magnetic field created by a permanent magnet. The comparison between both models allows to clarify Horton's model and notably to establish the constant of proportionality between the magnetic field generated by the magnet and the one generated by the magnetized string. The comparison of the theoretical results with the experimental data leads to a discussion about the modeling of the string magnetization suggested by Horton's model.

Kinematics Analyses of Strings and Upper Limbs During Harpist's Performance

Jean-Loïc Le Carrou,[1†] *Delphine Chadefaux*[2]

[1]Sorbonne Universités, UPMC Univ Paris 06, CNRS, LAM / Institut ∂'Alembert, Paris, France
[2]Aix Marseille Univ, CNRS, ISM, Inst Movement Sci, Marseille, France
[†]jean-loic.le_carrou@upmc.fr

ABSTRACT

While learning their instrument, musicians develop a significant gestural expertise. This allows them to perform with a great precision the movement required to produce the desired sound. For plucked strings instruments, the instrumentalist brings the string to a particular state defining its future free oscillations. Recent works on the concert harp showed that, for chords or arpeggio sequences, initial conditions of the vibrating string are characteristic of each musician, repeatable and dependent on the playing technique. For expanding these results to a more realistic playing context, we developed a specific experimental setup, which allows to measure simultaneously the plucking parameters for six strings of the instrument with opto-switch sensors properly calibrated and a specific signal processing, the kinematics of the harpist's upper limbs with inertial sensors, and the instrument's sound. Using ten harpists playing an imposed piece, results confirmed the instrumentalists' capability to reproduce the string trajectory at the plucking point and underlined the adaptation of this trajectory to the musical context, to the string played, and to the finger used. Finally, results indicated that the upper limbs kinematics tends to compensate the postural imbalance by adjusting the flexural movement magnitude of each shoulder and the right elbow and to participate to the sound production by facilitating the instrumentalist musical interpretation.

Towards Prediction of Violin Timbre from Vibrational Measurements

Massimiliano Zanoni, Fabio Antonacci, Augusto Sarti

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano - Piazza Leonardo da Vinci 32 - 20133 Milano, Italy

massimiliano.zanoni@polimi.it

ABSTRACT

This contribution investigates on the acoustics of violin, and more specifically on the relationship existing between vibrational impulse responses and the timbre of the instrument. With respect to previous publications on this topic, we tackle the problem using a feature-based approach. More specifically, we aim at finding the correlation between the features extracted from accelerometric measurements of the bridge mobility and from audio recordings of a prescribed set of performances. Results demonstrate that features describing to the global shape of the spectrum are strongly related. On these descriptors we also show the possibility of predicting the features of audio recordings from the vibrational ones. Experimental data are based on a set of 25 modern violins.

1. INTRODUCTION

The perceived quality of violins is mainly determined by its timbre. Many studies have been proposed in the literature[1, 2]. In this paper we investigate on the relationship existing between vibrational data and audio recordings of performances of violins. The Rayleigh integral [3] suggests that the emitted sound is related to the vibrational velocity field on top and back plates through an integral equation, which, however, cannot be used to predict deterministically the timbre from the vibrational analysis, due to the complexity of the physical and vibrational behavior of the instrument. Instead, works in the literature aim at finding a relationship between timbre, perception and vibrational response of the instrument using tools from machine learning. In an early study in [4] the author demonstrated that peaks in the bridge admittance between 1 kHz and 3 kHz impact on the perception of the timbre. In [5] the author observes that, despite the fact that the bridge response is characterized by relevant peaks and troughs, the perceived quality of the instrument does not change much from note to note played by the musician, suggesting that the timbral quality of the instrument must be attributed to the global shape of the spectrum rather than to its details. In [6] authors relate vibrational and acoustics data of violins by comparing the harmonics of the long-term spectrum of glissando performances with the bridge mobility transfer function, which interestingly turn out to exhibit similar features.

The characterization of the timbre, which is mostly subjective, using objective descriptors extracted from the sound, has gained interest in the research community [7], and is spreading also in Musical Acoustics [2]. In this work we adopt a feature-based approach in order to investigate on the relation between vibrational and audio recordings. As far as vibrational analysis is concerned, we measured the bridge mobility function [8]. Audio data, instead, consist in the recording of excerpts of songs, open strings, and musical scales. Many low-level timbral features have been extracted from both sound recordings and vibrational impulse responses. The Pearson correlation is then analyzed to assess which descriptors of the sound recordings can be predicted from the vibrational data. Results demonstrate that descriptors related to the overall shape of the spectrum exhibit a good degree of correlation, thus confirming the results in [5] and [6]. On these descriptors, we also implemented a curve fitting algorithm, and a good fitting accuracy is found. We believe that this paper sheds light on the possibility for violin-makers to fine tune some parameters of the timbre of the instrument at the different stages of the making process (e.g. additional thinning of plate).

2. DATA ACQUISITION

Our study is based on 25 violins: 16 from the "Triennale Violin-making Competition" held in Cremona, Italy, in September 2015 and 9 from the collection of the school of lutherie.

Vibrational acquisition: in order to conduct the vibration analysis we acquired the bridge mobility, or bridge admittance, of each violin. The instrument was suspended with nylon cables to have a free boundary condition measurement and decoupling the flexible vibration modes from the system rigid motion [9]. The bridge bass side was excited using a small impact hammer (5 g mass) and the response at the other side was captured with a piezoelectric uni-axial accelerometer (0.5 g mass). At least five hammer hits were averaged for each test to reduce the measurement noise [9]. A detail of the excitation and measurement positions is shown in Fig. 1, where also the little foam stripes used to damp the strings on both sides of the bridge are visible.

The H_1 estimator was adopted to derive the Frequency Response Function (FRF) data:

$$H_{oi}(\omega) = \frac{G_{oi}(\omega)}{G_{ii}(\omega)}, \tag{1}$$

where $G_{oi}(\omega)$ is the crossspectral density function between the input driving force i and the output measured vibration o, whereas $G_{ii}(\omega)$ is the autospectral density function of the input force. Fig. 2 shows, as an example, the bridge mobility of three tested violins. The coherence function γ, calculated as

$$\gamma(\omega) = \frac{|G_{oi}(\omega)|^2}{G_{oo}(\omega)G_{ii}(\omega)}, \tag{2}$$

Figure 1. Experimental set-up.

indicates the degree of correlation between input and output and defines the frequency range over which the FRF data can be considered reliable. For the presented data, the upper limit was set to be 2 kHz, whereas the sampling frequency was set to 25.6 kHz.

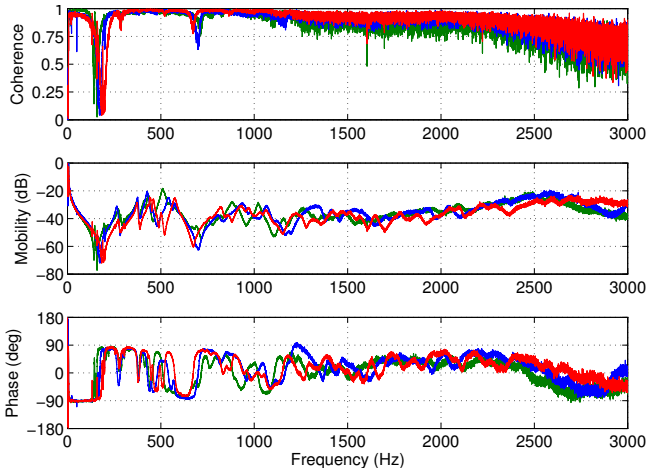


Figure 2. Examples of experimental FRFs for three tested violins; from the top: coherence function, mobility magnitude and phase.

Since some descriptors are extracted from the time domain representation of the signal, we also measured the impulse response $h_{oi}(t)$ as $h_{oi}(t) = \mathcal{F}^{-1}\{H_{oi}(\omega)\}$, where the symbol $\mathcal{F}^{-1}$ denotes the inverse Discrete Fourier Transform.

Audio recordings: instruments were played by a unique professional musician in an acoustic dry room, in order to be independent on the acoustic environment. The musician was asked to play the same pieces of music for all the instruments. The pieces have been properly selected in order to best emphasize the different aspects of the sound of violins: open strings (G,D,A,E), major scale, six musical pieces in different styles (*deciso*, *arpeggiato*, *legato*, *sforzato*, *pizzicato*, *appasionato*, *etereo*). The notes were played at nearly constant intensity, with the bow moving slowly both upward and downward, every time waiting for the sound to completely release before moving on to the next note. For the recordings we used the Beyerdymanic MM1 omnidirectional microphone placed at about 40cm from the instrument. Audio was recorded with an Aphex 188 preamplifier and an Apogee 32 A/D converter at a sampling frequency of 48 kHz and at a bit rate of 24 bits.

3. FEATURE-BASED CHARACTERIZATION

For each instrument, we extracted the set of audio descriptors from audio recordings and vibrational data using a windowing technique. In order to be coherent with the frequency range where the FRF is reliable (see coherence in Fig. 2) we low-pass filtered audio recordings with a cutoff frequency of 2 kHz.

In this study we extend the feature set proposed in [7] by including other descriptors. More specifically, the additional features have been found to be effective in music [10] and sound analysis [11, 2]. A complete description of the additional descriptors is in [10, 12, 13]. The total number of features resulted to be 40. In the following we shortly describe each category.

Noisiness measure: In order to provide a measure of the noisiness of sound we use Zero Crossing Rate (ZCR). In addition, we compute the Spectral Flatness and Spectral Crest to measure the similarity between the magnitude of the spectrum under analysis and a flat one. **Statistical moments:** in order to compactly represent the shape of the spectrum, we compute Spectral Centroid, Spectral Spread, Spectral Average Deviation, Spectral Skewness, and Spectral Kurtosis. **Spectral shape:** besides the statistical moments, also other descriptors provide a compact representation of the shape of the spectrum. We include Spectral Rolloff, Spectral Smoothness, Spectral Slope, Spectral Irregularity K [13] and Spectral Irregularity J [14]. **Summary of the spectrum:** in order to provide a generic summary of the spectral components, MFCCs (20 coefficients) and Spectral Contrast (7 descriptors) are also considered in this study. MFCCs are based on Mel-Frequency scale, based on the auditory model. They are obtained as the coefficients of the discrete cosine transform (DCT) applied to a reduced Power Spectrum, which is derived as the log-energy of the spectrum. Spectral Contrast [12] is defined as the dynamics of spectral peaks and valleys separated into different frequency sub-bands of interest. Samples from each sub-band are sorted in descending order and peaks and valleys of the k-th sub-band are computed as

$$\mathrm{P}_k = \log\left\{\frac{1}{\alpha N_k}\sum_{j=1}^{\alpha N_k}|S_{k,j}|\right\}, \tag{3}$$

$$\mathrm{V}_k = \log\left\{\frac{1}{\alpha N_k}\sum_{j=1}^{\alpha N_k}|S_{k,N_k-j+1}|\right\}. \tag{4}$$

Finally, the Spectral Contrast is computed as

$$\mathrm{SC}_k = \mathrm{P}_k - \mathrm{V}_k, \tag{5}$$

where α is a corrective factor used in order to ensure the steadiness of the feature [12], $S_{k,j}$ is the j-th sample of the DFT within the k-th sub-band and N_k is total number of samples in the k-th sub-band.

4. RESULTS

Correlation analysis: fig. 3 shows the Pearson correlation index between vibration- and audio-based features. As for

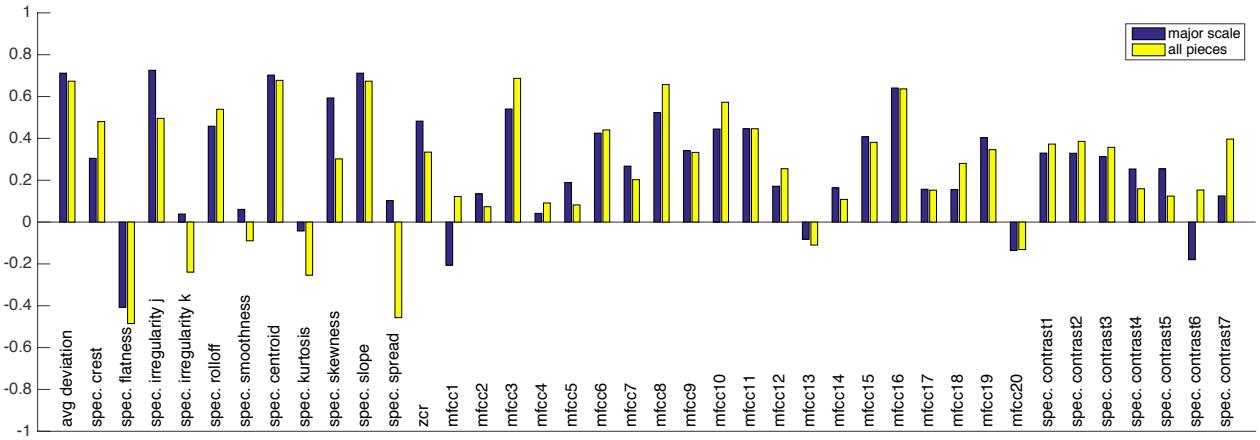


Figure 3. Pearson Correlation Index on the major scale and on all pieces performances.

the audio features, we performed the analysis for each of the performances described in Section 2. However, for reasons of space, only the correlation of descriptors extracted from the major scale and from all performances are shown.

Among the 40 features, several show a good correlation index. These are Average Deviation, Spectral Irregularity J, Spectral Centroid, Spectral Slope, and third, eighth and sixteenth coefficients of MFCC. It is worth noticing that these features convey global information about the spectrum, thus confirming the results in [5] and [6]. It is worth noticing that, despite a slight decrease, the set of most relevant descriptors remains unaltered when the analysis is extended to all the music pieces. **Curve Fitting Analysis**: in order to provide further insight on the correlation analysis, we also tested the accuracy of different curve fitting models on the features listed above. The adopted fitting models are linear, quadratic, cubic, and quartic polynomials and the exponential. In preliminary results we noticed that the use of higher-order polynomials does not bring relevant advantages in terms of fitting accuracy, thus we limited our analysis to the fourth order. Data have been previously normalized between 0 and 1. In order to be more robust to the presence of outliers in the dataset, we performed outlier identification based on the residuals of the fitting curves. Let $O = \{(x_i, y_i) : 1 \leq i \leq N\}$ be the dataset, being N the cardinality, and y and x the audio- and vibration-based features. Let $f(x)$ be the fitting curve and $f(x_i)$ be the corresponding point in the model for x_i, the residual r_i is computed as $r_i = x_i - f(x_i)$. Points in the distribution are considered outliers if $|r_i| > \mu_{|r_i|} + 1.5\sigma_{|r_i|}$, where $\mu_{|r_i|}$ and $\sigma_{|r_i|}$ are the mean and the standard deviation of the absolute values of r_i respectively.

In order to evaluate the quality of the fitting we have adopted two metrics, which are the ordinary Coefficient of Determination (R^2) and the Root Mean Square Error ($RMSE$), given by

$$R^2 = 1 - \frac{\sum_{i=1}^{N} |y_i - f(x_i)|^2}{\sum_{i=1}^{N} |y_i - \overline{y}|^2}, \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} |y_i - f(x_i)|^2}{N}}, \quad (7)$$

where $\overline{y}$ is the mean value of the considered vibration-based descriptor. R^2 ranges from 0 (fitting is not possible) to 1 (data can be fitted without any error), while $RMSE$ has the lower bound of 0, meaning a perfect fitting. We conducted the analysis both removing and keeping outliers. For reasons of space, we present here only the fitting with features extracted on the major scale performance. Fig. 4 shows the R^2 for the seven selected descriptors and for the different fitting models. The best accuracy is observed for the quartic curve. This is true for all the descriptors, but especially for MFCC 8 and MFCC 3. We repeated the experiment when data that have been recognized as outliers have been removed from the dataset. Results are reported in fig. 5. It is possible to observe a general increase of the R^2 metric.

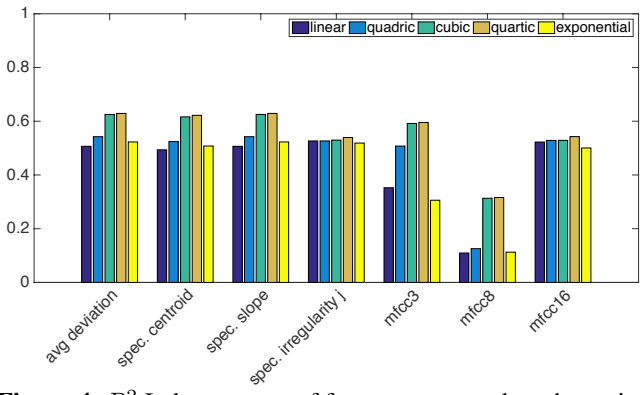


Figure 4. R^2 Index on a set of features extracted on the major scale performance.

Finally, Fig. 6 and Fig. 7 show the Root Mean Square Error with and without outliers. Even if the two metrics are not completely independent (see eq. (6) and (7)), it is worth noting that the RMSE and the R^2 metrics suggest that the best prediction is achievable for the MFCC16 and Average Deviation, respectively.

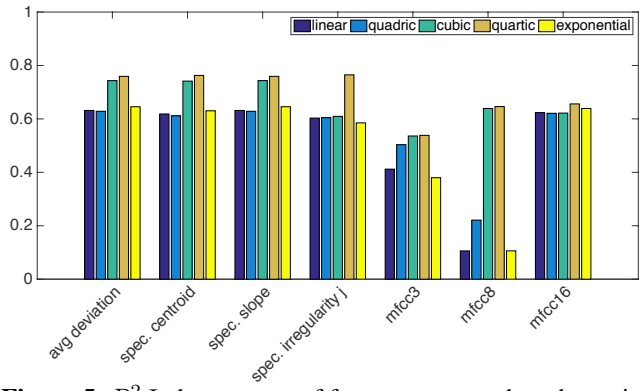


Figure 5. R^2 Index on a set of features extracted on the major scale performance without outliers.

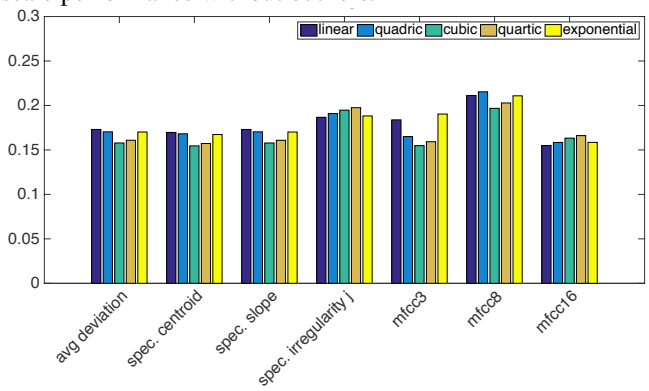


Figure 6. Root Mean Square Error (RMSE) on a set of features extracted on the major scale performance.

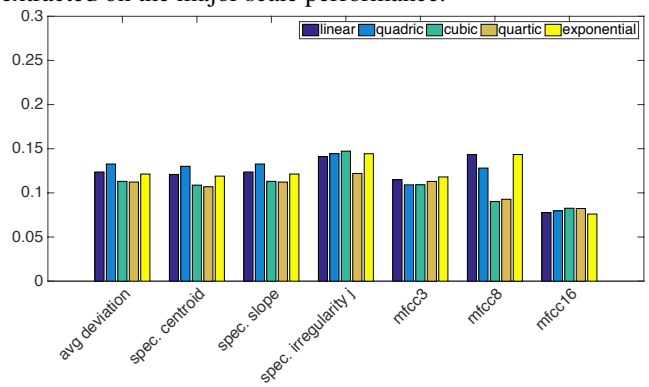


Figure 7. Root Mean Square Error (RMSE) on a set of features extracted on the major scale performance without outliers.

5. CONCLUSIONS AND FUTURE WORKS

In this contribution we investigated on the relationship existing between vibroacoustics and audio recordings of violins. Using a feature-based approach we demonstrated that the global shape of vibrational and audio spectra exhibit a good degree of correlation. This result paves the way to the possibility of predicting the timbre of the instrument also from accelerometric measurements. Some results in this direction are included in this paper.

REFERENCES

[1] J. Štěpánek, Z. Otčenášek, and A. Melka, "Comparison of five perceptual timbre spaces of violin tones of different pitches," *The Journal of the Acoustical Society of America*, vol. 105, no. 2, pp. 1330–1330, 1999.

[2] M. Zanoni, F. Setragno, F. Antonacci, A. Sarti, G. Fazekas, and M. Sandland, "Training-based semantic descriptors modeling for violin quality sound characterization," in *in Procideengs of the Audio Engineering Society Convention (AES)*, Warsaw, Poland, 2015.

[3] E. G. Williams, *Fourier Acoustics: sound radiation and nearfield acoustical holography*. Academic Press, 1999.

[4] M. Hacklinger, "Violin timbre and bridge frequency response," *Acta Acustica united with Acustica*, vol. 39, no. 5, pp. 323–330, 1978.

[5] C. Gough, "The violin: Chladni patterns, plates, shells and sounds," *The European Physical Journal Special Topics*, vol. 145, no. 1, pp. 77–101, 2007.

[6] B. Elie, F. Gautier, and B. David, "Acoustic signature of violins based on bridge transfer mobility measurements," *The Journal of the Acoustical Society of America*, vol. 136, no. 3, pp. 1385–1393, 2014.

[7] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The timbre toolbox: Extracting audio descriptors from musical signals," *The Journal of the Acoustical Society of America*, vol. 130, no. 5, pp. 2902–2916, 16 November 2011.

[8] J. Woodhouse, "The acoustics of the violin: a review," *Reports on Progress in Physics*, vol. 77, no. 11, 2014.

[9] R. Corradi, A. Liberatore, S. Miccoli, F. Antonacci, A. Canclini, A. Sarti, and M. Zanoni, "A multidisciplinary approach to the characterization of bowed string instruments: the musical acoustics lab in cremona," in *Proceedings of the 22nd International Congress on Sound and Vibration (ICSV22)*, 2015, pp. 2633–2640.

[10] H.-G. Kim, N. Moreau, and T. Sikora, *MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval*. John Wiley & Sons, 2005.

[11] Y. Kong, A. Mullangi, J. Marozeau, and M. Epstein, "Temporal and spectral cues for musical timbre perception in electric hearing," *Journal of Speech Language and Hearing*, vol. 54, no. 3, Jun 2011.

[12] D. N. Jiang, L. Lu, H. J. Zhang, J. H. Tao, , and L. H. C. in Proc. of IEEE International Conference Multimedia Expo, "Music type classification by spectral contrast features," in *Proc. of IEEE Internationa Conference Multimedia Expo*, 2002, p. 113116.

[13] J. Krimphoff, S. McAdams, and S. Winsberg, "Caracterisation du timbre des sons complexes. ii analyses acoustiques et quantification psychophysique." *Journal de Physique IV, Colloque C5*, vol. 4, 1994.

[14] K. Jensen, "Timbre models of musical sounds," Ph.D. dissertation, University of Copenhagen, 1999.

Evaluation of a Miniature Accelerometer with a Laser Doppler Vibrometer to Study Vibrations at the Neck of a Violin in Realistic Playing Scenarios

Matthias Flückiger,† Tobias Grosshauser, Gerhard Tröster

ETH Zurich, Electronics Laboratory
†mflueckiger@ife.ee.ethz.ch

ABSTRACT

This paper evaluates a miniature accelerometer to measure vibrations at the points where the violinist is in contact with the violin. The violinist perceives these vibrations as vibrotactile feedback, which plays an important role for the quality perception and the precise control of the instrument. The performance of the accelerometer is compared to a laser Doppler vibrometer in measurements with a loudspeaker and a violin. It is shown that the displacement calculated from the accelerometer signal has a resolution of 0.1 μm in the frequency range from 50 Hz to 1 kHz and that the added mass of the accelerometer does not influence the vibrations of the instrument. Finally, the capability of the approach is demonstrated in a violin neck vibration measurement, where horizontal and vertical vibrations were recorded with two accelerometers in a realistic playing scenario.

1. INTRODUCTION

Playing a musical instrument is a multi-modal task, where the musician controls the instrument based on sensory feedback. Vibrations of the instrument's body are perceived by the musician as vibrotactile feedback through the contact points with the instrument. The relevant frequency range of human vibration sensation extends from 40 Hz to 1 kHz; one of the most sensitive sites is the fingertip that can detect displacements as small as 0.1 μm [1].

In violin playing vibrotactile feedback is important for the quality perception, for the "feel" and for the precise articulation of the instrument [2, 3, 4, 5, 6]. The contact points between the violinist and the instrument are the thumb of the left hand holding the neck, the fingers of the left hand pressing the strings, the fingers of the right hand holding the bow, the chin, and the shoulder holding the violin body [3].

Marshall [2] analyzed the vibration behavior of a violin and found that the neck participates strongly in the lower order modes [2]. He suggested that these low order resonances must be related to the coveted "feel" of the great violins. Later, Askenfelt and Jansson [3] presented an extensive series of measurements at the contact points of string instruments and compared the results to the vibration sensation thresholds for the finger reported by Verrillo [1]. For the violin, the measurements were made with an accelerometer and for excitation the lowest note G3 (194 Hz) was played on the violin at fortissimo level. In a comparison experiment with four violins they found that there is no clear relationship, such as higher neck vibration levels indicate higher quality.

Extending on the work of Askenfelt and Jansson, Wollman et al. [5] investigated the vibrotactile feedback of the left hand in violin playing. He showed that the perception thresholds for the left hand in violin playing are lower than those reported by Verrillo, due to the greater contact area of fingers and thumb holding the neck, as well as the influence of the applied pressure. In the same study, based on the expert judgment of an experienced professional violinist, several violins were categorized into "vibrating" and "non-vibrating" instruments. The differences of the neck vibrations were investigated and they found that the "vibrating" violins vibrate significantly more in the frequency range from 600 Hz to 1 kHz. They suggest that these differences in vibration behavior might lead to the perceived characteristics of the individual instrument.

However, quality classification of violins based on vibrational properties remains limited and might depend on how perceptual criteria are individually evaluated by violinists during assessment, as pointed out by Saitis et al. [7].

To measure vibrations during violin playing and to further explore the role of vibrotactile feedback, we are developing a measurement method to measure vibrations at the contact points between violinist and the instrument during realistic playing scenarios. In this paper we evaluate a miniature accelerometer with a laser Doppler vibrometer and demonstrate the potential of the approach by presenting a measurement of the horizontal and vertical neck vibrations during playing of a chromatic scale.

2. SENSOR EVALUATION

To evaluate the miniature accelerometer for vibration displacement measurements of a violin, we compared it to a laser Doppler vibrometer (LDV). We focus on vibration displacement because perception data is reported as displacement values (e.g. see [1]). In the first experiment two examples of the accelerometer are compared to the LDV by measuring the motion of a loudspeaker driver membrane. Thereafter, the acceleration sensor is evaluated with vibration measurements on the top plate of a violin, where the violin was held in a fixed position. The violin used for the experiments is a professional quality instrument that was made by O. Spidlen in 1946 (Prague).

2.1. Equipment

We also tried piezo disc elements and piezo thinfilm sensors to measure vibrations at the contact points of a violin. The miniature accelerometer we evaluate hereafter was most promising with regard to resolution, mass, and size. The miniature accelerometer is a Knowles BU-21771-000 piezo-based miniature 1-axes accelerometer designed for vibration transduction. The size of the accelerometer is 7.9 x 5.5 x 4 mm and its weight is 0.28 g. According to the data sheet[1] the sensor has a sensitivity of -45 $\pm$ 4.5 dB relative to 1 V/g at 1 kHz. The frequency response of the sensor is flat up to 2 kHz. The accelerometer is connected to a preamplifier that we designed with a gain of 21 dB. The battery driven preamplifier also supplies the sensor. The accelerometer is attached to the vibrating surface with double-sided tape.

The performance of the miniature accelerometer was compared to an LDV consisting of a single-point laser head, a controller, and a velocity decoder[2]. The precise alignment of the laser beam is crucial for LDV measurements, because only the projection of the vibration velocity vector from the target surface onto the laser beam is captured. To increase reflection and to improve signal-to-noise ratio (SNR), we used pieces of retroreflective tape with a size of 5 $\times$ 5 mm that were attached to the target surface. The measurement range of $\pm$ 50 mm/s with a resolution of 0.6 μm/s proved to be a good compromise for the measurements presented in this paper.

A computer with an external soundcard[3] recorded the signals from the velocity decoder and the miniature accelerometer synchronously. The soundcard has a sampling rate of 48 kHz and a resolution of 16 bit. The measurements were analyzed with computer software we specifically made for this purpose.

2.2. Evaluation with a loudspeaker

In the first experiment, the accelerometer and the LDV measured the motion of the 8-inch woofer membrane of a 3-way hi-fi loudspeaker[4]. The setup is shown in Fig. 1.

Two configurations were measured. In the first configuration the accelerometer was attached to the loudspeaker membrane and the LDV was measuring the motion of the accelerometer. In the second configuration only the LDV was measuring the vibration of the membrane to assess the influence of the added mass of the accelerometer on the loudspeaker membrane. The loudspeaker was fed with a swept sine signal from 20 Hz to 20 kHz and the frequency response of the loudspeaker was estimated with the method presented in [8]. The results are shown in Fig. 2. For displacement calculation, the accelerometer signal was integrated twice and the vibrometer signal was integrated once[5]. To remove offset and low frequency noise, the signal was filtered using a fourth order high pass filter with a cutoff frequency of 20 Hz after each integration.

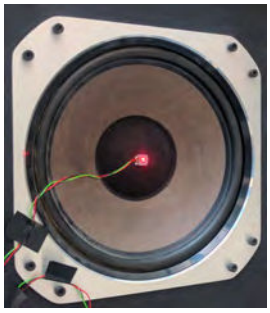
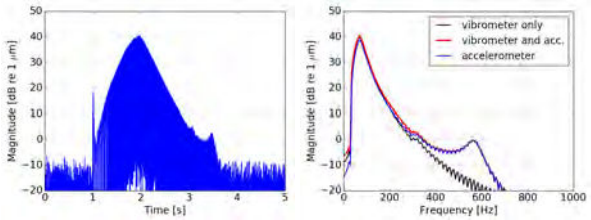


Figure 1. Setup for the sensor evaluation on the 8-inch woofer of a hi-fi loudspeaker. The accelerometer and the vibrometer are measuring the motion at the center of the membrane. Operation frequency range of the woofer is 20 Hz to 650 Hz.

Despite a level deviation of 1.6 dB the frequency response of the accelerometer and the LDV match perfectly. A deviation was expected from the manufacturing tolerance of the accelerometer sensitivity. We repeated the same measurement for a second example of the miniature accelerometer, where the level difference was 2.7 dB. The noise level after displacement calculation for the accelerometer is around -10 dB (see Fig. 2). This implies a resolution of approximately 0.3 μm. Note that due to the integration, the cumulated noise level in the displacement time values depends on the cutoff frequency of the high pass filter. By increasing the cutoff frequency to 50 Hz, the noise level was pushed below -20 dB resulting in a resolution of 0.1 μm.



(a) Displacement time signal calculated from the accelerometer. (b) Extracted frequency responses from the swept sine measurements.

Figure 2. Results of the woofer membrane measurement. A swept sine from 20 Hz to 20 kHz was played back through the loudspeaker.

Comparing the vibrometer measurement with and without the accelerometer in Fig. 2, the responses start to diverge above 400 Hz. The presence of the accelerometer influences the motion of the loudspeaker membrane, which in this case results in an additional resonance around 600 Hz. In the following, the cutoff frequency of the high pass filter applied after each integration was set to 50 Hz and the measured level differences of the accelerometers were taken into account. To restrict the comparison to the relevant frequency range for vibration sensation and to include a fair margin, the signals were also low pass filtered with a cutoff frequency of 2 kHz.

[1] `http://www.knowles.com/download/file?p=BU-21771-000.pdf`

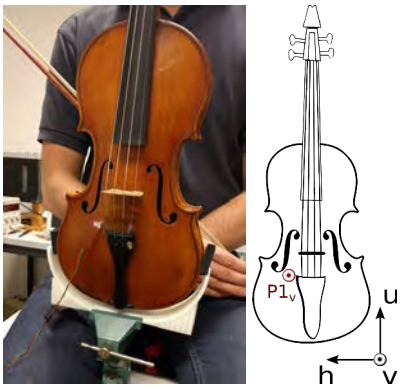
[2] Polytec OFV-303 head, OFV-3001 controller, OVD-01 decoder

[3] Sound Devices USBPre 2 audio interface

[4] Philips 22AH587, Manual: `http://docs.mfbfreaks.com/um/User_Manual_22AH587.pdf`

[5] For integration the scipy.integrate.cumtrapz function from the SciPy library 0.14.0 was used

2.3. Evaluation with a violin in a fixed position



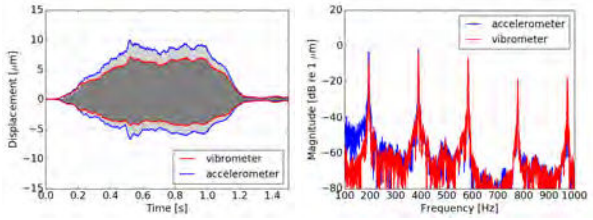
(a) (b)

Figure 3. Setup picture (a) and measurement position (b) for the evaluation of the accelerometer. Open strings were bowed from behind. The index v refers to vibrations measured along the vertical axis.

For an application based evaluation we created a setup to measure the vibrations of a violin during playing. The violin was held in a fixed upright position by the use of a 3D printed stand, such that the vibrations could be measured with the LDV and the violinist could bow open strings with downstroke from behind the instrument. The violinist was instructed to bow the strings in mezzo-forte (mf), although this setup is obviously not a realistic playing scenario. The purpose of this experiment is to exclude deviations due to the small added mass of the accelerometer, as well as to compare the sensitivity and the frequency range of the acceleration sensor and the LDV.

Measurement setup and measurement position $\mathrm{P1_v}$ are illustrated in Fig. 3. The measurement position was chosen because the top plate might be most sensitive to added mass. The distance between the laser measurement point and the accelerometer was roughly 1 cm.

The vibration displacement signals for the open G string (G3, 196 Hz) are shown in Fig. 4.



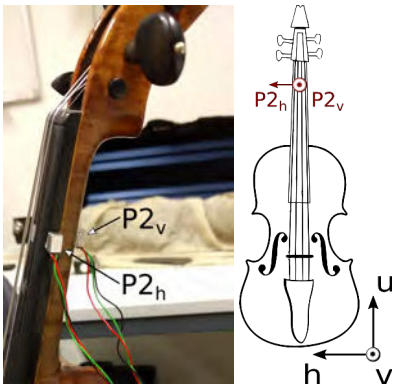
(a) Time signals with highlighted envelope (b) Frequency spectra

Figure 4. Comparison of laser Doppler vibrometer and accelerometer measurement on position $\mathrm{P1_v}$. The time signal shows the measured displacement when the lowest note G3 was played by the violinst. The frequency plot shows the corresponding spectrum.

The level difference of the time signals in Fig. 4 is around 12 % with regard to the displacement signal calculated from the accelerometer. The major deviations in the frequency spectrum are that the level of the LDV is 6 dB lower for the first partial and 2 dB lower for the second partial. The deviations arise from the laser beam misalignment in this case, but apart from the mentioned deviations, the frequency spectra closely match each other. We also calculated the Pearson correlation coefficient of the displacement signals. The correlation coefficient is above 0.96. Therefore it seems safe to conclude, that the accelerometer captures the vibrations up to 1 kHz. The accelerometer mass of 0.28 g attached to the violin body and the cables do not influence the vibration.

Summarizing, the accelerometer can measure vibration displacement with a resolution of 0.1 μm in the frequency range from 50 Hz to 1 kHz. We were surprised that the accelerometer performs comparable to the vibrometer for the desired application. In contrast to the LDV, the accelerometer is suited for measurements at the violin under realistic playing conditions, which is demonstrated in the following section.

3. VIBRATION MEASUREMENT DURING A REALISTIC PLAYING SCENARIO



(a) (b)

Figure 5. Setup picture (a) and measurement positions (b) to measure the vibrations at the neck. In a realistic playing scenario, a chormatic scale was played in mf. The index h or v refers to vibrations measured along the horizontal or vertical axis. Position P2 corresponds to the finger pressing position of a minor third on the fingerboard.

To unveil the potential of the mobile measurement setup, we measured the vibrations of the violin neck during a playing scenario. We attached two BU-21771-000 accelerometers to the neck of the violin as shown in Fig. 5. The accelerometer at position $\mathrm{P2_v}$ measured the vertical vibrations of the neck on the opposite side of the fingerboard and the accelerometer at position $\mathrm{P2_h}$ was attached to the side of the fingerboard to measure the vibrations in horizontal direction. The location alongside the fingerboard that corresponds to the sensor positions is the finger pressing position of a minor third. In contrast to before, the violinist was now performing in a real-

istic playing scenario and played a chromatic scale in mf from G3 (196 Hz) to A#5 (932 Hz) with downstroke. The sensors minimally disturb the violinist, because the violinist usually does not touch these positions of the neck with the fingers.

