Acoustical Phenomena of Indigenous Instruments: Intercultural Music Immersion Through the Lens of Physics

Abstract

Traditional musical instruments are often passed down and taught from generation to generation without essential information ever being documented. These instruments may have been in use for hundreds of years, yet if one desired to gain a better understanding of their acoustic properties, they would be left without crucial resources. In collaboration with the Department of Music, the traditional Chinese keyed soprano suona was studied. The audio signals of this musical instrument were recorded and analyzed to compare their acoustic properties to those of several contemporary western instruments. The audio waveforms were converted from the time domain into the frequency domain using a Fast Fourier Transform (FFT) algorithm, thereby generating the spectrum of frequencies that compose each sound. The frequency and relative intensity of the different spectral peaks determine the harmonic spectra, which were then compared with the instrument's qualitative timbral qualities. These relative amplitudes and time behavior of the harmonics (overtones) were related to their corresponding effects in terms of brightness, tone, formant, etc. This was done in order to visualize the similarities and differences between several contemporary western instruments (saxophone, bass guitar, and harmonica) and the traditional Chinese suona wind instrument. Results of these investigations will contribute to the development of a book describing instrumentation and orchestration techniques for instruments of the Chinese orchestra. This work provides a valuable, accessible resource for musicians, composers, music historians, and anyone interested in the physics of music.

I. Introduction

Music and sound have been an important part of many cultures throughout history. The songs, instruments, and musical traditions of a culture can give valuable insight into the lives of those before us. While the recording and analysis of these musical traditions is often the work of historians, musicians, and human behavioral researchers, there is a great deal of value that can be provided by examining our musical history through the lens of physics. By applying the types of analysis often employed in the investigation of physical phenomena, we can develop reliable, quantitative metrics to describe the qualities of different instruments and the sounds they produce.

From a physical standpoint, sounds are pressure waves propagated through some medium (usually air). The organs within our ears allow us to perceive these pressure waves as sound. The qualities of a sound are determined by the wave properties, specifically the frequency and amplitude as a function of time. Higher frequency sounds are perceived as higher in pitch, and higher amplitude pressure waves correspond to louder sounds. In practice, a sound is usually composed of many different wave frequencies, the composition of which constitutes the quality or timbre of a sound [1]. Devices like microphones convert sound waves into electrical signals, allowing us to store and process sound waves digitally.

Analyzing how sounds are produced, propagated, and perceived is an incredibly complex topic at the intersection of physics, engineering, physiology and psychology. In the context of musical instruments, the most important information for analyzing sounds is the relative intensities of different frequencies in the wave, often called the frequency spectrum or harmonic spectrum. When a musician plays an 'A' (220 Hz) on an instrument, the sound wave produced is not just a simple $f = 220$ Hz sine wave (the 'fundamental' frequency). Due to resonance properties of wind and string instruments, there are also frequencies at different multiples of the fundamental f called harmonics at frequencies nf, where n is an integer. The relative amplitude of all these higher harmonics is how we can differentiate between the same pitch played on two different instruments. Certainly, playing an A-220 note on a guitar and a tuba sound extremely different. Characterizing the harmonic spectra of different instruments allows us to quantitatively assess the properties of their respective tones.

Analyzing the frequency spectra of sounds requires a method for separating out the component frequencies of a complex waveform. The Fourier transform is one such method that is commonly used in various signal processing applications to transform time domain data to frequency domain data. However, the Fourier transform is a complex mathematical operation involving a difficult to compute integral over all time, which quickly becomes challenging with complex soundwave data. Thankfully, there are discrete Fourier transform algorithms that use numerical methods to quickly compute the frequency spectra of a finite signal. In particular, the Fast Fourier Transform (FFT) algorithm provides an incredibly fast method for computing frequency spectra when working with a finite number of data points in time of the form 2^m, where m is an integer. This allows researchers and musicians to determine frequency information quickly and accurately from a short sound recording.

II. Mathematics of fast Fourier transform

The experiments here rely on the Fast Fourier Transform. When conducting this analysis, it is important to take into consideration the number of data points that are analyzed, the 'frame size'. The larger the frame size, the higher the frequency clarity (resolution. However, due to the nature of the Fourier time-frequency uncertainty ($\Delta t \Delta f \geq 1/4\pi$), [2] a longer continuous note having a larger frame size must be collected to achieve a high resolution of low frequencies. To account for this, the instrumentalists recorded in this study were instructed to sustain each note they played for 1 to 2 seconds to ensure enough data would be available for detailed analysis. Figure 1 illustrates how frequency resolution changes with the number of sine wave cycles collected (frame size). The narrow peak (higher resolution) was generated by collecting 100 cycles, while the much broader peak (lower resolution) was generated by collecting only 10 cycles.

