UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Analysis of Vibrato in Human Vocals and String Instruments

PHYS 398: Design Like a Physicist

Sergi Castells, Stephen Elliott,
Ramone Randle, Noah Rogers

May 1, 2020

Abstract

Vibrato is a technique that is used in nearly all musical pieces, ranging from orchestras, operas, and other classical performances to contemporary music. Vibrato is a mechanical, musical technique that causes oscillations in frequency about a given note which, when executed properly, also causes equivalent oscillations among a note's overtones. All sung notes naturally have some vibrato and the oscillations present are consistent among overtones when vibrato is naturally occurring rather than in cases of false vibrato. String instruments, however, can produce cleaner, crisp notes without the presence of vibrato. The backbone of our analysis is the Fast Fourier Transform algorithm as we are primarily interested in the oscillations of a note's frequency. We analyze short audio samples, typically three to six seconds long, consisting of human vocals and stringed instruments to characterize vibrato from varying sources. A negligible difference in fundamental frequencies of notes is apparent between stringed instruments and human vocals. Analyzing the spectrograms, we found an average percent difference of notes' amplitude in frequency of the fundamental harmonic between notes containing vibrato and their non-vibrato counterparts of 17.9% for human vocals and 15.6% for cello.

Contents

1 Introduction

Vibrato is a common musical technique that is used throughout all musical genres and styles. This ranges from classical pieces and operas to contemporary music. It is commonly thought of as warbling of notes in vocal or instrumental sounds. In more scientific terms, vibrato is a mechanical, musical technique that causes oscillations about a given note which, when executed properly, also causes equivalent oscillations among a note's overtones [9]. Overtones are notes of any frequency greater than the fundamental frequency that may be heard with the original note. The oscillations about a note do not exceed a semitone[1] in all correct executions of vibrato. An explanation of semitones can be found in Fig. 1. In all sung notes vibrato is present, however instances of false or faked vibrato can vary the pitch widely since humans can force the modulation of the diaphragm or jaw at frequencies that result in pitch modulation of greater than a semi-tone [10]. On the contrary to vocals, string instruments can produce cleaner, crisp notes without the presence of vibrato without much training of the player. This is not to say that stringed instruments cannot produce notes with vibrato as they can and often do in contemporary music.

There is a point of contention when defining vibrato and tremolo in the musical community [9, 15, 16]. The definition of vibrato is the modulation of the pitch about a note while tremolo, a notably distinct technique, has two definitions depending on the instrument. The first definition of tremolo is the rapid repetition of a note, on a stringed instrument for example, and the second definition is the modulation of volume, which is primarily used on electronic instruments [16]. The other, less common class of thought, which is primarily only a position of electric guitar players, is that vibrato and tremolo are different types of the same technique, namely pitch vibrato and amplitude or volume vibrato. The common misunderstanding that tremolo is a type of vibrato is larger the fault of Leo Fender, the creator of Fender guitars, who mistakenly coined the name tremolo bar, with other common names including whammy bar and vibrato bar, and gave rise to nearly a century of incorrect interpretation of vibrato and tremolo [2]. This started the colloquialism of defining tremolo and vibrato as the same technique, which was continued throughout the 20th and 21st centuries by guitar players. This was exacerbated by the popularity of Fender guitars. Thus, for the sake of both the accuracy and simplicity of this analysis, we distinguish between the two techniques and classify them as separate entities. This is to say that we only analyze the vibrato in

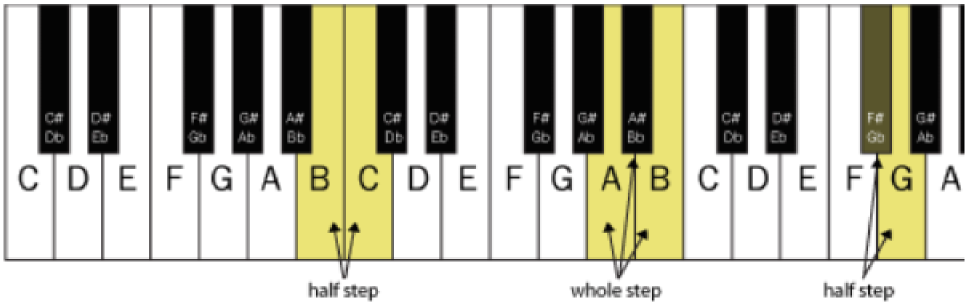
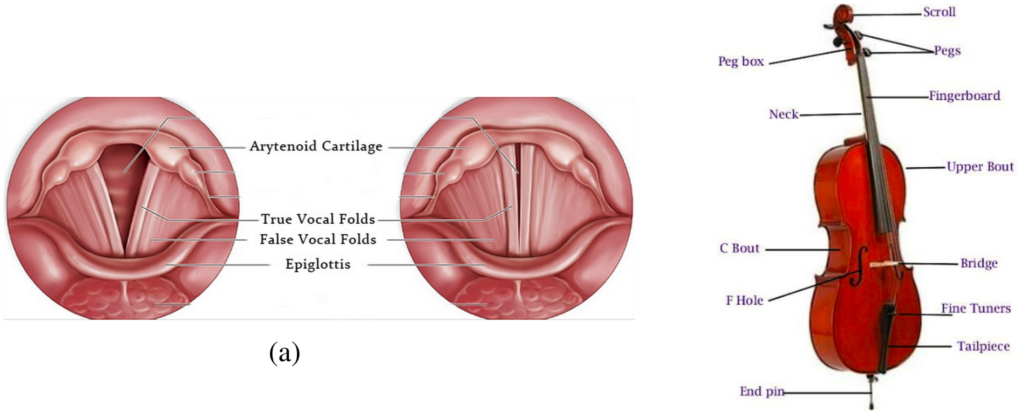


Figure 1: A demonstration of whole and half steps with respect to the keys of a piano. Whole steps are a white keys to white keys except B to C and E to F. Half steps are white keys to black keys and the aforementioned exceptions to whole steps [12].