The calculated displacement time signals from the accelerometer signals are shown in Fig. 6. For the majority of the played notes the levels of the horizontal vibrations are higher than those of the vertical vibrations. For example, the level of the lowest note (G3, 196 Hz) is more than two times (6 dB) higher in the horizontal than in the vertical direction. This is consistent with the result from [3], where the horizontal displacement levels for this note were 5-10 dB higher than the vertical ones. The signal levels in Fig. 6 vary quite substantially depending on the played note, although the chromatic scale was played with the same dynamics. The vibrations generated by notes played on the D string have the lowest levels. For the horizontal vibrations the notes with the highest displacement levels are B3 (247 Hz) and A#4 (466 Hz) and for the vertical vibrations C4 (262 Hz).

These maximum values must be related to the body modes of the violin. According to [2, 3] these are typically found around 200 Hz and 300 Hz. Also the first acoustic mode (or Helmholtz mode) is expected around 300 Hz [2], which is close to the fundamental frequency of the open D string (D4, 293 Hz). However, the relation of the modes and the results in Fig. 6 has to be clarified in further measurements.

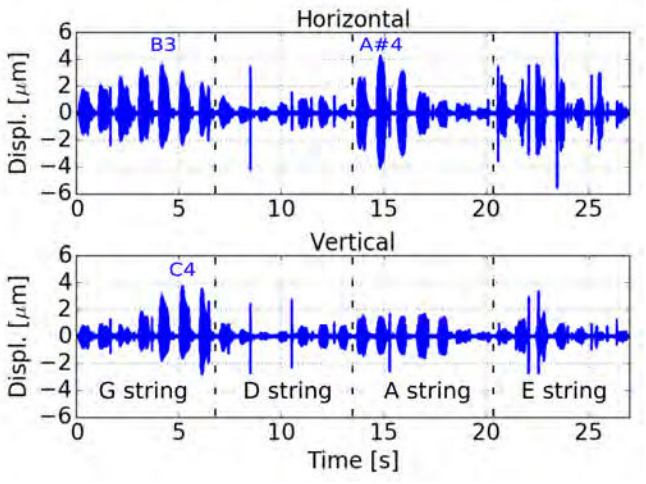


Figure 6. Time signals of the horizontal and vertical vibrations measured at positions $P2_h$ and $P2_v$, when the violinist was playing a chromatic scale from G3 (193 Hz) to A#5 (932 Hz) in mf. The levels vary substantially depending on the note. Notes, where the vibration levels are highest are indicated.

4. CONCLUSION

We demonstrated that the evaluated miniature accelerometer can be used to measure the vibrations of a violin in the relevant frequency range of human vibration sensation. With the vibrometer as reference, we found that the acceleration sensor measures with an accuracy of 0.1 μm from 50 Hz to 1 kHz. Further, it was showcased that multiple sensors can be used in a realistic playing scenario to record vibrations simultaneously.

Future work will be headed towards applying the method to study the vibrations on the fingerboard and also to measure the influence of finger touch, along with further investigations on the vibration behavior of the instrument. The method will be applied also to study vibrations of other string instruments and instruments that were augmented with actuators.

ACKNOWLEDGMENT

This research is pursued as part of the "Musician's behavior based on multi-modal real-time feedback" project, funded by the Swiss National Science Foundation (SNSF). Additional thanks go to Prof. Jürg Dual and Mario Weder from IMES of ETH Zurich for providing the vibrometer, as well as to our colleague Alberto Calatroni for suggesting the evaluated accelerometer.

REFERENCES

[1] R. T. Verrillo, "Vibration sensation in humans," *Music Percept.*, vol. 9, no. 3, pp. 281–302, 1992.

[2] K. D. Marshall, "Modal analysis of a violin," *J. Acoust. Soc. Am.*, vol. 77, no. 2, pp. 695–709, 1985.

[3] A. Askenfelt and E. V. Jansson, "On vibration sensation and finger touch in stringed instrument playing," *Music Percept.*, vol. 9, no. 3, pp. 311–349, 1992.

[4] C. Chafe, "Tactile audio feedback," in *Proc. of the Int. Comp. Music Conf.* Int. Comp. Music Assoc., 1993, pp. 76–76.

[5] I. Wollman, C. Fritz, and J. Frelat, "On the characterization of vibrotactile feedback in violinists' left hand: a case study," *Acta Acust. unit. Acust.*, vol. 101, no. 2, pp. 360–368, 2015.

[6] M. E. Altinsoy, S. Merchel, and S. Tilsch, "Perceptual evaluation of violin vibrations and audio-tactile interaction," in *Proc. of Meetings on Acoustics ICA2013*, vol. 19, no. 1. Acoust. Soc. Am., 2013.

[7] C. Saitis, B. L. Giordano, C. Fritz, and G. P. Scavone, "Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players," *J. Acoust. Soc. Am.*, vol. 132, no. 6, pp. 4002–4012, 2012.

[8] M. Rébillat, R. Hennequin, E. Corteel, and B. F. Katz, "Identification of cascade of hammerstein models for the description of nonlinearities in vibrating devices," *J. Sound Vib.*, vol. 330, no. 5, pp. 1018–1038, 2011.

Comparison of Tones and Physical Properties among Cello Strings of Different Materials

Masaki Mohri,[†] *Naoto Wakatsuki, Koichi Mizutani*

University of Tsukuba
[†]mohri@aclab.esys.tsukuba.ac.jp

ABSTRACT

In the performance of bowed-stringed instruments, it is well-known among players that tones are different depending on the brand of strings. It takes a lot of time and efforts to find strings suitable for their demands. However, it has not been clarified what kind of physical properties of strings affect the change of tones. By clarifying this, makers will be able to make a string which has characteristic of their wants. The difference of tones were observed in any conditions of bowing and to figure out the mechanism; which physical properties of cello strings lead to tone differences. In this report, the sounds of a cello with strings made of different materials were recorded, and physical properties of strings such as stiffness and frictions between strings and a bow were measured. In one recording, string was bowed a hundred stroke. Furthermore, the experiments were carried out on several conditions such as speeds of a bow and the load between the string and the bow. Besides, a tension of the bow and applied amount of a rosin were kept constant. For each recording, the sounds of different strings were analyzed statistically from view point of frequency response. As a result, the magnitudes of harmonic components differ over 600 Hz despite conditions of bowing. Many physical factors which lead to such difference of tones were considered. We especially focused on stiffness and maximum static frictional force between strings and the bow. To measure stiffness of strings, the relationships between tension and strain were observed. A stiff string generates high frequency components over 600 Hz. On the other hand, the property of friction was evaluated by the proportion of force to displacement when a string began to slip on the bow. The difference of friction was not observed among strings. The effect of maximum friction to the tone was still not clear. Thus, it was experimentally shown that the strings made of different materials produce apparently different tones. Stiffness might be one of causes to change the tone. However, other properties such as mechanical loss factor should be considered.

Exploring Simulation-Based Playability Metrics of 9 Perceptually Evaluated Violins

Quim Llimona,[1†] *Charalampos Saitis,*[2] *Esteban Maestre,*[1] *Gary Scavone*[1]

[1]Computational Acoustic Modeling Laboratory, McGill University
[2]Audio Communication Group, Technical University Berlin
[†]ql@lemonzi.me

ABSTRACT

The bridge admittance is a commonly used acoustic characterization of bowed string instrument bodies. While most studies focus on its frequency response, we analyze it in conjunction with the feedback loop created during the vibration of the string using numerical simulation. A measured admittance can be approximated in a compact form as a group of second-order resonant filters. These filters can then be incorporated into the numerical model of the bow-string interaction and of the string vibration in order to investigate how different violin bodies behave when they interact with the stick-slip motion coming from the string. As a case study we used bridge driving-point admittance measurements of 9 violins, which had previously been rated on ease of playing and dynamic range by experienced musicians during a perceptual test. Digital simulation allowed the collection of hundreds of hours worth of playback, from which we derived playability-related metrics, such as the empirical limits of the Schelleng diagram. The metrics were found to mostly differ in the minimum bow force, which was also shown to be in agreement with previously published theoretical limits that take the bridge admittance into account. Preliminary results from comparing different metrics to violinists' evaluations indicated a possible link between the simulated minimum bow force and ease of playing ratings.

Feature-Based Timbral Characterization of Historical and Modern Violins

Francesco Setragno,[1†] *Massimiliano Zanoni,*[1] *Fabio Antonacci*[1] *Augusto Sarti*[1]

[1]Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB) - Politecnico di Milano - Piazza Leonardo da Vinci, 32, 20133 Milano

[†]francesco.setragno@polimi.it

ABSTRACT

Violin timbre is a very complex case of study. The sound properties that distinguish an historical violin from a modern one are still not clear. The purpose of this study is to understand what are these properties, by means of feature-based analysis. We extract audio features related to timbre and we exploit feature selection techniques in order to investigate what are the most characterizing ones. We compare different feature selection algorithms and we illustrate how we applied their outcome to a classification task with historical and modern instruments. Results show that the classification performance improves when using the selected features.

1. INTRODUCTION

Violin has been a subject of research for decades. This complex instrument has been examined from several points of view (vibro-acoustic, chemical, microscopic etc.). Among them, timbre is certainly very important.

One interesting aspect in the study of violin sound quality is the timbral characterization of historical instruments. Although the ability of discerning a historical violin from a modern one is not easy for human listeners [1], it is know that many factors influence the sound of historical instruments making them, in most cases, distinguishable from modern ones. Well known techniques in the literature for musical instrument sound characterization rely on feature-based analysis, that allows to extract a set of descriptors from the audio waveform able to capture specific aspects of the sound.

Feature-based methodologies have already been considered for the analysis of the violin tone. In [2] Łukasik et al. analyzed the dissimilarity factors of the timbre of various master violins, i.e. the features that enable to automatically identify the individual instruments in a pool. Feature-based timbral analysis that relies on feature-based representations is also applied to musical instrument identification. In [3] the authors use a set of multiscale Mel-Frequency Cepstral Coefficients for the pairwise discrimination of musical instruments. In [4] the author used Long Term Cepstral Coefficients to characterize the subtleties of violin sound, while in [5] several harmonic features are extracted from a collection of 53 violin recordings. The study also showed that the four strings exhibit different values for the same feature. Taking into account this result, in our work, we study each string separately.

Feature-based analysis is also used as the base for semantic description as introduced in [6], where a set of bipolar descriptors from natural language are modeled by means of large sets of acoustic cues extracted by 28 historical and modern violins.

In this paper we investigate the timbral properties of violins that help in discerning historical and modern violins by means of feature-based analysis. That is, if we look the problem by the information analysis view point, find the subset of features that best separate the two classes. Several feature selection algorithms are present in the literature. For this is paper we considered two feature selection and two feature ranking techniques. We then study the consistency of their results in order to provide with a final set of selected features. The selection and the understanding of discriminative proprieties, allows to implement applications able to automatically discern the two classes (historical and modern) of instruments by means of machine learning techniques. We used the feature as input for a binary classifier. In the results we show the performance of the classifier before and after the application of the feature selection procedure. We recorded 22 violins: 9 average modern violins and 13 historical violins from the collection of the Violin Museum in Cremona.

2. FEATURE ANALYSIS AND SELECTION

In this section we describe how the feature analysis and selection has been carried out.

Low-level features are objective descriptors that can be extracted from a signal by means of some signal processing techniques. Each feature captures a specific property of the audio signal. For this reason they are suitable for our purposes.

For this study, we extract low-level features that are well known in the Music Information Retrieval (MIR) field. In particular, we use the four statistical moments of the audio spectrum (centroid, spread, skewness and kurtosis), several features related to the shape of the spectrum (slope, rolloff, sharpness, smoothness, crest, flatness, irregularity), the spectral average deviation, the spectral entropy, several features related to harmonics distribution (tristimulus 1, 2 and 3, odd-even ratio) and the zero crossing rate (ZCR). Finally, we also extract two vectorial features: the Mel-Frequency Cepstral Coefficients (MFCC) and the Spectral Contrast (SP), which have been successfully used in the context of timbre analysis [7, 8]. The total number of features used is 108 (by considering each element of the MFCC and the Spectral Contrast as one feature). We refer to [9, 10, 11, 12, 13, 14] for a detailed explanation of these features. We use the Matlab MIR Toolbox [11] for extracting the features. Feature algorithms that were not provided by the toolbox are computed using ad-hoc

developed software.

Each audio file is divided into 40 ms overlapping frames. Each frame is multiplied by a Hanning window and processed in order to extract the mentioned features. Therefore, for each frame we obtain a vector of length M, where M is the number of features.

To provide a more complete timbre characterization as possible, we extract a large set of audio features. This preliminary characterization is the input of the feature selection algorithm. Several algorithms are present in the literature based on specific selection measures. Tough many are effective, their results can be slightly different: even starting from the same input, the outputs of the different methods can be different. For this reason, in this study, we apply five algorithms, three for feature-selection and two for feature ranking, provided by Python's *sklearn* framework [15].

Feature selection algorithms output a subset of features according to specific criteria: SelectKBest and SelectPercentile select the K features and a given percentage of the features with the highest score, respectively, according to an ANOVA test. This test is useful because a feature which varies more from one group to another, rather than within the same group, represents a statistically relevant factor in discerning the two groups of instruments. False Positive Rate (FPR) selects the p-values below a threshold α based on a FPR test. Feature ranking algorithms produce a rank of features: the first we use is based on a Forest-Of-Trees, illustrated in [16], and the second is called Relieff [17].

We then compare the output of the selection methods in order to assess the consistence of their output. Features resulted to be selected by more algorithms are good candidates to compose the final set of features.

3. CLASSIFICATION

In order to prove that the selected features are actually relevant in the timbral characterization of historical violins, we setup a classifier to automatically discern modern and historical violins.

A classifier takes a vector of features as input and predicts the class of an instrument, according to a trained model. We use the well known Support Vector Classifier [18], that divides the feature spaces in such a way that the gap between each class is as wide as possible. We use a 3 order polynomial kernel in order to make the boundary nonlinear.

As reported in the next section, we record different sections ranging from open strings to musical excerpts. The classification task is performed separately for each section, since we want to test it for different type of content. In order to build the dataset, for each section the short frames are grouped into 5 seconds segments. Each segment contains the mean and the variance of every low-level feature and, for the frames belonging to the training set, a label which indicates if it is an historical or modern instrument (ground truth). In the training phase the extracted features are fed to the training algorithm, together with the correct label for each instrument (ground truth). In the test phase, the features are extracted from new audio segments and the trained model predicts the class of each segment.

The dataset is split as follows: 70% for the training set and 30% for the test set. The classification error is computed as the percentage of misclassified samples and this procedure is repeated five times. The classifier is trained with both the whole feature set and with the reduced one, that includes only the features selected by the feature selection algorithms. In this way we are able to investigate the impact of feature selection on the classification performance.

4. EXPERIMENTAL SETUP AND EVALUATION

In this section we provide the outcome of the feature selection procedure, comparing the result of each algorithm and showing the features that have been chosen most times. We then validate the set of selected feature in an historical versus modern automatic classification application.

Data Collection. In order to populate the dataset we recorded a set of instruments: 13 historical violins form the collection of the Violin Museum in Cremona and 9 modern violins from the violin making school Istituto Antonio Stradivari in Cremona.

The recordings were performed in a semi-anechoic room, using a measurement microphone always placed in the same position with respect to the instrument. In particular, the microphone was placed at a distance of approximately 40 cm, perpendicularly to the bridge. The audio was acquired with a sample rate of 44100 Hz. Since we did not want to include spatial information in this study, we acquired mono recordings. All the recordings were performed by the same professional musician that uses the same bow. For each instrument, the musician played the four open strings, a sequence of staccato notes on each string, a legato major scale from G4 to F#7 covering all the strings and 6 pieces of classical music including several styles and techniques. Therefore, each recording resulted in 15 sections: *1:Open G string; 2:Open D string; 3:Open A string; 4:Open E string; 5:Notes on G string; 6:Notes on D string; 7:Notes on A string; 8:Notes on E string; 9:Scale; 10:Excerpt 1; 11:Excerpt 2; 12:Excerpt 3; 13:Excerpt 4; 14:Excerpt 5; 15:Excerpt 6.*

Consistency between Feature Selection algorithms. In this section we investigate the consistency between the outputs of the adopted feature selection algorithms. We present the results by grouping the recorded sections into three subsets according to their musical content: open strings, single notes and scale/excerpts. We believed that the features that characterize the timbre of historical violins have different impact according to the nature of their content.

In order to compare and validate the results of the Relieff and Decision Tree ranking algorithms, we used the Spearman's rank correlation coefficient, that is able to compute how much two rankings are similar:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \tag{1}$$

Section	Spearman's coefficient
Open strings	0.80
Single notes	0.80
Scale and excerpts	0.69

Table 1. Spearman's rank correlation coefficient between the Relieff ranking and the forest-of-trees ranking, for each subset .

Section	Fraction of common features
Open strings	0.80
Single notes	0.68
Scale and excerpts	0.70

Table 2. Fraction of features that all the three selection algorithms choose most times.

where d_i is the ranking difference for the i-th feature and n is the number of elements in the rankings.

We obtained good results on the correlation between the two ranking methods for almost every section. This means that the rankings were consistent. We reported them in table 1.

In order to study the consistency between the other methods, for each group of sections we took the 20 features that have been chosen most times by each algorithm. Then we intersected the three lists in order to find the features on which the three methods agreed. The result is showed in Table 2 as the number of common features divided by 20. The open strings, where the audio content is simpler and steadier, show the best consistency.

Results on Feature Selection. Features that are most discriminant for our task are those which are most frequently selected by the selection algorithms. Results of the selection is presented in Figure 1 where we report the most frequently selected features from the five methods.

It can be seen that the four statistical moments of the spectrum resulted discriminative, in particular for the sections that are characterized by steady sounds (open strings and single notes). Two MFCC coefficients, two Spectral Contrast coefficients and the first coefficient of the Tristimulus resulted important for the scale and the excerpts. This could be explained by the fact that, in presence of a more complex and variable audio content, more general features are needed, which do not represent just one specific aspect of the spectrum and better capture the complexity of the content.

As an example of discriminative ability of the features, in Figure 2 we show the bidimensional feature space represented by MFCC2/MFCC7.In this space, the samples relative to the open strings are arranged. Even using only a couple of features, a separation can be noticed between the two classes (in particular for G string and A string).

Classification results.

In this section we show the results of the historical versus modern violin classification in the case when all the features are used and in the case when only selected features are. The set of feature used in the latter case are the ones reported in 1. Table 3 summarizes the result by reporting the classification errors of the Support Vector Classifier. The best results are

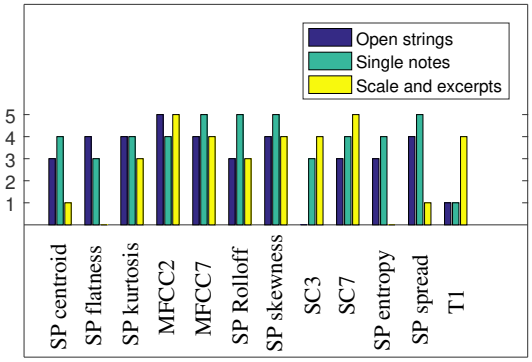


Figure 1. Number of algorithms that chose each feature for different sections. SC stands for Spectral Contrast, while T1 stands for Tristimulus (first coefficient)

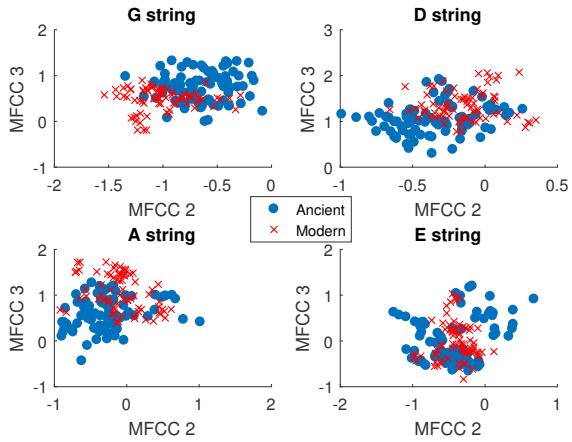


Figure 2. Values of the 2nd and 7th MFCCs for the open strings

achieved for the open strings. This could be due to the fact that the classification task is easier with steady notes, where there are not pauses and transients that affect the statistics of the low-level features. Table 4 shows the confusion matrix of the classification performed on the open strings.

In general, feature selection improves the classification results, meaning that the selected features are more relevant for this task.

5. CONCLUSIONS

In this work we investigated on the timbral aspects that are able to discern the sound of historical violins from the sound of contemporary ones, from a feature-based perspective. We applied different feature selection algorithms and we found out the most discriminative ones. In particular, the statistical moments of the spectrum and some MFCC and Spectral Contrast coefficients appeared to be important from this point of view. As we imagined, the relevance of such features depend

Section	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AF	0.06	0.03	004	0.04	0.13	0.17	0.17	0.18	0.11	0.12	0.06	0.13	0.13	0.06	0.21
FS	0.02	0.01	0.03	0.02	0.20	0.15	0.18	0.16	0.10	0.11	0.10	0.10	0.18	0.03	0.19

Table 3. Classification error with the Support Vector Classifier, both with All features (AF) and with Feature Selection (FS)

	Historical	Modern
Historical	0.93	0.07
Modern	0.17	0.83

Table 4. Confusion matrix of the historical vs modern automatic classification.

on the audio content.

As a validation of the feature selection, we tested a classifier and proved that a classification model that is able to discern a modern violin from an old one can be implemented and its performances improve with feature selection.

This work constitutes the first step in understanding which are the sound properties that characterize the timbre of violins. The next step is to perform a qualitative and quantitative analysis on the audio features that appeared to be more relevant, and integrate this data with vibroacoustic measurements, in order to understand what are the structural factors that most impact these features. This will be addressed in future works.

REFERENCES

[1] C. Fritz, J. Curtin, J. Poitevineau, H. Borsarello, F.-C. Tao, and T. Ghasarossian, "Soloist evaluations of six old italian and six new violins," *Proceedings of the National Academy of Sciences*, vol. 111, no. 20, pp. 7224–7229, 2014.

[2] E. Lukasik and R. Susmaga, "Unsupervised machine learning methods in timbral violin characteristics visualization," in *Proc. Stockholm Music Acoustics Conference SMAC*, vol. 3, 2003, pp. 83–86.

[3] B. L. Sturm, M. Morvidone, and L. Daudet, "Musical instrument identification using multiscale mel-frequency cepstral coefficients," in *Signal Processing Conference, 2010 18th European*. IEEE, 2010, pp. 477–481.

[4] E. Lukasik, "Long term cepstral coefficients for violin identification," in *In proceedings of the Audio Engineering Society Convention 128 (AES128)*, 2010.

[5] ——, "Matching violins in terms of timbral features," *Archives of Acoustics*, vol. 31, no. 4, p. 227, 2006.

[6] M. Zanoni, F. Setragno, F. Antonacci, A. Sarti, G. Fazekas, and M. B. Sandler, "Training-based semantic descriptors modeling for violin quality sound characterization," in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[7] H. Terasawa, M. Slaney, and J. Berger, "Perceptual distance in timbre space." Georgia Institute of Technology, 2005.

[8] D. Jang, M. Jin, and C. D. Yoo, "Music genre classification using novel features and a weighted voting method," in *2008 IEEE International Conference on Multimedia and Expo*. IEEE, 2008, pp. 1377–1380.

[9] T. S. H.G. Kim, N. Moreau, *MPEG-7 Audio and Beyond. Audio Content Indexing and Retrieval*. John Wiley & Sons Ltd, 2005.

[10] D. N. Jiang, L. Lu, H. J. Zhang, J. H. Tao, and L. H. Cai, "Music type classification by spectral contrast features," in *Proceedings of IEEE International Conference Multimedia Expo*, 2002.

[11] O. Lartillot and P. Toiviainen, "Mir in matlab (ii): A toolbox for musical feature extraction from audio," in *2007 International Society for Music Information Retrieval conference (ISMIR)*, 2007.

[12] K. Jensen, "Timbre models of musical sounds, tech. rep. rapport 99/7," University of Copenhagen, Tech. Rep., 1999.

[13] R. Plomp and W. J. M. Levelt, "Tonal consonance and critical bandwidth," *Journal of the Acoustical Society of America*, vol. vol. 38, pp. pp. 548–560, 1965.

[14] P. Juslin, "Cue utilization in communication of emotion in music performance: relating performance to perception." *Journal of Experimental Psychology: Human Perception and Performance*, vol. vol. 26, no. no. 6, pp. pp. 1797–1813, 2000.

[15] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler *et al.*, "Api design for machine learning software: experiences from the scikit-learn project," *arXiv preprint arXiv:1309.0238*, 2013.

[16] P. Geurts, D. Ernst, , and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63(1), pp. 3–42, 2006.

[17] K. Kira and L. A. Rendell, "The feature selection problem: Traditional methods and a new algorithm," in *AAAI*, vol. 2, 1992, pp. 129–134.

[18] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.

KEYNOTE ADDRESS

Collisions in Musical Instruments: a Numerical Modelling Perspective

Maarten van Walstijn[†]

Sonic Arts Research Centre, Queen's University Belfast, UK
[†]m.vanwalstijn@qub.ac.uk

ABSTRACT

The vibrational behaviour of musical instruments often involves some form of collision. The nature of the contact can vary from compact in both space and time, as with hammer-string interaction, to more distributed and sustained forms, such as string-membrane collisions in a snare drum. The former can usually be modelled in lumped form, suppressing the computation of the interaction forces to a single point, whereas the latter requires considering variations along spatial coordinates. In most cases, the impactive interaction represents an important nonlinear element that is closely linked to the characteristics and/or the gestural control of the instrument. Given the non-smooth nature of collision forces, incorporating such phenomena into time-stepping simulations can be challenging due to potential instability and further issues around computability, accuracy and efficiency. Approaches to address these challenges are discussed in this talk, starting from considering collision between a single mass colliding with a barrier and culminating in the simulation of full distributed interaction between a nonlinearly vibrating string and a barrier of arbitrary spatial profile. The presented numerical methods can serve as a basis or tool for time-domain study and sound synthesis of a range of stringed instruments. Specific attention will be given to so-called flat-bridge instruments such as the sitar and the tanpura, which produce spectrally rich drone sounds through sustained impactive string-bridge interaction.

On the Contact Duration Accuracy of Discrete-Time Collision Models

Vasileios Chatziioannou,[1†] *Maarten van Walstijn,*[2]

[1]Department of Music Acoustics, University of Music and Performing Arts Vienna, Austria
[2]Sonic Arts Research Centre, Queen's University Belfast, UK
[†]chatziioannou@mdw.ac.at

ABSTRACT

Collisions are a vital part of the function of musical instruments. They occur in various forms and during different phases of oscillation, including transient and steady-state regimes. The non-smooth nature of the nonlinearity inherent to objects coming into contact and decoupling again poses several difficulties when it comes to formulating numerical models. Besides issues regarding numerical stability – that have recently been handled using energy methods – questions arise about how accurately different numerical schemes can approximate the trajectories of colliding objects. This paper presents a comparative analysis of two particular types of time-stepping algorithms (namely a two-point, two-variable and a three-point, one-variable scheme) employed to simulate contact between a mass and a barrier. Focusing largely on cases for which the exact solution is known, the schemes are evaluated in terms of their ability to simulate the correct duration of contact.

1. INTRODUCTION

Many musical instruments feature mechanical collisions that can be assumed elastic or semi-elastic, with the contact characteristics determined mostly by the impact velocity and the material and geometrical properties of the colliding objects. Such collisions can be modelled by defining a repelling force that is non-zero only for negative inter-object distance, which corresponds to object compression. In music acoustics, this contact force is often defined in power law form, the parameters of which determine the duration and level of the compression. Having a physical basis in Hertzian contact theory, it provides simple and effective control over the collision characteristics, ranging from soft (e.g. hammer-string interaction [1]) to hard collisions (e.g. string-bridge contact [2]).

Incorporation of power-law contact into time-stepping schemes requires to observe the energy balance of the underlying continuous-domain system, as otherwise numerical stability issues may arise. Several methods for deriving energy-stable schemes have been proposed recently, with the discretisation performed over either a first-order [3, 4, 5] or a second-order [6, 7, 8] form. A recent study on modelling distributed string-barrier contact [9] has reported indications that the latter type is more prone to spurious oscillations. The present study makes further inroads into the understanding of the accuracy of these schemes by investigating impactive contact between a single mass and a barrier, focusing on errors in the contact duration.

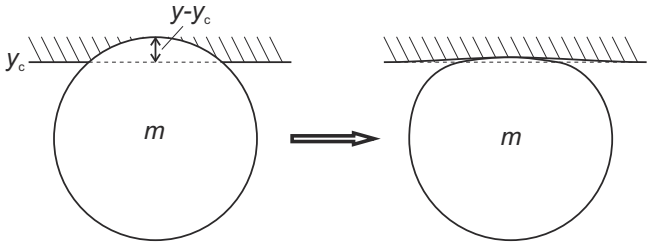


Figure 1. Left: mass-wall collision where the penetration argument $y - y_\mathrm{c}$ is defined; Right: the equivalent physical setting, showing the compression of the mass and the wall.

2. MASS-WALL COLLISION

Consider a (lumped) mass approaching a wall from below. The collision force can be modelled using a one-sided power law [10]. This may be formulated as

$$f(y) = -k_\mathrm{c}\lfloor (y - y_\mathrm{c})\rfloor^\alpha, \tag{1}$$

where y is the mass displacement, k_c the collision stiffness, α a power-law constant, y_c the location of the wall and $\lfloor \chi \rfloor^\alpha = h(\chi)\chi^\alpha$, $h(\chi)$ being the Heaviside step function. Hence $y - y_\mathrm{c}$ corresponds to the compression of the colliding objects (see Figure 1, noting that both the moving mass and the wall are assumed to get compressed during collision). The net amount of compression is controlled in the model by the contact parameters k_c and α. For simplicity, in what follows y_c is set equal to zero. Defining the collision force potential as

$$V(y) = \frac{k_\mathrm{c}}{\alpha+1}\lfloor y \rfloor^{\alpha+1} := \frac{k_\mathrm{c}}{\alpha+1} G(y) \tag{2}$$

the equation of motion for the mass m reads

$$m\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = -\frac{\partial V}{\partial y} \tag{3}$$

hence the Hamiltonian (total energy) of the system is

$$H(y,p) = \frac{p^2}{2m} + \frac{k_\mathrm{c}}{\alpha+1}\lfloor y \rfloor^{\alpha+1}, \tag{4}$$

where p is the conjugate momentum, with

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial H}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} = k_\mathrm{c}\lfloor y \rfloor^\alpha \frac{p}{m} + \frac{p}{m} m \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} \overset{(3)}{=} 0 \tag{5}$$

so, in the absence of damping terms, the energy of the system remains constant.

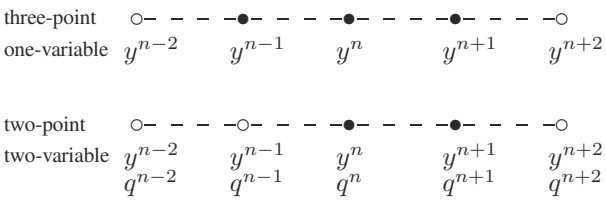


Figure 2. Three-point (top) and two-point (bottom) temporal stencils.

2.1. Discretisation

Two different starting points may be used for the discretisation of the above problem. Either Newton's equation of motion (3) can be directly discretised, or the problem can be first written using Hamilton's equations [11]. The former approach requires the numerical solution of a single, second-order differential equation, whereas the latter is based on the solution of two first-order equations, namely

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial H(y,p)}{\partial p} = \frac{p}{m}, \tag{6a}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H(y,p)}{\partial y} = -\frac{\partial V(y)}{\partial y}. \tag{6b}$$

Note that this first order form can be also obtained from the Newtonian description, by reducing the single second-order differential equation (3) to two first-order ones (see, e.g. [12]).

Discretising Newton's equation of motion (as in [6]) yields the following numerical scheme

$$y^{n+1} - 2y^n + y^{n-1} = -\beta_3 \frac{G(y^{n+1}) - G(y^{n-1})}{y^{n+1} - y^{n-1}}, \tag{7}$$

where $\beta_3 = \big(k_\mathrm{c}\Delta t^2\big)/\big(m(\alpha+1)\big)$ and y^n denotes the value of variable y at time $n\Delta t$, with $\Delta t = 1/f_s$ (f_s being the sampling rate). Equation (7) can be said to form a three-point, one-variable scheme, since its temporal stencil encompasses three grid points (see Figure 2 top). On the other hand, discretising Hamilton's equations using mid-point derivative approximations and setting $q^n = p^n\Delta t/(2m)$ (see [4]) yields

$$y^{n+1} - y^n = q^{n+1} + q^n \tag{8a}$$

$$q^{n+1} - q^n = -\beta_2 \frac{G(y^{n+1}) - G(y^n)}{y^{n+1} - y^n}, \tag{8b}$$

which is a two-point, two-variable scheme[1] (see Figure 2 bottom) with $\beta_2 = \beta_3/2$. Both schemes (7) and (8) are written in terms of two parameters, namely α and $\beta_\kappa, \kappa = 2, 3$, and can be shown to be second order accurate[2] for $\alpha > 1$, as well as numerically stable for $\beta_\kappa \geq 0$ (see [4, 10]) which holds unconditionally. The equations are nonlinear due to the presence of the collision term and may be solved using an iterative method (such as Newton's method). Figure 3 shows simulations of a softer ($\alpha = 1$, $k_\mathrm{c} = 10^9$) and a harder collision ($\alpha = 1$, $k_\mathrm{c} = 10^{10}$) of a 1 kg mass approaching a barrier from below with initial displacement $y(0) = -0.1$ mm and velocity $v(0) = 1$ m/s. The sampling rate is $f_s = 50$ kHz.

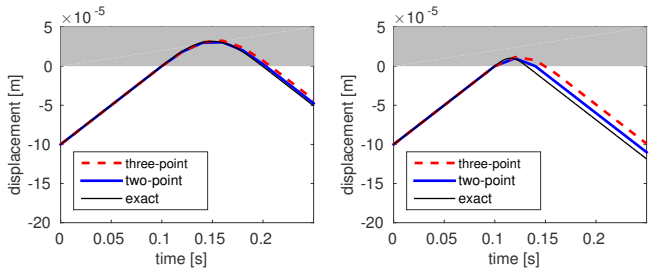


Figure 3. Mass displacement for a 'soft' (left) and a 'hard' collision (right). The red dashed curve corresponds to the three-point scheme and the blue curve to the two-point scheme, whereas the black line represents the exact solution.

In both plots, a difference in the contact duration τ_c is observed. This appears to be a systematic phenomenon that is particularly significant for hard impacts. One way to get insight into what causes these differences is to consider, for each scheme, whether it is possible to 'exit' the wall immediately after entering it. For the three-point scheme, this amounts to investigating the mapping in equation (7). Exiting the wall one time step after entering implies $\{\ y^{n-1} \leq 0,\ \ y^n > 0\ \} \Rightarrow y^{n+1} \leq 0$. Defining $y^{n-1} = -\delta$ and $y^n = \epsilon$, with $\delta \geq 0, \epsilon > 0$, equation (7) can then be written as

$$y^{n+1} = 2\epsilon + \delta - \beta_3 \frac{G(y^{n+1}) - G(-\delta)}{y^{n+1} + \delta}. \tag{9}$$

Considering the form of the collision potential and requiring y^{n+1} to be nonpositive yields $y^{n+1} = 2\epsilon + \delta > 0$ which is a contradiction. Hence, under the three-point scheme the mass-wall interaction can never be confined to a single time-step. On the other hand, if both $y^{n-1}, y^n > 0$ then the resulting equation (assuming $y^{n+1} \leq 0$ and setting $y^{n-1} = \eta > 0$) reads

$$y^{n+1} = 2\epsilon - \eta - \beta_3 \frac{G(\eta)}{\eta - y^{n+1}}, \tag{10}$$

which is of indeterminate sign. Hence, after having remained within the barrier for two time steps, the mass has the possibility to exit at the next time step. Hence, the minimum contact duration that this scheme can simulate is equal to $2\Delta t$, (i.e. $\tau_\mathrm{c} \geq 2\Delta t$).

For the two-point scheme a similar calculation can be performed to show that $\tau_\mathrm{c} \geq \Delta t$. The numerical scheme (8) may be written as

$$y^{n+1} = y^n + 2q^n - \beta_2 \frac{G(y^{n+1}) - G(y^n)}{y^{n+1} - y^n}. \tag{11}$$

The condition for a collision to occur is that, in the absence of a contact force, the mass will 'enter' the barrier at the next time step, which is equivalent to

[1] Note that, as explained in [4], under linear conditions this scheme is equivalent to both the trapezoidal rule and the implicit mid-point rule. However, in the presence of nonlinear forces, such as those acting during collisions, all three schemes are distinct (and only (8) is energy conserving).

