BEYOND KEY VELOCITY: CONTINUOUS SENSING FOR EXPRESSIVE CONTROL ON THE HAMMOND ORGAN AND DIGITAL KEYBOARDS

GIULIO MORO

BEYOND KEY VELOCITY: CONTINUOUS SENSING FOR EXPRESSIVE CONTROL ON THE HAMMOND ORGAN AND DIGITAL KEYBOARDS

GIULIO MORO

Submitted in partial fulfilment of the requirements of the
University of London Degree of Doctor of Philosophy
School of Electronic Engineering and Computer Science
Queen Mary University of London

SUPERVISORS:
Prof Mark B. Sandler
Dr Andrew P. McPherson

May 2020

Giulio Moro: *Beyond key velocity: continuous sensing for expressive control on the Hammond organ and digital keyboards* © May 2020

Unlike an empty oil barrel, you can bang on a synthesizer any number of ways, the resulting sound will always be within the realm of its specified, "programmed" behavior, intended or not. The oil barrel delivers an infinite number of variations for each "whack".

—- Zack Settel and Cort Lippe, 2003

ABSTRACT

In this thesis we seek to explore the potential for continuous key position to be used as an expressive control in keyboard musical instruments, and how pre-existing skills can be adapted to leverage this additional control. Interaction between performer and sound generation on a keyboard instrument is often restricted to a number of discrete events on the keys themselves (notes onsets and offsets), while complementary continuous control is provided via additional interfaces, such as pedals, modulation wheels and knobs. The rich vocabulary of gestures that skilled performers can achieve on the keyboard is therefore often simplified to a single, discrete velocity measurement. A limited number of acoustical and electromechanical keyboard instruments do, however, present affordances of continuous key control, so that the role of the key is not limited to delivering discrete events, but its instantaneous position is, to a certain extent, an element of expressive control. Recent evolutions in sensing technologies allow to leverage continuous key position as an expressive element in the sound generation of digital keyboard musical instruments.

We start by exploring the expression available on the keys of the Hammond organ, where nine contacts are closed at different points of the key throw for each key onset and we find that the velocity and the percussiveness of the touch affect the way the contacts close and bounce, producing audible differences in the onset transient of each note.

We develop an embedded hardware and software environment for low-latency sound generation controlled by continuous key position, which we use to create two digital keyboard instruments. The first of these emulates the sound of a Hammond and can be controlled with continuous key position, so that it allows for arbitrary mapping between the key position and the nine virtual contacts of the digital sound generator. A study with 10 musicians shows that, when exploring the instrument on their own, the players can appreciate the differences between different settings and tend to develop a personal preference for one of them. In the second instrument, continuous key position is the fundamental means of expression: percussiveness, key position and multi-key gestures control the parameters of a physical model of a flute. In a study with 6 professional musicians playing this instrument we gather insights on the adaptation process, the limitations of the interface and the transferability of traditional keyboard playing techniques.

ACKNOWLEDGEMENTS

This piece of work came together over an extended period of time, and as such it has seen the direct or indirect contribution of several people, several of whom are listed below.

My supervisors, Andrew and Mark, who have allowed me to follow my research interests with a lot of freedom throughout my PhDs, steering me back in the right direction whenever I was about to derail off the track.

With great supervisors come great research groups. Here's to the Augmented Instruments Laboratory, who has been a true source of inspiration and fun. I am particularly fond of the lab "Mk I", as I found it when I joined: Duncan, Laurel, Chris, Astrid, Liam, Robbie and Andrew.

The Bela project and the Bela team for the passion and the exciting times, the Bela users, and everyone who asks me questions, for being a continuous source of motivation and challenges.

Some of my former colleagues at C4DM that have helped me at various point in the process, providing advice on specific parts of this project: Bob Sturm, Adib Mehrabi, Brecht De Man, Fabio Morreale.

Victor, who took a whole evening to flash an SD card. Chris H., who handed me the stone with which I got two birds.

Matteo for words and illumination (and ups and downs). Keunwoo for being Unique. Delia and Gijs *por cuidarme tanto*. Floriano for reciprocally taking care of each other. Adan and Sophie for being *amici* and proofreaders. Marco for being above everything. Manos for good reasons.

The ***vulf*** for bringing music back to life and distracting me while listening and watching for too many low volume hours.

The songs that I could listen to on endless repeat in order to concentrate while writing this document: King Crimson's Islands (40$^{\text{th}}$ Anniversary Series version), The Velvet Underground's Pale Blue Eyes, Fleetwood Mac's Albatross (in the blend by Keith Fullerton Whitman).

Mi querida Juliana for all the love in the world.

Last, thanks to my parents and RyanAir and everyone who fed me in these long years, without whom I would not be here today.

Special thanks to Tiny and Ron at the Hammond Organ Restoration Company (Coventry, UK) for the advices on the Hammond and the great stories about vintage keyboards.

This work was funded by EPSRC grants EP/L019981/1("Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption") and the Queen Mary, University of London, Postgraduate Research Fund.

DECLARATION

I, Giulio Moro, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Date: May 12, 2020

Giulio Moro

PUBLICATIONS

Some ideas and figures have previously appeared in the following publications:

1. Moro, Giulio, Andrew P. McPherson, and Mark B. Sandler. 2017. "Dynamic temporal behaviour of the keyboard action on the Hammond organ and its perceptual significance" *The Journal of the Acoustical Society of America* 142 (5): 2808–2822.
2. Moro, Giulio and Andrew P. McPherson. 2020 "A platform for low-latency continuous keyboard sensing and sound generation" *Proceedings of the International Conference on New Interfaces for Musical Expression*, Royal Birmingham Conservatoire, United Kingdom.
3. Moro, Giulio, Astrid Bin, Robert H. Jack, Christian Heinrichs, and Andrew P. McPherson. 2016. "Making High-Performance Embedded Instruments with Bela and Pure Data" *Proceedings of the International Conference on Live Interfaces*. Brighton, United Kingdom.
4. McPherson, Andrew P., Robert H. Jack, and Giulio Moro. 2016. "Action-Sound Latency: Are Our Tools Fast Enough?" *Proceedings of the International Conference on New Interfaces for Musical Expression*. Brisbane, Australia.
5. Morreale, Fabio, Giulio Moro, Alan Chamberlain, Steve Benford, and Andrew P. McPherson. 2017. "Building a maker community around an open hardware platform" *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6948–6959.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ADC Analog to Digital Converter

CPU Central Processing Unit

DAC Digital to Analog Converter

DMA Direct Memory Access

DMI Digital Musical Instrument

FHV Final Hammer Velocity

GPIO General Purpose Input/Output

I2S Inter-IC Sound

IR Infrared

IRQ Interupt ReQuest

MRP Magnetic resonator piano

MPE MIDI Polyphonic Expression

MIDI Musical Instrument Digital Interface

OSC Open Sound Control

PRU Programmable Real-Time Unit

RAM Random Access Memory

RISC Reduced Instruction Set Computer

RTDM Real-Time Driver Model

SAR Successive Approximation Register

SoC System on Chip

SPI Serial Peripheral Interface

1 INTRODUCTION

In this research we explore the role of touch on a keyboard instrument and the potential for expanding its potential in the design of digital musical instruments. For many years it has widely been accepted that the scalar parameter of onset velocity is enough to characterise the qualities of a note for the purposes of synthesising or analysing a performance [Ortmann, 1925, Moore, 1988]. This is, however, not entirely true for many acoustic and electro-mechanical keyboard instruments which offer a more or less subtle position-dependent control on the timbral and temporal characteristics of a note, and the literature shows that such assumption does not always hold true even in the case of the piano [Goebl et al., 2014].

Continuous key control, the sensing of the key position throughout the key throw as opposed to a discrete velocity event, acts as a common thread across our research. We start by studying the case of the Hammond organ, for which the continuous position of the key is of primary importance in generating subtle differences during a key onset. This motivates us to investigate further how the subtle changes in the control of the key onset are perceived by the performers of this instrument, and how such control can be magnified in a newly designed instrument.

1.1 MOTIVATION

1.1.1 *The instrument as an action/sound transducer*

A musical instrument can be seen as a type of transducer which converts a performer's body movements into sound [Baily, 2008, p. 123]. For some instrument classes, this allows an intimate, continuous connection to the sound, as in the case of many wind, strings and percussion instruments, where the interaction between the performer and the instrument is very direct. There is often a direct physical contact between parts of the body of the performer and the part of the instrument devoted to producing the sound, allowing for a multitude of multidimensional interactions.

By contrast, most keyboard instruments present a very sophisticated and convoluted mediation between the interface exposed to the user (the keyboard) and the sound generation. On an acoustic piano, a complex and finely-tuned system comprising over 20 parts transforms the downward motion of the finger on the key into the motion of the hammer, which eventually decouples from the keys and hits the string [Askenfelt and Jansson, 1990]. As a first approximation, the speed of the finger as it presses the key is translated into the speed of the hammer which in turn affects the amplitude of the resulting sound and its timbre. Electromechanical pianos (such as the Wurlitzer and the Rhodes) have a simplified piano action which works on similar principles, however the resonating elements hit by the hammer are reeds and bars, respectively. The harpsichord's plucking mechanism is such that the loudness of the note is largely independent of the force used in its actuation [MacRitchie and Nuti, 2015]. In pipe organs with electronic consoles, the performer closes a single electric switch at some point during its motion. The electric signal then travels to the pipes, possibly remotely located, where an electro-pneumatic valve opens and closes the flow of air through the pipe [Le Caine, 1955]. Sound generators that are entirely electronic or digital often take a measure of the key velocity on the keyboard and use that as the only note-specific control to drive some parameter in the sound generator.

Other keyboard instruments offer a more direct connection to the sound, maintaining some sort of uninterrupted connection between the performer's finger and the sound generator [McPherson, 2015]. In the clavichord, which dates back to the XV century, the *"tangent"*, a tab attached to the back of the key, would hit the string at the onset and be held against it for all the duration of the note, effectively acting as one of the terminations of the string, thus allowing a gentle vibrato effect by altering the pressure. The Hohner Clavinet, an electromechanical keyboard produced in the second half of the twentieth century, would similarly use a *"tangent"* at the bottom of the key to hold the string against an *"anvil"*, acting as a termination for the string, thus allowing for a subtle vibrato along with some percussive staccato effects. In pipe organs with tracker action, the valves controlling the air flow to the pipe are directly controlled by pressing on the key, allowing for some subtle control on the attack and release stages of the note [Le Caine, 1955]. The electromechanical Hammond organ, produced in large numbers for 40 years in the twentieth century, has a stack of multiple contacts under the key, each controlling a different harmonic component of the note, which close at different points in the key-throw.

The Ondioline is an electronic synthesizer invented in 1941 by Georges Jenny [Fourier et al., 1994]. It stands out as perhaps the only keyboard instrument to have been produced in relatively large numbers[1], for which continuous key position was one of the main expressive features. This, combined with side-by-side keyboard vibrato allowed for a remarkably expressive instrument, even by today's standards: multi-instrumentalist Wally de Backer in 2018 described it as *"[an incredibly] expressive and timbrally versatile electronic instrument [...] the unique mechanics for playing it allows you to create sounds very sensitively and with a musical deftness I just feel isn't present on most other electronic instruments from the '40s – or decades since"*[2].

1.1.2 *Key velocity to encode a gesture*

The preparation gesture for executing a key press on the piano requires the coordination of movements of the upper body, arm, wrist and finger joints [MacRitchie, 2015]. Such a complex gesture is, however, transmitted to the instrument only through the fingertip, exerting a simple downward force on the key surface. The resulting downward motion of the key is, by comparison, a small part of the whole gesture, but only what happens during the few millimeters of the key-throw ultimately determines the velocity of the hammer as it hits the string. Given how it has been known for a long time that the hammer velocity is the only parameter to affect the intensity of the produced sound, but the ways in which a given hammer velocity value is achieved may vary widely [Ortmann, 1925], we may wonder whether the complex gesture that leads to the key press has any influence at all on the produced sound. A pianist can produce almost identical final hammer velocities by distributing the acceleration differently during the key press [Askenfelt and Jansson, 1990]. This ultimately encompasses the use of struck (percussive) and pressed (legato) touches, and all the fine variations between the two, the preparatory gesture, the use of the weight of the arm and the body posture, all through to the weight and compressibility of the intricate mechanisms of the piano action, to end with the hammer escaping the jack before reaching its final impact velocity and hitting the string [Askenfelt and Jansson, 1990, 1991, 1993]. However, some residuals of the complex gesture that accompany the key press manage to make their way into the generated sound [Goebl et al., 2014]. The "accessory noise" of

1 *"Over a thousand instruments were sold in the USA alone"* [Fourier et al., 1994]
2 "Gotye Makes His Return with an Ondioline. Sorry, a What?": `https://www.broadsheet.com.au/sydney/entertainment/article/gotye-makes-his-return-ondioline-what`. All online resources in this thesis have been last accessed on 12/4/2020.

the finger hitting the key during a struck touch, or that of the key impacting the keybed for a *forte* key press, can be perceived by the listener as part of the onset sound, somehow giving a sonic result, however quiet, to the more intricate set of actions that lead to the production of the note. We can therefore conclude that the gesture of the performer does have a (small) effect on the sound.

When controlling a digital synthesizer there is a need to translate the performer's physical gesture into a digital control that can be sent to the sound generator. With the advent of digital synthesizers, in the early 1980s, the Musical Instrument Digital Interface protocol (MIDI) was introduced, primarily aimed at the interconnection of keyboards and sound generators by different manufacturers [IMA, 1983]. The limited capabilities of digital devices at the time set some of the requirements of the standard, including low bandwidth and low resolution [Moore, 1988]. It was at that time that the already widely adopted practice of using the "velocity" of the key during the onset as the only parameter to transmit from the controller to the generator for a note onset, was codified. MIDI quickly became a de-facto industrial standard and it is still omnipresent today in the musical instruments industry.

1.1.3 *The bottleneck in the keyboard*

In their research on digital musical instruments (DMI), Jack et al. [2017] introduce the notion of *bottleneck*, to indicate the way a DMI *"constrains the gestural language of the performer, and encourages them to play in certain ways"*. The complex gestural language of the body of the performer is reduced in dimensionality and bandwidth to fit through the affordances provided by the interface, projected down through the bottleneck and then expanded out again into the parameters that control the sound generation. As a case study, they describe how the richness of the gestural language present at the initial approach with a percussion DMI is progressively reduced while exploring the instrument, as a consequence of its bottleneck. Within a short practice session, players adjusted their behaviour to exploit the possibilities enabled by the instrument, while disregarding those gestures or means of interaction that did not produce a useful musical result. Those actions that have no effect on the produced sound, because they did not pass through the bottleneck, were discarded (*push* effect), while the performer concentrated on those that effectively make them feel in control of the instrument and of the generated sound (*pull* effect).[3]

The presence of a bottleneck is inherent in the DMI design process: the physical world has to be simplified, constrained and digitised, before it can be interfaced to the sound generator. The propagation of the data from the physical domain through to the sound generator is a transmission of a selected set of data with limited bandwidth: in no way the continuous multi-dimensionality of the physical world can be constrained through that bottleneck. Any data not actively selected for "digitisation" will be lost in the process. The design of this bottleneck, in terms of which data should flow through it, offers the instrument designer room for exploration.

We find that Settel and Lippe [2003]'s words are well suited, if a bit extreme, to denote the limitations of the bottleneck in DMIs: *"Unlike an empty oil barrel, you can bang on a synthesizer any number of ways, the resulting sound will always be within the realm of its specified, 'programmed' behavior, intended or not. The oil barrel delivers an infinite number of variations for each 'whack' "*. The keyboard in a non-electronic instrument is already a very limited interface between the performer's gesture and the sonic outcome. However, there is a component of physical interaction that can take place beyond the designated interface, such as directly plucking the strings of a piano, or placing objects on its vibrating strings. Even when interacting solely with the keyboard, in some instruments there are some ways for gestures to transcend the constrains of the designated interface and produce meaningful sonic results, through the exploitation of the instrument's physicality. When a keyboard DMI is designed to let only the information relative to the note pitch and velocity go through its bottleneck, all those more or less subtle forms of control available on acoustic and electromechanical keyboard instruments which are not velocity-based disappear.

The implications of the push and pull effects on the impoverishment of the gestures used, found by Jack et al. [2017], therefore pose a concern on the future evolution of musical performance as a whole: if some techniques no longer affect the sonic output, because they are left out of the bottleneck, they will eventually – consciously or unconsciously – be discarded, potentially leading to a general impoverishment of the gestural vocabulary of performers. It would not be surprising to discover that some gestures are discarded when playing digital keyboards because the instrument "does not support" them.

3 Tuuri et al. [2017] use the term *push effects* when the user of an interface feels forced, guided or disempowered, while *pull effects* indicates that the user feels in control of the interface.

Investigating expression in digital musical instruments, Dobrian and Koppelman [2006] suggest that *"The most basic need for a controller is that it accurately capture the data provided to it by the human interface"*. While analysing the biomechanics of piano gestures, and building on Bernstein [1967]'s seminal work, MacRitchie [2015] claims that *"to produce a keypress [...] there are infinite possibilities for the movements needed to produce this action"*. Goebl et al. [2014] show evidence that the gesture choice may have an impact on the produced sound, which suggests that the type of gesture used is therefore worth capturing. Yet, it is the norm for keyboard DMIs to only sense timing and velocity of key presses, while there is a whole gestural richness in the performer's action that remains out of the picture. The "MIDI" keyboard is therefore not meeting what Dobrian and Koppelman [2006] call the "basic need", not even with regard to the relatively basic characteristics of the piano. Its inappropriateness is even more obvious when considering that for many other keyboard instruments, such as the Hammond organ, some pipe organs, or the Ondioline, the velocity and timing paradigm is plainly inadequate. A simple example to demonstrate this is a recent demo video[4] of a software emulation of the Ondioline. In the video, the digitally generated sound and the sound of the original instrument are compared side by side, and there is a remarkable resemblance between the two, as a testimony of the high quality of the software. However, when it comes to actually performing with the digital emulation, the demonstrator is using a common MIDI keyboard. To make up for the lack of continuous sensing on the MIDI keyboard, the player has to rely on two ribbon controllers. The gestural control of the original instrument and the intuitive immediacy of the interaction with the key is therefore completely lost, and the result is at best mediocre.

We believe it is up to the keyboard designer to ensure that subtle nuances in a gesture can translate in changes of the sonic outcome, in order to keep high the value of gestural richness, and the potential for expression. A redesign of the keyboard bottleneck to allow more complex gestures to be captured and represented accurately would surely help achieving this, especially when coupled with sound generators that are able to make use of it.

1.1.4 *Implications*

The work in this thesis is inspired by the possibility of expanding the bottleneck represented by a MIDI-like keyboard interface, allowing for a richer set of ges-

4 https://www.youtube.com/watch?v=wCfYMu8VZxE

tures when interacting with digital sound generators through a keyboard. We build our work around a piece of technology, the keyboard scanner [McPherson, 2013], which, instead of providing velocity readings, allows to sense the position of each key continuously. On the one hand, we expect this to provide a better control interface for software emulations of existing keyboard instruments. On the other hand, this opens up a wide range of expressive possibilities for entirely new keyboard-based DMIs.

We therefore start our research by studying the behaviour of one keyboard instrument for which we think continuous key position is a more adequate control than discrete key velocity, namely the Hammond organ. Next, we build a digital Hammond emulator which can be controlled by continuous key position. Last, we create a physical modelling synthesizer controlled by continuous key position. In this thesis we describe the analysis of the Hammond, the technologies we developed, the design of the instruments and the performance studies we ran with them.

1.2 RESEARCH QUESTIONS

We aim to bring continuous key sensing to keyboard-based digital musical instruments, expecting that this will increase the intimacy between a performer's gesture and the sound generation, breaking away from the long-standing paradigm of discrete note onsets characterized only by their velocity. Our overarching research question can therefore be formulated as:

- **What are the implications of continuous key sensing in the context of keyboard-based digital musical instruments?**

More specifically, we are interested in two classes of problems:

- **How does the subtle manipulation of key-sound mapping affect the performer's perspective on a keyboard DMI?** Some keyboard instruments offer a subtle control over the sound of the onset through the temporal evolution of the key position, in a way that is more complex than a single velocity measurement. In many cases, such control is a by-product of the sound generation process, and is not necessarily intended by the designer. A velocity-based keyboard will have some objective shortcomings when paired with a digital sound generator emulating these instruments, in that it is inadequate to represent the full gesture. However, given how the effect on the original instrument was initially small, the entity of the impact of said shortcomings is unknown. By using continuous key sensing and an adequate mapping strategy, we can assess the impact of the limitations of traditional controllers on the performer.

- **How can we leverage continuous key sensing to increase gestural richness and control capabilities in keyboard DMIs?** The control possibilities allowed by continuous key sensing have the potential of adding a new dimension to keyboard playing. Continuous key sensing was an integral part of the character of the Ondioline, a synthesizer from the 1940s, whose expressiveness is, to this day, remarkable. With the advent of digital technology the possible applications of continuous key sensing, through complex mappings between gesture and sound, can be expanded even further. We want to investigate the potential of such applications, and the extent to which a keyboard player can acquire the new skills required to master the new capabilities of the instrument, and how these relate to their pre-existing training.

1.3 thesis outline

The remainder of this thesis is structured as follows:

Chapter 2 background We start with a review of the literature of studies of touch and continuous control on keyboard instruments, also surveying sensing technologies and their limitations.

Chapter 3 dynamic behaviour of the hammond We perform a detailed study of the behaviour of the electronic key contacts on the Hammond organ, showing how the continuous temporal evolution of the key position during the note onset affects the transient of the produced sound (key-click), allowing subtle expressive control at each key press. A listening test shows that these minute differences can be perceived by the listener in an isolated context.

Chapter 4 position-based control of a hammond emulation We describe the technology we created to realize an embedded platform for sound generation based on continuous key angle sensing. Using this platform, we develop a Hammond organ emulator that can be controlled with continuous key angle and we run a performance study to investigate how performers react to different conditions of the mapping between key position and sound.

Chapter 5 controlling a physical model with key position We design a keyboard-based DMI using continuous key position, where the basic assumptions of discreteness of the note onset, and the sonic meaning of touch are challenged. In a performance study, we evaluate how traditional keyboard skills generalize to this instrument, and the potential for continuous key control in keyboard DMI.

CHAPTER 6 CONCLUSION We reflect back on the outcomes of our work and discuss the implications of continuous key sensing in the design of keyboards instruments.

2 BACKGROUND

This chapter contains a review of the literature which constitutes the general background for our research.

We discuss in Section 2.1 the effect of touch on the piano, which is a well-studied problem, with a particular focus on the definition and measurements of key percussiveness. Section 2.2 presents a brief historical excursus of continuous control on keyboard instruments. In Section 2.3 we discuss sensing strategies to detect gestures on the keyboard. Finally, in Section 2.4 we reflect on the limitations of MIDI in the context of continuous keyboard sensing, and in Section 2.5 we anticipate the potential implications of continuous sensing on the keyboard.

2.1 TOUCH ON THE PIANO

When discussing the effect of piano touch, most of the literature focuses on the distinction between:

- pressed (also called legato, non-percussive) touch, when the finger is resting on the surface of the key before pressing it.
- struck (also called staccato, percussive) touch, when the finger is moving as it engages the surface of the key.

An early study on the effect of touch showed that the psychological factors involved in different types of key press are different, but concluded that ultimately there is a one-to-one correspondence between the intensity of the touch and the tonal quality of the produced sound [Ortmann, 1925]. This correspondence is strictly true only for the sound produced by the string, but a key press is often accompanied by additional sound components. During a struck onset, for instance, an "early noise" is produced by the finger hitting the key. A listening test showed that when the early noise, which precedes the actual note onset, is excluded from a recording, a listener cannot infer the type (pressed or struck) of touch used to produce the tone [Goebl et al., 2004]. In real playing conditions, this noise will be part of the sound of the acoustic instrument, and will to some extent reach the listener. However, being much quieter than the tone produced by the string, and very close to it in time, it may be hard to distinguish it as a distinct event [Kinoshita et al., 2007]. In Goebl et al. [2014], the early noise, along with the more subtle "key-bottom" noise, caused by the key hitting the felt on the keybed, are shown to be perceptually relevant to the listener in a controlled experimental situation. Detailed simulations of the mechanics of flexible hammer shanks have shown that touch has the potential to affect the produced sound by inducing micro-oscillations on the hammer, in turn affecting the strike point and the hammer-string interaction forces [Vyasarayani et al., 2009]. Chabassier and Duruflé [2014]'s simulation shows the effects of the interaction forces on the spectrum of the produced sound, with struck touches generating "brighter" tones than their pressed counterpart for the same final hammer velocity.

The acceleration of the key during a key press is continuously under the control of the player. Even non-professional players can vary the way they distribute the acceleration, in order to control parameters other than simple velocity, such as percussiveness, weight and depth of a key press [McPherson and Kim, 2011]. Trained pianists, on the other hand, routinely semi-unconsciously control these dimensions as part of an expressive performance [McPherson and Kim, 2013]. These studies suggest that reducing gesture on the piano to a single scalar parameter, such as the velocity of the key or of the hammer, does not fully represent the expressive intention of the performer.

Goebl et al. [2005] conducted a thorough study on the relative timings of different sections of piano action as they relate to the type of touch used. The time between the finger-key actuation and the production of the sound changes greatly with the velocity of the touch, with pressed notes exhibiting a longer delay than struck notes with a similar hammer velocity, meaning that the action-to-sound latency is dependent on the dynamic used. Birkett [2014] suggests that the key-to-hammer-to-string interaction is consistently repeatable, in that a given key motion will consistently produce very similar sonic results.

When testing the response of an instrument to the performer's key press, there is need to find a strategy to create a dataset large enough to be statistically meaningful and to reproduce the exact gesture multiple times. Typical self-playing pianos use solenoids to press the tail of the key, and their performance has been evaluated in Goebl and Bresin [2003]. Hayashi et al. [1999] designed an active, voice-coil motor driven, mechanical finger which presses the front of the key, but is unsuitable for percussive touch. Askenfelt and Jansson [1990] used a rubber-tipped pendulum which occasionally causes undesired rebounds of the tip on the key. These methods tend to be expensive, cumbersome, or have drawbacks, so that the most widely used method is often to have a human player press the key repeated times, and then using measurements of continuous key position, final hammer velocity, produced loudness, or a combination thereof, in order to group similar presses together.

For an extensive review of the studies on piano touch, we refer the reader to MacRitchie [2015]. Touch has also been studied on other keyboard instrument, such as the Ondes Martenot [Quartier et al., 2015] and the harpsichord [Gingras et al., 2009, MacRitchie and Nuti, 2015].

2.1.1 *Percussiveness*

A percussive, or struck, key press occurs when the finger is already moving downwards before hitting the key. Ortmann [1925] observes that in the case of a non-percussive motion, the acceleration of the key is constant, while for a percussive press, the key motion shows an initial velocity jerk, due to the finger sending off the key during the initial impact, and later catching up with it.

McPherson and Kim [2011] formalize Ortmann's observations by modelling the finger-key-hammer action system as two masses (the finger and the hammer action) and a massless lever of the first type (the key). In such system, when the finger applies a constant force on the key, starting from the resting position, it will result in a linear increase in velocity. In a percussive touch, the elastic collision between the finger (with a non-zero initial velocity) and the key will cause an initial spike in the key velocity, proportional to the momentum of the finger. If the finger continues in its downward motion, it will eventually catch up with the key, and the velocity will henceforth increase linearly.

Ortmann [1925] further observes that, following the initial velocity jerk, the key will not be under the direct control of the finger for a certain portion of its downward motion and that the point in the key-throw at which the finger re-engages the key is dependent on the dynamic of the press. This will happen earlier in the key-throw for quieter dynamics, and later for louder dynamics. Since the finger spends less time overall in contact with the key in a percussive touch, Ortmann argues that key control is harder for percussive touches.

The classification of percussive and non-percussive touches is not always clearly defined. For instance, a key press where the velocity of the finger at the impact is slow will have characteristics very similar to that of a non-percussive touch, which leads to considering percussiveness as a continuous quantity, where different *levels* of percussiveness can be achieved. McPherson and Kim [2011] quantify the percussiveness of a key onset by the magnitude of the velocity spike: the absence of a spike indicates a non-percussive touch, and a larger spike denotes a more percussive touch. Bernays and Traube [2012] use two descriptors to describe percussiveness: the ratio of the key depression at half the attack duration to the maximum key depression, and the average of the key depression curve.

2.2 CONTINUOUS CONTROL ON KEYBOARD INSTRUMENTS

The capability of shaping the tone of a note on a keyboard instrument is typically confined to the instants of its onset and release. On most instruments, at the onset the player can affect the loudness, and/or the timbre, of the sound according to the velocity of the key. A complete review of keyboard instruments which provide continuous control on their keys can be found in McPherson [2015].

Few acoustic keyboard instruments allow continuous shaping of notes throughout their duration. The clavichord allows to slightly change the pitch of the note throughout its duration, as the "tangent" of the key is itself resting on the string, acting as a bridge, and so varying pressure during the duration of the note allows to achieve a vibrato effect [Kirkpatrick, 1981]. On tracker pipe organs, the key opens the valve that controls the airflow into the pipe, thus allowing to continuously control the emission to a certain extent. Some pipe organs with electric action, which therefore do not have this control, tried to mimic this effect by having multiple pipes opening at different points in the key throw [Le Caine, 1955]. In the acoustic piano, the release of a key can be used to continuously control the return of the felt dampers, thus allowing the change the sound of the release transient, or effectively moving from a "release instant" to a more prolonged "release gesture".

As for electromechanical instruments, the Hohner Clavinet, introduced in the 1960s, was based on a concept similar to that of the clavichord, with the tangent striking directly the string, but in this case the string is held against a bridge (the "anvil"). Some vibrato is achievable on the Clavinet, but to a much more limited extent than on the clavichord. The Hammond organ, an electromechanical organ introduced in 1934, has 9 contacts per key that close at slightly different points in the key throw, allowing some control during the onset timbre [Vail, 2002]. We study the effect of touch on the Hammond more in detail in Chapter 3.

The first electronic keyboard instrument to be manufactured and distributed in large numbers was the Ondes Martenot, a monophonic synthesizer introduced in 1928 [Laurendeau, 1990]. The first version of the instrument did not include a keyboard, but simply a resistive wire to control pitch, a number of switches for timbre and an isolated "intensity key", shaped in the form of a piano key, placed on the left hand side of the instrument, for controlling dynamics. When pressed, the intensity key compresses a bag of carbon powder, changing its electrical resistance and in turn affects the loudness of the instrument [Quartier et al., 2015]. Later versions of the Marthenot added an 83-key keyboard which, perhaps drawing inspiration from the earlier wire interface, allowed to perform vibrato on the keys by oscillating the key side-by-side. The expressiveness of the Martenot earned it popularity among classical composers, with over 1500 works composed for the instrument, which still lives on today within a small but dedicated group of performers and estimators [Schampaert, 2018].

Another keyboard monophonic synthesizer, George Jenny's Ondioline, introduced in 1941, also featured remarkable expression control [Fourier et al., 1994]. Similarly to the Martenot, pitch vibrato is achievable by oscillating the key side-by-side. Additionally, the player is able to obtain dynamic control directly on the playing keys, as the amplitude of the produced sound is dependent on the position of the key, thus allowing expressive one-handed operation Paradiso [1997]. The Ondioline owes a lot of its fortune precisely to its expressive capabilities, making for credible recreations of acoustic instruments, as it can be inferred from some of the existing footage of Jean-Jacques Perrey demonstrating it[1]. While the sounds it produces by means of analog, tube-based subtractive synthesis are technically much more limited than those obtainable with modern synthesis techniques, the *performability* of the instrument, with its modulation and articulation capabilities, still makes it stand out today.

Hunt [1999, pp. 9-10] notes how the new, innovative interface of early electronic instruments such as the Theremin[2] and the first iteration of the Martenot, were freeing musicians from the constraint of discrete notes and allowed continuous control in a way that was no longer possible when, shortly thereafter, the keyboard became the interface of choice for electronic instruments. We can speculate that the reason of such a change could reside in instrument designers trying to

1 "Jean-Jacques Perrey on 'I've Got a Secret', 1960, Part Two - Ondioline Demonstration" `https://www.youtube.com/watch?v=05sAxt8zNZI`

2 The Theremin, introduced in 1920, is an electronic instrument that is played by moving the hands in the air in its proximity [Paradiso, 1997]

leverage the existing expertise (and market) of keyboard players worldwide. Even if the Martenot and the Ondioline, two of the earliest examples of keyboard-based electronic instruments featuring continuous control on the key itself, this feature mostly disappeared from later instruments.

In electronic synthesizers, continuous control can be available, but not in the keys themselves. Potentiometers or switches on the front panel of the instrument have often been used for shaping the evolution of sound. Dedicated performance-oriented modulation sources, the pitch bend wheel and the modulation wheel, first appeared on the Moog Minimoog in 1970, and have since been a staple of the interface of analog and digital keyboard synthesizers [Paradiso, 1997]. Other interfaces for continuous control have surfaced on the keyboard market over the years, such as ribbon controller of the Moog Micromoog, the breath controller of the Yamaha VL-1 [Jones, 2008, p. 23-24], or the capacitive and pressure sensing and haptic feedback of the Touché[3], as well as several examples of expression pedals. Some examples also come from the academic world, Wierenga [2012] augmented a keyboard with all sorts of gestural controllers, including some novel ones such as crank, ratchet, and even a Wii remote, while Yang and Essl [2014], Granieri and Dooley [2019] explored the use of free-hand gestures to control sound parameters.

Using the key itself as a modulation source has long been limited, as far as commercial devices go, to the use of aftertouch, obtained pressing into the keybed once the key has reached the end of its travel. This was featured on many monophonic synthesizers, as early as 1972 on the ARP Pro-Soloist[4]. Polyphonic aftertouch was famously available on Yamaha's flagship polyphonic synthesizers, the GX-1[5] and the CS-80[6]. Aftertouch and polyphonic aftertouch became a part of the MIDI standard [IMA, 1983], however the polyphonic version is rarely implemented to this day [McPherson, 2015]. Yamaha was also one of the few manufacturers to incorporate a sensor to detect horizontal oscillations of the keyboard assembly (similar to the Ondioline), which Yamaha called "touch vibrato", as early as 1975 on the SY-2[7], but this technology was not very widespread, and only reached the YC-20 series of combo organs and the GX-1.

Several prototypes of keyboards capable of continuously sensing the key position surfaced over the years, such as Le Caine's electrostatic keyboard [Le Caine, 1955], and the multiply touch keyboard [Moog, 1982, Eaton and Moog, 2005].

3 https://www.expressivee.com/
4 http://www.vintagesynth.com/arp/prosoloist.php
5 http://www.vintagesynth.com/yamaha/gx1.php
6 http://www.vintagesynth.com/yamaha/cs80.php
7 http://www.vintagesynth.com/yamaha/sy1.php

Moog's notebender allowed to bend a note up or down by sliding the key forwards or backwards [Moog, 1987]. McPherson [2013] describes a portable device for continuous key sensing on the piano keyboard. None of these devices, however, has become commercially available.

Some products recently reached the market which, while not providing sensing of continuous key position, allow to produce modulation while keeping the fingers on the keys. TouchKeys[8] is a capacitive-sensing overlay for the keyboard that tracks the position of fingers on the key surface McPherson [2012]. Enancia's Neova[9] is a ring that is worn by the performer and contains an accelerometer that can detect oscillating motion, sideways and front tilt during key presses, but also allows free-space gestures. Sonic Manipulation's Vibe Bar[10] is a bar that is placed in front of the keyboard and allows to perform pitch bending gestures with the palm of the hand.[11] Keith McMillen's K-BOARD PRO 4 [12] is a MIDI controller whose silicone keys are arranged in a conventional keyboard layout, but without a keyboard action. Keys respond to pressure as well as sensing position on the longitudinal and transversal axes.

2.3 SENSING TECHNOLOGIES FOR KEYBOARD INSTRUMENTS

The problem of sensing gestures at the keyboard is multi-faceted, and the ideal solution depends on the task at hand.

2.3.1 *Acoustic pianos*

Gesture acquisition on an acoustic piano can be necessary for didactic, augmented performances, performance analysis, performance reproduction or research purposes. In an augmented performance, extracting real-time gesture data from the piano keyboard can be used as a controller for additional synthesized sounds, or augmented acoustic sounds [McPherson and Kim, 2010], or again to simplify machine-listening tasks when interacting with a computer interface [Kallionpää et al., 2017]. A self-playing piano equipped with performance acquisition sensors can be used to replicate a performance on the instrument at a later time. Moog

8 http://touchkeys.co.uk
9 https://www.enhancia.co
10 http://sonicmanipulations.com
11 "Vibe Bar Demo - Brassy Sound", https://www.youtube.com/watch?v=7DbwAcXYEfg
12 https://www.keithmcmillen.com/

and Rhea [1990] mention one such example, where the symbolic data from a live piano performance is acquired to be re-used later for an audio recording of the performance without the noise normally associated with the presence of the audience. Researchers in the field of music performance, musical acoustic and machine learning can also benefit from an accurate transcription of timing and velocity information of a given performance [Goebl et al., 2008]. Physiological studies on the biomechanics of the key press make use of accurate descriptions of the movements and forces involved in the piano press [Furuya and Altenmüller, 2013, Metcalf et al., 2014].

As the player presses a key on the piano, the hammer action accelerates the hammer to the point of "let-off", after which the hammer, now in free flight and mechanically decoupled from the key, hits the string, marking the beginning of a note. For this reason, many sensing systems focus on tracking the position of the hammer, often just its speed at the time it hits the string (the "Final Hammer Velocity", FHV). However, the accurate timing of the key movement is also of high relevance in several contexts, for instance when studying asynchronies between voices [Goebl, 2001], for certain augmentation strategies [McPherson and Kim, 2010], and to extract features of the key press [McPherson and Kim, 2011, Bernays and Traube, 2012].

As early as 1895 scientists and musicologists started devising bespoke technologies for measuring and recording key and hammer positions. Goebl et al. [2008] presents a complete historical account of such sensing technologies, while here we concentrate on the techniques of more relevance in the present days.

Two manufacturers of acoustic pianos, Bösendorfer and Yamaha, have commercialized "reproducing pianos" since the 1980s. These incorporate hammer detection capabilities as well as a self-playing mechanism based on solenoids or linear motors. The Bösendorfer SE290 uses two photocells and two shutters mounted on the hammer shank in order to detect the position of the key at two fixed points: that of hammer-string contact, and another one 5 mm earlier. A dedicated digital counter measures the time interval between the two events, yielding a 10-bit value for FHV [Stahnke, 1981]. An additional shutter and photocell are mounted underneath the key for calibration purposes and to detect key release [Goebl, 2001]. The Yamaha Disklavier pianos use a similar apparatus to detect FHV. Earlier Yamaha models use a 7-bit value for encoding FHV, while Disklavier Pro models (introduced in 1998) use a 10-bit value. The Bösendorfer CEUS, introduced in 2006, adds tracking of the key position using optical reflectance sensors mounted underneath the keys at a 500 Hz sampling rate. These pianos are mainly designed to be used for self-playing applications, but they have largely been used for research applications at the few institutions that could afford them [Goebl and Bresin, 2003, Goebl et al., 2008].

QRS music's PNOScan[13] is a commercially available retrofit that is installed on the keybed underneath the keys which uses optical reflectance sensors to track continuous key angle, and no mechanical modification to the action are required [Starkey, 1984]. The PNOScan and other similar devices, such as Willis [1995], can be fitted to acoustic pianos as well as electro-mechanical keyboards (such as the Rhodes, Wurlitzer, or Hammond), in order to add MIDI output capabilities to pre-existing instruments.

The Moog Pianobar [Mowat, 2005] is a portable device designed to be used on any 88-key grand piano in order to generate MIDI data from a piano performance. It can be installed and calibrated in a matter of minutes. The Pianobar sits on top of the keyboard and is equipped with optical reflectance sensors that measure key-angle by shining an infrared light on the surface of the white keys, and measuring the amount of light reflected by the key with a phototransistor. At the rest position, most of the emitted light is reflected into the sensor. As the key angle departs from the rest position, more and more light is reflected away from the sensor. Black keys would absorb most of the infrared light that is shone on them, greatly reducing the dynamic range of the optical reflectance measurements, so the Pianobar uses optical interruption sensing for the black keys, by shining an infrared beam across the back of the key to a phototransistor detector at the other side [McPherson, 2015]. Even in this arrangement, it produces less accurate readings on black than on white keys.

Despite the fact that the sensing technology on the Pianobar would in principle allow to measure continuous key-angle, the device itself only outputs *NoteOn* and *NoteOff* MIDI messages.[14] Similarly, QRS's PNOscan is using continuous key position internally, but ultimately only outputs discrete MIDI Note messages, or a low-resolution version of the internal position that is not very informative. As of the beginning of the 21st century, there was no commercially available device capable of sensing continuous key position, and even today, this only comes with

13 http://www.qrsmusic.com/PNOScan.asp

14 In MIDI terms, a *NoteOn* message is sent when a note is started (e.g.: a key is pressed), and it comprises two pieces of information: the note number and the onset velocity. Similarly *NoteOff* is sent when the note ends (e.g.: a key is released), it includes the note number, and it may or may not include an associated velocity information.

the Bösendorfer CEUS, which is a premium instrument, with a price tag that does not make it easily available to many research communities and musicians.

When Goebl needed to compute the finger-key contact times in order to study the asynchronies between keys when playing notes with different dynamics, they managed to compute the time of initial finger-key contact from a Bösendorfer SE290, by means of a clever hack. They accessed the calibration data of the instrument, which contains information about the typical time intervals between the beginning of the key press and the hammer hitting the string for different dynamics, and they were thus able to compute the times of key onset even though they were not directly accessible to the user by design [Goebl, 2001].

McPherson and Kim [2010] modified a Pianobar in order to provide continuous key angle to a host computer over USB. The initial use case for this was to use continuous key position to control the sound generation in the Magnetic Resonator Piano. McPherson later presented their own portable measurement system in a similar form factor to the Pianobar [McPherson, 2013]. For their own keyboard scanner, they turned to optical reflectance sensing for both white and black keys, made possible by "preparing" the piano by applying a short piece of white tape at the far back of the black key. The device transmits raw key-position data sampled at 1000 Hz to a host computer, where they can be used directly, or a dedicated software can turn these into MIDI NoteOn/NoteOff messages, and detect percussiveness. In our studies in Chapters 4 and 5 we will use McPherson's keyboard scanner.

When researching the hammer action behaviour, dynamics and acoustics of a piano, individual keys can be instrumented so that the position and relative timing of individual moving parts in the hammer action and the key can be measured accurately. Usual approaches include accelerometers [Goebl et al., 2003] and adding conductive surfaces to parts of the action so that their interaction can be measured electrically [Askenfelt and Jansson, 1990].

2.3.2 *Electronic keyboards*

The most common approach to sensing key motion on electronic keyboards is via switches closing during the key throw. As mentioned, the Hammond has a stack of nine contacts closing as the key is pressed down. Many electronic organs had a similar contact stack to allow several registers to play at the same time. Synthesizer keyboards would normally use one or two contacts per key. While in principle one contact per key would be enough to generate a voltage output corresponding to the pitch of the note being pressed, when using two contacts, one which triggers slightly earlier and one slightly later in the key throw, the one that triggers later can be used as a gate, yielding a more accurate pitch during the release of the key. Many synthesisers, including the Minimoog, had two contacts. A similar two-contact keyboard assembly was used on the first electronic piano to feature velocity sensitivity, the Roland EP-30 from 1974.[15] In the EP-30, a capacitor would charge in the time intervening between the closing of the two contacts, and its voltage would be used to determine the loudness of the corresponding sound. To this day, most velocity-sensitive keyboards still use two switches that close at different points, and the resulting time difference is used to compute a discrete velocity parameter. On some keyboards, it is possible to trigger the note on the topmost switch (ignoring the second one, and therefore the velocity parameter), so that the note starts earlier in the key throw. This setting is particularly appreciated by Hammond players, as the key on the Hammond starts emitting sound earlier in the key throw than most digital keyboards. Some digital pianos feature a simplified reproduction of the keyboard action of the piano, including a hammer, and would sometimes track the hammer instead of the key itself. Accurate tracking of hammer by using optical reflectance sensor is described in Kawamura and Muramatsu [1993]. The Yamaha AvantGrand series uses optical tracking of the hammer.[16] Casio and Kawai offer several piano actions with three switches tracking the key position, which improve responsiveness when repeatedly striking a key.[17] Three-contact actions are also featured in organ-style keyboards, as seen on Hammond-Suzuki's XK-5[18], which we discuss more in detail in Section 3.2.5. Other technologies to detect key position include linear Hall-effect sensors, as seen on Infinite Response's VAX77[19], or capacitive rubber, as seen on Kurzweil's MIDIBOARD [Moog, 1987]. On these same instruments, the Hall-effect sensors and capacitive rubber respectively are also used to detect aftertouch. Capacitive plates

15 https://astronautapinguim.blogspot.com/2012/02/roland-ep-30-electronic-piano-english.html

16 Yamaha AvantGrand product page: https://uk.yamaha.com/en/products/musical_instruments/pianos/avantgrand/index.html

17 Kawai keyboard actions: http://kawaius.com/technology/wooden-key-actions/; Kawai MP7SE product page: http://www.kawaimp.com/mp7se/detail/touch/; Casio Celviano product information: https://www.casio-intl.com/asia/en/emi/celviano/info/

18 Hammond XK-5 product page: http://hammondorganco.com/products/portable-organs/xk-5/xk5/

19 VAX77 product page: http://www.infiniteresponse.com/page8.html

can also be used for aftertouch [Buchla, 1985], as well as pressure sensors under the keys.

Continuous key-position tracking has been achieved with electrostatic sensing [Le Caine, 1955, Moog, 1982, Fourier et al., 1994], and more recently with optics-based solution. Freed and Avizienis [2000] use optical interruption sensing to measure key position, employing one LED and one phototransistor per key. Their sensors have to be installed permanently inside a keyboard, requiring modification to the key action, with the addition of a spring-loaded mechanism. Other solutions using infrared sensing include the aforementioned QRS PNOScan and McPherson [2013]'s piano scanner. Medeiros and Wanderley [2013] compare infrared, strain gage and Hall effect sensors to measure deflection of a cantilever beam and suggest a sensor fusion approach for optimal results. The authors later suggested that a similar approach could be used for the detection of the position of piano keys [Medeiros and Wanderley, 2014].

Finger position on the key has been measured with capacitive sensing [Moog, 1982, McPherson, 2012], high-speed imaging [MacRitchie and Bailey, 2013], or combinations thereof [MacRitchie and McPherson, 2015]. The use of matrices of force-sensitive resistors has also been explored [Grosshauser and Tröster, 2013] for the same purpose. The force exerted on the key is measured with force transducers, piezoelectric [Harding et al., 1989, Askenfelt and Jansson, 1991], or strain-gauge [Kinoshita et al., 2007], while [Goebl and Palmer, 2008] infer the acceleration, and thus the force, from motion capture data.

2.4 LIMITATIONS OF MIDI

MIDI 1.0 has been the de-facto standard for communication between controllers and sound generators for over three decades [IMA, 1983]. However, issues with its limitations in terms of resolution and bandwidth have been raised several times in the years following its release [Moore, 1988], which in turn fostered the development of alternative protocols, such as ZIPI [Wright, 1994] and OSC [Wright and Freed, 1997]. The bandwidth issue has partly been addressed by using the USB protocol as a transport layer, which enables much faster data rates than the UART transmission prescribed by the original standard.[20] Coenen [1992] reviewed reproducing pianos in the early 1990s. The devices in exam were MIDI-compatible, yet some of them could internally perform better than MIDI, both in terms of resolu-

20 "Basics of USB-MIDI", https://www.midi.org/articles/basic-of-usb

tion (the Bösendorfer SE290 has a 10-bit internal data representation for velocity, while MIDI only uses 7-bit values), or bandwidth (using parallel data internally is faster than MIDI serial). This is an indication of how MIDI can effectively act as a gating factor, as it enforces the sending and receiving parties to reduce the quality of the information they transmit and process in order to be compatible with other devices.

Wright [1994] lists as one of the limitations of MIDI that of having to *"commit"* to the pitch of a note at the moment the note starts. When MIDI data is generated by a pitch tracker on an acoustic instrument, the authors argue, this forces a delay in the start of the note, while a better option would be to send distinct messages for NoteOn, at the moment the onset of the note is detected, and pitch, a few milliseconds later, when the pitch of the note has been inferred. The underlying problem is that, in the MIDI specifications, notes are discrete events, and pitch bend and modulation controllers, which can be sent continuously, cannot be paired with individual notes, rather they are associated with a specific "channel"[21]. Thiebaut et al. [2013] investigate the issues of this paradigm when dealing with Multi Dimensional Controllers, such as the Roli Seaboard or the Haken Continuum. On these instruments the same nature of the instrument's interface allows to pitch bend individual notes, and said limitations make it hard to connect these instruments to sound generators other than their own. The MIDI Polyphonic Expression (MPE)[22], introduced in 2017, was created partly to address these concerns. MPE does not fundamentally change the MIDI protocol, in fact the data being transmitted are MIDI 1.0 compatible, however it uses a clever management of the MIDI channels to ensure that only one note is played on one channel at a time, so that pitch bend information, and other controllers, can be sent per-note.

Even MPE, however, does not address the fundamental limitation of considering notes as discrete events associated with a velocity value. For instance, let us consider the case of a keyboard capable of sensing continuous key position. The device can send positional data for the key (continuously, or upon change), but will also have to send a NoteOn and a NoteOff event. *When* should these events be sent? And what should be their associated velocity? If a NoteOn message has to be sent as soon as the key starts moving, then it may be too early to be able to accurately compute the velocity information, and also the key motion may just be the result of an unintended gesture, which will not be followed up by a full key

21 In typical use, one channel corresponds to one instrument

22 "MIDI Polyphonic Expression (MPE) Specification Adopted!", `https://www.midi.org/articles/midi-polyphonic-expression-mpe`

press. On the other hand, if a NoteOn message is sent after the key has travelled for a few millimeters (as it is often the case for regular keyboards), then the sound generator will have trouble understanding the continuous position data sent until that moment, because it would have no pitch value associated with it, as the pitch value is provided by the NoteOn message.

From our review of commercial devices and sensing technologies, it emerges that continuous key position sensing is not widely available. Even when it is used internally by a device, the continuous data often remains within the device, and its output consists of discrete events (for instance, NoteOn and NoteOff). Whether this is a direct consequence of the MIDI specifications is hard to tell, but definitely the lack of a suitable representation of the data in MIDI terms would make it hard to interface continuous controllers with generic sound generators. The Magnetic Resonator Piano, which uses key position for real-time continuous modulation, uses a customised approach at transporting the data internally from the controller to the sound generator via OSC and MIDI [McPherson and Kim, 2012]. One of the strengths of MIDI is its generality and universality: "any" controller can be connected to "any" sound generator and produce a meaningful result. While both MPE and the Magnetic Resonator Piano demonstrate that MIDI *can* be exploited to transport continuous data for each note, these are in practice little more than bespoke hacks that rely on some strong assumptions between the controller and the sound generator, and these fail to generalise to all devices. The upcoming MIDI Profiles[23] present themselves as a promising evolution in this sense, that will allow to codify different assumptions about what data a controller can be expected to generate, and a sound generator can be expected to respond to, however their eventual effectiveness, diffusion, and overall impact are yet to be evaluated.

2.5 POTENTIAL OF CONTINUOUS SENSING FOR KEYBOARD AUGMENTATION

We have seen in the previous sections that gestures on keyboards are a complex, multi-dimensional phenomenon, which mostly ends up being simplified for digital keyboards to a scalar velocity value. The purpose of this thesis is to investigate the expressive potential of digital keyboards when retaining a much larger bandwidth between the key movement and the sound generator. One of the reasons for augmenting the keyboard, as opposed to creating a completely new interface, is to

23 "Details about MIDI 2.0, MIDI-CI, Profiles and Property Exchange": https://www.midi.org/articles-old/details-about-midi-2-0-midi-ci-profiles-and-property-exchange

add to its control capabilities, while capitalising on the pre-existing sensorimotor skills. Cook [2001] states, among their principles for designing music controllers, that *"Some players have spare bandwidth, some do not"*, and encourages the instrument designer to leverage that spare bandwidth. In exploring suitable mappings for TouchKeys, McPherson et al. [2013] raise the issue of human factor limitations when augmenting a familiar interface. Keyboard technique is an art in itself, and it comes with its own set of constrains and degrees of freedom. McPherson et al. find that the absolute positioning of a finger on the key in regular keyboard playing is dependent on fingering and context, such as the preceding and following note. Using absolute positioning as a source of modulation would in turn require the player to be conscious of the placement of each finger the whole time, which is considerably more complicated, as it would lead to revolutionise their technique, and is also not ergonomic. McPherson et al. therefore find the *"space between the notes"*, to ensure the disruption to regular playing techniques is minimized, and at the same time find where the *spare bandwidth* of the player is, so that it can be leveraged for the purpose of the augmentation.

One place where we hope to find the spare bandwidth when augmenting the keyboard with continuous key sensing is the attack of the note. We have seen that a significant amount of attention is devoted to touch as part of piano pedagogy [Bernays and Traube, 2013, MacRitchie, 2015], however the type of touch used rarely makes a difference on the produced sound [Ortmann, 1925, Goebl et al., 2014]. McPherson and Kim [2011] showed that several different gestures during note onset can be identified from continuous positional data, so that a player could effectively assign different meanings to each note depending on how they attack it. An augmentation could therefore add a distinct sonic effect to each of several gestures. Further sources of inspiration can be found in the behaviour of the Ondioline (key position to volume), the Hammond organ (multiple switches, one for each harmonic), and the Magnetic Resonator Piano (key position to volume and multi-key gestures) [McPherson, 2010].

3 DYNAMIC TEMPORAL BEHAVIOUR OF THE KEYBOARD ACTION ON THE HAMMOND ORGAN AND ITS PERCEPTUAL SIGNIFICANCE

This chapter contains significant material from the article "Dynamic temporal behaviour of the keyboard action on the hammond organ and its perceptual significance" by Moro, McPherson and Sandler, originally published on The Journal of the Acoustical Society of America, 142 (5): 2808–2822, 2017 [Moro et al., 2017].

3.1 MOTIVATION

The Hammond organ occupies a prominent position in popular music. After its introduction in the 1930s, it was widely used in popular music since the 1950s and its sound has been heard on countless recordings.

Where the piano and most digital keyboards have a clear relation between velocity and the produced sound, the Hammond shows a more subtle effect: the key-click, a distinctive transient in the sound at the beginning of every note. This click is the result of the behaviour of nine different contacts embedded in the keyboard mechanism. As we show in this chapter, these contacts are not ideal switches: they do not close at the same time and they exhibit bouncing. These characteristics change in response to both the speed and the quality of the key press, giving the instrument its own distinctive form of touch response.

We conduct two studies to characterise the behaviour of the Hammond organ keyboard. Our first study analyses the dynamic electromechanical behaviour of the key action and its main components. Through position and electronic measurements we show that the moving parts in the action react to the details of the key press, producing measurable differences in the timing and characteristics of the onset transient. The second study is a listening test to assess to what extent the measured differences in contact behaviour are relevant to the listener. We find that, although the key-click is a short burst of a few milliseconds at the note onset, a statistically significant number of participants are able to correctly infer the type of touch used to generate a note.

3.2 THE HAMMOND ORGAN

3.2.1 *History*

Laurens Hammond started prototyping the Hammond organ in 1929, patented it in 1934 and first made it available to the general public in 1935, making it one of the first commercial pipeless electronic organs [Hammond, 1934, Roads, 1996].

The Hammond organ was originally designed and sold as a cheaper substitute for church organs. While offering a wide palette of sounds, its timbre was less rich in harmonics than pipe organs and the attack of the note on the Hammond was much faster and sharper, somehow limiting its realism as an emulation. Regardless, many church communities were willing to accept this trade-off given the lower cost of the new instrument [Ng, 2015, 23-24]. Since the moment of its introduction, the Hammond organ was used in a variety of genres, far beyond the original aims of its inventor. It was used in classical, gospel, rock, jazz and all sorts of popular music, as well as radio and TV shows, theatres, stadiums and other public venues, cruise ships [Vail, 2002, 13-24]. Some of the reasons for this success can be traced back to the attack sound itself: the fast attack allows playing faster tempi with accurate rhythmic precision [Ng, 2015, 36-37].

3.2.2 *Principle of operation*

The Hammond organ, as patented by Laurens Hammond, is an electromechanical keyboard music instrument. It is in essence a polyphonic additive synthesizer, whose oscillator bank consists of 91 quasi-sinusoidal signals. Each of these is generated by a mechanical dented wheel (tonewheel), spinning driven by a synchronous AC-motor. Each tonewheel induces a sinusoidal signal at the output of an electromagnetic pickup placed in front of it[1].

1 The profile of the tonewheel is designed and cut in such a way that it induces a sinusoidal signal in the pickup. However, manufacturing and mechanical tolerances may lead to the generated sound including higher order harmonics [Hammond Instrument Company, 1941, sec. 2, p. 10]. In order to remove the higher components, a passive filter consisting of capacitors and transformers is added to some of the tones [Hammond Instrument Company, 1941, sec. 2, pp. 11-12]. The harmonic content

Contact number	Harmonic	Interval
9	eighth	+3 octaves
8	sixth	perfect fifth +2 octaves
7	fifth	major third +2 octaves
6	fourth	+2 octaves
5	third	perfect fifth +1 octave
4	second	+1 octave
3	fundamental	unison
2	sub-third	perfect fifth
1	sub-fundamental	- 1 octave

Table 3.1: Harmonic ratios and intervals of the frequencies routed to the contact switches of each playing key.

The Hammond C-3 used in this study, one of the most popular models, has two 61-note (C2 to C7) keyboard manuals, as well as pedals and an expression pedal. Each key on the keyboard manuals closes multiple contacts: seven were present in the original patent, but there are nine in most tonewheel Hammonds, including the C-3. Each contact is connected to one of the sinusoids from the tone generator, and each of these corresponds to the frequency of one of the harmonics or sub-harmonics of the note, as outlined in Table 3.1 [Hammond Instrument Company, 1941, sec. 2, p. 17-19]. The bottom 12 tonewheels from the generator are reserved for the pedals, and only 79 are routed to the playing manuals. As these do not cover the entire range of frequencies needed for the playing manuals, some of the keys from the bottom octave and some from the topmost two octaves use the signal from a tonewheel one octave below or one or two octaves above the nominal frequency ("foldback") [Wiltshire, 2008]. An intricate web of 549 resistive wires for each of the two manuals routes the generator tones to the contact switches.

Every time a key is pressed, this causes each contact to close against one metal bar ("busbars"). When a contact is closed, this connects the signal from the generator to the busbar through the resistive wire, which allows for passive summation of several signals on the same busbar. The output of each busbar is connected to one tap of a matching transformer which sums the signals coming from the bus-

of the tones has not been carefully studied in the literature, however [Werner and Abel, 2016] and [Pekonen et al., 2011] treat them as sine waves.

bars before feeding them to the preamplifier. The relative levels of the harmonics of the notes on each manual are adjusted by selecting the tap of the matching transformer used by each busbar. Ten "preset keys" (reverse-colored keys with locking mechanism at the left end of each playing manuals) allow quick access to predefined harmonic combinations, while a set of nine[2] levers ("drawbars") allows the performer to adjust the harmonic mix to taste.

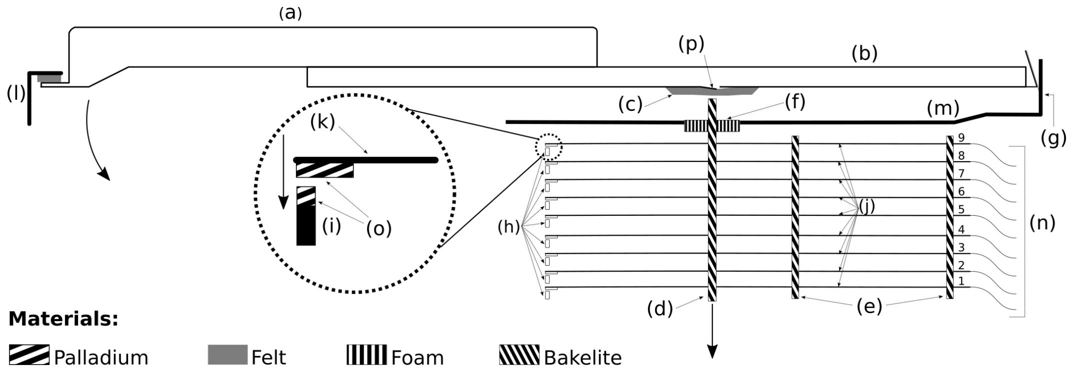
A general overview of the routing mechanism can be found in Wiltshire [2008]; a detailed technical description is in the service manual [Hammond Instrument Company, 1941, p. 17-19], while a simplified mathematical model can be found in Werner and Abel [2016].

In this chapter we use the term "Hammond organ" to refer to tonewheel-based organs manufactured by Hammond, of the likes of the classic models B-3/C-3/A-100, as opposed to later models from the company which featured solid-state or digital generators. On these instruments Each of the two organ keyboards has 61 playing key (C to C) and 12 inverted color "preset" keys (C to B), at the bottom end of the keyboard.

3.2.3 *Key mechanics*

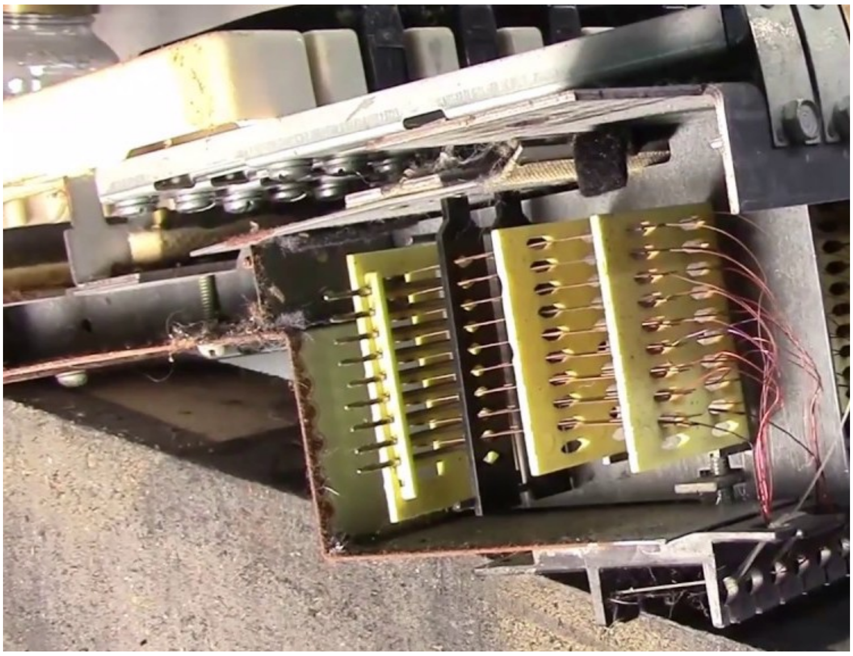
Here and in the remainder of this chapter we will refer to a 1967 Hammond C-3 available at our lab, which we used for our measurements and recordings. The organ was in working order and was recently serviced by a specialised technician. A diagram of the key action and a picture of the contacts can be found in Fig. 3.1. The keyboard action features square-front ("waterfall") key caps mounted on metallic shafts, equipped with a return spring which is responsible for returning the key to the home position once it is released. A bakelite contact pusher is positioned vertically below the key shaft, at about 3 cm from the pivotal point at the back of the key. Where the bottom of the key shaft makes contact with the contact pusher there is a tiny strip of felt which couples the metal key shaft to the bakelite strip. A metal tab is cut out on the bottom of the key shaft and allows adjustment of the distance at which the felt engages the contact pusher. The contact pusher has nine horizontal openings, each holding one bronze contact spring. Each contact spring is connected via a resistive wire to one tone from the generator. The busbar and the spring contact are covered in palladium to ensure optimal conductivity [Ham-

2 The number of drawbars is always the same as the number of busbars and as the number of contacts per key, nine in the case of the C-3, as noted above.



(a) Schematic representation of a typical tonewheel Hammond key and the underlying contacts. Solid arrows indicate the directions of the motion of the key, the contact pusher and the tip of the spring contact.