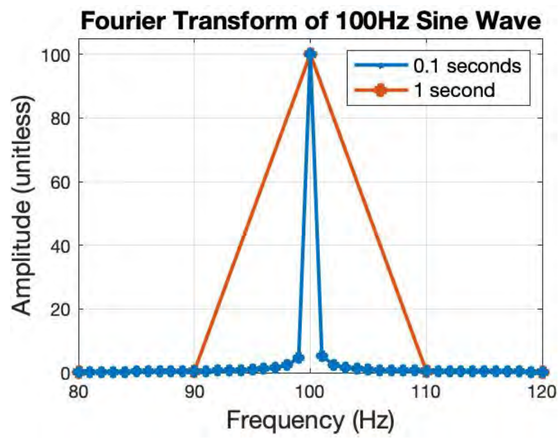

Figure 1: *The Fourier transform frequency spectra of a 100 Hz sine wave. The narrow peak is sampled at a higher rate than the wider peak, illustrating the lower frequency resolution that arises with lower sample rates.*

Before the recordings were taken, it was important to choose a sample rate and bit depth high enough to ensure no aliasing or quantization error. The sample rate is the number of amplitude data points taken per unit time, typically measured in Hz, in units of 1/seconds. As the upper limit of human frequency perception is typically around 20 kHz, a sample rate of 44.1 kHz was used to ensure a minimum of two samples per cycle could be recorded for any frequency up to 20.5 kHz—a value referred to as the *Nyquist frequency* (f_N). The bit depth determines the number of bits used to store each amplitude data point taken, where 1 bit can store 1 binary number, 2 bits can store 4, and so on in an exponential relationship. The minimum bit depth to effectively cover the range of human hearing is 24-bit, and as such, this was the depth used for all recordings.

Finally, before the FFT is applied to the data, windowing must be used to prevent false frequencies that arise from the periodicity of the transformation window. Windowing functions generally fade out the samples toward the beginning and end of the bin, while keeping those in the center at full amplitude, and are a necessity for an accurate analysis. The Hann window was used in this experiment, as it is a commonly used window that is effective for a wide range of purposes. [3]

III. Contemporary instruments

In this section, three contemporary instruments were analyzed for both frequency and amplitude responses in loud and soft dynamic ranges across the pitch range of each respective instrument. The performers were instructed as follows:

- Play through the entire range of your instrument in intervals of roughly 3 to 5 half-steps.
- Play each note twice, at the highest and lowest comfortable volumes.
- Every note should be sustained for at least 1 to 2 seconds.
- Be sure to record in a quiet environment to minimize background noise pollution.

These recordings were then imported into Adobe Audition (2022), a professional-level audio editor and analysis software. For each note, the letter pitch, exact frequency, average RMS amplitude, and dynamic range were recorded. In digital systems, amplitude is measured using negative decibels, with 0 dB being the universal maximum amplitude above which clipping will take place. RMS (root mean square) amplitude is used as it is a measure that takes the average over a period to give a better indication of how loudness of a sound is perceived. Dynamic range is the dB variation measured in the sound and is represented in the intensity graphs as colored line thickness. The amplitude variation was recorded for each note, and an average was taken and used to scale the thickness of the colored lines. The amplitudes contained within the line represent the amplitudes performed by the musician aiming to perform at a consistent volume. The data for each instrument was imported into Matlab (version R2021b), and graphs were generated to highlight the sound intensity as a function of pitch over the range of the instrument [4], as well as an FFT of a single note for each instrument showing their unique timbrel qualities [5]. The note highlighted in each respective FFT is not the same across instruments, but is an arbitrary note chosen to show a representative spectrum of each unique instrument.

While these graphical representations reveal a great deal of information about the various sounds, [6] it is important to note that all the qualities are not addressed. The attack, decay, sustain, and release of a note are commonly used to describe its sound, and can be collectively referred to as the "envelope". The envelope of a sound is what makes percussive instruments sound so aggressive with their extremely short attacks, and what makes reverberating notes so calming with their long releases. These elements can be seen by looking at the waveform of a sound in the time domain as opposed to the frequency domain, but were omitted in this study.