[2] for $\alpha = 1$ only first order accuracy can be inferred, since $G(y)$ is not twice differentiable at $y = 0$.

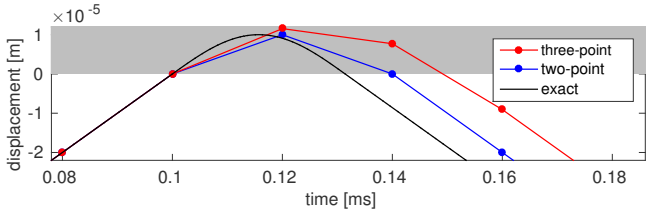


Figure 4. Detail of the mass displacement across the barrier in the case of a hard collision. Notice that only the two-point scheme simulates an impact force of a single time step.

$\{\, y^n \leq 0,\ \ q^n > 0 \,\} \Rightarrow \hat{y}^{n+1} > 0$, where $\hat{y}^{n+1}$ is the virtual displacement at the next time step had the collision force not been present (i.e. for $\beta_2 = 0$). Defining $y^n = -\zeta$ and $q^n = \mu$, with $\zeta \geq 0, \mu > 0$ leads to

$$\hat{y}^{n+1} = 2\mu - \zeta > 0. \tag{12}$$

We proceed to show that in the presence of a contact force the mass will spend at least one time instant within the barrier (i.e. $y^{n+1} > 0$). Assume that this is false and $y^{n+1} \leq 0$. Then

$$y^{n+1} = 2\mu - \zeta - \beta_2 \frac{G(y^{n+1}) - G(-\zeta)}{y^{n+1} + \zeta} = 2\mu - \zeta \overset{(12)}{>} 0, \tag{13}$$

which is a contradiction. Hence $y^{n+1} > 0$ and $\tau_c \geq \Delta t$.

Furthermore, it can be shown that after having spent one time instant within the barrier, the mass has the possibility to exit the barrier (with the following displacement value being of indeterminate sign). So any contact duration higher than Δt may be simulated by the two-point scheme. This is visualised in Figure 4, which shows the details of a contact with ($\alpha = 1$, $k_c = 10^{11}$), the remaining parameters being as in Figure 3. The fact that the two-point scheme may simulate shorter contact durations results in a smaller error compared to the three-point scheme. This error comparison is further discussed in the next section where the possibility of simulating the correct contact duration is explored.

3. EXACT SOLUTION FOR $\alpha = 1$

For $\alpha = 1$, and initial conditions $y(0) = -d$, $v(0) = v_0$, an exact solution can be obtained in the form of a continuous, twice differentiable piecewise function

$$y(t) = \begin{cases} -d + v_0 t & : \quad 0 \leq t < t_1 \\ \dfrac{v_0}{\omega_c} \sin\left(\omega_c (t - t_1)\right) & : \quad t_1 \leq t < t_2 \\ -v_0 (t - t_2) & : \quad t \geq t_2 \end{cases}, \tag{14}$$

where $t_1 = d/v_0$ and $t_2 = t_1 + \tau_c$ are the time of impact and release, respectively, and where $\omega_c = \sqrt{k_c/m}$ is the natural frequency that would occur if the potential in (2) was always 'active', i.e. if $V(y) = \frac{1}{2} k_c y^2$. The contact duration is $\tau_c = \pi/\omega_c$, which corresponds to a half cycle of a sinusoidal waveform of frequency ω_c.

Both schemes presented in the previous section can – within certain limits – be made exact in terms of this contact duration (i.e. having the correct natural frequency ω_c). For the two-point scheme this is achieved via the substitution

$$\beta_2 \to \beta_2^* = \frac{1 - \cos(\omega_c \Delta t)}{1 + \cos(\omega_c \Delta t)}, \tag{15}$$

with the constraint that $\omega_c < \pi/\Delta t$, as otherwise the natural frequency exceeds the Nyquist frequency (i.e. is aliased). This substitution, that can be derived using frequency-domain analysis (see, e.g. Sec. 2.3 in [5]), ensures that the numerical scheme has the correct contact duration. Note that numerical stability is not affected, because the condition $\beta_2^* \geq 0$ is still satisfied unconditionally. The non-aliasing constraint translates to a lower bound $\tau_c > \Delta t$ on the contact duration.

For the three-point scheme the equivalent substitution to obtain the correct contact duration (again for $\alpha = 1$) is

$$\beta_3 \to \beta_3^* = \frac{1 - \cos(\omega_c \Delta t)}{\cos(\omega_c \Delta t)}. \tag{16}$$

However another, more severe constraint $\omega_c < \frac{1}{2}\pi/\Delta t$ applies now, because otherwise β_3^* is not guaranteed to be non-negative. This stability constraint translates to a lower bound $\tau_c > 2\Delta t$ on the contact duration. Hence both schemes can be made exact in terms of their oscillatory behaviour (see Figure 5) up to a minimum contact duration (for a given sampling rate) which is twice as small for the two-point scheme, in comparison to the three-point scheme. When it is required to simulate contacts with durations shorter than this lower bound this substitutions can only be performed to the limit that aliasing is avoided and β_κ remains non-negative. Hence an error is introduced which can be quantified as the difference between the exact and the simulated contact duration.

Note that these lower bounds on τ_c are in agreement with those calculated in the previous section from the discrete update equations, for arbitrary values of α. Figure 6 shows (for the parameters of the hard collision in the previous section) the minimum contact duration that may be exactly simulated using each scheme, as a function of the sampling rate. The bottom plot shows the error that is introduced by each approximation when shorter durations need to be simulated.

From the above analysis it can be extracted that (1) the two-point scheme can exactly simulate contacts of durations twice as short as the three-point scheme and (2) when even shorter contact durations need to be simulated the error introduced by the three-point scheme is significantly larger.

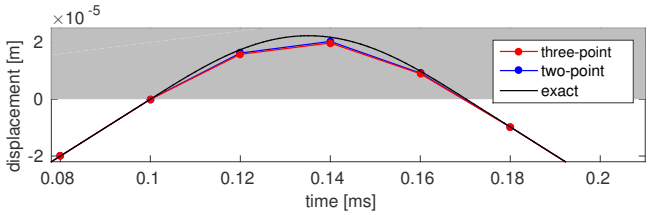


Figure 5. Detail of the mass displacement across the barrier when the corrected values β_2^* and β_3^* are used for $\omega_c = 44700$ rad/s.

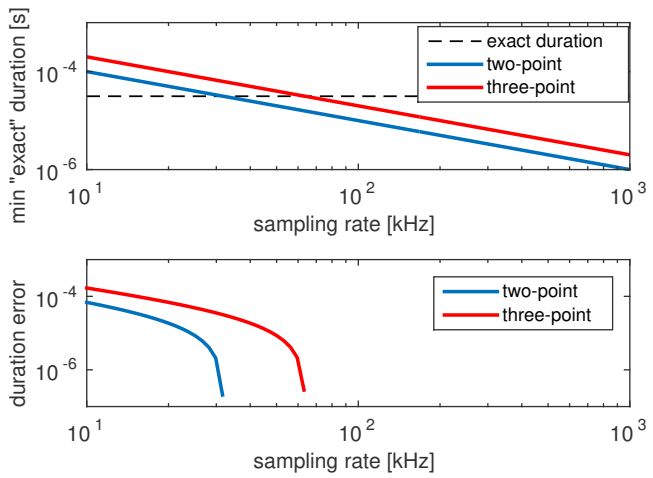


Figure 6. Top: the minimum contact duration that each scheme can simulate exactly for different sampling rates (for the parameters of the hard collision of Figure 3). The dashed line indicates an example target duration. Bottom: the error introduced in the contact duration when it is too short to be simulated exactly for a given sampling rate.

4. CONCLUSIONS

In this paper two different schemes (that have recently seen much application in simulation of musical instrument contact dynamics) have been analysed, based on their ability to simulate the colliding mass trajectory. In particular, it has been shown that different results are obtained, especially for hard collision cases, where the reference contact duration is of the order of the sampling interval.

In the context of music acoustics research, the ability of a numerical scheme to accurately model the contact duration time can be important. Note onsets are often related to impactive excitations, so simulations of transients may be affected by inaccuracies in the underlying collision model. Furthermore, when repeated impacts are simulated (as is the case for example in instruments involving string-barrier collisions, such as a sitar or a snare drum) the periodicity and/or spectral evolution are affected by the contact duration (see, for example, [13]). So in general, the time step has to be chosen sufficiently small to enable to compute the results for a particular study without significant error. As such it is advantageous if the required accuracy can be achieved with a lower sample rate, which is what the two-point scheme offers over the three-point scheme.

An interesting future research direction would be to investigate whether and how model parameters could be set to better approximate contact durations for cases when $\alpha > 1$. Another question that arises from this study is to what extent the differences between two-point, two-variable and three-point, one-variable schemes are potentially significant for other types of musical instrument nonlinearity and/or for time-variant components of musical instruments.

5. ACKNOWLEDGEMENTS

This research is supported by the Austrian Science Fund (FWF): P28655-N32.

REFERENCES

[1] A. Chaigne and J. Kergomard, *Acoustics of Musical Instruments*. New York: Springer, 2016.

[2] M. van Walstijn and V. Chatziioannou, "Numerical simulation of tanpura string vibrations," in *Proc. Int. Symp. Musical Acoustics*, Le Mans, 2014, pp. 609–614.

[3] V. Chatziioannou and M. van Walstijn, "An energy conserving finite difference scheme for simulation of collisions," in *Proc. Sound and Music Computing (SMAC-SMC 2013)*, Stockholm, 2013, pp. 584–591.

[4] V. Chatziioannou and M. van Walstijn, "Energy conserving schemes for the simulation of musical instrument contact dynamics," *Journal of Sound and Vibration*, vol. 339, pp. 262–279, 2015.

[5] M. van Walstijn, J. Bridges, and S. Mehes, "A real-time synthesis oriented tanpura model," in *Proc. Int. Conf. Digital Audio Effects (DAFx-16)*, 2016.

[6] S. Bilbao, A. Torin, and V. Chatziioannou, "Numerical modeling of collisions in musical instruments," *Acta Acustica united with Acustica*, vol. 101, no. 1, pp. 155–173, 2015.

[7] M. Ducceschi, S. Bilbao, and C. Desvages, "Modelling collisions of nonlinear strings against rigid barriers: Conservative finite difference schemes with application to sound synthesis," in *Proc. 22nd Int. Congress on Acoustics*, Buenos Aires, 2016.

[8] J. Chabassier, A. Chaigne, and P. Joly, "Modeling and simulation of a grand piano," *Journal of the Acoustical Society of America*, vol. 134, no. 1, pp. 648–665, 2013.

[9] M. van Walstijn and J. Bridges, "Simulation of distributed contact in string instruments: a modal expansion approach," in *Proc. Europ. Sig. Proc. Conf. (EUSIPCO2016)*, 2016.

[10] S. Bilbao, *Numerical Sound Synthesis*. Chichester, UK: Wiley & Sons, 2009.

[11] V. Arnold, *Mathematical methods of classical mechanics*. New York: Springer, 1978, vol. 60.

[12] D. Greenspan, "Conservative numerical methods for $\ddot{x} = f(x)$," *Journal of Computational Physics*, vol. 56, no. 1, pp. 28–41, 1984.

[13] C. Issanchou, S. Bilbao, J. Le Carrou, C. Touzé, and O. Doaré, "A modal-based approach to the nonlinear vibration of strings against a unilateral obstacle: Simulations and experiments in the pointwise case," *Journal of Sound and Vibration*, vol. 393, pp. 229–251, 2017.

Realistic Guitar Modelling using a Dynamical Multibody Approach

Vincent Debut,[1†] *Jose Antunes*[2]

[1]INET-md, Universidade Nova de Lisboa
[2]C2TN, Instituto Superior Tecnico
[†]vincentdebut@fcsh.unl.pt

ABSTRACT

Most musical instruments consist on a set of dynamical subsystems connected at a number of constraining locations, through which the vibratory energy flows or tuning can be achieved. Coupling is therefore an essential feature in instrument modelling and, when addressing physically-based synthesis, most modelling and computational difficulties are connected with the manner in which the coupling constraints are enforced. Typically, these are modelled using standard techniques such as Lagrange multipliers or penalty methods, each one with specific merits and drawbacks. In this paper we explore an approach based on the Udwadia-Kalaba (U-K) formulation, originally proposed in the early 90s for discrete constrained systems, which is anchored on analytical dynamics but avoids the use of Lagrange multipliers. Up to now, this general, very elegant and appealing formulation has been nearly exclusively used to address conceptual systems of discrete masses or articulated rigid bodies, namely in robotics. In a recent publication - Antunes & Debut, JASA (2017, in print) - we developed an extension of the U-K formulation to deal with flexible systems modelled through their unconstrained modes. We explored the potential of combining the U-K formulation for constrained systems with the modal description of flexible structures, in order to achieve reliable and efficient computations of the dynamical responses. This modelling approach was shown to be particularly effective, in particular for simulating the transient responses of musical instruments, as demonstrated by computing the dynamical responses of a guitar string coupled to the instrument body at the bridge. In the present paper we further generalize our computational model by incorporating all the guitar strings with both their motion polarisations, as well as the geometrical nonlinear string effects using the Kirchoff-Carrier (K-C) simplified approach. The illustrative results presented highlight the coupling of the various strings and body dynamics through the instrument bridge, and also emphasize the significance of string nonlinearity in the responses of plucked string instruments, which often lead to audible modal coupling terms and frequency gliding effects. The results presented complement extensive work already performed in the past by the authors on guitar string modelling using penalty methods, thus enabling an interesting comparison between the computational efficiency using different modelling techniques.

Physical Modelling and Sound Synthesis of a Viola Caipira

Guilherme Paiva,[1,2†] *Frédéric Ablitzer,*[2] *François Gautier,*[2] *Mathieu Sécail-Géraud,*[2]
José Maria Campos Dos Santos[3]

[1]University of Campinas
[2]Laboratoire d'Acoustique de l'Université du Maine
[3]LVA-UNICAMP
[†]pitupaiva@gmail.com

ABSTRACT

The Viola Caipira is a folk guitar played in traditional and modern Brazilian music. In general, it has ten metal strings arranged in five courses of two strings with the thinnest string located in the middle. Two pairs have identical strings tuned in unison and three pairs have strings with different thicknesses tuned in an octave. This work presents an experimental study of the Viola Caipira pluck by means of a high speed camera (5000 frames per second), which allowed to reveal some specificities of the instrument. It is found that the instrument is characterized by a double pluck excitation : The two strings of a given course are plucked sucessively and rapidly. Sympathetic phenomena and collisions between strings are also identified. A physical modelling based on a hybrid modal approach including string/string collisions is developped for the sound synthesis of the viola caipira sounds. Such approach is hybrid since it combines two different ways for describing strings and body dynamics: Analytical expressions of mode shapes, natural frequencies and damping factors are used to obtain the modal basis for each string while body modal parameters are extracted from mobility measurements at the bridge. In order to couple both strings and body, a finite difference scheme is used to calculate the coupling forces at each time-step, which permits a set of sound simulations. It is found that the model reproduces the main features of the Viola Caipira sounds, among which the sympathetic string resonances and the string/string collisions observed in the video analysis.

Lip Dynamics in a Physical Model of the Trumpet

N. Giordano [1†]

[1]Department of Physics, Auburn University, Auburn AL, 36849, USA

[†]njg0003@auburn.edu

ABSTRACT

A model of the trumpet has been studied in which the Navier-Stokes equations are used to calculate the air pressure and velocity, while the lips are described using the swinging lip model of Adachi and Sato. This approach allows us to determine the pressure throughout the lip region, and thus calculate the lip motion without the approximations necessary in previous modeling work.

1. INTRODUCTION

Models of musical instruments have become increasingly sophisticated in recent years, addressing questions that are difficult or impossible to tackle with experiments. Wind instruments such as the trumpet pose particular challenges, since modeling must include the dynamics of the air through the instrument as well as the motion of the player's lips. This paper describes an initial attempt to model the aeroacoustics of a brass instrument in a quantitatively accurate way using a direct application of the Navier-Stokes equations. While the instrument geometry we consider is admittedly simplified, our results should pave the way to the study of realistic geometries.

Essentially all previous modeling of brass instruments has assumed one-dimensional airflow and a Bernoulli-like approximation for the pressure near the lips (e.g., [1-4]). These approximations have been paired with various lip models, with the one devised by Adachi and Sato [2] appearing to be the simplest approach that captures the essential features of a "blown open" lip reed. While previous modeling of the trumpet has led to important insights, that work has significant limitations due to the approximate treatment of the pressure in the lip region. In this work we use a first-principles model of the airflow which allows us to compute the forces on the lips, and hence determine the lip motion, without many of the approximations made in previous work. We begin by describing our model and compare our results with previous work. We then present the first results for an important case that cannot be studied with previous modeling approaches - the behavior of asymmetric lips.

2. THE MODEL

Figure 1 shows our simplified model of the trumpet. It consists of a conical bore of length 5 cm with a diameter at the open end of 1.4 cm attached to rectangular lips that are able to swing in the direction of the cone axis in Fig. 1 and undergo stretching and compressive motion in the transverse direction. Our choice of such a small instrument reduces the computational time required for the aeroacoustic calculations.

Lip motion is described using the model of Adachi and Sato [2], according to which each lip experiences a Hooke's law restoring force when displaced from its equilibrium position, along with forces on the lip surfaces due to the air pressure.

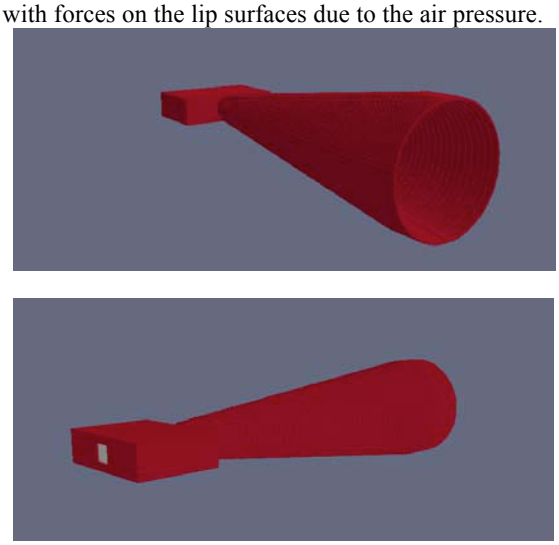


Figure 1. Model of the trumpet studied in this work. Top: Showing the conical bore open at the lower right. Bottom: The opening at the lower left leads inside to the mouth and lips, which are based on the description in [2].

There are a number of lip parameters in the model, including the mass m, stiffness (i.e., spring) constant k, and damping factor Q, along with the dimensions of the lips. For this study the lips were each 2.0 mm long (along the flow direction), 1.0 mm thick (in the flexible transverse direction denoted below by y), and 2.0 mm wide in the third dimension, and the nominal lip opening with no air flow was 0.6 mm along the y direction. The nominal cross sectional area of the lip channel was 1.2 mm^2 while that of the mouth (upstream from the lips) was 8.0 mm^2. The corresponding volumes were 2.4 mm^3 and 28 m^3. The lip mass for most of the calculations presented here was determined by assuming a lip density of 1000 kg/m^3, the approximate density of human tissue. The value of k then determined the natural frequency of each lip, and the damping was set by Q = 3, as suggested in [2]. Not surprisingly, the largest lip oscillations and largest sound pressure was found when the natural frequencies of the lips were close to one of the resonant frequencies of the bore, which was near 1 kHz. In such cases the lip oscillations could become very large and to maintain stability an extra restoring force was imposed when the lips touched (a similar approach was taken in [2]).

The air flow through the instrument was calculated using the compressible Navier-Stokes equations with the numerical approach described in [5]. The algorithm was a three-dimensional finite-difference time-domain calculation that

yielded the air velocity and density on a grid with spacing 0.1 mm in the vicinity of the lips and inside the instrument. The results of studies of the recorder in [5] validated that our numerical method gives an accurate description of the flow at the length scale appropriate for the present work. The lips were treated using the immersed boundary method developed originally by Peskin [6] (see also [7,8] for a description of the method). With this method the lip edges are not limited to grid points, but are able to move continuously so the resolution of the lip motion is not limited by the grid spacing used in the Navier-Stokes calculation. The trumpet model was contained in a closed region with walls that reflect and absorb sound, as would the walls in a typical room [5]. The total number of grid points for the model considered here was 4.4×10^7.

In our algorithm [5] the instrument is "blown" by imposing a constant air velocity u that follows Poiseuille's law in a channel leading to the mouth cavity that then leads to the lips. That velocity was zero at $t = 0$ and increased linearly with time until a final value was reached at $t = 5$ ms; the blowing speed was then kept constant for the rest of the simulation. The blowing speeds required to produce lip oscillations and tones were somewhat larger than for a real trumpet, due to the small size of our model. For our chosen mouth geometry the pressure in the mouth was typically 10 kPa, a typical value for real playing, and the large mouth volume (relative to the lips) acted to somewhat buffer the flow as would be the case for a real player. The oscillations were stable for the duration of each simulation, with no hint of instabilities at long times. Steady state was generally reached after about 10 or 20 ms, and was then maintained for the duration of each simulation (typically 100 ms).

3. BEHAVIOR WITH SYMMETRIC LIPS

Some typical results for the sound pressure as a function of time with symmetric lips, i.e. with the two lips having the same dimensions, mass, stiffness constant, and other properties, are shown in Fig. 2. Corresponding results for the width of the opening between the lips are given in Fig. 3. In this case the lip oscillation was relatively large, and the lips sometimes closed for a short time during each oscillation cycle.

Figure 4 shows the motion of the two lips along y, the direction perpendicular to the net flow, where we plot the y coordinate of the center of each lip edge. Here $y = 0$ corresponds to the center of the channel between the lips; when the lips are in their equilibrium positions, each lip edge is 0.3 mm from the center of the channel. In Fig. 4 we expand the time scale slightly to show how the motions of the lips are synchronized as the lips move apart and then together (and nearly touch) each cycle. Figure 4 also shows that the amplitude of the motion is slightly larger for the upper lip in this case. We found that the lip oscillations were sometimes slightly asymmetric even though the lips themselves were nominally symmetric. There are several possible causes for this symmetry breaking. (1) While the lip parameters (mass, stiffness constant, and dimensions) were the same, the overall geometry was not perfectly symmetric as the instrument was placed slightly off center in the computational volume. (2) Two coupled oscillators that are degenerate can, if they interact, form two coupled modes with slightly different frequencies. (3) Such symmetry breaking could also be produced by small asymmetries in the numerical algorithm. The reason for the asymmetry found here is not clear and will be investigated in the future. We did observe that the best symmetry was generally found when the lip oscillations were large and with the lip frequency close to a resonant frequency of the bore.

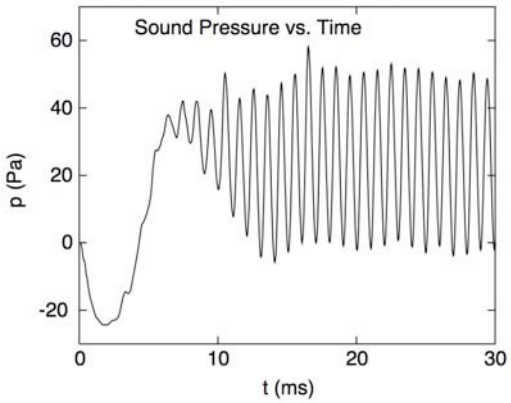


Figure 2. Sound pressure as a function of time at a location outside the trumpet and off axis relative to the bore. Lip parameters were $m = 4 \times 10^{-6}$ kg, $k = 125$ N/m, and $Q = 3$. The blowing speed at the center of the channel leading to the mouth was 200 m/s.

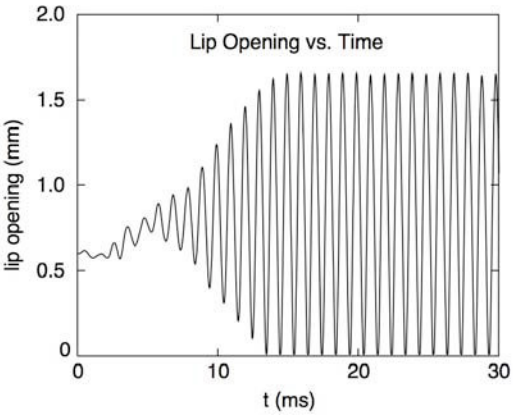


Figure 3. Variation of the width of the lip channel with time, for the simulation in Fig. 2.

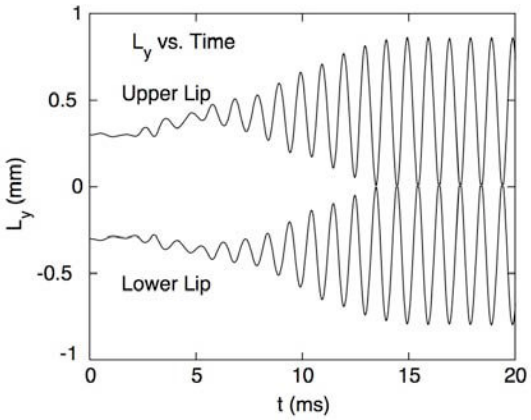


Figure 4. Motion of the upper and lower lips along the y direction (perpendicular to the net flow) for the simulation in Figs. 2 and 3.

Figure 5 shows the lip motion in a different way, plotting the lip displacement along y (the direction perpendicular to the net flow) versus the displacement along x (the direction parallel to the net flow) after the steady state oscillation was reached. While there is no indication of time in Fig. 5, the oscillations of the two lips were very close to being mirror images with respect to the $y = 0$ axis (Fig. 4). The lips thus nearly touch when x~+0.1 mm. The equilibrium lip positions are at $x = 0$ so the lips spend most of each cycle blown open, as expected, although they do swing back a small distance counter to the direction of net flow.

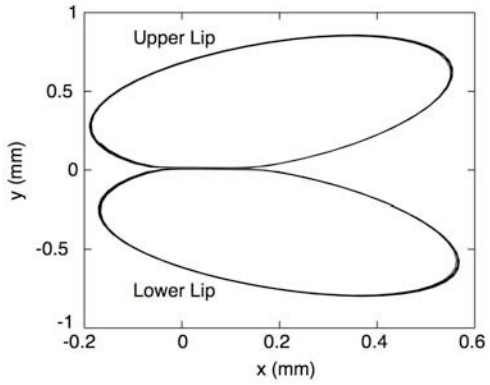


Figure 5. Motion of the center of the edge of each lip during successive oscillation cycles. Motion along y corresponds to squeezing and stretching of the lips, while motion along x corresponds to swinging motion parallel to the direction of net flow. Compare with Fig. 5 in [2].

4. THE CASE OF ASYMMETRIC LIPS

The results in Figs. 2-5 were all obtained with lips that were symmetric, that is, with the same mass, stiffness, and dimensions, and we found that the lip motion was approximately symmetric. To the best of our knowledge, all previous trumpet modeling has assumed symmetric motion of the lips. However, the lips of a real player are not symmetric, and our use of the Navier-Stokes equations to treat the air motion and pressure in three dimensions within the instrument allows us to readily consider the case of non-symmetric lips. In general one would expect both the mass and stiffness to be different for the upper and lower lips of a real player, leading at least to different oscillation amplitudes (as is indeed found experimental studies of the lip motion for real players of brass instruments [9-13]). For simplicity we will in this section consider just two cases: (1) lips with the same masses and different stiffness constants, and (2) lips with the different masses and the same stiffness constants. In both cases the two lips will be taken to have the same dimensions as in model considered above, and all other dimensions of the model were the same.

We first consider case (1), with lips having the same mass as in Figs. 2-5 but different stiffness factors k. The ratio of the stiffness factors was chosen to be 2, so since the natural frequency of a lip varies as $k^{1/2}$, the natural frequencies of the two lips differ by about 40% in this case. Figure 6 shows results for the width of the lip opening as a function of time; we find an oscillation at a single frequency that is approximately the same as the resonant frequency of the bore and the same as seen in Figs. 2-5, suggesting that coupling through the resonant motion of the surrounding air has caused the lips to move in synchrony. This is confirmed in Fig. 7, which shows the lip motion along y; indeed, the results in Figs. 6 and 7 are quantitatively nearly identical to the case with symmetric lips in Figs. 4 and 5. Note also that the two lips again move in phase; that is, they reach their maxima (of the lower lip) and minima (of the upper lip) at the same time. This synchrony and phase locking of the motion of the two lips is maintained for very long times (100 ms or more) which is easily long enough to reveal a 40% difference in the frequencies of the two lips if that were the case. Our results show that the Q factor is low enough that the lips are able to lock to the frequency of the conical bore.

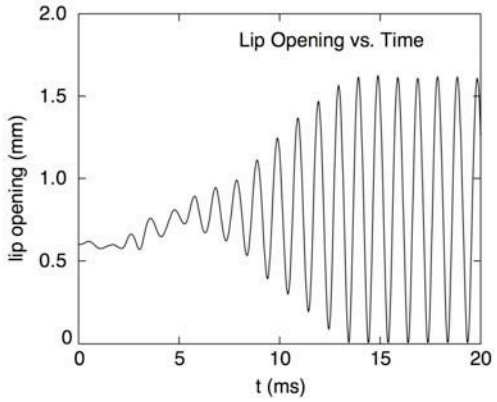


Figure 6. Variation of the width of the lip opening with time, for the case of lips with different stiffness factors: $m = 4\text{x}10^{-6}$ kg, $k_{\text{lower}} = 125$ N/m, $k_{\text{upper}} = 60$ N/m.

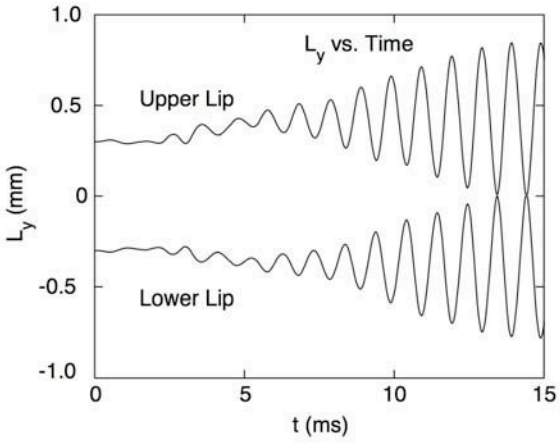


Figure 7. Lip motion along y for lips with different stiffness factors as given in the caption for Fig. 6.

Figures 8 and 9 show results for lips with different masses and the same stiffness. The ratio m(upper)/m(lower) = 2, so the natural frequencies of the two lips differed by about 40%, as for the case considered in Figs. 6 and 7. In Figs. 8 and 9 we use the same scales as in Figs. 6 and 7 to emphasize that the results are now very different from the behavior found with symmetric lips. The resulting lip vibration amplitude is much smaller than found with either symmetric lips or with lips of the same mass but different stiffness. We also see from Fig. 9 that the two lips do not even lock to a common frequency.

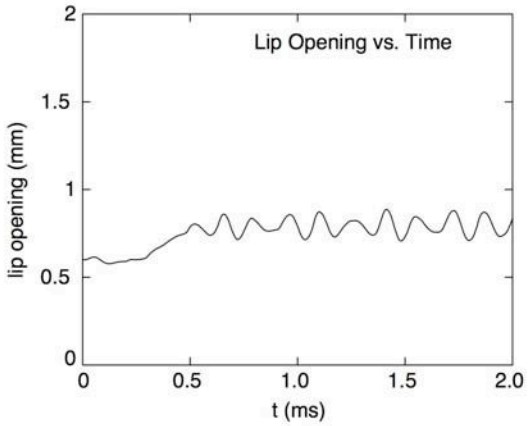


Figure 8. Variation of the width of the lip opening with time, for the case of lips with different masses and the same stiffness: $m_{\text{lower}} = 4\text{x}10^{-6}$ kg, $m_{\text{upper}} = 8\text{x}10^{-6}$ kg, $k = 125$ N/m.

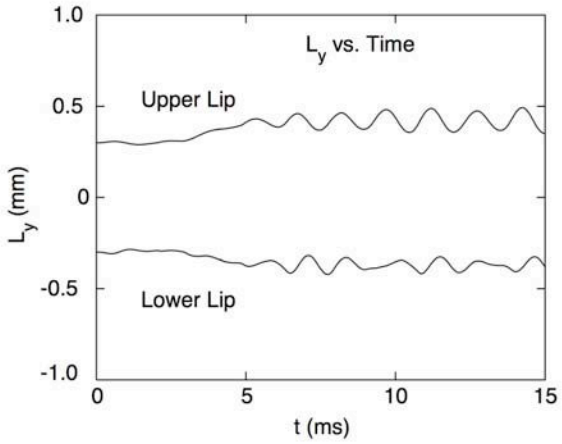


Figure 9. Lip motion along *y* for lips with different masses as given in the caption for Fig. 8.

5. CONCLUSIONS

This paper presents first results for a trumpet model in which the air is treated using the compressible Navier-Stokes equations in three dimensions. For the case of symmetric lips we find results that are similar in many respects to that found previously in modeling that treated the air flow in a more approximate way. We have also obtained results for the more realistic case of asymmetric lips. For lips in which the asymmetry is due to different stiffness factors, the lips move in synchrony and in phase at a single frequency producing what should be a good musical tone. When a roughly comparable asymmetry is present with lips of differing mass, the lips move slightly out of phase resulting in a much reduced sound amplitude.

Our results suggest many avenues for future work. The behavior for asymmetric lips with masses that are more nearly equal will be studied to determine if and when synchrony occurs, and the case of lips which differ in both mass and stiffness must be explored. It will also be interesting to extend our work to include more complicated lip models, like those used in studies of phonation [14]. In addition to the lip motion, our modeling approach yields the spatial and temporal variations of the pressure and air velocity throughout the mouthpiece-lip region, which will be compared with experiments in future work. It is also possible to make movies of the lip motion and flow velocity, and these will be presented elsewhere.

6. ACKNOWLEDGEMENTS

I thank S. Adachi for a useful discussion. This work was supported by US National Science Foundation grant PHY1513273. The computations were carried out at the Rosen Center for High Performance Computing at Purdue University and with computational facilities supported by the Office of Information Technology at Auburn University.

REFERENCES

[1] S. J. Elliott and J. M. Bowser, "Regeneration in brass instruments," *J. Sound Vib.* **83**, 181-217 (1982).

[2] S. Adachi and M. Sato, "Trumpet sound simulation using a two-dimensional lip vibration model," *J. Acoust. Soc. Am.* **99**, 1200-1209, (1996).

[3] M. Campbell, "Grass instruments as we know them today," *Acta Acust. Acust.* **90**, 600-610 (2004).

[4] J. A. Kemp, S. Bilbao, J. McMaster, and R. A. Smith, "Wave separation in the trumpet under playing conditions and comparison with time domain finite difference simulation," *J. Acoust. Soc. Am.* **134**, 1395-1406 (2013).

[5] N. Giordano, "Simulation studies of the recorder in three dimensions," *J. Acoust. Soc. Am.* **135**, 906-916 (2014).

[6] C. S. Peskin, "Flow patterns around heart valves: a numerical method," *J. Comput. Phys.* **10**, 252-271 (1972).

[7] R. Mittal and G. Iaccarino, "Immersed boundary methods," *Annu. Rev. Fluid Mech.* **37**, 239-261 (2005).

[8] J. Pederzani and H. Haj-Hariri, "A numerical method for the analysis of flexible bodies in unsteady viscous flows," *Int. J. Numer. Methods. in Eng.* **68**,1096-1112 (2006).

[9] D. W. Martin, *J. Acoust. Soc. Am.* **13**, 305-308 (1942).

[10] D. C. Copley and W. J. Strong, *J. Acoust. Soc. Am.* **99**, 1219-1226 (1996).

[11] S. Yoshikawa and Y. Muto, *Acta Acustica United with Acustica* **89**, 145-162 (2003).

[12] S. Bromage, M. Campbell, and J. Gilbert, *Acta Acustica United with Acustica* **96**, 603-613 (2010).

[13] H. Boutin, N. Fletcher, J. Smith, and J. Wolfe, "Relationships between pressure, flow, lip motion, and upstream and downstream impedances for the trombone," *J. Acoust. Soc. Am.* **137**, 1195-1209 (2015).

[14] F. Alipour, C. Brücker, D. D. Cook, A. Gömmel, M. Kaltenbacher, W. Mattheus, L. Mondeau, E. Nauman, R Schwarze, I. Tokuda, and S. Zörner, "Mathematical models and numerical schemes for the simulation of human phonation," *Curr. Bioinform.* **6**, 323-343 (2011).