(a) plastic key, (b) metal shaft, (c) contact pusher felt, (d) contact pusher, (e) insulated support, (f) foam, (g) pivotal point, (i) and (h) busbar, (j) and (k) contact spring, (l) and (m) support frame, (n) resistive wires from generator, (o) precious metal points, (p) adjusting tab



(b) Picture of the contacts of a tonewheel Hammond, taken from `https://youtu.be/_h6CKRzkhTs`

Figure 3.1

mond Instrument Company, 1941, sec. 2, p.17]. The top of the contact pusher is inserted in a slot cut out from a layer of foam which prevents it from moving sideways in the horizontal plane.

The pivotal motion of the key is transformed in a vertical motion of the contact pusher and of the spring contacts, with the key assembly effectively acting as a

9-pole-single-throw switch. When the key is at rest, the felt is resting about 2 mm above the contact pusher. During the key travel, the felt engages the contact pusher and pushes it down. This in turn actuates the stack of spring contacts by pushing each of them against its corresponding busbar, thus connecting the tones from the generator to the output. The maximum displacement of the front of a white key is 9 mm, while the maximum displacement of the shaft at the felt is about 4 mm. We did not investigate the mechanical details of the contact switches, however more details on their design can be found in Hammond [1934, p. 7-8].

In spite of the precious metal coating, dirt, oxide and dust will often make the contact less than ideal over time [Vail, 2002]. Additionally, contacts may not close simultaneously and each of them may bounce multiple times. When the signal at the input of the contact is non-zero and the contact is switched, this causes a transient in the output signal, thus giving rise to the key-click.

3.2.4 *Touch on the Hammond organ*

The literature on the effect of touch on the piano and its effects on the produced sound is abundant, as we showed in Section 2.1. However, no formal study on the player's touch on the Hammond organ can be found in the literature.

When a key is pressed, the key contacts are connected to the busbars, generating an onset transient, after which the amplitude and harmonic characteristics of the sound remain constant for the entire duration of the note. As soon as the key is released, the contacts are disconnected from the busbars and the note terminates with an offset transient. The sound produced by a key press has – as a first approximation – a rectangular amplitude envelope, gated only by the key press, with the key-click marking the beginning and ending of the note. The volume of a note can be varied only through the use of the registration drawbars or the volume pedal, the former affecting all the notes on the corresponding manual and the latter affecting all the notes being played on the organ. The velocity at which a key is pressed will not affect the amplitude or the harmonic content of the steady-state part of a note.

Laurens Hammond always considered the attack click of his instrument more as a defect, rather than as an expressive feature and tried to make it less prominent with low-pass filtering in the power amplifier and on dedicated speakers. Regardless of the designer's opinion on the subject, the "key-click" is, to date, one of the most appreciated characteristics of the Hammond organ by players, so much that when newer technologies were introduced to completely eliminate the key-click from fully electronic organs, musicians objected to the consequent lack of articulation [Vail, 2002, p. 44-45].

The particular contact stack in use also allows some extended playing techniques. Progressive key presses consist in slowly depressing a key, so that individual harmonics will start playing one at a time, as soon as each contact touches the corresponding busbar. Partial key presses are also possible, when the key is not pressed all the way down and it does not trigger all of the harmonics. These are commonly used to obtain percussive "non-pitched" sounds. The squabbling technique[3] consists in playing a chord with one hand where one or more of the intermediate keys are partially pressed.

3.2.5 *Digital emulations of the Hammond organ*

There are in the academic literature very few publications on Hammond emulation, while more space is given to the emulation of the Leslie rotary speaker. A computationally-efficient emulation of the key-click sound was proposed in Pekonen et al. [2011] using an Attack-Decay-Sustain-Release envelope applied to the sixth harmonic of the fundamental note. Reid [2003], attempts the exercise of emulating a Hammond organ with a traditional synthesizer, and uses an envelope modulation of the Voltage Controlled Filter cutoff frequency to model the keyclick.

A large number of software and hardware clones are available on the market. These are sometimes referred to as "clonewheel", a word generated by the union of "clone" and "tonewheel". A review of clonewheels was published in Vail [2002], but, being it now over 15 years old, it is now mostly superseded. We present here a brief review of currently available state of the art Hammond emulators.

One of the most renowned virtual instrument which emulates the Hammond sound is the VB3-II[4], which is also used as the sound generator for the hardware emulator Crumar Mojo[5]. The VB3-II emulates the 9-contacts onset computing a "contact closure time" from the MIDI velocity parameter. An open-source software project, `setBfree`[6], provides a comprehensive emulation of the instrument.

3 "Must-Know Techniques for Hammond Organ": https://www.keyboardmag.com/lessons/must-know-techniques-for-hammond-organ

4 "VB3-II Tonewheel Organ Simulator": http://www.genuinesoundware.com/?a=showproduct&b=24

5 http://www.crumar.it/?a=showproduct&b=4

6 https://github.com/pantherb/setBfree/

The reference Hammond emulator stand-alone sound generator is an FPGA-based module, the HX-3[7], also available as the embedded sound generator on many third-party controllers. The only digital emulator to date which properly recreated the multiple contacts per key as in the original Hammond organ is the "New B3"[8], produced by Hammond-Suzuki in the early 2000s, and now discontinued. The New B3 replaces the electromechanical tone generator with an entirely digital "virtual-tonewheel" oscillator bank, with 91 DACs which act as the tone generators. Tones are then routed to the keyboard assembly just like in a tonewheel Hammond organ, so that each key would actuate a 9-contact stack similar to the one in tonewheel Hammonds. Key click and touch responsiveness come automatically as part of this hybrid digital-mechanical-analog design. The high retail price and considerable weight of these instruments lead to modest commercial success, despite collecting favourable reviews from professional Hammond players . In an interview[9], player Tony Monaco admitted that the most prominent feature of the New B3, which is missing from all the other clones in the market, is the faithful reproduction of the touch response of the original instrument. In 2016, Hammond-Suzuki introduced the XK-5 model which features a traditional electronic keyboard keybed with a total of three mechanical contacts per key, each triggering three contacts in the synthesis engine.[10] The ongoing effort by Hammond-Suzuki and other companies to reproduce the details of the control and the sound of the note onset to the highest degree of detail indirectly shows their relevance to the musicians.

While most aspects of the Hammond sound and interface are covered by dedicated emulators, none of those currently available on the market allows to control the touch similarly to the original instrument and this is a major cause of complaints from experienced players of the original instrument. While the multiple-contacts-per-keys was used in some analogue transistor organs (e.g. Crumar Organizer) and on the New B3, it is a very expensive technology, and so more traditional 2-contact keybeds are used by most emulators today (again, except the XK-5). This leaves performers on digital emulations without the possibility to perform a series of gestures only possible with a multi-contact action or a continuous position controller, paired with an appropriate software implementation. Besides

7 http://wiki.keyboardpartner.de/index.php?title=Hauptseite
8 "Hammond B3 Modelled Electromechanical Tonewheel Organ": https://www.soundonsound.com/reviews/hammond-b3
9 "The Reinvention of the Hammond B3 ": https://web.archive.org/web/20160204104035/http://jazztimes.com/articles/14128-the-reinvention-of-the-hammond-b3
10 http://hammondorganco.com/products/portable-organs/xk-5/

preventing techniques like the percussive non-pitched staccato [Madera, 2011, p. 121], or any gesture that does not bring the key to the bottom of the keybed, this also takes away from the player the subtle control on the attack of individual notes.

3.2.6 *Effect of contact bounce on the audio signal*

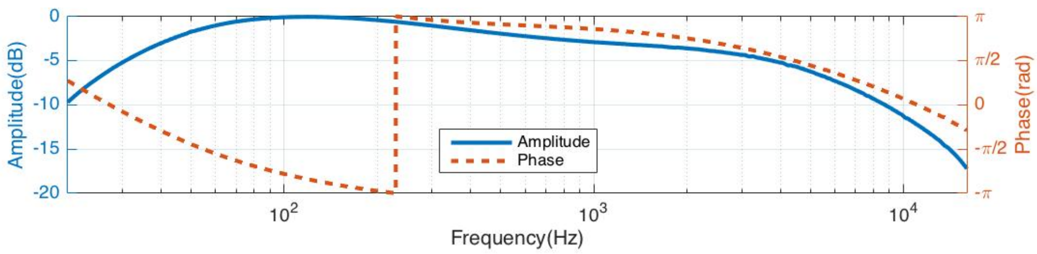


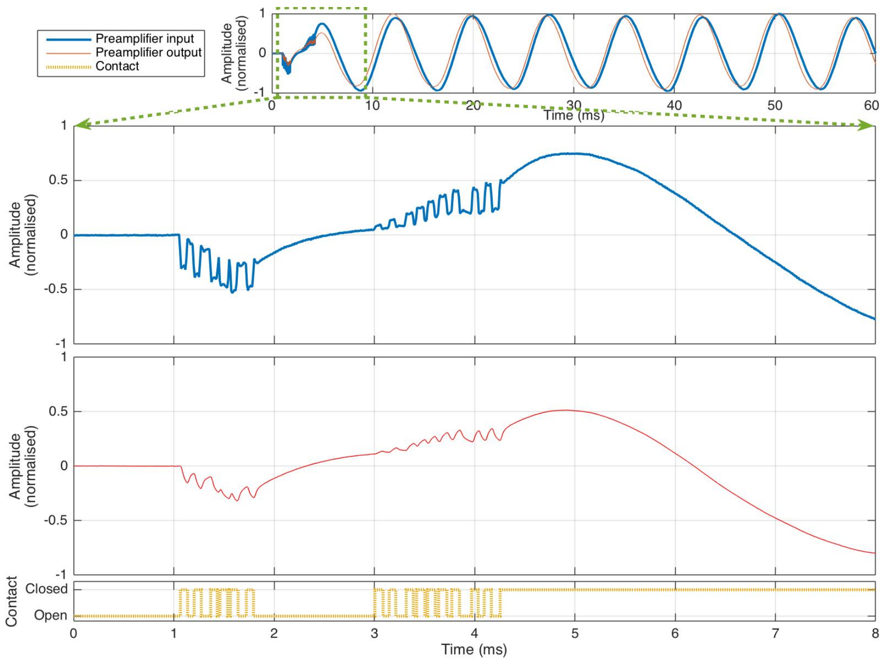
Figure 3.2: Transfer function of the organ preamplifier, measured with a test signal consisting of a sine sweep between 2 Hz and 20 kHz with an amplitude of 0 dBu.

In order to get an indication of the effect of the contact bounce on the audio signal produced by the instrument, we took paired recordings of the audio signal at the input of the preamplifier and at its output (the line output of the organ) during a key onset, obtaining the waveforms in Fig. 3.3. The transients and high-frequency oscillations in the initial part of the signal are due to the bouncing of the contact associated with each active drawbar. The signal only reaches full amplitude several milliseconds after the contact settles in the closed state. As the signal passes through the preamplifier, whose frequency response is shown in Fig. 3.2, the high-frequency components are attenuated and the transients are smoothed.

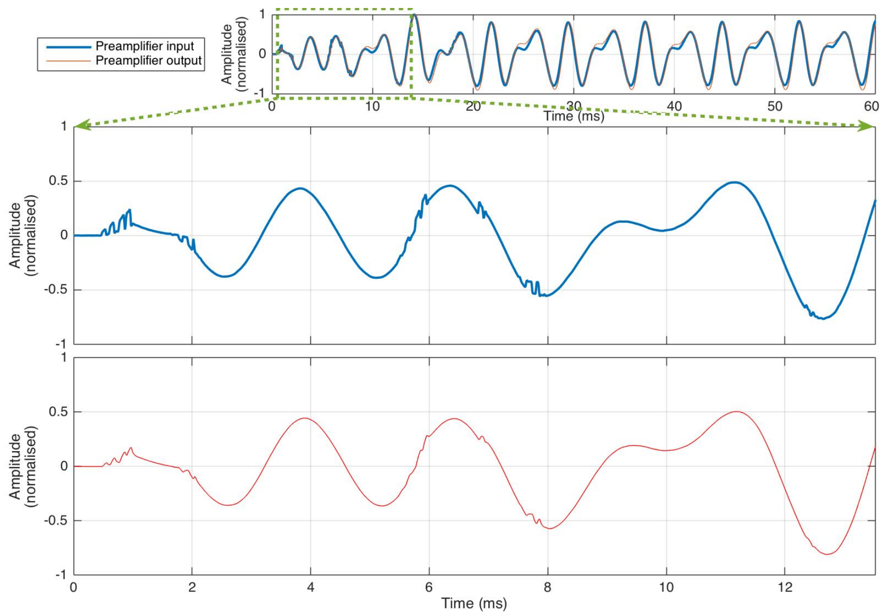
In Fig. 3.3a, a single drawbar is active, therefore all the transients are due to a single contact. By inspecting the waveform at the preamplifier input, we were able to infer and manually annotate the open/closed state of the contact, shown at the bottom of Fig. 3.3a. In Fig. 3.3b multiple drawbars are active, producing a more complex tone. As the contact associated with each drawbar closes, the corresponding frequency is added to the output signal after a brief transient noise caused by the bouncing of the contact.

3.2.7 *Implications*

The Hammond organ features a multi-contact keyboard and shows no relation between the velocity of the key-stroke and the loudness of the produced sound, yet



(a) Single drawbar



(b) Multiple drawbars

Figure 3.3: Comparison of the signal at the output of the generator and the signal modulated by the key contact during an onset, measured at the output of the matching transformer. The note played is a C3, with a fundamental frequency of 131 Hz. In (a) only the drawbar 1 was active, while in (b) drawbars 1, 2, 3 and 4 were active. The contact state at the bottom of (a) is inferred from the discontinuities in the audio signal.

some clear consequences of the multi-contact array allow the extended techniques mentioned above. Other, less-obvious dependencies may reside in the subtle control available in the shape of the key profile during a note onset, and in the way it affects the key-click. We therefore investigate if and how the velocity and the type of touch used can affect the sound of a note and how this effect is perceived by a listener.

3.3 DYNAMIC MECHANICAL BEHAVIOUR

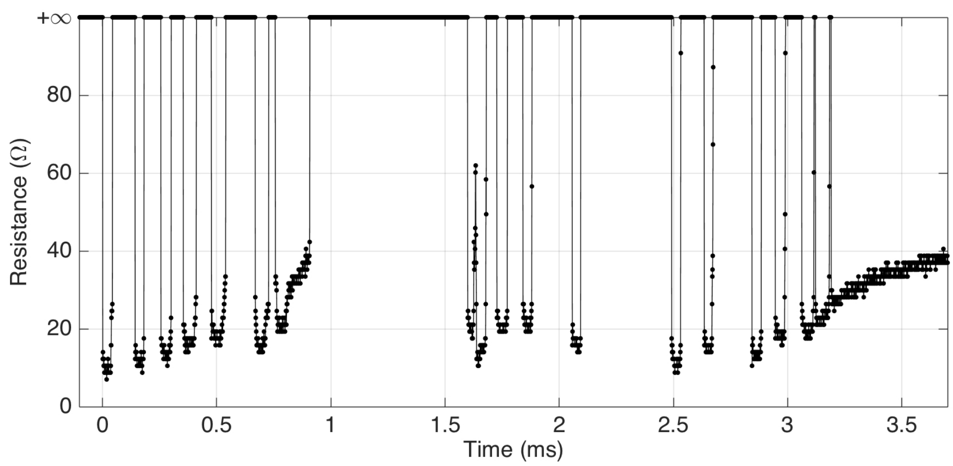


Figure 3.4: Detail of the measured resistance during the onset bouncing of a key contact for a struck touch, sampled at 1 MHz with an Agilent MSO-X-3054A oscilloscope. Each dot represents one sample.

During a note onset, up to nine tones from the generator are connected to the output circuit through the contacts at the back of the key[11]. While the key is at rest, all the contacts are open. When a contact first touches the corresponding busbar, it will usually bounce a few times before it settles in the "closed" position (see Fig. 3.4), affecting the signal as explained in Section 3.2.6. The characteristic transient in the audio signal of each note onset on the Hammond organ, the key-click, ultimately results from the overlapping of the effect of these bounces across all the contacts. One of the consequences of this is that the note's onset transient will begin when the first contact starts bouncing and will stop once all the contacts settle in the "closed" state and the steady state part of the note begins.

11 The actual amplitude of the tones connected to the generator depends on the current drawbar registration.

In this experiment we recorded the continuous position of the key and the electrical state of each of the nine contacts activated by that key, in order to understand the relation between the gesture and the characteristics of the contact bounce. The design of the instrument prevents the capturing of the sound output whilst simultaneously monitoring the contact state.

3.3.1 *Experimental setup*

3.3.1.1 *Key angle*

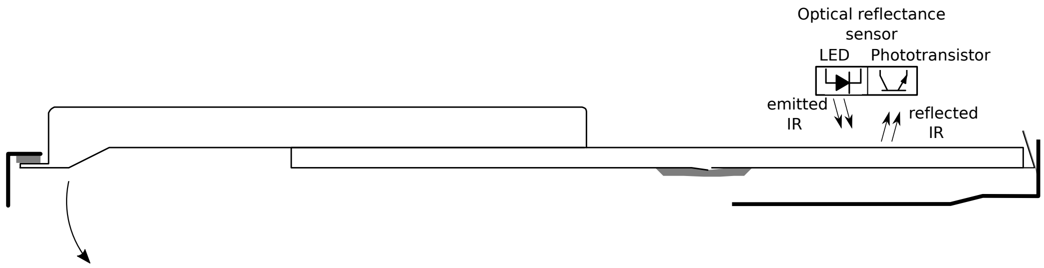


Figure 3.5: Placement of the optical sensor on top of the key shaft.

The key shaft on the Hammond keyboard assembly is U-shaped and its opening is facing upwards. To measure continuous key position, an optical reflectance infrared (IR) sensor was used.[12]. The sensor was placed at the top of the key shaft, pointing down towards the shaft, as shown in Fig. 3.5. To avoid reflections from the internal sides of the shaft, a thin piece of white paper was glued on top of the shaft opening, so to form a uniform, reflective, flat surface to allow for more accurate measurements.

The signal from the optical sensor was buffered, amplified and scaled to a suitable voltage range with a derivation of the circuit in McPherson [2013]. The bandwidth of the sensor and the analog preamplifier was measured with a test signal to be approximately 16 kHz.

Optical reflectance sensors typically exhibit an inherent non-linearity when used to measure distances [Pardue and McPherson, 2013]. This is accentuated in this application by the fact that the surface does not move perpendicularly to the IR beam and therefore the angle between the beam and the reflective surface changes as the key is moved. The reading of the sensor in the range of interest is monotonic

12 Omron EE-SY1200 Photomicrosensor (Reflective), datasheet avaiable at `http://omronfs.omron.com/en_US/ecb/products/pdf/en-ee_sy1200.pdf`

with respect to key position, so that it was possible to compensate for the non-linearity through a calibration procedure consisting of paired readings of a digital caliper and the sensor's output at multiple intermediate key position.

3.3.1.2 *Contacts*

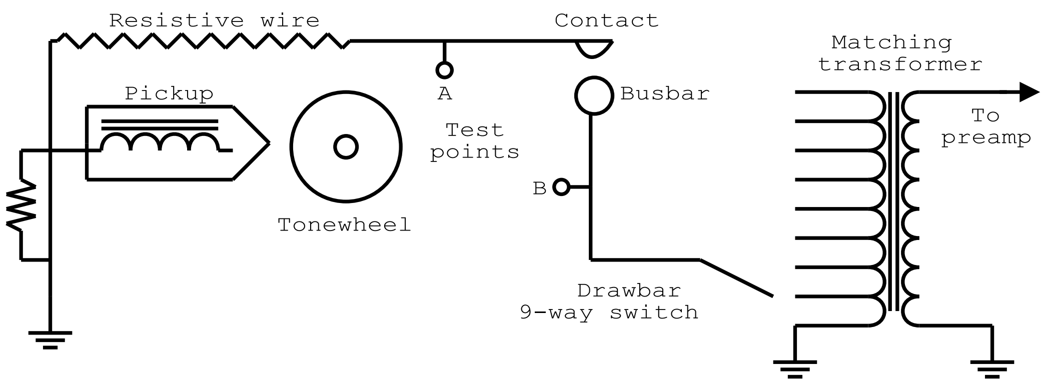


Figure 3.6: Simplified electric diagram from the tonewheel to the matching transformer. Full circuit diagram can be found in Hammond Instrument Company [1941].

In order to avoid the interaction of the circuit under test with the impedances of the generator's pickups and of the output transformers, both poles of each of the nine contacts were isolated from the rest of the Hammond organ and were connected to the test circuits described below. The wires carrying the tones to the key were disconnected at the tone generator end, while the wires from the busbars were disconnected before the matching transformer. These wires are marked as A and B, respectively, in Fig. 3.6.

A preliminary test consisted in measuring the resistance of the switch using a circuit with a passive voltage divider sampled with a digital oscilloscope at a sampling rate of 1 MHz. The behaviour of a switch during a typical key onset is reported in Fig. 3.4. This showed that the contact bounces multiple times and that the duration of each of these bounces is of the order of tens of microseconds. Each bounce brings the measured resistance from $+\infty$ (open circuit) to a small resistance value of about 10Ω (closed circuit), with very few intermediate resistance values. The "closed circuit" resistance (greater than zero) is determined by the resistance of the resistive wire connecting the tone generators to each contact.

The resistance values suggest that the dominant audible effect here is likely to be the wide range open-close discontinuity rather than the finer details of the actual resistance value when it is low. As such, we assumed that thresholding the resistance value so to only distinguish between "open" and "closed" states would not cause a significant loss of information. For each switch, a pull-up resistor circuit was used, connected to a comparator to generate a digital signal which represents the open/closed state of the contact, with a threshold of 120Ω.

3.3.1.3 *Data acquisition*

A Bela single-board computer was used to acquire the signals described above [McPherson and Zappi, 2015]. The output of the preamplifier of the optical sensor was connected to a 16-bit analog input, while the digital signals from the switches were connected to the digital inputs of the board. The Bela board sampled all the inputs at 44.1 kHz and logged them to its internal memory.

3.3.1.4 *Key presses*

To generate the data for this study we chose to repeatedly press the key with a finger, as alternative options such as the robotic finger or the pendulum come with drawbacks that would have not been suitable for this task (see Section 2.1).

In total, we collected data for about 800 key presses for each of 8 keys (E3, A♭3, C4 (= middle C), F4, A♭4, C5, E5, F5) on the upper manual of the organ. The performer used pressed and struck touches and tried to produce with the widest possible range of velocities.

3.3.2 *Results and discussion*

3.3.2.1 *Key profile*

Details of key onsets obtained with a pressed and a struck touch are shown in Fig. 3.7; the velocity and key position profiles of these two plots are representative of the respective types of touch. The pressed touch (left) starts from a null velocity which increases steadily, reaching maximum velocity just before key bottom. The struck onset (right) shows a spike in the velocity at the beginning of the onset, due to the inertia of the finger and arm which are already moving as they engage the key. The velocity of the key increases quickly to the peak value and the impact also triggers a resonant behaviour in the finger-key system. The velocity then slightly decreases during the remainder of the key press.

We expect that the portion of the key travel that has a wider influence on the behaviour of the key contacts is the one during which the key contacts close. Therefore, in order to compute a discrete onset velocity value for each onset, we chose

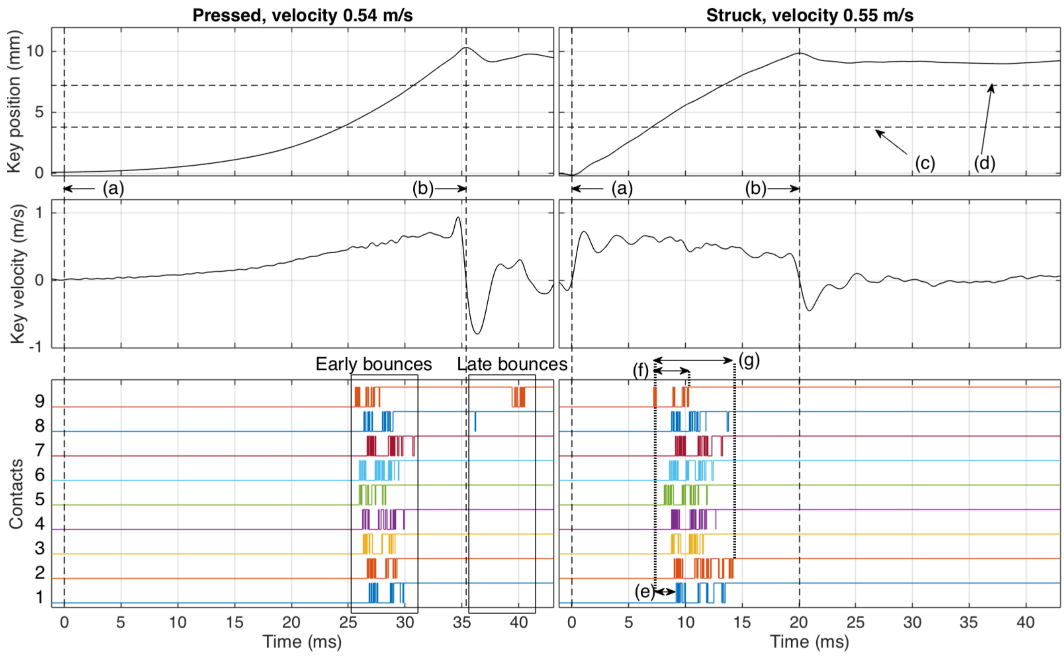


Figure 3.7: Details of the key profile and the contact state for two touches with similar onset velocities and different types of touch. Key position increases from key at rest (0 mm) to key bottom (11 mm). In the plot at the bottom, each of the nine contacts is represented with a line. When the line is low the contact is open, when the line is high the contact is closed.
(a) onset start, (b) onset end, (c) and (d) fixed points for computing discrete onset velocity, (e) contact closing offset (time range), (f) contact bounce duration for contact 9 (time range), (g) overall bounce duration (time range).

to compute the average velocity between those two points d_0 and d_1 in the key throw within which 95% of the contacts close across the whole dataset of presses for that key. The discrete onset velocity metrics displayed at the top of Fig. 3.7 and used in Fig. 3.9 are therefore computed as the average velocity of the key between d_0 and d_1:

$$v = \frac{d_1 - d_0}{t_1 - t_0} \tag{3.1}$$

where t_0 and t_1 are the times corresponding to key positions d_0 and d_1 respectively. For key F5, these points are: $d_0 = 3.7$ mm and $d_1 = 7.18$ mm, and are represented in Fig. 3.7 by lines (c) and (d) respectively.

DISCUSSION The different way in which acceleration is distributed in the two types of touch leads the two key presses in Fig. 3.7 to have a similar value of discrete onset velocity but a different duration. The duration of each key onset, measured as the time from when the key is at rest to key-bottom, is 35 ms for the pressed one and 20 ms for the struck one. The key-travel at the front of a white key is 11 mm, so by using $d_1 = 11$ mm and $d_0 = 0$ mm, Eq. (3.1) gives us average velocity values of 0.31 m/s and 0.55 m/s respectively. When measuring velocity using Eq. (3.1), two very similar values of 0.54 m/s and 0.55 m/s, respectively, are found. However, if we were to compute a discrete velocity measurement based on the final velocity (the velocity 2 mm before key bottom), we would obtain 0.93 m/s and 0.45 m/s respectively.

Comparing the continuous velocity profiles in Fig. 3.7 with those of a piano action, such as those in Goebl et al. [2005, figure 1] and McPherson and Kim [2011, figure 4], the most significant difference is that the rebounds of the finger on the key for a struck touch, visible as dips in the early part of the velocity curve, are less deep in the case of the organ. This is due to the spring-loaded action of the organ responds more quickly to changes in finger pressure than the weighted action of the piano, following the finger more closely in its rebounds.

3.3.2.2 *Contact closing distance*

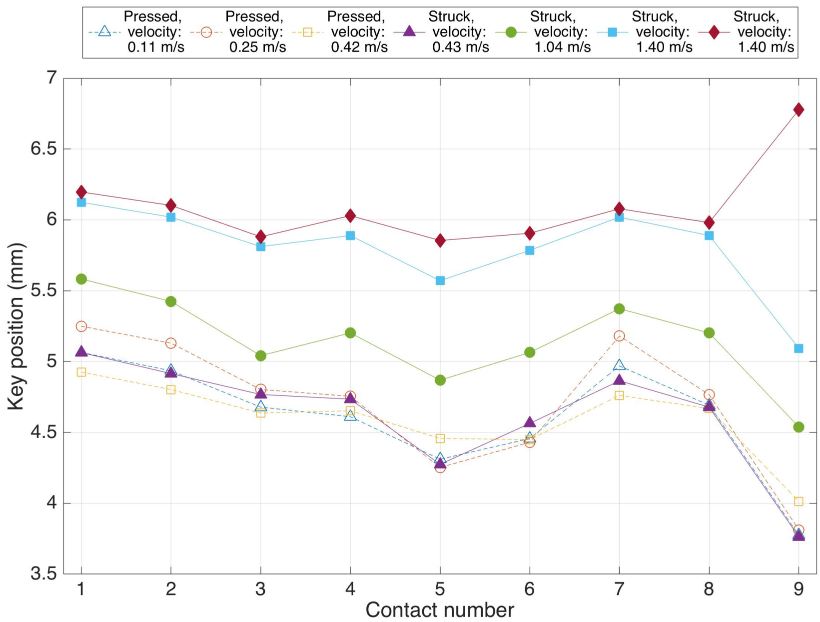


Figure 3.8: Closing distance for each contact for different key presses on a F5 key.

The instant when the spring contact first touches the busbar determines the beginning of a transient for the generator tone carried by that contact. For any given key, this does not happen at the same point in the key travel across all contacts.

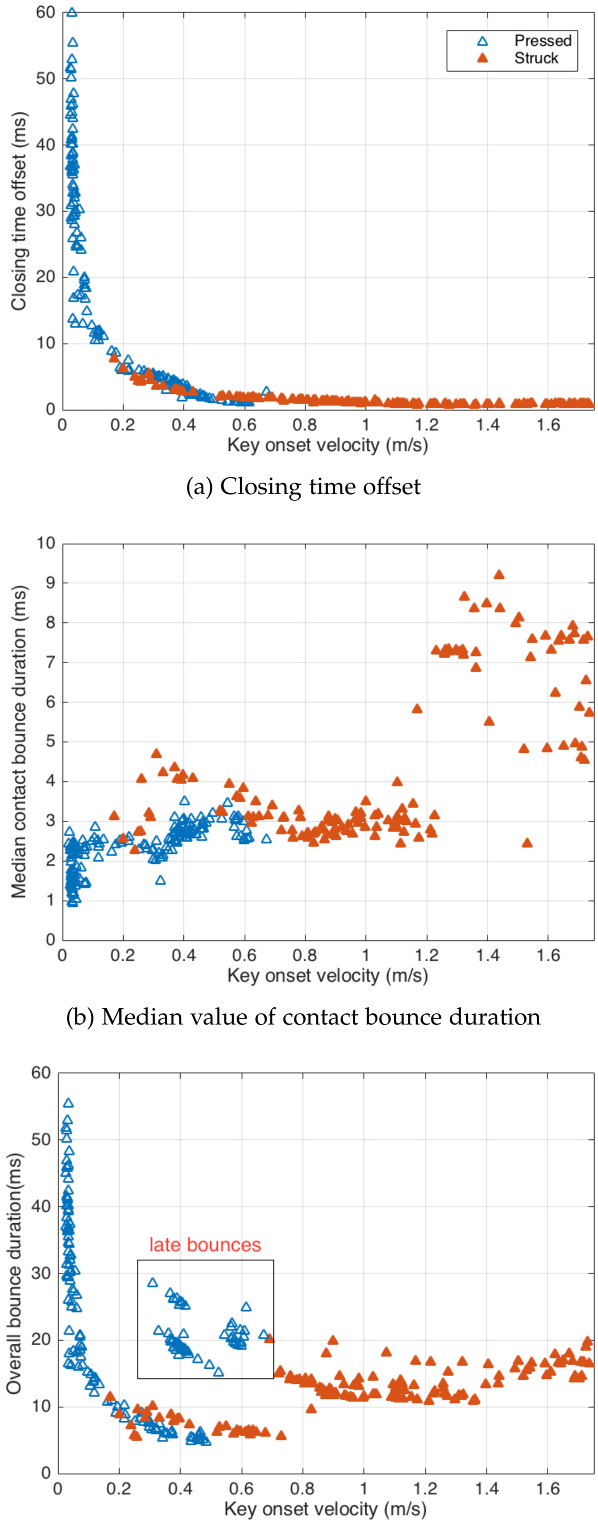
Fig. 3.8 shows the point of the key travel at which each contact of an F_5 note first touches the busbar, for different key presses. In most cases, all the contacts start making contact with the busbar within the space of about 1.5 mm from the earliest to the latest. The order in which they make contact and the spacing between them remains similar for different velocities, but there is an overall offset which is affected primarily by the velocity of the key press, with higher velocities causing the contacts to close at a later point in the key travel. Lower velocity key presses generate similar closing patterns among contacts, translated along the vertical axis as a function of velocity. For higher values of the velocity, outer contacts (1,2,8,9) occasionally break free from the pattern and close later than expected. The two struck key presses with a velocity of 1.4 m/s in Fig. 3.8 are an example of this behaviour: they share a mostly similar contact-closing pattern, but contact 9 closes much later in one than in the other.

DISCUSSION We find that contacts do not all close at the same point in the key travel, rather they close within a range of a couple of millimeters. This very characteristic is the one that allows the extended techniques mentioned in Section 3.2.4: if all the contacts were closing at the same point in the key travel, then progressive or partial key presses would not be possible. As the closing pattern changes on a key-by-key basis, partial key presses cannot be used as an alternative to drawbar registration to programmatically select which harmonics should play.

We do not have an explanation for the translation in the closing pattern depending on velocity, but we do not expect it to have an impact on the resulting sound or interaction.

3.3.2.3 *Timing properties of the onset transient*

We define the **closing time offset** as the time between the beginning of the first contact onset and the beginning of the last contact onset. It is the time that it takes for all the contacts to start producing sound. The **contact bounce duration** is the time interval during which a given contact is bouncing before it settles to the closed state. This is the duration of the onset transient for the audio signal carried by that contact. The **overall bounce duration** is the time during which at least one of the contacts of the key is bouncing, from when the earliest contact starts bouncing to when the last contact stops bouncing. This is the overall duration of



(a) Closing time offset

(b) Median value of contact bounce duration

(c) Overall bounce duration

Figure 3.9: Onset metrics for 370 distinct keys presses on a F5 key.

the onset transient for the note. In Fig. 3.9 these metrics are plotted against the average velocity of the onset.

The closing time offset in Fig. 3.9a exhibits a relation of inverse proportionality with the onset velocity, which is coherent with its definition: the higher the velocity, the shorter it takes for all the contacts to start closing. The duration of the individual contact bounces varies widely with the key velocity. The overall bounce duration, displayed in Fig. 3.9c encompasses the time interval across all contacts during which at least one contact is bouncing.

Most of the pressed onsets with higher velocity (between about 0.3 m/s and 0.67 m/s) exhibit a significantly longer overall bounce duration than struck onsets in the same velocity range. Only struck onsets with much higher velocity will reach similar values of overall bounce duration.

The median value of the bounce duration across all contacts for each key onset is displayed in Fig. 3.9b. In the range of velocities where there are both pressed and struck touches, struck touches show higher contact bounce duration than pressed ones. Only very small values of key onset velocity result in significantly smaller median contact bounce duration. Occasionally for these low velocity presses one or more contacts would exhibit no bounce at all.

DISCUSSION While recording the key strokes shown in Fig. 3.9, the player tried to cover the entire onset velocity range for both pressed and struck touches. However, only in the region between 0.17 m/s and 0.67 m/s did they manage to produce both pressed and struck touches. Velocity values below this range were only achieved through pressed touches and values above only through struck touches, suggesting that struck touches allow the production of higher velocity values, confirming similar findings on the piano keyboard in [Goebl et al., 2005].

The fixed points we chose to compute the discrete onset velocity value were chosen as those within which the contacts are more likely to close (see Section 3.3.2.1). Given the non-uniform distribution of the acceleration and the differences between pressed and struck key profiles, choosing different points would change the shape of the plots in Fig. 3.9, mainly affecting the overlap on the velocity axis between pressed and struck key presses.

Upon close inspection of the behaviour over time of bounces of individual contacts, we observed that an onset bouncing is characterised by a first part, which we call “early bounces”, usually less than 5 ms long, during which the contact quickly alternates between the open and closed position. After the early bounces, the contact is pushed against the busbars and stays closed. For some of the key presses, "late bounces" can be observed 3 ms or more after the end of the early bounces. The presence of late bounces on one or more contacts may increase the overall bounce duration. In Fig. 3.7 we labelled early bounces and late bounces in the time-domain representation of contact bounces, while in Fig. 3.9c) we highlighted those touches for which the overall contact bounce duration is affected by late bounces.

We find that late bounces are correlated with the rebound of the key after key bottom, which causes some of the contacts to be temporarily released from the busbar and left free to bounce again. Late bounces are usually less dense than early bounces but could last longer, depending on the final key velocity. Moreover, late bounces are more likely to occur on those contacts that are at the outer ends of the contact pusher (contacts 1,2,8,9). The rebound of the key is ultimately affected by the key velocity at the moment when the key hits the keybed (the final key velocity), but a high final key velocity does not deterministically produce late bounces. Rather, what we observe is that the *likelihood* of late bounces is higher for presses with higher final key velocity. This can be explained by the fact that the key rebounding on the keybed and the contact pusher form a dual-pole resonant system and therefore the phases of the two oscillations will determine whether the contacts are released from the busbar and produce late bounces.

As observed earlier, pressed touches have a steadily increasing key velocity, while the velocity of struck touches starts with a spike and then slowly decreases. Therefore, for two key presses with the same value of key velocity onset, as measured by us between two fixed points, the final key velocity will be higher for a pressed touch than for a struck touch, and the former will be more likely to exhibit late bounces.

Key presses with similar values of overall bounce duration may exhibit widely different contact behaviours, according to the distribution of velocity along the key throw. For instance, the dominant factor on a onset with high overall bounce duration may be a long contact closing time offset due to a slow velocity, or - alternatively - a single contact exhibiting late bounces, due to a high final velocity.

To summarise, we find three primary phenomena which could affect the character of a note onset. The contact closing offset is caused by the fact that contacts do not all make contact with the busbar at the same point in the key travel. This is directly related to the onset velocity, and is shorter for higher velocities. The duration of the early bounces is determined by the velocity of the spring contact when they hit the busbar: the higher the velocity, the longer they will bounce before settling in the closed state. Late bounces may occur if the final velocity is high enough to cause a rebound of the key.

3.3.2.4 *Variability between keys*

On the Hammond organ the key action is the same across octaves, so that, unlike in the case of the piano, there is no expected systematic variation between different registers, although there may be some variation due to manufacturing tolerance.

There is no systematic way of adjusting the vertical position of individual busbars, thus affecting the triggering point for a given contact across the keyboard. However, vertical offset of the drawbars due to manufacturing tolerances or bent drawbars can affect the triggering point for a given contact systematically across keys. Superimposed to this offset there are any additional local variations due to key felt, contact pusher and individual contacts.

For the eight keys we measured on the upper manual of the organ, contact 9 is always the first one to close in the key-throw and contact 1 is always the last one. We then measured the contact closing distance on five keys on the lower manual and found that contact 6 is always the first to close and contact 1 is always the last. These findings suggests the possibility that each manual of each instrument may have a distinctive contact closing distance pattern, with additional variations due to each key.

Results in Fig. 3.9 are for key F5 on the upper manual; all of the other white keys we measured show a similar overall behaviour, matching our observations earlier in this section. As for the two black keys we tested, they do not show the phenomenon of late bounces. Comparing the observed final velocity values, we find that the upper limit is around 1 m/s for pressed touches on the white keys and around 0.6 m/s on the black keys. The difference in the observed final velocities for the two types of keys can be explained in terms of the different key-throw (10 mm for the white keys, 6 mm for the white keys): given that in pressed touches the velocity of the key tends to increase during the press (see Fig. 3.7), the final velocity value will tend to be smaller if the overall distance is smaller. We showed in Section 3.3.2.3 that the presence of late bounces is associated with high final velocity values; the lack of late bounces in the pressed touches on black keys can therefore be ascribed to the smaller final velocity values obtainable on the black keys.

The $A\flat_2$ key also showed another singular behaviour, in that contact 1 tends to bounce for long periods for struck touches of velocity comprised between 0.5 m/s and 0.9 m/s. This in turn causes a higher overall bounce duration in this velocity range for this type of touch, in the 10-60 ms range, while most other keys would have figures below 10 ms under similar conditions.

3.3.3 *Implications*

We show how the continuous evolution of the key position affects the behaviour of the key contacts. The early part of the bounces is conditioned by the velocity when the contacts engage the busbar, while the late part of the bounces depends on the velocity just before the impact with the keybed. The contact closing behaviour displays a complex velocity-dependent pattern of asynchrony and bouncing which is dependent on velocity, but the velocity measured at different points in the key travel affects different aspects of the bouncing. Therefore, a single scalar velocity measurement is not enough to represent the multidimensionality of different types of touch, which distribute the velocity differently along the key travel.

3.4 LISTENING TEST

The mechanical behaviour of the keyboard contacts demonstrates that different aspects of the onset transient on the Hammond organ are affected by the velocity and the type of touch in use. However, these changes only cover a small period of several milliseconds at the beginning of each note, while the sound of the sustained part of the note is not correlated with the transient stage and will always exhibit a consistent behaviour regardless of the gesture used to produce the note, as long as the key is fully depressed. We therefore set out to determine whether these changes in the transient were audible and whether they could be reliably associated with the particular type of touch that produces them.

A listening test was designed to validate or reject the following null hypotheses with regard to notes played on the Hammond organ:

1. listeners are unable to distinguish between notes played with different types of touch and velocity
2. listeners cannot distinguish what touch was used to produce a given sound
3. the accuracy of listeners performing the two tasks above is independent of the level of familiarity with the Hammond organ

The test was designed in such a way that it was possible to undertake it locally, under the direct supervision of one of the authors, or remotely, over the internet, using a bespoke online service based on the Web Audio Evaluation Tool [Jillings et al., 2016].

3.4.1 *Experiment design*

3.4.1.1 *Stimuli*

A dataset of over 2000 key presses was recorded from the monophonic line output of a 1967 C-3 Hammond organ. The Hammond is an electromechanical instrument and is always played through a loudspeaker or recorded via a line output, therefore we disregarded any acoustic recording of the finger and key noise.

We recorded the sound outputs generated by a total of eight different keys spanning the whole range of the keyboard, namely C_1, $A\flat_1$, E_2, C_3, $A\flat_3$, E_4, C_5, $A\flat_5$. The tone produced by each key press was at least 2 seconds long, but it was faded out with a logarithmic fade of duration 0.5 s, starting 1 s after the beginning of the onset transient, so that the release transient was not included in the stimulus. Using the same sensing circuit described in Section 3.3.1.1, a discrete average velocity value was computed for each of the recorded key press. The notes were played with all the drawbars pulled out. The signal from the line output of the organ was recorded with a sampling rate of 44.1 kHz and a bit depth of 24 bit using a Motu 828 Mk-III soundcard.

Four combinations of velocity ranges and touch ("touch classes") were chosen for the test (slow-pressed, fast-pressed, slow-struck, fast-struck). Stimuli were selected which had a velocity value of 0.2 ± 0.05 m/s, 0.45 ± 0.05 m/s, 0.7 ± 0.1 m/s, 1.4 ± 0.1 m/s respectively. With this method, a total of 64 unique stimuli were selected: 8 stimuli per key, 2 for each combination of slow/fast and pressed/struck.

3.4.1.2 *Structure*

The test consisted of a training section followed by four test sections, one of which was an A/B/X test and three of which involved A/B tests. Before the test the subject had to go through a short survey asking about their familiarity with the Hammond organ. An optional survey at the end allowed us to gather feedback from the participants.

Touch classes	Question
slow-pressed, fast-pressed	*Which of these pressed notes has a FASTER velocity?*
slow-struck, fast-struck	*Which of these pressed notes has a FASTER velocity?*
pressed, struck	*Which of these notes was played with a STRUCK touch?*

Table 3.2: Touch classes and questions for each A/B section

The A/B/X section consisted of 72 trials. For each trial, participants were asked to listen to three stimuli, all generated from the same note, labelled A, B and X. They could listen to each stimulus as many times as they liked. Stimuli A and B belonged each to one of the four touch classes and the class of A was always different from the class of B. The X stimulus was a duplicate of one of A or B. Participants had to select the stimulus that better answered the question *"Which of these sounds matches the reference X?"*

Each A/B section consisted of 24 trials. For each trial, participants had to listen to two stimuli, labelled A and B, which were generated from the same note. They could listen to each stimulus as many times as they liked. For each section, stimuli were selected which belonged to two touch classes and the participant was informed what these classes were. In each trial, then, each class would be represented by exactly one stimulus. The same question was asked for each trial throughout a section. After listening to the stimuli, the participant had to select the stimulus that better answered the question. The touch classes used and the question asked in each section are summarised in Table 3.2.

Each set of 24 stimuli for each category consisted of 3 pairs of stimuli for each of the 8 notes. Within the 3 A/B pairs for each note, 2 of them consisted of the same pair of recordings. All the 48 unique pairs from the A/B tests were collated together and used in the A/B/X trials. Additionally, 24 of these pairs were presented twice in the 72 A/B/X trials.

The A/B/X section would always be the first section in the test, immediately after the training, followed by the three A/B sections. For each participant, the following variables were randomised: the order in which the three A/B sections were presented, the order in which the trials were presented within each section, the A/B labels assigned to each stimulus, the stimulus labelled X in each of the A/B/X trials.

3.4.1.3 *Training*

Some basic training was given prior to the test in order for the participant to understand the basics of the effect of touch on the Hammond organ.

A brief video demonstrated visually and aurally the difference in the physical action between a pressed and a struck note[13] . The aim of the video example was to help the participant get a better understanding of the physical action associated with the sound, hopefully helping them to create a stronger link between the type of key press and the associated sound.

All participants were then presented a set of training stimuli which included one example stimulus for each of the touch classes used in the remainder of the test (slow-pressed, fast-pressed, slow-struck, fast-struck) for each of 3 notes (A♭2, C4, E5). Participants could listen to each stimulus as many times as they wanted.

The participant was forced to go through the training once at the beginning, but they were then allowed to go back to it at any later time during the test.

3.4.2 *Results*

A total of 50 participants completed at least one of the sections of the test and 46 of these completed all of the sections. 27 participants undertook the test locally at our research facilities, while the remaining 23 did it remotely online. The local participants were recruited among the postgraduate students at Queen Mary University of London, age range = 25-39. The age data was not collected for online participants. Most (42) of the 50 participants had experience playing an instrument, and 37 of these had played their main instrument for more than 5 years, 6 of them at a professional or semi-professional level. The listening test was approved by the Queen Mary University of London Ethics of Research Committee, with approval code QMREC1691a. and followed the institution's guidelines in participant data collection. Local participants were provided with a set of Bose QT-25 headphones, while remote participants were encouraged to use headphones for the test.

13 "The effect of touch on the Hammond organ": `https://youtu.be/kcRtyvQxbgU`

	A/B/X	A/B pressed: slow/-fast	A/B struck: slow/-fast	A/B pressed/struck
pessimistic	46(0)	45(2)	34(20)	23(2)
optimistic	50(0)	45(0)	32(11)	34(0)
consistency	44	41	28	29
participants	50	47	46	48

Table 3.3: Results of the listening test. For each test condition, we report the number of participants for whom we can reject the null hypothesis with $p < 0.05$. Numbers in parentheses indicate how many of these had reversed the labelling.

We performed a statistical analysis on the results of the listening test. To test the first of our hypotheses, that listeners are unable to distinguish between notes played with different types of touch and velocity, we used the results of the A/B/X test. The second hypothesis, that listeners cannot distinguish what touch was used to produce a given sound, was tested under three different conditions with the A/B tests. The data from all the test conditions combined with the self-reported familiarity of the participant with the Hammond organ were used to evaluate the third hypothesis, that the performance obtained by listeners at these tasks is independent from their familiarity with the instrument.

Each A/B/X and A/B trial can be considered a Bernoulli trial, with a probability $p = 0.5$. Assuming the trials are independent, we can analyse the collection of results using null-hypothesis statistical tests under the binomial distribution [Boley and Lester, 2009]. Duplicate trials are, by definition, not independent, and have to be removed before the analysis. In order to remove duplicates, we used a dual optimistic/pessimistic approach. We reduced each pair of duplicated trials (which yielded outcomes r_1, r_2) to a single trial, of outcome r_d. In the case of the *optimistic* approach, "success" if the outcome of *at least* one of two trials was "success", "failure" otherwise: $r_d = (r_1 \text{ OR } r_2)$. In the case of the *pessimistic* approach, the outcome of *both* trials was "success", "failure" otherwise: $r_d = (r_1 \text{ AND } r_2)$

In this test the listener was asked to make judgements on sounds - actually a specific characteristic of the sound - that they potentially never heard before, or had never considered to such level of detail. Despite the training provided, it is reasonable to expect that some of the participants may have learned the labels wrong and/or reversed their decision criteria during the test. We therefore took into account the reversal effect [Boley and Lester, 2009] if a participant is able to label sounds belonging to a given class of touch in a consistent way, it means that they can discriminate between the classes, regardless of the fact that the labelling itself is correct or wrong.

In the context of a binomial distribution, the cumulative distribution function gives the probability that a certain number of "success" outcomes from a number of Bernoulli trial is not the result of randomised answers [Boley and Lester, 2009]. We used a minimum of 95% confidence level, so that when a participant has cumulative binomial probability above 95% or below 5% (accounting for the reversal effect) under a given test condition, this indicates a relevant perceptual difference. Results of the test are summarised in Table 3.3.

For each of the test conditions we also tested for self-consistency of the listeners, leveraging the duplicated trials. In order to do so, we considered each pair of duplicate trials as a single consistency trial whose outcome is "success" if the outcomes of the two trials are the same or "failure" if the outcomes of the two trials differ, thus making the consistency trial a Bernoulli trial. We then computed the cumulative distribution function for each test with a threshold of $p < 0.05$. The number of participants who passed the consistency test are in Table 3.3 Out of the 46 participants who completed all of the four tests, only 1 was not consistent in any of the tests, 3 were consistent in one test only, 12 in two tests, 11 in three tests and 19 in all the four tests.

Participants were asked to report their familiarity with the sound of the Hammond organ and with the technical aspects of the instrument, each as a numerical rating on a scale between 1 and 5. By averaging together the numerical ratings from the two questions, we grouped participants in three groups, according to their familiarity with the instrument: "low" (average $\leqslant 2$), "mid" ($2 <$ average $\leqslant 4$), "high" (average > 4). In a different analysis, we split the participants in two separate groups between those who had a significant experience in playing the instrument and those who never or almost never played it, on the basis of their answer to a dedicated question in the survey. The results, expressed as relative number of correct outcomes, for each of these groupings are in Table 3.4.

In all the tests, participants were allowed to listen to the samples as many times as they liked and in any order. The samples in each pair were taken from the same note and with the same drawbars setting and the same volume, therefore there is no difference in loudness levels between the two notes, which rules out possible

Familiarity	A/B/X	A/B pressed: slow/fast	A/B struck: slow/fast	A/B pressed/ struck
low(14)	85.7%	85.7%	**78.6**%	35.7%
mid(17)	**100.0**%	**100.0**%	64.7%	**58.8**%
high(19)	89.5%	84.2%	63.2%	42.1%
not played(32)	90.6%	**90.6**%	**71.9**%	**50.0**%
played(18)	**94.4**%	88.9%	61.1%	38.9%

Table 3.4: Percentage of subjects in each group who passed the test, according to the pessimistic evaluation. Within parentheses is the number of participants in each group.

Notes	A/B/X:	A/B pressed: slow/fast	A/B struck: slow/fast	A/B pressed/ struck
C4	189	111	115	86
A♭6	194	116	111	78
A♭2	191	130	114	92
A♭4	200	135	109	97
C2	198	129	83	85
C6	194	139	109	115
E3	196	136	93	101
E5	191	135	89	99
Trials per note:	225	141	138	144

Table 3.5: Number of successful trials for individual notes in each test.

effects of forward masking. Chi-squared tests did not show any significant effect of the presentation order for the A/B/X test [percent correct: 86.1%, $\chi^2(2) = 1.04$], the A/B pressed test [percent correct: 91.4%, $\chi^2(2) = 0.74$] or the A/B struck test [percent correct: 74.55%, $\chi^2(2) = 0.87$. However, the chi-squared value shows an apparent effect of the presentation order in the A/B struck/pressed test, [percent correct: 65.4%, $\chi^2(2) = 5.42, p < 0.02$], where participants tended to select "A" when the correct response was "B" more often (229 times out of 607, 37.8%) than they would select "B" when the correct response was "A" (170 times out of 545, 31.2%).

Results in Table 3.4 show that there is no clear difference in the results achieved by participants with different levels of familiarity with the instrument. Of the participants with a high familiarity with the instrument, only 2 took the test locally. A chi-squared analysis could not find any significant difference in the overall performance of those who attended the test locally and those who attended it remotely [$\chi^2(2) = 0.042$], so there seems to be no effect due to non-uniform testing conditions.

A breakdown of the number of correct outcomes for each note is shown in Table 3.5. A chi-squared independence test shows no effect across notes for the A/B/X test, however it found that for each of the A/B tests there is a significant difference between the notes ($p < 0.001$).

3.4.3 *Discussion*

From the outcome of the A/B/X test, we conclude that for at least 46 out of 50 participants there is a perceptually relevant difference between stimuli produced with different types of touch and velocity. This outcome is, by itself, enough to reject our first null hypothesis. 40 participants went through the section in less than 25 minutes and only 3 took more than 40 minutes, but there is no way to control whether they had breaks during the session.

Most participants were also able to reliably distinguish between the slow-pressed and fast-pressed touches. This test was the one on which participants spent the least amount of time (about 2 minutes and 10 seconds on average), while the other A/B sections took about 3 minutes each. The comments of 6 participants specifically refer to this task as being the easiest one, with one participant mentioning that in the slower presses *"it was almost possible to hear harmonics coming in"*.

There were fewer successful cases in the slow-struck/fast-struck A/B test, yet at least 32 participants could distinguish between the two. A large number of reversal effect cases was registered in this particular section, suggesting that, while participants were able to consistently distinguish between transients, they struggled to remember which class a specific type of transient belonged to.

We found in Section 3.3.2.4 that key A♭4 shows a behaviour different from all the other keys under test in that the struck touches with lower velocity (slow struck) show larger overall bounce duration than the other keys. The A♭4 has the lowest number of correct trials in the A/B struck slow/fast test, suggesting that the unusual behaviour of the key makes the identification task harder for the listener.

The pressed/struck A/B section is the one with the lowest number of successes. This test condition is the only one that presents stimuli from all four touch classes. The test set therefore contains a more diverse set of samples, though the participant is still asked to chose between two (pressed and struck). A recurring comment, mentioned by 9 participants, is that differentiating between the three cases of fast-pressed, slow-struck and fast-struck was very difficult, which is in line with the results for this test condition.

In all the tests, according to both the pessimistic and the optimistic evaluation, the number of participants for whom we can conclude there is a relevant perceptual difference is well above 5%, thus suggesting that the statistically relevant result is that listeners are able to classify notes played on the Hammond organ according to the type of gesture used to generate them, thus rejecting our second null hypothesis.

While differences due to familiarity are often small, there is no clear evidence that people familiar with the Hammond organ performed better at this test, thus confirming our third null hypothesis. This finding suggests that the sonic differences due to the touch may not be trivial to tell apart for Hammond organ players, or at least not when they are decoupled from the actual gesture which produces them.

3.5 CONCLUSION

3.5.1 *Findings*

In this chapter we analysed the details of note onsets on the Hammond organ. The instrument does not show a relation between key velocity and the amplitude of the produced sound. Yet, the way a key is pressed affects the onset transient of the note.

A first study shows that the behaviour of the key contacts, which are ultimately responsible for generating the onset transient, is affected by the continuous key position and key velocity during the key press. The velocity measured around the points in the key travel where key contacts close affects the spread over time of the contact transients and the duration of the early part of each contact's bounces. The velocity measured immediately before key-bottom affects the probability that late bounces appear due to the rebound of the key on the keybed. Pressed and struck touches show two clearly distinct velocity profiles over time, which means that the above measurements are both needed and one cannot be inferred from the other without previous knowledge of the type of touch in use. Additionally, the instantaneous position of the key is also relevant, as it ultimately determines which of the key contacts are active at any point time.

Our second study, a listening test, shows that combinations of different types of touch and velocities produce different sounds, and that these can be perceived as such by the listener. A statistically significant number of our test subjects managed to classify a set of recordings according to the touch and velocity used to produce them. This indicates that the key gesture has a perceivable effect on the onset transient of the generated sound.

These observations make the instrument not only touch-responsive, but they make it so in such a way that cannot be captured with traditional velocity-based keyboard sensing. Reducing the richness of the key gesture to a single velocity parameter causes a loss of information, losing details on the original intent of the player and making it impossible to fully describe the sonic outcome. Most digital emulations of the Hammond organ do not allow the player to control the sound generator with the continuous position of the key, as they mostly use regular MIDI keyboard controllers with a single discrete velocity measurement.

This dimensionality reduction is similar, in a certain sense, to the one that occurred on pipe organs when electro pneumatic valves were introduced to replace direct control of air flow. On organs with direct control, the player retains a certain degree of control on the shaping of the transient onsets which improves phrasing and articulation, but is lost when the key acts as an electronic switch controlling the valve [Le Caine, 1955]. For both the pipe organ and the Hammond, the control at the musician's fingertips is subtle and not such that it allows to the player to vary the loudness of the produced sound. Yet, in both cases, players tend to react negatively to limitations imposed by simplifications in the response of the keyboard action enforced by the advent of a new technology.

3.5.2 *Recommendations*

Our findings suggest some general recommendations for creators of digital emulations of the Hammond organ, in order to replicate the amount of control available on the original instrument. Keyboard controllers based on switches are very common; these compute a velocity parameter from the time interval between the closing of two switches placed at different points in the key throw. Most controllers have two contacts per key, but keyboards with three contacts per key have recently surfaced on the market. The position of the contacts along the key-throw is critical, and there is a trade-off between their position and the velocity metrics that can be obtained. Being able to access the two contacts individually, and trigger separate harmonics on each contact, has the potential to give better results than using a single triggering point.

A keyboard controller that provides continuous tracking of the key-position, such as the one described in [McPherson, 2013], is best suited to capture all the subtleties of the gesture on the key, and it is the only way to accurately simulate the effect of very slow key presses, which cannot be be tracked accurately with traditional technologies based on two or three discrete switches.

3.5.3 *Limitations*

In the two studies presented in this chapter we focussed on individual key presses, on a single instrument, in a non-performance context. The action of tonewheel Hammond organs has not changed much over the years, so we would expect to find similar results on different instruments from different years, but a comparative study is required to ascertain this. Extensions of the listening test we performed would also investigate how different combinations of drawbars and partial key presses would impact the final result. We did not specifically perform audio analysis of the onset transient, but rather we inferred some of its characteristics and its perceptual significance from the two studies.

3.5.4 *Implications*

Whether in a performance context the touch-responsiveness of the instrument still makes a difference to the listener remains an open question. Le Caine [1955], Moog [1987] report that the playing of pipe organ players is affected at a macroscopic level by a small amount of touch-sensitivity, so it is not unreasonable to expect similar results for the Hammond. In the study we present in Chapter 4, we investigate the effect on the player of subtle changes to the amount and type of control available during the onset of the notes while playing a Hammond emulator.

4 POSITION-BASED CONTROL OF ONSETS ON A HAMMOND EMULATION

Some of the concepts from this chapter are summarised in [Moro and McPherson, 2020].

From our work in Chapter 3 it emerged that on the Hammond the touch and velocity on a key press can cause audible effects on the initial transient of each note, whose duration is typically between 5 ms and 60 ms, because of the presence of nine contacts that close at different times during the key throw. We are now interested in studying if variations of the subtle amount and type of control on the onset transient can be appreciated by players. An instrument which uses continuous key position paired with a software emulation can be used to digitally recreate the multiple contacts of the Hammond, and can be programmed to provide several different key responses, ranging from those typical of MIDI keyboards to that of the real Hammond. In this chapter we describe one such instrument we built and a performance study with ten players comparing different touch conditions on the instrument.

We start by laying out the motivations of our work in Section 4.1, followed by an overall description of the instrument (Section 4.2). In Section 4.3 we then describe the development of an embedded platform for continuous keyboard sensing and sound generation. In Section 4.4 we describe our modifications to an existing Hammond emulator sound engine in order to allow continuous control and the dynamic generation of contact bounces. Section 4.5 briefly describes the process of combining the platform and the digital sound generator with the physical organ. Section 4.6 and Section 4.7 cover the design and the results of the performance study, respectively, which we discuss in Section 4.8.

4.1 MOTIVATION

Our study in Chapter 3 considered only listeners and focused on isolated single notes, which is not representative of how players encounter the instrument, where correspondences between tactile and audio feedback shape the experience, and where control of individual note onsets is situated within a broader musical context. A suggestion that these differences can be relevant also for players can be found by observing discussions on a long-standing issue among Hammond aficionados, that of whether the digital emulators can get close to emulating the "real thing". As an example, we extracted some relevant quotes from a conversation which took place in the February of 2019 on the "Hammond tonewheel organ clones forum" (aka "CloneWheel"), one of several online discussion groups on the topic.[1] The conversation is started by a user who asks for opinions over the differences between two emulators: the Hammond XK-5 and the Crumar Mojo, which we reviewed in Section 3.2.5. The XK-5 has three physical switches per key, and to each switch the user can assign zero or more of the virtual switches of the key in the sound generator. This technology can deliver an experience somewhere in between the full 9-contact stack of tonewheel Hammonds, and the single triggering point of virtually every other clone (including the Mojo) or software-based virtual instruments.

A first user answered: *"The sound of the XK-5 and the Mojo are both excellent.. if you asked 100 players in a blind listening test which sounds they preferred... you'd probably get a 50/50 split or somewhere close. [...] The virtual key contact scheme on the XK-5 is theoretically good, but comments from players range from 'best thing since sliced bread' to 'didn't notice any difference'... [...] Obviously some people prefer the authentic style interface of the XK-5, but with that goes a lot of weight [... and cost]"*.[2]

Another user added shortly thereafter: *"I always find it interesting that we all hear the same things and interpret them so differently! [...] it is safe to say the current crop of Clonewheel organs [...] is pretty excellent across the board [...] I love the XK-5's 3 contact action – it gives you the option to use the 3 contacts or the highest or lowest – I think it makes a world of difference to use the three and I didn't think it would!"*[3]

A third user replied: *"at the hands of a proficient player you can sense that the triggering off the XK5 is a lot more intuitive than a clone without the keystack. [...] play-ability is*

1 The original conversation started on the 9th of February, 2019, and can be retrieved from the group's homepage (`https://groups.io/g/clonewheel/`). Quotes reproduced here with permission from the original authors.

2 `https://web.archive.org/web/20200412225624/https://www.csrhome.online/CWSG/2019/Feb/msgidx_84967.html`

3 `https://web.archive.org/web/20200413072850/https://www.csrhome.online/CWSG/2019/Feb/msgidx_84970.html`

something that was left out at the design stage mostly because of cost [...] on clones before this, and is important enough to be missed, imo" .[4]

A fourth user commented: *"[The XK-5] was a definite improvement that I think will be most appreciated by players familiar with playing the real thing. [...] [on the XK-3c, an older clone] there is definitely a different feel when playing the real thing. The XK-5 was a definite step in the right direction. [...] Things like the intro to Deep Purples 'Hush' intro sounded much more authentic having more of that texture that seems lacking in clones. [...] I also like the XK-5 having 4 sets of drawbars like the real thing"*.[5]

These are brief extracts from a longer discussion thread, and, while not exhaustive, they highlight themes that, in the author's experience, are some of the most recurring ones in this and other mailing lists, when discussing the latest Hammond clones:

THE SOUND IS GREAT the sound of the latest generation of emulators is very good and very close to that of a tonewheel Hammond, so that even an experienced listener would have a hard time telling them apart.

THE INTERFACE MATTERS the SK-5 can command a significantly higher price, not only because of the brand it bears or the sensing technology, but also because all the additional controls in the interface (e.g.: to control presets, chorus/vibrato, percussion, etc) are very similar to those of a classic tonewheel console.