A. Alto saxophone

The saxophone is a woodwind instrument, and one of the latest additions to the chamber ensemble after its invention in 1840s. It has a soft, gentle tone in its quieter register, but a harsh, reedy sound when played loudly (see Figure 2). In the FFT analysis shown in Figure 3, it can be seen that there is a significant presence of upper harmonics that seem to drop off around the 5 kHz range (around the 9th harmonic)—this is likely responsible for the instruments bright tone quality, which aided its success in formats ranging from smooth jazz to hard rock.

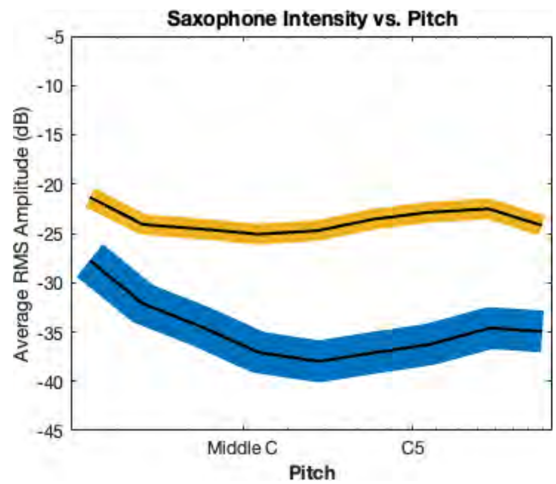

Figure 2: *Dynamic range of an alto saxophone; the upper band shows the intensities of the loudest tone produced, as a function of pitch; the lower band is made up of the softest tone produced.*

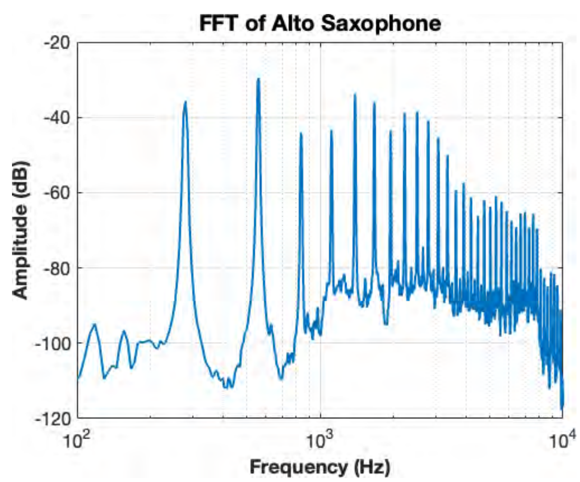

Figure 3: *Frequency spectrum of an alto saxophone taken at an FFT size of 8192 bins.*

B. Bass guitar

The bass guitar is a string instrument commonly found in live performance, especially in rock–adjacent genres. It has a very warm, full sound, but sits comfortably in a lower range, shown in Figure 4. The FFT graph in Figure 5 shows that, while many upper harmonics are present, they decay in volume as they increase in pitch, allowing the bass to achieve its signature dark timbre.

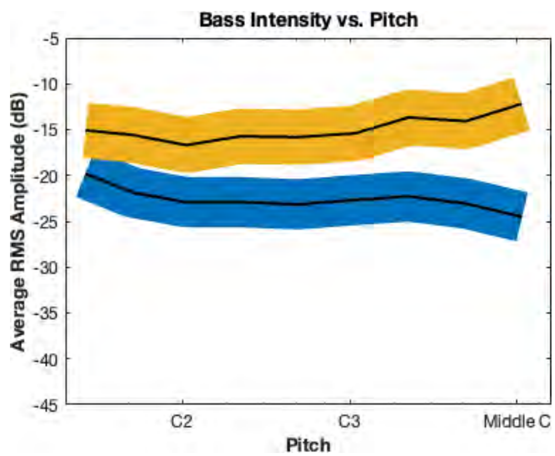

Figure 4: *Dynamic range of a bass guitar; the upper band shows the intensities of the loudest tone produced, as a function of pitch; the lower band is made up of the softest tone produced.*

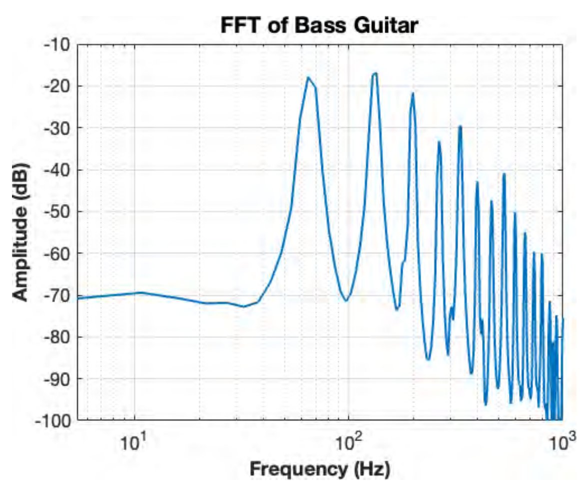

Figure 5: *Frequency spectrum of a bass guitar taken at an FFT size of 8192 bins.*