Analysis-Synthesis of Saxophone Input Impedances via Recursive Parallel Filters

Esteban Maestre,[1†] *Julius O. Smith,*[2] *Gary P. Scavone*[1]

[1]Computational Acoustic Modeling Laboratory, McGill University
[2]Center for Computer Research in Music and Acoustics, Stanford University
[†]esteban@music.mcgill.ca

ABSTRACT

We introduce an analysis-synthesis framework for simulation of wind instrument input impedances as recursive digital filters in parallel form, derived from measured data. For analysis, we iteratively minimize the error between the frequency response of a saxophone input impedance measurement and that of a synthetic impedance model constructed from a digital filter structure akin to the discretization of a modal expansion. The input impedance model is used to derive a reflectance filter and used for efficient sound synthesis via a reed model based on non-linear scattering.

1. INTRODUCTION

As efficient sound synthesis schemes, most digital waveguide models [1] of wind instruments approximate their air columns as being cylindrical. In a typical digital waveguide model, the air column of an ideal instrument constructed from a cylindrical pipe and a bell can be represented by a pair of delay lines simulating pressure wave propagation inside the pipe, and a termination that includes two digital filters: one that lumps frequency-dependent propagation losses and dispersion, and another one emulating the frequency-dependent bell reflectance. In these efficient schemes, the reed-valve end termination of the pipe is often modeled via a non-linear scattering element that is interfaced to the air column model through decomposed pressure traveling waves P^+ and P^-, respectively going into and reflected back from the pipe input interface. Approximations with conical elements are possible but often result in inharmonic resonance structures that are difficult to tune for sound synthesis.

To account for realistic, non-ideal instrument air column shapes, one could treat the entire air column as a resonant load, observe its linear behavior from frequency-domain experimental data, and design an *air column load* impedance filter model $Z(z)$ (i.e., an input impedance filter) for simulation so that the reflected pressure wave $P^-(z)$ can be obtained from the incident wave $P^+(z)$ via $P^-(z) = R(z)P^+(z)$, with $R(z)$ being a digital reflectance model derived from $Z(z)$. In a previous work [2], a frequency-domain measurement of an air column input impedance is used to construct a discrete time reflection function $r[n]$ that is suitable for a traveling wave numerical scheme based on convolution. Authors propose a method to evade the time-aliasing and other numerical problems that naturally arise from estimating $r[n]$ via inverse Fourier transform of a frequency-domain measurement signal.

The motivation of this work lies in avoiding the aforementioned problems by proposing a methodology for translating input impedance measurements directly into recursive digital filters of moderately low order, with the added advantage that efficiency is improved with respect to discrete convolution. We construct a digital model $Z(z)$ of the air column input impedance by means of recursive digital filters in parallel form, so that a corresponding digital reflectance $R(z)$ can be connected to a reed-valve non-linear scattering model under the traveling wave formulation. Although applicable to other wind instruments, the design of such impedance filter is carried out here by fitting its coefficients to a measured saxophone input impedance obtained from experimental data.

Alto saxophone input impedances were measured using a six-microphone probe, calibrated with three non-resonant loads, via a least-mean square signal processing technique as described in [3]. In Figure 1 we display the magnitude of an alto saxophone input impedance measurement (grey curve), obtained for a 'Bb3' fingering, and normalized to the characteristic wave impedance of the air column input. As the resonance amplitudes decrease with frequency, the normalized impedance tends to a value of 1, i.e., total transmission.

2. THE INPUT IMPEDANCE AS A RECURSIVE PARALLEL FILTER

By attending to the resonance structure of input impedance measurements, it appears clear that we can benefit from proposing a digital filter formulation that is akin to the discretization of a modal expansion. In fact, in a previous paper we addressed a similar problem for the case of bridge input admittances in string instruments [4]. Here we construct an input impedance digital model $Z(z)$ by creating a digital modal basis over which impedance measurements are projected. Each m-th impedance modal basis function $H_m(z)$ is defined as

$$H_m(z) = \frac{1 - z^{-1}}{(1 - p_m z^{-1})(1 - p_m^* z^{-1})}, \qquad (1)$$

which corresponds to a two-pole resonator with one added zero at DC. The resonator is defined by a pair of complex conjugate poles p_m and p_m^*, which we relate to the corresponding modal frequency f_m and bandwidth β_m by $2\pi f_m / f_s = \angle p_m$ and $\beta_m = -\log|p_m|/\pi$, with f_s being the sampling frequency in Hertz. The impedance model $Z(z)$ is formulated in parallel as

$$Z(z) = \sum_{m=1}^{M} (r_{0,m} + r_{1,m} z^{-1}) H_m(z), \qquad (2)$$

where $r_{0,m}$ and $r_{1,m}$ are real-valued coefficients that allow control of both the amplitude and the phase of the the m-th resonator. The main reason behind the choice for our parallel resonator structure is that, while enabling the control of the relative phase between resonators, it imposes a gain of zero at DC irrespective of the coefficients $r_{0,m}$ and $r_{1,m}$.

3. IMPEDANCE FILTER DESIGN

Departing from a frequency-domain measurement of the target impedance $\hat{Z}$, the problem of designing the coefficients of a recursive digital filter model (2) that approximates the measurement can be stated as the minimization of an error measurement $\varepsilon(Z, \hat{Z})$ between the measurement and the model, with parameters being a vector $\boldsymbol{p} = \{p_1, \cdots, \omega_M\}$ of complex poles each corresponding the m-th resonator of of the model, and vectors $\mathbf{r}_0 = \{r_{0,1}, \cdots, r_{0,M}\}$ and $\mathbf{r}_1 = \{r_{1,1}, \cdots, r_{1,M}\}$ of respective numerator coefficients. We solve this problem via sequential quadratic programming [5]. At each iteration only pole positions are exposed as the variables to optimize: once they are decided, zeros (i.e., numerator coefficients) are constrained to minimize an auxiliary quadratic cost function, resulting in a simple closed-form solution. The positions of the poles are optimized iteratively: at each step, an error function is successively evaluated by projecting the target frequency response over a basis of frequency responses defined by the pole positions under test. We add a set of linear constraints to guarantee feasibilty and to ease convergence. This optimization routine is devised as an extension to the digital filter design technique proposed in [6].

3.1. Impedance measurement pre-processing

As it can be observed in the grey curves of Figure 1, the high-frequency region of an impedance measurement tipycally presents artifacts caused by noise and limitations of the measurement method. It is important to remove those artifacts so that the target normalized impedance effectively tends to 1 as frequency increases. This is needed to help the fitting process in providing an impedance model design for which the normalized impedance also tends to 1 in the high frequency region; otherwise, a derived air column reflectance filter would deliver reflected pressure waves with significant energy around Nyquist, and therefore cause undersired behaviors in the reed-valve non-linear scattering model. To this end, we perform cross-fading between the normalized impedance measurement and a constant value of 1, as illustrated in Figure 1, where we display the magnitude response of a pre-processed measurement.

3.2. Initialization and problem statement

We initialize the model parameters via finding a set of initial pole positions by attending to the magnitude response of the impedance measurement. First, resonance peak selection in the low-frequency region is carried out through an automatic procedure that iteratively rates and sorts spectral peaks by attending to a salience descriptor. For estimating modal frequencies, three magnitude samples (respectively corresponding to the corresponding maximum and its adjacent samples) are used to perform parabolic interpolation around selected peaks. For estimating bandwidths, the *half-power* rule [1] is applied using a linear approximation. For the high-frequency region we spread an additional set of poles, uniformly distributed on a logarithmic frequency axis. This leads to a total M modes, each expressed in terms of complex pole angles and radiae in the z-plane.

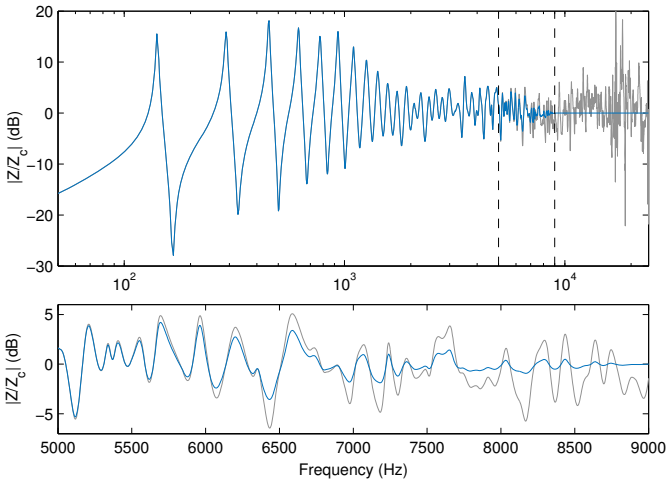


Figure 1. Magnitude response of an alto saxophone impedance measurement ('Bb3' fingering), normalized by the characteristic impedance of the input of the air column. Grey and blue curves are respectively used for raw and pre-processed data. Top: full band, with cross-fading region delimited by vertical lines. Bottom: cross-fading region.

We parametrize the initial set of M modes by representing each respective m-th complex pole pair in terms of its angle parameter $w_m = |\angle p_m|$ and its radius parameter $s_m = -\log(1 - |p_m|)$. This leads to two parameter sets: a set $\boldsymbol{w} = \{w_1 \cdots w_M\}$ of angle parameter values, and a set $\mathbf{s} = \{s_1 \cdots s_M\}$ of radius parameter values. With the new parametrization, we state the problem as

$$\begin{aligned} \underset{\boldsymbol{w},\mathbf{s}}{\text{minimize}} \quad & \varepsilon(Z, \hat{Z}) \\ \text{subject to} \quad & \mathbf{C}, \end{aligned} \tag{3}$$

where $\mathbf{C}$ is a set of linear constraints, and numerator coefficients have been left out as they are not exposed as variables in the optimization.

A key step before contraint definition is to sort the pole parameter sets so that linear constraints can be defined in a straightforward manner to ensure that the arrangement of poles in the unit disk is preserved during optimization, therefore reducing the number of crossings over local minima. Elements in sets $\mathbf{w}$ and $\mathbf{s}$ are jointly sorted as pairs (each pair corresponding to a complex-conjugate pole) by ascending angle parameter w_m.

3.3. Constraint definition

From ordered sets $\boldsymbol{w}$ and $\mathbf{s}$, linear onstraints $\mathbf{C}$ are defined as follows. First, feasibility is ensured by $0 \leq s_m$ and $0 \leq w_m \leq \pi$. Second, to aid convergence we constrain the pole sequence order in set $\boldsymbol{w}$ to be respected. This is expressed by $w_{m-1} < w_m < w_{m+1}$. Moreover, assuming that initialization provides an already trusted first solution, we can bound the search to a region around the initial pole positions. This can be expressed via the additional inequalities $w_m^- < w_m < w_m^+$ and $s_m^- < s_m < s_m^+$, where '$-$' and '$+$' superscripts are used to respectively indicate lower and upper bounds.

3.4. Error computation

At each i-th step of the optimization, the error $\varepsilon(Z, \hat{Z})$ is measured as follows. Given K samples of the target impedance frequency response $\hat{Z}(\omega)$ and the set $\mathbf{p}$ of M complex poles defining the modes at the i-th step, numerator coefficient vectors $\mathbf{r}_0$ and $\mathbf{r}_1$ can be obtained via least-squares by

$$\underset{\mathbf{r}}{\text{minimize}} \|\mathbf{H}\mathbf{r} - \hat{\mathbf{z}}\|^2, \tag{4}$$

where $\mathbf{r} = [\,\mathbf{r}_0^T\, \mathbf{r}_1^T\,]^T$ is a real-valued vector; $\hat{\mathbf{z}}$ contains K samples of the target frequency response at frequencies $0 \leq \omega_k < \pi$, i.e., $\hat{z}_k = \hat{Z}(\omega_k)$; and $\mathbf{H}$ is a $K \times 2M$ matrix of basis vectors constructed as

$$\mathbf{H} = [\mathbf{h}_{0,1} \cdots \mathbf{h}_{0,m} \cdots \mathbf{h}_{0,M}\, \mathbf{h}_{1,1} \cdots \mathbf{h}_{1,m} \cdots \mathbf{h}_{1,M}\,]$$

with column vectors $\mathbf{h}_{0,m}$ and $\mathbf{h}_{1,m}$ containing the sampled frequency responses of $H_m(z)$ and $z^{-1}H_m(z)$ respectively. With numerator coefficients, we evaluate the frequency response of the model and compute the error measure as the l_2-norm of the difference vector, i.e., $\varepsilon(Z, \hat{Z}) = \|\mathbf{H}\mathbf{r} - \hat{\mathbf{z}}\|^2$.

3.5. Final solution

Once poles have been optimized, numerator coefficients of model (2) are found by solving again problem (4). In Figure 2 we display the frequency response of an example impedance model, obtained from a normalized saxophone impedance measurement after cross-fading to a value of 1 above 8 kHz. Although in principle the model (2) is not guaranteed to be positive-real, fitting to measurements of positive-real functions generally provides positive-real designs. This is important for stability of the sound synthesis model because we need to construct an air column reflectance model that is passive.

4. REALIZATION AS A DIGITAL WAVEGUIDE REFLECTANCE

We treat the air column as a digital impedance load to which we interface the reed model via a pair of decomposed traveling waves. From the input impedance model (2), we construct a reflectance that keeps the state of the air column as a resonating element, and allows us to obtain reflected waves from its interface. Following the digital waveguide formulation for loaded parallel junctions [1], we can compute the scalar flow

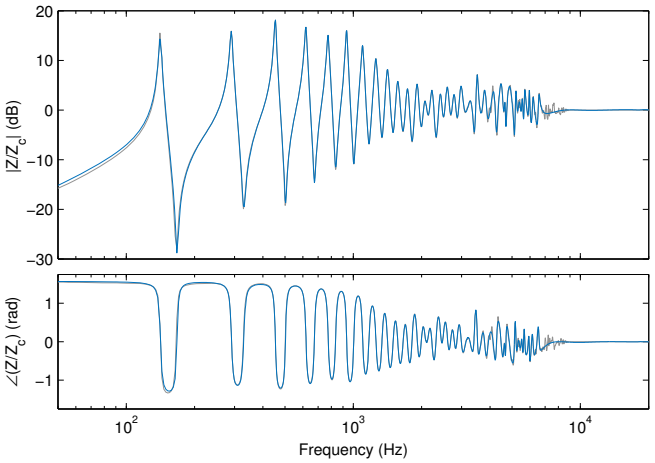


Figure 2. Frequency response of an example saxophone impedance model ('Bb3' fingering), normalized by the characteristic impedance of the input of the air column. Grey and blue curves respectively used for the pre-processed measurement and the model ($M = 35$).

$U(z)$ at the input of the air column solely from the input pressure wave $P^+(z)$ as

$$U(z) = \frac{2Y_c P^+(z)}{1 + Y_c Z(z)} \tag{5}$$

where Y_c is the characteristic admittance of the input of the air column, and $Z(z)$ is the digital impedance model. From the flow $U(z)$, it should be straightforward to compute the scalar pressure $P(z)$ at the input of the air column via

$$P(z) = Z(z)U(z). \tag{6}$$

Finally, from the air column pressure $P(z)$ it is possible to obtain the (reflected) outgoing pressure wave $P^-(z)$ by means of

$$P^-(z) = P(z) - P^+(z). \tag{7}$$

Because the formulation of the impedance model (2) presents a parallel structure that we want to maintain, inverting $Z(z)$ as it appears in equation (5) is impractical. To overcome this problem in the realization of the reflectance, we reformulate the impedance in a similar fashion as we proposed in [4] for the admittance of a string instruments. First, we rewrite each resonator $H_m(z)$ of equation (2) as

$$H_m(z) = 1 + z^{-1} H_m^p(z), \tag{8}$$

with

$$H_m^p(z) = \frac{c_{0,m} + c_{1,m} z^{-1}}{1 + a_{1,m} z^{-1} + a_{2,m} z^{-2}}, \tag{9}$$

$c_{0,m} = -1 - a_{1,m}$, and $c_{1,m} = -a_{2,m}$. Note that denominator coefficients are related to pole radius and angle by $a_{1,m} = -2|p_m|\cos(\angle p_m)$ and $a_{2,m} = |p_m|^2$. We now rewrite the impedance model as

$$Z(z) = B_0 + z^{-1} B_1 + z^{-1} H_0(z) + z^{-2} H_1(z), \tag{10}$$

with

$$B_0 = \sum_{m=1}^{M} c_{0,m}, B_1 = \sum_{m=1}^{M} c_{1,m}, \quad (11)$$

$$H_0(z) = \sum_{m=1}^{M} c_{0,m} H_m^p(z), H_1(z) = \sum_{m=1}^{M} c_{1,m} H_m^p(z). \quad (12)$$

With this new formulation, we rewrite (5) and (6) as

$$U(z) = \frac{2Y_c P^+(z) - z^{-1} Y_c \big(B_1 + H_0(z) + z^{-1} H_1(z)\big) U(z)}{1 + Y_c B_0} \quad (13)$$

and

$$P(z) = B_0 U(z) + z^{-1} \big(B_1 + H_0(z) + z^{-1} H_1(z)\big) U(z). \quad (14)$$

It is important to point out that now the parallel structure appears in the numerator terms $H_0(z)$ and $H_1(z)$, making possible its implementation. Moreover, $H_0(z)$ and $H_1(z)$ can be jointly implemented as a sole bank of parallel resonators. Finally, it is worth mentioning that the term $z^{-1}\big(B_1 + H_0(z) + z^{-1} H_1(z)\big) U(z)$ appears in equations (13) and (14) but does not need to be implemented twice–once it has been computed to obtain $U(z)$ via equation (13), it can be reused to compute $P(z)$ via equation (14).

5. SOUND SYNTHESIS

We construct an efficient sound synthesis scheme by interfacing our air column reflectance model and a modified version of the reed scattering model used in [7] as follows. At each iteration, two main computations are interleaved: the reed scattering update and the air column reflectance update. During the reed scattering update, the differential pressure driving both the reed motion and the reed channel flow relation (see [7]) is first computed as the difference between the mouth pressure and the value of the scalar air column pressure obtained in the previous reflectance update (see Section 4). Then, the pressure wave obtained from the reed scattering is used to feed the next reflectance update. In Figure 3 we display the reed channel flow (see [7]) and the scalar air column pressure P of a synthesis example obtained from an alto saxophone 'Bb3' fingering model, driven by a piecewise linear mouth pressure signal. It can be observed how, during the initial transient, the model first jumps to an upper octave and then transitions to its nominal regime. This model runs about 40 times faster than real-time in one core of a current laptop computer.

6. OUTLOOK

Albeit still preliminary, results shed light on a promising route for efficient, yet realistic sound synthesis with potential applications both in rendering music and in analysing the timbre and playability of real air column prototypes. Among our planned future steps are a more complete exploration with impedances over the full range of notes and of different wind instruments, the development of strategies for transitioning between fingerings, and the modeling of measured radiativities.

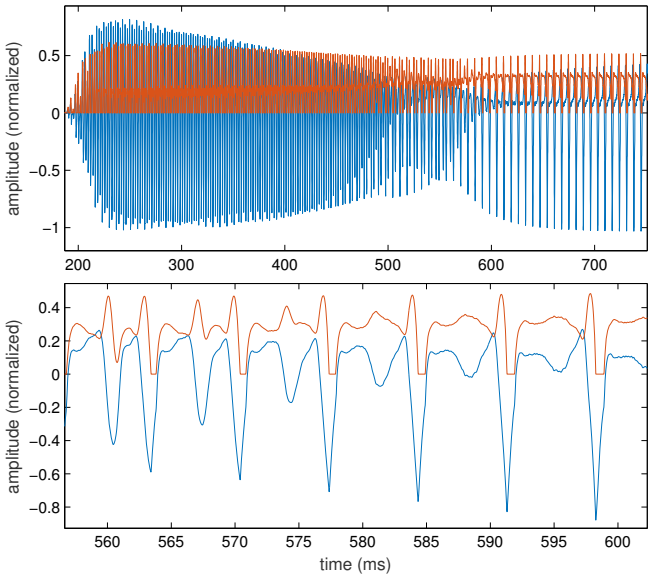


Figure 3. Synthesis example for an alto saxophone 'Bb3' fingering model, driven by a piecewise linear mouth pressure signal. Orange and blue curves respectively used for reed channel flow and air column pressure. Top: initial upper-octave transient. Bottow: detail of the transition towards the nominal regime.

REFERENCES

[1] J. O. Smith, *Physical Audio Signal Processing*. W3K Publishing, 2004. [Online]. Available: http://ccrma.stanford.edu/~jos/pasp/

[2] B. Gazengel, J. Gilbert, and N. Amir, "From the measured input impedance to the synthesis signal: where are the traps?" *Acta Acustica*, vol. 3, pp. 445–472, 1995.

[3] A. Lefevbre and G. P. Scavone, "A comparison of saxophone impedances and their playing behavior," in *Proc. of the Forum Acusticum Conference*, 2011.

[4] E. Maestre, G. P. Scavone, and J. O. Smith, "Digital modeling of bridge driving-point admittances from measurements on violin family instruments," in *Proc. of the Stockholm Music Acoustics Conference*, 2013.

[5] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 2006.

[6] E. Maestre, G. P. Scavone, and J. O. Smith, "Design of recursive digital filters in parallel form by linearly constrained pole optimization," *IEEE Signal Processing Letters*, vol. 23:11, pp. 1547–1550, 2016.

[7] G. P. Scavone and J. O. Smith, "A stable acoustic impedance model of the clarinet using digital waveguides," in *Proc. of the International Conference on Digital Audio Effects*, 2006.

Synthesis of the Human Voice Using a Simplified Physical Model

Xavier Pelorson,[†] *Anne Bouvet, Annemie Van Hirtum*

Gipsa-lab, France

[†]xavier.pelorson@gipsa-lab.grenoble-inp.fr

ABSTRACT

Synthesis of the human voice by means of physical models is difficult and challenging. This can be understood by the complexity of the physics of the human voice, involving the biomechanics of the vocal folds, the fluid mechanics of the glottal airflow and the resulting fluid-structure and aeroacoustical interactions. Direct Numerical Simulation of all these physical effects is still impossible at present time. Further, recent attempts show that days of computation on supercomputers are necessary to simulate a few seconds of signal, which makes DNS useless for synthesis purpose.

In this contribution we present a simplified physical model for sound synthesis which aims at making a trade-off between physical complexity and computational costs. Our goal is thus to identify what are the major physical effects involved during voicing and to model them theoretically. The relevance and accuracy of the theoretical models is tested against experiments on mechanical replicas of the vocal folds and of the vocal tract.

The mechanics of the tissues is described using the model of Avanzini et al (2002). Theoretically, this model accounts for a single mechanical mode and, compared with other models (like the popular two-mass model proposed by Ishizaka and Flanagan (1972)) has the advantage to require a limited set of parameters, most of them being inferable by the experiment. A particular attention will be given on the physical description of the collision of the vocal folds. Different collision models will be presented and discussed.

Concerning the fluid mechanics of the airflow through the glottis, it will be shown that the closure of the vocal folds is also critical. Except when a very high fundamental frequency of oscillation is expected, viscous effects appear to be predominant. Different flow models of increasing complexity will be presented and discussed in term of their accuracy but also in terms of their computational costs.

Lastly, the acoustical coupling between the vocal folds and the sub and supra glottal cavities will be discussed. This presentation will be illustrated with sound samples obtained with the synthesizer. It will be shown that what can be thought as very small details of the physics can have spectacular effects on the synthesized signals.

Experimental Investigations of the Effects of Fast Slide Movements in the Trombone on the Behaviour of an Artificial Mouth.

Amaya López-Carromero ,[1†] *Jonathan A. Kemp,*[2] *D. Murray Campbell* [1]

[1]University of Edinburgh
[2]University of St. Andrews
[†]A.Lopez-Carromero@sms.ed.ac.uk

ABSTRACT

Previous studies have shown that fast trombone slide movements, such as the ones performed by the player during glissando, can introduce infrasound pressure variations approximately proportional to the slide acceleration, and which can reach 100Pa. These transients introduce instabilities and potentially affect the performance. To characterise the behaviour of the lips under these conditions, an artificial mouth coupled to a King Trombone is used. The physical effects on the artificial lips are evaluated by performing inwards and outwards glissandi at different speeds. This evaluation is carried out using an optical light transmission setup to track the opening of the lips, pressure transducers which measure variations at the mouth cavity, mouthpiece and water key of the instrument, and a laser sensor able to measure the displacement of the slide over time. These results may help in the implementation of transients in physical modelling synthesis and in the educational ambit by guiding instrumental teachers and performers.

1. INTRODUCTION

The introduction of the slide represented a considerable evolutionary step in brass wind acoustics. Before its development around the fifteenth century, the playable notes on tubular brass instruments without side holes were limited to the approximate natural harmonic series available from the rigid geometry of the resonator.

However groundbreaking the invention of the slide was, it left performers facing new challenges such as limited execution speed and the difficult task of performing a legato between notes which are only available in widely different slide positions [1]. It could be argued that these aspects are intrinsic to the instrument, but in some cases the performer will try to overcome these limitations by performing extremely quick slide gestures, at speeds which can surpass 2 ms^{-1} [2]. These fast gestures have been proven to introduce infra sound pressure fluctuations which travel towards the mouthpiece of the instrument [3] but the effects of this over pressure on the dynamics of the lip reed are still unknown.

1.1. Pumping effect of playing gestures

When the trombone slide is either extended or contracted, sudden variations in volume and consequently in what would otherwise be a static pressure, occur. These disturbances in the conditions of the air column are propagated, and can be predicted analytically by assuming that the transformations taking place inside the resonator are adiabatic [3].

The generated pressure differential can be approximated as

$$p_{gen} \approx -2v\rho c, \tag{1}$$

where v is the velocity of the slide, ρ is the equilibrium density of air and c is the speed of sound. This pressure pulse propagates inside the instrument following the clindrical waveguide model described in [3], and for terminations resembling a wall, such as the one described in this same work, the reflection from the mouthpiece end of the instrument can be assumed to be constant, positive and close to one. Assuming a negative reflection from the bell and taking into account time delays for propagation, the observed pressure in the slide is approximately proportional to the negative of the slide acceleration. [3].

However, the validity of this model is under question when the mouthpiece end termination is exchanged for a complex dynamic system such as the player's lips, as the lips may not be completely closed and may be moving under the influence of mean flow and pressure differentials.

While the general behavior of lips at acoustic frequencies has been successful modeled in the literature using oscillators with varying numbers of degrees of freedom, the influence of infra-sound components on the behavior of the lips and reflection of such signals from the lips is a topic requiring further investigation before it can be integrated into such models.

Even for infra sound components, way below the resonant frequencies of the lips, their effect could introduce a deviation from predictions based on the assumption of a rigid, flat termination [3].

2. EXPERIMENTAL METHODOLOGY

2.1. Technical setup

Experiments were carried out in an acoustics laboratory at James Clerk Maxwell building in the University of Edinburgh. The instrument, a King trombone, was connected to an artificial mouth via a transparent mouthpiece, and all the equipment was mounted on an optical rail to allow measurements of the lip opening via a light transmission setup such as the one described in [4], [5]and [6]. Very detailed descriptions and

sketches of the artificial mouth can be found in [5]. The artificial mouth is an airtight rectangular box with glass walls where one of the lateral sides receives the artificial rubber lips which are coupled to the mouthpiece. The box has three more openings; one to which a flat response loudspeaker is screwed, another orifice to introduce a microphone and an entrance connected to an air blower. The instrument plays when this blower increases the pressure upstream from the lips and they oscillate at a frequency which can be sustained by the resonant air column in the instrument coupled to them.

A general schematic showing all the apparatus used can be seen in figure 1. The instantaneous lip opening was measured using a class IIB laser, located at one of the ends, and a custom made linear photometer at the other end of the optical rail. The laser beam was expanded via a divergent lens with a focal length of -8 mm and a diameter of 30 mm. After crossing a polarising filter used to adjust the intensity, the beam reached the artificial mouth and was converged at the other side by a lens of both focal distance and diameter 50mm on to the photometer's sensor.

The pressure in the mouthpiece was measured using a high pressure prepolarised G.R.A.S. 1/2″ microphone, while the pressure in the artificial mouth was measured with a prepolarised G.R.A.S. 1/4″ microphone. The radiated sound was also measured outside the bell, using a prepolarised G.R.A.S. 1/2″ microphone.

Inside the slide section, the pressure was measured with a Sensor Technics HCXM020D6V differential pressure transducer connected via a 5cm long rubber pipe to the water key. This transducer could not measure large negative components which limited the usability of the data collected with it.

The physical extension of the slide during the performance of the glissandi was measured using a Baumer OADM 20I4471 laser sensor. The laser sensor was attached firmly to the mouthpiece end of the instrument and remained static during the whole process. The laser beam was intercepted by a small reflector attached to the mobile part of the slide.

All transducers requiring external power supply, such as the laser and the pressure sensors, were powered via a custom made supply box.

All signals were acquired via a program written with Signal Express software, using N.I interfaces, which included a PXIe-4464 and a PXI-4461 card, both mounted on a PXIe-1073 chassis.

All microphones were equipped with TEDS technology, thus ensuring direct calibration when using N.I. interfaces. The rest of the transducers were calibrated during the postprocessing, using information from their technical specifications documentation. For frequencies below the ranges included in the documentation the transducers were tested in the laboratory by evaluating their response to an artificially generated sine played through the loudspeaker. This response was found to be satisfactory for the means of this work, although the precise calibration has not yet been determined.

All data was processed in MATLAB, including up sampling corrections to the slide position measurements and adequate filtering when necessary. This software was also used to generate all plots included in this work.

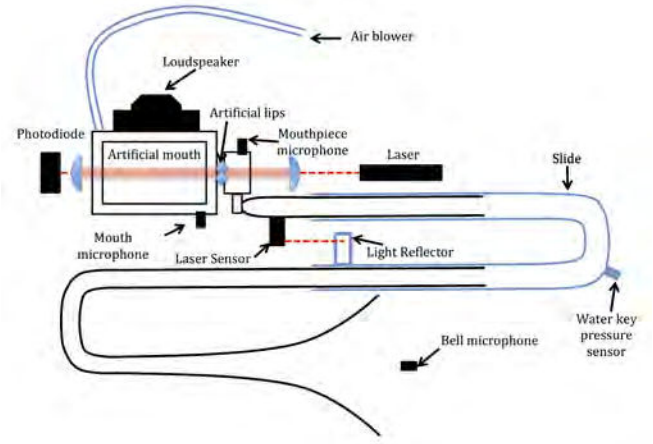


Figure 1. Schematic showing the light transmission setup, allowing the lip opening to be tracked when the slide undergoes a fast gesture. The blue section of the instrument indicates the moving slide.

2.2. Experiment design

Measurements were made with one of the authors performing glissando slide movements in different directions and at different speeds.

First the artificial lips were adjusted both in spacing and internal pressure, so that the embouchure was able to play a musical note around the second position of the trombone slide when the air blower injected a constant flow into the artificial mouth. Then, the slide was displaced from this position either inwards (towards the first position) to perform an outwards movement with positive velocity, or outwards (towards the third position) to perform an inwards gesture with negative velocity, so that during the recording of the gestures the lips would not stop buzzing at any time.

Variables measured included the pressures inside the artificial mouth, the mouthpiece and the water key, the lip opening (via the light transmission method previously mentioned), the position of the slide over time, and the pressure level and pitch of the radiated sound.

3. RESULTS

3.1. The effect of pumping on the radiated sound

Slide displacements in the trombone change the resonating length of the instrument and therefore affect the efficiency of the coupling between the artificial lips and the resonator. This coupling has been investigated thoroughly in the literature [7] where the artificial lips are shown to play above and below their main mechanical resonances due to a large Q factor that allows them to "lip" a note. This deviation from the position of their main mechanical resonance, however, involves a reduction in the amplitude of their oscillation, which can be observed in all the measurements presented here at the beginning and end of

each gesture, and which explains the differences in radiated pressure when the slide positions are stable. However, fast slide gestures have the potential of shifting this amplitude due to hysteresis in the lip dynamics, which no longer are able to adapt in the linear fashion shown in [7], causing unexpected variations in the radiated amplitude.

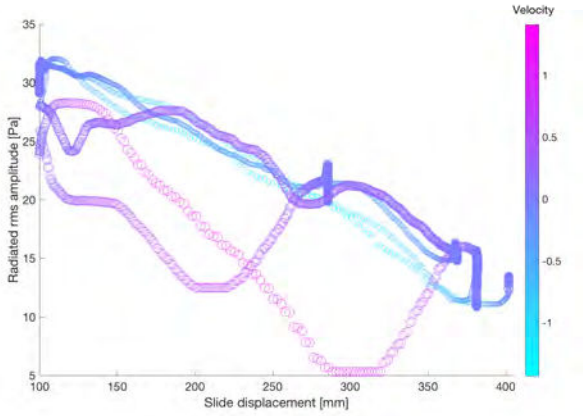


Figure 2. Six slide movement events featuring four fast slide glissandi (two moving outwards and two moving inwards), and two slow slide movements, one in each direction. Positive velocities, in the magenta range, are shown with a big circular marker while the negative velocities, in the cyan range, are shown with a similar, but smaller marker.

Figure 2 shows six representative events featuring four fast slide glissandi (two moving outwards and two moving inwards), and two slow slide movements (one in each direction). Here, the signal amplitude is shown to drop significantly when doing a downward glissando (positive velocity). This is an asymmetrical effect, as large negative velocities, with essentially the same accelerations as a function of slide position, don't produce a noticeable adjustment in amplitude when compared to the slow motions.

3.2. The effect of pumping on the lip behaviour

Further insight on this effect and on the behaviour of the lips is given by Figures 3 and 4 where two of the fastest slide movements are analysed in greater detail.

When the pressure in the mouthpiece decreases due to a fast slide extension, the amplitude of the lip oscillation reduces and the pressure in the mouth builds up temporarily. In fast slide retractions, the opposite effect takes place in the mouthpiece initially. However, the lips also observe a reduction in the amplitude of their oscillations, with a similar overpressure in the mouth taking place.

This can be explained by understanding the over pressures acting on the lips as destabilising forces which shift the lips from their position of equilibrium, and cause them to become less efficient oscillators. After the gesture is finished and the standing wave is restablished, the lips start oscillating normally again with the expected amplitude for that slide position.

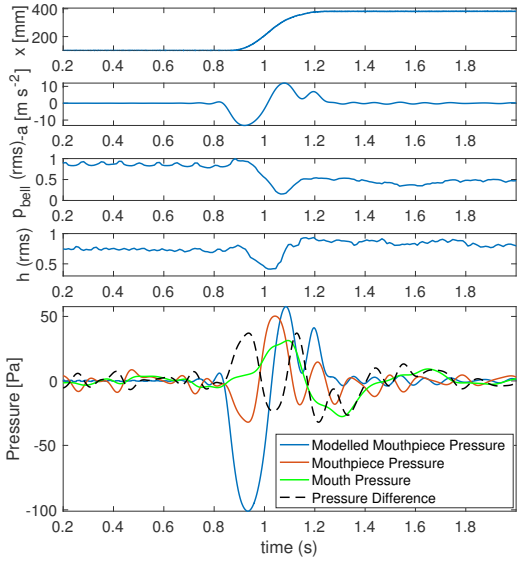


Figure 3. Effects of an outwards fast glissando. From top to bottom: Slide displacement x. Negative slide acceleration $-a$. Rms radiated signal normalised by its maximum value. Rms lip opening normalised by its maximum value. The lower subplot shows the mouthpiece pressure predicted by the cylindrical waveguide model with a rigid termination at the lips (blue line), the measured low-pass filtered mouthpiece (orange line) and mouth (green line) pressures, and the difference between mouth and mouthpiece pressures (dashed black line).

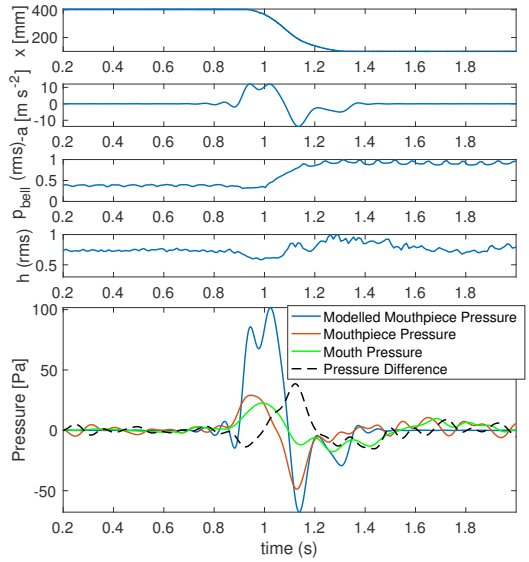


Figure 4. Effects of an inwards fast glissando. See: Figure 3

4. CONCLUSIONS AND FURTHER WORK

An extension of the cylindrical waveguide model seems necessary in order to include the effects of the lips, and properly predict the effect of glissando gestures on them. This extension could be achieved by substituting the currently stationary reflection coefficient at the mouthpiece end of the instrument by a time-varying term able to update in each time iteration and which is models the response of the lips under these infrasound components. Further characterisation of the mechanical properties of the lips at low frequencies is required to this end.

5. ACKNOWLEDGEMENTS

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ un- der REA grant agreement No. 605867 supporting the BATWOMAN ITN Project. Thanks to Alan Woolley for technical support and providing part of the apparatus.