THE EXPERIENCE OF PLAYING WITH THREE CONTACTS PER KEY IS SUBJECTIVE The XK-5's signature feature is to have 3 physical contacts per key, to each of which the user can arbitrarily assign the 9 virtual contacts of the sound generator. It can make a lot of difference, or no difference, depending on the player. More experienced players seem to find it more natural, more similar to a tonewheel Hammond, and are deemed more capable of appreciating the differences.

What emerges from these observations, is that the sound of the instrument is not necessarily the most important element when evaluating a clonewheel, given also how there are many with an excellent sound, and that the familiarity of the interface is also very valuable. The opinions on the effect of the three contacts per key, and its relevance, or lack thereof, to some players, is what is particularly interesting to us.

4 https://web.archive.org/web/20200412225451/https://www.csrhome.online/CWSG/2019/Feb/msgidx_84975.html

5 https://web.archive.org/web/20200413072856/https://www.csrhome.online/CWSG/2019/Feb/msgidx_84976.html

These anecdotes suggest that there is space for exploring the role of key onset when this is no longer a single event, but a rapid succession of multiple discrete events in a small space. At a macroscopic level, the behaviour of the instrument is not different, and the cognitive approach taken by the player does not need to change: in both cases they will often just press the key all the way down, and they can still think of the action of attacking the note as a discrete event. The fact that it has more than one contact becomes, in a sense, an *implementation detail* of an otherwise discrete onset, that the player *need not* worry about in order to obtain a sound. Some players do, however, report differences, but whether that is result of Hammond's marketing hype, the player's self-suggestion, or an actual perceived difference, cannot be concluded from the above.

In this chapter we investigate the effect of variation in the key responses that range from discrete (single-triggering point) to quasi-discrete (several triggering point, but similar macroscopic behaviour), which result in small differences in the onset transient of the note. The case of the Hammond is a compelling one, both because of the examples above that testimony the relevance of the topic, and because our earlier findings show that audible effects of the timescales involved are indeed perceivable. We therefore created a digital emulation of the Hammond organ which can be controlled with continuous key position. In our instrument, we can customise the mapping between the key position and the closing of each of the nine virtual contact switches in the sound generator, thus arbitrarily changing the amount and type of control on the key onset available at the player's fingertips.

4.2 THE HYBRID HAMMOND EMULATOR

This study requires a sound engine capable of simulating the sound of a Hammond organ which can respond, in real time and with low latency, to a device capable of scanning the continuous position of the keys on the organ keyboard. A selectable mapping layer will map the continuous key position data to triggering individual contacts of the Hammond simulator, allowing to simulate several keyboard responses. Additionally, a complete log of the sensor data and the audio should be recorded for further analysis. No commercially available device or sound generator exists that satisfies the requirements above, therefore we set out to create a bespoke combination of hardware and software.

In order to maximize the realism of the simulation, we re-used as much as we could of the real Hammond keyboard, interface and electronics. We modified a working Hammond C-3 console (the same as used in Chapter 3) and built a hybrid digital-analog instrument around it. The high-level block diagram of our instrument is in Fig. 4.1. The performer plays the keyboard on the organ, but the internal link between the drawbars and the preamplifer circuit is disconnected. An optical scanner is fitted to the organ keyboard, and its readings are sent to a single-board computer (Bela[6]), which also generates the sound. The sound generator is based on the `setBfree`[7] open-source Hammond emulator, which we improved for our purposes, as we describe in Section 4.4. `setBFree` is capable of performing a full simulation of the sound generation mechanism on the organ, including a virtual tonewheel generator, virtual contacts and virtual drawbars, scanner vibrato/chorus, preamplifier, as well as a rotary speaker, with MIDI input controls for keyboards and expression pedal. However, in order to maximize the realism of the simulation, we output the signal before the preamplifier simulation and we feed it back into the Hammond's actual preamplifier, at the point where the signal coming from the drawbars would normally be connected. This way the player will hear the sound after it passes through the tube preamplifier and they can use the organ's expression pedal, connected to preamplifier itself, as in a regular Hammond performance.

The next three sections document the technical effort that was required in order to build this instrument.

4.3 A platform for continuous keyboard sensing and embedded sound generation

The first step towards building this instrument was to build an embedded hardware and software environment for low-latency sound generation controlled by continuous key position. Our implementation builds on a keyboard scanner to detect key position and a Bela embedded computer board for handling sound processing.

Combining them was not possible with the out-of-the-box capabilities of the two devices, so we developed a real-time communication between the devices in order to be able to control the audio processing with the scanner data. The keyboard scanner provides the filtered raw readings from its sensors, however these are not linearly related to the position of the key, because of the nature of the sensing

6 https://bela.io
7 https://github.com/pantherb/setBfree

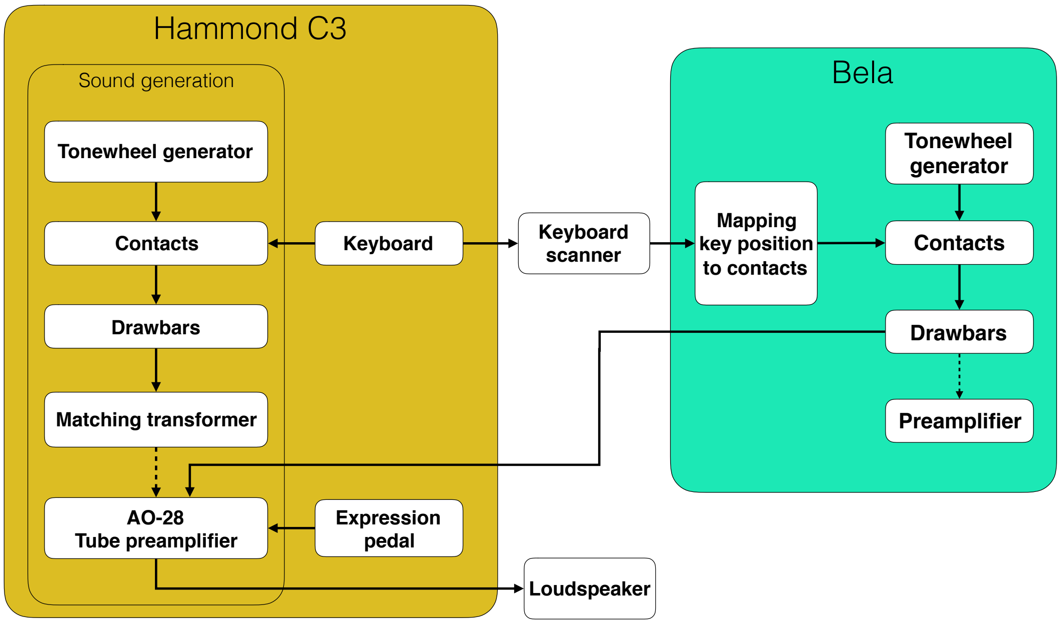


Figure 4.1: Block diagram of the hybrid analog-digital Hammond emulator. The dashed line denotes an internal connection of the C3 organ that is disconnected so that only the signal from Bela is processed through the preamplifier.

technology employed, as explained in Section 3.3.1.1. We therefore developed a linearisation algorithm and associated calibration procedure.

4.3.1 *Technologies used*

4.3.1.1 *Keyboard scanner*

The keyboard scanner [McPherson, 2013] uses infrared sensors to detect the distance of the key from its rest position. It is installed vertically on top of the keyboard, at the far back, causing minimal disturbance to the player. The position of each key is scanned with a sampling frequency of 1 kHz. The scanner is composed of up to 4 boards, each of which can sense up to 25 keys. There are three sizes of boards available: 25 keys (C to C), 24 keys (C to B), 15 keys (A to B). A typical configuration uses four boards of 15, 24, 24, 25 keys respectively, covering the whole 88 keys of a grand piano.

An STM32F103 ARM Cortex-M3 microcontroller on each board handles data acquisition, inter-board communication and communication to a host computer. The boards communicate with each other over a Serial Peripheral Interface bus (SPI), with the lower boards acting as SPI slaves, and the topmost board acting as the SPI master while at the same time handling the USB connection to the host. Each board samples its sensors every 1 ms, using an internal clock counter. Every 1 ms, the SPI master requests data from the slave boards. Every 5 ms, the SPI master sends a synchronization message to all the boards to avoid clock drifts over extended periods of time. The master board can also transmit data to a host PC through a virtual serial connection over USB. Further details can be found in [McPherson, 2013].

4.3.1.2 *Bela*

Bela is an embedded platform for audio and sensor processing built around a BeagleBone Black computer on a board, which uses a Texas Instrument Sitara AM3358 system-on-chip (SoC). The SoC features a single-core ARM Cortex-A8 1 GHz CPU and additionally includes 2x 200 MHz Programmable Real-Time Units (PRU), as well as a number of communication peripherals, including SPI, GPIO and I2S. The PRUs are RISC microcontrollers that are optimized for real-time operations can directly access all the peripherals of the SoC and share the RAM with the ARM core. Additionally, the Bela cape (expansion board for the BeagleBone Black) provides a stereo audio codec (communicating to the SoC over I2S), an 8-channel successive approximation register (SAR) ADC and an 8-channel string DAC (both communicating to the SoC over a shared SPI bus). In Bela's jargon, the SAR ADC and string DAC are referred to as "analog", while the codec's channels are the "audio" channels. Additionally, some of the SoC's GPIO are used as Bela's "digital" channels

In the standard Bela environment, the audio processing code and the operating system run on the ARM core, while one of the PRUs is used as a sophisticated Direct Memory Access (DMA) controller, which reads and writes data from and to RAM, to and from the audio, analog and digital inputs and outputs. The operating system on the ARM core is Linux with the Xenomai co-kernel, which allows to achieve lower and more consistent thread wakeup latencies than regular Linux, allowing the Xenomai process with the highest priority (typically, the Bela audio thread), to run within 30 to 60 µs from the moment new data becomes available and the PRU issues an interrupt request (IRQ) to the ARM core.

The combined use of Xenomai and the PRU allows the board to achieve reliable hard real-time, low-latency performance and to dedicate virtually all of the CPU time to audio processing. Bela is capable of achieving round-trip latencies from analog input to analog output smaller than 100 µs, and from audio input to audio

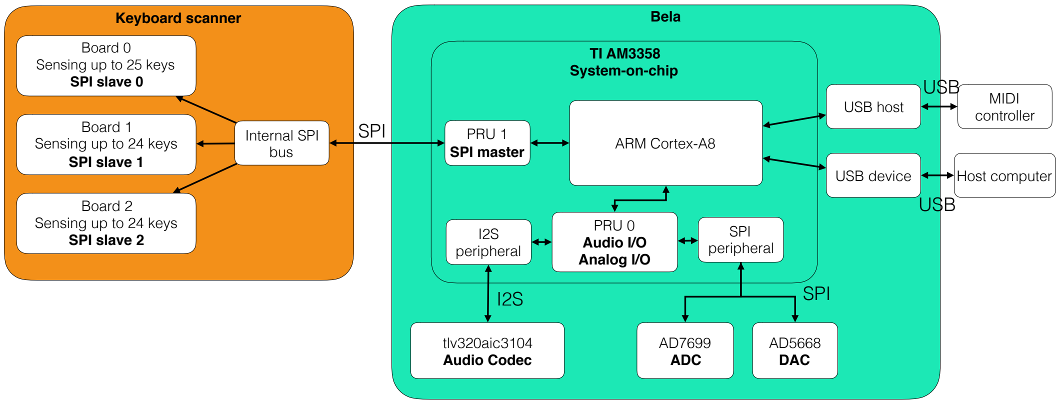


Figure 4.2: Block diagram of the system comprising the keyboard scanner and the Bela board

output smaller than 1 ms (with most of the latency due to the codec's internal filters). Xenomai's high-priority threads can occupy up to 100% of the CPU time, eventually preventing the regular Linux kernel from running at all. While this can potentially lead to some of Linux's services (such as USB or storage) to stop working reliably, the audio thread will be unaffected, as it does not rely on any of them. The PRU, on the other hand, is totally independent from Linux, once it is started, so the I/O is also guaranteed to continue working even when the Xenomai threads take over 100% of the ARM CPU. Bela's real-time capabilities are described in more detail in McPherson and Zappi [2015], Moro et al. [2016], McPherson et al. [2016].

4.3.2 *System architecture*

In our application we used an SPI connection between the topmost board of the scanner and Bela, we then developed a real-time SPI driver to communicate to the scanner from the ARM CPU, so that the position data, appropriately linearised, could be used to control the sound generator.

The Bela environment has no provision for connecting to arbitrary peripherals through the UART, SPI, I2C or USB ports with hard real-time performance. These peripherals are still accessible over the regular Linux driver, but can not guarantee timing accuracy, and their jitter and throughput could become significantly worse under heavy CPU load, and even stop working altogether. For our use case, we want to ensure that the connection to the keyboard scanner guarantees good latency and jitter performance even when the audio thread generates a heavy CPU load. The only viable solution is therefore to handle the connection to the scanner on the second PRU, exchanging data with a Xenomai thread on the ARM core, effectively replicating the principle already used by the core Bela audio and sensors environment. A block diagram of the system comprising the keyboard scanner and the Bela board and all the relevant peripherals and communication buses is detailed in Fig. 4.2.

The keyboard scanner's native interface to the host is via its USB device port, however there is no easy way of accessing the USB bus in a real-time safe way from Bela. Given how for our initial application, the hybrid Hammond instrument, we only needed to sense the 61 playing keys of one of the Hammond manuals, we only needed three scanner boards. We therefore decided to have Bela behave as the SPI master for the scanner, effectively replacing the role of the master board in the original scanner configuration, minus the data acquisition. This solution requires no change to the firmware of the scanner boards, and leaves the USB host port on Bela available to connect a MIDI USB controller to be used for non real-time critical applications, such as adjusting parameters of the sound engine. There is no additional SPI peripheral available on the SoC, so we used the PRU to bit-bang the SPI protocol to communicate to the scanner boards.

The three boards of the scanner are connected in parallel on the same SPI bus, and they share the serial clock (SCK), master input/slave output (MISO) and master output/slave input (MOSI) lines. Each device additionally has one dedicated chip-select line (CS0, CS1, CS2), that allows the master to address each device individually. As an extensions to the SPI protocol, the scanner boards have an additional read/write (RW) line, which is used by the master to notify the devices that a message is being broadcasted and that they should not try to write to the MISO line. We designed and fabricated a printed circuit board to connect the keyboard scanner to the Bela cape. This provides power for the scanner (from Bela's 5V supply rail), and the relevant communication signals, broken out on a 24-pin 0.5 mm-pitch flat flex cable (FFC) connector, suitable to connect to the scanner.

4.3.2.1 *I/O and processing in the Bela core*

The programming structure employed in the Bela core to perform I/O and processing consists of a pair of periodic threads running in parallel on separate cores, exchanging data via shared memory in a two-buffer arrangement, and where one thread performs input and output from a peripheral and acts as a master, relying on an internal or external clocking mechanism, and signals the slave thread when

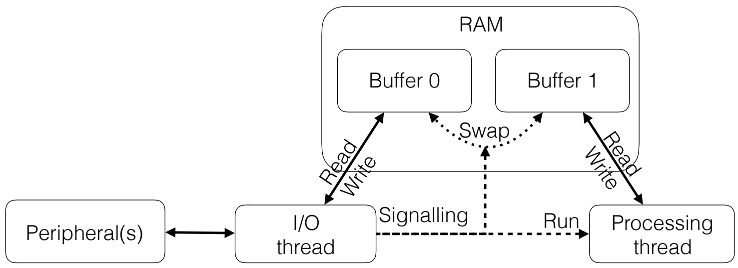


Figure 4.3: Block diagram of the Dual-Thread Double Buffering arrangement for data processing.

new data is available for it to process. We call this a Dual-Thread Double Buffering (DTDB) arrangement, and a block diagram is shown in Fig. 4.3.

The AM3358 SoC only has one ARM core, therefore in order for the threads to run in parallel, one runs on the ARM core, and the other one runs on one of the PRU cores. PRU and ARM will each have exclusive access to one of the two buffers for the duration of one period. For instance, the PRU will spend one period operating on buffer 0, while ARM at the same time operates on buffer 1. At the end of the period, the PRU will signal ARM that it should move on to the other buffer, and so for the next period, the PRU operates on buffer 1, while ARM operates on buffer 0. The I/O thread communicates with the peripherals every few microseconds, and it is crucial that its operation is never stopped (preempted). For this reason, it has to run on the PRU. The processing thread runs on the ARM CPU, and it only keeps it busy for the time needed to perform the processing. It then sleeps, releasing the CPU for other processes or threads to run. The processing thread on ARM benefits from real-time performance through the Xenomai co-kernel, which provides low-latency wakeup, ensuring that it resumes execution quickly, and arbitrarily high priority, to ensure it does not get preempted.

4.3.2.2 *Improvements to the Bela core architecture*

In the original Bela implementation, signalling from the PRU to the ARM thread in the DTDB was done by means of a flag in the shared memory. Once the thread running on ARM would terminate executing one iteration of its loop, it would go to sleep and wake up at regular intervals (every 25% of its period time) and poll the flag to detect whether new data was available. This implementation is simple but has two important drawbacks. First of all, each time a thread wakes up, it uses some CPU time, and if it wakes up and there are no data ready to process, this means that precious CPU cycles are being wasted. Second, the jitter between the data being ready and the thread waking up can be as large as 25% of a period, meaning that the CPU could not consistently spend more than 75% of the time processing audio without the risk of missing the next deadline.

This behaviour is acceptable when there is only one high-priority thread running on the CPU, however it becomes particularly inefficient when two real-time threads concur for CPU time. We therefore designed a more efficient notification mechanism, that allows the real-time thread to be woken up upon receiving an IRQ from the PRU. We wrote a device driver using Xenomai's Real-time Driver Model (RTDM)[8]. A Xenomai thread can safely make calls to a RTDM driver without losing its real-time characteristics. The driver programs the interrupt controller of the PRU so that it can trigger one of the IRQ lines of the interrupt controller on ARM. It then registers the interrupt handler and enables the corresponding IRQ line. Upon opening the driver's device file, it can be configured when the application starts through the *ioctl()* interface. At run-time, a call to *read()* will block until the IRQ line is triggered. When the IRQ is received, the thread becomes runnable and the scheduler will schedule it based on its priority. In our case, the PRU can trigger the IRQ with a simple write to specific bits in its own R31 register. We incorporated the `rtdm_pruss_irq`[9] driver in the Bela core environment, as well as using it for our own application.

4.3.2.3 *Real-time threads*

Thread name	**Core**	**Priority**	**Clock in source**	**Period**	**Clock out divider**
Bela I/O	PRU0	n.a.	I2S frame clock	$1/(2F_s)$	$2n$
Audio	ARM	95 (Xenomai)	Bela I/O thread	$2n/F_s$	n.a.
Scanner I/O	PRU1	n.a.	Cycle counter	1 ms	1
Scanner Processing	ARM	90 (Xenomai)	Scanner I/O thread	1 ms	n.a.

Table 4.1: Summary of the real-time threads in our system

In our system, there are two instances of the DTDB architecture, for a total of 4 threads: one instance is part of the Bela core environment, described above, to which we add another one to handle the data from the keyboard scanner. The Audio and Scanner Processing threads are running on the ARM CPU, and the cor-

8 https://xenomai.org/documentation/xenomai-3/html/xeno3prm/group__rtdm.html
9 https://github.com/BelaPlatform/rtdm_pruss_irq

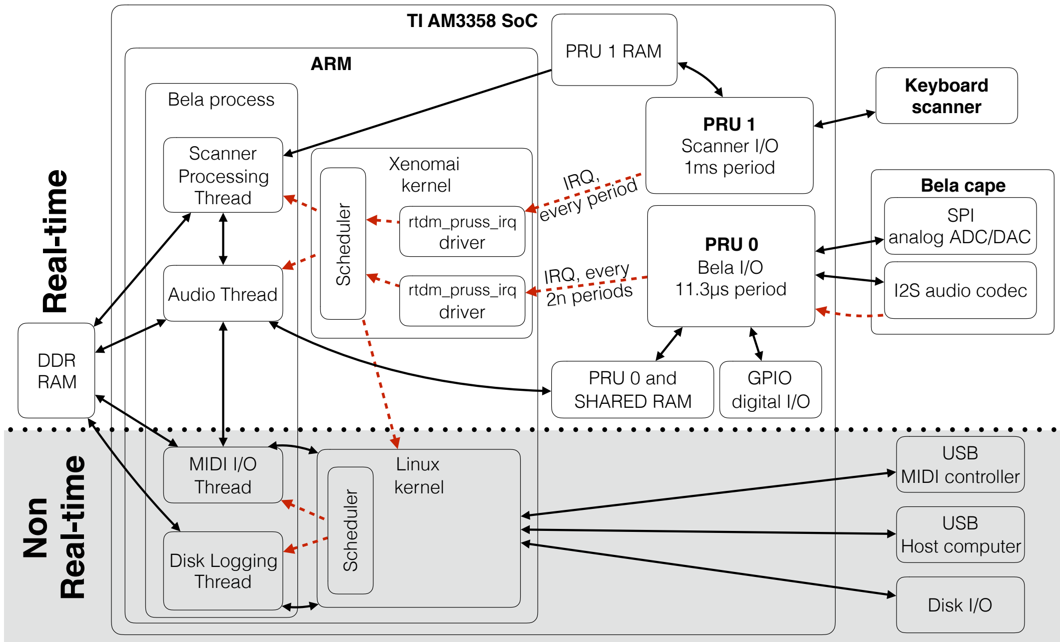


Figure 4.4: Block diagram of the threads in our system and their functions. Black arrow lines indicate data flow. Dashed red arrow lines indicate the direction of clock, IRQ, or scheduling control.

responding PRU threads are the Bela I/O and the Scanner I/O threads, running on PRU0 and PRU1, respectively. The Scanner Processing thread performs linearisation of the sensor data and runs a user-defined sensor processing callback. The Audio thread is the Bela core processing thread, and runs a user-defined audio processing callback. An example application would detect note onsets an velocity in the Scanner thread, and use those to control a synthesizer, whose sound is generated in the Audio thread.

The characteristics of each of the threads are summarised in Table 4.1, and Fig. 4.4 shows the threads in the context of the wider system architecture. In the table and figure and in the passage below we use the following notation:

- F_s is the audio sampling rate. The default value is 44.1 kHz.
- n is the blocksize, that is the number of audio frames passed to the Bela audio callback in the audio thread. This can assume one of the following values: 2, 4, 8, 16, 32, 64, 128.

The Bela I/O thread on PRU0 runs an internal loop with a fixed period corresponding to $1/(2F_s)$, typically 11.33µs, and runs once per each transition of the frame clock of the I2S peripheral. For each iteration of the loop, the PRU reads data from the shared buffer in PRU0 RAM to write them to the audio, analog and digital output peripherals, and it reads data from the input peripherals to write them into the buffer. The execution time for one iteration of the loop is approximately 8μs, and once one iteration is completed, the thread waits for the I2S peripheral to signal that a new channel is ready to be read or written, and it starts another iteration. Every $2n$ iterations of the loop, the PRU sends an IRQ to the ARM CPU, and updates a flag in memory to notify the Audio thread of the buffer swap.

The Audio thread on the ARM runs with the highest priority on the board, and is guaranteed not to be preempted by any other thread, or by any IRQ requests, thanks to Xenomai. It runs an internal loop with a period corresponding to n/F_s; with the default value $n = 16$, this corresponds to 362.8μs. For each iteration of the loop, the thread reads from PRU0 RAM which buffer it should operate on, and copies the corresponding audio, analog and digital data from the PRU RAM into DDR RAM. It subsequently calls a user-defined callback, which will produce some audio, analog and digital outputs that are copied back into the PRU RAM. Once one iteration is completed, the thread blocks on a *read()* call on the *rtdm_pruss_irq* driver, waiting for an IRQ from the Bela I/O thread to resume.

The Scanner I/O thread on PRU1 runs a loop with a fixed period of 1 ms, and is scheduled according to an internal counter. The communication is bi-directional between the PRU and the keyboard scanner via the SPI bus, and uni-directional from the PRU to the ARM thread via shared memory. For each iteration of the loop, the PRU gathers the scan results from the keyboard scanner and writes them into the RAM buffer. The execution time for one iteration of the loop is approximately 100μs, and once one iteration is completed, the thread waits until 1 ms has elapsed from the start of the last scan, and then it starts another iteration. As each of the boards of the scanner has their own internal clock generator, they would tend to drift over extended periods of time, therefore, every 5 iterations of the loop, the PRU additionally sends a synchronization message to the boards avoid clock drifts over extended periods of time, similarly to what the master board of the scanner would do. At the end of each iteration of the loop, the PRU also asserts an IRQ line on the ARM CPU, starts working on the buffer that was previously used by the Scanner Processing Thread, and updates a flag in memory to notify the Scanner Processing Thread of the buffer swap.

The Scanner thread on the ARM CPU is a Xenomai thread with lower priority than the Audio thread, and it also runs a loop with a period of 1 ms. For each iteration of the loop, the thread reads from the shared memory with the PRU which buffer it should operate on and copies the corresponding input data from the PRU

RAM into DDR RAM. It validates the input from each of the boards, interpolating when missing or corrupted data are received, applies the linearisation algorithm, described in Section 4.3.3, to the sensor data and subsequently calls a user-defined callback. Once one iteration is completed, the thread blocks on a *read()* call on the *rtdm_pruss_irq* driver, waiting for a new IRQ from the PRU Scanner I/O thread to resume.

The Xenomai threads are scheduled by Xenomai with a first-in, first-out (FIFO) policy, and the scheduler implements a pre-emptive, priority-based scheduling [Gerum, 2004]. This means that the currently running thread will always be the thread with the highest priority among those threads that are ready to run. In the case of our system, this means that the Audio thread, which runs with the highest priority, may preempt the Scanner Processing thread during its execution, because the two threads are not synchronized, as they are driven by independent clocks (see Table 4.1).

The main means of communication between the two threads consist of a shared memory buffer representing the current position of each key. This buffer is filled in by the Scanner thread, and accessed from the Audio thread. We did not implement more complicated solutions such as a mutual exclusion mechanism, as those come with CPU overheads due to memory copying and use of mutexes. The end result is that the Audio thread "samples" the key position buffer at irregular intervals, and potentially it can "see" the same buffer multiple times (if running faster than the Scanner thread), or miss one or more Scanner frames altogether (if running slower than the Scanner thread). The shared buffer contains key position information stored as 32-bit `float` values. The Scanner thread ensures that the buffer thread is always in a valid state by writing values into it sequentially. If the thread is preempted while writing, the buffer will contain partly old data and partly new data As the Cortex-A8 can perform single-cycle writes for 32-bit values, this guarantees that there will never be corrupted data in the buffer.

The process of extracting features from the key position data often relies on having access to measurements performed at a constant sampling rate. For this reason, tasks such as computing velocity, percussiveness, or keeping a state machine of the key movement, or, more in general, anything that relies on the temporal evolution of the key profile are better placed within the Scanner thread. These features can then be communicated to the Audio thread via global variables, or lock-free queues. On the other hand, when the instantaneous position of the key is all that matters, for instance in the case where it is used to directly control the amplitude of the generated audio signal, then the Audio thread can access the position information directly through the shared buffer, and apply the necessary filtering to smooth the data, if needed.

4.3.2.4 *Non real-time threads*

We allow some lower-priority threads to run alongside the Audio and Scanner threads on the ARM CPU. These provide facilities such as MIDI I/O, to connect a USB MIDI controller or receive MIDI data from the host computer, and disk output, to write logging audio and key position data to disk for later analysis. These threads use Xenomai's Real-Time Inter-process protocol[10] to communicate to and from the real-time threads.

4.3.2.5 *Action to sound latency*

We performed a series of measurements to evaluate the action-to-sound latency of the platform. The test setup involves an IR source connected to one of the analog outputs of Bela. The IR source is placed in front of one of the sensors on the keyboards scanner. A program running on Bela periodically writes a square wave to the analog outputs and measures the amount of samples it takes for the signal to be read back through the sensor. The latency figures thus obtained will be the sum of the inherent latency of Bela (one processing block) and the latency of the keyboard scanner itself. The actual action-to-sound latency will be the latency of the scanner plus the duration of one processing block plus the group delay of the audio DAC (21 samples)[11]. The summary of the computed action-to-sound latency for block sizes of 16 and 64 samples and different CPU usages are shown in Table 4.2. We performed at least 1600 measurements for each condition. The mean latency is 3.33 ms for block size of 16, and the jitter, that is the maximum variation between the mean and the worst case scenario, is below ±0.5 ms. At 64 samples per block, the mean latency is 4.15 ms, and the jitter is ±0.7 ms, so that the worst case scenario is always below 5 ms. CPU usage does not seem to have a considerable effect on the latency figures, however we would expect that when the CPU usage of the audio thread approaches 90%, the performance may degrade severely, as the scanner thread will start missing some of its deadlines. Wessel and Wright [2002] recommended 10 ms with a jitter of ±1 ms as the recommended

10 https://xenomai.org/documentation/xenomai-3/html/xeno3prm/group__rtdm__ipc.html

11 Datasheet of the Texas Instruments TLV320AIC3104 codec: http://www.ti.com/lit/ds/symlink/tlv320aic3104.pdf

worst case latency for a responsive instrument. Jack et al. [2018] confirms those findings, in particular showing that larger jitter values can be detrimental to the perceived quality of the instrument. Our results confirm that the performance of our platform well exceeds these criteria and can consistently achieve better results than most devices commonly used for DMIs McPherson et al. [2016].

Block size	**CPU usage**	**Latency(ms)**			
		Mean	σ^2	Min jitter	Max Jitter
16	25%	3.33	0.19	-0.37	0.44
16	60%	3.33	0.19	-0.37	0.35
64	25%	4.15	0.28	-0.65	0.62
64	60%	4.17	0.28	-0.67	0.69

Table 4.2: Action to sound latency figures for our platform.

.

Given the non real time scheduling of the MIDI thread, the action-to-sound latency of MIDI inputs is larger than that of data received from the keyboard scanner, and subject to a much higher jitter, especially when the CPU load is high, and we would expect latency figures comparable to those reported by McPherson et al. [2016] for USB MIDI on macos (mean latency of about 9 ms with an additional jitter of ±3 ms).

4.3.3 *Sensor linearisation*

The keyboard scanner uses near-field optical reflectance sensors (Fairchild QRE1113), which combine an infrared LED and a phototransistor in the same package. As the LED is powered on, it emits infrared (IR) light, which is reflected back into the phototransistor if a reflective surface is placed in front of the sensor. The amount of IR light reflected into the phototransistor, and therefore its collector current, is a function of the distance and/or angle of the reflecting surface. On the keyboard scanner, there is one phototransistor per key, and the scanner is installed so that the sensor is facing down towards the surface of the back of the key. Black keys are naturally less reflective than white keys, and therefore a piece of white tape has to be applied at the back of the black keys in order to ensure a good dynamic range in the sensor reading. The sensing circuitry and the acquisition strategy are detailed in McPherson [2013].

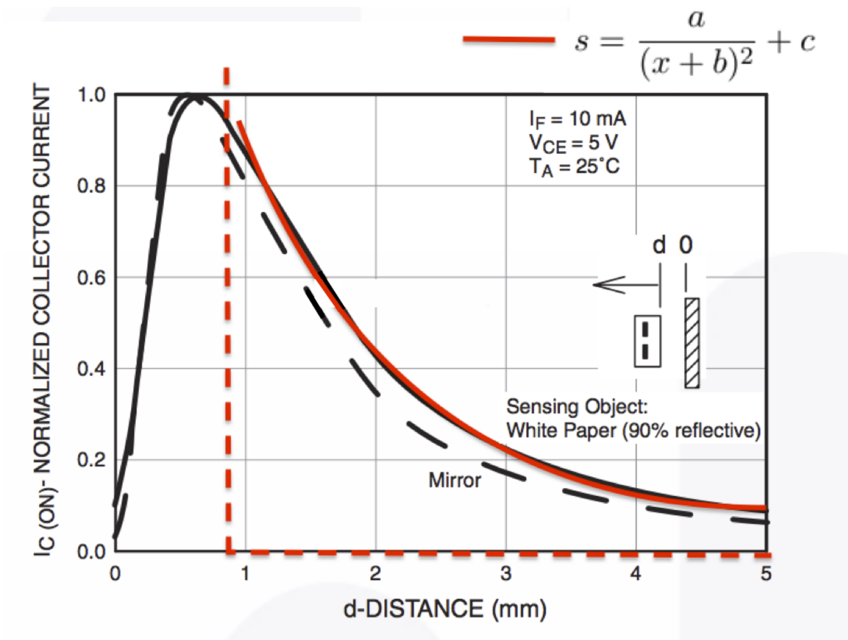


Figure 4.5: QRE1113 Normalized collector current vs Distance between device and reflector. This image is our overlay (red) on top of an image from the QRE1113 datasheet (see Footnote 12)

The collector current of the phototransistor is proportional to the amount of IR light reflected into the phototransistor, which in turn is a non-monotonic function of the distance of the reflective surface, as reported by the datasheet for the part[12] (see Fig. 4.5). For a distance comprised between 0.6 mm and 5 mm, the behaviour is monotonic, and roughly approximates the inverse square function. The keyboard scanner employs a transconductance amplifying circuit to turn the collector current into a voltage which is in turn digitized by the on-board ADC. The ADC reading provided by the keyboard scanner is ultimately proportional to the collector current, thus exhibiting the same non-linear behaviour in function of the distance of the key. In order to obtain the key position from the scanner reading, there is need for a linearisation algorithm.

In Section 3.3.1.1 we also used a near-field optical-reflectance sensor, which had a similar non-linear curve to the QRE1113. In that case, we were operating on one key at a time and required high precision, therefore we calibrated the sensor by taking in excess of twenty measurements along the whole key throw, using a digital caliper and a Bela, and then generated a look-up table that was then used to linearise the sensor reading. The process was quite tedious and it would take

12 See "Figure 2. Normalized Collector Current vs. Distance between Device and Reflector" of `https://www.onsemi.com/pub/Collateral/QRE1113-D.PDF`,

about 5 minutes per key. In this case we have the necessity of calibrating up to 73 keys at a time, and the procedure should be easy and take no more than a couple of minutes for the whole keyboard, so that it would be convenient to repeat it as frequently as needed. In the monotonic region, the relation between the sensor reading s and the key displacement x can be described by the following formula:

$$s = \frac{a}{(x+b)^2} + c \tag{4.1}$$

This is the formula for the inverse square function, scaled by a, translated on the x axis by $-b$ and with a vertical offset c. For $x = 0$ the key is at rest (key top) and for $x = 1$ the key is fully depressed (key bottom). If a, b, c are known, the position of the key can be computed from the sensor reading as:

$$x = -b \pm \sqrt{\frac{a}{s-c}} \tag{4.2}$$

where we choose the positive sign as it is the one that yields a meaningful result. In order to obtain an estimate for the parameters a, b and c, we need to take three paired measurements of the sensor reading and key displacement at different points of the key throw and feed them to a nonlinear curve fitting algorithm.

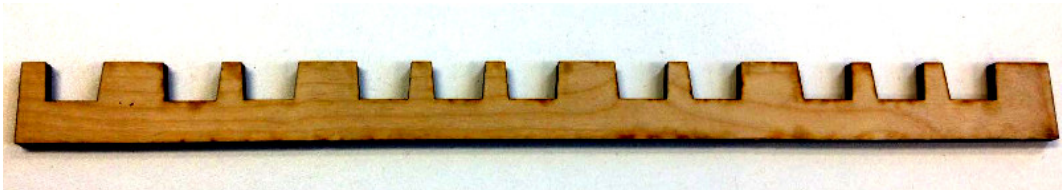


Figure 4.6: The "calibration comb"

In order to obtain these three measurements, we rely on the reasonable assumption that all the white keys and all the black keys on a given keyboard have the same shape, the same rest position and the same key bottom position. For each key, we can easily obtain two paired measurements, by measuring the sensor reading s with the key at rest ($x = 0$) and with the key fully depressed ($x = 1$). In order to obtain a third measurement for each key, we designed a custom calibration tool that we call the "calibration comb" (Fig. 4.6). This is a piece of rigid 8 mm pine plywood spanning 28 keys (A to C), where 14 mm-deep slots have been cut out in correspondence of the sharp keys. If the comb is positioned vertically at the front edge of the black keys and pressed down, the following behaviour can be observed:

- initially, only the white keys move

- when a black key starts moving, the displacement of the neighbouring white keys is approximately x_{mw}
- when a white key is fully depressed, the displacement of the neighbouring black keys is approximately x_{mb}

If our assumptions hold, the values x_{mw} and x_{mb} will be constant across the whole keyboard and they can be measured once with a caliper and reused for all successive calibrations. The three pairs of measurements we need to collect for each key are listed in Table 4.3.

Sensor reading	Key position	Measurement
s_t	0	Top (key at rest)
s_m	x_{mb} or x_{mw}	Midpoint (obtained with calibration comb)
s_b	1	Bottom (key fully depressed)

Table 4.3: Paired sensor and position measurements for each key

4.3.3.1 *Calibration procedure*

The calibration procedure acquires data from the keyboard scanner on the Bela board, the data is then transferred to the host computer where a Matlab script performs data segmentation and curve fitting, generating a calibration file that is then copied back to the Bela board. The operator starts the procedure by starting a data logging program on the board which reads the sensor data and writes them to disk. The calibration procedure is divided in four sections:

1. As the procedure starts, all the keys need to be at the rest position for 1 second, to calibrate the position of the key at rest.
2. The operator then fully depresses all keys, one at a time, to calibrate the key bottom position.
3. If the calibration procedure requires calibrating aftertouch, the operator can optionally press all the keys again, this time adding extra pressure after the key bottom is reached. This step is optional.
4. The operator then places the calibration comb at the top of the topmost octave, and uses it to depress the 28 keys at once several times (5-10). The calibration comb is then moved down in steps of one octave and this operation repeated until all the keys have been covered.

The operator then stops the data logging program and transfers the generated data to the host computer. For each key, multiple measurements of s_m are obtained, one for every time the comb was pressed on that key, and only the median s_m value is considered for the final curve fitting step, performed using MATLAB's *lsqcurvefit()*[13], which finds the parameters for the provided function that best fit the data in the least squares sense. This gives us estimated parameters a, b and c for Eq. (4.1). The fitted parameters are then exported from Matlab and stored back on the board, where they are used to compute normalized, calibrated positional data for each key at runtime, using Eq. (4.2).

The calibration procedure only needs to be repeated when the scanner is moved, as changes in ambient light and in the offset of the amplifying circuit due to thermal drift will be automatically compensated for by the differential reading implemented in the scanner board.

4.4 HAMMOND SOUND ENGINE

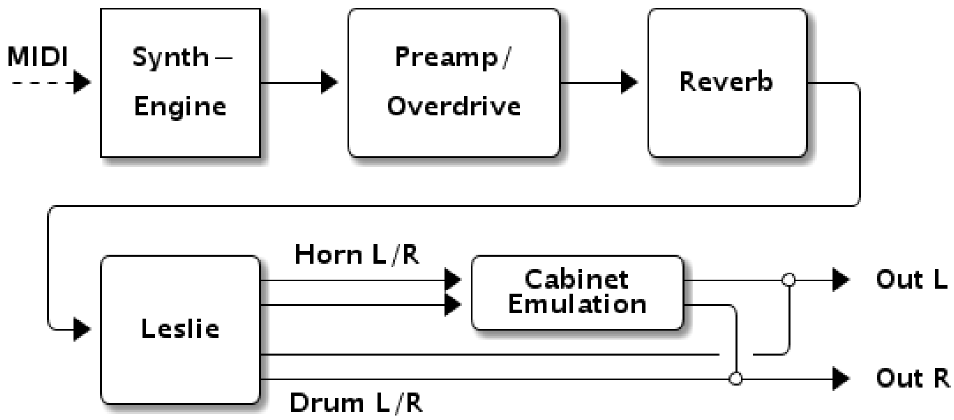


Figure 4.7: Signal flow diagram of setBfree, reproduced from https://github.com/giuliomoro/setBfree/blob/dynamic-envelopes/doc/signal_flow.png.

In its original implementation, *setBfree*, which is written in C, is a fully-featured Hammond emulator that receives MIDI note messages to activate each key, and implements static or random envelopes to emulate contact bouncing. The original block diagram of the software is displayed in Fig. 4.7. The part we are interested in, the virtual tonewheel generator, contacts and drawbars are comprised in the "Synth-engine". The synthesizer engine is mostly entirely comprised in the long

13 https://uk.mathworks.com/help/optim/ug/lsqcurvefit.html .

`oscGenerateFragment()` function[14]. A simplified description of the function follows. We can consider `oscGenerateFragment()` as comprised of three main parts:

PROCESS THE MESSAGE QUEUE A lock-free queue is used to communicate between the MIDI and the audio thread. `oscGenerateFragment()` retrieves messages from the queue to update an internal state (the "activated list"), which stores the currently playing notes, with those notes that just started or just stopped playing.

PROCESS THE ACTIVATED LIST The activated list, newly updated, is read to find which notes are currently playing, and which of these have just started playing. Each of the tonewheels associated with each of the contacts of the notes that are currently active is assigned a gain value depending on the voicing of the instrument (representing the resistive wires in the Hammond manuals). The notes that just started playing are assigned a pre-computed envelope (to simulate contact bounce). At this step a series of "core instructions" are generated. These consist of a gain value and a pair of pointers to the pre-computed tonewheel lookup table and, if appropriate, to the envelope table.

EXECUTE THE CORE PROGRAM INTERPRETER The core interpreter executes the core instructions by reading values from the tonewheel table and applying the prescribed gain and, if appropriate, envelope. The results of each multiplication are added into an output buffer.

We modified the program for the purpose of our study, so that it can now connect to the keyboard scanner, receiving continuous key position and individually triggering each of the nine tones associated with the key. We added complete control over each individual attack transient, designing and implementing a dynamic parametric model for contact bounce. This required modifying several parts of the program. With regard to the `oscGenerateFragment()` function itself, we added provisions for receiving not only *note* on/off messages, but also individual *contact* on/off messages, associated with a velocity parameter. Furthermore, the envelopes are now computed dynamically on the basis of the velocity of each contact onset. The software can log to disk the position of all the keys (sampled at 1 kHz) as well as the generated audio signal (sampled at 44.1 kHz).

The following sections describe our most salient contributions to the software: a dynamic contact bounce model (Section 4.4.1), its mapping possibilities (Section 4.4.2), and a method for semi-automatically calibrating the voicing of the

14 The original can be found here: https://github.com/pantherb/setBfree/blob/master/src/tonegen.c

sound engine based on recordings of a Hammond (Section 4.4.3). Our final version of the code is available online.[15]

4.4.1 *Synthesis of the contact onset*

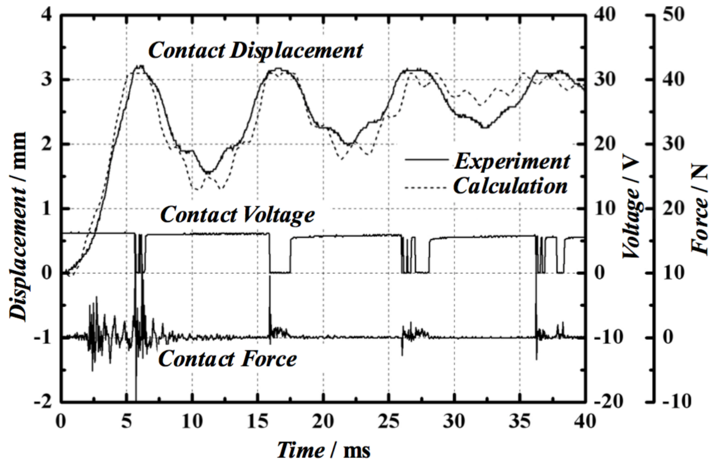


Figure 4.8: Contact displacement over time for contacts in a relay. Image reproduced from Jun et al. [2008].

In this section we present an empirical approach at the digital simulation of the bouncing of contacts during the note onset on the Hammond organ. Our purpose is to obtain an insight in the contact bounce phenomenon, and to produce a model that can adequately respond to a single control parameter such as the initial velocity of the contact. On the other hand, we are not aiming at a physically accurate simulation, as that would require a deeper analysis of the materials and weights involved in the process, coupled with some sophisticated measuring equipment such as high-frequency laser displacement sensors, as used in [Jun et al., 2008]), and a scientifically accurate simulation of the phenomenon transcends the scope of our research. Starting by observing the recorded bounces, we built a simplified model based on qualitative observations and justified by a simplified physical representation of the underlying phenomenon.

4.4.1.1 *Empirical observations*

During the experiments described in Chapter 3 we collected a number of measurements of the state (open or closed) of contacts on the Hammond during key

15 https://github.com/giuliomoro/setBfree

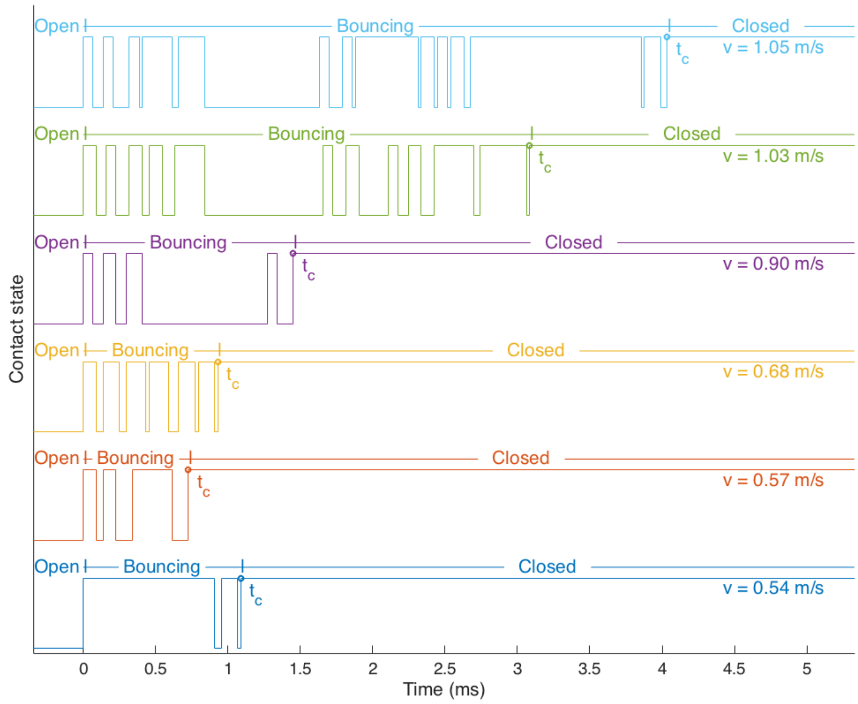


Figure 4.9: Recorded contact bounces for different values of instantaneous key velocity on the Hammond. The measurements were obtained by reading the voltage across the switch and thresholding it. The contact is closed when the line is high, open otherwise. The signal was sampled at 44.1 kHz

onset. We showed some of their macroscopic characteristics, and their dependence on the gesture used, in Section 3.3.2.3. In that occasion, the speed of the contact at the impact was not measured, however we did record the position of the key. From the key position, it was possible to compute the instantaneous speed of the key, by taking the numerical derivative of the position. The speed of the key at the instant when the contact first touches was used as a proxy for the initial speed of the contact when it hits the busbar.

We inspect here the details of how an individual contact bounces over time, and its high-frequency oscillations. Fig. 4.9 shows the state of the contact during several onsets, each obtained with different values of instantaneous keys velocity. After the contact first enters the closed state at time $t = 0$, it then alternates between the open and closed states before settling in the closed state. We identify for each of the onsets displayed in Fig. 4.9, t_c as the time when the contact enters the "closed" state for the last time. We distinguish three different time intervals for each onset:

the "open" interval(when $t < 0$), the "bouncing" interval (for $0 \leqslant t < t_c$), and the "closed" interval, for $t > t_c$. The duration of the bouncing interval, and therefore t_c, tends to increase with velocity.

Observing the pattern of the bouncing interval we notice that it alternates periods of fast transitions to periods when the contact stays in the same state for a prolonged time. We call these states "temporarily open" and "temporarily closed". During the periods of "fast transitions", the oscillations have a period of approximately 200μs.

4.4.1.2 *Physical justification*

The motion of the contact spring as it hits the busbar can be approximated as a harmonic motion where energy is dissipated at every bounce of the contact on the busbar. Barkan [1967] shows that when a contact switch closes, the moving contact follows, as a first approximation, the rules of a spring-mass system with inelastic impacts. As a consequence of the inelastic impact, a dissipation of energy takes place at each bounce, which in turns reduces the amplitude of the successive bounce. The coefficient of restitution (COR) is an indication of the energy that is dissipated during an impact and it can be obtained experimentally. For an inelastic impact between two bodies, one of which is either not moving or used as the system reference, the COR e is given as:

$$e = -\frac{v_f}{v_i} \tag{4.3}$$

where v_i is the velocity of the body before the impact and v_f is the velocity of the body after the impact. For an elastic impact, $v_i = v_f$ and $e = 1$. For a fully inelastic impact, $v_f = 0$ and $e = 0$. For partially inelastic impacts the COR is $0 < e < 1$ and it is a measure of how much energy is dissipated in the impact (with smaller values indicating larger values of energy dissipation).

The equation for simple harmonic motion describes the position x and the velocity v as a function of the time t for an angular velocity ω and an amplitude A, with phase ϕ at $t = 0$, as:

$$x(t) = A\cos(\omega t + \phi) \tag{4.4}$$

$$v(t) = -A\omega\sin(\omega t + \phi) \tag{4.5}$$

where ω depends on the mass and elastic constant of the system and A depends on the initial displacement.

At the time of first impact, $t = 0$, the contact is at position $x = 0$ and has instantaneous velocity v_0. For $t > 0$, as long as the key holds the contacts pressed down, we can describe the motion of the bouncing contact as a piece-wise concatenation of simple harmonic motions under the constraints that every time the position reaches the rest point ($x = 0$), the contact bounces back with a partially inelastic impact. It derives from Eq. (4.4) that the initial phase for the first bounce will be such that $A\cos(\phi) = 0$. If we take A to be positive, and the contact bounce covering positive values of x, then $\phi = -\pi/2$. The first bounce will start with velocity $v_f = -ev_0$ (from Eq. (4.3)), and the oscillation will have amplitude $A = v_f/\omega$ (from Eq. (4.5), for $t = 0$). After half a period of harmonic oscillation, the contact will be back at $x = 0$, and impact the busbar again, interrupting the current harmonic oscillation and starting a new one.

At each bounce, we compute new values for A and ϕ such that:

$$A := Ae$$
$$\phi := \omega t - \phi - \pi/2$$

thus effectively resetting the oscillation to an initial phase of $-\pi/2$, with a new amplitude obtained from the previous one scaled by the COR (combining Eq. (4.5) and Eq. (4.3)). The actual vibrating pattern will be more complicated than this, as the contact beam will be also subject to some additional oscillations due to its free oscillation modes [Jun et al., 2008].

4.4.1.3 *Proposed model*

For the purpose of this study we need a contact bouncing model that can be controlled by the initial velocity of the contact. We use the simplified model of a bouncing contact presented in Section 4.4.1.2 to compute the displacement of the contact over time. We use the velocity of the key at the moment when the contact closes as a proxy for the velocity of the contact itself and we use that value to compute the initial amplitude A in Eq. (4.4).

We then threshold the contact displacement with two thresholds, th_h, th_l, such that $th_h > th_l$. Comparing the thresholds with the contact displacement x we can infer the behaviour of the contact at each time sample. When $x < th_l$, the contact reed is pushed against the busbar and therefore the switch is in the "temporarily closed" state, when $x > th_h$, the contact reed is far from the busbar and therefore the switch is in the "temporarily open" state. When $th_l < x < th_h$ the contact is in the "fast transition" state. This state is an approximation we introduce to account

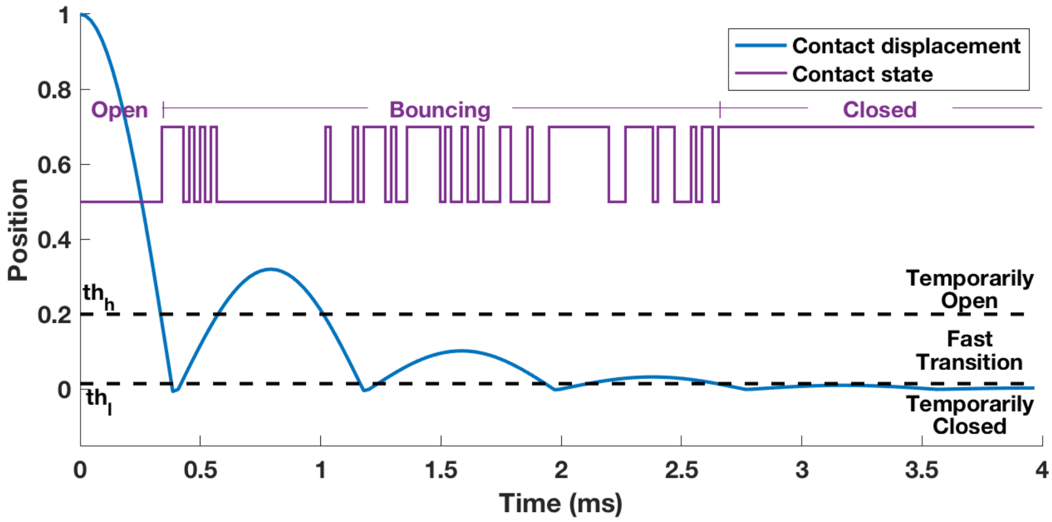


Figure 4.10: Synthesized contact bounce. The continuous position of the contact is displayed, along with its electrical state. Parameters used in the simulation: $A = 1$, $\omega = 4000$, $th_l = 0.015$, $th_h = 0.2$.

for the combined interaction of a higher frequency modal oscillation of the contact and the non-ideal electrical characteristics of the contact surface. When in "fast transition", we use a randomized function to determine the electrical state of the contact. The result is in accordance with our observation in Section 3.3.2.3 that the higher the velocity at the moment of the first impact, the longer the bounce: a higher initial velocity implies higher kinetic energy in the system, which takes longer to dissipate for a given oscillating frequency and COR.

Once the electrical state of the contact over time is obtained, this could in principle be used to gate the signal at the input of the contact to obtain the output signal. However, after we implemented it this way, early informal playing/listening tests revealed that the attack of the note tended to sound too sharp, noisy and unpleasant.

After a visual inspection of the audio recordings of the actual instrument (such as the ones in Fig. 3.3a and Fig. 3.3b), we noticed that the amount of amplitude modulation caused by the contact bounce was smaller than the full-range value it was expected to have. This is probably a combined effect of the capacitive coupling between the contact and the busbar while the contact is bouncing and of the inductance of the coupling transformer. We took this into account by introducing a tunable parameter `depth`, which denotes the actual amplitude excursion of the bounces in a normalized 0:1 range. The resulting bounce jumps from 0 to 1 as soon as the virtual contact first hits the busbar, but successive amplitudes are limited between 1 and $(1 - \mathtt{depth})$. We experimented with different values for this $\mathtt{depth}$ parameter and settled on 0.15 for our Hammond emulation.

This model offers a simple, computationally efficient, physically-justified approach to an algorithmic synthesis of the contact bounce which can be controlled by a single real-time input parameter, the speed of the contact at the time of the first impact, and tunable parameters such as the COR and the *width* of the modulation.

Informal listening tests have shown that the model seems to behave realistically in the domain of application, after the amplitude limitation correction. However, unless the sound generated by the model is validated by either a more comprehensive listening test or a perceptually-informed DSP analysis technique, we can only speculate as to whether any of the limitations above impacts the realism effect.

4.4.2 *Control of contact onsets*

We modified the *setBfree* program to accept as an input the position of the key and customize how the key position and the key velocity affect the bounce of each contact, as detailed below. Examples of the mappings that can now be achieved with the software, and their sonic possibilities are described in Section 4.6.1.

MAPPING BETWEEN KEY POSITION AND CONTACT TRIGGERING

There are three different mappings between key position and contact triggering available. For each of them, the triggering points are adjustable.

- **Single triggering point, no velocity**: the onsets of all the contacts are triggered when the key crosses a given threshold. A randomized contact closing offset value in a custom range will be generated for each onset.
- **Single triggering point, velocity**: the onsets of all the contacts are triggered after the key crosses a given threshold. The contact closing onset ranges from 1 ms to 127 ms and is depending on the velocity of the onset. As a result, fast velocities will trigger the contact onsets very close together, whereas slower velocities will spread them over a longer period of time.
- **Individual triggering points**: each contact is set to trigger at a given threshold in the key-throw. This can be customized per-contact and per-key.

With the *Single triggering point, velocity* mapping, the generator produces a velocity-dependent temporal staggering (see contact closing offset in Section 3.3.2.3), of the virtual contacts to replicate the temporal staggering due to triggering the three individual contacts. For those key presses that have a constant acceleration and go all the way to the bottom of the key, this behaviour should be virtually indistinguishable from that of the *individual triggering points* mapping. However, for partial presses, or when the acceleration is distributed non-uniformly during the key press, then the difference between the last two mappings becomes more obvious.

mapping between velocity and contact bounce

As an additional parameter, selectable independently from the above, the instantaneous velocity of the key at the triggering point can be used to control the initial velocity of the contact in the contact bounce synthesizer. Velocity is computed using Eq. (3.1) based on the position values at the current and previous time instants. When the mapping is disabled, a randomized value for each key press is used to initialize the contact bounce synthesizer.

4.4.3 *Voicing of the sound engine*

The process for adjusting the relative gain of the tones connected to each individual contact on a real tonewheel Hammond organ is tedious but crucial. Its ultimate purpose is to ensure that the tone of the instrument is balanced, across registers and across drawbar settings. These adjustments are performed by setting the length, and therefore the overall resistance, of the resistive wire that connects the tone generator's pick up to each contact. This voicing procedure was routinely performed before the organ left the factory, and rarely, if at all, repeated during the lifetime of the instrument.

In the *setBfree* software, the user can manually specify, via a configuration file, the gain of the signal from each tonewheels that reaches each individual contact, obtaining an effect akin to that of the resistive wire in the real instrument. We found that with the default settings, *setBfree* sounds unbalanced, with a weak bottom end and a shrilling high end.

We used the sound of our Hammond C3 organ as a reference for re-voicing the digital emulation through a semi-automated procedure. The operator would play a chromatic scale on all the keys of the organ, for nine times, each time with only one drawbar activated, while recording the audio output of the organ before the pre-amplifier. This way, each key press would connect the signal from exactly one tonewheel to the output busbars, and its amplitude would be scaled by the resistive wires. The resulting audio signal would contain mostly the signal from that tonewheel, plus any crosstalk from other tonewheels picked up either in the generator itself or elsewhere inside the organ. Assuming the effect of the crosstalk on the amplitude of the signal to be negligible, the amplitude of the generated signal will be determined by the combined effect of the amplitude of the signal generated by the tonewheel and the gain reduction due to the resistive wires. A *Matlab* script was used to analyse these recordings and extract the amplitude of the signal for each tonewheel in order to generate a configuration file for *setBfree*.

4.5 THE COMPLETED INSTRUMENT

Figure 4.11: Picture of the completed instrument.

The instrument in its final form is shown in Fig. 4.11. The scanner is mechanically attached with Blu Tack[16] to the vertical panel at the front of the organ which separates the two manuals, and it scans the position of the keys on the lower manual. The top manual and the controls are unused, and they can be covered with a black cloth for the experiment in order to hide from sight Bela and the other components. The audio output of Bela is routed into the AO-28 tube preamplifier on the Hammond, at the same point where the audio input from the matching transformer of the drawbars is normally connected. This is the same point in the circuit from which the "preamplifier input" signal is taken in Section 3.2.6. An additional QRE1113 sensor is used to track the position of the expression pedal of the organ. This measures the position of the terminal part of the rod that connects the pedal to the variable capacitor in the preamplifier.

The *setBfree* software running on Bela responds to key position data to trigger contacts. The chorus, reverb, pre-amplifier and rotary speaker simulators of *setBfree* are disabled, so that the audio output of Bela corresponds to the output

16 https://en.wikipedia.org/wiki/Blu_Tack

of the virtual drawbars, as required by the diagram in Fig. 4.1. Audio processing takes place in blocks of 64 samples, sampled at 44.1 kHz, which, as we reported in Section 4.3.2.5, results in an overall action-to-sound latency always smaller than 5 ms. The jitter figure (±0.7 ms) is particularly relevant to our application, as it is much smaller than the closing time offset on the Hammond (see Fig. 3.9a), indicating that our instrument can provide a reasonably close match to the real Hammond in terms of the relative closing timing of the contacts.

setBfree can also respond to MIDI messages, either from the host computer or from a MIDI controller plugged into the USB host port. The MIDI controls allow to adjust the virtual drawbars and select a number of preset mappings between key position and contact onsets. The software logs to disk the positional data from the keys, and the generated audio. An example of the sensor data generated alongside the audio is shown in Fig. 4.12.

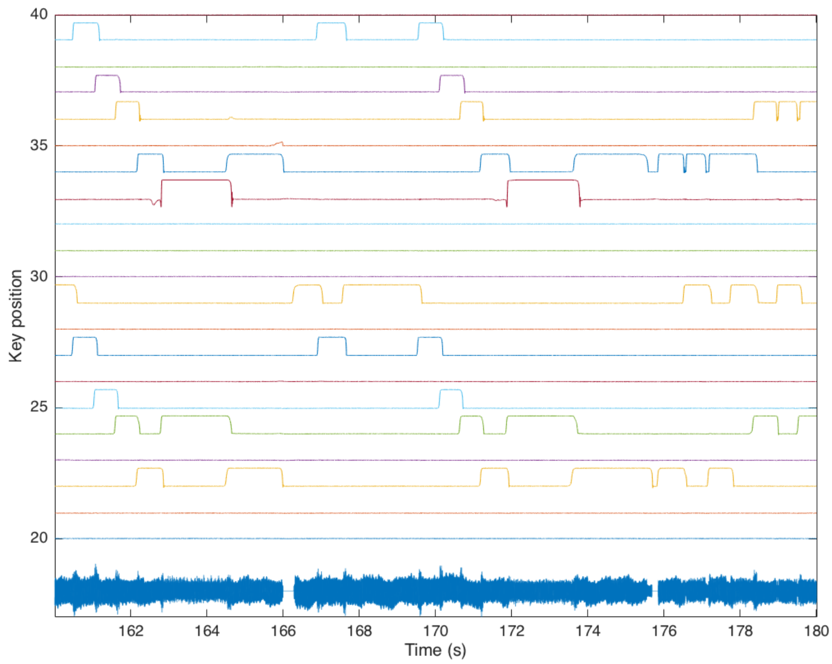


Figure 4.12: Temporal evolution of key position for keys 20 to 40, alongside the waveform of the generated audio signal.

4.6 STUDY DESIGN

In Section 4.1 we stated the motivations that brought us to designing the hybrid Hammond organ we described in the intervening pages. With the present section we start the description of a performance study using this instrument. The objective of this study is to evaluate how subtle changes in the mapping of the key onset affect the actions and the experience of keyboard players. During the study, players compare four different touch conditions on the instrument. We will assess the effect of the key contact behaviour by obtaining subjective assessments and objective measurements.

4.6.1 *Touch conditions*

Four different touch conditions on our instrument were obtained combining the controls described in Section 4.4.2, and were used throughout our study. These are described in Table 4.4.

Touch condition	**Key position to contact triggering**	**Velocity to bounce**	**Triggering range (relative units)**
C1	Single triggering point, no velocity	No	0.33
C2	Single triggering point, velocity	Yes	0.33
C3	Individual triggering point	Yes	0.33 to 0.48
C4	Individual triggering point	Yes	0.19 to 0.75

Table 4.4: Touch conditions on the hybrid Hammond instrument used for the performance study.

C1 has a single triggering point, and its behaviour is completely independent from the velocity of the key press. When the key crosses the triggering point, a randomized contact closing offset value is generated in the range of 1 ms to 5 ms, so that all onsets will be short, and there will not be much variation between repeated onsets.