C. Harmonica

The harmonica is unique as double–reeded instrument, which does not get a great deal of attention in mainstream pop music, but is a staple in "blues" music. It has an extremely reedy tone, and as such, it is no surprise to see that its FFT shown in Figure 7 is rife with prominent upper harmonics that rival the volume of the fundamental until around 10000 Hz (around the 14th harmonic). Note that the dynamic ranges shown in Figure 6 by the broad curves are especially large.

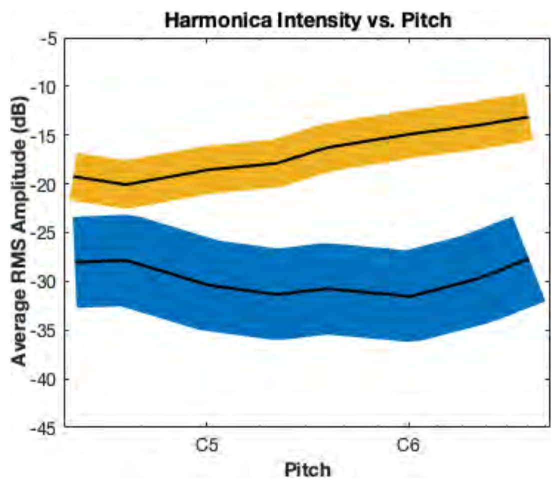

Figure 6: *Dynamic range of a harmonica; the upper band shows the intensities of the loudest tone produced, as a function of pitch; the lower band is made up of the softest tone produced.*

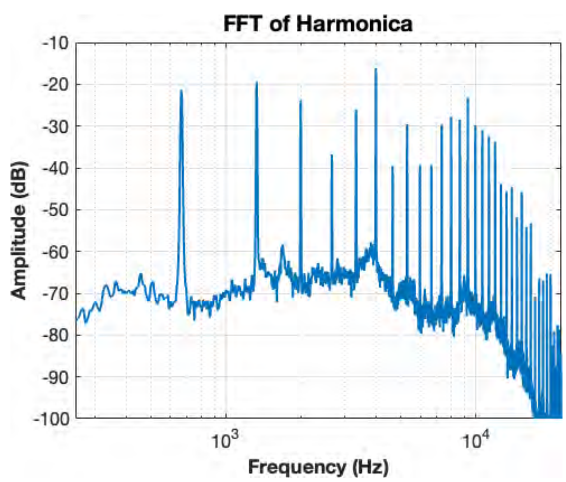

Figure 7: *Frequency spectrum of a harmonica taken at an FFT size of 8192 bins.*

IV. Indigenous instruments

We will use the steps laid out for the above classical instruments to analyze the complex orchestra of traditional Chinese instruments, as well as indigenous instruments from other cultures when available, such as the Afro–Brazilian berimbau. To take these results and compare them with similarly constructed, but better documented instruments from the contemporary Western orchestra is not only an important step towards keeping the intangible cultural heritage of these unique instruments alive, but also provides an avenue for new audiences, and composers interested in writing new music for non-Western traditional instruments. This will enhance the repertoire for the performers of these instruments, and perhaps lead to new innovations in cross-cultural creativity and expression.

A. Chinese suona

The Chinese suona is a double–reeded wind instrument similar in design to the contemporary western oboe. Both instruments have a distinct reedy sound that bears a striking resemblance to the harmonically rich square wave when performed correctly. It can be seen in Figure 9 that, as one would find in the FFT of a square wave, this instrument's spectrum contains notable peaks well into the 10 kHz range. This includes four harmonics which can be seen to sound louder than the fundamental, giving an extremely pointed and bright quality to the sound.

It is important to note that the so–called "softer" line in Figure 8 crosses above the "louder" line on the high-frequency end. This occurs because the upper range of the soprano suona begins to destabilize above a G#5 (830 Hz). These pitches become difficult to achieve even for a professional, due to their instability, and especially difficult to achieve at a consistent loudness level, even with a performer's best efforts at playing equally loud (or soft) tones. As a result, we see strange circumstances such as those in Figure 8. Because of this, the upper register of the instrument is used sparingly in performance, and especially in situations requiring precise dynamic control. However, it was included in this experiment in order to provide the fullest documentation possible as we begin to explore the instrumentation and orchestration possibilities of this non-Western traditional instrument [4].