6. REFERENCES

REFERENCES

[1] T. D. Rossing and N. Fletcher, "The physics of musical instruments," 1998.

[2] G. Ekdahl, "A simple model of the mechanics of trombone playing," 2001.

[3] J. Kemp, A. Lopez-Carromero, A. Woolley, and M. Campbell, "Experimental investigation of doppler shift and infrasound generation during wave propagation within the bore of the trombone during slide movement," *Proceedings of the Third Vienna Talk on Music Acoustics*, vol. 16, pp. 9–14, 2015.

[4] S. Bromage, "Visualisation of the lip motion of brass instrument players, and investigations of an artificial mouth as a tool for comparative studies of instruments," Ph.D. dissertation, University of Edinburgh, 2007.

[5] O. F. Richards, "Investigation of the lip reed using computational modelling and experimental studies with an artificial mouth," Ph.D. dissertation, University of Edinburgh, 2004.

[6] M. J. Newton, M. Campbell, and J. Gilbert, "Mechanical response measurements of real and artificial brass players lips," *The Journal of the Acoustical Society of America*, vol. 123, no. 1, pp. EL14–EL20, 2008.

[7] M. J. Newton, M. Campbell, J. Chick, and J. A. Kemp, "Predicting the playing frequencies of brass instruments," in *Proc. Forum Acusticum, Krakow, Poland*, 2014.

Relationships between Abdominal and Around-Lip Muscle Activities and Acoustic Features when Playing the Trumpet

Megumi Satou,[1†] *Tetsuro Kitahara,*[2] *Hiroko Terasawa,*[3] *Masaki Matsubara,*[3]

[1] Graduate School of Library, Information and Media Studies, University of Tsukuba
[2]Department of Information Science, Nihon University College of Humanities and Sciences
[3] Faculty of Library, Information and Media Science, University of Tsukuba
[†]megumi@slis.tsukuba.ac.jp

ABSTRACT

When playing the trumpet, embouchure and breath control are the most important factors in achieving stable performance. To investigate these factors in greater detail, we asked 11 amateur trumpet performers to play tones of different pitches, intensities, and durations and analyzed the variations of the activities of their abdominal muscles and the muscles around their lips, which were measured via surface electromyography at the participants' abdominal oblique muscle (AOM) and depressor anguli oris muscle (DAO). The measurements were conducted during the periods of preparing and sustaining the tones. Our main experimental results are as follows. (1) As the pitch increased, the activities of both the DAO muscle and AOM increased during both the preparation and sustain periods. (2) As the intensity increased, only the activity AOM increased in both periods. (3) As the tone length increased, the AOM activity increased, while the DAO muscle was not affected. (4) The DAO muscle and AOM activities were lower and higher, respectively, during the sustain period than they were during the preparation period.

1. INTRODUCTION

Both embouchure and breath control are important in playing brass instruments but most of existing studies focused only on embouchure [1, 2]. Embouchure refers to the use of the facial muscles to shape the lips to the mouthpiece of a brass instrument. The sound of the instrument is driven by the vibration of the lips, which are created by the rapid flow of breath from the gap between the two lips. Embouchure plays a significant role in creating the proper lip vibrations. Therefore, many researchers have investigated the activities of the muscles around the lips when playing a brass instrument [2-9].

Breath is also an important factor and is controlled by the activity of abdominal muscles. In particular, breathing during performance differs from breathing at rest due to the fact that larger volumes of the lungs must be used while performing. Player's abdominal muscles therefore work more actively to support the breath during performance. However, brass players' abdominal muscles have been investigated in only a few studies. Cossette *et al*. reported the relationship between lung capacity and muscular activity during flute performance [10], but they did not focus on brass instruments.

Nonetheless, it is informative to investigate brass players' abdominal muscles since the abdominal muscle activity begins before the instrument starts to sound. Rather, the activity during the preparation phase is important. Generally, performers prepare to exhale just before playing to achieve the proper tone pitches or intensities, which requires the use of the abdominal muscles. Consequently, the abdominal muscle activity during preparation is important in producing correct tones.

In this paper, we discuss the relationships between trumpet players' muscle activities and the acoustical features in the following regards: (1) how the activities of the abdominal oblique muscle (AOM) and depressor anguli orins muscle (DAO) change relative to variations in the acoutical features (e.g. pitches, intensities, and durations), (2) the differences between the activities of these muscles during the periods of preparing and performing a tone, and (3) whether these changes have different tendencies between the AOM and DAO. We aim to provide evidence regarding breath control while playing a brass instrument based on our analysus of the objectively measured muscle activities.

2. RELATED WORK

White and Basmajian reported on the relationship between trumpet performance proficiency and the activities of the orbicularis oris muscles in the upper and lower lips, the levator anguli oris muscle, and the DAO muscles of 18 trumpeters [4]. According to their study, muscle activity varied when the pitch or intensity changed, regardless of the player's proficiency. In addition, it was observed that the beginners had greater degrees of muscle activity in their upper lips than in the lower lips. The muscle activities of the novice players were also found to vary widely in general.

Matsukata et al. also analyzed the muscle activity around the lips during trumpet playing [5]. The amount of muscle activity was found to be greater when playing high pitches than when playing low pitches. In addition, the players who could perform over wider pitch ranges exhibited smaller muscle activity differences between the high and low pitch ranges.

Hirano et al. studied the skin surface movement and muscle activity around the lips before and after tone production in horn playing [6]. Ten proficient horn players were asked to played sounds of various pitches, intensities, and durations while wearing markers and electrodes around their lips. The authors analyzed the positions of the markers and electromyography (EMG) signals and found that the performers had great control over both their muscle activities and skin surface movements (i.e., that these activities and movements were performed with high stability), resulting in no differences between before and after tone production. These results suggest the importance of proper embouchure before tone production.

Moreover, Bianco et al. investigated the temporal evolution and interrelationships among intraoral pressure, mouthpiece force, and activity in two groups of facial muscles [7]. Three professional players were asked to perform 10 isolated C5 quarter notes and a series of 10 concatenated C5 quarter notes, using pp, mf, and ff dynamics. The results showed that the experimental conditions affected the amount and locations of the variations that occurred during performance as well as the intraoral pressure during tone production.

3. EXPERIMENT

3.1 Participants

Eleven amateur trumpet players (seven men and four women) with more than 4 yr (mean, 8.5 yr) of training in playing brass instruments participated in this study. Informed consent was obtained from all the participants, and the experiment was approved by the institutional review board at the University of Tsukuba.

3.2 EMG and sound data recording

We collected EMG signals from the DAO muscle and AOM on the right side of each player by using a wireless EMG logger (Logical Product Co.), myoelectric sensor (Oisaka Electronic Equipment Ltd.), and disposable electrode (Nihon Kohden Co.). Considering the area of each participant's face, the electrodes used to measure the participant's DAO muscle were cut to the appropriate sizes.

To perform EMG on each participant's DAO muscle, we attached electrodes obliquely from the jaw to below the corner of the mouth of the participant. To conduct EMG on the AOM in each case, electrodes were attached to each subject at the intersection between the umbilicus and the anterior axillary line, where the layer portions of the external and internal oblique muscles are located.

We also recorded the performed sounds by using a microphone with an audio interface (both by Harman International Japan Co., Ltd.). The microphone was placed 1 m from the bell of the instrument in each case. We detected the onsets of the performed sounds and used the onset data to extract the relevant EMG data periods.

3.3 Tasks

The participants played in a semi-anechoic chamber using the instruments that they usually played. Our stimuli consisted of six sequences, which covered the following pitches, intensities, and durations:

- pitch: B ♭ 3 (233.08 Hz), F4 (349.23 Hz), B ♭ 4 (466.16
- Hz), F5 (698.46 Hz), and B ♭ 5 (932.33 Hz)
- intensity: pp, mf, and ff
- duration: short note (one beat), long note (eight beats)

The sounds in each sequence were arranged randomly. Each participant played the set of six sequences three times. The order of the sequences was randomized in each set and for each participant. The tempo was set to 80 beats per minute and was shown to each participant by an illuminated electronic metronome (Yamaha). Each participant started playing when given a cue.

4. ANALYSIS OF EMG AND SOUND DATA

4.1 EMG signals

We removed the DC offset of each EMG signal and then filtered the signal using a fifth-order Butterworth band-pass filter, which had a pass-band of 30–400 Hz. Next, we calculated the root mean square (RMS) of the signal, using a window width of 300 ms.

We computed the mean EMG data during periods of 375 ms before the onset of sound production and 750 ms starting 3 s after the onset. In the following paragraphs, the former and latter periods are referred to as the "preparation" and "sustain" periods, respectively (Fig. 1). In each one-beat case, we computed the mean EMG data only for the preparation period.

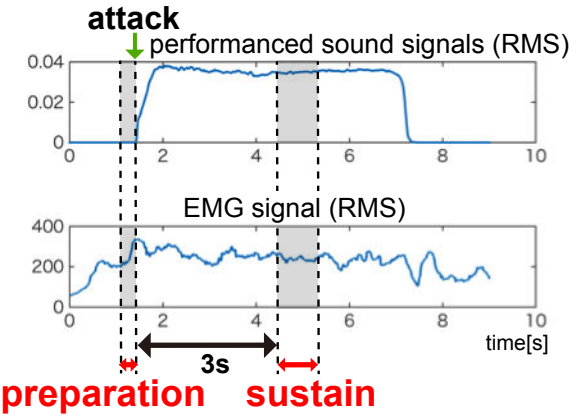


Figure 1. Preparation and sustain periods.

The EMG data were normalized to the muscle activity at the time of maximum voluntary contraction. In this report, we refer to the normalized values as %EMG values.

We calculated the average %EMG values using three trials for each participant and inspected the differences between these averages due to the variations of features such as pitch and intensity during the preparation period in the one- and eight-beat cases and during the sustain period in the eight-beat cases. In addition, to verify the variations caused by changes in duration, we compared the average %EMG values during the preparation period in the one- and eight-beat cases.

4.2 ANOVA

Three-factor analysis of variance (ANOVA) was used to examine the effects of pitch, intensity, and duration, and their interactions during the preparation period. Two-factor ANOVA was conducted to examine the effects of pitch and intensity during the sustain period. Furthermore, for the preparation and sustain periods in the eight-beat cases, three-factor repeated-measure ANOVAs were performed to examine the effects of pitch, intensity, and period (preparation and sustain). As mentioned in Section 3.3, the pitch and intensity categories were defined by performance instructions, rather than physical measures. Finally, post-hoc analyses with multiple comparisons, namely, t-tests with Holm's adjustments, were conducted.

5. RESULTS

Since the entire analysis is comprehensive and too lengthy to describe in its entirety in this paper, four key findings are described in the following sections.

5.1 Pitch and muscle activities

As the pitch of the performed sound became higher, the DAO muscle and AOM activities significantly increased during tone production just before and during the sustain period ($p < 0.001$ in each case). This trend is illustrated for the AOM in Fig. 2. Thus, the higher the pitch, the more active both the muscles around the lips and the abdominal muscles are.

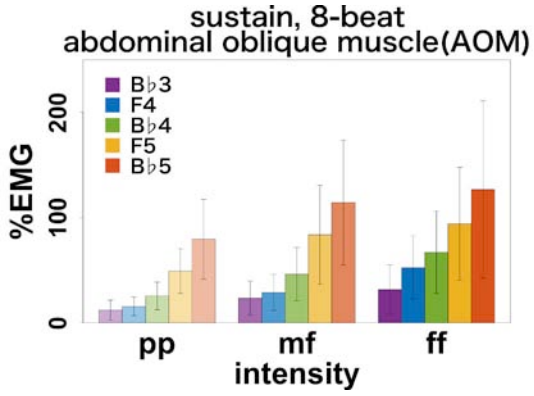


Figure 2. Mean and standard deviation (SD) %EMG values for the AOM during the sustain period in the eight-beat cases.

5.2 Intensity and muscle activities

As the intensity of the performed sound increased, the AOM activities changed, although the DAO muscle didn't increase during both the preparation and sustain period. This trend is observable for the AOM in Fig. 2. The intensity only significantly affected the AOM activity during both the preparation ($p < 0.01$) and sustain ($p < 0.001$) periods. This result suggests that performing a tone with a greater intensity requires more abdominal muscle activity than does performing a tone with a lower intensity, although the activities of the muscles around the lips are not affected by intensity changes.

5.3 Duration and muscle activities

When the duration of the performed sound increased, the activity of the DAO muscle did not change during the preparation period, while that of the AOM decreased with statistical significance ($p < 0.05$), as shown in Fig. 3. These findings suggest that only the abdominal muscles are activated while performing longer tones, but that the muscles around the lips are not affected by the tone duration.

5.4 Muscle activities during preparation and sustain periods

The DAO muscle activity was found to be greater during the preparation period than during the sustain period, while that of the AOM was greater during the sustain period than during the preparation period. In the long-tone cases, the AOM activity differed significantly between the sustain and preparation periods ($p < 0.001$ in each case), while also being affected by the

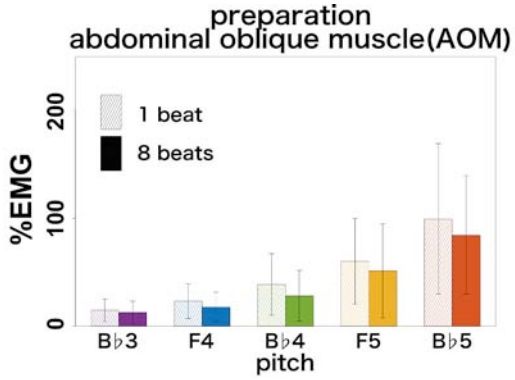


Figure 3. Mean and SD %EMG values for AOM during the preparation period in the one- and eight-beat cases.

pitch and intensity. In addition, interactions were observed between pitch and intensity ($p < 0.01$) and between intensity and period ($p < 0.001$). The pitch, intensity, and period all significantly affected the DAO muscle activity as well (with $p < 0.001$, $p < 0.05$, and $p < 0.01$, respectively).

These results suggest that when performing long tones, the abdominal muscles are relaxed during the preparation period and activated during the sustain period. On the other hand, the muscles around the lips are activated during the preparation period and relaxed during the sustain period.

6. DISCUSSION

6.1 Muscle activity variations with tone pitch

The muscle activity differences observed when the participants performed different pitches would have been caused by the corresponding intraoral pressure changes. Intraoral pressure refers to (1) the feeling of resistance caused by breath flowing out rapidly from the gap between the lips that is reflected from the mouthpiece, or (2) the pressure applied to the oral cavity in order to open the lips against the pressure from the air column inside the instrument.

When performing a high-frequency tone, the oral cavity pressure must be greater than it is when performing a low-frequency tone [9]. The increase in pressure inside mouth causes the AOM and DAO muscle to become more active. It seems that to realize a greater flow of breath, the AOM must support the diaphragm more strongly, resulting in a higher oral cavity pressure. Meanwhile, the DAO muscle becomes active to maintain stable embouchure by tightening the cheeks against the increasing oral cavity pressure.

We observed less activity in both the AOM and DAO muscle when lower pitches were played than when higher pitches were played, in order to realize the same interval differences. In other words, more muscular activity (i.e., more energy) is required in higher pitch ranges to perform the same jump between pitches than is required in lower pitch ranges (e.g., a jump from B ♭ 4 to B ♭ 5 requires more muscle activity than a jump from B ♭ 3 to B ♭ 4 does). Therefore, the muscle activity required for performance is considered to increase exponentially (yet within the realistic range) as the pitch increases.

6.2 Muscle activity variations with tone intensity

The muscle activity differences observed when the participants played tones of different intensities would also have been

affected by the corresponding intraoral pressure changes. When playing a tone on a brass instrument, the performer's intraoral pressure depends on the intensity with which the tone is played [6, 7]. Specifically, the intraoral pressure increases as the intensity of breathing increases. Consequently, the AOM activity must be greater to support stronger breathing.

The amount by which the sound intensity of a brass instrument changes depends on the expiratory volume. A larger difference in sound intensity requires a larger expiratory volume than a smaller difference in sound does, resulting in a more rapid decrease in lung volume. The AOM contributes to realizing this decrease in lung volume by pushing up the diaphragm.

6.3 Muscle activity variations with tone duration

Significant variations were not observed in either the DAO muscle or AOM activity, but the mean %EMG value for the AOM was greater in the short-tone cases than in the long-tone cases. We suspect this difference was caused by the participants' consciousness of the required tone duration. Many of the participants reported in follow-up interviews that the degree of psychological stress involved in playing the target sounds differed between the short- and long-tone cases. Higher psychological stress was involved in performing the short tones. We suspect that the AOM activity during the preparation period was higher when the psychological stress was greater.

6.4 Muscle activity differences between before and after production

The DAO muscle activity during the sustain period was less than it was during the preparation period. This finding contradicts that of Hirano et al., who observed no differences between the preparation and sustain periods in horn performance [6].

Unnecessary straining while sustaining a tone leads to fatigue, making it difficult to perform long tones [7]. Therefore, it is desirable to minimize the amount of muscle activity during the sustain period when performing.

Moreover, we observed that the AOM activity was greater during the sustain period than during the preparation period. This difference could have been caused by the AOM being used to push up the diaphragm to squeeze the last remaining air out from the lungs.

7. CONCLUSIONS

In this study, we focused on the activities of the DAO muscle in embouchure and the AOM in breathing during trumpet playing and analyzed the relationships between the levels of muscle activity and acoustic features such as the pitch, intensity, and duration of the performed sound. Based on the experimental results, the following conclusions can be drawn:

1. As the pitch of the performed sound increases, both the muscles around the lips and the abdominal muscles become more active.
2. As the intensity of the performed sound increases, the abdominal muscles become more active. However, the activity of the muscles around the lips does not change with the intensity.
3. When performing longer tones, the abdominal muscles become more active, while the muscles around the lips are not affected.
4. When performing long tones, the abdominal muscles are relaxed during the preparation period and activated during the sustain period. On the other hand, the muscles around the lips are activated during the preparation period and relaxed during the sustain period.

We wonder if performers were aware of these complex muscle activity patterns themselves. Since performers need to conduct sequences of such coordinated movements smoothly, we anticipate that EMG measurements would be beneficial to performers, since such measurements would enable performers to recognize their own muscle activities and make connections between their physical movements and the resulting sounds. Thus, EMG feedback systems for practice sessions may be useful to performers who wish to understand their sound production processes more objectively.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Numbers 16K16180 and 26240025. We would like to thank to the people who participated in the experiment.

REFERENCES

[1] P. Farkas, *The art of brass playing: a treatise on the formation and use of the brass player's embouchure*. Bloomington: Brass Publications, 1962.

[2] B. C. Papsin, L. a. Maaske, and J. S. McGrail, "Orbicularis oris muscle injury in brass players," *The Laryngoscope*, vol. 106, no. 6, pp. 757–760, 1996.

[3] P. W. Iltis and M. W. Givens, "EMG characterization of embouchure muscle activity: Reliability and application to embouchure dystonia," *Medical Problems of Performing Artists*, vol. 20, no. 1, pp. 25–35, 2005.

[4] E. R. White and J. V. Basmajian, "Electromyography of lip muscles and their role in trumpet playing," *Journal of Applied Physiology*, vol. 35, no. 6, pp. 892–897, 1973.

[5] S. Matsukata, H. Terasawa, M. Matsubara, and T. Kitahara, "Muscle Activity in Playing Trumpet: The Dependence on the Playable Pitch Region and the Experience of a Non-trumpet Brass Instrument Player," in *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*, Stockholm, Sweden, 2013, pp. 529–534.

[6] T. Hirano, K. Kudo, T. Ohtsuki, and H. Kinoshita, "Orofacial muscular activity and related skin movement during the preparatory and sustained phases of tone production on the French horn." *Motor Control*, vol. 17, no. 3, pp. 256–72, 2013.

[7] T. Bianco, V. Freour, I. Cossette, and F. Bevilacqua, "Measures of facial muscle activation, intra-oral pressure and mouthpiece force in trumpet playing," *Journal of New Music Research*, vol. 41, no. 1, pp. 49–65, 2012.

[8] N. H. Fletcher and A. Tarnopolsky, "Blowing pressure, power, and spectrum in trumpet playing," *Journal of the Acoustical Society of America*, vol. 105, no. 2, pp. 874–881, 1999.

[9] I. Cossette, P. Monaco, A. Aliverti, and P.T. Macklem, "Chest wall dynamics and muscle recruitment during professional flute playing," *Respiratory Physiology & Neurobiology*, vol. 160, no. 2, pp. 187–195, 2008.

Vibroacoustics of Brass Instruments: Analysis of the Coupling Mechanisms Thanks to a Double Bell Experiment

Mathieu Secail-Geraud, Lévy Leblanc, Joël Gilbert, François Gautier[†]

Laboratoire d'Acoustique de l'Université du Maine
[†]francois.gautier@univ-lemans.fr

ABSTRACT

The wall vibrations of brass instruments are coupled in a complex manner to the internal acoustic pressure. In the trombone case, the vibration level is particularly high in the bell region, where the wall is thin and where the radius is large. The importance of those vibrations for the musician is a long standing question, which is today still open. In this paper, an analysis of the involved vibroacoustic couplings is proposed thanks to a double bell experiment.

A double bell consists in the association of two concentric bells, defining an annular region which can be filled with water. Such a bell is associated to a blowing machine to produce sounds. This configuration permits to vary the structural modal basis without modifying the excitation conditions or the internal geometry of the vibrating waveguide. The coincidence conditions which are observed between acoustical and structural modes can be controlled by the water level, which leads to variations of the vibroacoustic couplings.

In this configuration, different physical mechanisms are identified : 1/ perturbation of the input acoustic impedance by the vibroacoustic coupling between the walls and the air column, 2/ direct external radiation by the walls, 3/ vibroacoustic coupling localized at the bell extremity, 4/ transmission of the body vibrations to the mouthpiece and to the musician's lips. The relative importance of these various mechanisms is discussed.

Hybrid Modeling of Cristal Baschet Sounds

François Gautier,[1†] *Frederic Ablitzer,*[1] *Jean-Loïc Le Carrou,*[2] *Liujian Zhang*[1]

[1]Laboratoire d'Acoustique de l'Université du Maine
[2]Institut Jean le Rond d'Alembert
[†]francois.gautier@univ-lemans.fr

ABSTRACT

The Cristal Baschet is a musical instrument created in the fifties by F. and B. Baschet. The produced sounds are resulting from friction-induced vibrations due to the interaction between the wet fingers of the player and a large number of glass rods called glass bows. Each glass bow is connected to a resonator consisting in an assembly of two beams linked by a mass which is tuned in order to reach a targeted pitch. Vibrations are then transmitted to thin plates acting as sound radiators. Analysis of the glass bow instability using an High speed Camera (7000 frames per second) shows that a slick-slip mechanism occurs at the finger/rod interface : sliding and sticking phases follow each other while the glass rod exhibits a solid body motion coupled to the bending motion of the resonator. An hybrid model is proposed to describe the instability: the finger/rod interaction is described analytically by a simplified friction law while the motion of the resonator (beam assembly) is described by its measured input mobility. The musician gesture is described by the velocity of finger and the shear force applied on the glass bow. Such characteristics are used to provide time domain simulations of the sounds produced by a Cristal Baschet. Cristal Baschet makers are sometimes reporting the fact that when the mobility of one beam assembly is too low, the sound is difficult to produce. Such configurations are investigated : the model is used to investigate the role played by the input mobility of the resonator. The analysis of the conditions needed to reach periodic oscillations is performed using the time domain simulations and a linear stability analysis, which permits a Schelleng type diagram for the Cristal Baschet to be proposed.

The Birl Physical Model

Dharit Tantiviramanond[1†], *Jeffrey Snyder*[1]

[1]Princeton University
[†]dharit.tan@gmail.com

ABSTRACT

This paper outlines a project to couple a woodwind physical model that incorporates a continuously variable tonehole model [1, 2] with a new experimental electronic wind controller called the Birl, and to develop an algorithm to automatically adjust the parameters of the model based on the desired tuning of the final instrument.

1. INTRODUCTION

The Birl project [3, 4, 5] led by Jeff Snyder, attempts to create an expressive electronic wind instrument that allows continuously variable tonehole control and more complex embouchure detection than that available on existing commercial wind controllers. This particular contribution to the Birl project focuses on implementing a digital waveguide physical model [6] as a strategy for mapping the continuous tonehole control to sound.

Toneholes, especially in open-holed acoustic wind instruments such as the bansuri, shanai, or recorder, provide a wide range of sonic possibilities through the variety of ways that a performer can interact with them. Completely covering a tonehole effectively increases the length of the instrument bore, dropping the pitch to a lower frequency, while opening a tonehole does the opposite. Partially covering it results in interpolation between these two pitch boundaries. A combination of higher breath pressure and certain coverage values can produce unusual acoustic effects such as multiphonics, which create the perception of multiple pitches sounding at once. Different fingerings that produce the same pitch may nevertheless produce varying timbres, allowing for the use of an effect called a "timbral trill", in which the performer switches quickly between two fingerings to create an oscillation in tone color. These techniques are all possible due to the physical characteristics of the tube and the unique way that air pressure waves blown into it interact with the bore cavity, and how the toneholes dynamically alter these properties. However, much of this flexibility is lost in current commercially available electronic wind instruments such as the Akai EWI or the Yamaha WX series, which have only binary inputs for the toneholes (open or closed state). These instruments output a MIDI pitch value based on the combination of open or closed keys, which simplifies their communication with synthesis engines.

If one desires to take variations in continuous toneholes into account, then this mapping from the state of the toneholes to the synthesis becomes more complex. There have been some projects by other authors that involved the creation of electronic wind instruments with continously variable toneholes, most notably the PIPE, by Scavone [7], and the Epipe, by Hughes, Cannon and Modhrin [8]. In the case of the Birl, several mapping strategies have been attempted, including rule-based methods, and using trained neural networks to create interpolated states between particular finger positions and a desired pitch output [4]. This paper details our experiment in using physical modelling synthesis to achieve useful mapping of continuous tonehole parameters to sound. Compared to the neural net or rule-based approaches, this solution comes with many limitations in the possibilities of pitch mappings, as it is limited to creating a model that is similar to a physical instrument with a limited number of toneholes in particular locations. Despite these limitations, a physical modelling approach allows for very interesting and acoustically-inspired sonic results, and very intuitive performance possibilities, as mentioned in an early paper describing the addition of tonehole models to digital waveguides [9]. For instance, we were particularly interested in allowing the execution on a digital woodwind instrument of extended acoustic wind instrument techniques such as half-holing, cross-fingering, overblowing, and multiphonics.

One downside inherent in a physical modelling approach that uses a more complex model with multiple toneholes, when compared to a simpler model that simply adjusts the length of a single tube, is the difficulty of setting the parameters of the model to achieve the desired tuning results for the pitches produced when the toneholes are closed and opened. Therefore, another goal of this project was the development of an algorithm to simplify the procedure of setting these parameters so that a musician could define any arbitrary musical scale and a model would be calculated to achieve that result. The goal is an electronic instrument that retains many interesting performance capabilities of acoustic instruments, but is not subject to the same physical constraints.

2. IMPLEMENTATION

2.1. Physical Model

The implementation of the Birl physical model is based on the BlowHole STK (Synthesis Toolkit) class [10], which is a single tonehole implementation of a digital waveguide algorithm with woodwind-like components. The reed table, tonehole filter, and termination reflection filter are drawn directly from BlowHole. The tonehole model was set up as a pole-zero filter, and the equations relating the filter coefficients and the tonehole radius were implemented as part of the tonehole

initialization according to [1, 2]. Furthermore, a three-port scattering junction as in [2] was also in place.

For the body, BlowHole uses an optimized single delay line (SDL) implementation, but we found it easier to conceive the model as a bidirectional waveguide so we converted it. Information on converting between SDL and bidirectional waveguide models can be found in [11]. Then, we removed the register vent for simplicity and extended the model to include 11 toneholes, to match the number of tonehole sensors on the Birl controller.

A design decision was made to use only integer-length delay lines and to correct for the resulting rounding error by adjusting the diameter of the next tonehole. This was done both to avoid signal distortion and for performance benefits. Fractional delays involve interpolation functions to yield an approximation of a continuous input signal. Even with complex interpolation functions, a small amount of signal distortion is inevitable. Interpolation also invariably incurs more computational overhead. These two factors are insignificant for a small number of delays, but in a model that comprises a series of delays feeding into each other like the Birl, distortion and computational work accumulate rapidly.

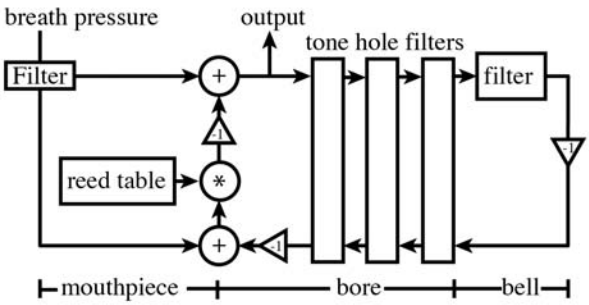


Figure 1. A basic schematic for the digital waveguide synthesis used in the Birl physical model [12].

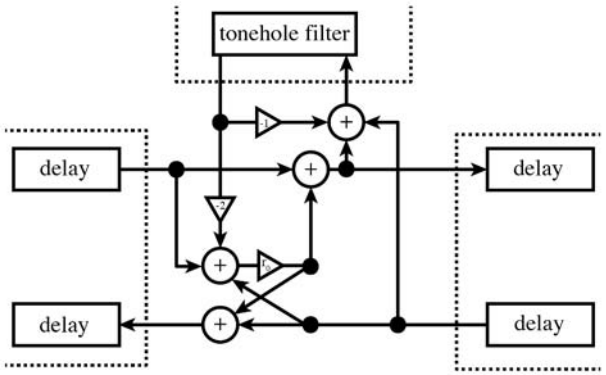


Figure 2. Three-port scattering junction. [2]

2.2. Tuning Algorithm

In order to allow the user to specify arbitrary tunings for the instrument, an algorithm was implemented that relates the desired tuning or scale to the delay line lengths and tonehole radii. A system of equations is given in [13] that determine the parameters of a delay line tube section and tonehole filter that must be appended to the model in order to produce each desired pitch.

The Birl model is a closed-open tube, so its output frequency is related to the tube length according to

$$F_1 = \frac{c}{4L} \tag{1}$$

where L is the effective tube length, F_1 is the fundamental frequency, and c is the speed of sound. Since air pressure waves actually travel a small distance beyond the open end of a tube, the discrepancy between the effective length and the cut (actual) length of the tube must be calculated. This difference is the end correction of the tube (Δl_T).

In a closed-closed tube $\Delta l_T = 0$, since the pressure waves reflect exactly at the ends of the tube. A closed-open tube of cut length l_T will produce a fundamental frequency equal to that produced by a closed-closed tube of the corresponding effective length L_S, so $L_S = l_T + \Delta l_T$. The following equation is used as an estimate of Δl_T:

$$\Delta l_T = 0.3 d_1 \tag{2}$$

where d_1 is the bore diameter.

The same principles apply to toneholes as well. The length from the mouthpiece end of the tube to the center of any given tonehole i is equal to the effective length of the tube when all toneholes before tonehole i are closed. This per tonehole cut length is labeled l_H, and the corresponding effective length $L_{S(h)}$. Their relationship is

$$\Delta l_H = z L_{S(h)} \tag{3}$$

$$z = 0.5g\sqrt{1 + 4\frac{L_{B(h)}}{gL_S}} - 0.5g \tag{4}$$

$$L_{B(h)} = (l_H + d_H)(\frac{d_1}{d_H})^2 - 0.45 d_1 \tag{5}$$

d_H refers to the tonehole diameter, and $L_{B(h)}$ is a variable that does not have a physical manifestation, but refers to "the length of a fictitious duct that has the same diameter as the main tube, and the same conductivity as the tonehole" [13]. g is the interval ratio between the current tonehole and the previous one (closer to the mouthpiece).

To solve for the position and radius of a tonehole, (3), (4), and (5) must be combined with (1) to find suitable values for l_H and d_H. $L_{S(h)}$ can be determined from the desired pitch using (1). Then, a reasonable starting value must be chosen for either d_H or l_H which can be used to solve (3).

Algebraic manipulation of (3), (4), and (5) yields:

$$L_{B(h)} = \frac{(L_{(h)} + 0.5gL_{S(h)} - l_L)^2}{gL_S} - \frac{gL_S}{4} \quad (6)$$

$$d_H = \frac{d_1^2}{L_{B(h)} + 0.45d_1} \quad (7)$$

The general algorithm for tuning the model at a tonehole position given a desired frequency for that tonehole is outlined here. Note that this algorithm must be run iteratively for each tonehole in the model.

1. Calculate the effective length $L_{S(h)}$ for this tonehole.
2. Calculate the interval ratio g by dividing two ratios. The numerator is the ratio of the current tonehole pitch to the fundamental, and the denominator is the ratio of the previous tonehole pitch to the fundamental.
3. Start with a reasonable diameter value d_H that is above a minimum threshold. Compute $L_{B(h)}$ using equation (6).
4. Solve for z using equation (4).
5. Compute l_H using $l_H = L_{S(h)} - zL_{S(h)}$. Round down to the nearest integer.
6. If l_H is the same length as a previous delay line, then increase oversampling rate and restart algorithm for entire model.
7. Compute d_H using equation (7). If d_H is too small, decrement l_H by 1 and repeat steps 4-5.

Tuning the fundamental is simpler as $L_{B(h)}$ is not involved. The integer tube length can be calculated directly from $L_S = l_T + \Delta l_T$ and (2).

The inclusion of minimum threshold values for d_1 and d_H is the result of observing detuning effects that occur when tonehole diameter values are too small. Intuitively, small radii may reflect an air pressure wave entirely. In the digital domain, this translates to the filter's coefficient values reflecting the signal completely, eliminating the effect of the filter. The exact minimum threshold at which unwanted detuning begins to occur is unknown, but values of 1.0 (in sample lengths) for both variables were deemed adequately effective in the Birl model.

Step 6 uses oversampling to work around the problem of delay line length clashes. If the user specifies two frequencies that are close enough together such that their corresponding wavelengths are within 1 sample length of each other, the tuning algorithm will return the same delay line length for both. Oversampling increases the granularity of the model so that the difference threshold at which this problem occurs is divided by the oversampling rate.

3. EVALUATION

3.1. Physical Model Evaluation

A formal user study of this research has not been conducted. However, as part of the evaluation process, a professional saxophonist was invited to informally explore the Birl while it was connected to the physical model. He noted that half-holing and cross-fingering felt natural, and multiphonics generally felt similar to the way that they are achieved in a saxophone, by using a "somewhat unpredictable embouchure/air stream" in conjunction with certain cross-fingerings. He was also able to achieve overblowing, but this was not realistic enough to perform on-demand. In acoustic wind instruments, overblowing occurs when the player blows firmly into the mouthpiece, but in the Birl a particular fingering configuration was also needed. Overall, the acoustics felt "very authentic." However, he also mentioned that the instrument felt " a little too sensitive" and that it was hard to know when the model would start squeaking. The biggest area of improvement he suggested was in providing the user with more control over the sound output by the model. Recordings of the model producing these effects can be found on this project's online git repository [14].

3.2. Tuning Algorithm Evaluation

While the tuning algorithm produced scales that sounded close to the intended results, there remain several problems. The most obvious defect of the current Birl model is an absolute pitch offset that is unsolved at the time of writing. When tuning the fundamental, the result is invariably sharp of the expected pitch. One possible explanation is equation (2). Forster acknowledges that this is not exact, but rather a "good estimation" for the relationship between d_1 and Δl_T, and "an exact equation for the end correction of open and closed tubes does not exist because the end correction depends slightly on wavelength" [13].

Relative tuning is also inaccurate, although less so: toneholes closer to the open end of the bore (lower-pitched) tend to be sharper. We have discovered that two bores with equivalent length, one with a closed tonehole and one simply contiguous, will not produce the same pitch. The addition of the tonehole filter slightly increases the output pitch. The compounding of this error may be the cause of the relative tuning error for toneholes further down the bore.

Another flaw in the current model is that oversampling rates above 2X result in high-frequency artifacts. Therefore step 6 in the tuning algorithm in Section 2.2 will only work once. The exact cause of this issue is unknown at time of writing, but we are exploring the possibility that the tonehole model needs to be adjusted for higher sampling rates beyond simply using the higher sample rate in coefficient calculation.