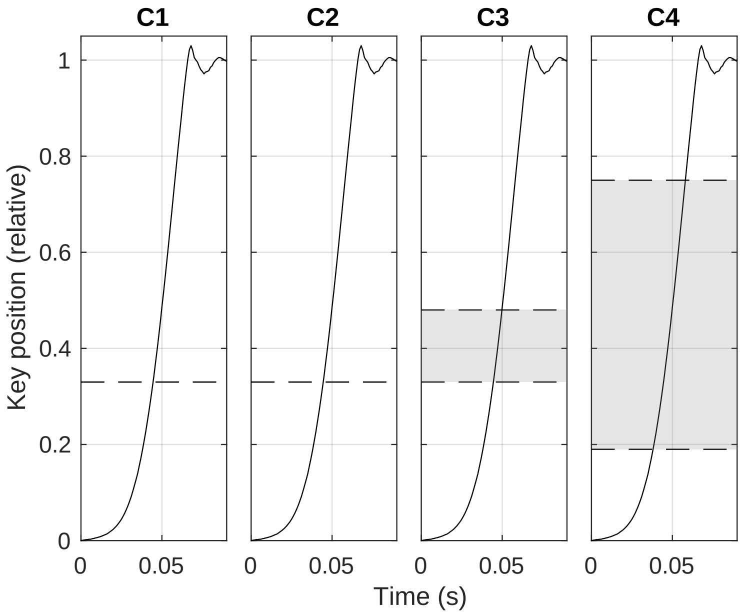


Figure 4.13: Triggering points for the four touch conditions. The grey rectangle represents the interval in which contacts may close, dashed lines represent the minimum and maximum contact closing point. C1 and C2 have a single triggering point for all contacts.

C2 still has a single triggering point, but the contact closing offset of the virtual contacts will depend on the actual velocity of the key press, ranging between 1 ms (for high velocity) and 127 ms (for low velocity). The velocity further affects the amount of bounce of each individual contact.

C3 triggers each contact individually depending on the position of the key, and the triggering point of each contact is assigned randomly to a point within the given range. Additionally, the instantaneous velocity of the key as it crosses each triggering point is used to affect the bounce of each individual contact.

C4 is very similar to C3, with the only difference being the wider triggering range.

We used two popular Hammond emulators as a source of inspiration for designing some of our test conditions. The Clavia Nord C2[17] features, according to their website, *"An ultra fast trigger-to-sound response time"*. Our understanding is that this behaviour is achieved by triggering on the upper contact of the otherwise standard two-contact key-bed it is equipped with. As a consequence of triggering on the first contact, we expect that there will be no velocity control available to

17 https://www.nordkeyboards.com/products/nord-c2

drive the sound engine. The GSi VB3 software[18], provides velocity-sensitive onsets, where the duration of the onset is affected by the velocity of the MIDI input. We have verified this behaviour with an informal test. The VB3 is the sound engine that runs on the Crumar Mojo keyboard[19], so we expect its velocity-sensitive keyboard to affect the sound generation in a similar way.

While we have not been able to gain access to these instruments to verify the behaviours we just described, we have used them as a reference point for conditions C1 (Clavia), and C2 (Crumar). The C3 condition recreates the behaviours we measured on our Hammond organ (see Section 3.3.2.2), however it does not explicitly recreate the effect of percussive touch. C4 is an exaggeration of the behaviour of the original instrument, where the triggering point of the contacts is distributed along a larger portion of the key-throw.

In this paragraph we describe and compare the behaviour of each condition using the terminology we introduced in Section 3.3.2.3. C1 will have very short closing time offset (the time it takes for all the contacts to start speaking), long contact bounce duration (the duration of the bounce of each contact), and average overall bounce duration (the overall duration of the onset transient). C2 will have variable values for each of these, depending on velocity. By pressing the key very fast with C2, an onset similar to those produced by C1 can be achieved. C3 will again have variable values for each metric, and is designed to behave similarly to the measurements in Section 3.3.2.3. C2 and C3 will perform very similarly to each other, except when the press is very slow, because C2 is limited to a maximum contact spread of 125 ms, as explained in Section 4.4.2, or when the press is incomplete, in which case with C3 only some contacts will close, while with C2 all contacts will close. C4 will have longer closing time offset than C3, because of the larger triggering range, and consequently larger overall bounce duration. C3 and C4 both allow for partial presses where only some contacts are triggered, and very slow presses where each contact can be heard coming in individually. A fast press with C4 may achieve a similar overall bounce duration as a slower press on C3. Conversely, when playing with the same velocity on C3 and C4, the overall bounce duration will be longer for C4.

The triggering range for C3 was modelled according to the measurements we performed on our Hammond in Fig. 3.8. The triggering point for C1 and C2 were subsequently chosen as the same as the upper limit of C3, so that the point in which sound production starts is the same for C1, C2 and C3. The triggering range

18 https://www.genuinesoundware.com/?a=showproduct&b=44
19 https://www.crumar.it/?a=showproduct&b=4

for C4 is an extension in both directions of that of C3 and is somewhat arbitrary, as this condition does not have a real-world counterpart.

4.6.2 *Pilot study*

We ran an exploratory pilot study with two participants, designated as PS1 and PS2 below, to inform the design of the main study. We asked the participants to sight read some simple organ pieces, for which we provided the score, using C1, C3, C4, and then to improvise freely while we were switching between the test conditions.[20] An informal open-ended interview followed.

PS1 is a pipe organ and piano player, who took formal classes on both instruments for over 10 years and currently mostly plays organ as a semi-professional musician. Asked about the difference between the test conditions, they admitted that while sight-reading the simple organ pieces they did not quite understand what the difference was between the conditions, though they could feel that something was changing. During the free exploration they described C1 as *"just horrible"*, while the difference between C3 and C4 was *"not so obvious"*. When asked for an explanation as to how the different musical materials elicited different responses, they pointed out that *"when playing larger chords"* – as they did in the free exploration *"you need to synchronize onsets, so you appreciate the "feedback of the gradual onset"*, also *"the musical material in the first part was very easy so the action did not make a huge difference"*. They then explained in detail their experience with different types of keyboard action on pipe organs, comparing tracker action, which has a mechanical coupling between the key and the pipe, and electric action, where a switch closed by a key press sends an electrical signal to an electro-pneumatic valve which opens the flow of air to the pipe. The tracker action, they explained, responds to the force exerted on the key: the player feels a resistance and as soon as they apply a sufficient force, the note onset happens. The electric action is lighter and does not give any resistance, and the player has to *"get an idea of how long it is going to take you to get from top of the key to the trigger point and [...] anticipate of that amount". C1 feels as if you do not have any kind of support"*, similarly to a pipe organ with electric action: *"You are always on edge, you do not know when you are going to trigger a note"*, while *"C3 and C4 give you more support"*. They recognized C3 as being overall more expressive than C1, and C4 to be *"too spongy"*.

20 At the time we had not yet implemented C2

The reported difference between C3 and C1 was mainly about physical quantities: in C1, there is a perceived loss of *"depth"* of the action, which makes the action feel lighter than the other conditions: *"In C3 you feel there is more weight because you need to do all the action of pushing down"*. They drew an analogy between the action of C3 and C4 and that of a tracker organ, which makes it easier to be rhythmically accurate.

PS2 took formal piano classes for 10 years, but currently they mainly play trumpet. When asked about the difference between the test conditions, they immediately reported that they felt a change in the action of the keyboard, as if the keyboard weight was changing between conditions: C1 felt lighter than C3 and C4 felt sloppier than C3. They made some judgements on the sound, recognizing the different "articulation" in the three test conditions.

4.6.2.1 *Implications*

Both participants placed an emphasis on how the "feel" of the keyboard action was the changing parameter between the conditions. In fact, throughout the conversation with PS1 there was no reference to the timbre of the produced sound, except when asked directly. PS1 was able to relate each test condition to the feel of specific keyboard actions as you would find on pipe organs. The outcome of these informal interviews suggests that the mapping between the sound and the physical gesture has an effect on the performer at a subconscious level. Surprisingly, the perceived change is not so much in the produced sound, but rather in the sensation of weight.

The tactile sensation of the instrument while playing has long been recognized as a key factor to the performer's relationship with an instrument [Askenfelt and Jansson, 1992]. In the realm of keyboard instruments, the tactile sensation can contribute to the perceived quality of the instrument [Fontana et al., 2017]. What we observe here is the complementary of what Fontana found, in that the haptic sensation is being enriched by the sound response.

4.6.3 *Procedure*

The participant performs the following tasks, in order:

- **Training** The participant sits on a piano and is given a leadsheet of jazz standard Blue Bossa (see Fig. A.1). They first practice it on their own, and then rehearse it alongside a backing track.[21] By practising on the piano they get to familiarize themselves with the piece on a familiar instrument, without starting to form an opinion on the instrument they will use for the rest of the training.

- **T1, performance** (50 minutes, plus training) The participant sits at the organ, and has to play Blue Bossa along the backing track. They can keep using the leadsheet if they wish to. They play 12 takes, each consisting of 6 choruses, evaluating one pair of conditions in each take. The two conditions are labelled A and B, and a computer screen placed behind the organ shows them what condition they are on. They are instructed to play two choruses of "Theme" (their interpretation of the Blue Bossa theme), two choruses of "Chords" (the way they would accompany someone else's solo), and two choruses of improvised "Solo". As they play, the test condition in use is switched every chorus, so that the structure of each take is as follows:

Chorus	**Material**	**Condition**	
1	Theme	A	
2	Theme		B
3	Chords	A	
4	Chords		B
5	Solo	A	
6	Solo		B

All the possible pairs of conditions are tested twice, once in the first 6 takes, and once in the second 6 takes. A 5 minutes break is scheduled to take place after the first 6 takes. The configuration of the drawbars is set once at the beginning of the task and then switched after take 6, so that each pair of conditions is encountered with a different drawbar setting (preset 1: 888800000 and preset 2: 800008888). For each take on a given pair of conditions, the order in which the conditions are presented is swapped, but the user always sees the first condition labelled as A and second labelled as B. For instance, the first time the participant encounters conditions C1 and C2, the take may start with C2 (labelled A), followed by C1 (labelled B); when later they encounter the same conditions again, the take would then start with C1 (labelled A), followed by C2 (labelled B). After each take, the participant un-

21 The audio of the backing track is taken from the video "Blue Bossa Backing Track for Piano/Guitar Comping"': https://www.youtube.com/watch?v=9p7sa0H1bIY

dertakes a survey where they are asked to compare the conditions A and B (described in Section 4.6.3.1). The orders in which the pairs and the drawbar presets are presented are randomized for each participant.

- **T2, rhythmic exercise** (4 minutes) The participant plays a 6-bar rhythmic exercise alongside a 120 bpm metronome track. The score for the rhythmic pattern can be found in Appendix A.1. Players are instructed to play the pattern above, but are free to choose their own fingering and notes, and can perform the exercise on a single note , or on multiple notes. The exercise is repeated 4 times, one for each condition. No survey data is collected.
- **T3, free exploration** (no limit of duration, we suggest between 10 and 20 minutes) This task is aimed allowing the performer to form their subjective assessment of each test conditions. They are given access to a MIDI controller (see Section 4.6.3.3) so that they can switch between conditions and adjust the drawbar levels at their own will. During this task they are instructed to fill in a survey to evaluate each condition. The survey, also described in Section 4.6.3.1, contains questions about all the conditions on a single page, and they are free to fill it in at any point during the task.
- **Interview** a semi-structured interview follows (see Section 4.6.3.2).

4.6.3.1 *Questionnaires*

During the test, participants are asked to evaluate touch conditions on the following descriptive terms: Weight, Responsiveness, Temporal control, Expressiveness, Fatigue, Difficulty, Realism, General preference. The graphical user interface consisted of a computer program displaying slider inputs using a Continuous Category Rating scale (CCR), as recommended by ITU-T P.800[22].

In T1 the questionnaire asks to comparing the two touch conditions using during the take in relation with each other. Quantitative answers in T1 use a dual-sided slider interface, which initially presents itself with the cursor in the middle, as shown in Fig. 4.14, and results are stored with a numerical range from 1 (leftmost) to 101 (rightmost). The central position of the cursor corresponds to the value of 51, but if the cursor is left untouched, a placeholder value is stored instead. As an example to illustrate the phrasing of each question, we use the value "Weight". The question would be: "How did you find the weight of the keyboard action under each setting?", and the two labels at the two ends of the slider would

22 https://www.itu.int/rec/T-REC-P.800-199608-I/

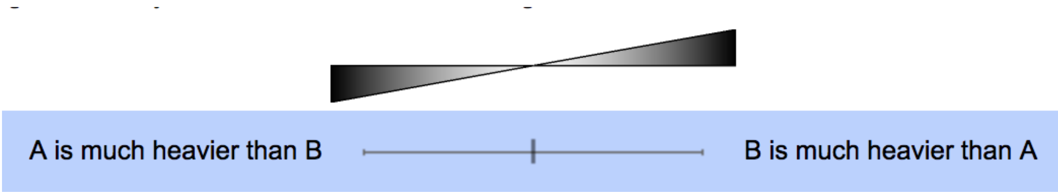


Figure 4.14: Example of dual-sided slider interface as used for the questionnaire in T1.

be "A is much heavier than B" on the left and "B is much heavier than A" on the right. The complete list of questions and labels used can be found in Table A.1.

The questionnaire in T3 asked the participant to evaluate each test condition on its own merit. The quantitative questions for T3 used a slider such as the one in Fig. 4.15. No cursor is visible initially, and the stored numerical values range from 1 (left) to 101 (right). As each condition is now evaluated individually, the phrasing is adjusted accordingly. Again, in the case of "weight", the question would be "How did you find the weight of the keyboard action under setting Cx?", and the labels at the two ends of hte slider would be "very light" on the left, and "very heavy" on the right. The complete list of questions and labels used can be found in Table A.2.

Figure 4.15: Example of the slider interface used for the questionnaire in T3.

4.6.3.2 *Interview*

A semi-structured interview was conducted at the end of the study, after T3 and the associated questionnaire were completed. We first asked about their impressions on T1, then what they discovered in T3, and how that relates back to their experience in T1. Only then we revealed to them what the differences were, and continued the conversation in a more free-form way. Throughout the interview process, the participant could go back to the keyboard as needed, to help with the explanation and explore further.

4.6.3.3 *Setup*

A sketch of the layout of the room during the experiment is in Fig. 4.16. The investigator's and the participant's computers are connected over a point-to-point ethernet connection. The second display of the participant's computer is located

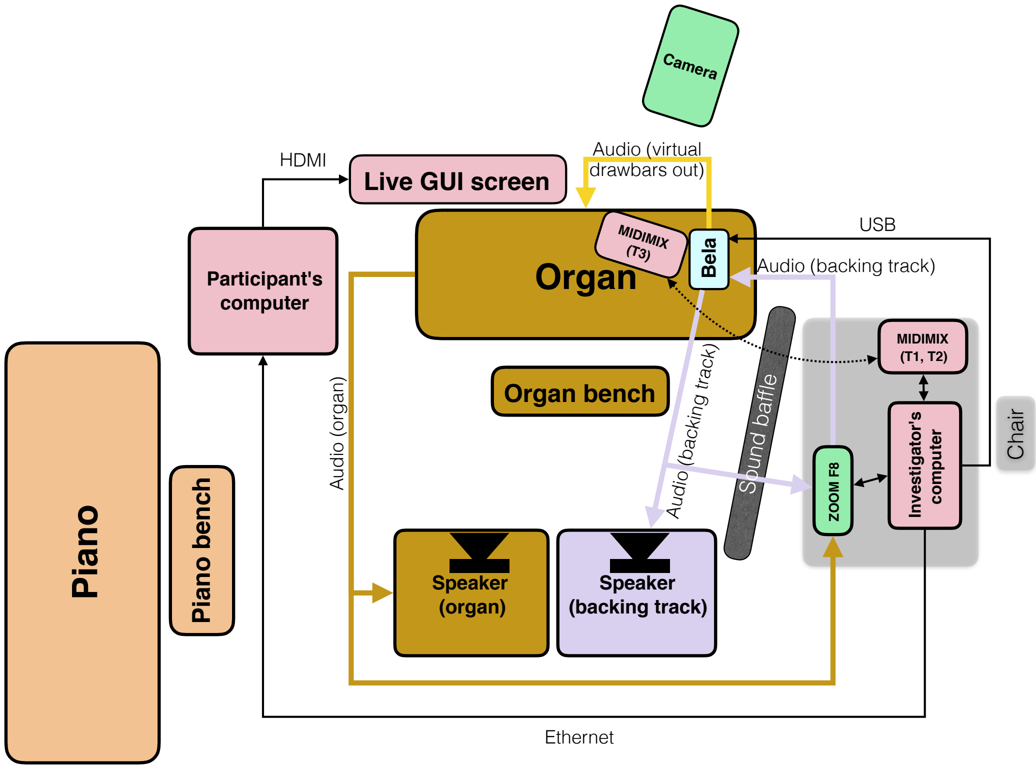


Figure 4.16: Sketch of the layout of the room during the experiment

behind the Hammond and is used to display the live GUI, indicating the current musical material and condition. The main monitor of the participant's computer, and its trackpad and keyboard, are used by the participant to fill in the questionnaires in between takes. To get access to the laptop, the participant has to stand up from the organ bench and take a couple of steps to their left. As the performer plays, the sound of the organ is played back through one RCF ART412A placed behind them slighty to their left, while the backing track (for T1) and the click track (for T2) are played back through another speaker of the same model placed behind them, slightly to their right. A sound baffle (height 1.5 m) is placed between the participant and the investigator's table until the end of T3. The investigator can still see through to the participant, but the latter cannot see the MIDI controller and the investigator operating their computer.

On the investigator's computer, the Reaper[23] software is used to control the playback of the backing track. Audio goes out through a Zoom F8[24] used as an audio interface. Besides outputting the audio, Reaper also sends MIDI messages

23 http://reaper.fm
24 https://www.zoom.co.jp/products/handy-recorder/zoom-f8-multitrack-field-recorder

to a Pure Data (Pd)[25] patch running on the same computer at the beginning of each chorus of the track. Pd maps the position in the track to the active condition, and sends this as a MIDI message to the Bela board over the USB port. Reaper additionally sends MIDI messages at every chorus change to a *node.js*[26] application, which in turn issues an HTTP GET[27] request to the participant's computer, where a *node.js* server updates the live GUI[28]. A MIDI controller (Akai MIDIMIX[29]) is also connected to Pd on the investigator's laptop, and can be used to manually control the currently active condition and the drawbar setting on Bela. Before T3, the MIDI controller is hidden from the sight of the participant. During T3, it is given to the participant to allow them to manually switch condition and sound. Bela receives the backing track from the computer via analog audio (using a DI box), and loops it back out through the right output, while the organ sound is sent out of the left output. The two outputs from Bela are then split and sent both to the dedicated speakers and back into the Zoom F8 for recording on the host. On Bela itself, the audio input is logged to disk alongside the generated audio and the sensor data. This allows to obtain sample-accurate logs of the audio, sensors and backing track for analysis. A video camera placed in front of the Hammond is recording video and audio (through a Earthworks QTC40[30] omnidirectional microphone) of the whole session.

4.7 RESULTS

A total of 10 participants took part in this study, all of them with a professional or semi-professional music career, 5 of them experienced Hammond players and 5 of them experienced piano players. Many of the Hammond players, as it often is the case, also had extensive piano training.

They were paid a professional rate for their time. All the Hammond players were familiar with the musical piece used in the performance section (T1), and spent little or no time in the initial rehearsal session, while 3 of the piano players spent over 20 minutes rehearsing the piece. Participants spent between 5 and 35 minutes on the free exploration section (T3).

25 http://puredata.info/
26 https://nodejs.org/
27 https://www.w3.org/Protocols/rfc2616/rfc2616-sec5.html
28 The client/server node application, also developed by the author, is available at https://bitbucket.org/giuliomoro/luana9001
29 https://www.akaipro.com/midimix
30 https://earthworksaudio.com/products/microphones/qtc-series/qtc40/

In the remainder of this chapter, we refer to participants in the form $Px(y)$, where x is the participant number (1 to 10), and y is h for Hammond players, and p for piano players. We refer to each individual take in the form *Cx-Cy,z*, where x and y are the touch conditions in the order they are presented, and z is the preset sound used during the take.

In Section 4.7.1 we present the data from the numerical questions of the questionnaire. We then collate the responses from the questionnaire and the interviews to present the overall findings in Section 4.7.2, and more in-depth profiles for each participant in Section 4.7.3. We conclude presenting the analysis of the sensor data in Section 4.7.4.

4.7.1 *Questionnaire*

A significant fraction (36%) of the quantitative questions for T1 were left unanswered (the cursor was not moved from the initial position and was left in the middle). Table 4.5 reports absolute mean deviation[31] and number of unanswered questions for each participant. When answers were given, they would often deviate only slightly from the central value of 51, with a mean deviation of 11.7 across all participants.

When participants did answer the questions, they would often be inconsistent when evaluating the same pair of conditions in two different takes. Fig. 4.17 shows whether there has been agreement between the two evaluations of the same pair, for each question. If a question was answered in at least one of the two takes, then we consider the player to agree with themselves only if they moved the slider's cursor towards the same condition in both cases, regardless of how far from the centre position they went. If the participant did not move the slider from the centre position for both question, it is marked "Unanswered". The "Overall (average)" plot represents the arithmetic mean of the others, once the unanswered questions have been taken out. Only P2(h), P5(p) and P9(h), were consistent across all the questions for at least one pair of conditions, as denoted by the white boxes in the "Overall (average)" plot, with P9(h) being perfectly consistent on 3 pairs of conditions. When looking at individual metrics, "Weight" is the one with the largest number of consistent answers (17), and "Realism" the one with the fewest (9). Overall, out of a total of 480 pairs of responses, only in 108 cases there was consistency be-

31 This is the mean of the deviation from the central value of 51, and not the mean of the deviation from the mean.

Participant	Unanswered (n = 96)	Mean deviation
P1(p)	23 (24.0%)	13.7
P2(h)	27 (28.1%)	16.3
P3(h)	67 (69.8%)	3.5
P4(p)	57 (59.4%)	11.0
P5(p)	7 (7.3%)	16.6
P6(p)	15 (15.6%)	16.5
P7(p)	53 (55.2%)	8.2
P8(h)	28 (29.2%)	17.3
P9(h)	24 (25.0%)	10.4
P10(h)	50 (52.1%)	3.9
Total (n = 960):	351 (36.6%)	11.7

Table 4.5: Summary of unanswered questions and mean absolute deviation for T1

.

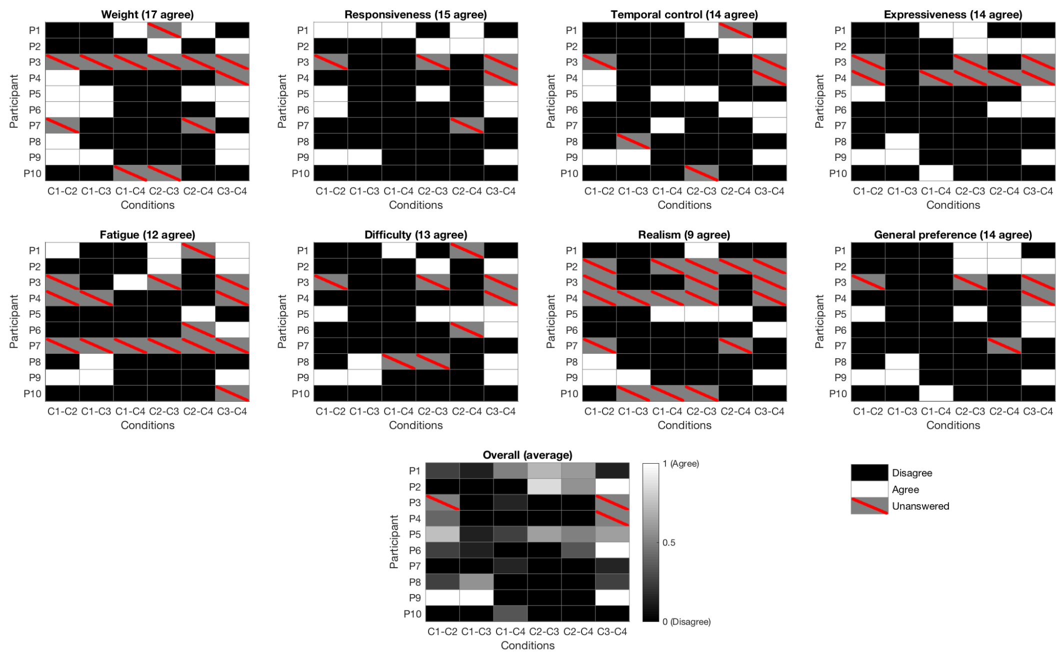
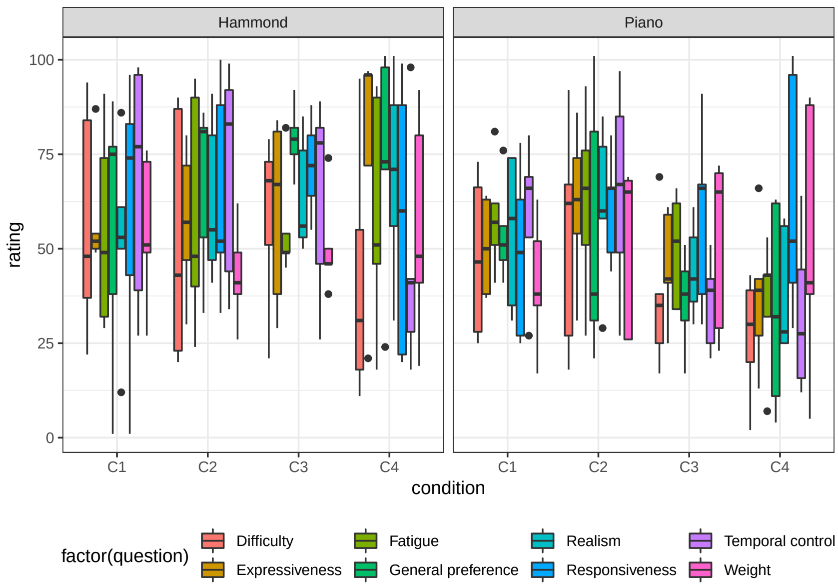


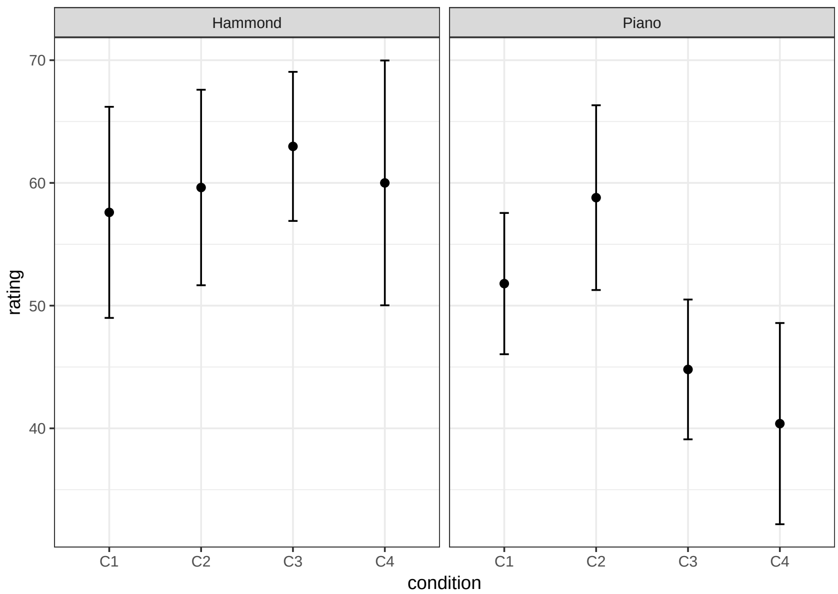
Figure 4.17: Agreement between questions on the same pair of conditions in T1.

tween the first and second take (22.5%). The computed Pearson's product-moment correlation coefficient across all questions that have been answered in both takes, is $r(472) = -0.155$ ($p = 0.0265$), showing a small negative correlation, where we would expect a strong positive correlation if the answers had been consistent.

There were only two unanswered questions, both from P4, the one about difficulty of C1 and the one about temporal control of C4. We analysed the quality ratings from the survey in T3 to assess statistical differences in subjective evaluations of the different conditions. The *fatigue* and *difficulty* questions from the survey have a clearly negative connotation, therefore prior to performing the analysis we reversed the ratings for these questions, so that the new rating would be $\mathtt{newRating} = 100 - \mathtt{oldRating}$. For instance, a value of 0 would become 100 and viceversa. The connotations of the *weight* is more ambiguous, as some players may associate higher weight with a positive sensation, while others would do the opposite, as it also emerged from the interviews, so we left it untouched. For each *group* (Hammond players, non-Hammond players) we fitted separate linear mixed effect regression (LMER) models with fixed effects of *question* (weight, responsiveness, temporal control, expressiveness, fatigue, difficulty, realism, general preference) and *condition* (C1, C2, C3, C4), and random intercepts for each *participant*. Models were fitted using the *lme4* package for R [Bates et al., 2015]. We conducted a full factorial Type III ANOVA on each LMER model, with Satterthwaites's degrees of freedom approximation from the *lmerTest* package [Kuznetsova et al., 2017]. For the Hammond *group*, we found no significant effect of *question* or *condition*. For the non-Hammond group, we found a significant effect of condition ($F(7, 143.02) = 6.67, p < 0.001$), and no effect of question. Fig. 4.18b shows that non-Hammond players rated C2 higher on average than all the other condition. A Tukey post-hoc analysis on condition shows that the only statistically significant differences are those between C2 and C4 ($p < 0.001$) and between C3 and C4 ($p = 0.0079$).



(a)



(b)

Figure 4.18: (a) shows the median and interquartile range of the ratings from all responses to the questionnaire in T3. (b) shows the mean averaged across all the ratings and confidence intervals .

Part.	Performance task (T1)	Free exploration (T3)	Favourite condition
P1(p)	Condition changes affect the state of "flow" of their performance	C4 feels unreliable and does not make them feel "safe"	C2, as it feels more natural
P2(h)	Difference in weight and articulation	The articulation is the most prominent difference	C4, as it has the most articulation
P3(h)	Difference in latency, the are distracted by the preset in use	They identify the triggering points are changing	C3, as it has the most realistic triggering
P4(p)	Difference in responsiveness	Difference in responsiveness	C1, as it has the right responsiveness
P5(p)	Differences in weight, responsiveness and tuning.	Difference in the transition between the notes	C2, as it is the easiest, most comfortable and most playable
P6(p)	Differences in timbre, weight and articulation	Differences in the weight of the action	C1, as it feels the heaviest
P7(p)	They have to press differently when playing chords depending on condition	No notable difference	C1, but they are all very similar
P8(h)	Differences in loudness, keyclick	Differences in the sound of the attack, and in the contact triggering	C4, as it is more authentic
P9(h)	Differences in weight	Differences in the attack sound of C4 and in the weight of the action	C1, as it is the most realistic
P10(h)	Differences in action weight, triggering point, articulation and sustained sound	Only difference is in the attack sound and triggering points of C4	C3. C1, C2 and C3 are very similar and better than C4 because of temporal control.

Table 4.6: Summary of the outcomes of the interview and survey for each player

4.7.2 *Interviews and textual responses*

A summary of the findings for each player, compiled from the interview data and the questionnaires, is presented in Table 4.6. One of the most recurrent comments was that participants struggled to find major differences between conditions during T1. Many reported they were unsure about the answers they gave in T1, and that some differences they perceived were very slight or may have even been the result of self-suggestion. P10(h) would often lead their textual answer with *"I thought [...]"*, denoting insecurity and further adding comments such as *"I may also be imagining this. Still unsure whether there was definitely a difference"*. P9(h) commented *"All of these things are incredibly subtle, as it has to do with when you first strike the note. You start questioning what you are hearing/feeling when doing 12 takes"*. During T3, P5(p) commented *"It's difficult to perceptualise, it's not like [the conditions] are very distinct"*. Occasionally, participants reported as changing some aspects of the sound that were not actually being modified between conditions, for instance the volume or harmonic content of the sound, or the tuning of the instrument: P1(p), in C3-C4,2: *"It sounded like a change in tone to a more subtle, muted tone in [C4]"*; P8(h) in C2-C3,2 *"There is a volume difference between the two settings"*, P5(p) in C2-C4,1 *"[C4] felt [...] quieter in upper registers"* P5(p) in C4-C1,1: *"subtle - if any difference. [C1] seemed more detuned / out of tune"*. P3(h), who scored the smallest absolute deviation and the largest number of unanswered questions in T1, had prior knowledge of the experimenter's previous work on latency, so after answering none of the questions in the first 3 takes, they started relating their observations to latency in take 4 (C1-C4,1): *"hard to say. maybe [C4] was more laggy than [C1]"*.

Most participants could, however, recognize some differences between conditions in T3. The criteria used in the assessment of the conditions were varied, with each participant referring back to their own experience and training, and their personal preference. The 5 piano players all rated as their favourite one of C1 or C2, while 3 out of 5 Hammond players preferred C3 or C4. P3(h) and P8(h) systematically explored in T3 the triggering behaviour of all the conditions, guided by their a-priori knowledge of the Hammond, while P4(p), P6(p) and P10(h) discovered accidentally while playing chords that, if the keys were not held fully down, there could be an unexpected re-triggering of the contacts in C4. Of these, only P10(h) systematically explored the transient behaviour of C4 during T3, and they did not notice that there was a similar effect in C3. Even when strong preferences were developed during T3, comparing these with the earlier comparisons in T1 often shows inconsistent results.

The perceived weight of the keyboard was a recurrent topic of discussion. P4(p), P5(p), P6(p) and P9(h) reported that the weight was was the main factor that was changing between conditions. The way weight was perceived was very subjective, with C1 and C4 being the conditions more often described as the heaviest, despite them being diametrically opposed in terms of behaviour. P1(p), P2(h) and P7(p) rated C1 as the heaviest, while it was rated as the lightest by P3(h), P9(h) and P10(h). Conversely, P3(h), P4(p), P6(p) and P9(h) rated C4 as the heaviest, while it was rated as the lightest by P1(p), P2(h), P7(p) and P8(h). P6(p) tried to rationalise the weight sensation: *"I rationally guess it's only the sound that changes and not the action, but it really felt like the action was changing, it felt heavier"*. P5(p) outlined that in some of the conditions, the weight sensation is strong: *"[C4] felt a bit more resistance, which made it harder, but interesting to play"*. P8(h) felt some changes in the weight during T1, but not in T3. P9(h) mentioned multiple times weight differences during T1, and later explained that sometimes they perceived the triggering point was changing. P5(p) also found that C3 *"feels harder, but also it's not comfortable, it's not feeling right"*, suggesting that they were perceiving both a change in the weight and a change in the sound, as two distinct effects. P1(p) and P10(h) were the only players to mention a change in triggering point during T1. While acknowledging that they were feeling differences in the initial part of the key gesture, P1(p) realized that such change was not in the weight, but in the *"sensitivity of how much I pressed"*, and P10(h) explicitly mentioned in the interview the triggering point and its effects: *"maybe the trigger point was different. Not so much about weight, although the trigger point may affect your perception about the weight of the keyboard"*. P3(h) rationally ruled out the possibility that the weight would be changing, being aware that it would be extremely complicated for the experimenter to be changing the mechanics of the action, however they reported changes in the perceived latency of the sound.

The only participant to regularly mention the sound of the onset transient during T1 was P10(h), while P2(h), P6(p) and P8(h), mentioned it only occasionally. In C3-C2,1, for instance, P6(p) commented *"[C3] was preferable, especially in the chords, but there was something different about the attack of that sound that I didn't like as much"*. In the free exploration task T3, the articulation was mentioned more often. P2(h)'s preference was strongly guided by the sound of the *"articulation"* of C4: *"It sings more"*. P9(h) reported a *"higher tone"* perceived during the attack when using C4 as the only difference in sound among the conditions. In the early part of their exploration, P8(h) would describe the sound of the instrument as varying between C1's *"flat"* to C4's *"dirtier [...] the sound isn't clear"*. P5(p)'s focus in T3 was on the *"flow"* when playing a melody, without mentioning any differences in the sound, except for *"in C4 there is a real click, a slight delay to come in or something. A real clicky sound"*. P1(p), P4(p) and P7(p) made no reference to the attack sound.

Specifically on the topic of rhythm, P1(p) reported that with C4 it is *"harder to get the rhythm"*, and the asynchrony of the closing contacts was also flagged by P6(p) as a, perhaps subconscious, cause of annoyance. By contrast, P3(h) found that it was *"harder to play rhythmically with [C1]"* during T1 in C4-C1,2. P8(h), the only one to report differences in T2, commented on the *"[in T2] one of the setting was easier to play than the others, and the timing was more accurate"*.

It was reported that some aspects of the test setting during T1 were distracting. P3(h) commented that the drawbar presets chosen for the experiment were a source of distraction, as they did not feel at ease playing with the resulting sound. In particular preset 1 is, in their experience, only suitable to be played with a fast Leslie, while preset 2 would normally be accompanied by harmonic percussion, or, alternatively, by the addition of the 9th drawbar (i.e.: 888800008) P8(h) would have wanted full control on the drawbars during T1, to dial in their own sound: *"I couldn't choose the sound of the drawbars. It was a full sound, a flat sound, and so not good to play chords"*. P4(p) commented on preset 1 as being "atonal" and unsuitable for a chord section, while P5(p) described it as "inharmonic". P3(h) and P5(p) also commented on how the backing track chosen was not the most suitable to explore the differences between the conditions and suggested that a slower, more "ambient" one would be best suited. To various extents, these distractions were seen by the participants as making it even harder to distinguish the subtle differences in T1.

4.7.3 *Participant profiles*

P1(p) is a piano player with degrees in composition and performance, who mostly plays popular music on piano and digital keyboards. In T1, they reported to be more "comfortable" playing with some conditions, and they noticed this mostly when changing from a "more" to a "less" comfortable one, but they would not notice when changing *into* a more comfortable one. *"Some places were definitely more comfortable. I wouldn't feel the change into them, but the change out of them"*. When they found themselves in a state of flow, a condition change to a less comfortable one would disrupt their flow. They noticed that C1 would trigger unwanted notes while resting the fingers on them, although they also mentioned that the instrument has, in general, a "high" triggering point. *"[The differences between conditions]* *are not drastic, but they really affect you. It's the nuances, when you want to really lock-in, it's something like having slippery fingers"*. In T3, they reported that with C4 it is hard to get the rhythm right, and it feels more difficult, *"Too hasty, jumps from under my fingers and takes away my freedom. Annoying"*. By contrast, the only comment about responsiveness in T1 was made about C1 in C2-C1,2: *"the freedom of idling fingers on keys was lost as it produced notes when not wanted"*. Their favourite setting in T3 was C2, because *"it feels easy, natural, easy to sync with"*. They made no comments on the sound differences between conditions.

P2(h) is a Hammond player who has been playing tonewheel Hammonds for over 25 years, and digital emulators for the past 15 years, and mostly plays blues, jazz and rock. In T1 they could not really understand what the differences were, as they were just enjoying playing, *"you are in the groove trying to make great music"*. Their experience playing several Hammonds over the years makes them aware that every instrument responds differently, in terms of action and sound, and so they have learned to adapt and not worry so much about the minute differences. In particular, they pointed out that while playing chords they would not notice any difference at all. In T3 they strongly preferred the sound of the articulation of C4, *"I can hear it straight away"*, *"C4 it's more like the attack of a sax or a trumpet. C1 is like the sound of a synth that is always 127[MIDI velocity]"*, *"[C4] picks up the little busbars variations, which the [digital] clones cannot do [...] there is a difference in sound every time you hit it"*. Despite showing awareness of the differences in the produced sound, they did not realize that the triggering points varied between conditions. They would occasionally refer back to the sound of a real Hammond: *"C4 makes the Hammond sound better, it's easier for those contacts to be heard"*, however they also admit that theirs is a very subjective opinion: *"C4 may be too articulate for some players, they wouldn't like it [...] C1 is more even [...] Some Hammond players would prefer that"*. Reflecting back to T1 during the final interview, they thought that C4 was their preferred setting even back then *"While playing, there were moments where I was 'I love this!' and I am pretty sure it was on the C4 setting"*. Additionally, when asked if they felt the different settings would have changed the way they played, *"When soloing and you hear the evenness of tone you would play slightly differently than on the C4, where I would feel more inspired. I am pretty sure when it was on the C1 setting, I would haven't felt like that on the solo"*. However, this is not always consistent with their responses from the questionnaire in T1; for instance, while in C3-C4,1 *"[C4] felt great easy to control dynamics"*, in C4-C1,2 condition C1 was rated as more articulated, and in C1-C4,1 very little difference was found. C1 was also rated as *"is great very articulate"* in C2-C1,1. They thought that they were playing the real Hammond sound all along, and that some adjustments were being made to the physical response of the keys.

P3(h) has been playing jazz Hammond organ for 30 years, both on toneweheel Hammonds and several digital emulators. The participant searched the internet for the name of the investigator before the experiment, and found a paper related to latency [McPherson et al., 2016], and as a consequence of this, they came in thinking that we would perhaps be studying the effect of latency. In T1, after struggling to detect any differences for the first few takes, they started reporting slight latency differences between conditions. For instance, in C1-C4,1 *"hard to say. maybe [C4] was more laggy than [C1]"*, in C2-C4,2 *"at first I noticed a slight delay to [C4]"*, in C1-C3,2 *""slight lag on [C3]?"*. They felt some slight changes to the weight of the action, but achieved the rational conclusion that the mechanics of the action could not be changing, and therefore they inferred that it could be the triggering point or the latency. During the free exploration in T3, they tried to slowly depress the key and observed how the triggering point changes between conditions. They identified C3 as the one that behaves most closely to a real Hammond, and, for this reason, it is their favourite, with C2 a close second *"[C2] has trigger depths in the range I'd expect from a vintage console"*. C1 is depicted as unrealistic, because all drawbars trigger high, but it is easier to play. C4 was the least favourite because of the excessive depths for the triggering points of the last drawbars. They stated that the different conditions would probably make a difference in a solo performance, but not so much in a band context: *"With a backing track and drums, the attack of the Hammond is lost"*.

P4(p) has been playing piano for 35 years and mostly performs classical and folk repertoire. In T1 they struggle to perceive differences, but they commented on the responsiveness of the action in a number of occasions, for instance: C1-C2,1 *"perhaps [C2] is more responsive than [C1]"*, C3-C2,1 *"I felt [C3] was a little heavier than [C2] on that trial"*. During the interview they would clarify *"By responsive I mean that you just literally have to touch the key and there is the sound"*. When they got to T3, they more directly referred to the weight of the key as the changing factor: the *"heaviness of the touch [...] did change in the conditions"*. They noticed that the drawbars trigger at different points one condition, but they did not explore that systematically across the conditions. Their favourite setting was C1 (*"Light touch - pretty responsiveness [sic]"*), contrasted with C3 (*"much more responsiveness, timing harder as a result"*), and C4 (*"much harder to play, as it's too responsive"*). Referring back to T1, *"the times when I said that I literally just had to place my hands on the keyboard and it was very responsive, that was probably towards C3 or C4"*. However, this is in plain contrast with their answer on C3-C1,2: *"I really noticed the responsiveness of [C1] this time"*. They also commented that preset 1 is not a sound that they would use for a chord section.

P5(p) is a piano player with 35 years of experience, mainly playing pop and jazz at present. In T1 they reported differences in weight and responsiveness, e.g.: C1-C3,1 *"found [C1] and [C3] both difficult, heavy and not as responsive as other takes"*, C3-C4,2 *"[C4] seemed slower action, heavier and more tiring"*, C3-C1,2 ""*[C1] felt heavier, and less responsive, but I still preferred it"*, C1-C4,2 *"[C4] felt much easier and more comfortable to play"*. In T3 they concentrated on the sound and "flow" and the perceived resistance of the keyboard. All conditions were initially found very similar in terms of sound, but C2, which they eventually ranked as their favourite, stood out, although without a conscious understanding of the timbral differences: *"there's something a bit warmer about the sound that feels just a little bit more flowing"*. Contrast that with C1: *"C1 is nice and comfortable, but there is something that doesn't quite ... the timing just feels a bit quick and a bit sudden, so it's quite difficult to have a flow"*. In general, they like more keyboards that offer resistance *"The resistance gives you the feel that you are playing an instrument, as opposed to just pressing keys, as if you were on a computer"*. In light of this, they responded positively to C4 *"quite realistic, felt a bit more resistance, which made it harder, but interesting to play"*, an interesting contrast with C3: "slower, less flowing and not very comfortable".

P6(p) is a classical and contemporary piano player with over 30 years of experience. In T1 they report a number of times some significant differences between the weight of the keyboard action: C1-C4,2, *"I prefer the action in [C4] although this could be due to my training as a pianist!"*, C3-C1,2 *"In this case the lighter setting [C1] seems to somehow work better in terms of control"*. Later in the task, they would report differences in sound as well: C4-C2,2 *"Personally, [C4] had a better deeper tone for me"*, C2-C3,2 *"[C2] had a 'rounder' tone (and 'feel')"*. Often, they would report on sound in association with weight: C4-C3,1 *"Interesting. [C3] was much nicer to play and 'felt' heavier, but I am now convinced that actually it's because I prefer the tone, which was less 'brash' and warmer"*, C1-C2,1 *"[C1] sounded (and subsequently felt) nicer / was more enjoyable"*. In the interview, they openly stated their perceived relation between sound and perceived weight: *"tone color really affects the feel of the keyboard. Sometimes I felt that the action was really heavy or really light, but it was just the tone control impression"*. Over the course of the experiment, they would refer to the "tone" or "color" of the sound with terms such as "warmer", "richer", "darker", "less bassy", "bite". Despite commenting on the attack sound in C3-C2,1 (*"there was something different about the attack of [C3] that I didn't like as much"*), they did not explore the attack sounds much in T3, and by the end of it they still struggled to pinpoint what the changes in the sound were: *"maybe the harmonic content?"*. In terms of action, some settings felt more "reliable" than others. Their favourite condition was C1, as it was the heaviest, and *"Rather rich, has some 'bite' to it"*. They would find C2 and C3 very similar to each other: *"Darker sound, but strangely doesn't 'feel' as if I can 'sink' as much [as C1] into the key, less bite"*. C4 was their second favourite, as it was perceived as the second heaviest (after C1). After we explained to them the difference between the conditions, they found a plausible explanation as to why they preferred C1 the most, relating to their piano training: *"when we play piano, if things are not together, it's not good. So if you play a chord, you need to work very hard to have it all play at the same time. My feeling of control [with C1] is connected to the fact that I want to hear things sounding together. If when I press something, it comes all immediately, then that is good, it gives more feeling of control and depth and good tone"*. Additionally, C1 has the quality of being repeatable: *"I can do fast attacks that are all the same. [...] C4 sounds more feeble, is not as predictable"*. They also found that C4 requires more control throughout the duration of the note, as a note could be accidentally re-triggered (at least for some contacts) if it is not held all the way down firmly, which is yet another aspect of keyboard playing that piano players tend not to have to think much about. Looking back at their questionnaire from T1, we found that only in one instance (C4-C1,1) C1 was perceived as the heaviest of the two conditions being compared, which is a marked contradiction with their statements in T3 and in the interview.

P7(p) is a piano player with classical training but mostly performing popular music. Throughout the study, they did not report many strong opinions and they struggled to understand T1, where they found no notable differences among conditions. In T3, they reported C1, C2 and C3 to be very similar: *"quiet [sic] hard to control the sound and not good for many style of music"*. C4 was found to be harder than the others in terms of touch and control. They stated that after having spent 10 hours a day for one year practising on a hybrid grand piano, they can now easily play on any keyboard without being very sensitive to differences in the action. The most noticeable difference they found between the conditions, as they reported in the interview, was in T1 during the chords sections: they would not necessarily press the keys all the way down, and they noticed that this would make a difference between different settings. They did not investigate this aspect during T3.

P8(h) has played on tonewheel organs and analog and digital Hammond emulators for the past 10 years, mostly performing jazz, funk and rock music. In T1 they reported differences in loudness between conditions (C2-C3,2), and a mention of *"more keyclick"* in the C3 condition (C3-C2,1). In T2, they found that with one setting it was easier to obtain an accurate timing, and after the exploration in T3 they suggested that that may have been C2. In T3 they quickly came to the conclusion that C1 had a more flat sound while C4 had *"a nicer, dirtier sound"*. Progressing in the exploration, they found out about how the triggering points differ between conditions. They ended up classifying C4 as the most authentic one (and their favourite), while C1, the least favourite, was described as being unresponsive to how fast you press: *"the worst emulator [...] full of power, but unrealistic, completely without control* ", C2 was ranked as the second best, because it is the easiest to play, and more precise, the way a good emulator would be. C3 was regarded as a poor attempt to emulate a real Hammond, thus falling short of both C4's realism and C2's playability. During T1 they felt changes in keyboard weight between conditions, but not during T3. While C4 was for them the clear winner in T3, during T1 it was rated best in terms of general preference only once (C4-C2,2). They commented that T1 was a demanding task, with many factors that were distracting them from the comparison between conditions: they were playing an instrument they never used before, they were not allowed to choose drawbar settings that would make them comfortable, and they had to follow along the backing track.

P9(h) is a jazz Hammond player with over 10 years of experience playing tonewheel Hammond and digital emulators. In T1 they mostly commented about subtle weight differences, but not in a consistent way. For instance, they found C4 *"harder"* than C2, C2 harder than C1, C1 harder than C4. In the interview they clarified that this was actually due to a perception of the triggering point moving in the key throw, and that they would adjust their playing accordingly. In T3, only C4 was found to be sonically different from the others: *"there was some higher tone [in C4] that was creeping in"*. Relating back to T1, they stated that what they had described as "harder action" in T1 was probably C1, and the lightest was C3, however this is not consistent neither with their responses on the questionnaire for T3, nor with those for T1. Referring to their experience on several Hammonds, they would say that the variability is such among different instruments that it is hard to think of a reference one, in terms of realism. After we explained to them what the differences were between the conditions, they admitted they had not thought about the differences in terms of triggering points, *"It's just a minute sensation when you play, it's gone in one microsecond"* and genuinely believed that something mechanical about the keyboard was changing. They concluded saying that the easiest settings to play are C3 and C4, as they are most realistic, however in the questionnaire during T3 they had previously selected C1 as the easiest to play and most realistic one (*"Quite like what I'm used to for a hammond organ"*), and C4 as the hardest of all settings.

P10(h) is a Hammond player with 10 years of experience on tonewheel Hammonds and clones. In T1 they noticed slight differences in the articulation: in C1 during C2-C1,2 they reported a louder keyclick, then in C1-C2,1 *"I found that the [C1] setting 'spoke' better than [C2] and seemed to be more responsive when soloing"*, in C3-C2,1 *"I thought there was a slight registration change for [C2] which facilitated better sounding fast lines"*, and in C4-C3,1 *"I thought I noticed a mechanical spring back type noise during fast lines in [C4] around the middle keys that I didn't notice during [C3]. I'm wondering if [C4] is real hammond and [C3] is a modelled equivalent?"*. They commented on triggering point during C1-C4,2 (*"During the soloing I imagined I could feel that the trigger point of the sound was higher in [C4], thus making it slightly more of a physical effort to play"*), and in several other takes: higher triggering point in C4 during C2-C4,1, higher triggering point in C3 during C1-C3,1. Additionally, some slight difference in the overall sound were reported, in C4-C2,2 *"I thought there might have been slightly more harmonic leakage under B"*, in C1-C4,2 *"I thought, during the comping and melody I could detect more mid frequencies in the sound during [C4]"*. They reported during the interview that in T1 the only concrete observation was an an unexpected behaviour in C4: *"I realized there was some sensitivity around this area, while my notes were resting on the keys, and I thought 'maybe this is the real hammond'. It only made a difference while resting my hands on it, not when playing a melodic line. If you play a line it doesn't really affect that so much, but if you are playing something really slow it could affect the timing"*. In T3, they further explored this and discovered that in C4 the contacts actually close at different points in the key-throw. C1 and C2 and C3 all sounded the same to them, with the only difference being C4. Despite being a professional Hammond player, P10(h) never owned a tonewheel Hammond, but regularly plays them in recording studios and live concerts. As such, they never got to notice the individual triggering behaviour on the Hammond, and were very surprised by discovering it in C4: *"it's potentially a quite expressive effect to have, it would be quite nice, it could be used interestingly in* *slow-tempo ballads, but it's harder to control tempo with it during normal playing [...] If I had a real Hammond at home I would have probably noticed that".*

4.7.4 *Sensor data*

4.7.4.1 *Rhythmic accuracy*

We evaluated the performer's rhythmic accuracy during T2, by measuring the synchronization error between the onsets generated by their key presses and the reference click track.

CLICK TRACK We first attempted to use general-purpose onset detection methods from the literature [Bello et al., 2005], implemented in the form of VAMP plugins[32] by [Cannam et al., 2018]. The click track was generated in Reaper, and each of the "clicks" is a short (4 ms) sinewave of frequency 1600 Hz. The onsets are spaced by 0.5 s (120 bpm click track). The best results we could obtain reported inter-onset intervals (IOI) oscillating between 0.4992 s and 0.5021 s, for a jitter of 0.0029ms, probably corresponding to the algorithm's internal window step. Given how existing literature on synchronization error shows values in the range of a few milliseconds (see, for instance, Jack et al. [2018]), we felt the need for a more accurate onset detector. By leveraging our prior knowledge of the signal, we implemented an onset detector in Matlab which uses a simple thresholding on the time-domain signal, and delivers almost sample-accurate results, which results a maximum jitter of the IOI of 90µs, corresponding to 4 samples.

ORGAN We initially attempted to use general-purpose onset detectors on the sound generated by the instrument. While these can be made to work when individual, non-overlapping, notes are played, they tend to become very unreliable when the release of one note overlaps with the onset of the next note. Additionally, these algorithms are still affected by the jitter issues mentioned above.

We therefore used the sensor data we recorded during the study in order to obtain the exact onset time of each note. This was achieved by simply detecting when the downward motion of the key crossed the triggering threshold for each condition (explained in Section 4.6.1). For conditions C3 and C4, which defined a range for the triggering points, we used the mean of the range as the reference threshold to compute the onset instant.

32 https://vamp-plugins.org/plugin-doc/qm-vamp-plugins.html

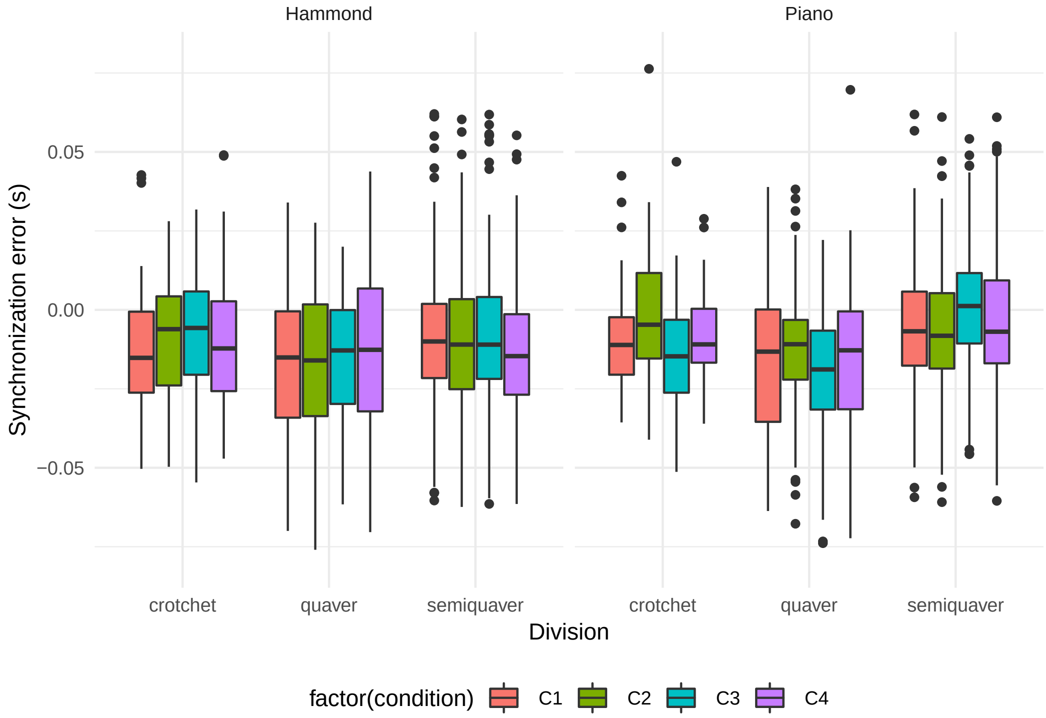


Figure 4.19: Interquantile ranges and median for the synchronization error during T2 for both groups.

SYNCHRONIZATION ERROR Once all onsets were detected, we inferred a higher-resolution grid from the click track in order to account for the subdivision (crotchet, quaver, semiquaver) of each section of the material performed (the rhythmic exercise in Fig. A.2). The synchronization error is then computed as the difference between a key onset and the nearest element of the grid.

The median values and interquantile ranges for all participants across all conditions are shown in Fig. 4.19. We fitted a linear mixed effect regression model to the synchronization error with fixed effects of *group* (Hammond, Piano), temporal *division* (crotchet, quaver, semiquaver), and *condition* (C1, C2, C3, C4), and random intercept for each *participant*. We used the *lme4* package in R [Bates et al., 2015], and the resulting model is `syncerror ~ group + division + condition + (1|participant)`. A full factorial Type III ANOVA on the LMER model, with Satterthwaites's degrees of freedom approximation, found a significant effect of *division* ($F(2, 4742) = 82.4611, p \ll 0.001$), but no effect of *condition* or *group*. A Tukey post-hoc analysis showed that the effect of *division* is driven by quaver-crotchet ($Z = -6.497, p < 0.0001$), and semiquaver-quaver ($Z = 12.796, p < 0.0001$).

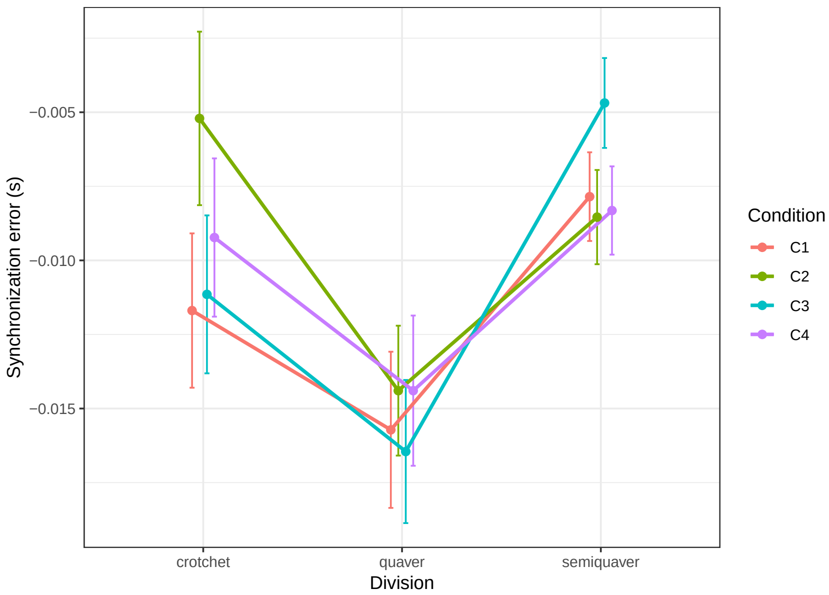


Figure 4.20: Mean and standard error of all responses from both participant groups, grouped by division.

We also tested for interactions between *division* and *condition* by fitting another model ($\text{syncerror} \sim (\text{division} * \text{condition}) + (1|\text{participant})$), which shows a significant effect of the interaction ($F(6, 4736.0) = 5.5049, p \ll 0.001$). Fig. 4.20 shows that the shape of the behaviour of the synchronization error with respect to the division is the same for all conditions, with the quaver having a larger (more negative) synchronization error on average than both crotchet and semiquaver. An interaction test with the *phia* package for R shows several significant interactions.

To assess the effect of *condition* for each *division* we then fitted separate models for each *division*, with fixed effects of *condition* and *division*, and again random intercepts for each *participant* ($\text{syncerror} \sim \text{group} + \text{condition} + (1|\text{participant})$). In the case of crotchets, we find a significant effect of *condition*. Fig. 4.21 shows that for crotchet, the mean of C2 is smaller than all the others, and a Tukey post-hoc analysis confirms a statistically significant difference between C2 and all other conditions (C2-C1, $p < 0.001$; C2-C3, $p = 0.028$, C2-C4, $p = 0.002$).

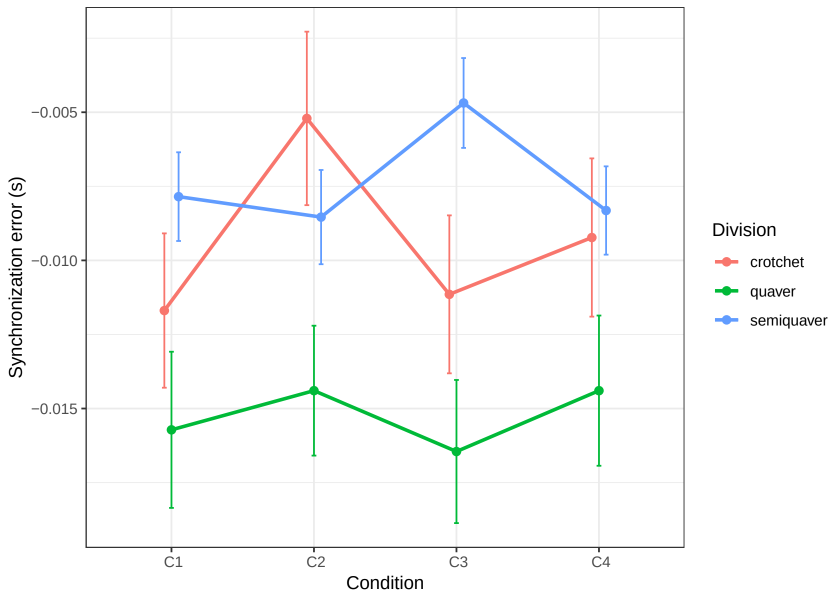


Figure 4.21: Mean and standard error of all responses from both participant groups, grouped by condition.

No significant difference is found for quavers. For semiquavers there is again a significant difference for *condition* ($F(3, 2754.05), p \ll 0.001$). Fig. 4.21 shows that C3 has the smallest average error of all the conditions for semiquavers, and a Tukey post-hoc analysis confirms that the difference between C3 and all the other conditions is statistically significant (C3-C1, $p = 0.0078$; C3-C2, $p < 0.001$; C3-C4, $p < 0.001$).

4.7.4.2 *Attack time*

We processed the sensor data from the keyboard scanner in order to segment them into notes and extract, for each note, the attack time. Bernays and Traube [2012] defines attack time as the time interval between the onset (the moment the key moves from the rest position) and the maximum depression point (the moment in which the key hits the key bottom). A drawback with applying this definition is that the point of maximum depression can actually occur after the key hits the key bottom, for instance in the case where the player puts extra pressure into the key. We therefore used an improved detection method whereby the instant of key bottom, which marks the end of the attack, is the moment where the key bounces back after hitting the felt at the end of the key throw. This can be detected numerically as the instant when the velocity of the key changes sign, as shown in Fig. 4.22.

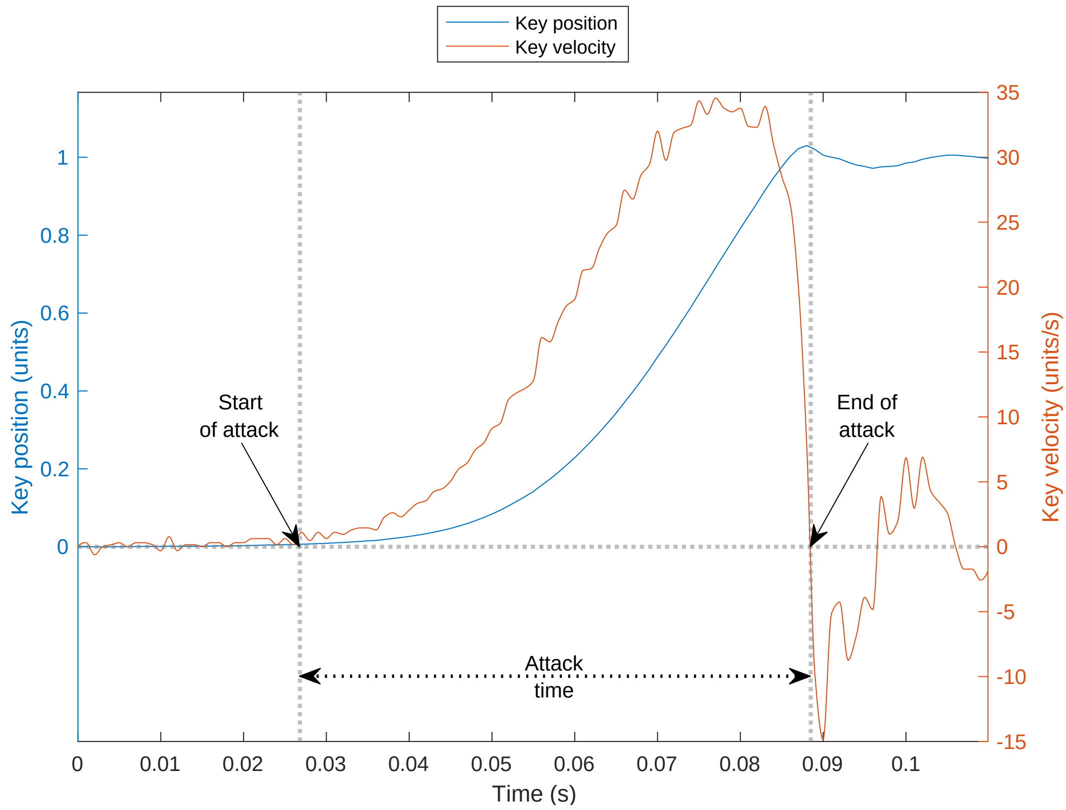


Figure 4.22: A note onset and the landmarks for computing attack time.