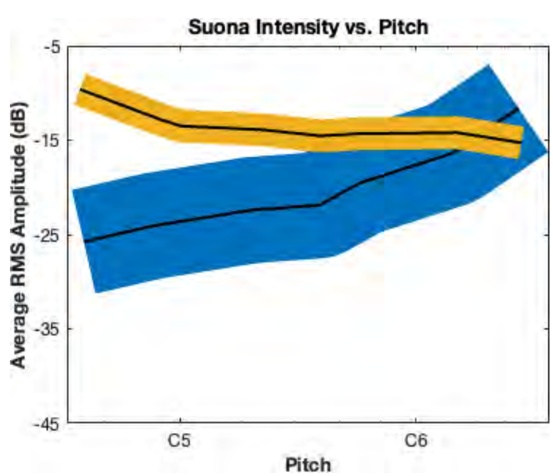

Figure 8: *Dynamic range of a Chinese keyed suona; the upper band shows the intensities of the loudest tone produced, as a function of pitch; the lower band is made up of the softest tone.*

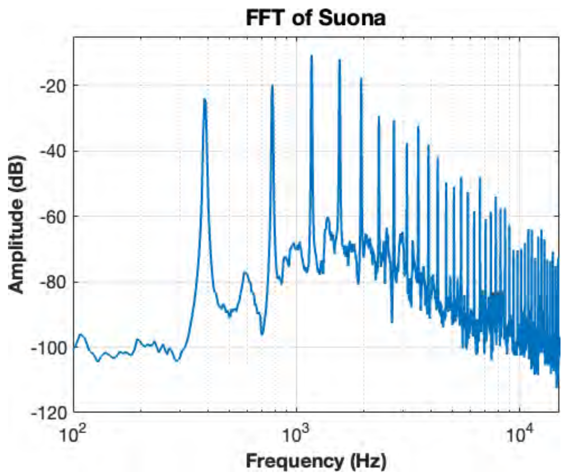

Figure 9: *Frequency spectrum of a traditional Chinese suona taken at an FFT size of 8192 bins.*

V. Conclusion

These acoustic experiments demonstrated that the timbral qualities and playing capabilities of musical instruments can be represented quantitatively through visualizations of data in the spectral domain as well as relative intensity curves. It was seen that instruments tend to have a wider dynamic range in their softer register and a narrower range in their louder register. Strong presence of upper harmonics was seen to align with a bright tone color as expected in both the alto saxophone and the harmonica. The mathematics behind a Fast Fourier Transform and general digital audio processing were adjusted to ensure that optimal recordings and processing could be achieved. These experiments will soon be repeated with traditional instruments in order to support a foundation for better documentation. Results of these experiments will be a key contribution to a currently in-process book on instrumentation and orchestration techniques for ancient Chinese traditional instruments.

Acknowledgements

We would like to acknowledge the contributions of the Northeastern Department of Music for their generous support in connecting us with musicians and acquiring instruments, as well as support from the Department of Physics and the College of Science. We would also like to acknowledge Professor Hubert Ho from the Northeastern Department of Music for his contribution to the process behind the acoustic analyses of the instruments. We thank Guo Yazhi for recording the suona sound samples used in this experiment. The experiments grew out of a physics lab experiment on acoustics in the course, *Advanced Physics Lab*, at Northeastern University. [7]

References

[1] W. Jian, J. Liu, X. Zhang, S. Wang, and Y. Jiang, "Analysis and Modeling of Timbre Perception Features in Musical Sounds," Appl. Sci. **10**(3), 789(2020); https://www.mdpi.com/2076-3417/10/3/789.

[2] J. N. Oppenheim and M. O. Magnasco, "Human Time-Frequency Acuity Beats the Fourier Uncertainty Principle," Phys. Rev. Lett. **110**, 044301 (2013).

[3] J. Robinson Pierce, *The Science of Musical Sound*, p. 126

[4] H. F. Olson, *Music, Physics, and Engineering*, p. 224-226.

[5] Hong Kong Chinese Orchestra, "Standard Orchestra Instrumental Range Chart" (2022).

[6] M. Schutz, "Sculpting sound: A new tool for visualizing musical timbre," https://maplelab.net/maestro/.

[7] D. Heiman and H. Kumarakuru, "Acoustics and Fourier Transform," https://web.northeastern.edu/heiman/3600/index.html.

Alto Saxophone	Bass Guitar	Harmonica	Soprano Suona
			

Table 1: *Appendix of instruments included in the experiments.*