4. CONCLUSION AND FURTHER WORK

The Birl physical model is a proof-of-concept that certain unique acoustic properties of woodwinds can be achieved by using a digital waveguide with continuously variable toneholes coupled with a controller that allows for continuous tonehole control. However, its current drawbacks prevent it from being ready for a performance setting. The source of the global and relative tuning errors in the tuning algorithm, as well as the oversampling issues, need to be determined and resolved before the instrument is ready for performance. We also need

Specified Pitch (Hz)	**Output Pitch (Hz)**	**Absolute Error (cents)**	**Relative Error (cents)**
440.00	462.20	85.22	21.95
493.88	513.90	68.79	5.52
554.37	578.10	72.56	9.29
587.33	612.90	73.78	10.51
659.25	687.05	71.51	8.24
739.99	765.70	59.13	-4.14
830.61	858.50	57.18	-6.09
880.00	916.50	70.36	7.09
987.77	1019.20	54.23	-9.04
1108.73	1150.00	63.27	0.00

Table 1. Absolute and relative tuning error between one set of desired input frequencies and the actual frequencies produced by a model with parameters calculated by the tuning algorithm. The relative error is calculated in reference to the highest pitch, as that tonehole has the least amount of toneholes in front of it, and so should have the least error.

to build into the model an effective way for the performer to control overblowing.

One approach we are currently exploring as a solution to the tuning error is the use of pitch-detection within the tuning algorithm. During initialization, the error between the actual and desired pitch output can be used to inform a searching mechanism that would adjust the tonehole radius until the desired pitch is reached. We are currently working on incorporating the YIN pitch detection algorithm [15] as an additional step in the tuning algorithm to adjust for the discrepancies.

We also have plans to further extend the Birl physical model. For example, a register vent could be added to allow for octave jumps (one already exists in the STK BlowHole). This would follow the specification of a two-port scattering junction as outlined in [1]. Other extensions such as tonehole chimneys and non-constant bore diameter or curvature could be added to implement models of more complex instruments.

Work on the Birl is ongoing, and the results so far of this physical model implementation are encouraging. We also hope that others will find our repository of code useful for their own explorations.

REFERENCES

[1] G. Scavone and J. O. S. III, "Digital waveguide modeling of woodwind toneholes," in *International Computer Music Conference, Thessaloniki, Greece*, 1997.

[2] P. C. Gary Scavone, "Real-time computer modeling of woodwind instruments," in *International Symposium on Musical Acoustics, Leavenworth, WA*, 1998. [Online]. Available: https://ccrma.stanford.edu/~jos/tonehole/tonehole.html

[3] J. Snyder, "Exploration of an adaptable just intonation system," in *Columbia University Graduate School of Arts and Sciences Doctoral Dissertation*, 2010. [Online]. Available: http://www.scattershot.org/full-dissertation.pdf

[4] D. R. Jeffrey Snyder, "The birl: An electronic wind instrument based on an artificial neural network parameter mapping structure," in *International Conference on New Interfaces for Musical Expression, Goldsmiths, University of London, UK*, 2014. [Online]. Available: http://www.nime.org/proceedings/2014/nime2014_540.pdf

[5] J. Snyder, "The birl: Adventures in the development of an electronic wind instrument," in *Musical Instruments in the 21st Century*, H. E. S.-I. H. S. W. Till Bovermann, Alberto de Campo, Ed.

[6] J. O. S. III, "Physical modeling using digital waveguides," in *Computer Music Journal*, 1992. [Online]. Available: https://courses.cs.washington.edu/courses/cse467/05wi/pdfs/pmudw.pdf

[7] G. Scavone, "The pipe: Explorations with breath control," in *New Interfaces For Musical Expression*, 2003. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1.9899&rep=rep1&type=pdf

[8] S. O. M. Stephen Hughes, Cormac Cannon, "Epipe : A novel electronic woodwind controller," 2004. [Online]. Available: http://medialabeurope.org/research/library/Hughes_Epipe_2004.pdf

[9] T. I. L. Matti Karjalainen, Vesa Valimaki, "Modeling of woodwind bores with finger holes," in *International Computer Music Association*, 1993. [Online]. Available: http://quod.lib.umich.edu/i/icmc/bbp2372.1993.007/1

[10] G. S. Perry Cook. (1995) Stk. [Online]. Available: https://ccrma.stanford.edu/software/stk/index.html

[11] T. T. Matti Karjalainen, Vesa Valimaki, "Plucked-string models: From the karplus-strong algorithm to digital waveguides and beyond," in *The Computer Music Journal*, 1998. [Online]. Available: http://users.spa.aalto.fi/vpv/publications/cmj98.pdf

[12] J. O. S. III, "Digital waveguide architectures for virtual musical instruments," 1993. [Online]. Available: https://ccrma.stanford.edu/~jos/asahb04/asahb04.html

[13] C. Forster, *Musical Mathematics*. Chronicle Books, 2010.

[14] Birl physical model git repository. [Online]. Available: https://bitbucket.org/dharit_tan/birlsynth

[15] H. K. Alain de Chevigne, "Yin, a fundamental frequency estimator for speech and music," in *Acoustical Society of America Spring Meeting*, 2001.

An Electromagnet Based Wine Glass Instrument

Lior Arbel,[1†] *Yoav Y. Schechner,*[1] *Noam Amir*[2]

[1]Technion - Israel Institue of Technology
[2]Tel Aviv University
[†]lior51@yahoo.com

ABSTRACT

Wine glass instruments such as the glass harp are typically played by rubbing the glass rim with a moistened finger. This playing form limits both musical expression and instrument design. We explore a wine glass instrument involving electromagnetic excitation. The glasses are fitted with small magnets and excited using solenoid electromagnets. Several such glasses are connected in an array, driven by a single standard analog signal. The glass array is excitable by any amplified musical instrument or signal source and is capable of producing wine glass sounds with unique properties, unattainable on existing instruments. Several excitation configurations are suggested. The primary configuration is by driving the glass array with an amplified string instrument. For this configuration, noise reduction and sound enhancement digital filters were explored and customized. In addition, the produced wine glass sound characteristics as decay, attack, and harmonic spectrum were altered using various digital effect plugins, such as Swell, Wah-wah and compression. An additional proposed configuration is by driving the glass array with the output of virtual instruments or sample libraries, played by a MIDI controller. The response of the glass array to several sample library sounds was explored. In addition, a virtual instrument was custom designed with the specific intent of glass array excitation. The acoustic properties of the virtual instrument are designed to fit the glass array acoustics. Several technology-oriented musicians evaluated the instrument hands-on, and provided ideas and suggestions for improvement.

Multiphonics in the Trombone, Measurements and Time-Domain Simulations

Lionel Velut,[1] *Christophe Vergez,*[1] *Joël Gilbert*[2†]

[1]Laboratoire de Mécanique et d'Acoustique, CNRS
[2]Laboratoire d'Acoustique de l'Université du Maine, CNRS
[†]joel.gilbert@univ-lemans.fr

ABSTRACT

Multiphonic sounds of brass instruments are produced by playing a note while simultaneously singing another note. This results in a peculiar sound, heard as a chord or a cluster of more than two notes in some cases. Measurements of the mouth pressure, the pressure inside the mouthpiece, and the radiated sound are recorded while a trombone player performs a multiphonic, first by playing an F3 and singing a C4, then playing an F3 and singing a note with a decreasing pitch. Comparison between simulated and measured signals is discussed. Complementary results can be found in the following recent publication: Velut, L., Vergez, C., and Gilbert, J. (2016). "Measurements and time-domain simulations of multiphonics in the trombone," J. Acoust. Soc. Am. 140(4), 2876–2887.

Optimization of Brass Instruments Based on Physical Modeling Sound Simulations

Robin Tournemenne,[1] *Jean-François Petiot,*[1†] *Joël Gilbert*[2]

[1]Ecole Centrale de Nantes
[2]Laboratoire d'Acoustique de l'Université du Maine
[†]jean-francois.petiot@irccyn.ec-nantes.fr

ABSTRACT

Sound simulations can be interesting for instrument design as far as they are representative of the functioning of the real instrument. This work presents how physical modeling based simulations are integrated in the optimization process of the shape of a trumpet. The novelty of the approach lies in the fact that the optimized objective function concerns the sound of the instrument, simulated when it is in interaction with a virtual musician. The physics-based model consists in an acoustic model of the resonator (input impedance), a mechanical model of the excitator (the lips of a virtual musician) and a model of the coupling between the excitator and the resonator. The harmonic balance technique is used to compute sounds in permanent regime, representative of the shape of the resonator according to control parameters of the virtual musician. An optimization problem is formulated, in which the objective function to be minimized is the overall intonation of the different notes played by the instrument (deviation from the equal-tempered scale). The design variables are the physical dimensions of the resonator. Given the computationally expensive function evaluation and the unavailability of gradients, a surrogate-assisted optimization framework is implemented using the mesh adaptive direct search algorithm (MADS). Surrogate models are used both to obtain promising candidates in the search step of MADS and to rank-order additional candidates generated by the poll step of MADS. Examples of design spaces of various sizes are presented to demonstrate the validity of the approach. Two design problems concerning the optimisation of the intonation of four notes (regimes 2, 3, 4, 5) of a trumpet are presented, with two (2D) and five (5D) design variables, located in the trumpet leadpipe. The intonation of the optimized bore, compared to that of the standard bore, shows that significant improvements of intonation can be achieved at reasonable computational cost (average improvement of 5.4 cent in the 2D case, and 5.8 cent in the 5D case).

The implementation of this method for computer-aided instrument design is discussed, considering different objective functions or constraints, based on the intonation but also on the timbre of the instrument.

Harmonic Balance and Continuation Applied to a Model of Brass Instruments

Lionel Velut,[1] *Christophe Vergez,*[1†] *Pierre Vigué,*[1] *Joël Gilbert*[2]

[1]LMA, CNRS, UPR 7051, Aix-Marseille Univ., Centrale Marseille, F-13453 Marseille cedex 13, France.
[2]Laboratoire d'Acoustique de l'Université du Maine CNRS
[†]vergez@lma.cnrs-mrs.fr

ABSTRACT

Harmonic balance method is useful to find periodic solutions of a nonlinear dynamical system modeled by a continuous time system. The main idea is to replace unknown variables of the continuous time system by their Fourier series. Therefore, the continuous time system, including possibly time derivatives, is transformed into an algebraic system. The unknowns of the algebraic system are the Fourier coefficients and, when the system is autonomous (case of brass instruments), the oscillation frequency as well. The stability of each periodic solution can be assessed.

Harmonic balance method is applied to a physical model of brass instruments. Lips oscillation is reduced to one-mode mechanical oscillator, whether the measured input impedance of the bore of the instrument is projected onto the truncated modal basis of the air column. The air flow entering the instrument is written according to the usual instantaneous model relying on the Bernoulli theorem applied between the mouth and the lips channel, as well as turbulent mixing inside the cup of the mouthpiece without pressure recovery. For a given set of parameters values, harmonic balance gives the spectrum of each periodic solution found as well as its playing frequency. Moreover, the stability of each periodic solution is found.

In order to deepen the understanding of the model with respect to the value of the blowing pressure, continuous branches of periodic solutions are sought, rather than collections of solution points for different values of the blowing pressure. This is done by coupling the harmonic balance method to a continuation method, namely the asymptotic numerical method in our case. Particular attention is paid to the birth and the extinction of the sound when the blowing pressure increases and decreases respectively.

Resolving Rhythm Ambiguities with Timbre

Punita G. Singh

Sound Sense
punita@gmail.com

ABSTRACT

The perception of rhythm and meter depends critically on some sort of differentiation between the sound events comprising a sequence. Differences in pitch, loudness, duration and timbre can all serve as markers imbuing accents that can demarcate temporal sub-groups and create tension and release via syncopation and cyclical resolution. The role of timbre as a breaker or coherer of rhythmic groups has been studied in earlier research (Singh, 2013). Experiments with controlled stimuli designed with specific spectral and temporal attributes indicated that timbre differences based on spectral density and locus are stronger accent markers than those based on temporal envelope slopes (Singh, 1999). Ambiguous rhythms based on simultaneously embedded triple and quadruple metrical events could be resolved effectively with strong spectral accent markers. Building on that work, experiments are in progress with more realistic 'sampled' sounds. Sequences with straightforward or mixed metrical structure have been recreated with different levels of timbre contrast based on instrument family kinship. Listeners are asked to report if they perceive a triple, quadruple or ambiguous meter, or an isochronous series of events with no meter at all. Results to date indicate that metrical ambiguity in sequences with undifferentiated timbre accents is the most difficult to resolve. When distinctive timbres are available as accent markers, listeners are able to latch on to underlying rhythmic groups to resolve metrical ambiguity. Timbres within an instrument family provide lower levels of contrast than timbres across families. The dissimilarity of timbres in a sequence is evidently a vital cue for parsing temporal groups and can serve to resolve rhythm ambiguities.

References:
Singh, P. G (1999) 'Parsing complex rhythmic structures: The contribution of spectral and temporal dimensions of timbre', J. Acoust. Soc. Am., 105(2), DOI: http://dx.doi.org/10.1121/1.426093
Singh, P. G (2013) 'The contribution of spectrum and tempo to auditory streaming of simple and compound "bols" in tabla rhythms', J. Acoust. Soc. Am., 134(5), DOI: http://dx.doi.org/10.1121/1.4831234

Assessment of Auralizations of Piano and Guitar for Music Practice Rooms

Konca Saher, [1†] *Yalin Ozgencil,* [2]

[1]Kadir Has University, Interior Architecture and Design, Istanbul, Turkey
[2]Yildiz Technical University, Architecture, Istanbul, Turkey
[†]konca.saher@khas.edu.tr

ABSTRACT

The acoustic quality of music practice rooms are of great significance to a musician for perceiving the sound of the instrument, the nuances in intonation, the frequencies and the pitch. In the scope of this paper, two conservatories in Istanbul were chosen and three different music practice rooms (for piano and guitar rehearsals) with varying design and reverberation time were analysed by acoustic measurements, simulations and auralizations. The survey results indicated that the existing reverberation time values vary between 0.4 sec and 1.4 sec in the mid frequency range. The subjective response of the musicians indicated that the music practice rooms were perceived to be too live. The measured music practice rooms were acoustically modelled and three alternative designs were developed for each room resulting in total twelve acoustic models. Listening tests prepared from auralizations of guitar and piano instruments in twelve acoustic models were assessed by musicians of the conservatories. The results verified the assumption that different reverberation time preferences are made for different musical instruments.

1. INTRODUCTION

Music practice rooms are probably the type of rooms which have the highest occupancy time among the rooms which are specifically designed for music. Musicians might spend up to 40 hours a week, therefore, the perception of the sound in the room and "critical listening" becomes very important and crucial for the musician's professional development [1, 2].

The acoustical design of the music practice rooms are important for the musician to determine the sound, intonation, articulation, pitch and variations in the frequency in regard to the music instrument he is playing [3]. In the rooms that only utilize high frequency absorption, flutes, violins, and other high frequency sounds, including high overtones of most instruments can be lost. Intonation, articulation and timing can be blurred, and critical listening becomes impossible [4, 5].

The optimal acoustic design for a music practice room might depend on the type of music, instrument and the musician. [5,6,7]. Most of the standards and guidelines for music practice rooms describe the acoustical quality in terms of reverberation time. [8,9,10]. However, none of these standards describe the requirements or possible variations for different types of musical instruments. This paper seeks to investigate the reverberation time requirements for two different types of musical instruments; piano and guitar.

The reverberation time values recommended in these standards and guidelines for various room sizes are between 0.3-1.5 seconds [8,9,10]. These reverberation time recommendations are usually in the mid frequency range. (500 Hz, 1000 Hz, 2000 Hz) Table 1 below summarizes the reverberation time values for related standards.

Room Type	**Fl. Area (m^2)**	**AS107. 2000 [8]**	**DfE .2002 [9]**	**ANSI S12.60 [10]**
Music theory room	50-70	0.5 - 0.6	0.4 - 0.8	< 0.6
Group music room	16-50	0.7 - 0.9	0.5 - 1.0	< 0.6
Recital room	50-100	1.1 - 1.3	1.0 - 1.5	N/A
Practice room	6 - 10	0.7 - 0.9	0.3 - 0.6	< 0.6

Table 1. Recommended reverberation time values in guidelines for various music practice room types and dimensions.

There is no reverberation time recommendation for practice rooms in Turkey. A research prior to this study examined six practice rooms in two conservatories in Turkey by interviews with musicians and concluded that the musicians are not satisfied with the acoustic conditions of their practice rooms [11]. Most of the musicians indicated that the music practice rooms are too live [11].

The aim of this paper is to determine the most appropriate reverberation time and room type for the practice of two musical instruments; piano and guitar. Therefore, the range of reverberation time (RT) is chosen as the acoustical criteria and the floor area has been chosen as the architectural criteria. Three practice rooms with different floor area and reverberation time were chosen from two conservatories in İstanbul, Turkey. Impulse response measurements of the practice rooms were made, then acoustic models of rooms were developed (in ODEON v12). Three new acoustic models were proposed for each practice room by adding absorptive and reflective panels, bass traps and reducing the volume; resulting in twelve acoustic models. (Three models are the original practice rooms, nine models are the proposed improvement models) The anechoic recordings of piano and guitar were convolved with the room impulses of these twelve models. The auralizations of each instrument for different room models were used to prepare the listening tests. The subjects have been asked to select 'the best' room for the designated instrument.

2. ACOUSTICAL MODELING

2.1. Measurements and Simulations of Existing Rooms

Three practice rooms are selected for this study based on their floor areas, volumes, and reverberation time and musician responses from the preceding study. [11] Detailed description of the practice rooms are given in Table 2 below. The first two rooms; MIAM 212 and MIAM 224 are practice rooms in Istanbul Technical University Centre for Advanced Studies in Music (MIAM) and ITU Recital Room is a practice and recital room in Istanbul Technical University Conservatory. The three practice rooms below are exhaustively used by musicians for various musical instruments including guitar and piano.

	Height (m)	**Floor Area (m^2)**	**Volume (m^3)**
MIAM 121	2.65	43.7	137.2
MIAM 224	2.72	16.2	44.13
ITU Recital Room	4.41	67.2	296.4

Table 2. The dimensions of the selected practice rooms

Impulse response measurements were done in these three practice rooms to determine the existing reverberation time values. At least two source and three receiver locations were used for each room measurements. Results were presented as the average of all measurements for each room. (Figure 1)

3d architectural models of three practice rooms were constructed based on the architectural survey. The architectural survey were made with the aid of a laser meter and a traditional measuring tape. The materials of each surface was also carefully noted during the architectural survey. The physical model of the practice rooms were developed in 3dMax and AutoCAD and then exported to a room acoustic modelling software. (ODEON v12)

One of the most important tasks during the acoustical modelling was the assignment of materials to the room surfaces and estimation of the absorption and scattering characteristics of the materials. The room acoustic models of three practice rooms were calibrated based on the measurements. Figure 1 below shows the simulated versus measured reverberation values for three practice rooms. The measured reverberation times in the mid frequency range is 0.9 sec for MIAM 121; 0.6 sec for MIAM 224 and 1.4 sec for ITU Recital Room. The reverberation time values are much higher in the low frequencies in two rooms (MIAM 224 and ITU Recital Room as can be observed from the Figure 1 below. Normally, common finishing materials such as carpets, curtains and upholstery absorb the high frequencies but not very effective in low frequencies and these results are to be expected. The Figure 1 below also shows a very good match between the measured and simulated values.

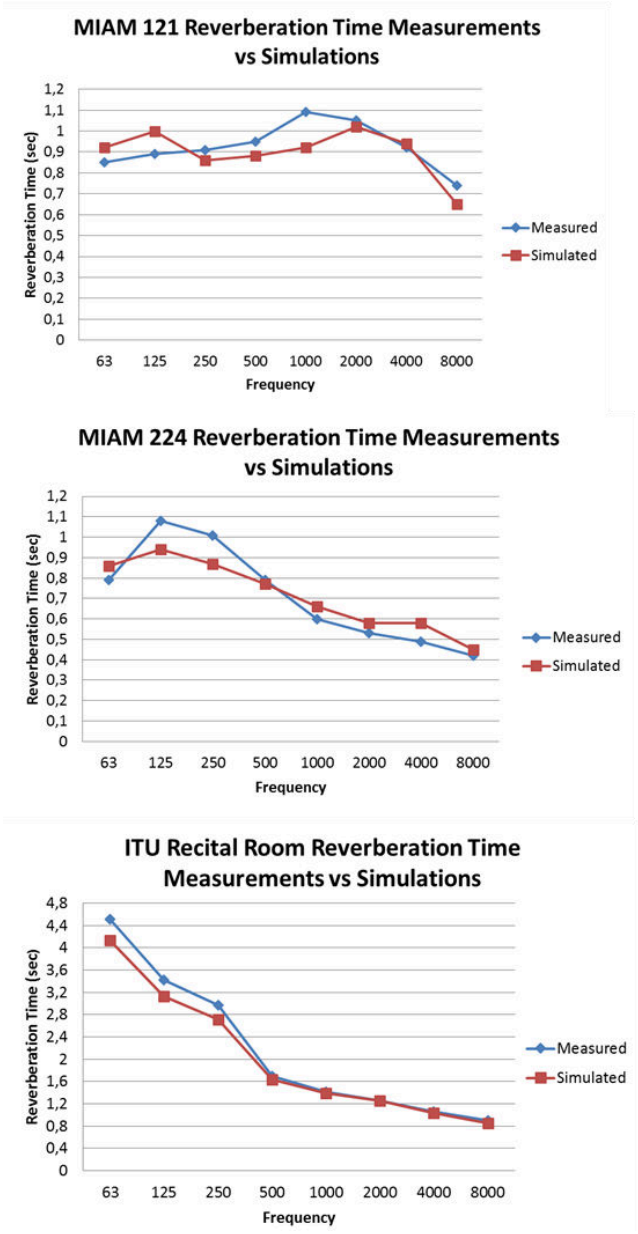


Figure 1. Measured vs simulated reverberation time values for three music practice rooms

2.2. Proposed Acoustic Models

Three room proposals with different acoustic characteristics were developed for each music practice room and these rooms were modelled in ODEON. These room models were developed by adding absorptive and diffusive surfaces inside the practice rooms, changing the volume of the room or changing the absorption characteristics of the surfaces. The detailed description of each proposed model for each practice room and reverberation time at 1000 Hz for measured and proposed models are given in Table 3 below. Detailed frequency response of proposed models and measured rooms are given in Figure 2 below.

		Proposed Change	**RT (1kHz)**
MIAM 121	Original	-	0.92
	Model 1	Reducing the volume	0.84
	Model 2	6 absorptive panels to the walls + 4 reflective panels to the ceiling	0.71
	Model 3	6 absorptive panels to the walls + 3 bass traps to the corners	0.69
MIAM 224	Original	-	0.66
	Model 1	4 absorptive panels to the walls	0.52
	Model 2	2 bass traps to the corners 4 reflective panels to the walls	0.66
	Model 3	4 absorptive panels to the walls + reflective panel	0.44
ITU Recital	Original	-	1.39
	Model 1	10 absorptive panels to the walls	0.95
	Model 2	10 absorptive panels to the walls 4 bass traps to the upper corners	0.89
	Model 3	10 absorptive panels to the walls 4 bass traps to the upper corners 4 absorptive panels to the ceiling	0.67

Table 3. Description of proposed acoustic models

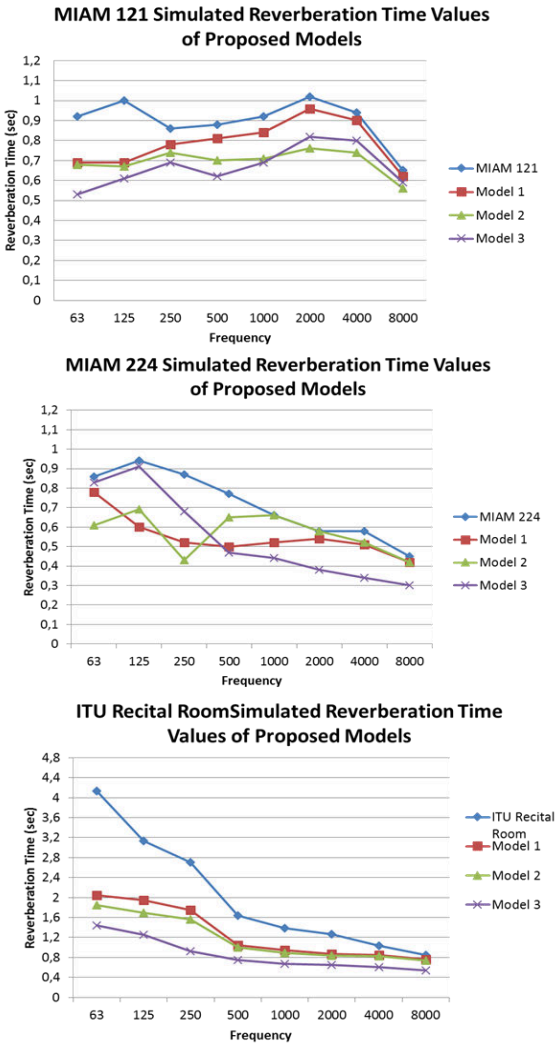


Figure 2. Simulated reverberation time values for proposed acoustic models

MIAM 121 practice room has originally a reverberation time of 0.92 sec. in 1000 Hz. Three acoustic models were proposed for this room; Model 1 by reducing the volume of the room; Model 2 by adding six absorptive panels to the walls and four reflective panels to the ceiling; and Model 3 by adding six absorptive panels to the walls and three bass traps to the corners. (Table 3) The reverberation time values reduced to 0.84 sec, 0.71 sec and 0.69 sec. in 1000Hz for Model 1, Model 2 and Model 3, respectively.

MIAM 224 practice room has originally a reverberation time of 0.66 sec. in 1000 Hz. Three acoustic models were proposed for this room; Model 1 by adding four absorptive panels to the walls; Model 2 by adding two bass traps to the corners and adding four reflective panels to the walls; and Model 3 by adding four absorptive panels to the walls and one reflective panel to the opposite of the instrument. (Table 3) The reverberation time values reduced to 0.52 sec, and 0.44 sec. in 1000Hz for Model 1 and Model 3; however, the reverberation time remained the same for Model 2 as 0.66 sec. for 1000Hz.

ITU Recital practice room has originally a reverberation time of 1.39 sec. in 1000 Hz. Three acoustic models were proposed for this room; Model 1 by adding ten absorptive panels to the walls; Model 2 by adding ten absorptive panels to the walls and four bass traps to the upper corners of the room; and Model 3 by adding ten absorptive panels to the walls, four bass traps to the upper corners of the room and four absorptive panels to the ceiling. (Table 3) The reverberation time values reduced to 0.95 sec., 0.89 sec. and 0.67 sec. in 1000Hz for Model 1, Model 2 and Model 3, respectively. Even though the reverberation time values were obtained for 63-8000 Hz, the reverberation time results are discussed as 1000 Hz for most part of the paper.

3. AURALIZATIONS AND LISTENING TESTS

The anechoic recordings of piano and guitar pieces were convolved with binaural room impulse responses of measured three rooms and three proposed models for each room; resulting in twelve auralizations of piano and twelve auralizations of guitar. The anechoic recordings of guitar and piano were taken from ODEON wave signals library. The directivity of piano and directivity of guitar were assigned to the sources in the auralizations. All auralizations were edited in Protools software so that they were all 20 seconds in durations. The listening tests were made with 28 musicians between the ages of 15-30; however; majority of the musicians were between the ages of 16-25. None of them reported any hearing problems. They were all musicians who use the practice rooms very often at the conservatories. Among all musicians 14 musicians were piano players and 14 musicians were guitar players.

The auralizations in the listening tests were presented through headphones, so special care must be taken to make the levels equal to the levels experienced in real situations. It was done by adjusting the levels in the computer to the levels the authors believed to be close to the real situation; however; for the next listening tests efforts will be made to calibrate the levels to the real levels. No information about the rooms were given to the subjects before and during the listening tests. The musicians were asked to listen to auralizations of the original room and

three proposed rooms in one session for each practice room (MIAM 121, MIAM 224 and ITU Recital Room) and they were only asked to choose the preferred room to practice among the four auralizations for each room. Guitar players listened to the guitar samples only; and piano players listened to the piano samples. They could listen to the auralization files as many times as liked. Table 4 below shows the preference of the musicians for guitar and piano practice among the simulated rooms. The results shown here are the majority's preference among the musicians. The preference scores are also included in Table 4.

Room	**Piano**	**Pref. Score**	**RT(1000 Hz)**
MIAM 121	Original	%57	0.92 sec.
MIAM 224	Original	%50	0.66 sec.
ITU Recital	Model 2	%66	0.89 sec.
Room	**Guitar**		**RT (1000 Hz)**
MIAM 121	Model 2	%75	0.71 sec.
MIAM 224	Model 3	%75	0.44 sec.
ITU Recital	Model 3	%57	0.67 sec.

Table 4. Preference by musicians for simulated models for piano and guitar practice

The results show that piano players prefer the original rooms (which have reverberation time of 0.92 sec. and 0.62 sec. respectively at 1000 Hz) for piano practice in MIAM 121 and MIAM 224; while they prefer Model 2 for ITU Recital Room which has a reverberation time of 0.89 sec. at 1000 Hz. The guitar players prefer Model 2 (which has 0.71 sec. reverberation time at 1000 Hz) for MIAM 121 and Model 3 (which has 0.44 sec. reverberation time at 1000 Hz) for MIAM 224 and Model 3 (which has 0.67 sec. reverberation time) for ITU Recital Room. The results indicate that musicians prefer much dry rooms for guitar practice and more lively rooms for piano practice. The musicians prefer rooms with reverberation time of 0.6-1.0 sec for piano and 0.4-0.7 sec. for guitar practice. All these reverberation time values are mid frequency values. The results also indicate that musicians who play different instruments have different acoustical preferences for the same practice room.

4. CONCLUSIONS

Most of the standards and guidelines recommend reverberation time values based on the floor area or number of people for a practice room; however; different acoustic requirements for different musical instruments are not fully explored in literature. In this study, three music practice rooms in two conservatories in Istanbul are investigated. Previous research by the authors have shown that musicians in these conservatories are not satisfied with their music practice rooms and they find their music practice rooms lively. The aim of the research was to determine the preferred reverberation time for guitar and piano practice. Acoustic measurements and simulations were made for three music practice rooms and three new acoustic models were proposed for each room by adding absorptive and reflective panels, adding bass traps or in one instance by reducing the volume of the room. Listening tests were prepared from auralizations of each room model with piano and guitar recordings. The results strongly support the presumption that the reverberation time requirements or preferences for musicians might vary depending on the musical instrument. The results indicated that musicians who play guitar and who play piano have different preferences of reverberation time for the same music practice room. Guitar players prefer much dry rooms; however; piano players prefer lively rooms.

The limited number of music practice rooms and musicians makes it difficult to generalize the results. It is required to increase the number of participants in the future listening tests to be able to have more statistically significant results. It may also be useful to extend the analysis in more detail in low frequencies. Future studies should extend this work to explore preference for other musical instruments such as flute, violin, etc. to examine the interaction between the acoustic properties of the room and the musician to achieve an optimal music practice room.

REFERENCES

[1] D. Lamberty, "Music Practice Rooms," *Journal of Sound and Vibration*, vol 69 (1), pp. 149-155, 1980.

[2] H. Laitinen, "Factors affecting the use of hearing protection among classical music players," *Noise Health*, vol 7, pp, 21–29, 2005.

[3] E. McCue & R. H. Talaske, Eds, *Rehearsal Room Acoustics, Acoustical Design of Music Education Facilities.* New York, Acoustical Society of America, 1990.

[4] Wenger Corporation, Planning Guide for School Music Facilities, 2008.

[5] K. Saher & J. H. Rindel, "How Does Architecture Sound for Different Musical Instrument Performers," in *Proceedings of the Architecture, Music, Acoustics, International Cross-Disciplinary Conference*, Toronto, Canada, 2006.

[6] G. Gunnlaugsdottir, "Optimization of Acoustic Conditions in Music Practice Rooms", *in Proceedings of Joint Baltic-Nordic Acoustics Meeting*, Reykjavik ,Iceland, 2008, pp 1-6.

[7] T.D. Rossing, F. R. Moore & P. A. Wheeler, *The Science of Sound*, 3rd Edition, SanFrancisco: Pearson Education, 2002.

[8] Australian Standard AS/NZS 2107 Acoustics- Recommended design sound levels and reverberation times for building interiors. (2000).

[9] Dfe (Department for Education, United Kingdom) Acoustic Performance Standards for the Priority Schools Building Programme, version 1.7, June 2013.

[10] American national standard acoustical performance criteria, design requirements and guidelines for schools, ANSI S12.60, 2002.

[11] Y. Ozgencil & K. Saher, "Konservatuar Prova Odalarının Akustik Özelliklerinin İncelenmesi" *in Proceedings of 11th National Turkish Acoustic Congress*, İstanbul, Turkey, 19-20 October,2015.

The Effect of Musical Abilities on Implicit and Explicit Memory

Rachel Miller,[†] *Daniel Fogerty*

Department of Communication Sciences and Disorders, University of South Carolina
[†]remiller@email.sc.edu

ABSTRACT

Musical training is associated with better performance on various cognitive and perceptual tasks. Previous research has suggested a musician advantage for auditory memory tasks or speech perception in noise. However, it is not clear how these two abilities might be involved in the memory and recognition of speech and music in noise. The musician advantage for these memory tasks has typically been assessed in quiet using a single talker. However, in noisy conditions, the perceptual processing and resolution of the auditory signal might become more important. Indeed, previous results support musicians' enhanced capabilities for speech perception in noise. The direct contribution of memory under these degraded listening conditions has not been examined in relation to musical ability. One hypothesis for the musician advantage is that musicians develop increased sensitivity to acoustic parameters through musical training. This may in turn facilitate speech perception and encoding of the stimulus for later retrieval in memory. Alternatively, or concurrently, musical training may also facilitate the development of robust working memory abilities. Thus, musicians may be better able to hold and manipulate items in memory. As greater working memory capacity is associated with better speech perception in noise, this may indicate cognitive rather than perceptual factors underlying the musician advantage. The present study aims to examine implicit and explicit memory capabilities for speech and music. Participants completed a task battery, which assessed memory capabilities for melodies, instruments, and speech. Participants also completed a speech in noise task, as well as several cognitive measures examining vocabulary and pattern recognition capabilities. Musical ability was assessed through the Advanced Measures of Musical Audiation (AMMA), and by years of musical training. Preliminary results suggest that memory for speech is unrelated to memory for music. Possible correlations were explored between memory performance, music perception, and cognitive measures. A notable relationship was observed between AMMA performance and visual pattern recognition abilities.

Transient Timbre and Pitch Changes in Musical Instruments and Voice

Malte Kob,[1†] *Armin Kohlrausch*[2]

[1]Detmold University of Music
[2]Human Technology Interaction, Technical University Eindhoven
[†]kob@hfm-detmold.de

ABSTRACT

The human perception of single tone pitches depends on the loudness of the sound source. There seems to be agreement that the pitch shift is frequency-dependent [Snow 1936, Terhard, 1979, Rossing 1990, Liljencrants 2003]. Based upon the observation of pure tones it is expected that for complex harmonic sounds the relation between level and pitch should be reflected in the perception of timbre. This effect is expected to be more subtle but has been observed in the decay of orchestra sounds. In this investigation the timbre and pitch perception is investigated from two sides: physically from the analysis of musical sounds using signal processing methods, and perceptually using adequate stimuli and listening tests set-ups. Since most analysis methods for timbre of musical instruments make use of sustained or at least windowed sound samples the transient nature of changes requires time-resolved analysis methods. A review and first implementations of model sounds and analysis methods are presented and and their usability for the analysis of transient pitch and timbre changes is discussed. A perspective of our study is the potential categorisation of listeners into two groups who determine pitch according to the fundamental frequency and those who determine pitch from the spectral distance of the harmonic components.

1. Snow, W. B.: "Change of pitch with loudness at low frequencies", Journal of the Acoustical Society of America, 8(1), 14-19, 1936, URL: http://www.aruffo.com/eartraining/research/articles/snow36.htm

2. Terhardt, E.: "Calculating Virtual Pitch", Hearing Research 1, 155, 1979.

3. Rossing, Thomas D.: "The Science of Sound", Addison-Wesley, 2nd ed., ch. 7, 1990.

4. Liljencrants, J.: Pitch Shift From Change In Loudness. URL: http://www.fonema.se/pilo/pilo.html

Measurements and Perceptions of Interactions between Musicians and Cellos

Timothy Wofford,[1†] *Claudia Fritz,*[1] *Benoît Fabre,*[1] *Jean-Marc Chouvel*[2]

[1]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7190, Institut Jean Le Rond d'Alembert, F-75005 Paris, France
[2]Sorbonne Universités, Paris-Sorbonne Univ Paris 04, CNRS, UMR 8223, IReMus, F-75002 Paris, France
[†]wofford.timothy@gmail.com

ABSTRACT

Measurement of physical properties alone has failed to predict how an instrument will be perceived by musicians. Physical properties are only capable of characterizing an instrument as an object. They are not enough to characterize the suitability of the object as an instrument for producing music by a particular musician. To evaluate an instrument, the musician must apply gestures to the instrument and interpret the response of the instrument to those gestures. Two musicians may use different gestures, receive different responses, and thus have a different set of interactions upon which to base their evaluation. This could explain why musicians disagree over instrument evaluations.