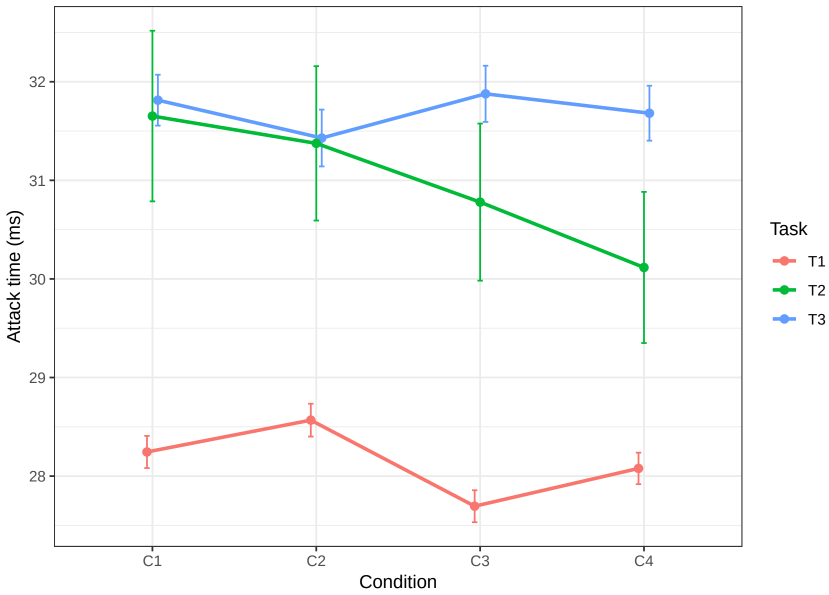


Figure 4.23: Mean and standard error of the attack times of the notes played by all participants throughout the study.

Mean and standard error of the attack times of the notes played by all participants in the four test conditions are summarised in Fig. 4.23.

We used linear mixed effect models for a statistical analysis of the attack times using the `lme4` package for R [Bates et al., 2015]. The analysis covers the full key presses with an attack time shorter than 100 ms from all the 10 participants, under the 3 tasks and 4 conditions, for a total of 154989 key presses. Using the participant as a random effect, we tested for the fixed effects of task and condition on the attack time, fitting two models, one with interaction term (`attackTime ~ condition * task + (1|participant)`) and one without (`attackTime ~ condition + task + (1|participant)`) A likelihood ratio test on the fitted models found a significant difference between the two ($\chi^2(6) = 30.762, p \ll 0.001$), which made us continue our analysis with the former.

A full factorial ANOVA was conducted using the model with interaction. We found significant effects of condition ($F(3, 154979) = 3.5855, p = 0.01312$), task ($F(2, 154982) = 841.2827, p \ll 0.001$) and their interaction ($F(6, 154980) = 5.1275, p \ll 0.001$).

A Tukey post-hoc analysis on condition of the fitted model showed significant differences between C1 and C3, C2 and C3, C3 and C4 (all $p \ll 0.01$), and between C2 and C4 ($p = 0.033$). When performed on tasks, significant differences were found between T1 and each of the other tasks ($p < 0.001$). A post hoc analysis of interaction contrast using the `phia` package showed a significant contrast between C1-C3 and T1-T3 ($p = 0.015$), C2-C3 and T1-T3 ($p < 0.001$), C1-C4 and T2-T3 ($p = 0.015$).

We then fitted mixed linear models to the data from each task separately, thus having 107667, 4866, 42456 observations for each of T1, T2 and T3, respectively. Again we used participant as a random effect in our model (`attackTime ~ condition + (1|participant)`). In all cases, we found significant effect of condition (T1: $p \ll 0.001$, T2: $p < 0.002767$, T3: $p < 0.002769$). Pair-wise Tukey post-hoc analysis found significant differences between C1 and C2 ($p = 0.01137$), C2 and C4 ($p = 0.00279$) for T1; C1 and C4 ($p = 0.00136$) for T2; C1 and C2 ($p = 0.0113$), C2 and C4 ($p = 0.0025$) for T3.

The results of the statistical analysis indicate that players varied the attack time of their notes depending on the task and condition, and on their interaction. The main effect of the task is that the attack times are overall shorter in T1 than in any of the other two tasks. It is interesting to find that in both T1 and T3 there is no statistically significant difference between C1 and C4, which are the two conditions that differ the most between themselves. This same pair of conditions is, however, the only one to show a significant difference in T2.

4.8 DISCUSSION

During Task 1, playing with a backing track, most participants struggled to perceive differences between the conditions, with some of them becoming frustrated as, take after take, they could perceive very little or no difference at all. While some players detected and reported meaningful differences, this resulted in a large number of unanswered questions overall, contradictory answers and the reporting of changes to aspects of the sound that were not actually changing. On top of this, the answers given in Task 3, the free exploration, and in the subsequent interview, were sometimes in sharp disagreement with those reported earlier in T1. All of this indicates that it is fairly hard for most musicians to perceive the subtle difference between our test conditions when playing along with an up-tempo backing track. Two possible reasons for this are the difficulty in hearing clearly the details of the onset transients due to the accompaniment, which makes differences between the conditions less perceivable, or the higher cognitive load required by the task, which distracts the performer. A possible alternative explanation for the disagreement within T1 could be found in the fact that drawbar settings were changed between repeated trials on a given pair of conditions. As for the disagreement between T1 and T3, a possible explanation the perceived attributes of the conditions may change depending on the musical context in which they are used. However, the data gathered in this study is not sufficient to confirm or reject these two possibilities.

By contrast, in T3, most players were able to express the differences between the conditions, with varying degrees of rationalization and understanding of the underlying phenomenon. Even though only two Hammond players understood fully, almost scientifically, the triggering behaviour of each condition, there was a clear pattern emerging from T3, where piano players would prefer conditions with single triggering point, while most Hammond players (3 out of 5) would prefer conditions with multiple triggering points. Possible insights as to why pianists preferred single triggering point conditions, beyond the obvious reason that that is the most similar to the behaviour of most other keyboard instruments, can be found in P6(p)'s observation, that pianists train to ensure notes sound together (when they are meant to), so C1 and C2 can give them a sensation of more control. Additionally, the initial triggering point of C4 is slightly higher than all the other conditions, which reportedly made it excessively responsive for some of the participants. However, some pianists rated C4 as their second favourite because of the sensation of weight that comes with the depth of its triggering points. The strongest preferences for C4 came from two of the Hammond players (P2(h) and P8(h)), who appreciated the "realistic" sound of its articulation and its overall feel.

4.8.1 *Perception of the attack*

An unexpected finding of our study, which surfaced as early as the pilot study, is that many participants perceived changes in the weight sensation of the keyboard action. The understanding of how the auditory feedback translates into weight is not, however, shared across the participants, and, further to that, the relation between the subjective perception of weight and, ultimately, the overall preference is not uniform. P3(h) rejected the idea that the weight of the keyboard could be changing, and they also were the only participant to mention latency as the changing parameter between conditions in T1, and possibly did so because of their previous knowledge of the author's work. In all of the three cases where they did, they indicated the condition with multiple triggering points as the one that was *"laggy"*. These conditions are characterized by a longer duration of the onset transient, which can lead to the effective beginning of the note to be perceived as delayed, as the "full sound" comes in later in the key throw, thus giving a plausible explanation to their sensation of latency. The literature shows that the effect of latency of tactile feedback affects the perceived weight of virtual buttons on touchscreens, with higher latencies consistently corresponding to heavier perceived weight [Kaaresoja et al., 2011], and that well-matched auditory feedback can induce haptic sensations of softness, elasticity and sense of movement when pressing a virtual button on a static rigid surface [Lai et al., 2011]. In our case, the changes were in the quality of the auditory feedback and, possibly, in its perceived latency, however the mechanics of the key clearly play a more complicated role than the static, flat surface of touchscreens, so it is hard to find a direct correspondence between our case and these.

Galembo and Askenfelt [2003] showed that the kinaesthetic sensation of playing a piano has an important role in the evaluation of the sound quality of the instruments, so much that such sensation can be sufficient for a performer to recognize an instrument when the visual and auditory feedbacks are removed. Fontana et al.

[2017] showed that the vibrotactile sensation perceived during piano playing can also affect the perceived quality of a piano sound. In both cases, the effects of the non-auditory sensation are unconscious, yet important. Our findings seem to indicate that the auditory feedback has the potential to be used for altering the perceived haptic response of a keyboard, suggesting the need for further investigation on the multi-directional effect in the cross-modal interaction of sound and haptic sensations.

An interesting anecdote came from P5(p), who mentioned a recent experience of them playing the Pianoteq[33] physical modelling piano plugin, where the feeling of the keyboard's resistance was affected by the produced sound: *"when you play it as a pianist, things are happening [in the sound] that you expect to happen. There's 'resistance', there's a sense with the sound ... which obviously there is no more or less resistance because it's the same MIDI keyboard ... but what you are hearing is a perception of a different resistance, which suddenly made it like I can just play it as a piano. [...] On the other hand, I played the most amazing piano samples but ... they are not more or less enjoyable to play on a MIDI keyboard than the MIDI keyboard itself"*. We can only speculate as to what is the phenomenon at the root of P5(p)'s experience with Pianoteq, however it is another interesting observation on how the mapping between gesture and sound has the potential to elicit physical sensations.

When P8(h) described C1 as *"completely without control"* they were referring to the inability to control the sound of the onset, which was therefore always *"full of power"*, while in C4, by contrast *"if you touch in a quicker way the key, the sound is quieter"*. They therefore seem to appreciate the dynamic control available at their fingertips, on the basis of the variety of the achievable sonic outcome, even though the difference is concentrated on the onset transient of the note, and does not affect the dynamic of the steady-state part of the note. P2(h) also went through a similar process. No other players, however, demonstrated such sensitivity for the attack sound. Among pianists, only P5(p) reported a difference specifically in the attack sound, singling C4 out, and yet they did not investigate specifically the attack sound of the other conditions, although they did feel a difference in "flow" between them when playing melodic lines. Even P4(p) and P6(p), who found out by chance in T1 about something uncommon going on with the triggering characteristic of C4, did not dedicate time to exploring it during T3. An open question is therefore: why did the only sonic parameter that was changing not get nearly as much attention as the physical sensation of weight that it, alone, generated?

33 https://www.pianoteq.com/

4.8.2 *Effect on timing*

The analysis of the attack time feature showed some statistically significant effects of condition and task. The effect of task seems to be the most prominent, with the mean attack time during T1 being 3.5 ms faster than during T3. A possible explanation for this is that in presence of a loud backing track, the player would tend to play "harder", as if they were reaching for a louder sound. Clearly, a faster velocity does not result in a louder sound on the Hammond or on velocity-insensitive keyboards, but it could well be a matter of habit for keyboard players. In Finney [1997], players performing a given passage without auditory feedback tend to play with slightly higher velocity overall than in normal auditory feedback conditions. While this is in principle in agreement with our observation, conclusive finding on the impact of relative level of the auditory feedback of the instrument and the accompaniment on key velocity of velocity-insensitive keyboards would require further study.

A key press during C4 and C3 will produce a temporal smearing of the keyclick of the produced sound, and more so for C4 than for C3. Some participants have lamented difficulties in controlling the tempo with C4, due to the fact that the sound is not "complete" with all its harmonic partials until the key reaches the bottom. The mean attack time in T2 is maximum in C1 (31.6 ms) and becomes shorter and shorter moving towards C4 (30.1 ms). The fact that the attack time reduces as the temporal smearing becomes more prominent during the tapping exercise (T2) could be due to the player trying to reduce the duration of the transient, and make all the partials sound together as close as possible to each other. The decrease of attack time with temporal smearing only happens in T2, and neither T1 nor T3 shows a similar pattern. This suggests that the musical context of the rhythmic task (T2) requires participants to optimize for timing accuracy, with consequent reduction of the temporal smearing, more than the other tasks do.

The statistical analysis of the synchronization error of the notes played during T2, showed that when playing crotchets, the slowest subdivision, C2 yields the minimum error (players play more in time), while when playing semiquavers, the fastest one, it is C3 which yields the minimum error.

4.8.3 *Ecological considerations*

Studies of music performance in a lab setting always come with trade-offs between ecological validity and control of the study setting [Runkel and McGrath, 1972]. In our case we tried to place our participants in a realistic context, by using the Hammond itself as an interface, and freely improvising to try to minimize their cognitive load, and let them concentrate on the task. On the other hand, we had to set some constraints, for instance the lack of a Leslie speaker, percussion, vibrato or even drawbar adjustment, in order to allow reproducibility across participants. While some of the Hammond players reported to have been distracted by the impossibility to adjust the sound to their taste during T1, and that this did not allow them to concentrate on the actual task at hand, we stand by our design decision: we believe that giving piano players control over all these parameters would have disrupted their performance more than *not* having control over them disrupted that of Hammond players. We report here a short anecdote which makes the case for the sufficient ecological validity of the playing experience on our instrument. When the author first met P2 three years prior to this experiment taking place, it was in the warehouse of a Hammond seller, a common acquaintance from whom we eventually bought the instrument we used for our studies. In that occasion, P2 pointed out how the liveliness of the real instrument, due to the contact bouncing and the tonewheels, were impossible to replicate on a digital clone. Surely this put some pressure on us when we invited them to join for the study. Surprisingly for us, for the whole duration of our experiment, they believed that the sound they were hearing came from the real Hammond sound generator. While anecdotal, this episode reassured us that the experience of playing our hybrid Hammond is somewhat realistic. As the digital components of our instrument were not immediately visible to the player for the duration of the experiment, with the exception of the MIDI controller in T3, this example also reminds us of the important role played by physical interaction and pre-conceptions in the debate of "real Hammond vs emulator" that is so important to the worldwide community of Hammond players.

Starting the study with the performance task was a deliberate decision, to avoid players getting familiar with the conditions before the performance, as we were afraid that this would have made them concentrate more on trying and guessing which condition was which, instead of concentrating on their sensations. This in turn is probably one of the causes for the difficulty in detecting differences in T1, and the large number of unanswered questions in the questionnaire. As some of the performers have suggested, perhaps using a slower, backing track, less busy with drums, would have allowed to hear more of the onsets of the organ sounds, thus making the task easier.

4.8.4 *Future work*

Our study was conducted on a small sample of players with a diverse musical background and diverse playing styles. Musicians who play Hammond regularly may also develop a very personal approach to exploiting the attack onset into their playing. This gave us interesting insights on the wide range of responses that variations in the key-sound mapping can elicit in a player, however the small number of subjects and their heterogeneity may have contributed to the weak significance of some of the statistical measures, and to the lack of emergence of clear preferences. A study where participants are pre-selected according to their playing style, whether through quantitative or qualitative means, in order to obtain a more homogeneous group, would possibly see clearer patterns emerge.

We have observed our participants in a controlled laboratory environment, under at times demanding cognitive load, and we limited their freedom of choice of repertoire and control of the settings of the instrument. A longitudinal study would be extremely useful to outline variations and adaptations in the playing style over a longer period of time, hopefully increasing their awareness of the differences in sound and feel between the conditions, and giving rise to development of strongly polarised personal preferences. We could then examine how a player's technique evolves when playing with a given touch setting extensively, and how their gestural vocabulary changes in response to different amounts of control.

4.9 CONCLUSION

In this chapter we designed a platform for real-time sound generation from continuous key position. We used this platform to create a digital emulator of a Hammond organ, for which we developed a dynamic contact bounce model and arbitrary mappings between the key position and the 9 virtual contacts of the sound generator, extending a previously existing piece of software. We integrated our hardware and software into the cabinet of a Hammond organ, using part of the original sound production circuitry and hardware where possible, aiming to maximally preserving the playing experience of the instrument for all aspects except the key-sound mapping. Our instrument has 4 different touch conditions, each associated with a different key-sound mapping, two of which emulate the response of commercially available Hammond emulators, with single triggering point, a third one which is modelled after the measurements we performed in Chapter 3, and the last one which is an exaggerated version of the latter. The last two conditions, have multiple triggering points, and each virtual contact closes at a different point in the key throw. In a performance study with five Hammond and five piano players, we investigated their response to the change of touch conditions, while they performed along a backing track, and while freely exploring the instrument.

Our results indicate that distinguishing between the conditions while playing with the backing track was very hard for the players, and only few players showed consistency in their responses. When exploring the instrument more freely, however, participants managed to hear more clearly the distinctions between the conditions, and most developed a strong preference for one of the conditions. Hammond players tended to favour conditions with multiple triggering points, while piano players favoured the reliability and reproducibility of those with single triggering point.

In Chapter 3 we showed that changes in the attack sound comparable in range to those produced by our four conditions can be perceived even from listeners that are not familiar with the sound of the Hammond. In the study we presented in this chapter, however, piano players reported more often differences among the conditions with regard to the perceived weight of the keyboard action, probably due to the different triggering points, rather than on the produced sound. This finding opens up the question of how the perception of the sound can be influenced by the experience of actually playing it, and, additionally, of how the generated sound interferes with the perception of the physical interaction.

In this chapter we investigated the effect of continuous key position control on the onset of notes, which has consequences for the performer in terms of the perceived sound and feel of the keyboard, however the overall behaviour of the instrument is not too different from a traditional velocity-based keyboard. In our next study we will leverage continuous control in a more dramatic way, by placing it at the foundational working principle of the instrument, in order to maximally engage with the player's perception and creativity.

5

CONTROLLING A PHYSICAL MODEL INSTRUMENT WITH CONTINUOUS KEY POSITION AND PERCUSSIVE FEATURES

Some of the concepts from this chapter are summarised in [Moro and McPherson, 2020].

In this chapter we discuss the design of a new keyboard-based digital musical instrument where the continuous key position parameter is leveraged to allow extended keyboard techniques, and a performance study on the instrument.

We start laying out the motivations of our work in Section 5.1, followed by a description of the different parts of the instrument: the sound engine (5.2), the control space on the keyboard (5.3), the mapping between the two and the sounds it can produce (5.4), and a brief summary of the implementation details (5.5). We then describe the performance study (Section 5.6), and its outcomes (5.7), which we then discuss in relation to the literature and our previous work (5.8).

5.1 MOTIVATION

The instrument we created in Chapter 4 has a quasi-discrete behaviour with respect to key position, so that a player does not need to be aware of the effect of continuous key position in order to play it. This is reflected in the outcome of the study, where differences in the mapping of continuous key position result in relatively few comments which denote an awareness of the triggering behaviour, while many identify changes to parameters such as keyboard weight and the produced sound. We are then left wondering what would happen in the case of a more obvious effect of continuous key position on the sound generation, especially considering the success of an illustrious precedent in the history of electronic keyboard instruments, the Ondioline [Fourier et al., 1994], which made of continuous key position one if its main expressive parameters.

We therefore set out to create a new keyboard instrument based on continuous key position, as a probe for studying the generalization of keyboard playing skills to changes in the interface. The instrument we describe in this chapter is a monophonic synthesiser where extended techniques based on the instantaneous position of the key, its temporal evolution, and interactions between neighbouring keys produce a characteristic sonic effect. This instrument challenges some of the basic assumptions underpinning most keyboard instruments. Traditionally, a key is a discrete velocity-sensitive switch which requires "bursts" of attention from the performer, in correspondence with attack and release. In our instrument the key becomes a continuous controller, which requires control and accurate movements throughout the duration of a note. We do, however, retain an element of discreteness, in that discrete percussive sonic events can be achieved by using a percussive gesture when attacking a note. Onset velocity, however, has no effect. Additionally, we change the relationship between keys. On polyphonic keyboards, keys are independent of each other, and can be activated individually, without, in first approximation, one affecting the sound of the other. On ordinary monophonic synthesizers, when multiple keys are pressed, only one would sound at any time, and the sound of the active key would not be affected by the other keys. In our instrument, also a monophonic synthesizer, pressing two neighbouring keys at the same time results in an interaction between them, producing a pitch bending gesture with the second key acting as a continuous controller on the pitch of the first. The continuous position control and the percussiveness gestures are inspired to the phenomena we observed in the Hammond in Chapter 3, though magnified to make them much more prominent, while the pitch bending draws inspiration from a similar behaviour found on the Magnetic Resonator Piano [McPherson and Kim, 2012].

In order to make the maximum of the extended techniques, the sound engine of the instrument should have a sustained tone that can be shaped continuously, and allows continuous control of its internal parameters. We found physical modelling synthesis particularly attractive for this project. Physical models have, by nature, a large number of internal parameters, and they can reproduce, to a certain extent, the "misbehaviour" of acoustic instruments, as well as some unexpected behaviours of their own, thus producing remarkably rich, naturally sounding results [Borin et al., 1992, Castagne and Cadoz, 2003]. Their sound is therefore often *physically plausible*, in that even if the sounding object being emulated does not necessarily have a counterpart in the real world, it will still appear to the listener as if it was produced in a physical way [Castagne and Cadoz, 2003]. While we are not concerned by the realism of the model we use, we expect that a connection to reality through physical plausibility can help players to understand the behaviour of the instrument more intuitively.

In principle, the problem of mapping controller inputs to physical models could be easily solved by leveraging the real-world counterparts of the exposed parameters, for instance mapping virtual toneholes in a clarinet simulation to toneholes on a clarinet-shaped gestural controller. However this comes at the cost of developing a dedicated controller, and taking the time to master it [Miranda and Wanderley, 2006, p. 15]. In our approach, we are trying to minimize the necessity to learn new skills by re-using an existing controller which is normally used for discrete events, repurposed in a continuous fashion.

Some previous attempts to overcome the discrete characteristics of the keyboard interface, such as the Seaboard [Lamb and Robertson, 2011] and the Continuum [Haken et al., 1998], did so by completely transforming the mechanics of the instrument and its haptic and tactile response, eventually preserving only the spatial location of the notes. As a consequence, they require a considerable adaptation and relearning effort for players who are already experienced with a conventional keyboard. Using Krakauer et al. [2006]'s terminology, we would say that many of the skills from the context of keyboard playing do not *generalize*[1] to Seaboard or Continuum playing. In our case we use a regular weighted keyboard, an interface familiar to most keyboard players, and we therefore hope to see a *transfer*[2] of existing skills, especially with regard to spatial navigation. At the same time, we are somewhat defamiliarising the keyboard, by using it in a continuous fashion, thus extending the concept of keyboard beyond its common understanding. It is consequently very likely that for some of the skills required to play our instrument, we will not observe any generalization of keyboard skills, and possibly we will observe *interference*[3], in the case where pre-existing training becomes an obstacle to learning the new techniques.

The high-level block diagram of our instrument, including its interface and main components can be found in Fig. 5.1. The instrument is implemented using the platform we presented in Section 4.3.2. Three functionally distinct modules run on Bela: the sound generator, a "non-linear waveguide flute" physical modelling synthesizer, the "keyboard controller", which transforms the continuous key sensing coming from the keyboard into continuous and discrete control parameters, and an intermediate mapping layer.

1 *" Generalization of motor learning refers to our ability to apply what has been learned in one context to other contexts"* [Krakauer et al., 2006].
2 *"When generalization is beneficial, it is termed transfer"* [Krakauer et al., 2006]
3 *"when [generalization] is detrimental, it is termed interference"* [Krakauer et al., 2006]

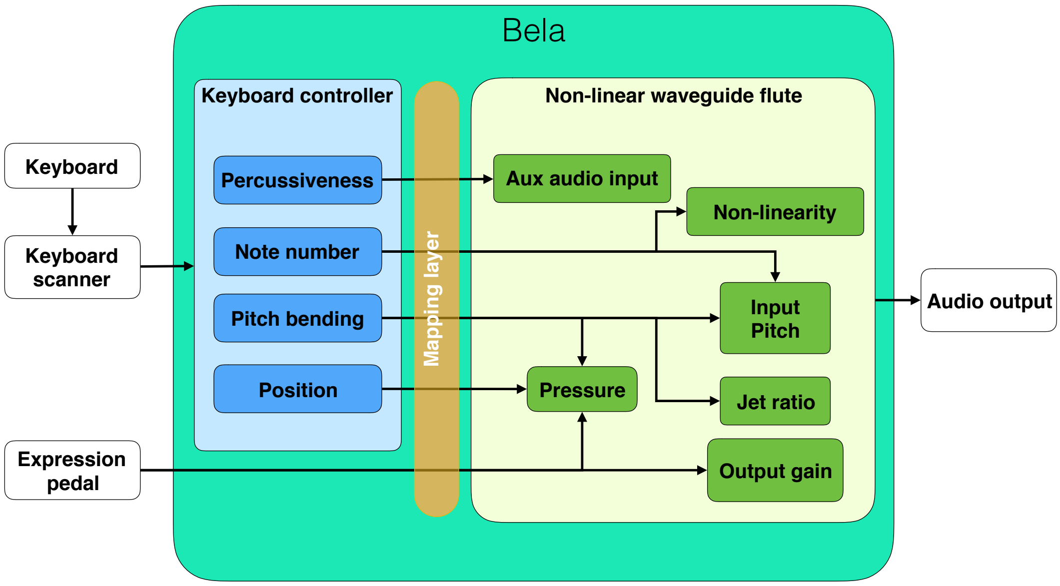


Figure 5.1: Block diagram of our instrument, a physical modelling synthesizer of a flute controlled with continuous key position, running on Bela.

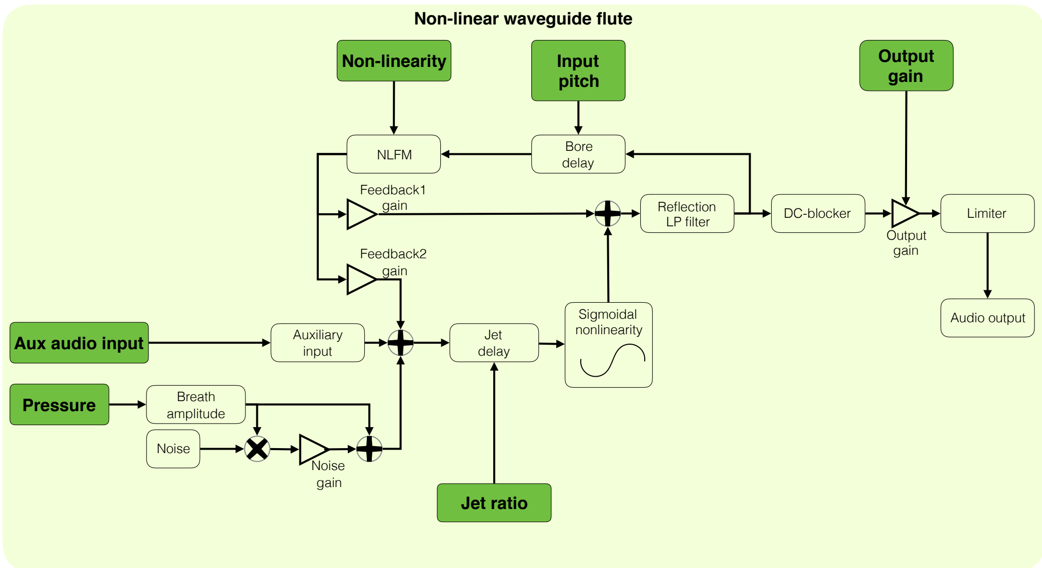


Figure 5.2: Block diagram of the non-linear waveguide flute model. Dark green boxes indicate the parameters exposed for real-time control.

5.2 THE NONLINEAR WAVEGUIDE FLUTE

The starting point for the sound engine of our instrument is a nonlinear waveguide physical model of a flute developed by Romain Michon in the FAUST [4] program-

4 FAUST: Functional AUdio STream is a programming language that proposes an abstract, purely functional approach to signal processing. http://faust.grame.fr/

ming language, which is currently distributed with FAUST [Michon and Smith, 2011]. This model does not include toneholes, so it effectively emulates a slide flute, and is based on the one presented in Cook [1992], with an added non-linear ladder filter modulator (NLFM) [Smith and Michon, 2011]. We further modified the model to provide control over the length of the air jet between the mouth and the mouthpiece, improved the stability of the model, and added a DC-blocker and a dynamic range limiter to the audio output. A flow diagram of the final model and the control parameters that we expose to other parts of the software for real-time manipulation can be found in Fig. 5.2.

5.2.1 *Working principle*

The bore of the instrument is modelled as a waveguide , that is a fractional delay line ("Bore delay" in Fig. 5.2) whose length in samples d_b corresponds to the period of the resonant frequency f_b of the physical bore, which is controlled by the "Input pitch" parameter:

$$d_b = F_s/f_b \tag{5.1}$$

for a sampling rate F_s [Smith, 1992]. A low-pass filter ("Reflection LP filter") and a fixed gain factor ("Feedback1 gain"), together, represent the frequency-dependent reflection loss at the open end of the bore. The excitation signal is modelled as a signal comprising a white noise source ("Noise"), scaled by and summed to a DC component ("Breath amplitude"), plus an "Aux audio input" to add an arbitrary signal generated externally. The resulting signal is coupled to the opening of the bore by a smooth non-linearity ("Sigmoid non-linearity"), which in Cook's model is a polynomial of the form $y = x^3 - x$, and is called a "sigmoid"[5] . The "Jet delay" is a delay line that represents the propagation time of the jet reed, and is due to the hydrodynamical disturbances that arise in the air jet between the mouth of the performer and the mouthpiece of the instrument. These disturbances travel much slower than the speed of sound [McIntyre et al., 1983], and the length of the jet delay is, in non-overblowing conditions, the same as that of the Bore delay [Karjalainen et al., 1991].

5 Cook [1992] calls such non-linearity a "sigmoid", while this term is more often used to denote monotonic functions (see, e.g.: https://en.wikipedia.org/wiki/Sigmoid_function). In its place, Välimäki ([Välimäki et al., 1992, 1996]) uses a proper sigmoid function, $y = h + k \tanh(x)$, as given by the physical analyses of McIntyre et al. [1983] and Fletcher [1999].

The model described so far closely resembles that presented by Cook [1992]. Michon and Smith [2011] added a non-linear ladder filter modulator ("NLFM") to the bore delay line. The allpass filter coefficients in the NLFM are modulated dynamically by a function of the input signal in order to add interesting timbral evolutions [Smith and Michon, 2011].

Our contributions to the model include an improved output stage, with a high-pass filter ("DC-blocker") and dynamic range compression ("Limiter"), the input for external audio signals mentioned above, improvements to the overall stability of the model, and fine tuning of the model in order to improve usability. Additionally, we added fine control for the length of the jet delay line in order to produce overblown tones and slight pitch deviations from the resonant frequency of the bore.

5.2.1.1 *Jet ratio*

A flute player can play tones that are harmonics of the fundamental frequency of the tube (overblowing) by changing the distance between the lips and the embouchure and increasing the blowing pressure, and, consequently the velocity of the jet, transitioning to a higher mode of oscillation [Fletcher, 1999, p. 748]. Equivalently, in our physical model, overblowing can be achieved by changing the length of the jet delay line. Shortening its length by an integer ratio, a corresponding upper harmonic pitch can be achieved [Välimäki et al., 1996]. Fractional changes to the length of the delay line result in slight pitch changes. We added to the model a parameter "jet ratio" (j_r) which determines the length of the jet delay line (d_j) in relation to that of the bore delay line (d_b):

$$d_j = d_b j_r \tag{5.2}$$

When j_r is a positive integer, the produced tone will be the corresponding harmonics of the fundamental frequency of the bore. When it is a fractional value, small pitch changes take place, accompanied by an increased turbulence in the produced sound. We analyse the behaviour of the model under these conditions in Section 5.2.2.5.

5.2.1.2 *Stability*

The original model proved to be unstable for large sustained pressure values. Inspecting the signal flow diagram, we found that the gain of the feedback loop was allowed to become larger than 1 due the large outputs of the unbound sigmoid non-linearity for large input signals[6]. We therefore added a clipping stage following the non-linearity in order to limit its maximum gain and restore stability.

5.2.1.3 *Auxiliary audio input*

We added an audio-rate input to the model, so that the incoming signal is summed to the airflow generated in response to the pressure parameter. This allows us excite the model with arbitrary audio input signals.

5.2.2 *Calibration of the physical model*

The sound model as described so far is not immediately playable because of some minor defects. The produced pitch is not linearly related to the input pitch parameters, so that it tends to be flat at higher pitches. The Nonlinear Ladder Filter Modulation (NLFM) inserted in the bore delay line sometimes prevents the establishing of a standing wave in the waveguide corresponding to the bore. This phenomenon is particularly obvious in the mid-high register of the instrument, where fast ramps in the pressure value result in "dead notes", ending up in a state where a sustained high pressure generates a breathy sound, with no harmonic component. We identified the parameters of interest that affect the intonation of the instrument, (jet ratio, pitch, pressure, non-linearity), and those that affect the occurrence of dead notes (pitch, pressure (speed and range), non-linearity).

In this section we present a set of measurements that we performed in order to investigate the behaviour of the model and we suggest remedies to normalize the behaviour of the instrument. We adopted a black box approach, where we measure the pitch and amplitude of the sound produced by the instrument for different values of the input parameters. For these measurements, we loaded our FAUST model into Pure Data[7] (Pd) using the `[faustgen~]` external[8], which allows to run FAUST code inside Pd. We can then use message boxes and regular Pd objects in order to alter the parameters of the FAUST model, and use the `[fiddle~]` object to measure the fundamental frequency and the amplitude of the generated sound [Puckette et al., 1998].

6 When introducing a non-linearity in a feedback loop, stability is preserved only as long as the gain of the non-linearity is smaller than one, see https://ccrma.stanford.edu/~jos/smithbook/Stability_Nonlinear_Feedback_Loops.html

7 Pure Data is an open source visual programming language for multimedia, http://puredata.info/

8 https://github.com/CICM/pd-faustgen

5.2.2.1 *Choice of reference values*

We experimented with the sonic outcomes of different values of the non-linearity control, and we settled for a value of 0.12 for our instrument, as it seemed to provide a good balance of tone richness and naturalness. Similarly, we experimented with different levels of pressure to find one that guarantees enough dynamic range, so that the pressure level can be increased progressively leading up to that point, but also such that the sound is not pushed too harshly into the distortion that occurs at higher pressure levels; we settled for a pressure level of 1.2.

5.2.2.2 *Effect of non-linearity on dead notes*

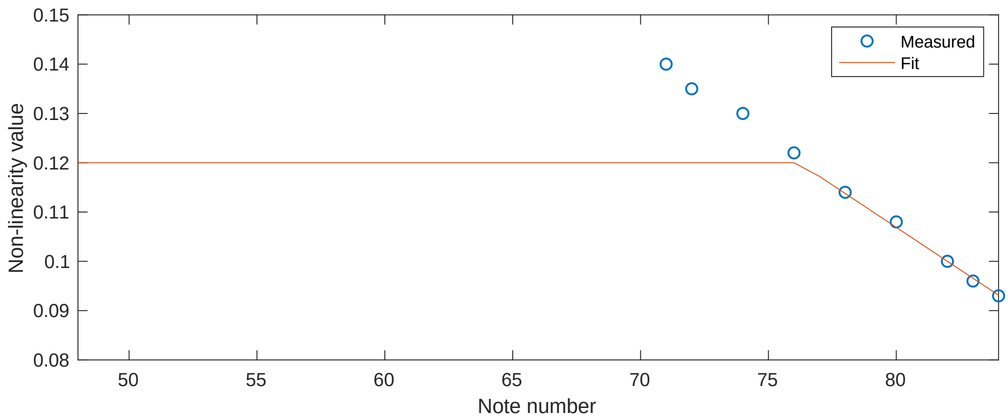


Figure 5.3: Maximum non-linearity value that allows to generate a non-dead note. The fit curve is the graphical representation of Eq. (5.3).

Values of non-linearity different from 0 have the undesired effect of preventing some notes in the higher register from sounding when the pressure control is ramped up too fast. We then decided to adaptively reduce the non-linearity value in the higher register to avoid this phenomenon.

Here we describe the experimental procedure we followed to obtain a frequency-dependent non-linearity curve which aims to prevent the occurrence of dead notes. For each note:

- Start with a non-linearity value of 0.12.
- Ramp up pressure from 0 to 1.4 in 10 ms, and hold it at 1.4.
- If the note doesn't speak, reduce the non-linearity in increments of 0.005 until the note starts speaking.
- When the note starts speaking, make a note of the non-linearity value.

- Revert the pressure to 0, then perform the onset ramp again as above, and verify that the note speaks correctly.

The values retrieved with this method for several different pitch input levels are shown as blue circles in Fig. 5.3. For pitches up to 75 (E♭5), the non-linearity value of 0.12 does not cause problems. For higher pitches, however the maximum acceptable non-linearity value seems to decrease linearly. The pitch-dependent non-linearity curve we adopt is the piecewise linear function (red line in Fig. 5.3)[9], defined as:

$$y = \begin{cases} 0.12 & , x \leqslant 75 \\ \min(0.12, \quad -0.00344x + 0.38226) & , x > 75 \end{cases} \tag{5.3}$$

This correction solves the issue of dead notes for all pitches up to the higher note we tested for on our instrument (83, B5)

5.2.2.3 *Effect of input pitch on produced pitch*

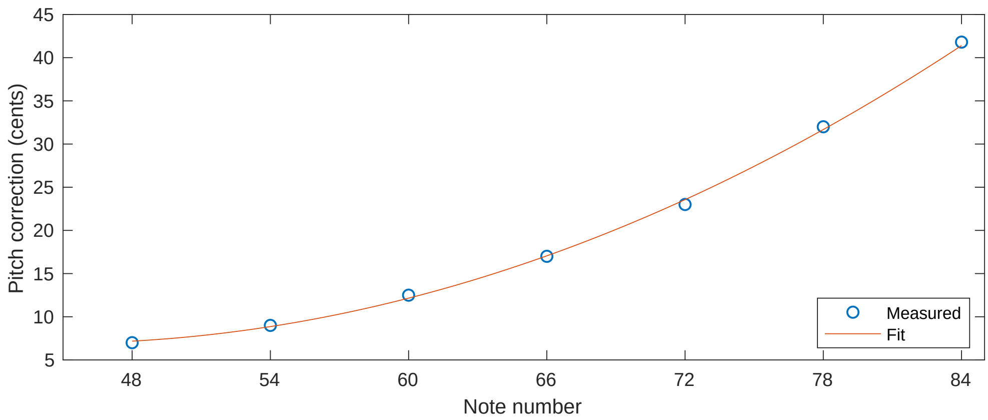


Figure 5.4: Measured pitch correction offsets needed to obtain an in-tune output from the sound engine for a fixed pressure value of 1.2. The fit is obtained using Eq. (5.4)

After setting a pressure value of 1.2, we send an input pitch value to the sound engine, applying the non-linearity correction from Eq. (5.3) as appropriate. If the tone produced as a result is slightly off-tune, we manually correct the input value until the output is in tune, and we record the pitch correction needed for that note. Repeating this procedure for several notes across the range of the instrument,

9 The $\min()$ expression ensures continuity in the curve

we obtain the circles in Fig. 5.4. We then used a polynomial least-squared fit to interpolate the observed pitch correction values, obtaining the following equation:

$$y = 0.00022321x^2 - 0.01995833x + 0.5155 \tag{5.4}$$

Fig. 5.4 also shows the fitted values.

5.2.2.4 *Effect of pressure on produced pitch*

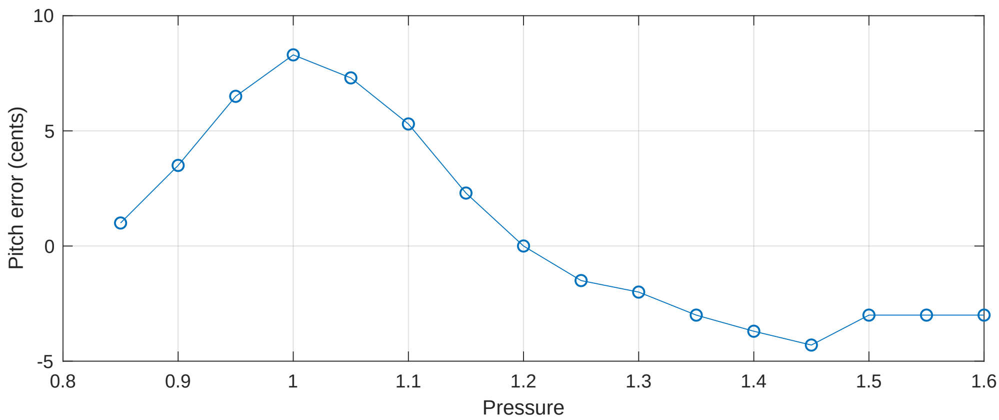


Figure 5.5: Measured pitch error for a C3 note, for a fixed pitch input, for changes in the input pressure.

The amount of pressure into the physical model affects the produced pitch Fig. 5.5. Fig. 5.5 shows the measured pitch error for a given note, for a fixed pitch input, as the input pressure changes. The input pitch value is determined applying Eq. (5.4), and consequently, the error for pressure 1.2 is 0 by definition. The error range is within +8 cents, -5 cents.

5.2.2.5 *Effect of jet ratio on produced pitch and timbre*

When the jet ratio is changed from the default value of 1, both the spectral content and the pitch of the generated signal change. In Fig. 5.6, are the spectra of two signals with the same output pitch, obtained through two different combinations of the pitch input control, and the jet ratio. The overall energy is lower for the one with non-integer jet ratio, and the relative intensity of the harmonics is noticeably different.

Fig. 5.7 shows the spectrogram and the output pitch of a tone for which the jet ratio is varied smoothly between 0.2 and 2. For jet ratios larger than 1, the pitch increases smoothly until it is about 200 cents (2 semitones) above the original

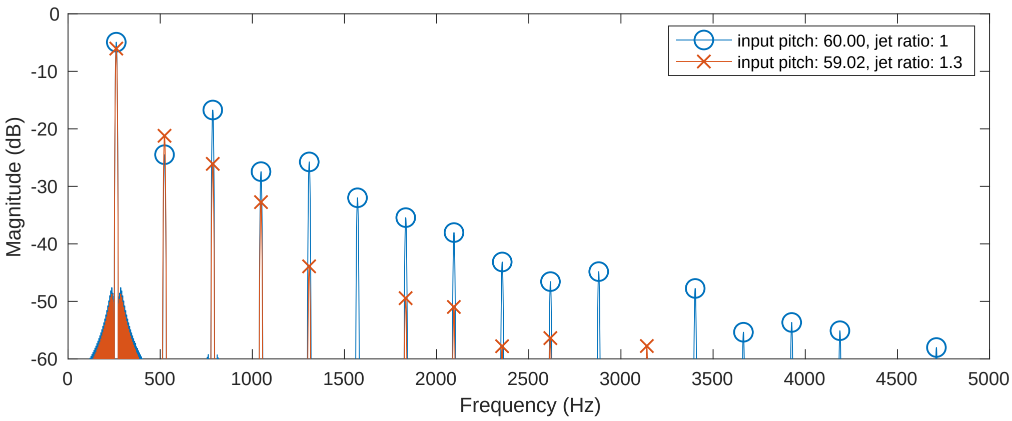


Figure 5.6: Spectra of a tone with a fundamental of 261.6 Hz, obtained with a jet ratio of 1 and 1.3, and a corresponding input pitch value. Pressure is 1.2 in both cases.

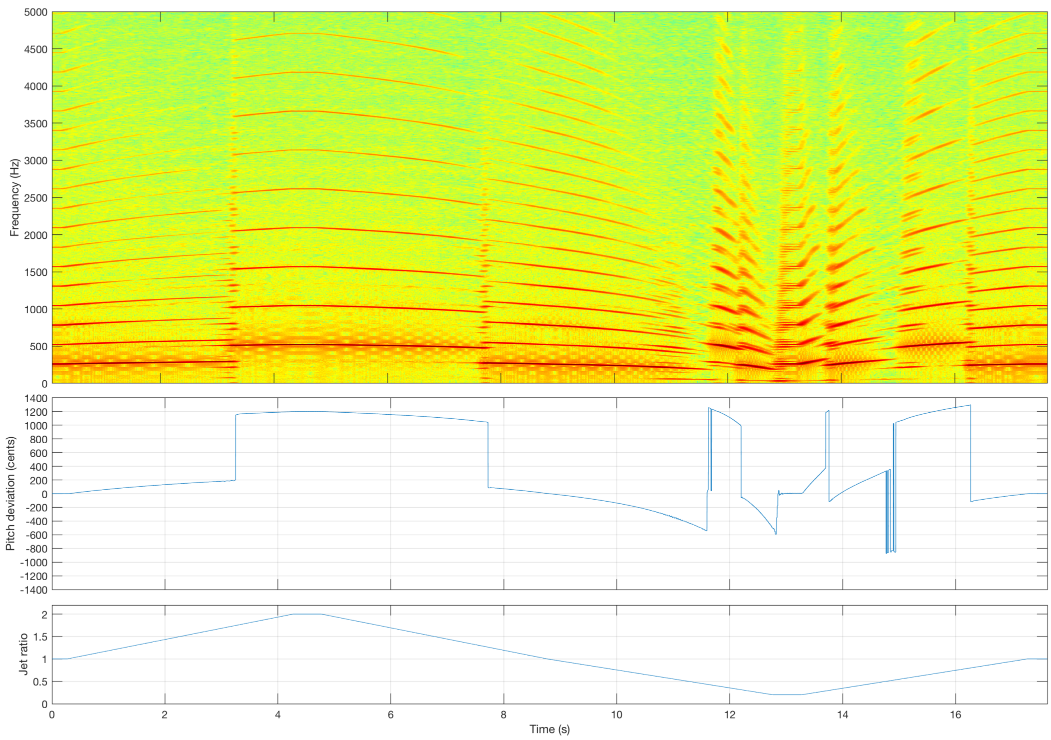


Figure 5.7: Spectrogram and pitch of the generated sound for a changing jet ratio. The jet ratio starts from 1, ramps up to 2, then ramps down to 0.2 and back up to 1. Pressure is fixed at 1.2 and input pitch fixed at 60. Pitch detection was performed with Pd's `[fiddle~]`

pitch, then it suddenly spontaneously jumps into "overblown" mode. At first, after the jump the note will be slightly flat, but it eventually reaches 1200 cents (one octave) above the initial pitch when the jet ratio reaches the integer value 2. For jet ratios smaller than 1, initially the pitch decreases smoothly until about -450 cents (over 4 semitones), but beyond that the behaviour becomes more chaotic, with more complex spectra and multiple transitions. We observe that the trajectories of the pitch for increasing and decreasing values of the jet ratios differ, which is an indication of the presence of hysteresis in the behaviour of the system. In Fig. 5.8 we plot the hysteresis loop for pitch in function of the jet ratio, where we notice two distinct hysteresis loops, one for values of jet ratio larger than about 1.2, and one for values smaller than about 0.8, and that there is no hysteresis around 1.

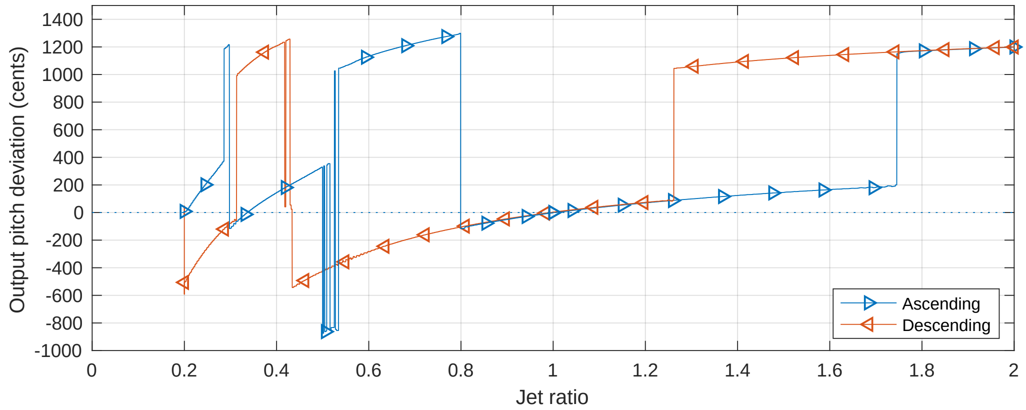


Figure 5.8: Hysteresis cycle of the output pitch for a changing jet ratio. This plot is obtained from the same signal as Fig. 5.7.

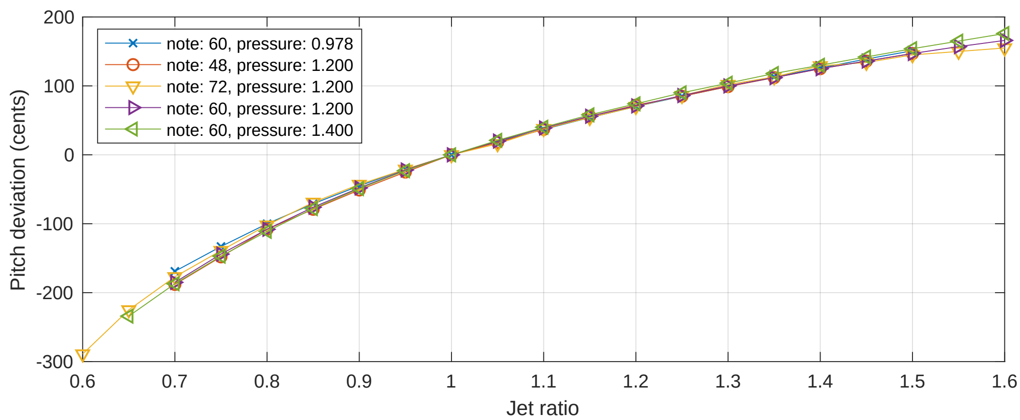


Figure 5.9: Measured pitch deviation for several pitch inputs and pressure values, for changes in the jet ratio.

In Fig. 5.9 we report some measurements obtained by modifying the jet ratio in small increments about the value of 1. The figure shows that for a fixed pressure of 1.2, deviations in the jet ratio for notes in different registers coincide across the jet ratio range, indicating that the effect of jet ratio on pitch deviation is independent of the base pitch. However, curves for different pressure values for the same note do not coincide, and when the pressure is higher, the effect of jet ratio on pitch deviation is larger.

These two observations suggest that for a given pressure, effect of the jet ratio on the pitch is the same across the range of the instrument. This property, combined with the earlier observation that there is no hysteresis around the jet ratio of 1, lays the foundations that will allow us to use jet ratio as part of pitch bending transitions between notes, as explained in detail in Section 5.4.3.

5.3 KEYBOARD CONTROLLER

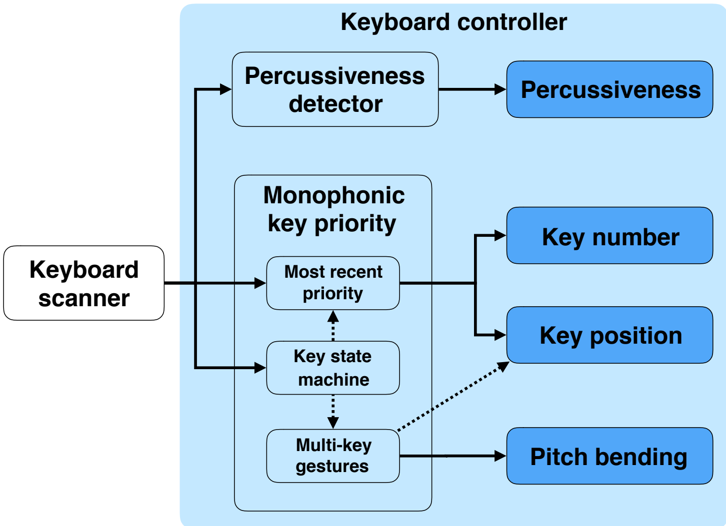


Figure 5.10: Block diagram of the keyboard controller. Dashed lines indicate minor interactions. Dark blue boxes indicate the outputs of the controller.

In this section we illustrate our keyboard controller, which processes in real-time the raw positional data from the keyboard scanner, evaluates interactions between multiple keys to determine a monophonic active key and possible bending gestures, and detects percussive key presses. A block diagram of the functions performed by the keyboard controller, and its inputs and outputs is in Fig. 5.10.

5.3.1 *Key state machine*

Observing the evolution of the trajectory of a key over time, it is possible to identify several distinct temporal segments corresponding to the different phases of the key

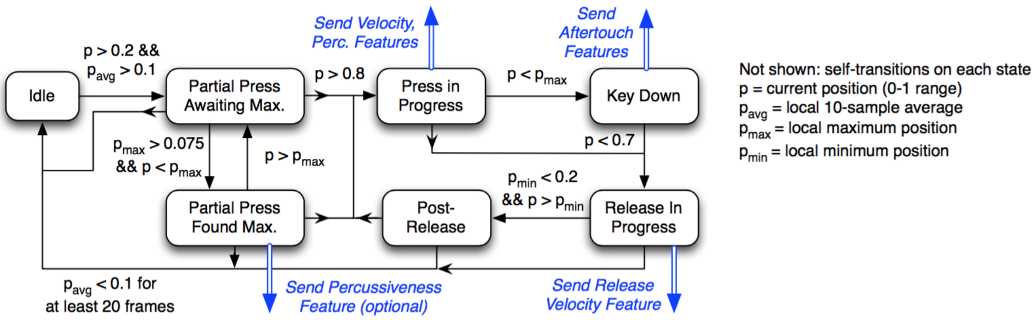


Figure 5.11: State machine diagram for the key position as implemented by [McPherson, 2013]. Figure reproduced with permission from the author.

press, and extract features that describe the behaviour of the press [Bernays and Traube, 2012, McPherson, 2013]. In McPherson's approach, a state machine is used to identify several states of the key, and transitions between states are triggered by either the crossing of a position threshold, or the detection of a local maximum or minimum in the position (velocity change), or based on a longer-term average of the key position, as illustrated in Fig. 5.11.

This state machine works very well to describe the state of discrete note events, however when used in a context where the player is actively using the continuous key position as a control, it shows some limitations in the sensitivity of the key after release, and in the detection of the "down" state, which in turn affect the computation of aftertouch.

We use this state machine, with the minor modification we describe below, as a helper for our multi-key gesture detection algorithm, and to detect percussiveness. We use a different approach to detect aftertouch, and velocity features are ignored.

Post-release

When a key is released, it normally bounces slightly. To prevent these bounces from triggering unwanted sounds, McPherson implemented a "post-release" state, from which the key could either transition to the "idle" state (20 ms after the key is stable around the rest position), or to the "press in progress" state, when the position gets close to key bottom ($p > 0.8$). This design is such that, in the context of discrete note events, as on the piano for which it was designed, no note-on events will be missed. The only drawback is that the new note will not have a percussiveness value associated with it, but this is a reasonable decision, in that it would be hard (and unlikely) for a player to perform a percussive gesture immediately after fully releasing a key.

In an instrument with continuous key control, however, an issue arises when the player starts pressing the key immediately after the release: the key will stay in the "post-release" state until the key gets close to the key bottom, while the intention of the player was to control the sound throughout the new attack gesture. One typical case when this would happen is if the player releases the key and they keep their finger on it, holding it slightly off the rest position, thus not allowing the transition to idle state, before pressing it again.

We implemented a dynamic threshold on the position of the key after release, so that the threshold to get out of the release state decreases over time, relying on the fact that successive bounces will be of decreasing amplitude We start from a lower threshold ($p = 0.4$), and at every bounce of the key, that is when a local maximum is detected, the threshold is updated with the local maximum value. Additionally, the threshold decays linearly over time, so that even if no local maximum is detected (that is, the key did not bounce), it will reach the value $p = 0$ within 0.4 s from the key being released, allowing a state transition to happen earlier in the key throw.

5.3.2 *Aftertouch*

Aftertouch is the term used to indicate the extra pressure put into the key once it reaches key bottom. The amount of aftertouch can be computed as the amount of key displacement beyond a baseline value, corresponding to the "regular" key bottom position, when no extra pressure is added. The baseline value has to be identified reliably with high accuracy, in order to avoid possible situations where a regular key press triggers a non-zero aftertouch value, or where the amount of extra pressure needed to reach a given aftertouch value is inconsistent between keys. If a static baseline is used, then the correct functioning of the aftertouch features heavily relies on an accurate calibration of the endpoints of the key excursion, and this is not always easily achievable. In McPherson [2013], the "Key Down" state is reached when the key is in the "press in progress" state and a local maximum is encountered. This local maximum is a consequence of the key hitting the keybed, and ever so slightly bouncing back. In their approach, the baseline value for the aftertouch is set dynamically for each key press as the key position at which the key bottom event is detected, thus removing the dependency on an accurate calibration.

Incomplete key presses, which go deep enough to enter the "Press in Progress" state may thus generate a local maximum which the state machine could interpret as a transition to the "Key Down" state, and set an associate baseline value for the aftertouch feature. If this is just a partial key press, and the key immediately releases, this will pose no problem. However, in the case where the local maximum is part of a gesture that is continuously modulating the position of the key, as it will be the case in an instrument like the one we are building, such approach may cause the key to remain in the "Key Down" state, reaching positions higher than the aftertouch baseline that was originally set, yielding positive aftertouch values before key bottom has even been reached. In order to avoid this possibility, and still achieve reliable aftertouch readings, we had to revert to position-based aftertouch detection. We added a "dead zone" around the calibrated key bottom value in order to avoid triggering aftertouch during a regular key press.

On keyboards not designed for aftertouch use, the extra displacement achievable by pressing into the key bed can be fairly small, and inconsistent between keys, as it depends on the properties of the felts and other mechanical parts of the keyboard action. We observed that the physical quantity most relevant to the player when performing an aftertouch gesture would likely be the pressure they put into the keybed, and not the amount of displacement of the key, which is often of the order of less than 1 mm in total. We therefore amended our calibration procedure in order to record the maximum amount of excursion achievable with aftertouch on each key for a given pressure, and we used that to normalize the aftertouch range across keys.

5.3.3 *Percussiveness detection*

A percussive key press occurs when the finger is already moving downwards before hitting the key, as we have reviewed in Section 2.1.1. For the purpose of our keyboard controller, we are aiming to obtain a discrete metric for each key onset, to quantify the amount of percussiveness. Bernays and Traube [2012] obtain a percussive metric from the ratio of the key depression at half the attack duration to the maximum key depression and the average of the key depression curve. Besides introducing a latency in the detection, by postponing the computation of the metric until the key has reached the key bottom, this approach would not work in the presence of incomplete key presses. McPherson and Kim [2011]'s approach considers the ballistic collision that causes the key to bounce off the finger shortly after the initial finger-key impact.

We show in Fig. 5.12 the key and velocity profiles of a percussive key press played on a Yamaha CP-300 electronic piano, and acquired with the piano scanner. As the key is hit by the finger, kinetic energy is transferred from the finger to the key, and the key starts a fast downward motion while it temporarily loses contact with the finger, which is still moving downwards but more slowly. The key is moving freely downwards, and the kinetic energy progressively dissipates until the key stops and eventually starts moving upwards. Shortly after that moment, the finger, which has kept moving down all along, catches up and the key starts moving downwards again, this time under the direct pressure of the finger. This behaviour translates in the velocity profile exhibiting an initial spike due to the impact, and the key profile exhibiting a local maximum during the early part of the onset, corresponding to the point where the key starts the upwards motion.

The percussion detection algorithm starts by detecting a local maximum in the key position during the early part of the key onset, while the key is in the Partial Press Awaiting Max state. When a maximum is found, the program looks back at

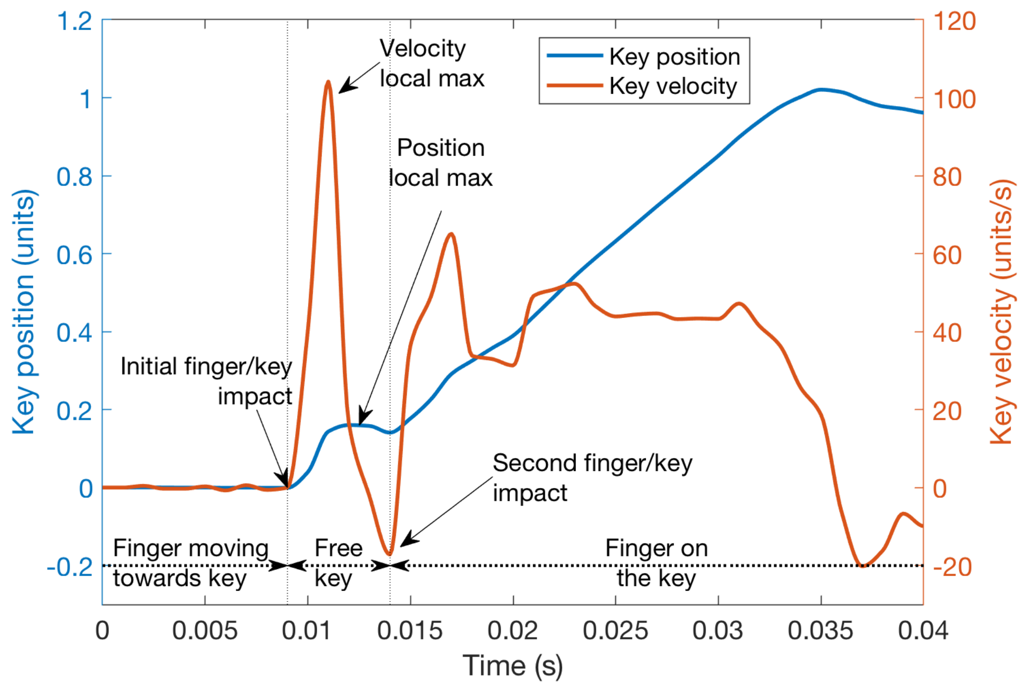


Figure 5.12: The position and velocity profile of a percussive key press.

the recent history of the key position to find the maximum value of the velocity, and that value is then used as the percussiveness metric.

5.3.4 *Multi-key gestures*

Keyboard instruments with limited polyphony need to adopt a strategy to decide what keys emit sound when more keys than the polyphony limit are pressed at the same time. In the case of monophonic synthesizers, the two most common strategies are low priority and high priority. With low priority, the leftmost note held down at any time is the one that controls the pitch of the instrument, while with high priority, it is the rightmost one. This approach is easy to implement even with analog circuitry, however it comes with a significant drawback: when playing ascending melodies with low priority, the player needs to fully release the previous key in order for the new one to start playing; and the same happens when playing descending melodies with high priority. Some synthesizers implement a "most recent" priority, which ensures that the note that was most recently pressed is the one that plays. These priority schemes only make sense in the context of discrete key states, where a key can only be "pressed" or "not pressed". However, in the case of an instrument where the key position continuously shapes the sound, like ours, a more complicated model is needed in order for the interaction to be intuitive.

5.3.4.1 *Most recent and deepest priority*

We created a priority scheme that can be defined as "most recent and deepest priority" in order to find the active key, that is the key that controls the sound generator. We formulated a set of rules to identify the active key when more than one key is pressed. These rules use the state machine described in Section 5.3.1 to inform the decision:

- The timestamps of when a key last entered the "Partial Press" or the "Key Down" states are recorded.
- If there is no key in the "Down" state, the active key is the one that is deepest in the key throw and is not in a "Release" state.
- If there is at least one key "Down", then the active key is whichever key has the most recent timestamp, regardless of whether they received their timestamp for entering the "Partial Press" or "Key Down" state, except in the case of bending gestures (see below).

- Keys in the “Post-Release” state are ignored.

We report some example behaviours to give a better understanding of the implications of the rules:

- If multiple keys are partially pressed, the one that is deepest in the key throw (i.e.: closest to the bottom of the key travel) will be the active one.
- If one key is “down” and a new key starts pressing, the new key becomes the active one. If the key never reaches key bottom it will remain active until it is fully released. This behaviour is partially overridden in case of bending gestures (see below).
- If more than one key is “down”, the one that reached the key bottom last will be the active one.
- The bouncing of a key after it is released will not cause spurious activations.

5.3.4.2 *Bending gestures*

Our instrument allows multi-key gestures to generate pitch-bending effects between keys that are within a specified interval from the active key. To enable this we need to partially override the key priority rules outlined above. If the active key is “down” and a key within the bending interval enters the “press in progress” state, a bending gesture is triggered, and we call the newly pressed key the “bending” key. If the active key enters the “Release in Progress” state, the bending gesture terminates. Similarly, when the bending key reaches the “Key Down” state, the bending gesture terminates, and the bending key becomes the new active key. As long as the original active key has not left the “Key Down” state, then releasing the original bending key will start the bending gesture again. When a bending gesture is active, the bending key is ignored by the rules for monophonic key priority.

The amount of bending is given by the position of bending key. We define the bending index i_b as:

$$i_b = \frac{p_b - t_o}{t_p - t_o}$$

where p_b is the current position of the bending key, t_o is the threshold at which a key enters the “Partial Press” state, and t_p is the threshold at which the key enters the “press in progress” state (typically, this is 75% of the key throw). Notice that in virtue of the bending gestures rules, the i_b value is constrained to lie between 0 (i.e.: $p_b \leqslant t_o$), when the bending key is only slightly depressed (as no bending occurs then), and 1, when the bending key is in the “down” state ($p_b \geqslant t_p$), as the bending gesture terminates then.

5.3.5 *Outputs of the keyboard controller*

The keyboard controller has outputs for key number, percussiveness, bending (range and index) and position. The note number output is simply the MIDI note number corresponding to the active key. The percussiveness amount is given by the percussiveness detection algorithm described above. Percussiveness can be detected on any key, even on an inactive or bending one. When no bending gesture is in progress, the position output is the position of the active key and the bending amount and bending index are zero, but when a bending gesture is in progress, these are handled differently, as explained below.

As mentioned earlier, when a bending gesture is taken to completion, the bending key reaches the "Key Down" state, and it becomes the new active key. If the position output during the bending was simply the position of the active key, then when the bending key becomes the active key, this could result in a discontinuity in the position output, which could in turn lead to audible clicks in the generated audio. To avoid this, we compute the position output (p_o) during a bend gesture as a weighted sum of the position of the active key (p_a), and that of the bending key (p_b), weighted by the bending index (i_b) with a tunable offset:

$$p_o = kp_a(1 - i_b) + kp_b i_b + (1 - k)$$

where $0 \leqslant k \leqslant 1$ allows to tune the behaviour between constant amplitude (when $k = 0$) and a crossfading with a corresponding amplitude dip (when $k = 1$).

5.4 MAPPINGS AND SOUNDS

The flute physical model in our sound engine, described in Section 5.2, exposes parameters for controlling the air pressure into the bore, the jet ratio, which corresponds to the distance between the mouth and the embouchure, the non-linearity of the bore, which affects the timbre of the sound, an auxiliary audio input to add an external signal to the air jet into the bore, and a control for the output gain. Our keyboard controller, described in Section 5.3, outputs the note number corresponding to the active key, the position of the key, the amount of pitch bending, and the percussiveness value of the current key press. This section describes the mapping layer anticipated in Fig. 5.1, which maps the outputs from the keyboard controller to the controls in the sound generator, and the sonic outcomes that can be obtained from the instrument.

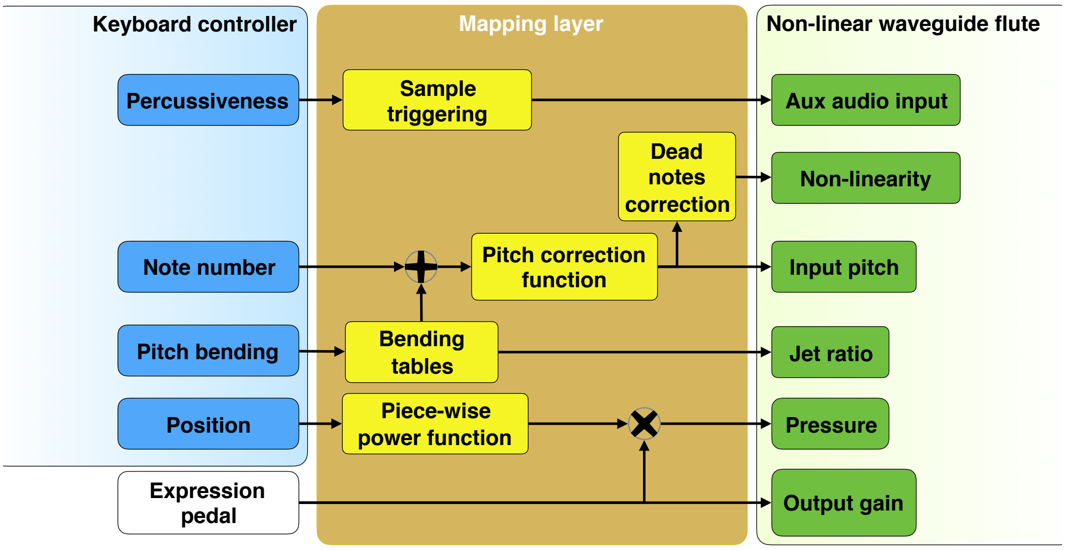


Figure 5.13: Block diagram of the mapping layer and of the relationships between the outputs of the keyboard controller and the inputs of the flute model. This represents the behaviour when the bending state machine is in the *normal* state.

In order to add a personal character to the instrument, we want the player to access some of the sonic effects based on altering the jet ratio parameter, as described in Section 5.2.2.5. The most obvious place where to add control over the jet ratio, because of its influence on the pitch of the instrument, is the pitch bending gesture. However, to keep the instrument playable and avoid sudden jumps in pitch (see Fig. 5.9), we have to restrict the available range of the jet ratio during "normal" playing, but we allow more extreme effects to be achieved by performing slower keyboard gestures. We implemented a state machine, controlled by the bending gesture, which allows to achieve a wider range of jet ratio values, making for the production of multiphonic and overblown tones.

The block diagram of the mapping layer during normal playing is shown in Fig. 5.13, and it is described in the following sections. Parts of the mapping layer are dynamic, as its behaviour depends on the state of the bending state machine, as explained in Section 5.4.4.

5.4.1 *Position and pressure*

The position output of the keyboard controller varies between 0 and 1 during the key throw, and is comprised between 1 and 1.1 in the aftertouch region. The pressure r is a function of the key position p

$$r = \begin{cases} r_o + r_r p^{e_t}, & 0 < p \leqslant 1 \\ r_o + r_r p^{e_a}, & p > 1 \end{cases} \tag{5.5}$$

where r_o is the pressure offset when the key is at rest, r_r is the pressure range, e_t is the exponent during the key throw, and e_a is the exponent in the aftertouch region. Note that this piecewise power functions is continuous at $p = 1$, where $r = 1.2$, the value we chose in Section 5.2.2.1.

We set $r_o = 0.5$, and $r_r = 0.7$, so that the key position can control the pressure range 0.5 to 1.2 during the key throw. The e_t parameter is adjustable in the range 0.5 to 1.1. The e_a parameter is adjustable in the range 2 to 6. We further clip the pressure to a value $r_{max} = 1.4$. Some of the pressure/position curves achievable with these parameters are shown in Fig. 5.14.

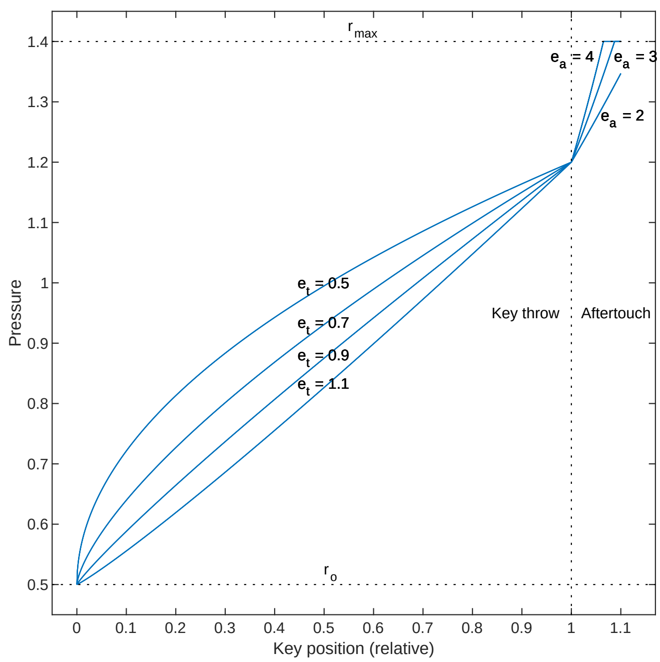
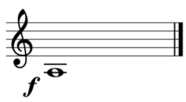


Figure 5.14: Pressure/position curves achievable for several values of parameters e_t and e_a.

Sounds

We present here some examples of the sounds and effects obtainable by leveraging the mapping between key position and the pressure of the flute model. Audio recordings of these and other examples can be found online.[10] Figures 5.15-5.20 display the pressure and key position (top), a time domain representation of the generated sound (middle), a frequency domain representation of the generated sound (bottom), as well as the notation we used to indicate the gesture (left).

In Fig. 5.15 is a regular key press that goes all the way to the bottom of the key, giving a full, rich tone. We use the dynamic notation *forte* in correspondence with the bottom of the key. Fig. 5.16 shows key press that fades in from nothing (*n*), down to the bottom of the key (*f*) and then fades out again. At the top of the key, the sound is quiet, and very breathy. As the pressure increases, the first harmonic comes in earlier, and then the others follow deeper in the key throw. Fig. 5.17 shows the effect of pressing deep into the key in the aftertouch region. When the pressure value exceeds about 1.3, it produces what we call a *growly* sound, which presents some inharmonic sidebands. When pressing into the keybed with a periodic motion, a gentle vibrato effect can be obtained when pressing lightly (Fig. 5.18), or a more intense one, which reaches the growl point, when pressing harder (Fig. 5.19). Notes of different dynamics can be obtained by pressing the key partially and sustaining it at that level Fig. 5.20



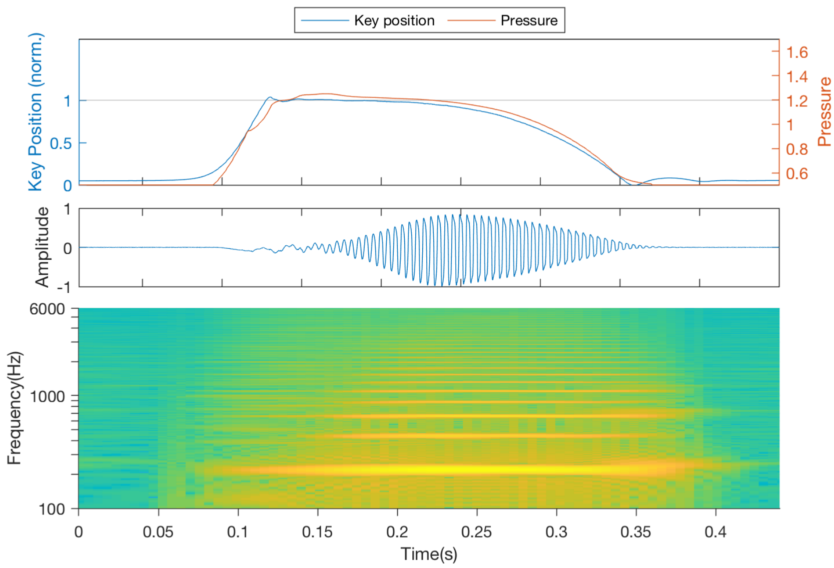


Figure 5.15: A fast key press, non percussive. The key goes all the way to the bottom.

10 "Physical modelling flute synthesizer with continuous key angle control": https://vimeo.com/352530701

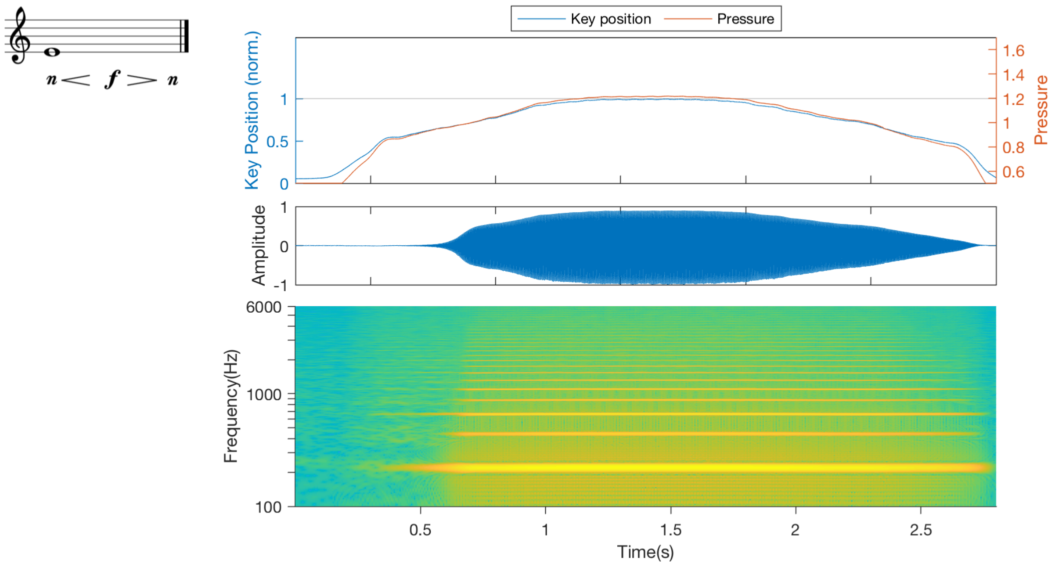


Figure 5.16: A note faded in and out by progressively pressing and releasing the key.

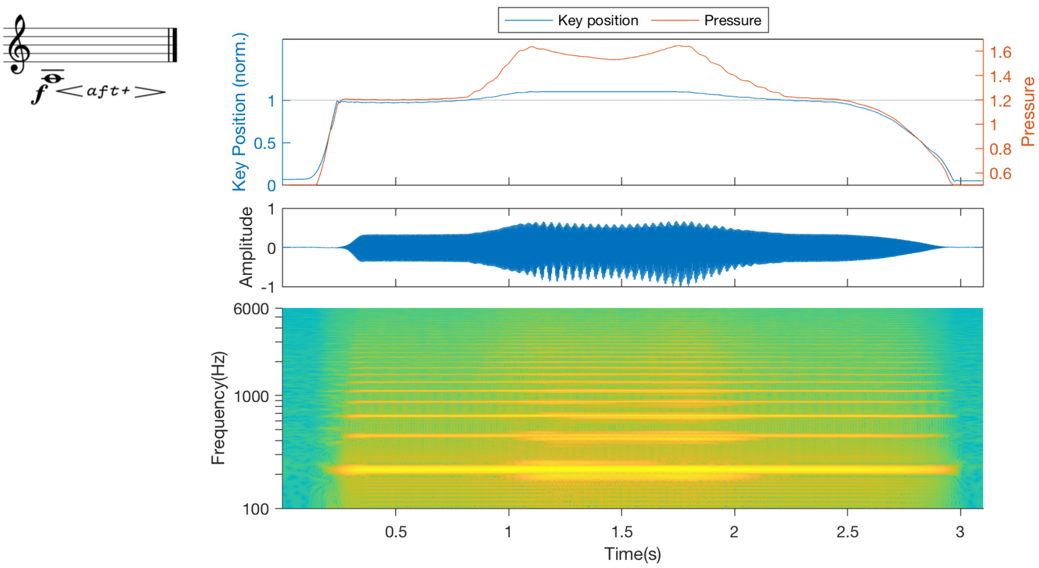


Figure 5.17: Pressing the key into the keybed (aftertouch), to obtain a "growl" sound. Notice that the mapping between key position and air pressure changes when entering the aftertouch region (between key position 1.0 and 1.1), as shown in Fig. 5.14.

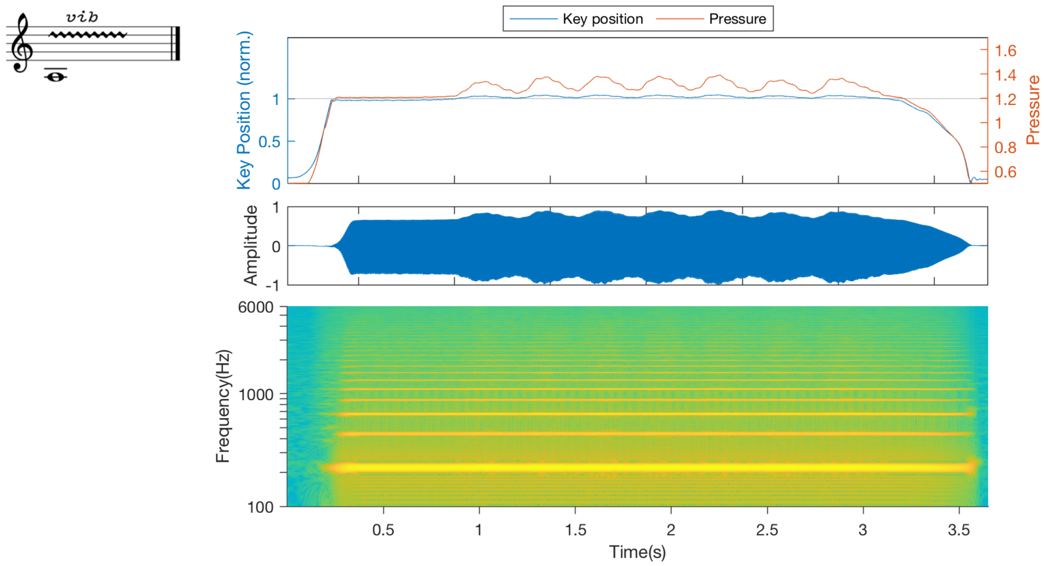


Figure 5.18: Pressure vibrato, obtained by pressing gently into the keybed repeatedly.

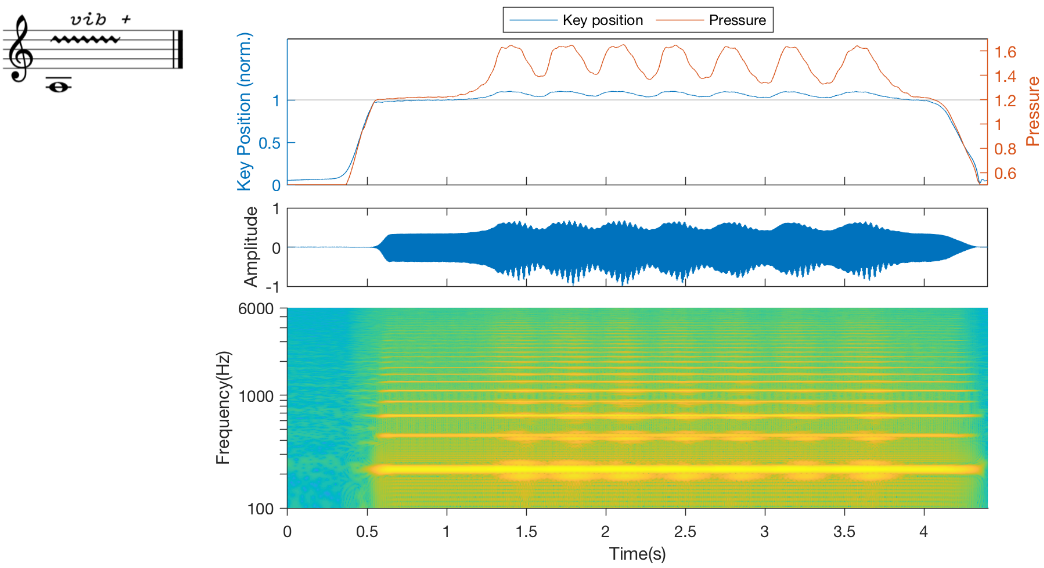


Figure 5.19: Growl vibrato, obtained by pressing heavily into the keybed repeatedly.

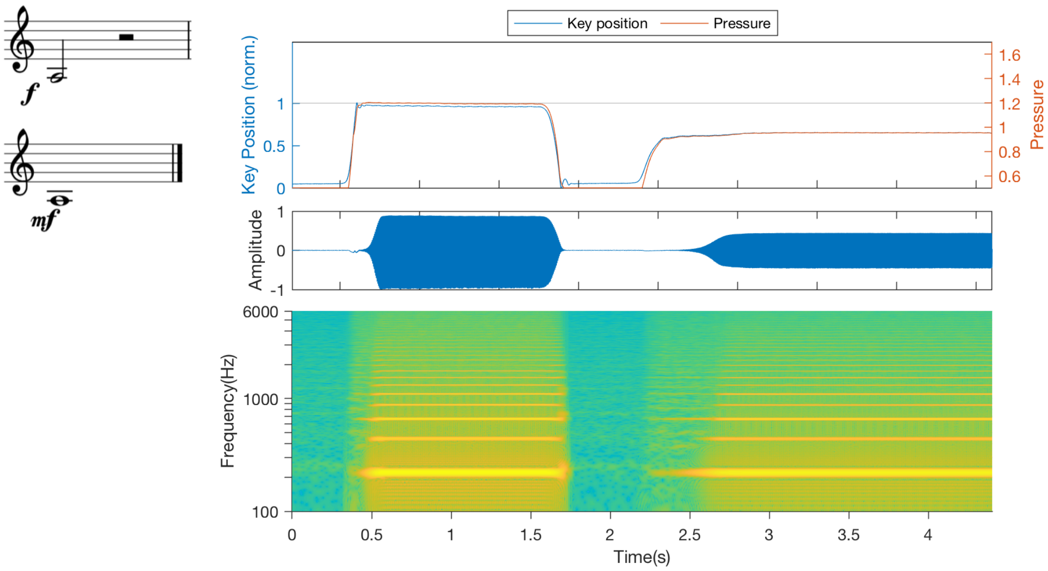


Figure 5.20: Two notes, the first fully depressed (*forte* dynamic), the second one partly depressed (*mezzoforte* dynamic).

5.4.2 *Percussiveness*

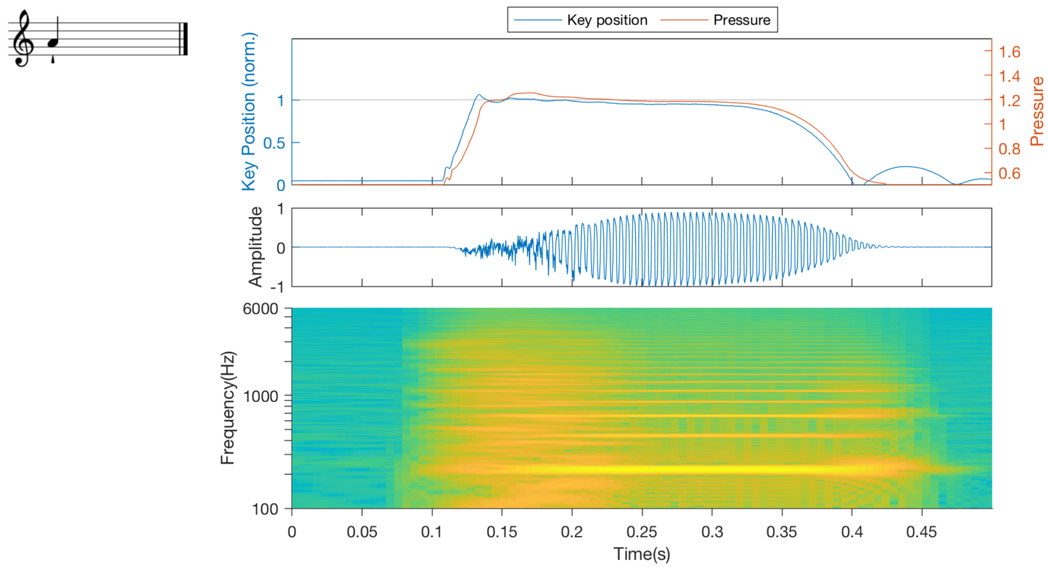


Figure 5.21: A percussive key press. Notice the spike of the key position at the beginning of the note and the consequent noisy burst in the audio signal.

When a percussive key press is detected, a percussive sound is injected in the model through the auxiliary audio input. The percussive sounds we use are taken from a collection of recordings of a person producing a "T" sound with their mouth into a microphone. This is not strictly equivalent to the effect of a flute performer pronouncing a "T" sound into the mouthpiece, however, the resulting "chuff" is reminiscent of the sound of a sharp attack on a flute. There are 6 samples loaded in our instrument, which are triggered in a cycle. The level of the sound is scaled by the magnitude of the output value of the percussiveness detector. Fig. 5.21 shows an example of percussive press.

5.4.3 *Pitch bend*

When the player holds down one note and presses a neighbouring one (up to one major third away), we want the sound to gradually bend from the old note to the new one. The keyboard controller outputs the bending range b_r (the number of semitones in the bending interval) and the bending index b_i (the current position in the bending gesture, normalized between 0 and 1). A straightforward mapping of these data to the sound engine would be to use these values in combination with the active key number (k_a) to obtain the (fractional) note number to use as the pitch input into the sound model (p_i):

$$p_i = k_a + b_i b_r$$

With this approach, the pitch bend would be obtained simply by changing the length of the bore in our slide flute model, however we found the sound resulting from this interaction uninteresting and artificial sounding. Bending a note on a transverse flute is done in practice by changing the distance between the upper lip and the mouthpiece, therefore resulting in a timbral change during the bending, and we have previously shown that with our model, a corresponding effect can be achieved by changing the jet ratio (see in Section 5.2.1.1 and Section 5.2.2.5). While we have not validated whether what we observed has an exact counterpart in an acoustic instrument, we will still exploit it for our purposes, because it produces a more lively sound during pitch bends.

We employed the following approach in order to obtain smooth bending curves combining the effect of jet ratio and input pitch correction to bend to notes up to a major third apart from the current one (± 4 semitones):

- the first half of the bending range ($i_b \leqslant 0.5$) is covered only adjusting the jet ratio.
- the second half ($i_b > 0.5$) is obtained reducing the jet ratio while at the same time increasing the input pitch correction.

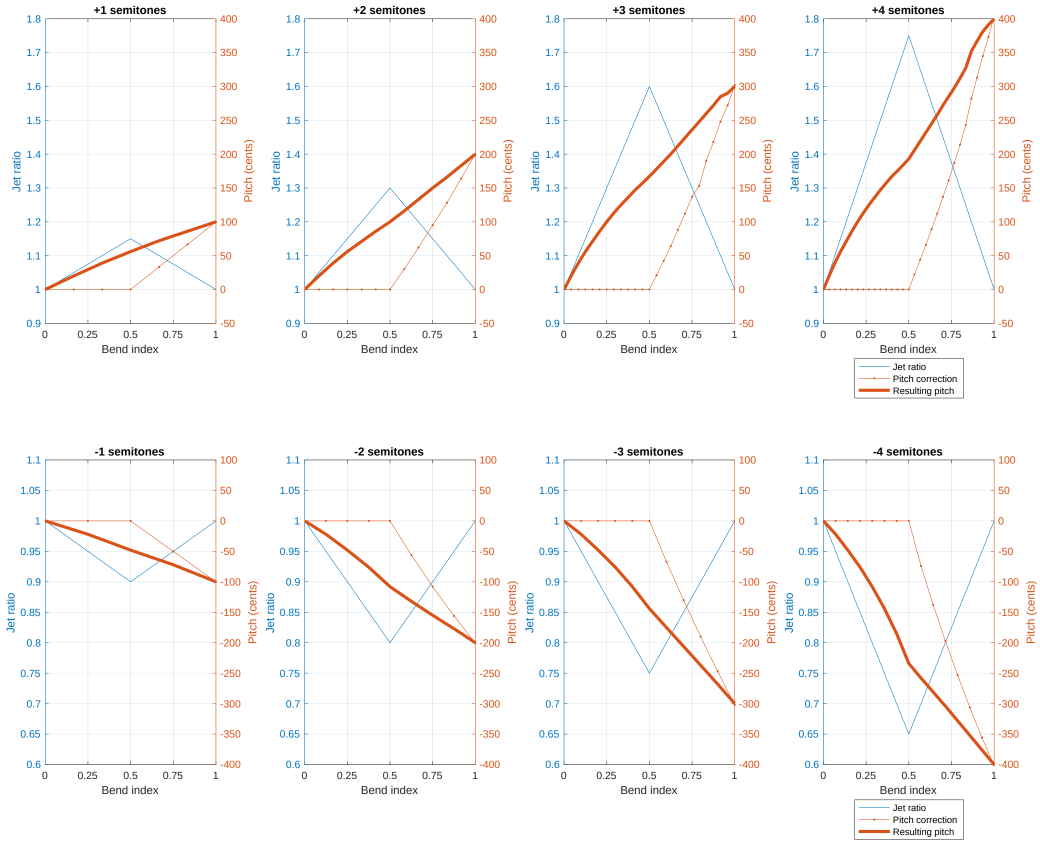


Figure 5.22: Pitch bending curves. The effects of the input pitch and the jet ratio are combined to obtain the desired output pitch.

- at the midpoint $i_b = 0.5$, the output pitch should be approximately to half of the bending range.
- the jet ratio as a dependency of i_b is piecewise linear, and symmetrical with respect to $i_b = 0.5$.
- the output pitch has to be monotonic, and approximately linear with respect to the bend index.

With this procedure we obtained the curves in Fig. 5.22. As the bending index approaches 1, the jet ratio is approaching the initial value of 1, so that when the bend gesture is completed, and the bending key becomes the new primary key, the transition is guaranteed to be smooth. Jet ratio values above 1.6 and below 0.7 are needed to achieve the larger bending intervals in Fig. 5.22. We showed in Fig. 5.8 that for such values the pitch may jump, but, as long as the gesture is swift, this will not happen.

5.4.4 *Bending state machine*

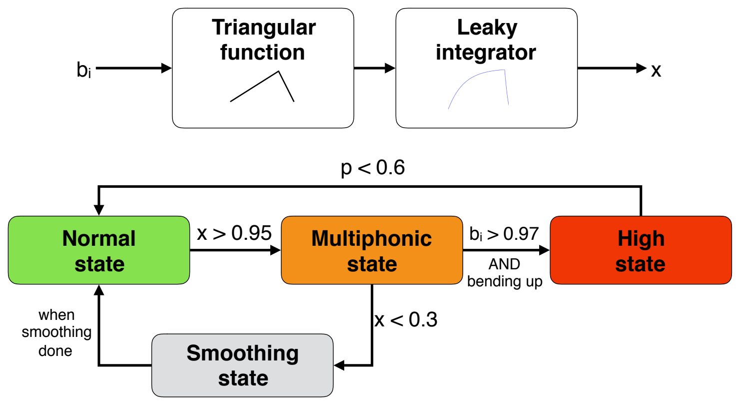


Figure 5.23: The bending state machine. b_i is the bending index, p is the normalized position of the bending key, x is the output of the leaky integrator.

On our instrument, swift bend gestures result in smooth pitch transitions, during which the frequency change is obtained as a combination of adjusting the jet ratio and the fundamental frequency of the resonant bore, so that the fractional jet ratio gives a characteristic timbre to the note during the bending. This behaviour is detailed in Section 5.4.3.

When a bending gesture is performed slower than normal, and the key spends a longer time in the middle of the bending, the instrument starts behaving in an unstable way, producing multiphonic sounds, and then possibly reach a more stable overblown state, where the resulting pitch is one octave above the key being pressed. Throughout the process, the player remains in direct control of the instrument through the position of the bending key. The states and transitions for the bending state machine are shown in Fig. 5.23.

The bending index b_i is fed through a triangular function:

$$\mathrm{tri} = \begin{cases} 1.2 b_i / 0.7 & , 0 \leqslant b_i \leqslant 0.7 \\ 1.2 \frac{0.9 - b_i}{0.9 - 0.7} & , 0.7 < b_i \leqslant 0.9 \\ 0 & , 0.9 < b_i \leqslant 1 \end{cases} \tag{5.6}$$

The value of tri in Eq. (5.6) will be higher when b_i is close to the mid value of 0.5, and smaller when b_i is closer to the extremes (0 or 1). The output of the triangular function is then clipped between 0 and 1 and fed into a leaky integrator with time constant 0.4 s. The output x of the leaky integrator will behave as a smoothed version of the output of the triangular function, on a longer time scale, and is used as one of the parameters that control the state machine.

The state machine starts in the *normal* state. If the output leaky integrator exceeds a threshold, the state machine enters the *multiphonic* state. When in the *multiphonic* state, if the leaky integrator value goes below a second, smaller threshold, x_{t2}, then the state machine goes back to the *normal* state, through a temporary *smoothing* state. If the gesture is an upwards bending, and, while in the *multiphonic* state, the bending key reaches key bottom ($b_i = 1$), and , then the state machine enters the *high* state. The state machine will remain in the *high* state until the bending key that triggered the transition is released below a threshold, at which point the state machine switches back to the *normal* state. Other notes can be played while in the *hight* state, as long as the transition key is held down.

NORMAL STATE This is the default state of the instrument. During a pitch bend, while in the normal state, the values of jet ratio and frequency are obtained as explained in Section 5.4.3. As a result, the output pitch of the instrument will increases roughly linearly with b_i.

MULTIPHONIC STATE When entering this state, the initial jet ratio and frequency values are stored, as well as the initial bending index b_i. As the bending key is moved from the initial position, the frequency value is interpolated linearly between the initial value to the bending range value, and the jet ratio is interpolated between the initial value and 2 when bending up, or 0.2 when bending down.[11] As the jet ratio value gets farther away from 1, multiphonic sounds are generated, as described in Section 5.2.2.5. When bending up, as the jet ratio offset approaches the value of 2, the sound returns to be monophonic and harmonic, and it approaches the pitch one octave above the bending key. When bending down, the sound will just stay multiphonic / unstable.

SMOOTHING STATE The purpose of this state is to smooth the transition from the *multiphonic* back to the *normal* state, avoiding the discontinuities that would be introduced by the jet ratio offset and frequency offset instantaneously jumping between the two.

11 The reason why we ramp towards 0.2 is that by going closer to zero the sound production tends to stop, as Eq. (5.2) diverges.

HIGH STATE In this state, the jet ratio is controlled by the position of the key that triggered the transition, even when the active key is another one. If the "transition" key is fully depressed, the jet ratio is 2, corresponding to an overblown tone, one octave above the pressed key. This effectively allows to play any note in the overblown register, as long as the initial key is held down. Partially releasing it will bend the tone down, as the jet ratio decrease, eventually jumping down to a non-overblown note when the transition to the normal mode takes place.

PITCH BENDING EXAMPLES An example of a pitch bend where the state machine stays in the normal state throughout is shown in Fig. 5.24. In Fig. 5.25 a slow bending is performed, so that the state machine enters the multiphonic state and then goes back to the normal state through the smoothing state. In Fig. 5.26 is a pitch bend where a swift motion of the key while in the transition stage completes the bending gesture ($b_i > 0.97$) and at the same time the output of the leaky integrator is still above the threshold ($x > 0.3$), thus reaching the high state.

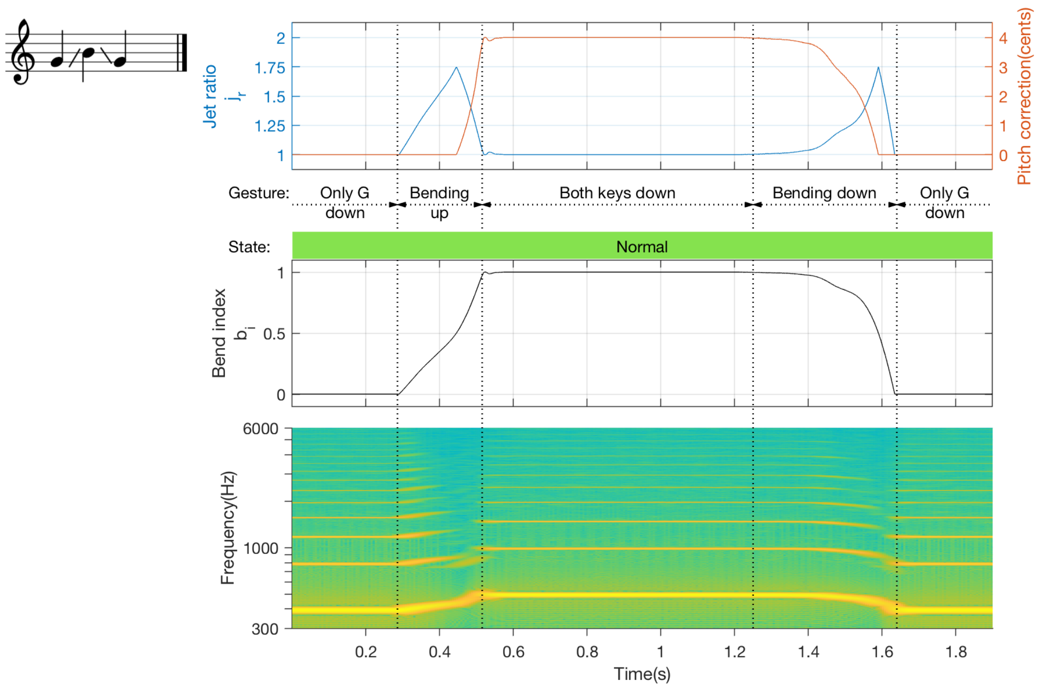


Figure 5.24: Pitch bend between a G_4 and a B_4 in the normal state.

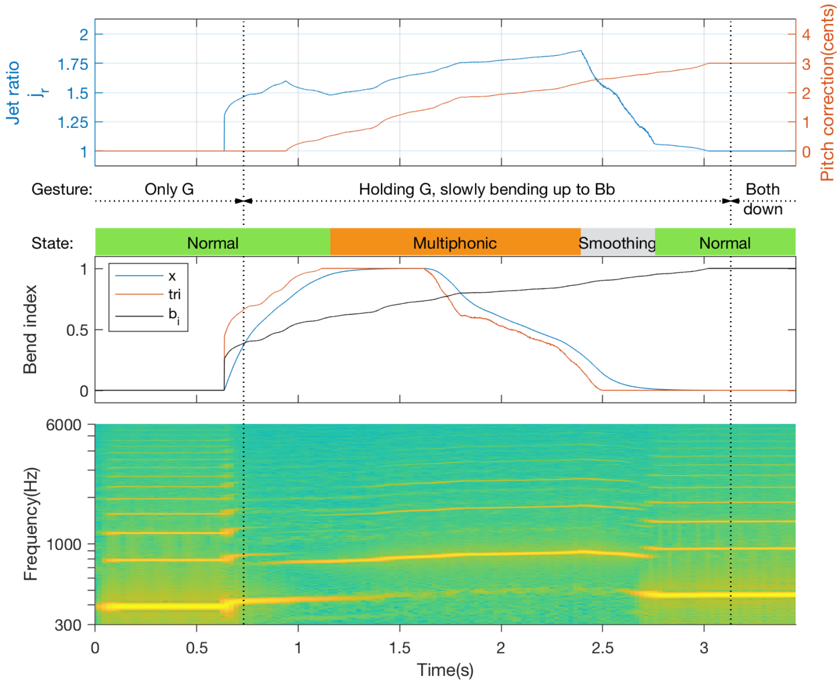


Figure 5.25: Pitch bend from a G4 and a B♭4 which enters the multiphonic state and then, through the smoothing state, back to the normal state. We did not develop a notation for this gesture.

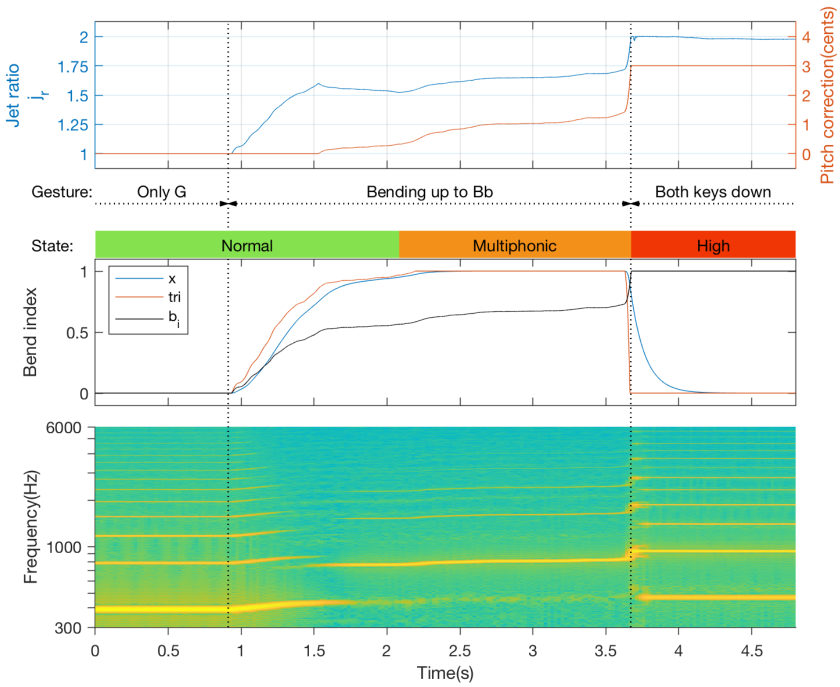


Figure 5.26: Pitch bend from a G4 and a B♭4 which reaches the high state.

5.4.5 *Expression pedal*

An optional expression pedal can be connected to the instrument to give a limited amount of control on the range of the pressure, and on the overall output gain of the model. By controlling also the pressure and not only the gain, we intend to give a more natural effect of dynamic. The effect of the pedal for several positions are reported in Table 5.1. For the pressure, intermediate pedal positions interpolate linearly between the two extreme values, whereas for the gain they follow a quadratic curve.

Position	Reading	Pressure scale	Gain
Heel	0	0.85	0.2 (-14 dB)
Mid	0.5	0.925	0.4 (-8 dB)
Toe	1	1	1 (0 dB)

Table 5.1: Effect of the expression pedal.

5.4.6 *Model corrections*

We showed in Section 5.2.2 some fine adjustments that can be performed on the sound model in order to prevent dead notes and to improve tuning. In the mapping layer, we apply pitch correction via Eq. (5.4) to the sum of the note number input received from the keyboard controller and the frequency adjustment generated by the pitch bend. We also apply non-linearity correction to avoid dead notes, using Eq. (5.3). We do not apply any correction for the effect of pressure on the pitch. As a consequence, notes will be exactly in tune only when fully depressed ($r = 1.2$), and slightly off when pressing the key partially or using the aftertouch.

5.5 IMPLEMENTATION

We implemented this instrument on the platform we previously developed in Section 4.3.2. We fitted two boards of the piano scanner on a Yamaha CP-300 digital keyboard[12], covering the range from B♭3 to B6 (37 notes). None of the sounds or electronics from the Yamaha were used. The sound is reproduced via a Yamaha MSP-5A speaker, placed in front of the keyboard, facing the performer. A Bela

12 https://uk.yamaha.com/en/products/music_production/stagekeyboards/cp300

Figure 5.27: The nonlinear waveguide flute synthesizer with continuous key control: the Yamaha CP-300, with two boards of the keyboard scanner installed, Bela (hidden underneath the black cloth), and the speaker.

board sits on top of the flat surface behind the keys and is hidden from sight with a dark cloth, and is connected to the piano scanner and to the speaker. A picture of the instrument is shown in Fig. 5.27.

The DSP code is written in FAUST, and it is largely based on Michon's code, on top of which we implemented the modifications to the flute model described in Section 5.2. The FAUST compiler produces a C++ file that contains the DSP code and provides wrapping code for the platform it will run on, using what is called, in FAUST jargon, an *architecture file*. We modified the default architecture file for Bela in order to interact with other parts of the software which we wrote in C++. The C++ code performs the following functions:

- it provides a low-level interface to the keyboard scanner (Section 4.3.2)
- it processes the raw data from the keyboard scanner to generate the pitch, position, percussiveness and bending outputs of the keyboard controller (Section 5.3)
- it maps the outputs of keyboard controller to the frequency, pressure, jet ratio, gain, and non-linearity of the sound model, and the bending state machine, and applies the correction curves, as detailed in Section 5.4

The code for the key state machine and percussiveness detection is based on the code written by Andrew McPherson for McPherson and Kim [2011], McPherson [2013]. Our full code is available online[13].

5.6 STUDY DESIGN

We designed and ran a performance study with our continuous key physical modelling flute synthesizer and gathered sensor, audio and video recordings of the participants during the approximately six hours they spent with the instrument.

5.6.1 *Procedure*

The study takes place in two sessions, to be held on different days, each divided into multiple tasks, as detailed below.

SESSION 1

FREE EXPLORATION 1 (15 MINUTES) Without instructions, the player explores the instrument on their own, the investigator leaves the room.

INTERVIEW 1 (10-20MINUTES) The player explains their findings and initial impression. The investigator describes the capabilities of the instrument in the "normal" state: key position and aftertouch affects pressure, pitch bending between neighbouring notes, percussiveness. The existence of the "multiphonic" and "high" states is not mentioned.

TRAINING (30-60 MINUTES) We prepared a set of training exercises[14] to guide the player through the techniques available on the instrument, whose scores are available in Appendix B.1. The materials are accompanied with pre-recorded audio example performed on the instrument.[15] The exercises presented in this task cover different types of touch (slow, fast, marcato, staccato), partial and progressive key presses to obtain crescendo and decrescendo, use of aftertouch for vibrato and growl, pitch bends (Figs. B.1 and B.2). These include longer melodic fragments where the techniques are combined together (Fig. B.3). The participant is presented with the score of each exercise and can listen to the examples as many times as they wish before or while they play it. Through-

13 https://github.com/giuliomoro/flute-key/

14 Kindly contributed by dr. Andrew McPherson

15 The pre-recorded audio example for the training materials are provided at the URL in Footnote 10 on page 147.

out the study, the investigator gives feedback and suggestions to the player.

FREE EXPLORATION 2 (10 MINUTES) Without instructions, the player explores the instrument on their own, the investigator leaves the room. The scores from the training material are left available to the player.

INTERVIEW 2 (15-30 MINUTES) A semi-structured interview where the investigator gets insights on the general impression, struggles, how the techniques are being learned, how they compare with traditional keyboard playing, and whether any additional techniques have been discovered.

BREAK (20 MINUTES)

PERFORMANCE EVALUATION 1 (30 MINUTES) We prepared eight études, which are slightly more difficult than the melodic fragments presented during the training. The first four études (Fig. B.4) have accompanying pre-recorded audio, and the player is given a score where only notes and rhythm have been notated, but none of the extended techniques of the instrument. Based on what they hear in the recording, the player first has to annotate the score with the extended techniques. The remaining four études (Fig. B.5 are fully annotated, and come with no pre-recorded audio, so the player has to perform them based uniquely on the techniques notated.

COMPOSITION (PART 1) (20 MINUTES) The participant is informed that they will have to produce a short composition (about 5 minutes) with the instrument, to be completed at the beginning of the next session. We ask that, by the end of the composition process, the piece will have to be notated with enough details for them to be able to perform it again in the near future. The purpose of this task is just to start collecting ideas for the composition, and perhaps exploring some more of the instrument. The investigator leaves the room.

SESSION 2

COMPOSITION (PART 2) (45-60 MINUTES) The participant has to complete their composition. They are given access for the first time to the expression pedal. The investigator leaves the room.

PERFORMANCE (15 MINUTES) The player performs at least twice the piece they composed and they indicate which take they consider to be their best.

INTERVIEW 3 (15-30 MINUTES) A semi-structured interview during which the player explains the compositional process, reflects on the affordances of the instrument, explains the notation.

BREAK (20 MINUTES)

PARAMETER ADJUSTMENT The player is given access to an interface that allows to adjust, separately, the response curve along the key throw and the response curve of the aftertouch. They are invited to find their favourite setting and comment on the decision. The settings they pick during this task are used for the remaining of the study.

PERFORMANCE EVALUATION 2 (30 MINUTES) Same as the Performance evaluation 1 task above, except that they start the annotation process of the first four études from the sheet they already annotated during the first session.

INTERVIEW 4 (20 MINUTES) During this semi-structured interview, the player is asked to describe the evolution of their technique in the course of the study and provide suggestions on how the instrument could be improved. Additionally, the topics from the previous interviews are re-visited.

5.6.2 *Setup*

A sketch of the layout of the room during the experiment is in Fig. 5.28. The investigator uses their computer to play back the pre-recorded audio examples during the training and performance evaluation tasks. The expression pedal is on the floor beneath the keyboard, and is kept out of sight throughout the first session. The Akai MIDIMIX is used by the participant for the parameter adjustment task, and is kept out of sight until then. The desk to the right of the participant can be used as a working space to write the score of the composition and annotate the scores during the performance evaluation tasks.

On the investigator's computer, the Reaper DAW is used to record the audio from the session using a DPA4060 lavalier microphone, which rests on the keyboard during playing, as is placed on the participant during interviews. On Bela itself, the sensor data and the generated audio are logged to disk. Separately, a video camera placed in front of the instrument is recording video and audio of the whole session.

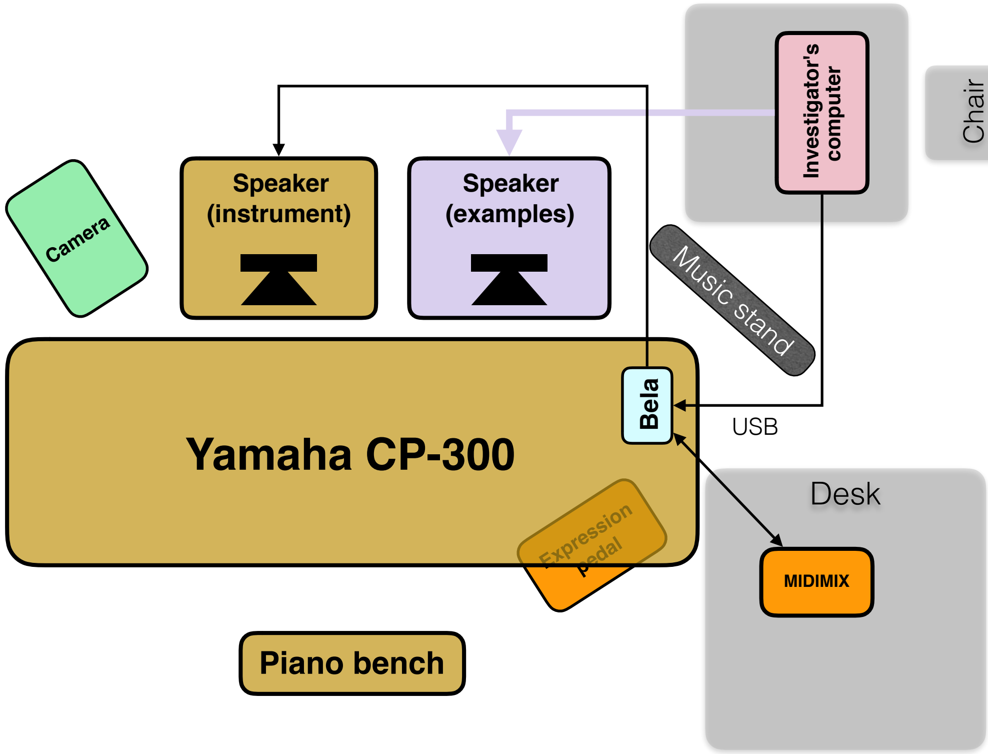


Figure 5.28: Sketch of the layout of the room during the experiment

5.6.3 *Participants*

Participants in our study have a professional or semi-professional music career, and they were paid a professional fee for their time. A total of seven participants were recruited, but only six completed the study. P1, P3 and P4 are contemporary piano players, P2 is a classical and jazz piano player, P5 and P6 mostly play jazz and rock music. P1 and P2 also took part in our Hammond study, where they were P6 and P8, respectively

5.7 RESULTS

5.7.1 *Initial discovery*

Table 5.2 summarizes the outcomes from the initial discovery, as deduced through observation of the video recordings of the initial free exploration task, and the follow-up interview. All players realized by the end of the task that the key position had a continuous control on the produced sound. While P1 and P6 discovered

	Continuous control	Pitch bend	Aftertouch	Percussiveness	Multiphonic	High state
P1	**U**[0:00]	**U**	**N**	**N**	**U**	**P**
P2	**U**[4:23]	**E**	**N**	**N**	**E**	**N**
P3	**U**[9:20]	**E**	**U**	**N**	**E**	**P**
P4	**U**[6:58]	**U**	**N**	**P**	**E**	**E**
P5	**U**[4:25]	**E**	**N**	**E**	**E**	**N**
P6	**U**[0:10]	**U**	**N**	**N**	**N**	**N**

Table 5.2: Summary of the results of the initial discovery. For each participant we indicate what which features of the instrument they discovered, and to what extent they explored them:
N: not produced. The effect was not audible at any point through the exploration.
P: produced. They produced the effect but did not actively explore it thereafter.
E: explored. They spent time investigating the effect.
U: understood. Following exploration, they manage to reliably produce the effect and understand the techniques involved.
In brackets, the time (mm:ss) from the beginning of the task when they first become aware of the effect of continuous control on the pressure of the sound.

it immediately, it took the others several minutes to realize it. P4 had already discovered the pitch-bending effect, and explored it for about 1 minute, before it occurred to them to explore the effect of key position in a single-key gesture. A handful of seconds into their exploration, P3 executed a short series of repeated partial key presses of increasingly greater depth, but from their reaction, we understand that they did not notice the effect of continuous key position. They subsequently discovered and explored aftertouch and pitch bending for several minutes before realizing the effect of key position in single-key gesture. P5 had an epiphany moment while playing in the lower register of the instrument, where the attack of the sound is by nature slower. From the video, it seems that they initially thought that the velocity would affect the ramp up time of the onset, and only after 3 repeated slow presses, they noticed that the control was position-dependent, and reacted by smiling visibly.

Most players discovered the pitch-bending gesture, although with different degrees of understanding. P2, P3 and P5, for instance, did not fully realize that the primary key had to be held down in order for the bending gesture to work. The limited range of the pitch bending was explored and understood by most. Several players also encountered the multi-phonic effect achievable with multi-key gestures, and often spent a significant amount of time playing with it, but only P1 would later formalise how to systematically achieve them: *"if I press the neighbouring key very, very slowly"*. While experimenting with the multi-phonic sounds, some players managed to enter into high state, but only P4 realized that holding down the transition key is needed in order to stay in this state. Confused by the apparent inconsistency in the reproducibility of the bending and multiphonic gesture, P5 would later report that they thought the settings of the instrument had been changing during the task.

P5 was the only one to notice the effect of percussive gestures during this task, although others occasionally triggered the effect without noticing. The discovery took place while playing two-handed fast repetitions on the high D♭ note. They quickly got a fairly reliable technique to achieve the effect, and immediately integrated it in a funk-style bass-line. P1 reportedly tried to look for effects of staccato notes during the initial free exploration: *"I tried to play short notes earlier, but didn't find anything special"*.

P3 was the only player to discover aftertouch, and also the only one to try and slide the finger on a key, which they did in the very early part of their exploration, to test whether the gesture produced any audible effect. They are also the only participant who spent a considerable amount of time trying to understand what note would play when two notes are pressed at the same time, and would look confused by the "most recent" priority we implemented, as it may seem to produce random results depending on the minute timing differences between the gestures on the two keys.

P1, the only Magnetic Resonator Piano player among our participants, was familiar with the visual appearance of the keyboard scanner and its capabilities. They therefore jumped straight into exploring the effect of key position, and the pitch bending techniques that are possible on the MRP, and quickly discovered the *"freaky multiphonics"* that occur during slow bending gestures. While aftertouch has an effect on the MRP, the player rarely employs it in their own performances, because they are not comfortable with the ergonomics it requires, and they did not explore its behaviour on our instrument.

All players associated the sound of the instrument with a woodwind instrument, mostly identifying it as a flute, although also Andean (pan) flute, clarinet, oboe and bassoon (in the low register) were mentioned. The effect of the key position during a single-key gesture was variedly described as controlling the *"air on the attack"*

(P3), the *"emission"*(P2), the *"breath"*(P6), or more simply the *"dynamic"* (P4) or *"volume"*(P5). P4 would describe the instrument as *"temperamental, very sensitive"*, because of its touch responsiveness. P5 demonstrated to have fully understood the effect of continuous key position, however, their initial explanation of the key response was given in terms of velocity (*"It's very velocity-sensitive. The velocity curve is very steep"*), but they struggled to fully articulate their thoughts.

5.7.2 *Learning process*

After the initial discovery and our explanation of the techniques, participants had a grasp of what the effect of individual on the instrument was, however they had not fully realised what potential musical results they could expect to achieve, when they would be used in the context of a musical phrase, or a larger piece. The training session thus helped to better understand the expressive potential of the instrument. At the end of Session 1, P2 commented:*"My impression is very different from before, because I learned there are many features, and it is not so easy to play. It's impressive"*. P3 remarked, during the training: *"It's like flautist training"* and by the end of it added: *"It's growing. So much possibilities and things to explore. It's good to discover [through the training] all the possibilities that I would have not otherwise encountered"*.

Most players spontaneously reported that the days intervening between the first and the second session helped them internalise the behaviour of the instrument, even though none of them practised the techniques on a keyboard in those days. P1 highlighted the increased feeling of security during the second session: *"I feel psychologically more secure, though I am not sure if I am playing more reliably. Revising today came much quicker than the first time it did"*. P6 reported *"it just worked better today than in the previous session, as I thought about the instrument a bit , and then the concept of it embedded in a bit more clearly. [...] There is a bit more instinctiveness to it. But far from being very natural and intuitive"*. P5 found it useful to film themselves during the last part of the first session and watch it back during the week: *"Between the sessions I have been thinking about it and how to achieve the sounds"*. Even P4, who undertook the two sessions only 24 hours apart, had acquired more confidence by the beginning of the second one: *"It felt better to come in today"*.

Technique	Session	TP	FP	FN	Precision	Hit rate	F-measure
Marcato	1	38	5	4	0.884	0.905	0.894
	2	39	6	3	0.867	0.929	0.897
Staccato	1	17	3	91	0.850	0.157	0.266
	2	17	4	91	0.810	0.157	0.264
Bend	1	29	6	1	0.829	0.967	0.892
	2	30	4	0	0.882	1.000	0.938
Aftertouch	1	11	1	1	0.917	0.917	0.917
	2	11	1	1	0.917	0.917	0.917
Vibrato	1	36	0	0	1.000	1.000	1.000
	2	36	0	0	1.000	1.000	1.000
Fade in/out	1	22	0	2	1.000	0.917	0.957
	2	23	0	1	1.000	0.958	0.979
Dynamic	1	2	2	4	0.500	0.333	0.400
	2	2	1	4	0.667	0.333	0.444
Bending trill	1	2	0	4	1.000	0.333	0.500
	2	2	0	4	1.000	0.333	0.500

Table 5.3: Results of the annotations performed by participants during the two sessions.
Legend: **TP**: True Positives, **FP**: False Positives, **FN**: False Negatives
Precision: $\frac{TP}{TP+FP}$
Hit rate, or recall: $\frac{TP}{TP+FN}$
F-measure: $2\frac{precision \cdot hitrate}{precision+hitrate}$

5.7.3 *Études*

We evaluated études 1-4 as annotated by the participants (Appendix B.2) against the musical score that was used to perform the corresponding pre-recorded audio example (Fig. B.4). The results, grouped by technique, are reported in Table 5.3. Most techniques have high precision and hit rate values, indicating that participants could reliably detect the sound produced by the new techniques. Staccato, however, was rarely annotated by participants, despite its prominent appearance in both étude 1 and étude 3. Only two participants realized that the trill at bar 3 of étude 3 was to be played holding on one note, thus generating a bending effect.

The "Dynamic" entry refers to dynamics annotation outside a fade in/fade out, which, in the études considered, only occurs once (the "piano" on the last note of étude 1), and was only notated correctly by two participants.

Only P3 made mistakes in annotating the marcato in étude 3, and they did so by assigning it to the note next to the one that was played marcato, possibly because of an oversight. P1 notated bends between all the notes in bars 4 and 5 of étude 4, event though only the first and last were actually performed with pitch bending in the original recording. They may have been prompted to do so by the overall legato feel of the descending phrase. Bar 5 of étude 2 includes a "long bend" across an interval of a major 5^{th}, from F4 down to B♭3. but this technique was not illustrated during the training, so that the first time the players encountered this étude (during session 1), they had not been trained as to how to perform it. While all participants notated a bend between the two notes during the first session, only P1, P2, P5 and P6 understood how to perform it during Session 1. P3 did not play the bending gesture, despite writing it down. P4 played the bending gesture between the two notes, even if no bending sound was produced, resulting in the F ending abruptly, and the B♭ fading in. The same technique was later presented, with a brief explanation, in étude 7, so that all players managed to successfully play the passage when they encountered étude 2 again in Session 2.

Most players used for the extended techniques the notation proposed during the earlier training. P1 and P2 alternated between using a oblique bar and a slur to indicate bending. P3 would use textual expressions for most of the techniques, except for the marcato, indicated with a ∨. P3 also included additional annotations such as "clean" or "det" (detaché) to explicitly indicate the transition between notes.

While we did not perform a full formal evaluation of the performances of the études, we observed that most techniques could be played reliably by all participants. The only one where all struggled significantly was the marcato (percussiveness), which also negatively impacted the rhythmic accuracy of the whole phrase. A brief analysis of the percussiveness in the études from the second session is presented in Section 5.7.5.1

5.7.4 *Compositions*

A brief description of each composition follows. The full scores created by the participants can be found in Appendix B.3. Recordings of the compositions of those players who gave us permission to share them can also be found at the URL in Footnote 10 on page 147.

P1's piece was completely improvised, and their score only comprises a list of annotations with instructions on what techniques or sounds to use in each section of the piece (Fig. B.13). They built the piece around the sounds that they found the most attractive. *"I like the breathy kind of tones, I like that you can sustain it, but also the fragility of it, the fact that it is unpredictable, vulnerable, it can break, it can stop abrupt or go to a full tone"*. They used most of the techniques extensively, particularly leveraging the ones that produce the most unusual sounds (growl aftertouch, multiphonics and high state). To work around their struggle in getting the percussive gestures reliably, they included it in their piece in a loud section (*"build up, adding percussive elements"*), where there was no requirement for all of the attempts at achieving it to succeed: *"[The percussive notes in that section] don't have to be reliable: I just know that I will get some in the mix"*.

P2's piece was composed as a study for clarinet or other woodwind. It is based around a simple melody, which is repeated and elaborated using staccato, marcato and aftertouch as variations Fig. B.14. In section 3, the percussive notes are incorporated naturally in the middle of the phrase. The "lento" part in section 6 uses progressive key presses and pitch bends.

P3 tried in their piece to explore the possibilities of the instruments. At the end of the session, they were thinking of writing a piece where they would try to emulate the behaviour of woodwinds, but during the intervening days they decided to do something completely original, because *"we already have woodwind instruments, so there is no point in [emulating] that"*. The piece is notated in standard notation, and includes a short improvised section Figs. B.15 and B.16. They made large use of those techniques that allow to shape the note continuously, and also explored the various types of attack (percussive/staccato). They carefully placed the marcato notes at the beginning or end of a phrase, where it was easier for them to achieve them. They extensively used the wilder sounds of the growl and multiphonics (which they called *"harmonics"*). Regarding the improvised melodic passage: *"I allowed myself a moment of freedom, where I would improvise without worrying about the technique"*. Interestingly, they did not use any of the techniques in that passage.

P4 built their piece around controlling the dynamic of the note, and was challenged by the limitation of monophony, so they tried to recreate a sensation of harmony by holding a pedal note and playing some short notes in the high register during the first part of their piece (Fig. B.15). They used the expression pedal, as a sort of "switch", keeping it fully back at the beginning and pushing it all the way forward half way through the piece. They did not use multiphonics or high mode. They commented that they would have approached the composition task differently without the extended techniques, and would have focused more on the melody and phrasing.

P5 wanted to use as much of the tonal variations as possible for their piece, to make it interesting and varied despite being on a monophonic instrument: *"I was too distracted with the noise of the instrument"*. The piece in its final state is more of a collection of sketches: *"I did not structure it until the end, I just wrote down some things I wanted to do [...] some melodic part, I wanted to use the bending, [...] I also wanted to use some dynamic range"*, and the performance consists in a fluid elaboration of the fragments notated in the score (Fig. B.18). They elaborated a new technique, described in Section 5.7.5.2 to obtain a coloured "air noise", and used it to open the piece. In one of the later sections of the piece, they used an arpeggio pattern, inspired by a Bach étude, or a cello suite. They devised their own two-handed technique to play this, also described in 5.7.5.2. They used percussiveness extensively in another section, where they played a syncopated rhythm with marcato accents (staff 6 of the score, but the marcato is not notated). During their performance, they left space for challenges and errors, and let these guide them partially: *"Trying to control the change from air noise to tone was a challenge that I wanted to have, I like the fact that it might surprise me [...] Sometimes I am trying to get something, but it doesn't work, and something else happens instead, so I just go with it"*.