To study the relationships between gestures, responses, and the perception of interactions in the context of evaluating cellos, we have put in place 1) a motion capture system for measuring the control parameters and 2) a set of piezoelectric sensors to measure the string vibrations of each string. Gestures and string vibrations were measured during the evaluation of two cellos by a few cellists. These measurements are interpreted in the light of the perceptual evaluations.

Acoustic Characteristics of Pressed and Normal Phonations in Choir Singing by Male Singers

Kenta Wakasa,[1†] Masaki Matsubara,[1] Yuzuru Hiraga,[1] Hiroko Terasawa,[1]

[1]University of Tsukuba, Japan
†wakasa.kenta.su@alumni.tsukuba.ac.jp

ABSTRACT

In this study, we investigate the vocal quality of pressed and normal phonations in choir singing by Japanese male singers. We recorded the vocal signals of 15 experienced male choir singers as they performed sustained vowels in two phonation modes: normal phonation, which is used in choir singing, and pressed phonation, which is undesirable for choir singers. All participants sang three vowels in three pitches with two dynamics, repeated three times. These sounds were recorded and several acoustic characteristics were measured. The results showed that pressed phonation had lower harmonics-to-noise ratio and higher singing power ratio than normal phonation, thus confirming hoarse vocal qualities.

1. INTRODUCTION

In choral performances, singers are advised to sing in *bel canto* style (normal phonation) and avoid pressed phonation; however, many singers experience confusion in distinguishing between these two styles. In this study, we investigated the timbre difference between these two singing styles among Japanese male choir singers.

According to Sundberg et al. [1], if one phonates with high subglottic pressure and adduction force in the larynx, the phonation is termed pressed phonation. However, pressed phonation is not preferable in classical choir singing; hence, it is mostly avoided by choir singers. What are the acoustic differences between normal and pressed phonation? Several studies have investigated the relationship between phonation modes and acoustic or physical characteristics.

Sundberg examined the formant frequencies of professional male singers and determined a peak called the singer's formant near 3 kHz [2]. Furthermore, Sundberg et al. examined the relationship between the phonation modes and the flow glottogram (i.e., the estimated source signal at the vocal cord) [1]. This study revealed that pressed phonation reduces the flow glottogram amplitude, and accordingly, the estimated amplitude of the fundamental frequency at the vocal source. Furthermore, the spectral slope of the flow glottogram is shallower in pressed phonation than in normal phonation.

As a physical measure of the vocal folds, several studies investigated glottal open quotients, which are related to voice qualities such as breathy and pressed [3, 4]. In [5], the modal and falsetto voices were compared and it was reported that the open quotients in the modal voice are lower and the spectral slope is smoother [5].

In the present study, we compare the acoustic characteristics of normal and pressed phonations in amateur Japanese male choir singers. The previous studies compared the normal and pressed phonations by observing the singers' glottal conditions. However, we attempted to differentiate these two phonations by observing only audio data. Therefore, the focus of this research is to investigate the acoustic characteristics related to pressed phonation.

2. METHOD

2.1 Environment

The subjects for this study included 15 male singers: 5 basses and 10 tenors. The singers have been singing in amateur Japanese choirs for an average of 4.2 years.

The recording experiment was performed in a quiet room. Vocal signals were recorded using a condenser microphone (AKGC391B) placed at a distance of 50 cm from the singers face. These signals were then digitized at 16 bits with a sampling rate of 44.1 kHz using an audio interface (MOTU UltraLite mk3).

2.2 Procedure

For this study, the subjects performed sustained-vowels in two modes of phonation: normal and pressed. Sustained-vowels consist of [a], [e], and [u]. These vowels were performed at two degrees of loudness (*mezzo-forte* and *forte*) and at three pitches, which differ depending on the singer's vocal range. The bass singers performed A2, D3, and G3. The tenor singers performed D3, G3, and C4. The length of the sustained-vowels was more than 2 s.

3. ANALYSIS

3.1 Data Pre-Processing

Prior to the analysis, we extracted sustain sections from the singing data because the onset and offset parts may exhibit distinct spectral patterns that need to be analyzed differently. The extracted sustain sections have a duration of 1 s, and they're extracted 0.5 s after the onset of singing data.

3.2 Characteristics

The following acoustic characteristics were used for the analysis. Note that these characteristics were first calculated over 1 s using a running window method and a Hanning window of length 1024. We then averaged the values over 1 s for each characteristic and employed the averaged value as our analysis result.

3.2.1. Harmonics-to-noise ratio

The harmonics-to-noise ratio (HNR) provides an index of the overall periodicity of the voice signal by quantifying the ratio between the harmonics and the noise components. In this study, HNR is measured using Praat, which is software for voice analysis [6]. Note that, lower the HNR, the more hoarse is the voice quality [7].

3.2.2. Jitter

Jitter represents a fluctuation of durations of consecutive periods in the voice. In this study, we used the neighboring five-point period perturbation quotient provided in Praat [6].

3.2.3. Shimmer

Shimmer represents a fluctuation of amplitudes of consecutive periods. In this study, we used the neighbor five-point amplitude perturbation quotient provided in Praat [6]. Shimmer and Jitter are both indices for roughness [8].

3.2.4. H1 − H2

H1 and H2 are the amplitudes of the fundamental frequency and the second harmonic, respectively. If the absolute difference between H1 and H2 (i.e., H1 − H2) is large, the slope of the spectrum envelope is steep and the voice quality is more breathy [9].

3.2.5. Spectral centroid

A spectral centroid indicates the location of the center of mass of the spectrum. If this value is large, the sound is sensed as bright [10].

3.2.6. Singing power ratio

Singing power ratio (SPR) is the ratio between the greatest harmonics peak between 2 and 4 kHz and the greatest harmonics peak between 0 and 2 kHz. SPR shows a significant relationship with perceptual scores of "ringing" quality [11].

4. RESULTS

In order to compare the acoustic characteristics of normal and pressed phonations, we conducted three-factor repeated-measure ANOVA. The three factors comprise two dynamics (*mf, f*), two pitches (D3, G3), and two phonation modes (normal, pressed). These pitches (D3, G3) were selected as a factor because both basses and tenors perform these pitches. This ANOVA was conducted for each vowel ([a], [e], and [u]).

We are interested in identifying if there is a significant difference between normal and pressed phonations. For example, we did not observe a significant difference in the centroid between the two modes. In the following section, we discuss four acoustic characteristics that exhibited noticeable differences between normal and pressed phonations.

4.1. HNR

A significant difference caused by phonation modes was observed in the HNR of [u] (p <0.05). Fig. 1 shows the mean and standard deviation of HNR for each vowel. For [u], the mean of HNR with pressed phonation is lower than that with normal phonation. Therefore, pressed phonation in [u] exhibits lower HNR and a more hoarse voice quality.

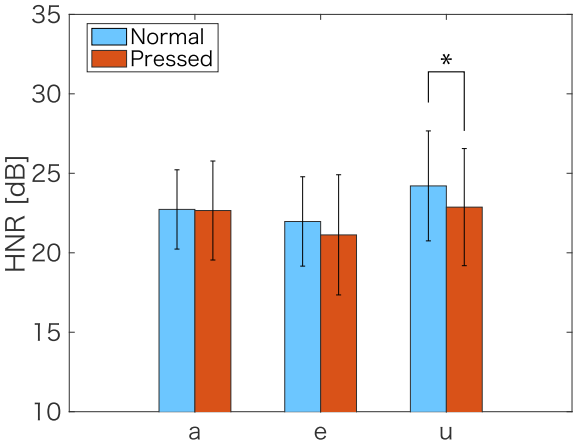


Figure 1. Mean and standard deviation of HNR for each vowel. For [u], the mean of HNR in pressed phonation is lower than that in normal phonation (*p < 0.05).

In addition, interaction was found between the dynamics and the phonation mode (p < 0.01) in the HNR of [u]. Fig. 2 displays the mean and standard deviation of HNR under each condition: dynamics (*mf, f*), phonation mode (Normal, Pressed), and vowel ([u]). When the dynamics is *forte*, the difference between the average values for the normal and pressed phonations is larger. In this dynamics level, the simple main effect of the phonation mode was confirmed (p < 0.01).

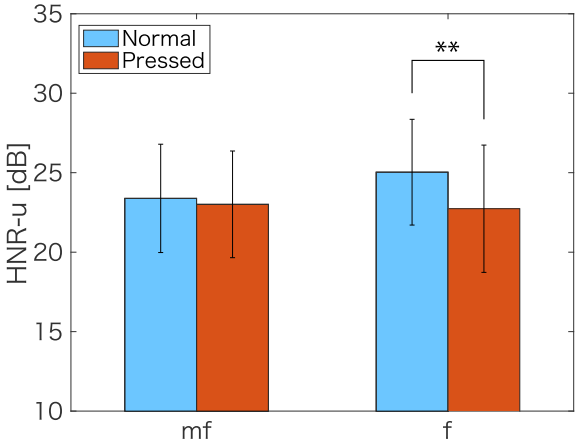


Figure 2. Mean and standard deviation of HNR-u in each dynamics. The difference between the average values for the normal and pressed phonations is larger in *forte* (**p < 0.01).

4.2. SPR

A significant difference caused by phonation modes was observed in the SPR for [u] (p < 0.01). Fig. 3 shows the mean and standard deviation of HNR in each vowel. For [u], the mean of SPR with pressed phonation is higher than that with normal phonation. Our initial expectation was that the SPR values for normal phonation would be larger; however, this result indicated an opposite tendency. We discuss this further in the next section.

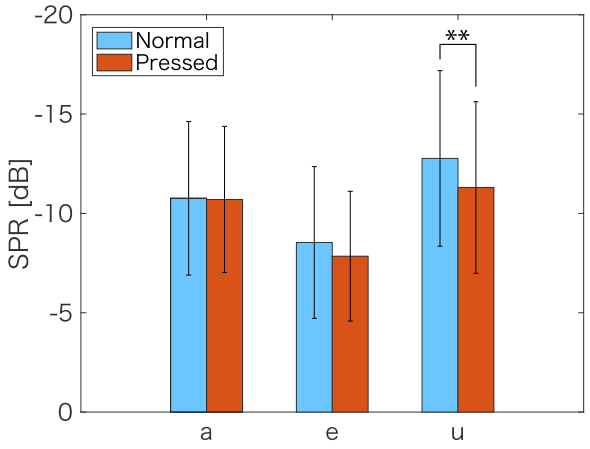


Figure 3. Mean and standard deviation of SPR for each vowel For [u], the mean of SPR with pressed phonation is higher than that with normal phonation (**$p < 0.01$).

4.3. Jitter

The phonation mode did not result in a significant difference in the jitter for all vowels. However, the relationship between pitch and phonation mode ($p < 0.05$) was observed in the jitter for [a]. Fig. 4 displays the mean and standard deviation of jitter under each condition: pitches (D3, G3), phonation mode (Normal, Pressed), and vowel ([a]). For the normal phonation mode, the difference between the average values of D3 and G3 is larger. Under normal phonation, the simple main effect of pitch was confirmed ($p < 0.05$). This result shows that normal phonation with a high pitch tends to lower jitter; however, pressed phonation does not achieve this.

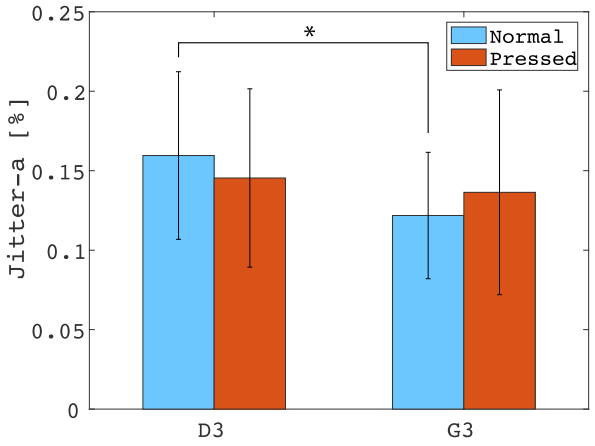


Figure 4. Mean and standard deviation of jitter-a for each pitch. The difference between the average values of normal and pressed phonations is larger in normal phonation (*$p < 0.05$).

4.4. Shimmer

The phonation mode did not cause a significant difference in the shimmer of all vowels. However, we observed a relationship between dynamics and phonation mode ($p < 0.05$) in the shimmer for [u]. Fig. 5 displays the mean and standard deviation of shimmer under each condition: dynamics (*mf, f*), phonation mode (Normal, Pressed), and vowel ([u]). When the dynamics is *forte*, the difference between the average values of normal and pressed phonations is larger. In this dynamics level, the simple main effect of the phonation mode was confirmed ($p < 0.05$).

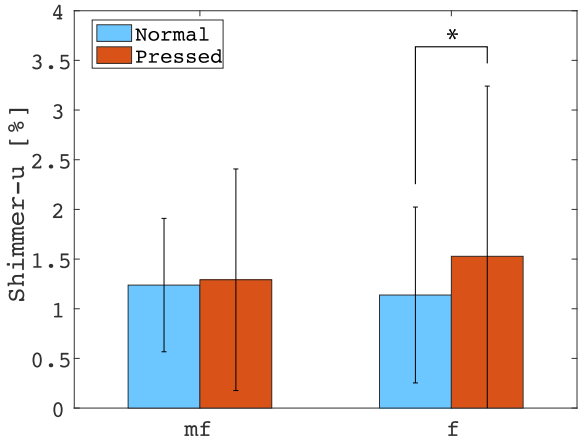


Figure 5. Mean and standard deviation of shimmer-u in each dynamics. The difference between the average values of normal and pressed phonation is larger in *forte* (*$p < 0.05$).

5. DISCUSSION

5.1. H1 − H2

Pressed phonation was predicted to be a less breathy voice obtained by the tightening of the throat. We first expected to observe a large difference in H1 − H2 for the two modes. However, in our analysis, we did not find any difference in the average of H1 − H2. This is possibly because the first formant frequency changed when subjects performed in normal phonation [12]. In other words, the changed first formant may have influenced the amplitudes of fundamental frequency and the second harmonic.

5.2. SPR values and the singer's formant

We expected higher SPR values for normal phonation because of the presence of the singer's formant. However, the average of the SPR in normal phonation was lower. We further observed the spectral slope, and the SPR value seems lower when the spectral slope is steeper. The SPR values can be influenced by both the singer's formant and spectral slope; for example, if the formant is fixed for all voices, the SPR values are affected only by spectral slope. Therefore, it is natural to assume that SPR values are influenced by both the singer's formant and the spectral slope with the interaction. In our case, we suspect that the SPR values of normal phonation were influenced more strongly by the spectral slopes than the singer's formant. In fact, we observed that only a few of the subjects demonstrated the singer's formant occasionally. Therefore, we suspect that their SPR values were more influenced by the spectral slope than by the singer's formant (which does not exist anyway).

5.3. Relationship of characteristics and intensity

We measured the intensity of each phonation using Praat and examined the correlation between intensity and each

characteristic. The result shows all characteristics treated in this study correlated with intensity. This correlation may make it difficult to compare normal and pressed phonation. In addition, we calculated delta dB, which is the difference intensity between two modes (*Pressed, Normal*). Table 1 shows the ratio of subjects who have positive delta dB under each condition. Under most conditions, this ratio is more than 50%. However, this ratio is less than 50% for the [u] vowel with G3. All characteristics were influenced by both intensity and phonation mode with interaction. For [u], the influence by intensity might have been weak; therefore, we confirmed a significant difference between normal and pressed phonation in HNR and SPR.

		a	E	u
D3	mf	66.7%	66.7%	60.0%
	f	73.3%	60.0%	53.3%
G3	mf	60.0%	53.3%	46.7%
	f	66.7%	53.3%	40.0%

Table 1. Ratio of subjects who have larger intensity in pressed phonation than normal phonation under each condition. The ratio for [u] with G3 is less than 50%.

6. CONCLUSION & FUTURE WORK

In this study, we investigated the relationship between several acoustic characteristics and pressed phonation in choral singing. The results show that pressed phonation in the [u] vowel has lower HNR values and higher SPR values. The pressed phonation in [u] confirmed the impression of hoarse voice qualities according to the acoustical feature analysis. In the future, we would like to find new acoustic characteristics instead of SPR to measure the singer's formant. Further, we would like to continue the measurement of the amateur choral singers. At that time, our issue will be control of intensity.

REFERENCES

[1] J. Sundberg and J. Gauffin, "Waveform and spectrum of the glottal voice source," *Frontiers of speech communication research* 1979, pp. 301–320.

[2] J. Sundberg, "Articulatory interpretation of the "singing formant"," *The Journal of the Acoustical Society of America*, vol. 55, no. 4, 1974, pp. 838–844.

[3] P. Alku and E. Vilkman, "A comparison of glottal voice source quantification parameters in breathy, normal and pressed phonation of female and male speakers," *Folia phoniatrica et logopaedica*, vol. 48, no. 5. 1996, pp. 240–254.

[4] D. H. Klatt and L. C. Klatt, "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *The Journal of the Acoustical Society of America*, vol. 87, no. 2. 1990, pp. 820–857.

[5] N. Henrich, C. d'Alessandro, B. Doval, and M. Castellengo, "Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency," *The Journal of the Acoustical Society of America*, vol. 117, no. 3. 2005, pp. 1417–1430.

[6] B. Paul and W. David, "Praat: doing phonetics by computer [computer program]. version 5.3.51," 2013. [Online]. Available:http://www.praat.org/

[7] E. Yumoto, W. J. Gould, and T. Baer, "Harmonics-to- noise ratio as an index of the degree of hoarseness," *The Journal of the Acoustical Society of America*, vol. 71, no. 6. 1982, pp. 1544–1550.

[8] M. Hyodo, "Diagnosis and treatment of speech disorders," *Nippon Jibiinkoka Gakkai Kaiho*, vol. 113, no. 10. 2010, pp. 818–821.

[9] M. Gordona and P. Ladefogedb, "Phonation types: a cross-linguistic overview," *Journal of Phonetics*, vol. 29, no. 383. 2001, p.406.

[10] J. M. Grey and J. W. Gordon, "Perceptual effects of spectral modifications on musical timbres," *The Journal of the Acoustical Society of America*, vol. 63, no. 5. 1978, pp. 1493–1500.

[11] K. Omori, A. Kacker, L. M. Carroll, W. D. Riley, and S. M. Blaugrund, "Singing power ratio: quantitative evaluation of singing voice quality," *Journal of Voice*, vol. 10, no. 3. 1996, pp. 228–235.

[12] S. Ternström and J. Sundberg, "Formant frequencies of choir singers," *The Journal of the Acoustical Society of America*, vol. 86, no. 2. 1989, pp. 517–522.

Effort in Interactions with Imaginary Objects in Hindustani Vocal Music – Towards Enhancing Gesture-Controlled Virtual Instruments

Stella Paschalidou,[1†] *Martin Clayton,*[2] *Tuomas Eerola*[2]

[1] Department of Music Technology & Acoustics Eng., TEI of Crete, Greece
[2] Department of Music, Durham University, UK
[†] pashalidou@staff.teicrete.gr

ABSTRACT

In Hindustani Dhrupad vocal improvisation singers often appear to engage with melodic ideas by manipulating intangible, imaginary objects and materials with their hands, such as through stretching, pulling, and pushing. This observation suggests that some patterns of change in acoustic features of the voice allude to the interactions that objects through their physical properties can afford and they are defined by the sensation of effort required for fighting against or yielding in to the resistance involved in the interaction.

The present study reports on the exploration of relationships between the voice, the interaction possibilities of malleable (elasticity) vs. rigid (weight/friction) objects and the physical effort they appear to require as perceived by an observer in Dhrupad vocal improvisation.

1. INTRODUCTION

With the recent shift towards embodied approaches in music performance, physical effort has often been highlighted as an important aspect of expressivity. In fact, the lack of expressivity in designed gesture-sound interactions has been attributed by [1] to the rupture of the so-called 'efforted-input paradigm'. However, systematic approaches to the role of effort still remain limited. In this study we understand effort as a concept which reflects the active or passive attitude of the person in fighting against or giving in to the physical conditions that influence the movement while trying to achieve an intentional task [2].

The current work is ultimately intended to enhance the design of mappings in Virtual Musical Instruments [3], in which imaginary objects are sculpted to produce sounds. It proposes to achieve this by taking advantage of the physical effort induced by familiar gestures associated with handling real objects in our natural environment and thus by making sound sculpting feel more 'physically plausible' [4].

To meet this end, the current paper uses the ecologically valid setting of real performances, in order to gain a deeper understanding of motor-based metaphors in the specific cases where a Dhrupad singer seems to manually interact with imaginary objects and materials (MIIOMs). Dhrupad is an 'oral' music tradition (in the sense that it does not depend on music notation) and relies heavily on improvisation, which is rule-based and conforms to the raga (modal) system. More specifically, the paper aims to examine whether physical effort and gesture interaction types are arbitrarily related to the voice or if they could be predicted through a small number of audio and/or movement features. MIIOMs offer a unique opportunity to study profound gesture-sound links that are not just related to the mechanical couplings of instrumental playing.

2. METHODOLOGY

The work makes use of a mixed methodological approach, combining qualitative and quantitative methods on original recordings of interviews, audio-visual material and 3D-movement data of vocal improvisations. The present study reports on the first part of the vocal improvisations by two different Dhrupad vocalists of the same music lineage, namely Afzal Hussain (raga Jaunpuri, scale: 1,2,*b*3,4,5,*b*6,*b*7, 18 minutes) and Lakhan Lal Sahu (raga Malkauns, scale: 1,b3,4,*b*6,*b*7, 23 minutes).

Hussain and Sahu were selected out of 15 performers who were recorded performing the slow non-metered section of an improvisation called alap, which is sung to a repertoire of non-lexical syllables. Only the alap improvisation was used in this study in order to concentrate on melodic factors without having to take into account the metrical structure or the lyrical content of the later stages of a raga performance. Real performances rather than designed experiments were used, in order to insure ecological validity in the approach of the subject. Thus, all recordings were conducted in domestic spaces in India (2010). The recording sessions included video, audio and motion recordings (10-camera Optitrack passive marker mocap system) and were followed by interviews.

First, we applied thematic analysis to the interview material to identify action-based metaphors, which informed the annotation process of the video material that followed. For the video analysis we relied on third-person observations that aimed at identifying, labeling and later classifying the audio-visual material in terms of recurrent types (categorical descriptors) of MIIOMs, as well as perceived effort levels (numerical) that appear to be exerted by the performer in a range between 0 (lowest) and 10 (highest).

The annotations (cross-validated by two choreographers /dancers) were used in the quantitative part of the analysis as the response values of models that were fit to the measured movement and sonic features in order to (a) estimate the annotated effort levels (numerical, through linear regression) and (b) classify gestures (categorical, through logistic regression), specifically discriminating between interactions with elastic vs. rigid objects. The features that were used for estimating the responses were computed by first extracting continuous (time-varying) movement and audio features from the raw movement and audio data respectively and then computing representative statistical global measures (such as mean, SD, min, max). The movement and audio features

reported in [5] were first used as a starting point, but then a number of alternative features were also explored that were meant to raise the explained variance of the estimated responses.

3. RESULTS

Two variations of models were developed for each task (effort estimation and gesture classification):
(a) a model that best fits each individual performer, thus better reflecting the idiosyncratic aspect of the individual singer and
(b) a model that can better describe shared cross-performer behaviours.

3.1. Effort estimation

3.1.1. Idiosyncratic schemes:

Different idiosyncratic schemes of associating the perceived physical effort with acoustic and movement features were identified through linear regression, that are based on the pitch space organisation of the raga and the mechanical strain of voice production.
- In the case of Hussain, the use of 5 non-collinear audio and movement features yielded a good fit (R^{2adj}) of about 60%. According to the results, higher bodily effort levels are required for slower movements that involve moving the hands wider apart and exhibit a larger speed variation. They are accompanied by singing melodic glides that start from lower degrees (closer to the lower tonic) and ascend to higher degrees (further away from the lower tonic) of the scale within the boundaries of each individual octave, in other words for larger melodic movements that move higher up in the octave. Pitch values are calculated here on a relative logarithmic scale and thus they describe glides associated with particular degrees of the melodic scale, better describing characteristic qualities of the specific raga.
- In the case of Sahu, the use of 4 non-collinear audio and movement features yielded an adequately good fit (R^{2adj}) of about 44%. According to the results, higher bodily effort is required for hand movements that exhibit a larger variation of hand divergence (change of speed in moving the hands apart or closer to each other), with a strong mean acceleration in the beginning (onset). They are accompanied by larger melodic glides that reach up to higher maximum pitches. In this case, pitch is expressed in an absolute logarithmic scale, which better reflects the mechanics of voice production. As the alap is organised based on a gradual ascent towards the pitch climax, pitch is here also representative of the alap macro-structure.

3.1.2. Generic scheme:

A more generic cross-performer estimation of effort was achieved by almost identical linear models for the two singers, yielding a good fit (R^2adj) of about 53% (with 5 features) for Hussain and 42% (with 4 features) for Sahu respectively. According to these models, high bodily effort levels are required by both singers for melodic movements that start from a lower and reach up to a higher pitch (on an absolute logarithmic scale, reflecting the mechanics of voice production). They are accompanied by movements which are slow on average but exhibit a large variation of speed, and in the specific case of Hussain by movements during which the hands move further away.

3.2. Gesture (MIIOM) classification

3.2.1. Idiosyncratic schemes:

Different modes of gesture class association to acoustic and movement features were identified through logistic regression, that are based on regions of particular interest in the raga pitch space organization and analogous cross-domain morphologies.
- In the case of Hussain, the use of 5 non-collinear audio and movement features yielded a high classification rate (AUC) of about 95%. According to the results, it is more likely that interactions with elastic objects are performed by hand gestures that exhibit a low absolute mean acceleration value and a large variation in the hands' divergence (speed in moving the hands apart). They are associated with slower and larger melodic movements that ascend to a higher degree of the scale. The highest degree of the scale (7^{th}) in raga Jaunpuri is the most unstable and therefore it dictates a change in pitch movement direction from ascent to descent (i.e. a double pitch glide, changing direction on the 7^{th} degree), which is similar to the direction change of the hands when interacting with an elastic object. It could be therefore suggested that the MIIOM type is associated with the grammatical rules of the raga.
- In the case of Sahu, the use of 4 non-collinear audio and movement features yielded a high classification rate (AUC) of about 80%. According to the results for Sahu, interactions with elastic objects are more likely to be performed with pitch movements of a larger interval and larger duration and with the hands moving faster and remaining bound to each other.

3.2.2. Generic scheme:

A more generic cross-performer MIIOM classification was achieved by almost identical general logistic models for the two singers, yielding a good fit (AUC) of about 86% (with 3 features) for Hussain and 78% (with 4 features) for Sahu respectively. According to these models, interactions with elastic objects are more likely to be performed at lower pitches for larger melodic movements (calculated on an absolute logarithmic scale), and with the hands moving further apart for Hussain and less apart but faster in the case of Sahu.

4. CONCLUSIONS

Based on the results, we rejected the null-hypothesis that gesture and effort are unrelated to the melody and found statistically significant movement and sound features that (a) best fit each individual performer and (b) describe the phenomena in the most generic way across performers in estimating effort and classifying MIIOMs.

Overall, models were more heavily dependent on acoustic rather than movement features and there was a bigger overlap of acoustic rather than movement features between models. This observation possibly highlights a stronger idiosyncratic factor in how MIIOMs are manually performed, which renders the task of describing them through a small number of movement features as non-trivial.

Findings indicate that despite the flexibility in the way a Dhrupad vocalist might use his hands while singing, the high degree of association between classes of virtual interactions and their exerted effort levels with melody provides good evidence for non-arbitrariness; this may reflect the dual nature of mapping in being associated with the mental organisation of the melodic context, the mechanical strain of vocalization, the alap

macro-structure and shared cross-modal morphologies. This conclusion leads us to further suggest that in designing an electronic musical instrument, a more flexible mapping scheme should be aimed for, which would enable a performer to switch between mapping modes.

By taking an embodied approach and mapping effort to a combination of features from both domains (auditory and movement), this work can contribute to the enhancement of mapping strategies in empty-handed artificial interactions on the grounds of physical plausibility and effort in sound control; novel interaction paradigms can be developed which are inspired by our interaction with the real world.

REFERENCES

[1] J. d'Escriván, "To sing the body electric: Instruments and effort in the performance of electronic music", *Contemporary Music Review*, vol. 25, no. 1-2, pp. 183-191, 2006.

[2] P. Hackney, *Making connections: Total body integration through Bartenieff fundamentals*. Routledge, New York, NY, 1998.

[3] A. Mulder, *Design of virtual three-dimensional instruments for sound control*, 1st ed. Ottawa: National Library of Canada = Bibliothèque nationale du Canada, 1999.

[4] N. Castagné, C. Cadoz, *A Goals-Based Review of Physical Modelling*, In Proceedings of the International Computer Music Conference, Barcelona, Spain, pp. 343-346, 2005.

[5] K. Nymoen, R. Godøy, A. Jensenius and J. Torresen, "Analyzing correspondence between sound objects and body motion", *ACM Transactions on Applied Perception*, vol. 10, no. 2, pp. 1-22, 2013.

The Role of Spectral Envelope Ratio within the Perception of Sound-Source Size: a Psychophysical and Affective Investigation

Joseph Plazak,[1†] *Stephen McAdams*[2]

[1]Illinois Wesleyan University
[2]McGill University
[†]jplazak@iwu.edu

ABSTRACT

Spectral envelope ratio (SER), amongst other acoustic features, has been found to play a key role for inferring the perceived size (i.e. largeness vs. smallness) of various sound sources, including speech, animal, and musical sounds. Further, it is thought that SER may play an important role in inferring the register, and thus implicit affective associations, of some musical instruments. In order to better understand the psychophysical and affective implications of SER manipulations within music perception, we devised an "ask and verify" experiment involving both expert and novice listeners. First, an expert panel listened to various musical instrument tone pairs containing various SER manipulations that were synthesized via the TANDEM-STRAIGHT vocoder. The expert panel reached consensus on eight affective terms for describing SER manipulations, which included: tension, complaintiveness, impatience, authoritativeness, pride, age/youthfulness, mocking, and calmness. In a follow-up study, 41 participants were presented with similar tone pairs representing five musical instruments and a variety of SER manipulations. Participants were tasked with verifying the relevance of the expert panel's affective descriptors. These results indicated SER manipulations that were consistent with an increasing sound source size were correlated with subjective ratings of patience, calmness, age, authoritativeness, and pride, whereas SER manipulations consistent with a decreasing sound source size were correlated with increased tension, complaining, and mocking. These results uncover several important considerations for understanding the role of SER within the perception of musical timbre. In particular, our results suggest that manipulations of SER may lead to patterned psychophysical and affective perceptions, and therefore, SER may be a useful metric for understanding the perception and affective associations of musical instrument registers. Further, SER may also provide useful insight for understanding the perception of various musical instrument combinations, as well as larger scale music orchestrations.

A Performance Evaluation of the Timbre Toolbox and the MIRtoolbox on Calibrated Test Sounds

Savvas Kazazis,† *Nicholas Esterer, Philippe Depalle, Stephen McAdams*

Schulich School of Music, McGill University
†savvas.kazazis@mail.mcgill.ca

ABSTRACT

We evaluate the accuracy of the Timbre Toolbox (v. 1.2) and the MIRtoolbox (v. 1.6.1) on audio descriptors that are putatively related to timbre. First, we report and fix major bugs found in the current version of the Timbre Toolbox, which have gone previously unnoticed in publications that used this toolbox as an analysis tool. Then, we construct sound sets that exhibit specific spectral and temporal characteristics in relation to the descriptors being tested. The evaluation is performed by comparing the theoretical (real) values of the sound sets to the estimations of the toolboxes.

1. INTRODUCTION

The Timbre Toolbox [1] and the MIRtoolbox [2] are two of the most popular MATLAB [3] toolboxes that are used for audio feature extraction within the music information retrieval (MIR) community. They have been recently evaluated according to the number of presented features, the user interface and computational efficiency [4], but there have not been performance evaluations of the accuracy of the extracted features. The aim of this paper is: (1) to detect and summarize the bugs in the current version of the Timbre Toolbox and (2) to evaluate the robustness of audio descriptors these toolboxes have in common and that are putatively related to timbre. For this purpose, we synthesized various sound sets using additive synthesis, calculated the theoretical (real) values of each descriptor tested, and compared these values with the estimations of the toolboxes. Section 2 summarizes the bugs found in the current publically available version of the Timbre Toolbox (v. 1.2). Section 3 describes the construction of the sound sets used for evaluating the performance of the MIRtoolbox (v. 1.6.1) and a beta version of the Timbre Toolbox, which fixes the reported bugs. Section 4 presents the results of the evaluation and Section 5 summarizes our findings.

2. POINTS OF CONSIDERATION AND BUG FIXING IN THE TIMBRE TOOLBOX

In this section, we report the bugs found in the current version of the Timbre Toolbox and some issues related to user interaction. The Timbre Toolbox incorporates the following sound models of the time-domain signal for extracting audio descriptors: the temporal energy envelope; the short-term Fourier transform (STFT) on a linear amplitude scale (STFTmag) and a squared amplitude scale (STFTpow); the output of an auditory model based on the concept of the Equivalent Rectangular Bandwidth, which is either calculated using recursive gammatone filters (ERBgam), or their finite impulse response approximation using the fast Fourier transform (ERBfft); and a sinusoidal harmonic model [1].

In some cases, especially when the amplitude of the lower frequencies is lower than the upper ones, the harmonic representation using the default amplitude threshold for detecting harmonics will not analyze even strictly harmonic sounds. Furthermore, the default analysis limit of 20 harmonics could also be problematic for analyzing low-frequency sounds having spectral energy that increases with harmonic number. However, this scenario is very unlikely to occur in natural sounds, but it is still possible with synthetic sounds used in psychoacoustic experiments (e.g., [5]) or in electroacoustic music. Another conceptual bug is the estimation of inharmonicity: according to Eq. 1, which is presented in [1], a signal with a fundamental frequency of 100 Hz and a partial at 150 Hz will be less inharmonic than a signal with the same fundamental and a partial at 190 Hz even though the partial of the second signal is only detuned by 10 Hz below the next harmonic.

$$\text{inharm} = \frac{2}{f_0} \frac{\sum_{h=1}^{\text{H}} (f_h - h f_0) a_h^2}{\sum_{h=1}^{\text{H}} a_h^2} \tag{1}$$

In v. 1.2, the end user only had access to summary statistics, and as such it was not possible to evaluate the time-varying patterns of audio descriptors. Furthermore, the export format of the results was a text file. This did not facilitate further processing of the results especially in the case of a batch analysis where the output consists of several text files. Also, MATLAB ran out of memory when the Timbre Toolbox processed long audio files.

According to Peeters et al. [1], the window that should be used for the harmonic analysis is a Blackman window. However, in the toolbox's implementation, the window is a boxcar (i.e., no window weighting at all), but we also noticed that the removal of the window's energy contribution to the input sound was implemented incorrectly. Furthermore, some calculations on audio descriptors returned the results in normalized frequency (including the spectral centroid) without warning the user and led to misinterpretations (e.g., [6]).

Although the actual sampling rate is read directly from the file, in some sound models it was not actually used: the parameters related to the FFT analysis were specified according to a fixed sampling rate of 44.1 kHz no matter the actual sampling rate of the input file. Finally, in most of the employed

sound models, the computations of spectral spread, skewness, kurtosis and spectral slope were implemented incorrectly.

The analysis results presented in this paper are based on a beta version of the Timbre Toolbox that fixes and takes into consideration all of the above-mentioned points except the calculation of inharmonicity and the threshold settings used in the harmonic representation.

3. CONSTRUCTION OF THE TEST SOUND SETS

The sounds were constructed using additive synthesis, which allows for a direct computation of the audio descriptors. Each sound set was designed to exhibit specific sound qualities that are directly related to the descriptors being tested. In this way, we are able to systematically test the performance of the toolboxes by tracking the circumstances under which certain audio descriptors are poorly calculated. All sounds were synthesized at 44.1 kHz with 16-bit resolution and peak amplitude of 6 dB relative to full scale (dBFS). To avoid spectral spread induced by an abrupt onset and offset when performing the FFT on these synthetic sounds, we applied a 10-ms raised inverse cosine ramp to all sounds except the ones used to test the attack time and the attack and decrease slopes. Durations were fixed at 600 ms and all sounds contained harmonics up to (but not including) the Nyquist frequency.

We used the following fundamental frequencies for all the sound sets except those related to the temporal energy envelope: C#1 (34.65 Hz), D2 (73.42 Hz), D#3 (155.56 Hz), C4 (261.63 Hz), E4 (329.63 Hz), F5 (698.46 Hz), A5 (880 Hz), F#6 (1479.98 Hz), G7 (3135.96 Hz) and B7 (3951.07 Hz). The C4 was slightly detuned from 261.63 Hz to 258 Hz in order to match exactly the frequency of an FFT bin and to test whether the estimations would be improved; for a sampling frequency of 44.1 kHz and an FFT size of 1024 samples (default setting of the Timbre Toolbox) the bins are spaced 43 Hz apart. We used such a wide frequency range because as the fundamental frequency increases and approaches the Nyquist limit, the number of "significant" FFT bins decreases, which may affect the accuracy of the results, especially in the presence of noise.