P6 started composing their piece aiming to combine melody, harmony and rhythm: *"I was interested in the idea of creating some kind of rhythm and being able to contain an ostinato and melody at the same time"*. The main section of the piece (Figs. B.20 and B.21) is built around a new technique they devised (described in 5.7.5.2), which produces *"glitchy drumkit effects"*. This was found to be better than what could normally be achieved on regular monophonic synthesizers, where *"you would not be able to control how loud the two parts would be"*. At the same time, it was different from what would be achievable on the piano: *"I like that you can get less of the note than you would on the piano"*. The introduction (Fig. B.19) was written last, in order to use some of the other effects, exploring the same melody with different techniques, including partial presses and aftertouch/growl.

5.7.5 *Use of techniques*

Overall, the gestures involving continuous key control were found easier to understand and execute than the percussiveness.

Participants found that holding a key in a partly-pressed state, while it is an easy concept to grasp, was not always easy to achieve. *"there is a lot of sensitivity in a very small space"*, reported P5. P3 found a parallel with a technique they had to use in a contemporary piano piece, which required an unusual amount of control on the onset, compared to most piano playing: *"I would press the key until I feel the resistance [close to the escapement point], and then play a pianissimo from there*. The composer had prescribed this technique because of the specific sound quality of the pianissimo achievable this way. Difficulties were reported when attacking partially pressed notes, because of the difficulty to avoid a long breathy onset transient, and get to the harmonic tone straight away: *It is difficult to get a pianissimo throughout without getting the breathy sound first, [...] you have to stop at a level, and it's difficult to get it consistently* (P4). When playing multiple partially-pressed notes in succession, the two problems are combined together. Participants would normally bring in the second note very slowly, and their tempo accuracy would suffer. Sometimes they faded out the previous note while bringing in the new one, thus generating a breathy sound during the note change.

One specific passage that was found particularly challenging to execute is the partial pitch bends between G and A in bars 5-6 of Fragment 2 (Fig. B.3). Here, the A key is only slightly depressed in order to obtain an A♭ pitch. Most players would find themselves using the 4^{th} finger on the G, thus having to perform the partial bend to the A key with their pinky. Besides the intrinsic difficulty gesture, the player also had to listen carefully for the pitch of the produced tone, which is something *"my piano ear is not trained for"* (P3). P4 commented that in order to make it as precise as possible they would use a *"very subtle change of arm weight"*.

Players would try to compensate for their uncertainty when performing partial and progressive key presses in several ways. P1, P2, P5, P6, would rest the palm of their hand on the frame of the keyboard, at the front of the keys, so to provide a stable anchoring point for their movements. *"I would never play piano that way"*, commented P1, *"I just find it easier this way for this instrument, so that the weight is not just hovering there"*. Several players (P2, P3, P6) also reported that they looked at the keys while playing more than they would normally do, trusting their eyes perhaps more than their ears when depressing the key partially. P2 commented

"I also learned that if you want to play at the same volume, you have to also use your sight, see how deep is the key and put the other one at the same position". We asked P2 and P3 to play again a short passage without looking at the keys and without resting their hand on the frame. The outcomes were excellent in both cases, and not very far from what they had previously achieved (see for instance P3's attempt in Fig. 5.29), suggesting that using visual cues and the keyboard's frame as a reference point could be avoided easily through further training and increased confidence. Another strategy frequently employed by the participants in order to exert more control on progressive key presses was to use more than one finger per key. P2 played the "lento" section of their piece with two fingers. P6 would place two or three fingers around the front edge of the black key when depressing it slowly, as if they were trying to pinch it. When playing a white key, they would sometimes use one finger from each hand. P4 played with multiple fingers *"just to have more control"*.

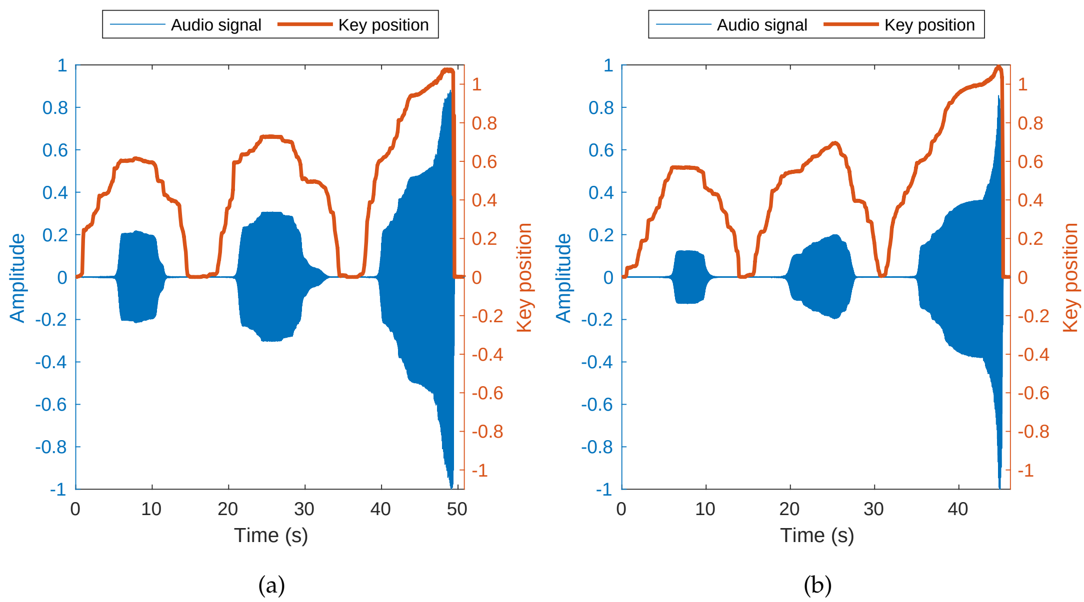


Figure 5.29: Performance of the first three bars of P3's composition, by looking at the keys and resting on the frame (a), and without looking at the keys or resting on the frame (b).

All players used two hands for most of their playing, especially from the end of the first session onwards. The second hand was often used to prepare in advance the following note in a slow passage involving partly pressed notes. P1 reported *"It is easier in terms of control, weight distribution, getting from one point to the next. I have hands ready in place, I prepare the amount of pressure ahead of finishing the previous* *note"*. Similarly, P4 *"I played two-handed a lot to have more control over the next note"*. P5 started using two hands only after the training *"I am going to get much more interesting options"*, and found it particularly useful for bendings: *"When bending, it is so much tiny intervals or movements on the hand that I don't have that. I tend to put my arm weight on things, but here it is really local muscle movements. It is particularly hard to hold one key partly down and pressing the one next to it"*.

All the participants were familiar with the concept of aftertouch, as present on some digital keyboards, however none of the players had used it regularly before the study. The pressure on the keybed required in order to obtain the aftertouch, and particularly the growl effect, on our instrument was judged to be excessive by all participants. A common approach to work around this difficulty was to use two or more fingers on the same key. P3 reportedly used more of the body weight and arm on these notes in order not to overload the wrist.

The only player to autonomously learn how to achieve the "high state" was P1, who discovered it while they were preparing the piece, and found a reliable technique to consistently achieve it: *"Depressing the [bending] key partially then switching fully, getting a high, distorted timbre"*. They then made it a central part of the composition, where they referred to its sound as *"bag pipe like"*. We described to P5 how to obtain the "high state", and they quickly managed to achieve that.

At the beginning of the second session, as the player was working towards their composition, we introduced the expression pedal. Only P4 used it for their composition, and they kept it back the whole time and pushed it forward at one point in the composition. P5 was initially excited by the idea of the pedal, but then found that it was not doing anything different from what they could already do. Similarly, P6 did not use the pedal because it was not useful enough: *"I listened to what it was doing and I could do more with the fingers"*. In general, the function of the pedal was not seen as a useful one, and it was more interesting, or challenging, to explore the sounds that were only accessible with the fingers.

5.7.5.1 *Percussiveness*

Percussiveness was by far the hardest technique to learn for most participants. There was general consensus that it was easier to obtain it on the black keys than on the white keys. The training exercises on this technique were found to be the most challenging by the players, especially number 24 (marcato on all notes), 26 (as the marcato note would fall on the 5^{th} finger), and 30 (marcato note at the end of a legato phrase), while placing the marcato on the first note of the phrase (29) was easier. Most players were more reliable with some fingers rather than others, with the 5^{th} finger, the pinky, often being the one with which struggled the most. In order to help those players who were struggling, we often resorted to explaining the details of how the key and the finger lose contact during a percussive touch, and this often helped them shape their gesture appropriately.

We performed a systematic analysis of the marcato notes played during the Performance Evaluation 2 task, which took place at the end of the second session, and is therefore a mirror of the final level of proficiency achieved by the participants. We annotated the recorded audio and sensor data from the participants rehearsing and performing études 1, 3, 5, 6 (those that include marcato notes). We identified the key presses that correspond to a marcato note on the score, and we labelled each of them as either true positive (TP), when the note was detected as percussive by our percussion detector, or false negative (FN), when the note was detected as non-percussive. During the task, performers typically repeated several times those notes or passages that were most challenging. In our analysis we included all the key presses that we could refer to a marcato note on the score, even when the performer was repeating only one note or a short phrase. The results are summarised in Table 5.4, where the hit rate ($\frac{TP}{TP+FN}$) is an indication of how reliably the participant can perform the gesture.

	Total	**TP**	**FN**	**Hit rate**
P1	203	124	79	0.61
P2	188	126	62	0.67
P3	64	22	42	0.34
P4	57	39	18	0.68
P5	145	124	21	0.86
P6	94	85	9	0.90

Table 5.4: Metrics of percussion accuracy per participant

P1 was the one who struggled the most. Their initial attempts did not trigger the percussiveness detector, and when they did manage to obtain a few percussive presses, they did so by using a very stiff wrist *"I think if I was using that sound on the piano it would produce a very harsh sound that I would personally avoid, which is probably why I am conditioned to approach the keyboard differently. It wouldn't be a good sound"*. Once we explained to them that the key needs to bounce off the finger in order for the percussiveness detector to work, they remarked that as part of their training they learned to always *"keep the weight on the key"*, and that consequently, it was unnatural for them to let the key bounce off. By the end of the study they could only partly reliably achieve it, and mostly only with the 2^{nd} or 3^{rd} finger. Additionally, they found that dedicating one of the two hands (typically the left) to percussive touches helped them achieve greater reliability.

P2 found that they could achieve a percussive tone by falling on the key from above, and they suggested it could be a good exercise for regular piano techniques: *"My teacher used to tell me that I have to 'fall' on the key, lifting my hand. But on the piano you don't hear the difference"*. The noticeable acoustic effect of this gesture on our instrument would, in their mind, make it easier to internalise and learn it. They were aware that the motion required to obtain a percussive effect is not directly linked to simple key velocity: *"It's not hard, it's not fast"*. By the second session, their percussive technique evolved into using a simpler, fine but decisive, motion of the finger. This allowed them to incorporate the gesture extensively in their composition, also using it in the middle of a legato phrase, and with remarkable accuracy. Their reliability playing the études was not very good, however, they used the technique with ease in their own composition: *"Because it's my piece, it's easier if it comes from me [...] Where the accents are in the études, they are not comfortable for the hand"*.

P3 initially approached the marcato with a full arm gesture, but it took them over 10 attempts before successfully getting a percussive sound in response to their action. Once they started understanding and recognizing the acoustic result of the gesture, it became easier for them to achieve it. During exercise 4 of the training, the first that encompasses the percussive gesture, they switched to using a flat finger and a wrist motion. Shortly thereafter they used a "whip" gesture with the wrist and the fingertip. They would use the same gesture on the piano for expressive purposes, even though it would not have a direct impact on the produced level: *"more at a psychological level, trying to visualize the sound that I want through my gesture, but I could achieve the same sound with a different gesture"*.

P4's described the arm gesture they used initially for percussiveness as *"the same I would use it for an accented tenuto with a dot"*, and later commented *"you have to use a lot of weight"*. They later settled for a different technique,*"It's something like a slap"*, which would give them fairly reliable results with fingers 2, 3 and 4. They would use this same technique on the piano for some very marcato playing, even though it would not make a difference in terms of sound, so that different techniques could be used in its place: *"On the piano I could use a range of techniques [with more* *or less arm weight] that would give the same result and which one I use would sometimes depend on my mood, strength, stamina, or fingering"*. Those other techniques, however, would not produce the desired effect on our instrument.

P5 encountered the percussive effect on their own during the initial discovery. They initially described their technique as *"hit it as fast and hard as I can. I play that a lot when I play Rhodes samples, where you get a distortion when you play it really hard"*, thus ascribing its causes only to the velocity of the note onset. Similarly to P2, they were overall more confident when using percussive tones in their improvisations, or compositions, than they were in the exercises and études. They would occasionally use a second hand to achieve percussive touches, especially for fast repetitions on a single key.

P6's initial approach was to *"try to give it more velocity, more arm weight"*. Towards the end of the first session they shifted to more of a wrist motion, trying to *"throw more weight at the key"*. Their technique proved fairly reliable throughout the study: they scored the highest hit rate in Table 5.4, and they even used it with their 5^{th} finger for the G and F♯ notes in étude 3, while most participants found the technique particularly hard with this finger. Despite mastering the technique, they only used one percussive note in their composition.

An interesting fortuitous discovery we made during the analysis of the interviews, while listening back to the recordings made through the DPA4060 microphone, is that the accessory acoustic noise produced by the finger-key and key-keybed impact was very noticeable. Our impression is that participants who were struggling the most seemed to generate a louder noise, as if they were trying to put more energy in the gesture than those who found a more reliable (and economic) technique. We are not, however, able to produce numerical evidence of the phenomenon, given that the microphone placement was not calibrated for this purpose.

5.7.5.2 *New techniques*

In their compositions, four participants leveraged in different ways the monophonic output of the keyboard controller, when pressing more than one key at a time. P3 would partially depress two keys (G3 and D4) and, with microscopic adjustments, alternatively make one or the other the deepest one (see the last bar of Fig. B.15). In this case, the keyboard controller's note number output will correspond to whichever key is the deepest, and it will therefore jump between the two notes, yielding a glitchy effect when alternating quickly, as P3 did. P4,

Figure 5.30: Transcription of one bar from P5's performance, with notation of the separate hands, as plated, and the resulting sound as heard.

throughout sections 1 and 2 of their composition (Fig. B.17), would hold one note partially or fully pressed, or even in aftertouch mode, while, with the other hand, fully depressing another note for a short period of time. P5 in the final part of their composition, held fully down a note with their left hand, while they played an arpeggio of semi-quavers with their right hand, as we transcribed in Fig. 5.30. During the semi-quaver rests of the right hand, the note held down by the left hand would then play. In this case, the behaviour of the keyboard controller was very similar to that of a monophonic synthesizer with highest or most-recent priority. P5 managed to carefully time the release of the right hand, so that the pedal tone would play on the tempo.

P6's multi-key technique was perhaps the most original one: they held a partially-pressed key in the high register while playing a staccato ostinato on fully-pressed keys in the low register. The high key would be initially pressed only slightly, so that it would not produce a periodic tone, and over time they would then change its depth, as shown in Fig. 5.31. Depending on the position of the high key, its result on the produced sound would vary between coloured noise (when lightly depressed), to pitched, decaying resonances (when depressed further), to fully sounding periodic tones (when fully depressed). They highlighted that this technique would not be achievable on a regular monophonic synthesizer, because there would be no easy way of obtaining different timbres for the two notes, the way it is possible here by controlling the depth of the held note. This technique is used throughout the "main" section of their composition (see Figs. B.20 and B.21).

P5 found that by pressing a key very fast, without percussion, and keeping the weight on it until the end, so that it is immediately pushed into the aftertouch region, they can get a note to produce no harmonic content but only some coloured noise, as shown in Fig. 5.32. If the key is slightly released, the note starts a harmonic oscillation at the expected pitch. The technique is not of simple execution, and during their performance P5 sometimes had to try multiple times in order

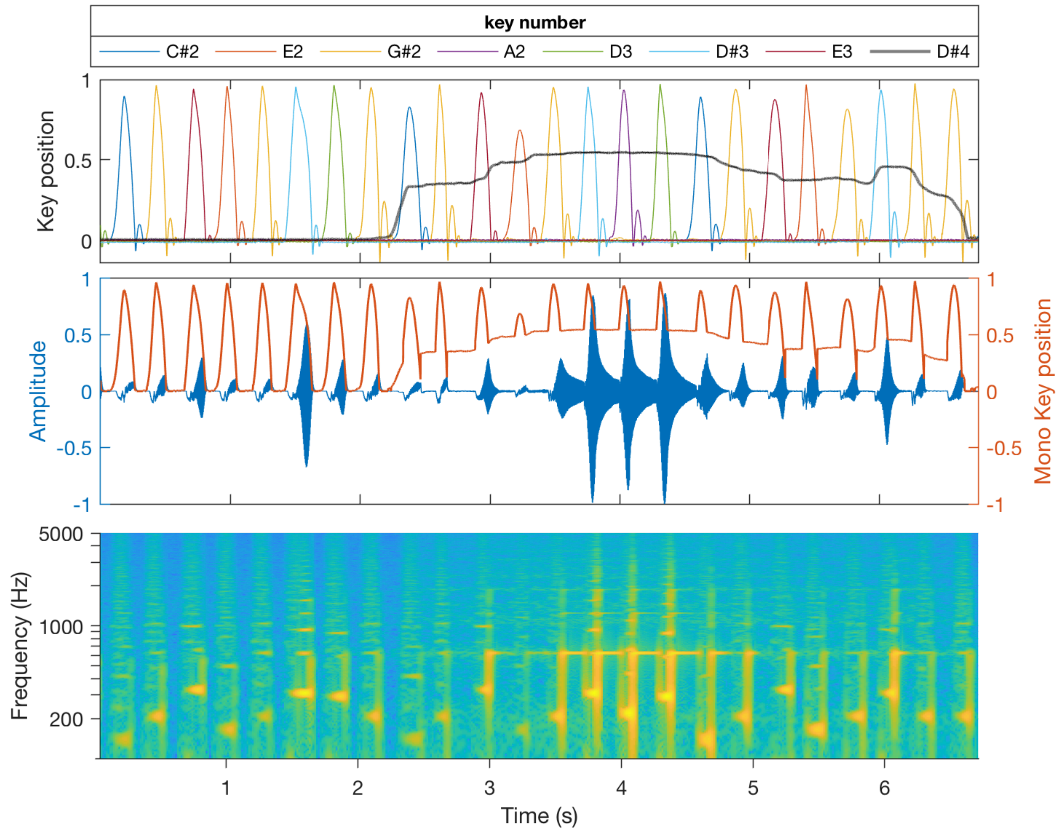


Figure 5.31: P6 holding one note partially in the high register (D#4) while playing an arpeggio in the low register. The top plot displays the position of each of the key involved in the gestures. The red line in the middle plot is the monophonic key position produced by the keyboard controller, which is what, ultimately, the sound engine uses to generate the audio (blue). This passage corresponds approximately to bars 2 and 3 of Fig. B.20.

to get it right. They notated it as "air noise" in the first bar of their composition (Fig. B.18). This technique is leveraging the effect of non-linearity in our physical model, causing what we earlier referred to as "dead notes", and for which we corrected in Section 5.4.6. Our corrections only applied to the high register, where the phenomenon was more obvious, and in fact P5 reported that they could not achieve the same result in the higher register. However, we did not apply any correction in the lower register, and that is where P5 used it (on a G3 note).

P6 developed a "fade-out" vibrato technique, as an extension to the regular "pressure vibrato ", first introduced in training exercise 12 (Fig. B.1). Their inspiration came from the last bar of étude 8 (Fig. B.5), where a pressure vibrato is followed by a fade-out. In our intent, the vibrato would stop when the note starts

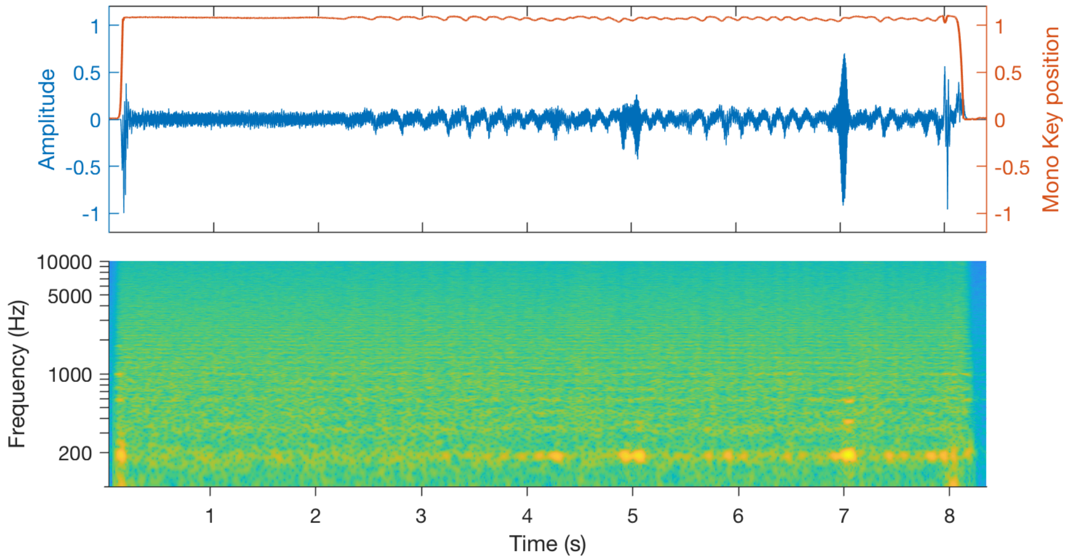


Figure 5.32: P5's "air noise" technique, obtained by producing a very fast attack, non percussive, which goes immediately in the aftertouch region.

fading out. P6's technique aimed at continuing the vibrato motion while progressively releasing the key, maintaining the oscillatory movement, but moving its oscillation point, as displayed in Fig. 5.33. They used this technique during the introduction of their piece (see Fig. B.19).

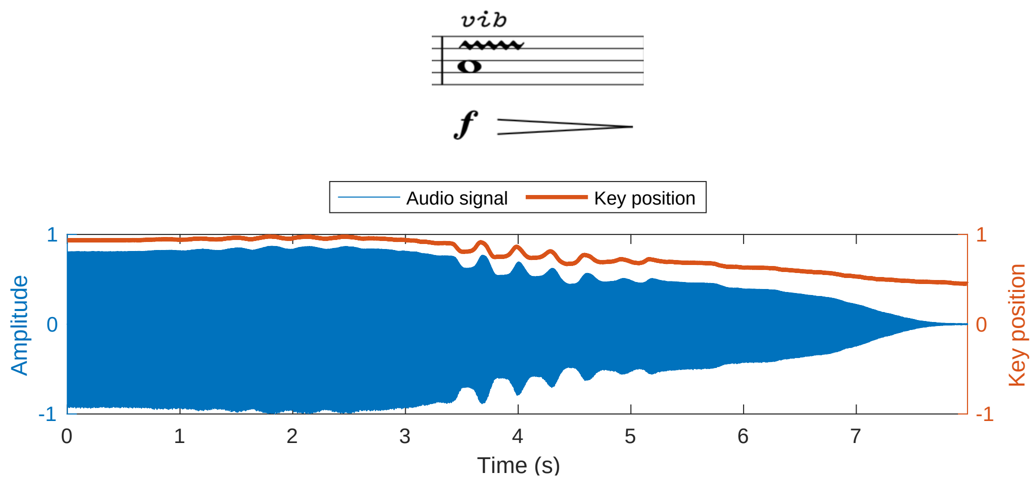


Figure 5.33: An example of fade-out vibrato by P6 (bottom), and notation (top).

5.7.5.3 *Parameter adjustment*

Parameter adjustment was performed after each player had played the instrument for more than 5 hours across the two sessions. Players could adjust parameters e_a

	e_t	e_a
P1	0.79	n.a.*
P2	0.67	6
P3	0.79	6
P4	0.70	6
P5	1.1	4.83
P6	1.1	7.5**

Table 5.5: Summary of the preferred setting after parameter adjustment.
*: We implemented aftertouch adjustment after P1 undertook the study.
**: P6 requested the aftertouch value to be extended beyond the originally available range.

and e_t, which control the mapping between key position and air pressure into the physical model, as described in Eq. (5.5). The available ranges were: $0.5 \leqslant e_t \leqslant 1.1$ and $2 \leqslant e_a \leqslant 6$. The settings the participants chose at the end of the adjustment task are reported in Table 5.5. P5 pointed out that the learning and practising process they had been going through was crucial in helping them understand the significance of the parameters being adjusted: *"if you had given me knobs straight away, I would have not known what to do with them at the beginning, would've not been able to tell the difference"*. However, many other players still struggled, at least initially, to understand what sonic effect was the result of tweaking the e_t parameter. After some experimentation, P1 said *"I don't know what I am changing ... is that the attack?"*. They eventually understood that when the e_t parameter was set to minimum, it gave a *"smoother attack"*. P3 could not find much difference at all, and eventually left the control around the default position. P4 also found it very subtle, but described the perceived effect on the interaction: *"It's easier to control the range after you hear the tone, so it's better if it speaks earlier"*. Other players had stronger opinions. P2 described the effect of the control in terms of the produced sound: *"With the control all the way up, if you press very fast, it takes longer to get the sound, it comes slowly. When it's all the way down it is too little responsive, more like a normal keyboard, but very easy to play fast without missing the sound"*. P5 and P6 both preferred the tunable parameter to be all the way up. P5, playing in this condition, found control of the dynamic to be much easier, perhaps because it reconnects the most to traditional keyboard playing: *"The portion where there is the tone now is nearer the bottom of the note, where I am more used to control the sound"*. P6 reported that they would occasionally slightly depress notes without intending to, which is a gesture normally occurring, and without effect, on the piano. Consequently, pushing the e_t setting all the way up was, to them, *"more forgiving for a piano player"*.

When playing the études again with the new setting, only P6 expressed the desire to change them again. During étude 5, they set the e_t control back to the middle in order to better play the crescendo on the C♭ in bars 3-4: *"For the steady crescendo it is actually easier"*. When playing étude 8, which consists almost entirely of long half partially pressed notes, they commented it was *"very hard"*. We therefore suggested that they bring the control back up to 127: *"It's easier to do, it favours the dynamics as well"*, although *"The volume change is more sensitive"*. The sensor and audio data from the two performances are displayed in Fig. 5.34. During the performance with the highest e_t setting (Fig. 5.34b), the player holds the key deeper in the key throw, while producing a slightly quieter tone.

As for the aftertouch control e_a, most players would quickly settle for having it all the way up, in accordance with what was reported throughout the study that, with the default setting, obtaining the growl sound with the aftertouch required to

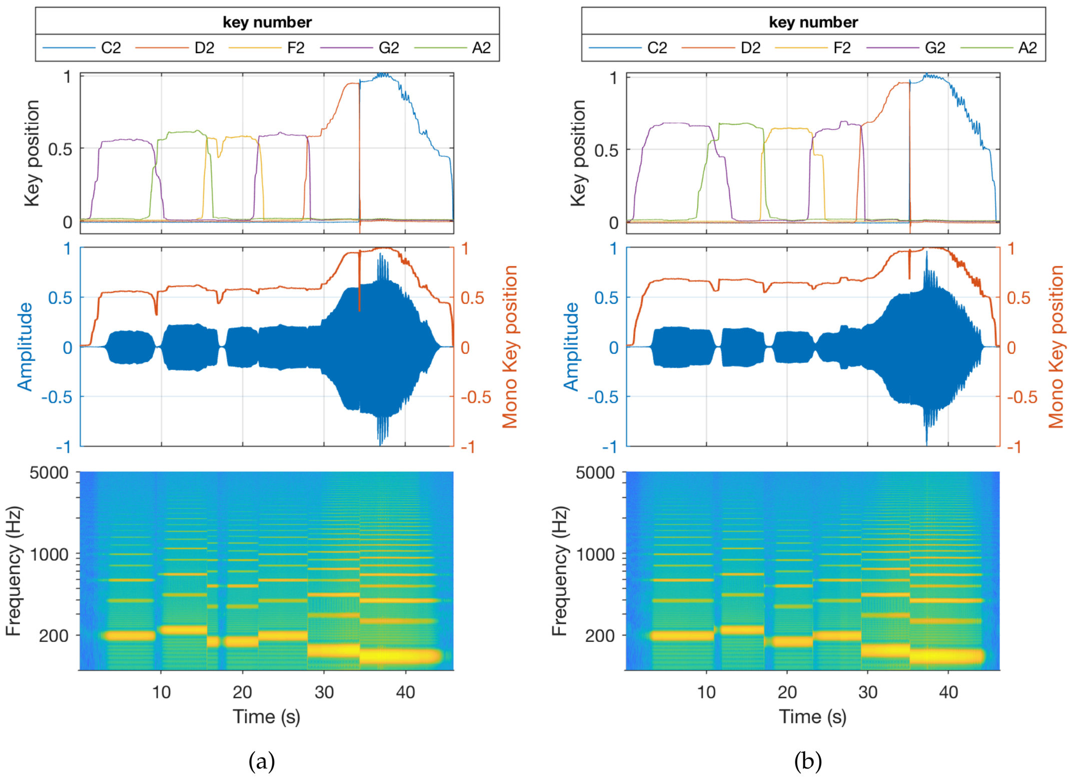


Figure 5.34: Performance of étude 8 by P6 with two different e_t settings: $e_t = 0.7964$ (a) and $e_t = 1.1$ (b).

exert an excessive amount of force on the key. P5 was not annoyed by the physical effort required: *"[the sound I can achieve with maximum pressure] is a horrible sound, so it **must** be hard to achieve"*. P6 was still unsatisfied with the pressure required on the key even when the setting was all the way up, so they asked us to modify the code to increase the range even more. They were satisfied with the extended range, and, besides being able to achieve wilder extremes, they could still easily control the early part of the aftertouch region, reliably achieving growl and non-growl gestures at their will.

5.7.6 *Comparison with traditional keyboard instruments*

The interface of our instrument is a traditional keyboard, however the continuous control it allows sets it apart from most other keyboard instruments. The extended techniques required to play the instrument made it not immediately accessible by pianists. *"When you are doing something that you are used to producing some sonic result but it produces a different sonic result it's very confusing"* (P1). There was a generalized feeling that the instrument had to be considered as something "else" in order to be fully understood: *"it's another instrument, it's not a normal keyboard"* (P2). P5 explained how the physical appearance and feel of the instrument was at odds with its functionalities, and how they were initially put off by that: *"Because I am a piano player, a keyboard player, when I am faced with this, I expect all the things that would work on a keyboard to work some way [...] maybe if it was a pink keyboard, or all the notes of a different color, it would just immediately make you go 'this is ... different!'. Or if they weren't plastic keys, they were rubber [finish] or wooden, you wouldn't immediately go 'I know what this [thing] does. Oh no wait it doesn't! What's going on?', you get that thing happening [currently]"*.

Several comments contrasted the gestural vocabulary with that of the piano: *"The whole body-and-arm gestural language from the piano would not be very useful here"* (P1), *"Some of these techniques seem on the surface quite straightforward, [but] it goes against the control that you are used on the piano. It's quite counter-intuitive to that, because your arm weights and your attack are used here for a completely different outcome. I am having a fight against my instincts, which is good and interesting"* (P4). P5 pointed out that the lack of mechanical support from the key made continuous gestures more complicated: *"Where is the focal point in the weight of my arm to hold that note and control it? It's not like playing piano at all"*. They contrasted this with their experience as brass player, where the mouthpiece provides the required support, and even if similarly very small movements affect the sound, *"they all happen in the same place, same mouthpiece. Here instead you have multiple physical locations [the keys], which is very difficult if you haven't done it before"*.

The process of learning the instrument often involved relearning techniques previously acquired on the piano, focusing the attention on previously ignored aspects of the gesture. As mentioned above, the technique P1 used for percussiveness was, to them, very unnatural, and went against the skills they acquired over thousands of hours of training: *"There is a bit of frustrations because some of the skills/gestures that are customary on the piano (and hard-wired in the performer's mind) do not work here, and viceversa"*. They suggested in several occasions that inexperienced players would find it easier to learn this and other techniques, as they would have no embodied pre-conceptions. P2 pointed out the higher conscious cognitive load required by the instrument *"I am now more conscious of what I do with my fingers on every single key. It has so many shades, you have to be very accurate playing it"*. Several players (P1, P3, P4, P6) reported the need to pay attention not only to the active sound-generating gestures, but also to some accessory involuntary gestures of the inactive fingers: *"You can also inadvertently trigger sounds just by resting your fingers on it"*, said P4, *"so there is need to re-jig my attention"*. Similarly, it would often be the case that the attack of a note would not be "clean" because of slightly depressing a neighbouring key. This gesture, which would normally not produce any sound on a piano, resulted in an unwanted pitch bending or a transient volume dip on our instrument, requiring participants to pay more attention to the cleanliness of their technique.

Some participants reported that a high-level mind shift helped them to come to terms with the instrument. *"The more I think it is not a normal keyboard, the more it makes sense. It is not so transferrable for someone who is fluent on the [traditional] keyboard to play this with expression"* (P5). P4 reported that by the end of the study their overall approach changed, as needed by a new instrument: *"I have a better idea of how to manipulate each key. I see each key as a separate entity, and not necessarily even looking horizontally [...] With this you also work in the vertical dimension and now I have a better grasp of that"*. P6 similarly had to re-think their mental image of the key: *"that aspect of [the key] being much less of a lever, but thinking of it throughout the key throw"*.

There was a general consensus among participants that the skills acquired and the time spent learning this instrument would bring improvements to their regular piano and keyboard playing. *"I am now more conscious of what I do with my* *fingers on every single key. It has so many shades, you have to be very accurate playing it"* (P2). This was true also for P1, despite the unnatural gestures they had to adopt during our study: *"I don't think this would disrupt my piano technique. If anything, it would help me improve my technique by improving my control and listening"*. Overall, the additional cleanliness and attention to unwanted movements required by the instrument were seen as improving the overall technique and control. P2 pointed out repeatedly that a piano somehow augmented with the sound would help understanding and paying more attention to some piano techniques, such as "falling" (through marcato) and legato (through pitch bending), improving the learning process for a beginner.

5.7.7 *General feedback*

Many participants shared the opinion that controlling the dynamic of the notes was not straightforward. While fade in and fade out could be achieved with good accuracy, attacking and sustaining a note at set level other than the "forte" dynamic corresponding to key bottom, was challenging: *"It's hard to play [individual] notes or a phrase with the same volume"* (P2). P1 pointed out how the most desirable key response setting may depend on the technique and passage in use *"the difficulty for the staccato is that it still takes time to speak in the initial attack. It would be good to make a sharper shorter initial attack selectively somehow. [...] Maybe it would be good to have the key response [e_t] adjustable, like a knob more to the left for shorter notes and more to the right for longer ones. For staccato is easier one way, for gradual presses is easier the other way"*. Similarly, as previously reported, P6 adjusted the e_t control while going through the études. P2 and P5 mentioned that they would want to be able to adjust in real-time the range of "dynamic" control available, to be able to cover a wider repertoire. Physical modifications to the instrument were also suggested in order to improve dynamic control. P3 and P5 suggested to adopt taller keys, thus increasing the tolerance for the very accurate movements that are currently required. P1 suggested that some sort of dynamic haptic feedback could be implemented in the key to facilitate maintaining a given intermediate position.

The lack of polyphony was reported as a limitation of the instrument by all players, for varied reasons. P1 was particularly interested in the possibilities of combining multiple pitch bend and multiphonic gestures to achieve *"microtonal sonorities and shimmering clashes"*. P5 would like the freedom to play two or three notes at a time occasionally, to reinforce a passage *"It wouldn't be all the time ... like* *playing lead guitar with occasional double stops"*. P3 and P4 found it limiting for not being able to establish a harmony as they are used to on the piano. When asked, in response, whether they would be able to control each finger on a polyphonic version of the instrument with the same accuracy they do now, most acknowledged it would be hard, but it was not impossible to learn. P6 suggested that fine individual control may be, in the polyphonic case, less important than a global sense of modulation and variation *"You would not be able to finely control each finger, but being able to add sensitivity even in a 5-voice chord could really add to the liveliness"* By the end of the study, some players had come to accept the monophony as a part of the character of the instrument: *"Polyphony is less of a problem now, as I have started accepting it as it is"* (P6).

Some participants (P1, P5, P6) noticed the change in pitch depending on the pressure (see Section 5.2.2.4), but it was not an issue for them, and they considered it to be part of the sound, and contribute to the liveliness of the instrument. P1 referred to the *"fragility"* of the instrument, remarking how its *"instability"* gives it *"life"*. In their view, the multiphonics and the high state sounds were an *"unexpected by-product"* of the instrument. P2 also perceived a difference in the loudness across the instrument range, where the low register sounds louder than the high register, and they wondered whether that is the case also in real wind instruments. P5 praised the natural feel of the instrument, contrasting it with other electronic instruments: *"I want to make it feel like the instrument speaks, and if an instrument sings, then it has to be able to scream. Most electronic instruments have an upper limit to the dynamic, unlike acoustic ones, where you can push it until it breaks. I like those things that are in this, that are a weird effect, like the multiphonics and the growl"*.

5.8 DISCUSSION

5.8.1 *Skills generalization*

Despite the fundamentally uncommon capabilities we built into it, our instrument presents itself on the surface as a remarkably "normal" keyboard. Most of our participants played it for several minutes before discovering that the key position controls the dynamic of the sound. They would start playing using their normal technique and expectations, and the instrument responded in an expected way, that is: it emitted a sound of the appropriate pitch. In a matter of seconds, they realized that the instrument could only play one note at a time and the velocity of the press would not affect the loudness of the resulting sound. In their eyes and ears, the initial experience must have not been very different from their previous experience playing regular monophonic synthesizers. Even when P5 autonomously discovered the "percussive" effect, their understanding of the technique was still heavily grounded in the common notion of key velocity (*"fast and hard"*), and so was their first attempt at describing the effect of key position. Simply observing these initial responses gives us a clear indication that pre-existing techniques can easily be used on our instrument, which in turns denotes the presence of a strong expertise *transfer* [Krakauer et al., 2006].

When the effects of continuous key position were discovered, autonomously or introduced by the investigator, players did not struggle to *understand* them. Gestures such as slowly depressing the key to fade the note in or out, or holding the key partially pressed to achieve a dynamic change, are fairly intuitive. *Performing* them accurately, however, comes with several difficulties, as the training and motor skills required to control the micro movements of the key for a sustained period of time substantially differ from those needed for obtaining discrete events, as it is common on regular keyboards. In other words, for these techniques piano skills do not necessarily *generalize* to the instrument. The aftertouch, growl and the vibrato gestures are easier to perform, because the key is resting against the felt at the bottom of the key, which offers a mechanical resistance, acting as a reference point for the performer's finger. P6's attempt at extending the vibrato technique to the key-throw (Fig. 5.33) indirectly confirms this, by showing that the amplitude of the oscillatory motion becomes more irregular as the key is lifted from the felt.

Partial pitch bending gestures also challenged the players. The notation we chose (see Fragment 2 in Fig. B.3), marks the resulting pitch instead of the key to be used to obtain it, and this subverts the notion that each key corresponds to a given location on the musical staff. For example, in the case of bar 5, we are asking a player to hold down the G, partially press the A (notated as A♭) to obtain an A♭ pitch, and ensure (using their ear) that the pitch they obtain is correct: that is a lot of asking, and most players on their first attempt pressed the wrong key. At the same time, the concept of one pitch per key is also subverted: the A key is used to produce an A♭ pitch. The minor impediment players found in locating the correct keys upon first reading the score and the struggle to get the pitch right can both be seen as a case of *interference* [Krakauer et al., 2006]. We are expecting keyboardists to regulate aspects of their playing, the intonation, which is not part of any conventional keyboard technique. We can hypothesize that a synth player accustomed with using the pitch bend wheel, or a guitar player, would have found the notation more intuitive, and would have been more accurate in the resulting pitch.

As for the percussiveness, many players found that this gesture was hard to perform. We have seen indications that understanding the gesture required to obtain it was a challenge in itself. Some players tended to think about it in terms of a "high velocity" gesture, or otherwise requiring a large amount of energy, whereas all that was needed was controlling the initial impact of the finger on the key. Explaining the expected mechanical behaviour of the finger-key system during a percussive touch, the way our algorithm expects it, seemed to help some performers understand it. We observed P1, P2, P3 and P4 evolving their percussive technique in the course of the study, each of them settling, in the end, for their own very personal approach. Several of the techniques they used were drawn from their piano experience, often accompanied by the remark that on the piano they would not really make a difference in the produced sound. P5 and P6, could mostly reliably achieve the technique in the études, and P2 was very reliable when playing their own musical material. P1, P3 and P4, on the other hand, struggled considerably, and by the end of the study they had not achieved even a passable reliability. From P1's comments it is clear that their training as a pianist was an obstacle in acquiring a percussive technique. At the very beginning, it took them several attempts, with several different techniques, in order to be able to find one that would produce the desired effect. Being used to *"keep the weight on the key"*, they would not easily let the key bounce off the finger, as required by our instrument. The technique they settled for in the end was not particularly reliable on our instrument, although it was better than the others they previously attempted, and it was a technique they would never use on the piano. P3's experience was not very different, as they went from technique to technique without ever achieving a reliable strike. Interestingly, P5 and P6 were those that had the least to say about their percussive technique, whose technique did not evolve much during the study, and also those that picked it up the fastest. One possible explanation for this could be that P5 and P6 were just lucky, and it just happened that their intuitive idea of "percussive" was the right one for our instrument. However, we could also hypothesize that these two players, who mostly play popular music, and perhaps have a less formalised technique, and fewer pre-conceptions about what is a "good" or a "bad" technique, felt more free to discover and adopt the technique that works. For P1 and P3, there is a clear indication of an interference of their piano training on the ability to play the percussive gesture on our instrument, while this was not the case for P5 and P6.

We know from the literature that on the piano, given the relatively low-bandwidth of the keyboard interface, a player is free to choose from a very large number of different gestures in order to obtain a desired sonic outcome [MacRitchie, 2015]. The choice of the gesture could depend on training, personal preference, musical context and musical momentum, but ultimately, is largely irrelevant for the sonic outcome. P4's comment reported above, about choosing a technique depending on "mood, strength, stamina, or fingering" is particularly revealing in this sense. When we start assigning special meanings to some of these gestures, as we did with the percussiveness, the degrees of freedom of the player decrease, and they have to find and learn what the "right" gesture is. Playing these new gestures is then difficult at two levels. In terms of execution, there is the intrinsic difficulty of learning a new gesture. However, at a higher conceptual level, there may be an even more fundamental problem: the player loses the freedom of choice in the moment, no matter the *stamina* or the *mood*, they will have to perform *that* technique. This could possibly have a bigger impact in the long term than learning and adopting a new gesture, as it requires a new, much stricter, performance discipline.

We observed accessory movements of the idle fingers, such as resting them on the keys, slightly depressing them, or slightly pressing a neighbouring key while attacking a note on a white key. Some similar observations were reported by McPherson et al. [2013]. These movements result in no audible effect on the piano and most keyboards, but they do affect the sound of our instrument, for instance causing unwanted notes to speak, or unwanted pitch bends. Players quickly learned to control their idle fingers and used a cleaner technique in order to overcome these minor interferences. In relation to this, P1, P2, P3 and P4 mentioned that the skills learned on our instrument can be enriching for their piano playing in terms of cleaner technique and increased awareness and control. This can be seen again as a sort of transfer, but this time taking the skills from our instrument back to the traditional keyboard. P2 also mentioned pedagogical applications, whereby the learning process for a specific technique could be improved by providing an unambiguous auditory feedback. Their observations resonate with earlier suggestions of using advanced gesture tracking for pedagogical applications [Hadjakos et al., 2009, McPherson and Kim, 2011].

5.8.2 *Appropriation*

The literature presents several examples of musical instruments whose limited affordances stimulate players to explore the constraints and develop new techniques to push the boundaries beyond the original intentions of the instrument designer. An example of this can be found in Gurevich et al. [2010], where a rich set of gestures, interactions and playing styles emerges from players engaging with a simple one-button instrument. [Magnusson, 2010] suggests that affordances in musical instruments tend to be more obvious (e.g.: a key is to be pressed) than constraints, and that exploring the latter tends to be a large part of the discovery process of an instrument. Zappi and McPherson [2014] suggest that constraints stimulate the exploration of the capabilities of an instrument, and ultimately lead to *appropriation*, that is *"adapt and adopt the technology around them in ways the designers never envisaged"* [Dix, 2007]. All of our participants initially lamented the lack of polyphony in our instrument as a limitation, possibly also because they rarely played unaccompanied monophonic keyboards in their previous experience. However, four of them elaborated, in the course of the study, their own original techniques precisely to overcome this limitation and be able to establish a sort of harmonic structure in their piece with multi-key gestures (Section 5.7.5.2). Interestingly, these gestures are not rooted in piano technique, and the sound they produce does not even have a counterpart in flute playing, they are entirely new techniques, developed specifically around our instrument. We can say that our players reacted to one of the constraints that the designers put on the affordances of the instrument by appropriating it.

We did not inform our participants about some of the affordances we built into the instrument, namely the "multiphonics" and the "high state". P5's "air noise" technique exploited an unintentional "error" that we made while programming the sound generator, thus revealing to us, the creators, an unknown affordance. We expect that, from the perspective of the player, all of these must have looked as unexpected behaviours, "glitches" in the instrument. Yet, each of these made their way into some of the pieces that were composed during the study, thus making the instrument's imperfections a signature characteristic of the instrument sound. This is another case of appropriation, and it adds to a long standing practice of leveraging non-ideal behaviours and technological failures for creative purposes, so much that they become part of the identity of the instrument, and of the repertoire, even when they were not part of the designer's original idea [Cascone, 2000, McPherson and Kim, 2012]. Emblematic in this sense is the case of the Hammond organ, where the creator even put active effort into removing the key-click, despite it being one of the most appreciated features by musicians [Vail, 2002].

5.8.3 *Additional remarks*

Our instrument was designed as a probe for studying the generalization of keyboard playing skills to changes in the mapping of the keyboard interface. We observed a significant transfer of skills, especially in the horizontal navigation of the pitch space, with a subject-dependent interference, at times strong, on a particular gesture (percussiveness). The continuous gestures, on the other hand, require a technique change where the piano's gestural language, involving upper body and arms weight, has to be adapted to a technique that is based on fine hand and finger movements. A similar adaptation is required from pianists moving to harpsichord [Martinez, 1990, p.37]), or organ [Steyl, 2018]. Continuous gestures did not suffer from interference, but also showed minimal transfer. In other words, they have to be learned. To what extent they can be learned, however, remains an open question. We can argue that the *"ceiling on virtuosity"*[Wessel and Wright, 2002] of our instrument is very high, in that it allows more complex performances than a regular synthesizer, and the *"entry fee"* is low, at least for players already familiar with keyboard instruments. However, some of the features of our instrument, those that really sets it apart from more traditional keyboards, may still be subject to an excessively slow learning curve. An indication of this risk comes from the fact that several of our participants have highlighted the difficulty to perform some of the continuous gestures, and that currently it is not easy to obtain notes of even loudness, attack quiet notes fast, and, more in general, master fine-grained control on the key position. As a possible workaround to this, some expressed the desire to have a global performance control to adjust the key response, or the overall dynamic level of the instrument. *"Good musical instruments must strike the right balance between challenge, frustration and boredom"*, writes Jordà [2004, p. 331]. A longitudinal study would be the most effective way of understanding how practical it is to learn and become proficient at these techniques, or whether the instrument is actually too complex and will eventually *"alienate the user before [its] richness can be extracted"* [ibid.].

Our flute synthesizer has a much more obvious mapping between continuous key position and the sound than the hybrid Hammond from Chapter 4 did. Yet, most of our participants initially played it for several minutes without realising that the keyboard responded to the continuous key position, indicating that the pre-conception of the keyboard as a controller of discrete events plays an important role in driving the exploration and understanding of a new instrument [Pigrem and McPherson, 2018]. The player's assumptions are so strong, that a substantial amount of evidence is required in order for these to be questioned, so much that the player may unconsciously ignore an auditory feedback that contradicts them. This can also help explaining the fact that most players of the hybrid Hammond did not realise that there were changes in the control on the onset of the note: they simply did not expect it. One of the two players that took part in both studies (P1 here, P6 there) was already familiar with the appearance of the keyboard scanner, and its applications, from playing the Magnetic Resonator Piano. For this reason, they started their free exploration on the flute synthesizer immediately exploring the effect of continuous gestures. However, when they played the Hammond in the earlier study, they did not explore its continuous behaviour, despite the fact that the keyboard scanner was also visible, which we would have expected to similarly stimulate their curiosity, and the fact that they perceived some difference in the attack of the note, which we would have expected to lead them to explore more in detail the onset of the key. When we asked them why they did not explore the Hammond in the same way, they answered *"I was very convinced by the Hammond and that was overriding everything else"*. In other words, after playing the Hammond as a conventional keyboard for over one hour, by the time they got the free exploration, their expectations were that the instrument was "normal", and therefore did not explore it further.

5.8.4 *Future work*

We briefly suggest here some possible improvements to the instrument to make it easier to play. Adding haptic feedback by means of physical modifications to the keybed is one possible pathway. Haptic keyboards have existed in the past [Oboe and De Poli, 2002, Cadoz et al., 1990, Gillespie, 1992], however the new challenge here would be to dynamically control the feedback depending on user action, for instance to help maintaining a stable position of the key half way through the key throw. Solutions that would not involve physical modifications are also interesting and worth exploring. The idea of two dynamic controls interacting together, one local (key position), and one global, which was suggested by some of our participants, was already present on the Ondioline, where a knee lever acted as a global volume control [Fourier et al., 1994]. An easy modification to our instrument to achieve this would make the role of the expression pedal more central, giving control over a wider dynamic range than it currently has. Another idea worth investigating is that the pedal would instead dynamically control the transfer function between key position and pressure of the physical model Section 5.4.1. One third option would be to make the transfer function between key position and pressure dynamic, setting dynamic operating points based on the recent history of the key position.

Another area of potential improvement is that of percussion detection. So far we have looked at the hit rates in Table 5.4 to get an idea of how reliable each player was in performing the percussive gesture, and the outcome is that only two out of six participants seem to have reliably learned how to play this gesture. This is a reasonable way of considering these results, and the question we have been asking so far is whether the performer can find the right gesture that will make the instrument play the expected way: can the performer learn how to play? Since P6 became very quickly proficient with a technique that satisfies the current requirements of our detector, it is legitimate to expect that anyone else could learn the same technique. Looking at it from a different perspective, however, the hit rate values in Table 5.4 can be interpreted as an indication of how good our percussiveness detector is, and the outcome is that it is not very good at detecting "percussive gestures" the way P1, P2, P3, P4 mean them. On the other hand, it is doing a good job at detecting "percussive gestures" the way P5 and P6 intend them. There is therefore room for improving the percussion detection algorithm in order to customise its response to the style of each individual player. This can be achieved by adding some tunable parameters or using machine learning.

The approach of fine tuning the parameter space of the instrument treating the sound generator as a black box, as described in Section 5.2.2 has given good results, with the only exception of the "loophole" found by P5 to produce the "air noise". We had identified a similar effect and corrected for it, however this other manifestation of the problem went unobserved until P5 found it. This outlines an issue with our black box approach: if, when we first observed "dead notes", we had investigated more carefully what the underlying issue in the physical model, that is answer the question "why does it break?", as opposed to simply "how do we fix it?", then we would have probably been able to prevent this issue from arising during the study. For our purposes, we are actually happy with the fact that this "defect" has been exploited for creative purposes, and we are therefore unlikely to remove it in a revised iteration of the instrument, however it is worth pointing out that, while our black box approach has the advantage of not requiring a deep understanding of physical modelling, it may reserve some more or less welcome surprises.

5.9 CONCLUSION

In this chapter we designed a keyboard-based musical instrument that can handle extended techniques in order to control the physical model of a wind instrument. The visual appearance and mechanical characteristic of the keyboard have not been modified, but the mapping between the keys and the sound generator has been subverted by adopting a paradigm where the instantaneous position of the key continuously controls the sound generator, as opposed to a more traditional approach based on discrete key presses. Multi-key gestures and percussive hits were also assigned new sonic meanings.

In a performance study, six keyboard players played our instrument and we found that the basic skills of their keyboard technique transferred to our instrument, because of the familiarity of its interface. The presence of continuous key control required a re-adjustment of their overall technique, but was ultimately intuitive to understand. The level of proficiency they achieved during the two sessions of the study is limited, and they highlighted some issues with specific techniques. A longitudinal study would be needed to assess the long-term controllability of the expressive potential of the instrument. The percussiveness gesture was remarkably hard to achieve for some of the players, because of interference of their pre-existing technique, and due the narrow set of gestures that would get successfully detected by the percussiveness detector. While improvements can be made to the percussiveness detection algorithm to be more generic or better adapt to individual styles, the long-term effects of the fundamental restrictions it imposes on the freedom of a player to adopt a gesture are worth investigating further.

6 CONCLUSION

This thesis has presented an in-depth investigation in the use of touch on keyboard instruments, analysing its effects on the Hammond organ and developing bespoke technology to allow real-time gesture sensing and sound synthesis in the digital domain, in order to explore the possibilities of new, richer mappings between gesture and sound. The contributions of this work are summarised in Section 6.1. In Section 6.2 we reflect on the outcomes and the implications of our work. Section 6.3 contains possible pathways for future research.

6.1 CONTRIBUTIONS

Effect of touch on the Hammond organ

We performed an analysis of the effect of touch on the produced sound. First, we found evidence that the bouncing characteristics of the key contacts are dependent on the velocity and percussiveness characteristics of the touch used to press the key. Through a listening test, we assessed that differences in the touch can reliably be detected by listeners in an isolated context. As a consequence of this, we conclude that a velocity-based controller is insufficient to encode a gesture on the Hammond, or to control a Hammond emulator.

Platform for continuous keyboard sensing and embedded sound generation

We built on two existing technologies, a keyboard scanner and a Bela board, in order to create a system that can generate sound in response to continuous key position. We described how we solved the associated engineering challenges to achieve reliable low-latency performance. The end result is a platform that is ideal for research applications, providing high-resolution sensing, deterministic behaviour, data logging and low-latency, which we used extensively in the rest of our work.

Hammond organ emulator with position-based control

We developed a dynamic model of contact bounce on the Hammond and incorporated it into an open-source Hammond emulation software, running on the platform we developed earlier. We then embedded our platform into an existing Hammond organ, replacing part of the analog sound generator with the digital emulation. The result is a hybrid Hammond organ emulator that looks and feels like a real Hammond, but whose sound generation is highly customisable, so that the mapping between key position and the virtual contacts of the sound generator can be altered at will to produce a number of different settings corresponding, for instance, to those of real Hammonds or digital emulations.

Performance study on the effects of variations in mapping discrete events to key position

We conducted a performance study on our Hammond emulator using several different mappings between key position and the sound generator, so that each mapping may produce slightly different acoustic results depending on the shape of the key onset. Our results indicate that players struggle to detect differences between the mappings when playing the organ along with a backing track. However, when playing on their own, freely exploring the instrument, most players would develop a strong preference for one of the settings.

Physical modelling synthesizer controlled by extended keyboard techniques

We designed a monophonic physical modelling synthesizer that uses continuous key position as its control input, again using the embedded platform we developed. The instrument allows several extended keyboard techniques based on continuous position, including single and multi-key gestures, and uses an algorithm for detecting percussive gestures on key onset. Each extended technique produces a clearly identifiable acoustic result in the produced sound. The instrument challenges some of the assumptions of traditional keyboard instruments, namely the discreteness of key presses, the independence of keys and the meaning of the percussive gesture.

Performance study on extended keyboard techniques

A performance study on our physical modelling synthesizer allowed us to investigate how keyboard playing skills generalise when the interface is defamiliarised by challenging some of its basic assumptions, and whether continuous key position control on an otherwise mechanically traditional keyboard can be used for expressive control purposes. We found a strong *transfer* of pre-existing technique, with some *interference* of previous training and gestural vocabulary specifically with regard to the percussive gesture. There is evidence of *appropriation* in the new techniques that participants developed during the study. Their use in the context of the pieces that were composed is an indication of the expressive potential of the instrument, and of continuous keyboard sensing at large.

6.2 reflections

Our hybrid Hammond instrument allows some gestures to produce an audible difference that would not otherwise do so on a velocity-based keyboard. In terms of Jack et al. [2017]'s concept of *bottleneck*, we made the bottleneck of our instrument slightly wider. The behaviour of the instrument is somewhere between discrete and continuous: a key press no longer amounts to a single instantaneous measurement, but at the same time the continuous behaviour unfolds over a timescale that is too short for conscious control of all the detail. As a result, the key press gesture itself may still be a discrete action, from the performer's point of view, but one which has several degrees of freedom that can translate to different profiles of key motion over a scale of tens of milliseconds. In this case, the wider bottleneck does not necessarily require the player to adjust their playing style, or even to be aware of it, in that the macroscopic behaviour of the instrument remains largely unchanged. However, it brings in a certain sonic richness associated with different key trajectories, which could lead performers to skilfully manipulate the character of the onset, perhaps even at a subconscious level, once they become completely familiarized with the response of instrument. Where Jack et al. observed a progressive impoverishment of the gestures used in an instrument with a narrow bottleneck, we expect a progressive enrichment as a consequence of a wider one.

The implications of a behaviour that is still discrete, but is sensed continuously, can extend beyond a Hammond emulation. In the case of digital pianos, for instance, the detection of a percussive onset by analysing the key profile, or the velocity of the key impact with the keybed, could be used to add finger noise or extra resonance to the synthesized sound [Goebl et al., 2014]. Another possible application would be that of varying the perceived mechanical response of the keyboard by manipulating its triggering point, as we have seen indication of such an effect in Chapter 4. On the spring-loaded Hammond keybed used in the study, the force of the key on the finger increases linearly with key position, while in the case of a weighted keyboard, the key response is instead discontinuous due to the presence of the escapement point [Oboe, 2006]. Whether and to what extent the manipulation of the triggering point can still elicit a mechanical sensation when presented in combination with the strong mechanical discontinuity embedded in a weighted keyboard would have to be verified in a further study.

Keyboard technique heavily exploits redundancies, in that several different gestures can lead to the same produced sound [MacRitchie, 2015]. Gestures are normally chosen on the basis of the context and musical intention, and they act more as an assistance to the player in delivering the overall musical phrase, but with little or no obvious effect on the sound of individual notes. By assigning stronger meanings to distinct gestures, which were previously producing the same sound, the player has to produce a more specific gesture to achieve the associated sound. This may prove challenging, first of all at a conceptual level: the player has to acknowledge that different gestures are no longer equivalent. They may then encounter issues at a biomechanical level, where the gesture required is one that they would not normally use, or that they have to learn anew. The biomechanical issue is, in our opinion, less of a concern, in that the right gesture can be learned, given enough practice, or the instrument can be reprogrammed to accommodate the player's technique, however the conceptual mind shift *needs* to happen.

The path to making a successful augmented keyboard goes, in our view, through an optimization problem of maximising skills *transfer*, by maintaining a familiar interface, or maximising playability, by subverting the interface. Instruments like the Seaboard [Lamb and Robertson, 2011] or the Continuum [Haken et al., 1998], whose design only preserves the spatial location of the notes of the traditional keyboard, are capable of very expressive performance, but their re-interpretation of the interaction surface heavily reduces the skills that a keyboard player can immediately re-use on the instrument. In order to maximise skills transfer, we approached the design of our extended techniques physical modelling synthesizer using an unmodified weighted keyboard, and constrained the interaction to the keys themselves. When the *locus of augmentation* is the same as the *locus of inter-*

action, the player can more easily leverage their existing sensorimotor skills [Morreale et al., 2019].[1] As a result, a player does not receive any sort of haptic support to their playing, which makes those continuous gestures that involve partial key presses rather hard to accomplish accurately on our instrument. We expect that, when first encountering the instrument, an experienced keyboardist would find these gestures easier on the Haken or Seaboard thanks the haptic response they provide, however more complicated passages involving full presses would be easier to perform on our instrument thanks to the familiarity of the interface. What the long-term potential, beyond the initial encounter, of either approach is remains to be established.

One of the most important advantages of MIDI is that of *generality*: as long as a sound generator can understand note and velocity information, then it can be played with a keyboard (or other MIDI controllers). In this sense, Rubine and McAvinney [1990], define MIDI as the *"crossbar switch between gesture and timbre"*. Jordà [2004] points out that when *"splitting the chain"* between controller and generator, *"It becomes hard – or even impossible – to design highly sophisticated control interfaces without a profound prior knowledge of how the sound or music generators will work"*. Our continuous key flute synthesizer requires the player to move the fingers on the keyboard in some new and unusual ways, due to the specific characteristics of the instrument as a whole. Therefore, even if our instrument could, in principle, be split into a *controller* and a *sound generator*, trying to replace that sound generator with a different one is a much more arduous task than it would be if the interface between the two was as simple and clearly defined as it is the case for MIDI-based instruments. In other words, using the key data from a performance on this instrument to control a different sound generator would most likely not yield a meaningful result. This is true even if, on our same instrument, we were to change the internal mapping of key position to pressure (see for instance Fig. 5.34). By making the bottleneck wider, much wider in this case than a velocity-based keyboard would allow, we have gained in the amount of control available, but we lost in the generality of the controller.

One of Perry Cook's principles for designing musical controllers in Cook [2001] is *"Copying an instrument is dumb, leveraging expert technique is smart"*. Following this idea, we successfully managed to *transfer* pre-existing skills to our instrument.

1 The expression pedal was not initially intended as a part of the instrument, and was added simply to support the player's compositional process, but we limited its control range and introduced it late in the study, in order to keep the player's focus mostly on the challenging new capabilities of the keyboard.

We may expect that those new skills that had to be learned specifically for our instrument will have very little application in other instruments leveraging continuous key position, in the case where these implement a mapping that assigns different meanings to the same gestures, or introduce new gestures. On the other hand, we could expect that some more general concepts, such as considering the importance of the key throughout the key throw, and understanding and internalising the reduction in the degrees of freedom of redundant gestures, could, indeed, transfer across instruments with continuous key sensing. We therefore speculate that, as a consequence of the loss of generality of the controller, skills generalization will have to happen at a higher level, and it becomes crucial for the player to be guided by the auditory feedback when adapting to a new instrument.

6.2.1 *Implications of continuous key sensing*

We summarise our findings here in relation to our initial research questions:

- **What are the implications of continuous key sensing in the context of keyboard-based digital musical instruments?**
- **How does the subtle manipulation of key-sound mapping affect the performer's perspective on a keyboard DMI?**
- **How can we leverage continuous key sensing to increase gestural richness and control capabilities in keyboard DMIs?**

We explored two possible uses of continuous key sensing, to provide a richer articulation on the short time scale of a note's attack, and as an explicit means of modulation. Treating each onset as a series of closely spaced discrete events, as in the case of the hybrid Hammond, gives the player more control on the sound without changing the overall behaviour of the instrument. The player may not explicitly be aware of the subtle control they have, and its effects are perceived at several levels, not only in terms of sound, but also in terms of the mechanical response of the action, or the accuracy and control they can attain with the gesture.

Skills transfer is a complex multifactorial issue that first manifests itself at a conceptual level (*understanding* the instrument) and only later at a biomechanical level (*playing* the instrument). After years of training and playing, a player's technique has come to specific expectations with regard to the capabilities and response of a keyboard, which are taken for granted, so that when they are no longer valid the player may not become immediately aware of it.

Mappings such as those we used on the hybrid Hammond instrument allow richer gestures to produce richer sounds without requiring an explicit understanding and a positive action from the player. By increasing the amount of control on the onset, we allow for the development of a richer gestural language over time, through an either conscious or unconscious process, in a way that a velocity-based keyboard could not offer.

The continuous mappings of our flute instrument come with greater challenges in terms of conceptual understanding, as some of the expected behaviours of the keyboard are completely subverted. One major hurdle for the player is to understand and accept that different gestures that would yield the same sound on other instruments are now associated with different sonic effects, so that the specific gesture used for a given note has a new level of importance. Most of traditional keyboard techniques still apply, ensuring that the performer can play straight away without having to relearn the fundamentals. However, the execution of some of the new extended techniques requires to learn to control the key in a new way, which may need a longer period of training in order to be mastered. Yet, we have indication, from some of the compositions made with the instrument and some of the new techniques that emerged, that continuous key sensing makes our instrument a much richer monophonic synthesizer than it could be in a MIDI world.

6.3 future work

In the performance study we conducted with the Hammond emulator, players seemed to have stronger opinions when freely exploring the instrument, without the constraint of playing along with the backing track. However, the time they spent exploring the instrument was very short. A follow-up study would investigate the longer-term implications of the different mappings, also observing how the technique evolves adapting to the subtle changes in key response.

The usefulness of continuous key sensing are not limited to new, richer control spaces. Some of the outcomes from our study suggest that changes to the triggering point of the key can impact the perceived weight of the keyboard action. One possible avenue of exploration would investigate the potential of adjustable triggering point in keyboard actions, and its impact on the perceived quality of the instrument. In this case, continuous sensing would produce a scalar velocity value as its final output, but the behaviours for *how* and *where* in the key throw this value is generated could be programmed in software in a very flexible way.

More in general, what we observed is a small example of how the perception of musical gestures can be influenced by multimodal interactions, and this is, on its own, a promising field for investigation.

Our study on extended techniques left the question open of what extent of proficiency can be achieved on the continuous techniques with ongoing training, and can only be answered by a longitudinal study in the wild. For this to be possible, the instrument's hardware needs to undergo a revision process to make it more portable and reliable, while making the calibration process more straightforward, so that it can live with player for extended periods of time, in the comfort of their studio, or on stage.

Technical improvements to the instrument, while not strictly needed for a longitudinal study, are surely possible, or even desirable. A "smarter" (i.e.: more flexible) percussiveness detector could improve the skills transfer of traditional keyboard technique on to our instrument; dynamic key mappings and haptic feedback could help making continuous gestures easier; polyphony would open up a new world of possibilities.

Instrument designers and music software makers have acquired their expertise over the years in developing sound generators that are suitable for discrete MIDI input. Designing sounds for a continuous keyboard, however, is an entirely new challenge, and it requires a shift in mindset to adapt to the new paradigm, and it could become its own field of exploration, investigating general strategies for sound models that take advantage of gestural interaction.

6.4 LOOKING FORWARD

In the few months intervening before the writing of this thesis and its final revision, two major announcements have been made which have the potential of changing the world of continuous control on digital keyboards: MIDI 2.0[2] and Expressive E's Osmose[3]. MIDI 2.0 supports higher resolution data and per-key controllers. It has therefore the potential, through device profiles, to support continuous key sensing. For the most part, however, it keeps an approach that is event-based. For instance, the NoteOn message has been expanded so that it can contain descriptors of the quality of the attack of the note. In the case of a software emulation of a violin , these could, for instance, be used to indicate one of several gestures, such as

2 https://www.midi.org/midi2
3 https://www.expressivee.com/discover-osmose

"staccato", "con legno", "martellato". In musical practice, however, achieving an attack matching one of these descriptors is the result of a continuous process that takes place over a relatively extended period of time (hundreds of milliseconds), and which can be performed in an infinite number of nuanced ways. Once again, a multi-faceted phenomenon with several possible temporal evolutions is discretised down to one scalar number by MIDI.

Osmose is a touch-sensitive keyboard with continuous control and gesture recognition, which was announced in November 2019, but will not get in the hands of players until mid-2020. The online demo videos show that besides continuous sensing of the key position, it can recognise the type of touch (pressed vs struck), and aftertouch. [4] They are using the sound engine developed by Haken for their Continuum. We are happy to see that their design pays attention to many of the aspects of continuous control that we have investigated and advocated for in this thesis, and we are really looking forward to playing it.

6.5 CLOSING REMARKS

The initial motivations of this work came from a very personal interest of mine, the passion for electromechanical keyboards, and the curiosity of understanding where the "magic" resides within them. While that question remains largely unanswered, my study of the Hammond has indeed shone light on where a small part of this magic may be, in the subtle sensitivity that exists right there in the key. The two instruments I built are my own modest attempt at bringing something magical into 21st century instruments.

I hope to have contributed to reviving the interest in continuous key sensing for performance analysis and performance augmentation, showing examples of its possible applications and of its expressive potential, and providing an accessible platform for the use of future researchers.

4 e.g.: https://www.youtube.com/watch?v=PkobxItFaNI

BIBLIOGRAPHY

Anders Askenfelt and Erik V Jansson. From touch to string vibrations. i: Timing in the grand piano action. *The Journal of the Acoustical Society of America*, 88(1): 52–63, 1990.

Anders Askenfelt and Erik V Jansson. From touch to string vibrations. ii: The motion of the key and hammer. *The Journal of the Acoustical Society of America*, 90 (5):2383–2393, 1991.

Anders Askenfelt and Erik V Jansson. On vibration sensation and finger touch in stringed instrument playing. *Music Perception: An Interdisciplinary Journal*, 9(3): 311–349, 1992.

Anders Askenfelt and Erik V Jansson. From touch to string vibrations. iii: String motion and spectra. *The Journal of the Acoustical Society of America*, 93(4):2181–2196, 1993.

John Baily. Ethnomusicology, intermusability, and performance practice. In *The New (Ethno)musicologies*, Europea: Ethnomusicologies and Modernities, pages 117–134. Scarecrow Press, 2008.

P. Barkan. A study of the contact bounce phenomenon. *IEEE Transactions on Power Apparatus and Systems*, PAS-86(2):231–240, Feb 1967. ISSN 0018-9510. doi: 10.1109/TPAS.1967.291840.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, Articles*, 67(1):1–48, 2015. ISSN 1548-7660. doi: 10.18637/jss.v067.i01. URL `https://www.jstatsoft.org/v067/i01`.

Juan Pablo Bello, Laurent Daudet, Samer Abdallah, Chris Duxbury, Mike Davies, and Mark B Sandler. A tutorial on onset detection in music signals. *IEEE Transactions on speech and audio processing*, 13(5):1035–1047, 2005.

Michel Bernays and Caroline Traube. Piano touch analysis: A matlab toolbox for extracting performance descriptors from high-resolution keyboard and pedalling data. *Proc. JIM*, 2012.

Michel Bernays and Caroline Traube. Expressive production of piano timbre: touch and playing techniques for timbre control in piano performance. In *Proceedings of the 10th Sound and Music Computing Conference*, Stockholm, Sweden, 2013. KTH Royal Institute of Technology.

N. Bernstein. *On coordination and regulation of movements*. Oxford: Pergamon Press, 1967.

Stephen Birkett. The slippery path from piano key to string. *The Journal of the Acoustical Society of America*, 136(4):2132–2132, 2014.

Jon Boley and Michael Lester. Statistical analysis of ABX results using signal detection theory. In *Audio Engineering Society Convention 127*, 2009.

Gianpaolo Borin, Giovanni De Poli, and Augusto Sarti. Algorithms and structures for synthesis using physical models. *Computer Music Journal*, 16(4):30–42, 1992.

Donald F Buchla. Velocity and aftertouch sensitive keyboard, December 17 1985. US Patent 4,558,623.

Claude Cadoz, Leszek Lisowski, and Jean-Loup Florens. A modular feedback keyboard design. *Computer music journal*, 14(2):47–51, 1990.

Chris Cannam, Emmanouil Benetos, Matthew EP Davies, Simon Dixon, Christian Landone, Mark Levy, Matthias Mauch, Katy Noland, and Dan Stowell. Mirex 2018: VAMP plugins from the centre for digital music. *Proceedings of the Music Information Retrieval Evaluation eXchange (MIREX)*, 2018. URL https://www.music-ir.org/mirex/abstracts/2018/CC1.pdf.

Kim Cascone. The aesthetics of failure:“post-digital” tendencies in contemporary computer music. *Computer Music Journal*, 24(4):12–18, 2000.

Nicolas Castagne and Claude Cadoz. 10 criteria for evaluating physical modelling schemes for music creation. In *In Proc. of the 9th Int. Conference on Digital Audio Effects*. Citeseer, 2003.

Juliette Chabassier and Marc Duruflé. Energy based simulation of a timoshenko beam in non-forced rotation. influence of the piano hammer shank flexibility on the sound. *Journal of Sound and Vibration*, 333(26):7198–7215, 2014.

Alcedo Coenen. Computer-controlled player pianos. *Computer Music Journal*, pages 104–111, 1992.

Perry R Cook. A meta-wind-instrument physical model, and a meta-controller for real time performance control. In *Proceedings of the International Computer Music Conference*, 1992.

Perry R. Cook. Principles for designing computer music controllers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 3–6, Seattle, WA, 2001. URL http://www.nime.org/proceedings/2001/nime2001_003.pdf.

Alan Dix. Designing for appropriation. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 2*, pages 27–30. BCS Learning & Development Ltd., 2007.

Christopher Dobrian and Daniel Koppelman. The e in nime: Musical expression with new computer interfaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 277–282, Paris, France, 2006. URL http://www.nime.org/proceedings/2006/nime2006_277.pdf.

John Eaton and Robert Moog. Multiple-touch-sensitive keyboard. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 258–259, Vancouver, BC, Canada, 2005. URL http://www.nime.org/proceedings/2005/nime2005_258.pdf.

Steven A. Finney. Auditory feedback and musical keyboard performance. *Music Perception: An Interdisciplinary Journal*, 15(2):153–174, 1997. ISSN 07307829, 15338312. URL http://www.jstor.org/stable/40285747.

N H Fletcher. The nonlinear physics of musical instruments. *Reports on Progress in Physics*, 62(5):723–764, Jan 1999. doi: 10.1088/0034-4885/62/5/202. URL https://doi.org/10.1088%2F0034-4885%2F62%2F5%2F202.

Federico Fontana, Stefano Papetti, Hanna Järveläinen, and Federico Avanzini. Detection of keyboard vibrations and effects on perceived piano quality. *The Journal of the Acoustical Society of America*, 142(5):2953–2967, 2017.

Laurent Fourier, Curtis Roads, and Jean-Jacques Perrey. Jean-jacques perrey and the ondioline. *Computer Music Journal*, 18(4):19–25, 1994.

Adrian Freed and Rimas Avizienis. A new music keyboard with continuous key-position sensing and high-speed communication. In *Proceedings of the International Computer Music Conference*, 2000.

Shinichi Furuya and Eckart Altenmüller. Flexibility of movement organization in piano performance. *Frontiers in human neuroscience*, 7:173, 2013.

Alexander Galembo and Anders Askenfelt. Quality assessment of musical instruments-effects of multimodality. In *5th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM5), Hannover, Germany*, 2003.

Philippe Gerum. Xenomai-implementing a rtos emulation framework on gnu/linux. *White Paper, Xenomai*, page 81, 2004.

Brent Gillespie. Touch back keyboard. In *Proceedings of the International Computer Music Conference*, pages 447–447. International Computer Music Association, 1992.

Bruno Gingras, Pierre-Yves Asselin, Meghan Goodchild, and Stephen McAdams. Communicating voice emphasis in harpsichord performance. In *ESCOM 2009 : 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, 2009.

Werner Goebl. Melody lead in piano performance: Expressive device or artifact? *The Journal of the Acoustical Society of America*, 110(1):563–572, 2001.

Werner Goebl and Roberto Bresin. Measurement and reproduction accuracy of computer-controlled grand pianos. *The Journal of the Acoustical Society of America*, 114(4):2273–2283, 2003.

Werner Goebl and Caroline Palmer. Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, 186(3):471–479, 2008.

Werner Goebl, Roberto Bresin, and Alexander Galembo. The piano action as the performer's interface: Timing properties, dynamic behaviour, and the performer's possibilities. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC'03), August 6–9*, volume 1, pages 159–162, 2003.

Werner Goebl, Roberto Bresin, and Alexander Galembo. Once again: The perception of piano touch and tone. can touch audibly change piano sound independently of intensity. In *Proceedings of the International Symposium on Musical Acoustics*, pages 332–335, 2004.

Werner Goebl, Roberto Bresin, and Alexander Galembo. Touch and temporal behavior of grand piano actions. *The Journal of the Acoustical Society of America*, 118 (2):1154–1165, 2005.

Werner Goebl, Simon Dixon, Giovanni De Poli, Anders Friberg, Roberto Bresin, and Gerhard Widmer. Sense in expressive music performance: Data acquisition, computational studies, and models. *Sound to sense-sense to sound: A state of the art in sound and music computing*, pages 195–242, 2008.

Werner Goebl, Roberto Bresin, and Ichiro Fujinaga. Perception of touch quality in piano tones. *The Journal of the Acoustical Society of America*, 136(5):2839–2850, 2014.

Niccolò Granieri and James Dooley. Reach: a keyboard-based gesture recognition system for live piano sound modulation. In Marcelo Queiroz and Anna Xambó Sedó, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 375–376, Porto Alegre, Brazil, June 2019. UFRGS. URL http://www.nime.org/proceedings/2019/nime2019_072.pdf.

Tobias Grosshauser and Gerhard Tröster. Finger position and pressure sensing techniques for string and keyboard instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 479–484, Daejeon, Republic of Korea, May 2013. Graduate School of Culture Technology, KAIST. URL http://nime.org/proceedings/2013/nime2013_286.pdf.

Michael Gurevich, Paul Stapleton, and Adnan Marquez-Borbon. Style and constraint in electronic musical instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 106–111, Sydney, Australia, 2010. URL http://www.nime.org/proceedings/2010/nime2010_106.pdf.

Aristotelis Hadjakos, Erwin Aitenbichler, and Max Mühlhäuser. Probabilistic model of pianists' arm touch movements. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 7–12, Pittsburgh, PA, United States, 2009. URL http://www.nime.org/proceedings/2009/nime2009_007.pdf.

Lippold Haken, Ed Tellman, and Patrick Wolfe. An indiscrete music keyboard. *Computer Music Journal*, 22(1):30–48, 1998.

Laurens Hammond. Electrical musical instrument, April 24 1934. US Patent 1,956,350.

Hammond Instrument Company. Service manual HOOO-OOO495. Technical report, Hammond Instrument Company, Nov 1941.

David C Harding, Kenneth D Brandt, Ben M Hillberry, et al. Minimization of finger joint forces and tendon tensions in pianists. *Med Probl Perform Art*, 4(3): 103–108, 1989.

Eiji Hayashi, Masami Yamane, and Hajime Mori. Behavior of piano-action in a grand piano. i. analysis of the motion of the hammer prior to string contact. *The Journal of the Acoustical Society of America*, 105(6):3534–3544, 1999.

Andrew David Hunt. *Radical user-interfaces for real-time musical control.* PhD thesis, University of York, 1999. URL http://www-users.york.ac.uk/~adh2/Andy_Hunt_Thesis.html.

IMA. Midi musical instrument digital interface specification 1.0. Technical report, International MIDI Association, 1983.

Robert H Jack, Tony Stockman, and Andrew McPherson. Rich gesture, reduced control: the influence of constrained mappings on performance technique. In *Proceedings of the 4th International Conference on Movement Computing*, page 15. ACM, 2017.

Robert H Jack, Adib Mehrabi, Tony Stockman, and Andrew McPherson. Action-sound latency and the perceived quality of digital musical instruments: Comparing professional percussionists and amateur musicians. *Music Perception: An Interdisciplinary Journal*, 36(1):109–128, 2018.

Nicholas Jillings, Brecht De Man, David Moffat, and Joshua D. Reiss. Web audio evaluation tool: A framework for subjective assessment of audio. In *2nd Web Audio Conference*, April 2016.

Randall Evan Jones. Intimate control for physical modeling synthesisls. Master's thesis, University of Victoria, Department of Computer Science, 2008.

Sergi Jordà. Instruments and players: Some thoughts on digital lutherie. *Journal of New Music Research*, 33(3):321–341, 2004.

Xiong Jun, He Jun-jia, and Zang Chun-yan. A dynamic model of electromagnetic relay including contact bounce. In *2008 International Conference on Electrical Machines and Systems*, pages 4144–4149. IEEE, 2008.

Topi Kaaresoja, Eve Hoggan, and Emilia Anttila. Playing with tactile feedback latency in touchscreen interaction: Two approaches. In *IFIP Conference on Human-Computer Interaction*, pages 554–571. Springer, 2011.

Maria Kallionpää, Chris Greenhalgh, Adrian Hazzard, David M. Weigl, Kevin R. Page, and Steve Benford. Composing and realising a game-like performance for disklavier and electronics. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 464–469, Copenhagen, Denmark, 2017. Aalborg University Copenhagen. URL http://www.nime.org/proceedings/2017/nime2017_paper0088.pdf.

Matti Karjalainen, Unto K Laine, Timo I Laakso, and Vesa Välimäki. Transmission-line modeling and real-time synthesis of string and wind instruments. In *Proceedings of the International Computer Music Conference*, pages 293–293, 1991.

Kiyoshi Kawamura and Shigeru Muramatsu. Keyboard musical instrument hammer movement detection apparatus, March 16 1993. US Patent 5,194,685.

Hiroshi Kinoshita, Shinichi Furuya, Tomoko Aoki, and Eckart Altenmüller. Loudness control in pianists as exemplified in keystroke force measurements on different touches. *The Journal of the Acoustical Society of America*, 121(5):2959–2969, 2007.

Ralph Kirkpatrick. On playing the clavichord. *Early Music*, 9(3):293–306, 1981.

John W Krakauer, Pietro Mazzoni, Ali Ghazizadeh, Roshni Ravindran, and Reza Shadmehr. Generalization of motor learning depends on the history of prior action. *PLOS Biology*, 4(10):1–11, 09 2006. doi: 10.1371/journal.pbio.0040316. URL https://doi.org/10.1371/journal.pbio.0040316.

Alexandra Kuznetsova, Per B Brockhoff, and Rune Haubo Bojesen Christensen. lmertest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 2017.

Chi-Hsia Lai, Matti Niinimäki, Koray Tahiroglu, Johan Kildal, and Teemu Ahmaniemi. Perceived physicality in audio-enhanced force input. In *Proceedings of the 13th international conference on multimodal interfaces*, pages 287–294. ACM, 2011.

Roland Lamb and Andrew Robertson. Seaboard : a new piano keyboard-related interface combining discrete and continuous control. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 503–506, Oslo, Norway, 2011. URL http://www.nime.org/proceedings/2011/nime2011_503.pdf.

Jean Laurendeau. *Maurice Martenot, luthier de l'électronique*. Louise Courteau, 1990.

Hugh Le Caine. Touch-sensitive organ based on an electrostatic coupling device. *The Journal of the Acoustical Society of America*, 27(4):781–786, 1955.

Jennifer MacRitchie. The art and science behind piano touch: A review connecting multi-disciplinary literature. *Musicae Scientiae*, 19(2):171–190, 2015.

Jennifer MacRitchie and Nicholas J. Bailey. Efficient tracking of pianists' finger movements. *Journal of New Music Research*, 42(1):79–95, 2013. doi: 10.1080/09298215.2012.762529. URL https://doi.org/10.1080/09298215.2012.762529.

Jennifer MacRitchie and Andrew P McPherson. Integrating optical finger motion tracking with surface touch events. *Frontiers in psychology*, 6:702, 2015.

Jennifer MacRitchie and Giulia Nuti. Using historical accounts of harpsichord touch to empirically investigate the production and perception of dynamics on the 1788 taskin. *Frontiers in psychology*, 6, 2015.

Christian Madera. From the churches to the jails: The majestic profanity of the Hammond Organ. *Discoveries*, 10:112–126, 2011.

Thor Magnusson. Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, 34(4):62–73, 2010.

Maria Dolores Martinez. *Psycho-physical aspects of the interchange between piano and harpsichord: a study of experiences and opinions of current performers and educators*. PhD thesis, Texas Tech University, 1990.

Michael E McIntyre, Robert T Schumacher, and James Woodhouse. On the oscillations of musical instruments. *The Journal of the Acoustical Society of America*, 74 (5):1325–1345, 1983.

Andrew McPherson. The magnetic resonator piano: Electronic augmentation of an acoustic grand piano. *Journal of New Music Research*, 39(3):189–202, 2010.

Andrew McPherson. Touchkeys: Capacitive multi-touch sensing on a physical keyboard. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012. University of Michigan.

Andrew McPherson. Portable measurement and mapping of continuous piano gesture. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 152–157, Daejeon, Republic of Korea, May 2013.

Andrew McPherson. Buttons, handles, and keys: Advances in continuous-control keyboard instruments. *Computer Music Journal*, 39:28–46, 2015.

Andrew McPherson and Youngmoo Kim. Augmenting the acoustic piano with electromagnetic string actuation and continuous key position sensing. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 217–222, Sydney, Australia, 2010. URL http://www.nime.org/proceedings/2010/nime2010_217.pdf.

Andrew McPherson and Victor Zappi. An environment for submillisecond-latency audio and sensor processing on beaglebone black. In *Audio Engineering Society Convention 138*, 2015.

Andrew McPherson, Robert Jack, and Giulio Moro. Action-sound latency: Are our tools fast enough? In *Proceedings of the International Conference on New Interfaces for Musical Expression*, volume 16 of *2220-4806*, pages 20–25, Brisbane, Australia, 2016. Queensland Conservatorium Griffith University. ISBN 978-1-925455-13-7. URL http://www.nime.org/proceedings/2016/nime2016_paper0005.pdf.

Andrew P. McPherson and Youngmoo E. Kim. Multidimensional gesture sensing at the piano keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2789–2798, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9.

Andrew P. McPherson and Youngmoo E. Kim. The problem of the second performer: Building a community around an augmented piano. *Computer Music Journal*, 36(4):10–27, 2012.

Andrew P. McPherson and Youngmoo E. Kim. Piano technique as a case study in expressive gestural interaction. In *Music and Human-Computer Interaction*, Springer Series on Cultural Computing, pages 123–138. Springer London, 2013. ISBN 978-1-4471-2989-9.

Andrew P McPherson, Adrian Gierakowski, and Adam M Stark. The space between the notes: adding expressive pitch control to the piano keyboard. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2195–2204. ACM, 2013.

Carolina Medeiros and Marcelo Wanderley. A comprehensive review of sensors and instrumentation methods in devices for musical expression. *Sensors*, 14(8): 13556–13591, 2014.

Carolina Brum Medeiros and Marcelo M Wanderley. Multiple-model linear kalman filter framework for unpredictable signals. *IEEE Sensors Journal*, 14(4): 979–991, 2013.

Cheryl D Metcalf, Thomas A Irvine, Jennifer L Sims, Yu L Wang, Alvin WY Su, and David O Norris. Complex hand dexterity: a review of biomechanical methods for measuring musical performance. *Frontiers in psychology*, 5:414, 2014.

Romain Michon and Julius O Smith. Faust-stk: a set of linear and nonlinear physical models for the faust programming language. In *Proceedings of the 14th International Conference on Digital Audio Effects*, Paris, 2011. IRCAM.

Eduardo Reck Miranda and Marcelo M Wanderley. *New digital musical instruments: control and interaction beyond the keyboard*, volume 21. AR Editions, Inc., 2006.

R Moog. A multiply touch-sensitive clavier for computer music. In *Proceedings of the 1982 International Computer Music Conference. San Francisco: International Computer Music Association*, pages 155–159, 1982.

Robert A Moog. Position and force sensors and their application to keyboards and related control devices. In *Audio Engineering Society Conference: 5th International Conference: Music and Digital Technology*. Audio Engineering Society, 1987.

Robert A Moog and Thomas L Rhea. Evolution of the keyboard interface: The bösendorfer 290 se recording piano and the moog multiply-touch-sensitive keyboards. *Computer Music Journal*, pages 52–60, 1990.

F Richard Moore. The dysfunctions of MIDI. *Computer music journal*, 12(1):19–28, 1988.

Giulio Moro and Andrew P. McPherson. A platform for low-latency continuous keyboard sensing and sound generation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Royal Birmingham Conservatoire, 2020.

Giulio Moro, Astrid Bin, Robert H Jack, Christian Heinrichs, and Andrew P McPherson. Making high-performance embedded instruments with Bela and Pure Data. In *Proceedings of the international conference on live interfaces*. University of Sussex, 2016.

Giulio Moro, Andrew P McPherson, and Mark B Sandler. Dynamic temporal behaviour of the keyboard action on the hammond organ and its perceptual significance. *The Journal of the Acoustical Society of America*, 142(5):2808–2822, 2017.

Fabio Morreale, Andrea Guidi, and Andrew P. McPherson. Magpick: an augmented guitar pick for nuanced control. In Marcelo Queiroz and Anna Xambó Sedó, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 65–70, Porto Alegre, Brazil, June 2019. UFRGS. URL http://www.nime.org/proceedings/2019/nime2019_013.pdf.

Will Mowat. Bob moog piano bar. *Sound On Sound*, Mar 2005. URL https://www.soundonsound.com/reviews/bob-moog-piano-bar.

Tiffany Kwan Ng. *The Heritage of the Future: Historical Keyboards, Technology, and Modernism*. PhD thesis, University of California, Berkeley, 2015.

Roberto Oboe. A multi-instrument, force-feedback keyboard. *Computer Music Journal*, 30(3):38–52, 2006.

Roberto Oboe and Giovanni De Poli. Multi-instrument virtual keyboard: the mikey project. In *Proceedings of the International Conference on New interfaces for musical expression*, pages 1–6. National University of Singapore, 2002.

Otto Ortmann. *The Physical Basis of Piano Touch and Tone*. Kegan Paul, Trenc, Trubner & Co., 1925.

Joseph A Paradiso. Electronic music: new ways to play. *IEEE spectrum*, 34(12): 18–30, 1997.

Laurel Pardue and Andrew McPherson. Near-field optical reflective sensing for bow tracking. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 363–368, 2013.

Jussi Pekonen, Tapani Pihlajamäki, and Vesa Välimäki. Computationally efficient hammond organ synthesis. In *Proceedings of the 14th International Conference on Digital Audio Effects*, Paris, 2011. IRCAM.

Jon Pigrem and Andrew P. McPherson. Do we speak sensor? cultural constraints of embodied interaction. In Thomas Martin Luke Dahl, Douglas Bowman, editor, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 382–385, Blacksburg, Virginia, USA, June 2018. Virginia Tech. ISBN 978-1-949373-99-8. URL http://www.nime.org/proceedings/2018/nime2018_paper0085.pdf.

Miller S. Puckette, Miller S. Puckette Ucsd, Theodore Apel, and David D. Zicarelli. Real-time audio analysis tools for pd and msp, 1998. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.40.6961.

L Quartier, T Meurisse, J Colmars, J Frelat, and S Vaiedelich. Intensity key of the ondes martenot: An early mechanical haptic device. *Acta Acustica united with Acustica*, 101(2):421–428, 2015.

Gordon Reid. Synthesizing tonewheel organs: Part 1. *Sound On Sound*, Nov 2003. URL https://www.soundonsound.com/techniques/synthesizing-tonewheel-organs-part-1.

Curtis Roads. Early electronic music instruments: Time line 1899-1950. *Computer Music Journal*, 20(3):20–23, 1996. ISSN 01489267, 15315169.

Dean Rubine and Paul McAvinney. Programmable finger-tracking instrument controllers. *Computer music journal*, 14(1):26–41, 1990.

Philip Julian Runkel and Joseph Edward McGrath. *Research on human behavior: A systematic guide to method*. Holt, Rinehart & Winston of Canada Ltd, 1972.

Dorien Schampaert. *The Ondes Martenot Network in the Twenty-First Century: The Co-Construction of the Ondes Martenot and its Users*. PhD thesis, University of Leeds, 2018.

Zack Settel and Cort Lippe. Convolution brother's instrument design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 197–200, Montreal, Canada, 2003. URL http://www.nime.org/proceedings/2003/nime2003_197.pdf.

Julius O Smith. Physical modeling using digital waveguides. *Computer music journal*, 16(4):74–91, 1992.

Julius O Smith and Romain Michon. Nonlinear allpass ladder filters in faust. In *Proceedings of the 14th International Conference on Digital Audio Effects*, pages 19–23, Paris, 2011. IRCAM.

Wayne L Stahnke. Method and apparatus for measuring the dynamics of a piano performance, December 29 1981. US Patent 4,307,648.

David T Starkey. Frequency independent ramp generator, October 9 1984. US Patent 4,475,427.

Marian Steyl. *Organ tuition for beginners without previous keyboard skills: A critical evaluation of four published methods against the background of initial piano tuition and general organ pedagogy*. PhD thesis, Stellenbosch: Stellenbosch University, 2018. URL http://hdl.handle.net/10019.1/103748.

Jean-Baptiste Thiebaut, Sean Soraghan, and Lauren Ianni. Multi-dimensional controllers, expressivity, software: an integration problem. In *Journées d'Informatique Musicale*, 2013.

Kai Tuuri, Jaana Parviainen, and Antti Pirhonen. Who controls who? embodied control within human–technology choreographies. *Interacting with Computers*, 2017.

Mark Vail. *The Hammond organ: beauty in the B*. Hal Leonard Corporation, 2002.

Vesa Välimäki, Matti Karjalainen, Zoltán Jánosy, and Unto K Laine. A real-time dsp implementation of a flute model. In *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 249–252. IEEE, 1992.

Vesa Välimäki, Rami Hänninen, and Matti Karjalainen. An improved digital waveguide model of flute-implementation issues. In *Proceedings of the International Computer Music Conference*, pages 1–4. Citeseer, 1996.

Chandrika P Vyasarayani, Stephen Birkett, and John McPhee. Modeling the dynamics of a compliant piano action mechanism impacting an elastic stiff string. *The Journal of the Acoustical Society of America*, 125(6):4034–4042, 2009.

Kurt James Werner and Jonathan S. Abel. Modal processor effects inspired by hammond tonewheel organs. *Applied Sciences*, 6(7):185, 2016. ISSN 2076-3417. doi: 10.3390/app6070185. URL http://www.mdpi.com/2076-3417/6/7/185.

David Wessel and Matthew Wright. Problems and prospects for intimate musical control of computers. *Computer music journal*, 26(3):11–22, 2002.

Red Wierenga. A new keyboard-based, sensor-augmented instrument for live performance. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012. University of Michigan. URL http://www.nime.org/proceedings/2012/nime2012_211.pdf.

Raymon A Willis. System for rejuvenating vintage organs and pianos, October 17 1995. US Patent 5,459,282.

Tom Wiltshire. Technical aspects of the hammond organ, 2008. URL http://electricdruid.net/technical-aspects-of-the-hammond-organ/.

Matthew Wright. A comparison of MIDI and ZIPI. *Computer Music Journal*, 18(4): 86–91, 1994.

Matthew Wright and Adrian Freed. Open soundcontrol: A new protocol for communicating with sound synthesizers. In *Proceedings of the International Computer Music Conference*, 1997.

Qi Yang and Georg Essl. Evaluating gesture-augmented keyboard performance. *Computer Music Journal*, 38(4):68–79, 2014.

Victor Zappi and Andrew McPherson. Dimensionality and appropriation in digital musical instrument design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 455–460, London, United Kingdom, June 2014. Goldsmiths, University of London. URL http://www.nime.org/proceedings/2014/nime2014_409.pdf.

APPENDIX

A

MATERIALS FROM THE HYBRID HAMMOND PERFORMANCE STUDY

A.1 SCORES

We present here the scores used for the study in Chapter 4.

Figure A.1: Lead sheet for Blue Bossa, by Kenny Dorham, which participants played during T1

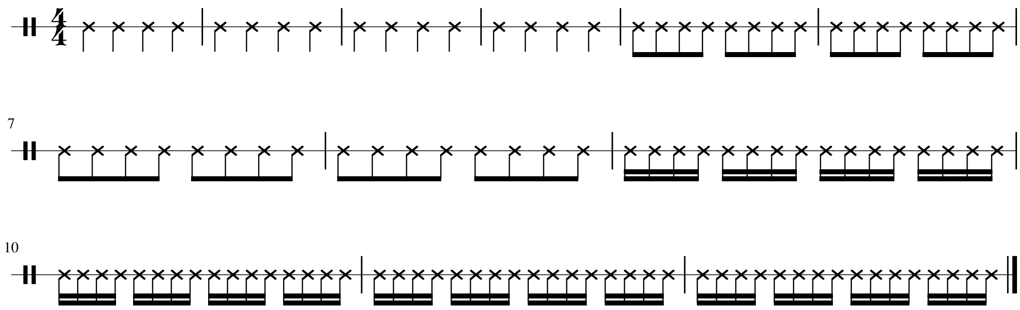


Figure A.2: The rhythmic exercise that participants played during T2

A.2 QUESTIONNAIRE FOR THE HAMMOND PERFORMANCE STUDY

We present here the questions asked in the questionnaire for the study in Chapter 4.

A.2.1 *Task 1*

During T1, participants were asked the questions in Table A.1. The interface presented a dual-sided slider, as shown in Fig. 4.14. For more details see Section 4.6.3.1.

Title	Question	Slider label (left)	Slider label (right)
Temporal control	To what extent were you able to control your timing with each setting?	A is much easier to control the timing of than B	B is much easier to control the timing of than A
Weight	How did you find the weight of the keyboard action under each setting?	A is much heavier than B	B is much heavier than A
Responsiveness	How responsive did you find the instrument under each setting?	A is much more responsive than B	B is much more responsive than A
Fatigue	How tiring was it to play under each setting?	A is much more tiring than B	B is much more tiring than A
Expressiveness	How expressive do you find each setting?	A is much more expressive than B	B is much more expressive than A
Difficulty	How difficult did you find it to play under each setting?	A is much more difficult than B	B is much more difficult than A
Realism	How realistic did you find the instrument under each setting?	A is much more realistic than B	B is much more realistic than A
General preference	What was your general preference of setting?	A is much more preferable than B	B is much more preferable than A
	What difference – if any – did you notice?	n.a. (textbox)	n.a. (textbox)

Table A.1: Questions asked after each trial in Task 1

A.2.2 *Task 3*

During T3, participants were asked the questions in Table A.2. Questions were answered through a slider interface, except "Description", which is in the form of a text box. Each question from Table A.2 was asked four times, one for each of C1, C2, C3, C4. All the questions were presented in a single page, grouped by condition. For more details see Section 4.6.3.1.

Title	**Question**	**Slider label (left)**	**Slider label (right)**
Temporal control	To what extent were you able to control your timing with setting Cx?	hard to control the timing	easy to control the timing
Weight	How did you find the weight of the keyboard action under setting Cx?	very light	very heavy
Responsiveness	How responsive did you find the instrument under setting Cx?	not responsive	very responsive
Fatigue	How tiring was it to play under setting Cx?	not tiring at all	very tiring
Expressiveness	How expressive did you find setting Cx?	not expressive at all	very expressive
Difficulty	How difficult did you find it to play under setting Cx?	easy to play	difficult to play
Realism	How realistic did you find the instrument under setting Cx?	not at all realistic	very realistic
Description	How would you describe setting Cx?	n.a. (text box)	n.a. (text box)

Table A.2: Questions asked for each condition in T3.

B MATERIALS FROM THE FLUTE SYNTHESIZER STUDY

B.1 TRAINING MATERIALS

We present here the training materials used for the study in Chapter 5. The pre-recorded audio example for the training materials are provided at the URL in Footnote 10 on page 147.

Figure B.1: Training materials, page 1 of 5

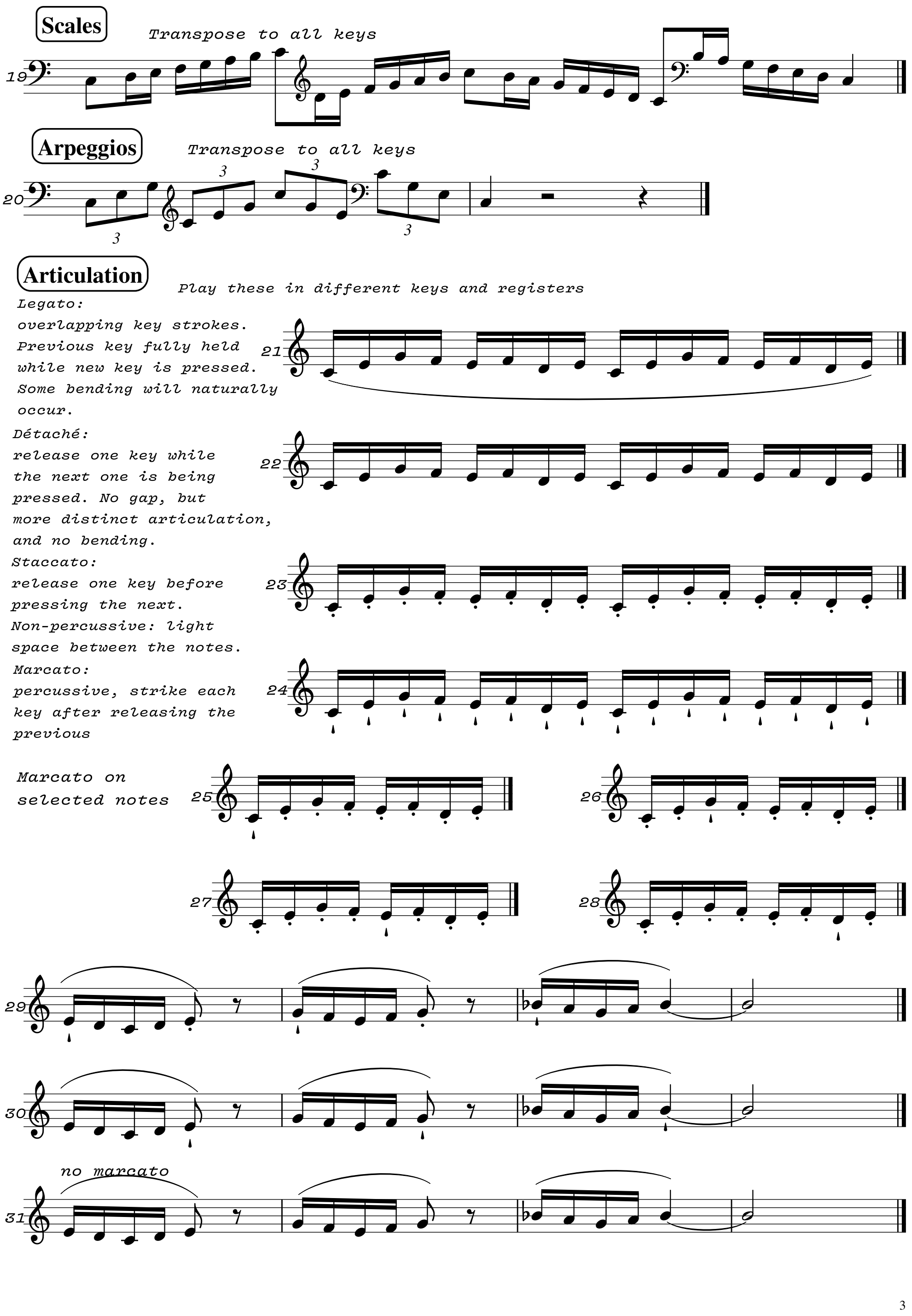


Figure B.2: Training materials, page 2 of 5

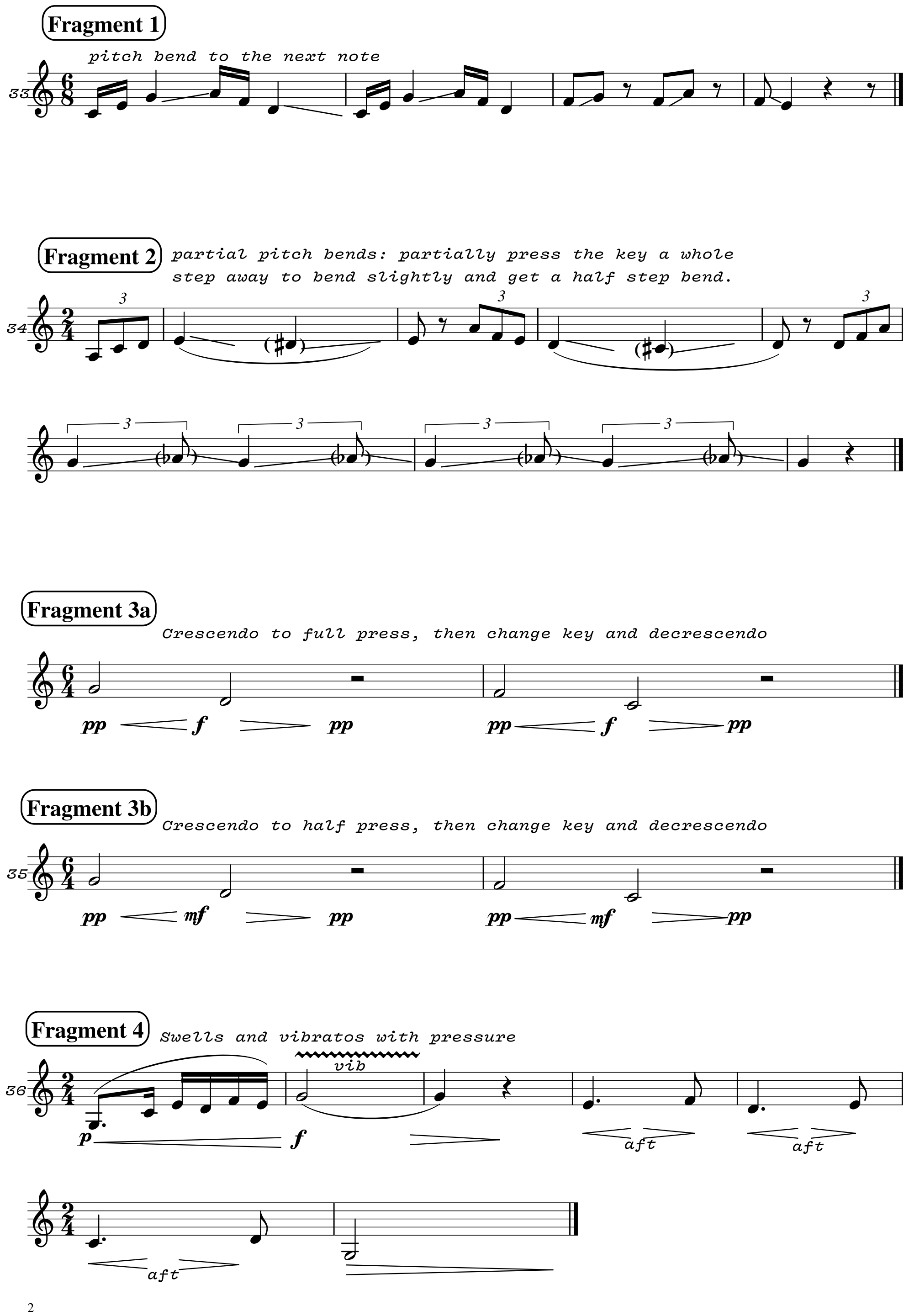
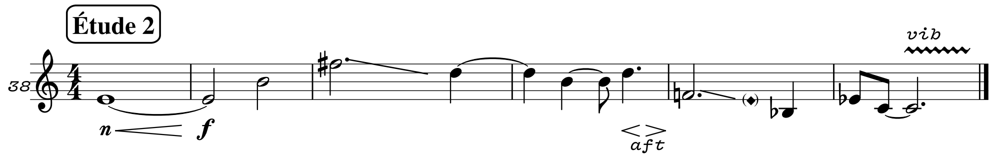
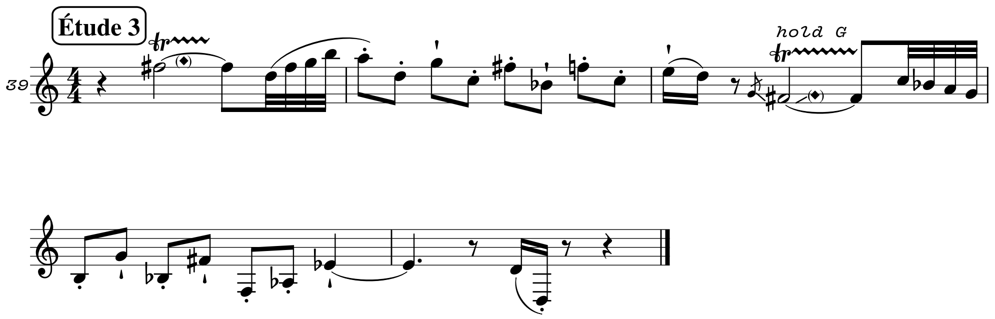
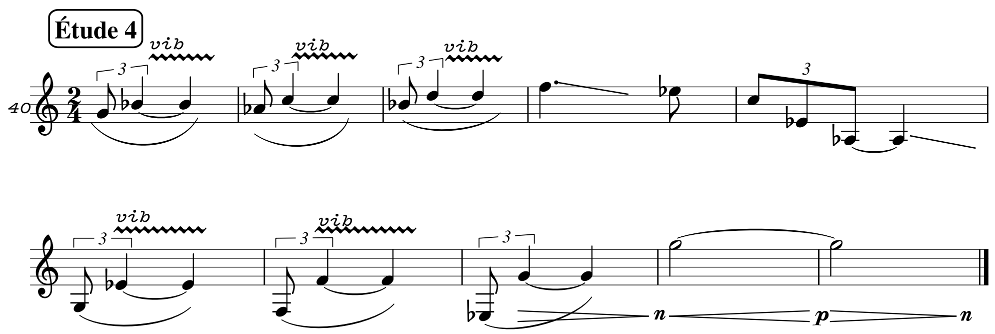


Figure B.3: Training materials, page 3 of 5

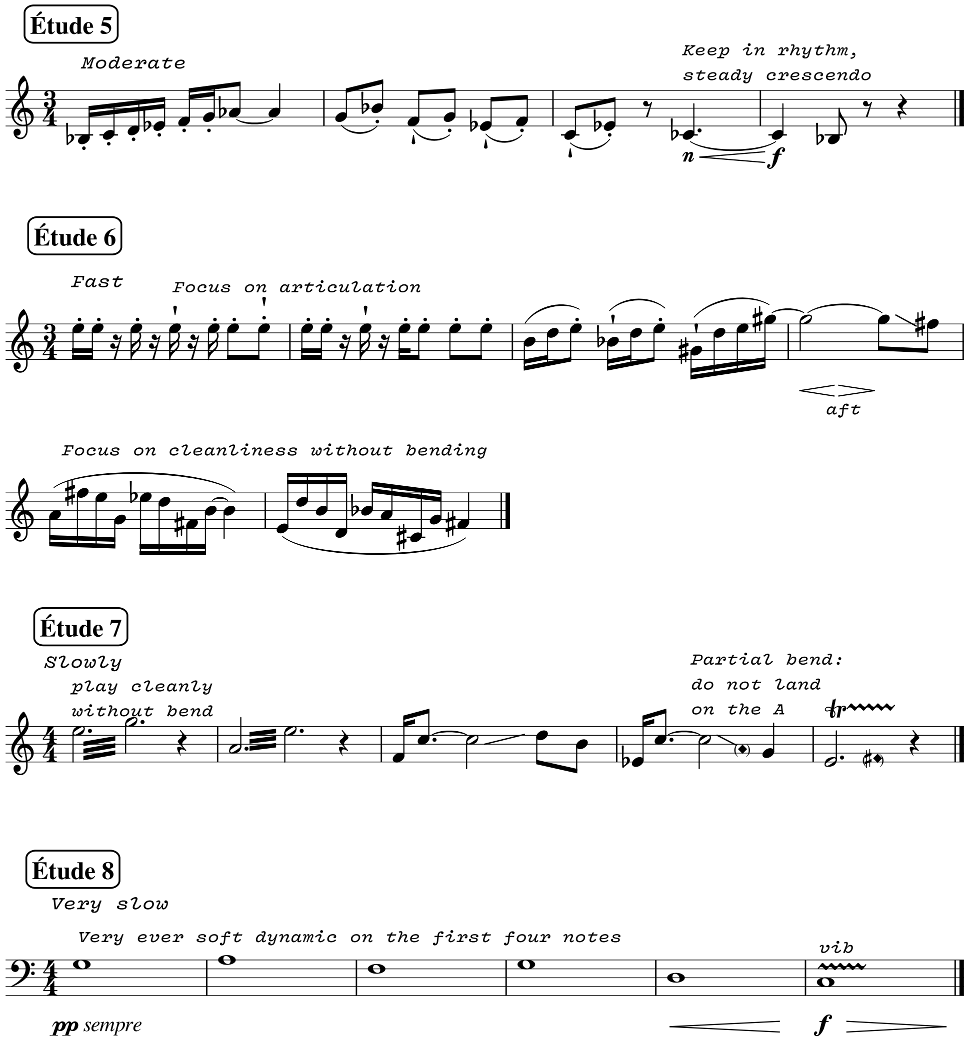






4

Figure B.4: Training materials, page 4 of 5



6

Figure B.5: Training materials, page 5 of 5

Figure B.6: Same score as Fig. B.4, but without the extended techniques notated. This was used as the base for the "Performance evaluation" tasks.

B.2 ANNOTATIONS

Here we present the Études annotated by the participants of the the study in Chapter 5 during the "Performance evaluation" tasks. The participant was presented with a non annotated score (Fig. B.6), and, listening to an audio recording of the pieces, performed with the extended techniques notated in Fig. B.4, was asked to write down the extended techniques they heard on the provided score. Annotations during the first pass (in Session 1) were performed with a pencil, and were later amended or corrected during Session 2 with a blue marker.

Figure B.7: Annotations of Études 1-4 by P1.

Figure B.8: Annotations of Études 1-4 by P2.

Figure B.9: Annotations of Études 1-4 by P3.

Figure B.10: Annotations of Études 1-4 by P4.

Figure B.11: Annotations of Études 1-4 by P5.

Figure B.12: Annotations of Études 1-4 by P6.

B.3 COMPOSITIONS

We present here the scans of the compositions produced by the participants to the study in Chapter 5. The performances of those participants who gave us permission can be found at the URL in Footnote 10 on page 147.

Breathy soft E : explore the edge of hearing.
Add G, "
Gradually increase tone, alternate between them.
Add pitch bends up / down
Introduce "multiphonic" by depressing adjacent key slightly
Gradually introduce melodic material, "bag pipe"-like,
holding E while playing with D, F, F#, D# (with "ornaments"*)
Build up, adding percussive elements
Stop abruptly, end with soft, long, breathy E

* There is a way to slowly activate note 3rd above by depressing key partially, then switching fully - getting a high, distorted timbre.

Aim for ca. 5'

Figure B.13: P1's composition, page 1 of 1

Figure B.14: P2's composition, page 1 of 1

Figure B.15: P3's composition, page 1 of 2

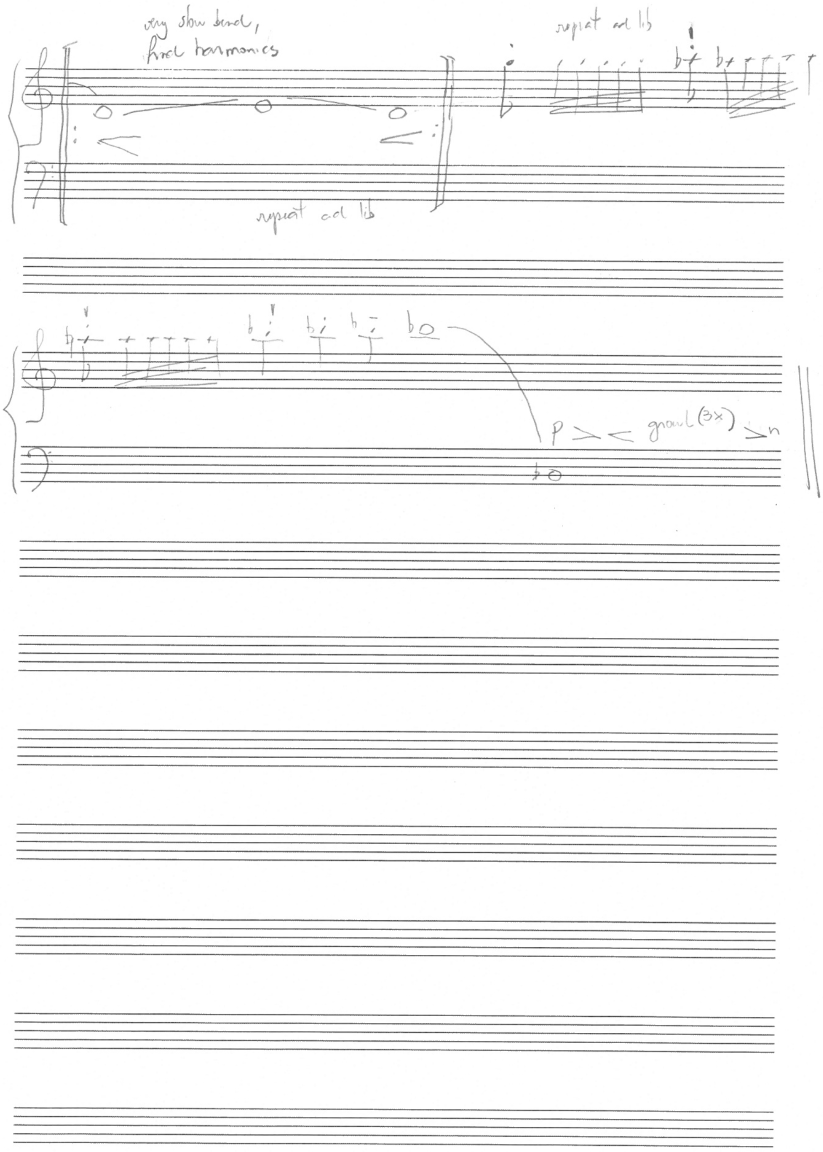


Figure B.16: P3's composition, page 2 of 2

Figure B.17: P4's composition, page 1 of 1

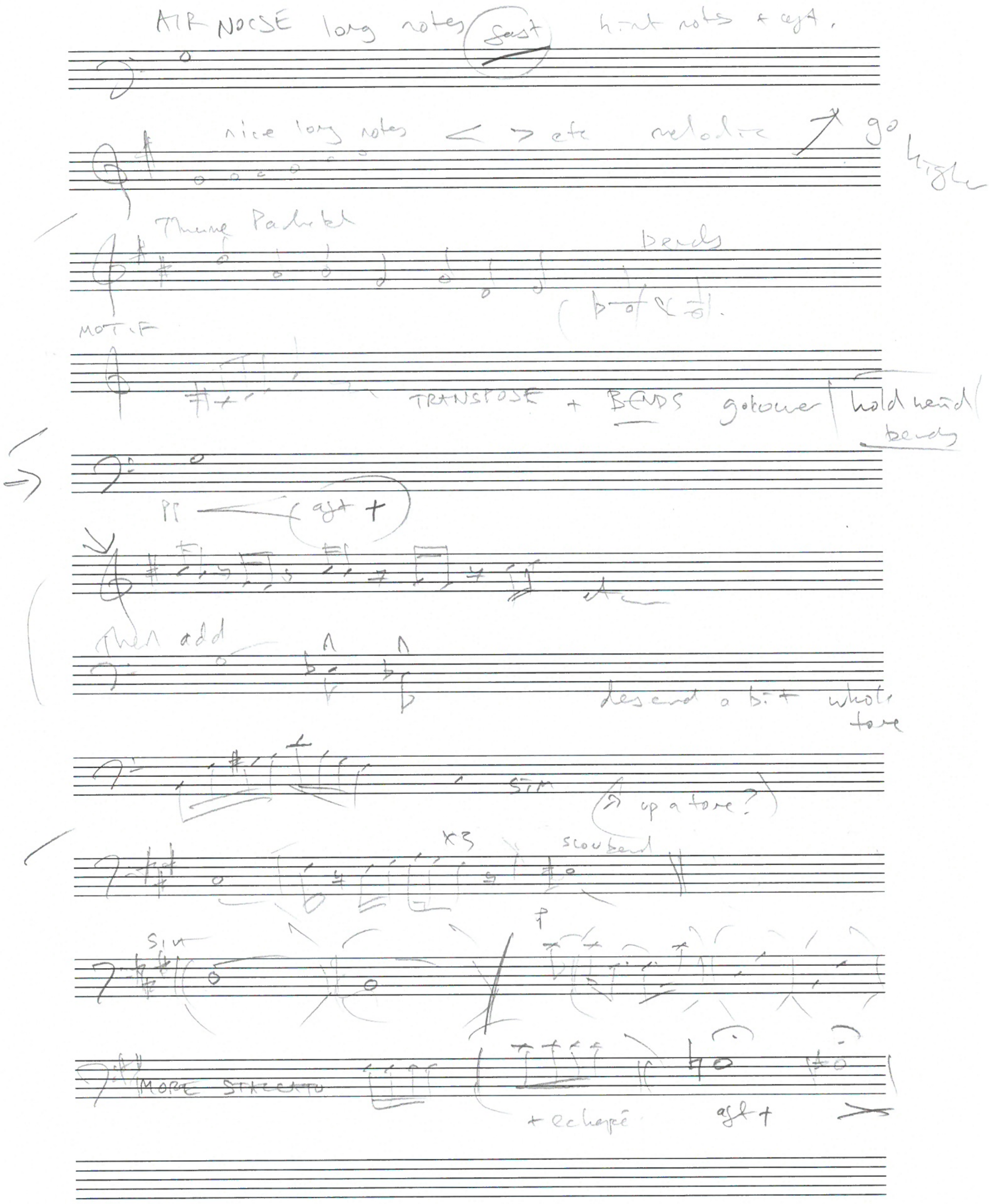


Figure B.18: P5's composition, page 1 of 1

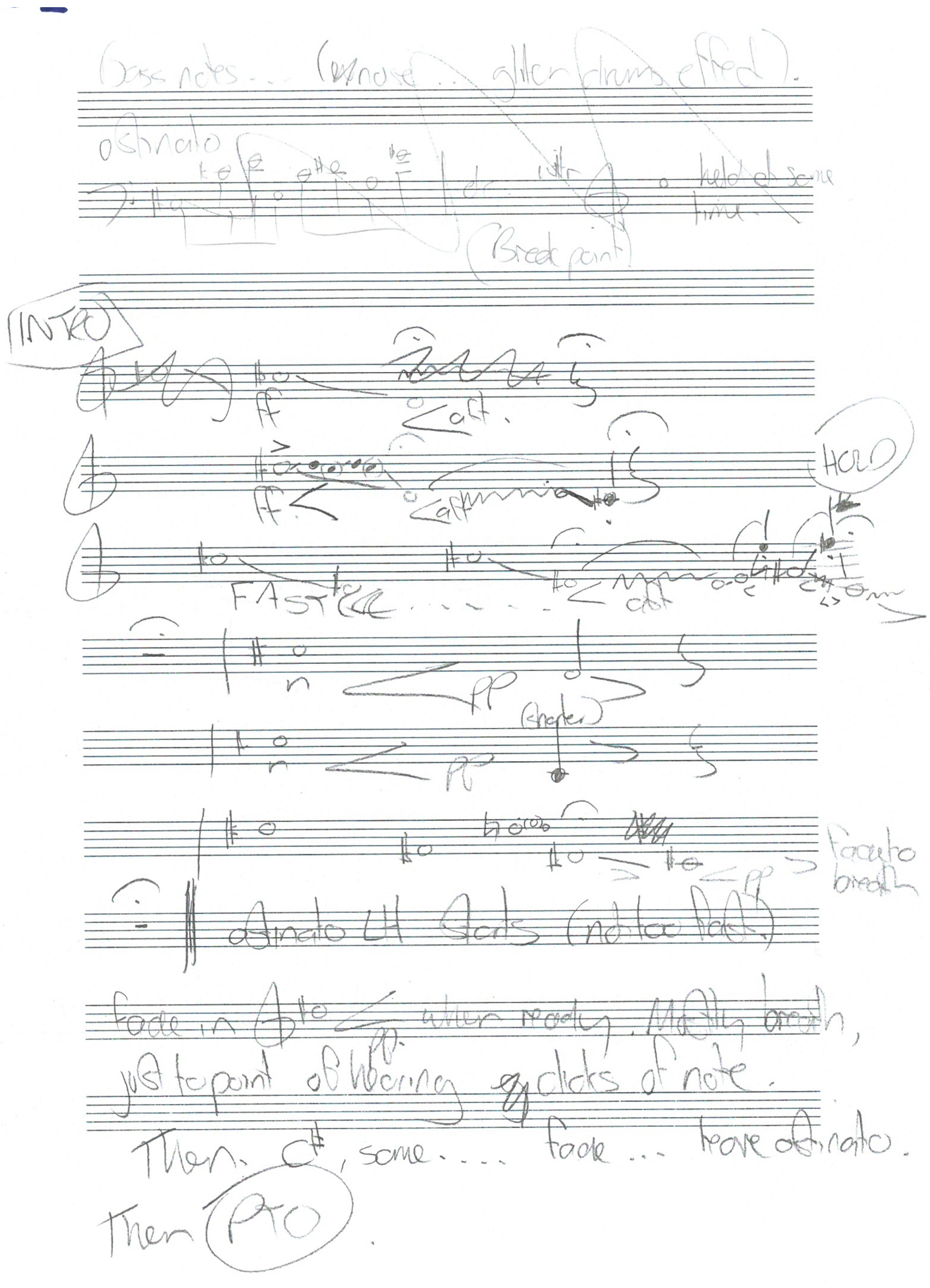


Figure B.19: P6's composition, page 1 of 3

Figure B.20: P6's composition, page 2 of 3

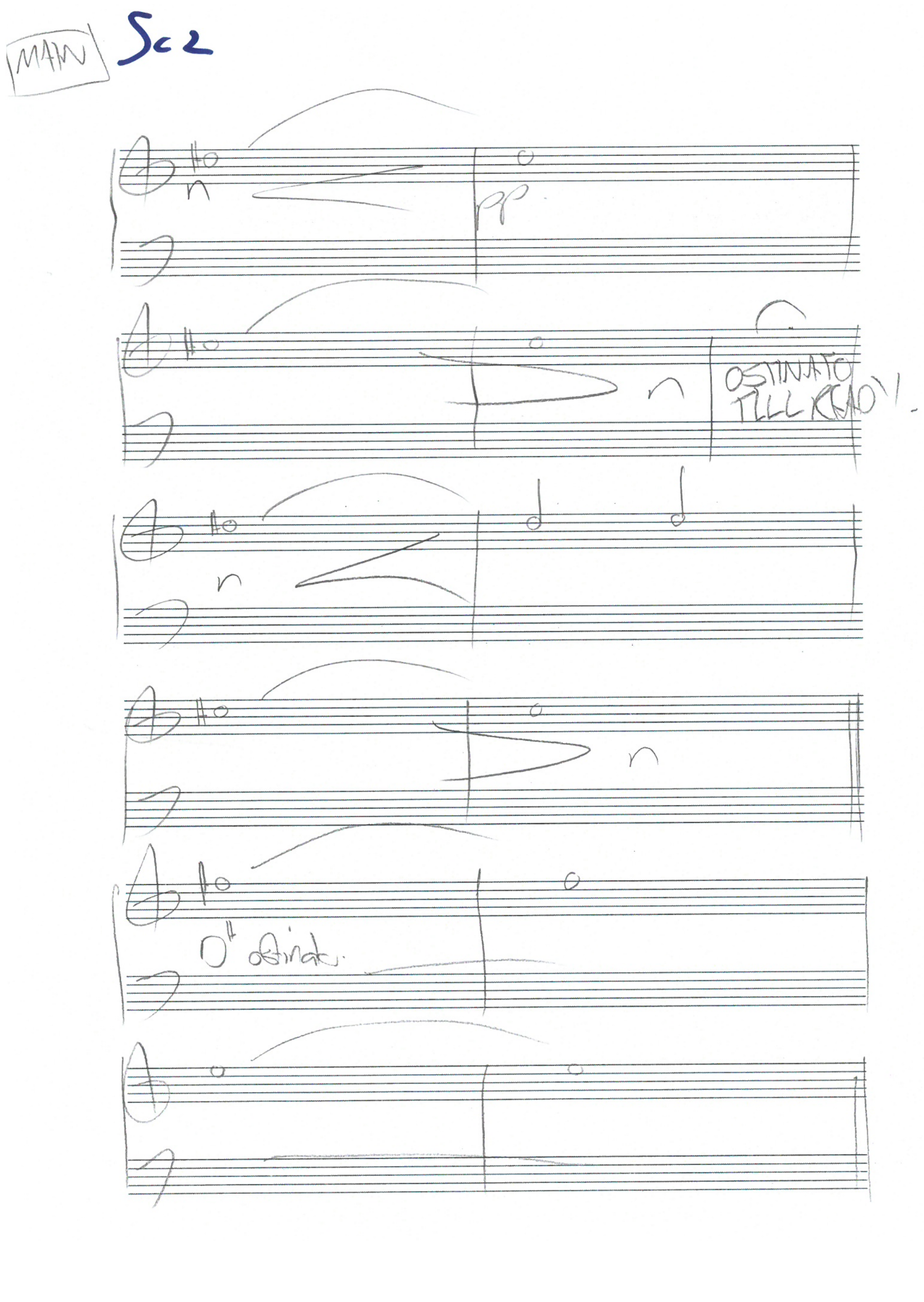


Figure B.21: P6's composition, page 3 of 3