3.1. Attack Time, Attack Slope and Decrease Slope

The Timbre Toolbox uses the "weakest effort method" for estimating the attack time and the attack and decrease slopes [1], whereas the MIRtoolbox uses a similar method based on Gaussian curves [2]. In these adaptive methods, the threshold energy level that the signal must surpass is not fixed, but is determined as a proportion of the maximum of the signal's energy envelope. The attack and decrease slopes are then estimated as the average temporal slope during the start and end times of the attack portion. An 'effort' is defined as the time it takes for the signal to go from one threshold value to the next. It is therefore logical to assume that if the signal varies rapidly and non-linearly during the attack time, the true attack time values may be poorly estimated.

For testing the accuracy of this method, we constructed nine attack envelopes for each of ten logarithmically spaced attack times ranging from 1 to 300 ms. The shape of the envelopes was determined by:

$$y(t) = mt^{\mathrm{b}} \tag{2}$$

where m controls the slope of the attack time and b is a curvature constant which was assigned the following values: 3, 2.5, 2 and 1.5 for an exponential shape; 1 for a linear shape; and 0.67, 0.5, 0.4 and 0.33 for a logarithmic shape. The attack envelopes were then applied to a flat harmonic spectrum with a fundamental frequency of 258 Hz and a total duration of 600 ms. A similar procedure was used for testing the estimations of decrease slope.

3.2. Spectral Centroid

For this sound set, we used a flat spectrum with octave-spaced harmonics and included in the above-mentioned set a lower fundamental of C0 (16.35 Hz). In order to systematically test the accuracy of spectral centroid estimation, we iteratively removed just one harmonic from the initial spectrum up to the last one for every fundamental. This way, the sounds generated from the last fundamental just contain a single frequency component, because there is only one harmonic present due to the Nyquist limit, and therefore the spectral centroid ideally should match the value of the fundamental frequency estimation.

3.3. Spectral Spread, Skewness, Kurtosis and Roll-off

For testing the estimations of spectral spread, skewness, kurtosis, and roll-off, we designed a sound set in which the sounds vary by fundamental frequency and according to spectral slopes. By precisely controlling the spectral slopes, we directly alter in a predictable way the higher statistical moments of the spectrum and the frequency below which 95% of the signal energy is contained. In our analysis, we took into account the fact that the MIRtoolbox uses a default value of 85%. For every fundamental, we constructed a spectrum that contained both odd and even harmonics with a $1/n^2$ power decrease, where n denotes the harmonic number. Then in nine steps we altered linearly the energy distribution of the harmonics until we reached a flat spectrum. The same procedure was repeated by starting from a flat spectrum and reaching in nine steps a positive slope of the harmonics which had an n^2 power increase.

3.4. Harmonic Spectral Deviation and Spectral Irregularity

Spectral deviation (in the Timbre Toolbox) and spectral irregularity (in the MIRtoolbox) are the same descriptors but are computed slight differently with respect to a scaling factor. MIRtoolbox offers two estimation methods based on Jensen [7] and Krimphoff et al. [8]. Here, we only tested the estimation based on Krimphoff's method (Eq. 3.), which is the only option available in the Timbre Toolbox. For every fundamental, we started from a flat spectrum that only contained the fundamental with even harmonics, and we gradually increased

the level of the odd ones until we reach a flat spectrum in ten steps.

$$\text{dev} = \sum_{h=2}^{\text{H-1}} \left| a_h - \frac{a_{h-1} + a_h + a_{h+1}}{3} \right| \quad (3)$$

3.5. Spectral Flatness

To evaluate the accuracy of the estimations of spectral flatness, we applied a Gaussian spectral window centered at the middle harmonic to a flat spectrum that contained both odd and even harmonics, and progressively altered its standard deviation in ten steps so that the last window resulted in an extremely peaky spectrum. For altering the width of the window we used the following coefficients, which are proportional to the reciprocal of the standard deviation: 0.5, 1, 1.5, 2, 3, 4, 5, 6, 7 and 8. This process was done for the whole range of fundamentals.

3.6. Inharmonicity

This sound set is similar to 3.2, but here we used inharmonic spectra. The inharmonic components were kept fixed in the whole sound set, and were spaced according to an inharmonicity coefficient that controlled the amount of deviation from each harmonic, which varied linearly from 0 to 0.5 with respect to the harmonic number. Inharmonicity was increased by gradually increasing the amplitude of the inharmonic components instead of increasing their deviation from the harmonics. The inharmonic components were initially attenuated with a $1/n^2$ envelope to reach a flat spectrum in ten steps μ by gradually increasing linearly their energy distribution.

4. RESULTS

We evaluate toolbox performance by analyzing the sound sets with each toolbox and calculating the normalized root mean squared (RMS) error between their output and the theoretical values. The theoretical values were calculated using either the power or magnitude scale depending on the input representation being tested. MIRtoolbox's default input representation using 'mirspectrum' is based on a STFT with a Hamming window and a half overlapping frame length of 50 ms, which is similar to the 'STFTmag' representation used in the Timbre Toolbox. For the Timbre Toolbox, we tested all the available input representations because there is no default option. For analyzing the sounds, we used the default settings of each toolbox, and the summary statistics from the frame-by-frame analysis were derived using the median values.

4.1. Temporal Energy Descriptors

MIRtoolbox uses two estimation methods for calculating the attack and decrease slopes: 'Diff', which computes the slope as a ratio between the magnitude difference at the beginning and end of the attack period and the corresponding time difference; and 'Gauss', which is similar to Peeters' method [1]. Table 1 shows the results of the error analysis. The observed general trend for both toolboxes was that short attack times (about less than 40 ms) were significantly overestimated, whereas longer attack times were mainly underestimated. The Timbre Toolbox also systematically estimated the exponential attacks as being longer than the logarithmic attacks.

Descriptors	**Timbre Toolbox**	**MIRtoolbox** (Diff / Gauss)
Attack Time	24.40	21.57
Attack Slope	36.85	36.15 / 36.82
Decrease Slope	37.31	37.53 / 37.36

Table 1. RMS error (%) of temporal energy descriptors.

4.2. Spectral and Harmonic Descriptors

Although we tested the accuracy of extracted descriptors on all sound sets, due to space limitations, the evaluation results presented in Table 2 are based only on the designated sets for each descriptor, which were presented in the previous section. Also, we only report the most accurate results (i.e., the minimum RMS error) among the Timbre Toolbox's different input representations. In the following, we present a qualitative inspection of the errors with respect to the sound sets.

Centroid: MIRtoolbox always overestimates slightly the centroids, whereas the Timbre Toolbox returns accurate results for fundamentals of 65.4 Hz and above.

Higher-order moments of the spectrum and roll-off: the MIRtoolbox was numerically unstable returning 'Not a Number' (NaN) in the estimation of spectral centroid for the sets with fundamentals of 34.65 Hz and 73.42 Hz. Table 2 summarizes the results after removing the sounds for which MIRtoolbox returned NaNs. Timbre Toolbox's STFTpow representation provides overall the most accurate estimations even when all sounds were included in the analysis, in which case it produced a 1.37% RMS error for spectral roll-off.

Spectral Flatness: MIRtoolbox again returned Not a Number for some of the sounds with fundamentals of 34.65 Hz and 73.42 Hz, and although this sound set was not designed to test the estimation of spectral irregularity, MIRtoolbox did not provide any results for the estimation of this descriptor and exited with an error message. We also noted that in both toolboxes, as the fundamental frequency increases and spectral spread decreases, the estimations of spectral spread become more erroneous, although limited within a small margin. Although the spectral centroid and higher moments were estimated quite accurately in both toolboxes, the estimation of spectral flatness was inaccurate.

Spectral irregularity (or *deviation*): Harmonic spectral deviation is only available in the harmonic representation of the Timbre Toolbox. However, we were not able to run the analysis on the whole sound set using the default amplitude threshold setting for harmonic detection: as the fundamental frequency increases, the settings should be lowered, otherwise the sound will not be further analyzed (in the beta version tested here, the user gets warned whenever this situation occurs). The MIRtoolbox also proved to be erroneous for the estimation of this

descriptor. However, both toolboxes returned quite accurate results for the spectral centroid and higher-order moments for this sound set.

Inharmonicity: The estimations of inharmonicity could not be quantitatively evaluated due to the current behavior of the Timbre Toolbox, as mentioned previously, and the unavailability of the precise equation used by MIRtoolbox. Qualitatively, and given the way this sound set was constructed (section 3.6), we expect the estimation of inharmonicity to increase for the subsets of each fundamental. Fig. 1 shows the estimations of the MIRtoolbox, which seem to be more plausible after the fifth set of fundamentals (i.e., from F5 up to B7, section 3).

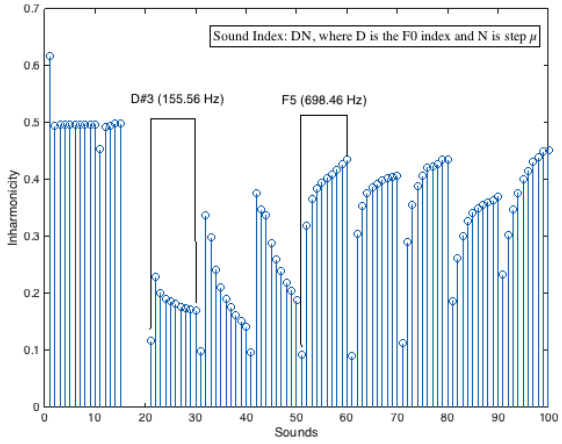


Figure 1. Inharmonicity estimation in the MIRtoolbox. The horizontal axis indicates the sound index DN, and the vertical axis the relative deviation of the partials from purely harmonic frequencies. The missing values correspond to NaN.

5. CONCLUSIONS

Before evaluating the accuracy of the toolboxes, we reported and fixed in a beta version the major bugs, configuration, and presentation issues that were encountered in the current version of the Timbre Toolbox (v. 1.2). Our evaluation on synthetic test sounds shows that for spectral descriptors, the Timbre Toolbox performs more accurately and on some sound sets outperforms the MIRtoolbox, with the short-term Fourier transform power representation (STFTpow) being overall the most robust. The estimations of spectral centroid and higher order moments of the spectrum were quite accurate with small errors except the estimation of spectral flatness, which both toolboxes estimated erroneously. The Timbre Toolbox failed to analyze some sounds using the harmonic representation with the default settings even though all sounds were strictly harmonic. In the beta version tested here, if this situation occurs, the estimation of fundamental frequency is automatically set to zero, which affects the calculation of all descriptors related to this representation. However, the user receives a warning message in order to alter the default settings appropriately. The MIRtoolbox's estimations of spectral centroid

Descriptors	**Timbre Toolbox**	**MIRtoolbox**
Centroid	01.21 (STFTpow)	03.56
Spread	00.00 (STFTpow)	02.95
Skewness	02.06 (STFTmag)	03.82
Kurtosis	04.31 (STFTmag)	06.87
Roll-off	00.00 (STFTpow)	01.57
Flatness	34.87 (ERBgam)	51.82
Irregularity	N/A	31.36

Table 2. RMS error (%) of spectral energy descriptors. In the Timbre Toolbox, spectral irregularity could not be evaluated after the fifth set of fundamentals (section 3).

on some sounds, and spectral irregularity and inharmonicity on a specific sound set proved to be numerically unstable returning NaN, or exiting with error messages without providing any results. For descriptors that are based on the estimation of the temporal energy envelope, both toolboxes perform almost equally but poorly. We noticed that in this case the errors depend both on the attack or decay times and on the shape of the slopes. The test sound sets are available at:https://www.mcgill.ca/mpcl/resources-0/supplementary-materials

REFERENCES

[1] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The Timbre Toolbox: Extracting Audio Descriptors from Musical Signals," *J. Acoust. Soc. Am.,* 130, pp. 2902-2916, 2011.

[2] O. Lartillot, *MIRtoolbox 1.6.1 User's Manual,* Technical report, Aalborg University, Denmark. December 2014.

[3] http://www.mathworks.com

[4] D. Moffat, D. Ronan, and J. D. Reiss, "An Evaluation of Audio Feature Extraction Toolboxes," in *Proc. DAFx,* Trondheim, Norway, 2015.

[5] A. Caclin, S. McAdams, B.K. Smith, and S. Winsberg, "Acoustic Correlates of Timbre Space Dimensions: A Confirmatory Study Using Synthetic Tones," *J. Acoust. Soc. Am.,* 118 (1), pp. 471-482, 2005.

[6] C. Douglas, *Perceived Affect of Musical Instrument Sounds,* Master's thesis, McGill University, Montreal, Canada. June 2015.

[7] K. Jensen, *Timbre Models of Musical Sound: From the model of one sound to the model of one instrument*, PhD thesis, University of Copenhagen. 1999.

[8] J. Krimphoff, S. McAdams, and S. Winsberg, "Caractérisation du timbre des sons complexes. II : Analyses acoustiques et quantification psychophysique," *Journal de Physique,* 4(C5), 625-628, 1994.

Meta-Analysis of Studies on Timbre Dissimilarity Based on Spectrotemporal Modulation Representations

Etienne Thoret,† *Philippe Depalle, Stephen McAdams*

Schulich School of Music, McGill University, Montreal, QC, Canada
†etienne.thoret@mcgill.ca

ABSTRACT

Many studies have investigated the perception of musical instrument timbre through dissimilarity ratings tasks (Grey, 1977, Grey & Gordon, 1978, McAdams et al., 1995, Iverson & Krumhansl, 1993; Lakatos, 2000; Krumhansl, 1989). In these experiments, subjects were asked to rate the dissimilarity of pairs of musical instrument sounds equalized in pitch, loudness and duration. It is now well known that audio descriptors such as attack time and spectral centroid, among others, account well for the dimensions of the timbre spaces underlying these dissimilarity ratings. Nevertheless, it remains very difficult to reproduce these perceptual judgments from distances computed on acoustical representations such as the waveform or the spectrogram. It has been shown recently that spectral and temporal modulations of a sound are also relevant to account for the perception of musical instrument timbre (Patil et al., 2012; Elliott, Liberty & Theunissen, 2013; Thoret et al., 2016). These acoustic features basically correspond to the spectral and temporal regularities such as the presence of a vibrato or harmonicity. Neurophysiological studies have revealed that these acoustic features are probably extracted at the level of the primary auditory cortex (Shamma, 2001), suggesting that they could be relevant for judging the similarity between two sounds. From a mathematical point of view, several representations are associated with spectrotemporal modulations. One is the Modulation Power Spectrum (MPS), which corresponds to the 2D Fourier transform of a spectrogram. Here we investigate the extent to which this representation can be used to account for the perceptual dissimilarity between musical instrument sounds. We re-analyzed the results of former studies on timbre spaces (Grey, 1977, Grey & Gordon, 1978, McAdams et al., 1995, Iverson & Krumhansl, 1993; Lakatos, 2000; Krumhansl, 1989) by using the MPS to compute distances between pairs of sounds. We found that the correlations between MPS-based distances (e.g., Euclidean, Itakura-Saito or thresholded/non-thresholded) and the perceptual dissimilarities are fairly high for certain timbre spaces (e.g. Iverson & Krumhansl, 1993) but low for others (e.g. Grey, 1977), which would suggest that relevant audio descriptors could be extracted from MPS, but that the full MPS representation is probably not used per se. Above all, these results reinforce the promising use of Modulation Power Spectrum representations to investigate the perception of musical instrument timbre (Patil et al., 2012; Elliott, Liberty & Theunissen, 2013; Thoret et al., 2016).

Relating Psychoacoustic Thresholds to Auditory Experience

Noam Amir,[1†] *Eitan Globerson,*[2] *Yael Zaltz*[1]

[1]Tel Aviv University
[2]Jerusalem Academy of Music and Dance
[†]noama@post.tau.ac.il

ABSTRACT

This paper describes an ongoing series of experiments in which we have been applying psychoacoustic tests to examine various capabilities of laymen, musicians in general, and specific classes of musicians. In previous work we have shown that overall, musicians (mainly pop/rock musicians in our experiment) achieve better thresholds in some tasks related to musical attributes of sound, such as pitch and duration, but not spectrum or intensity. This led us to pursue several further directions. In one experiment we tried to determine whether specific musical training improved performance on a specific psychoacoustic task. To this end we compared rock/pop guitarists to percussionists on the first two of the above tasks: pitch and duration. While both groups performed better than laymen, guitarists performed better than percussionists on pitch tasks but slightly worse on duration tasks. This demonstrates that improved performance is directly related to the specific domain of musical expertise. In another experiment we wished to observe the influence of the stimuli on experiment outcome: instead of using steady tones in a pitch threshold experiment, we used modulated tones resembling sung vibrato. Results of experiments on laymen showed that thresholds were about 2.5 times higher for vibrato stimuli, indicating that vibrato can serve to mask errors in tuning. It will be very interesting to see if musicians are influenced in the same way - initial results show that this indeed holds for musicians as well, though a full experiment on musicians has not been conducted yet. Finally, we present some initial results on pitch-matching tasks: both sung-tone matching and matching with sliders. Results show that in general, musical training improves these abilities, however different paradigms can produce interesting and often surprising results.

Generation of Perceptible Rhythmic Timbral Events

Felix Dobrowohl,[†] *Roger Dean, Andrew Milne*

MARCS Institute

[†]F.Dobrowohl@westernsydney.edu.au

ABSTRACT

Tonal stimuli are traditionally described by the relatively obvious dimensions of pitch and loudness, and some more ambiguous ones, such as brightness, roughness, fullness, intensity flux, density, which are generally grouped under the broad term of timbre. Most research so far has tried to approach timbre by focussing on the perceptual and/or the acoustical properties of sounds. Whilst obviously linked, few studies have actually looked at timbre from a musician's/composer's perspective, or used an approach that could make more sense in a musical setup. For musicians it is hard to translate perceptual based results of academic research into a musical interface. For example, a bass in an Electronic Dance Music (EDM) song or a Guitar-Drone/Pad in a Post-rock piece relies solely on timbral changes over an extended period of time. These changes are however controlled by seemingly arbitrary parameters of sound effects and manipulations. This study examines how timbre changes resulting from the manipulation of such parameters affects the perception of tonal events and how they relate to each other. More specifically these parameters include: the amount of partials in relation to the fundamental (which may help to determine "brightness"), the ratio of even- to odd-numbered harmonics (Square/Saw oscillation), the spectral shaping resulting from different plucking positions on a string (e.g.,"comb filtering"), the detuning of partials from perfect hamonicity and the combination of detuning and distortion/overdrive. These parameters are sonified by a self-built additive synthesizer and they are the timbre descriptors used for analytical purposes and to account for participants' responses.

The experiment looks at how much each of the chosen timbre manipulators has to be changed for a subject to notice a new event in a gap-free sound stimulus. Pilot studies so far have revealed perception asymmetries in the direction of stimulus change, as well as a notable sensitivity to phase position of partials (particularly in stimuli with slightly inharmonic partials) which contradicts the common belief that humans are insensitive to phase. Additionally the step sizes appear to be strongly dependant on the starting value of the chosen parameter. Overall, the results advance our understanding of sound perception and pave the way for future work designed to test listeners' ability to detect rhythmic sequences of timbral only stimuli (meaning no change in fundamental frequency or perceived loudness).

Musical Acoustics, Timbre, and Computer-Aided Orchestration Challenges

Aurélien Antoine[†] *and Eduardo R. Miranda*

Interdisciplinary Centre for Computer Music Research (ICCMR), Plymouth University, UK
[†]aurelien.antoine@postgrad.plymouth.ac.uk

ABSTRACT

Research on timbre has produced a lot of work over the last 50 years, whether on analysis of single tone, instrument timbre, synthetic timbre, or perception and emotion. However, there is still a lot to investigate for polyphonic timbre. This aspect is perhaps one of the biggest challenges in the field of computer-aided orchestration. This paper reports on the development of a system capable of automatically identify perceptual qualities of timbre within orchestral sounds. Work in this area could enrich current and future computing systems designed to aid musical orchestration.

1. INTRODUCTION

Timbre is a complex and multidimensional attribute of sound, whose definition has been largely discussed among the research community, see [1] or [2] for examples. However, a standard definition often cited in papers related to timbre is the one proposed by the American National Standards Institute [3]: *"Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar"*. In summary, timbre is defined as all the sound attributes, except pitch, loudness, and duration, which allow us to distinguish and recognize that two sounds are dissimilar [4]. However, some research challenges this definition and suggests that timbre, pitch, and loudness can interfere in the auditory sensation [5, 6, 7]. Furthermore, several works have demonstrated the importance of acoustic features in defining musical timbre [8, 9, 10].

We are particularly interested in the perception of musical timbre. Research into this field has identified various important audio features to represent timbral perception [11, 12, 13, 14, 15]. However, timbre can also be described using verbal descriptors [16, 4, 17]. For example, terms like brightness or roughness are words from everyday language used to describe perceived musical timbres. These terms can be more intuitive for musicians, composers, and audio engineers, than their acoustic correlates (e.g. spectral centroid, critical bands, etc.).

The system presented in this paper is built upon the existing research about the correlation between acoustic features and verbal description of timbre qualities. A similar approach has been used in [18] and [19], but they utilized a visual representation approach of timbre properties. We believe that using verbal descriptors of timbral attributes can aid in making the tool accessible to a broad audience by alleviating the need to have expertise in acoustics or psychoacoustics.

The aim of this paper and the system presented within is to propose an effective and direct application of the work done on semantic description of timbre perception. The development of such a system could aid in standardizing the metrics for perceived responses of timbral qualities, and also enrich computer-aided orchestration systems.

2. AUTOMATIC TIMBRE CLASSIFICATION SYSTEM

This section introduces the technical aspects of the automatic timbre classification system. First, we discuss the verbal descriptors currently implemented in the system, with details of their corresponding acoustic features. Then, we report on the design of our classification algorithms developed to identify the dominant timbral content within an audio file.

2.1. Verbal Descriptors and Acoustic Features

For our timbre classification system, we chose to represent timbral qualities by using verbal descriptors. There are currently five timbral attributes implemented in the presented system: *breathiness*, *brightness*, *dullness*, *roughness* and *warmth*. Corresponding acoustic features for each attribute are detailed below.

2.1.1. Breathiness

To identify the level of *breathiness* within an audio file, we need to calculate the amplitude of the fundamental frequency against the noise content, and also the spectral slope [20, 21]. The bigger the ratio between amplitude of the fundamental and the noise content, the breathier the sound.

2.1.2. Brightness

The acoustic correlates for the attribute *brightness* are the spectral centroid and the fundamental frequency [22, 23]. The higher the spectral centroid, the brighter the sound.

2.1.3. Dullness

Similar to brightness, to identify the *dullness* of a sound we need to calculate its spectral centroid. However, in this case, a low spectral centroid value suggests that the sound is dull [24].

2.1.4. Roughness

To determine the *roughness* index of a sound, we need to calculate the distance between adjacent partials in critical bandwidths and also the energy above the 6th harmonic [25, 26, 8].

2.1.5. Warmth

To calculate the *warmth* of a sound, we need to calculate its spectral centroid and retrieve the energy in its first three harmonics [4, 27]. A low spectral centroid and a high energy in the first three harmonics suggest that the sound is warm.

2.1.6. Acoustic Features Analysis

We have developed our acoustic features analysis algorithm within the `Matlab`[1] environment. We also utilized some functions from the `MIRtoolbox`[2] [28]. This toolbox proposes several `Matlab` functions designed specifically for retrieving and extracting various musical features from audio files.

Our system starts by computing the spectrum of the audio file input by the user, using the `mirspectrum` function from the `MIRtoolbox`. Then, the system calculates the acoustic features for each attribute as described previously. Finally, the analysis returns a value for each attribute, which are all stored in a singular vector that will be used by the classification algorithms, described in the following section.

2.2. Classification Algorithms

There is a significant amount of work about the perception and description of musical timbre, upon which this work is based. However, there are no universally agreed metrics for classifying audio samples according to perceived responses of timbre quality. Therefore, before the development of a classification algorithm, a comparative scale for each timbral attribute must be established.

2.2.1. Comparative Scale

We began with gathering data in order to establish the scale of each timbral attribute. Therefore, we collected over 250 audio recordings of various orchestral pieces and split each audio recording into 1, 2, 3 and 4 seconds long audio files. We chose to go up to 4 seconds as it could represent a bar's worth of music, or it can correspond with the length of a computer-aided orchestration system's output. The acoustic features analyzed are time-based and therefore longer durations could omit important data and produce inaccurate indexes. As a result, we analyzed over 236 000 audio files and collected values for each timbral attribute.

We analyzed the dataset in order to retrieve various statical values, such as minimum value, maximum value, standard deviation, and the distribution of the values for each attribute. This allowed us to establish a scale for each attribute and therefore develop a normalization algorithm in order to be able to compare the timbral values. The system continually calibrates the statistical values as new audio files are analyzed in order to adjust and improve the comparative scale. The dataset is then used by the classification algorithms we developed in `Python 3.5`.

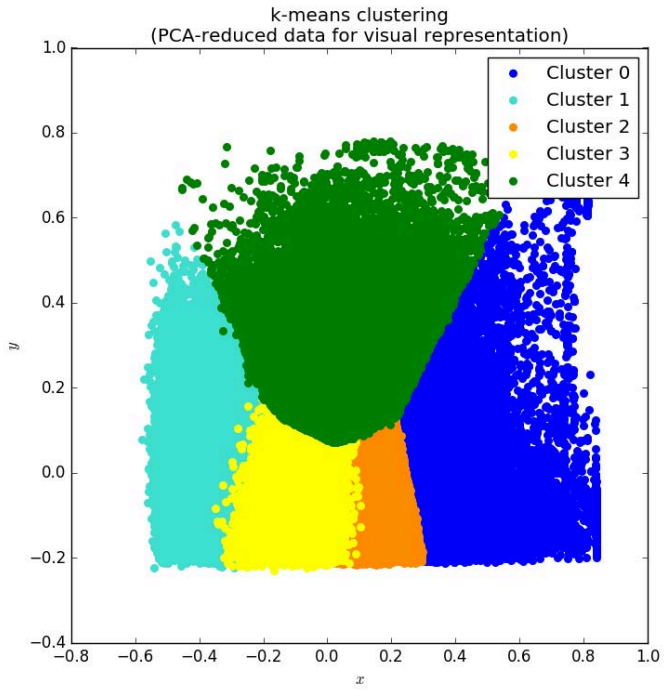


Figure 1. Graph showing the results for the *k*-means clustering performed on the 236 000 samples dataset.

2.2.2. K-Means Clustering

We first experimented with an unsupervised learning algorithm to identify a classification model. We wanted to divide the dataset into 5 parts—to represent the 5 timbral attributes. Therefore, we decided to perform a *k*-means clustering using Lloyd's algorithm [29] on our dataset.

We used the `KMeans` function from the `Scikit-Learn` package[3], with a `k-means++` initialization (which speeds up convergence [30]). Each entry of the dataset was input as a singular vector into the *k*-means algorithm. Results of the clustering is shown in Figure 1.

2.2.3. Support Vector Machine

Following the experiment with *k*-means clustering, we opted to implement a machine learning model usually utilized for data classification and regression analysis: Support Vector Machine (SVM) [31]. First, we needed to create a corpus training for the SVM algorithm. Therefore, for each timbral attribute we took and labeled the 250 highest index samples from the normalized dataset. Thus, the corpus training contained 1250 labeled samples.

[1] `http://www.mathworks.com/products/matlab/`
[2] MIRtoolbox is available at `https://goo.gl/d61EO0`
[3] `http://scikit-learn.org/stable/index.html`

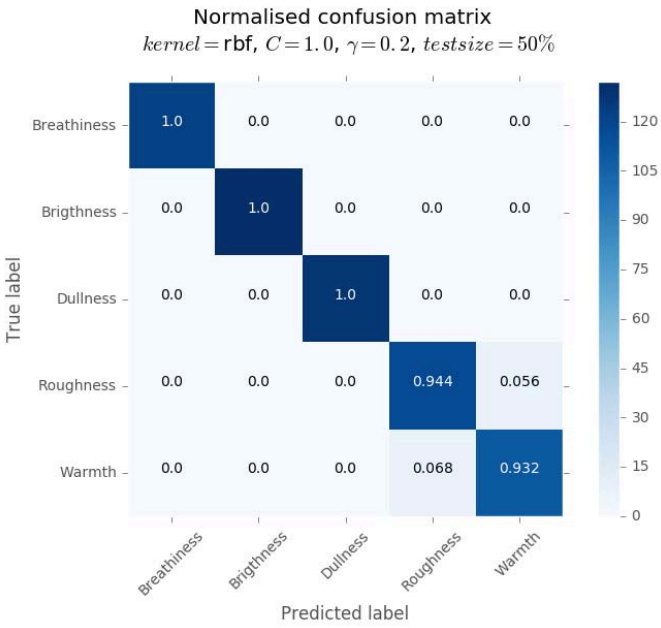


Figure 2. Representation of the normalized confusion matrix for the SVM algorithm, with test size = 50% (625 training samples, 625 testing samples).

We used the `svm.SVC` function from the `Scikit-Learn` package. After performing the `Scikit-Learn`'s functions for parameters' tuning, we selected a `rbf` kernel for our SVM algorithm, which performed a score of 0.976 (with 1.0 being the maximum). The resulted confusion matrix is shown in Figure 2.

3. RESULTS AND DISCUSSION

In [32], we reported on a pilot system that used the acoustic analysis described in section 2.1. We also reported on an experiment conducted with 20 participants to evaluate the accuracy of the analysis. Results indicated a correlation between the system's timbre analysis and human perception, which led us to develop the presented automatic classification system.

Timbral values being on different scales, the data gathering, and audio analysis helped us to evaluate the distribution of the values for each verbal descriptor, and therefore be able to normalize the data in the range (0.0–1.0), which then can be used in classification algorithms.

We performed a *k*-means algorithm on the 236 000 normalized samples in order to divide the dataset in 5 clusters, corresponding to the 5 timbral attributes. The result is shown in Figure 1, using a Principal Component Analysis (PCA) algorithm for visual representation. The *k*-means algorithm has been able to produce 5 distinct clusters, however due to the nature of the classification method the correspondent timbral attributes remains unknown until evaluated and labeled by the user.

To test the performances of the SVM algorithm, we divided the corpus training set into different batches of learning samples and testing samples. The SVM estimator with a `rbf` kernel performed a success score of 0.976. Figure 2 shows the normalized confusion matrix with a test size of 50%, in this case 625 samples for learning and 625 samples for testing. Although the SVM presents successful learning scores, it is dependent on a previously created corpus training. While this required action can be seen as a negative additional task, it could enable the user to create their own training data, which would represent their perceptual preferences.

4. CONCLUSION

In this paper, we have introduced the development of a computing system capable of automatically identifying the timbral qualities contained in orchestral audio samples. This system is built upon the existing research about the relation between acoustic features and verbal description of timbre qualities. It currently integrates five verbal descriptors: *breathiness*, *brightness*, *dullness*, *roughness* and *warmth*, introduced in section 2.1. Then, we detailed the creation of a comparative scale, based on audio recordings analysis. This scale enabled us to normalize data across the 5 timbral attributes, which is used for the classification algorithms presented in section 2.2.

Both *k*-means and SVM algorithms performed successful samples classification. However, a user action is still required, whether afterward for clusters labeling for *k*-means, or beforehand for corpus training's creation for the SVM. Nevertheless, these additional tasks can be used to calibrate the classification algorithms to the user's own musical perception, which could offer a solution to the challenging variation in music perception between individuals. Such developments could enrich computer-aided orchestration systems by harnessing perceptual aspects of polyphonic timbre within orchestral sound.

REFERENCES

[1] C. L. Krumhansl, "Why is musical timbre so hard to understand," *Structure and perception of electroacoustic sound and music*, vol. 9, pp. 43–53, 1989.

[2] D. Smalley, "Defining timbre—refining timbre," *Contemporary Music Review*, vol. 10, no. 2, pp. 35–48, 1994.

[3] American National Standards Institute, *Psychoacoustic terminology S3:20*. New York, NY: American National Standards Institute, 1973.

[4] R. Pratt and P. Doak, "A subjective rating scale for timbre," *Journal of Sound and Vibration*, vol. 45, no. 3, pp. 317–328, 1976.

[5] J. R. Platt and R. J. Racine, "Effect of frequency, timbre, experience, and feedback on musical tuning skills," *Perception & Psychophysics*, vol. 38, no. 6, pp. 543–553, 1985.

[6] R. D. Melara and L. E. Marks, "Interaction among auditory dimensions: Timbre, pitch, and loudness," *Percep-*

tion & psychophysics, vol. 48, no. 2, pp. 169–178, 1990.

[7] V. C. Caruso and E. Balaban, "Pitch and timbre interfere when both are parametrically varied," *PloS one*, vol. 9, no. 1, p. e87065, 2014.

[8] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and models*. Springer Science & Business Media, 2007, vol. 22.

[9] J. F. Schouten, "The perception of timbre," in *Reports of the 6th International Congress on Acoustics*, vol. 76, 1968, p. 10.

[10] J. M. Grey, "An exploration of musical timbre," Ph.D. dissertation, Stanford University, 1975.

[11] J. M. Grey and J. W. Gordon, "Perceptual effects of spectral modifications on musical timbres," *The Journal of the Acoustical Society of America*, vol. 63, no. 5, pp. 1493–1500, 1978.

[12] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff, "Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes," *Psychological research*, vol. 58, no. 3, pp. 177–192, 1995.

[13] R. A. Kendall, E. C. Carterette, and J. M. Hajda, "Perceptual and acoustical features of natural and synthetic orchestral instrument tones," *Music Perception: An Interdisciplinary Journal*, vol. 16, no. 3, pp. 327–363, 1999.

[14] V. Alluri and P. Toiviainen, "Exploring perceptual and acoustical correlates of polyphonic timbre," *Music Perception: An Interdisciplinary Journal*, vol. 27, no. 3, pp. 223–242, 2010.

[15] S. McAdams, *The Psychology of Music*, 3rd ed. Elsevier, 2013, ch. Musical Timbre Perception, pp. 35–67.

[16] G. von Bismarck, "Timbre of steady sounds: A factorial investigation of its verbal attributes," *Acta Acustica united with Acustica*, vol. 30, no. 3, pp. 146–159, 1974.

[17] K. Siedenburg, I. Fujinaga, and S. McAdams, "A comparison of approaches to timbre descriptors in music information retrieval and music psychology," *Journal of New Music Research*, vol. 45, no. 1, pp. 27–41, 2016.

[18] A. Tsiros and G. Leplâtre, "Animorph: animation driven audio mosaicing," in *Proceedings of the Conference on Electronic Visualisation and the Arts*. British Computer Society, 2015, pp. 115–116.

[19] S. Soraghan, A. Renaud, and B. Supper, "A perceptually motivated visualisation paradigm for musical timbre," in *Proceedings of the conference on Electronic Visualisation and the Arts*. British Computer Society, 2016, pp. 53–60.

[20] J. Hillenbrand and R. A. Houde, "Acoustic correlates of breathy vocal qualitydysphonic voices and continuous speech," *Journal of Speech, Language, and Hearing Research*, vol. 39, no. 2, pp. 311–321, 1996.

[21] M. A. Landera and R. Shrivastav, "Effects of spectral slope on perceived breathiness in vowels," *The Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 3339–3340, 2006.

[22] E. Schubert, J. Wolfe, and A. Tarnopolsky, "Spectral centroid and timbre in complex, multiple instrumental textures," in *Proceedings of the international conference on music perception and cognition, North Western University, Illinois*, 2004, pp. 112–116.

[23] A. C. Disley, D. M. Howard, and A. D. Hunt, "Timbral description of musical instruments," in *International Conference on Music Perception and Cognition*, 2006, pp. 61–68.

[24] D. B. Bolger, "Computational models of musical timbre and the analysis of its structure in melody," Ph.D. dissertation, University of Limerick, 2004.

[25] E. Terhardt, "On the perception of periodic sound fluctuations (roughness)," *Acta Acustica united with Acustica*, vol. 30, no. 4, pp. 201–213, 1974.

[26] v. W. Aures, "A procedure for calculating auditory roughness," *Acustica*, vol. 58, no. 5, pp. 268–281, 1985.

[27] R. Ethington and B. Punch, "Seawave: A system for musical timbre description," *Computer Music Journal*, pp. 30–39, 1994.

[28] O. Lartillot, P. Toiviainen, and T. Eerola, "A matlab toolbox for music information retrieval," in *Data analysis, machine learning and applications*. Springer, 2008, pp. 261–268.

[29] S. Lloyd, "Least squares quantization in pcm," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[30] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.

[31] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[32] A. Antoine, D. Williams, and E. Miranda, "Towards a timbral classification system for musical excerpts," in *Proceedings of the 2nd AES Workshop on Intelligent Music Production*, London, 2016.

Index of Authors