---

[1] Also commonly referred to as a half step. It is the smallest interval used in Western music.



(a)

(b)

Figure 2: a) A diagram of vocal cords situated at the top of the larynx [6]. b) A diagram of a cello with features denoted. Strings are held along the fingerboard and bowed above the bridge [19].

samples and entirely disregard any potential tremolo apparent in our samples.

Due to the repeated discussion of vibrato further into this report, it is prudent to discuss both the physiological and mechanical aspects of vibrato in vocals and stringed instruments, respectively. The physiological mechanism that allows humans to produce notes is a result of the larynx, which is colloquially known as the voice box. The cricothyroid muscle is the muscle specifically responsible for controlling pitch and stretching the vocal folds. It is with a combination of airflow and the use of the cricothyroid muscle that humans can produce notes. Vibrato arises due to the natural oscillations of tensed muscles within the human body [9]. These oscillations, which occur naturally when any muscle is tensed, vibrate the vocal folds thus creating what we hear as vibrato. The fact these oscillations occur every time a muscle is tensed is precisely the reason why all sung notes have some vibrato. Vibrato may be more present in a skilled vocalist's singing depending on preference, however it always a result of vibrations of the vocal folds. On the contrary to this naturally occurring vibrato, there is false or faked vibrato that arises from quivering of the jaw or diaphragm [10]. This is, as the name suggests, a false vibrato and is a forced and incorrect method to produce vibrato typically apparent in unskilled vocalists' singing,.

The physiological mechanisms that produce notes and vibrato within human vocals share some similarity with the mechanisms that produce notes on all stringed instruments. The vibration of strings produce notes similarly to the vibrations of the vocal folds. This is true for all stringed instruments. Sounds are produced by vibrations of the strings and certain notes are produced by certain lengths of the strings. The change in length of strings on all stringed instruments comes from pressing down on certain spots along the string and is the same process for all stringed instruments. Due to the availability of data and comparative simplicity of bowed instruments to guitars and other stringed instruments, we focus on bowed instruments and specifically the cello. For bowed instruments, playing a single note is rather simple. A string is held down at a given position to produce the desired note and either the bow is run across the string or the string is plucked. Adding vibrato to a note requires varying the pitch which is achieved by rolling the finger along axis of the string [13]. As stated previously, while all bowed instruments would be equally suited to study, the instrument of choice for this report is the cello.

The focus of our project is characterizing vibrato in both human vocals and stringed instruments. There are many techniques used to analyze sound with each being best suited for a given type of sample [14]. An obvious contender for our analysis is a Fourier Transform as we are primarily interested in the oscillations of a note's frequency. We analyze audio samples consisting of human vocals and cello notes to characterize vibrato in both cases qualitatively and quantitatively. The audio samples are intentionally cut into short time intervals in an attempt to reduce noise from extraneous notes. The hardware used for this project consists of a group of sensors that are compatible with the Arduino Mega 2560. For the software and data analysis, multiple common scientific Python libraries were utilized in our analysis; NumPy, SciPy, and Matplotlib are our main imported libraries, among others [7, 8, 11].

2 Hardware

There was a two-fold approach to creating the hardware for this project. Breadboards were used initially to aid with the experimental phase of circuit building and to allow for the testing of multiple sensors and tools. The final circuits built on the breadboards were then transferred to printed circuit boards (PCBs). There is one breadboard and one PCB for each of the group members along with one Arduino Mega 2560 per board. The breadboards were also used as backups in case the PCB failed or was unavailable. The PCBs were the primary equipment used for data collection and allowed for greater portability than the breadboards.

All of the sensors and equipment are compatible with the Arduinos. The sensors and tools used to complete both the breadboards and the PCBs are an LCD, contrast setting potentiometer, keypad, microphone, microSD breakout board, INA219 current/voltage regulator, and a battery pack. An Arduino acts as the processor for the board while everything else can be treated as extra components. All of these sensors required soldering to prepare them for installation on either of the hardware devices. There was also some soldering required to prepare the PCB for the installation of an Arduino and all of the other equipment.

2.1 Sensors and Equipment

The sensors and other various equipment require some explanation to make the analysis and discussion of data easier to interpret. Starting in no particular order, the liquid crystal display (LCD) is one of the tools used to make data acquisition easier. The LCD is not used to collect any data but is instead used to relay information. This information may be related to the sample, such as length or size, or may be used as a method to relay that the data acquisition is complete. The LCD is able to take in information from the Arduino through it's many pins and is then able to display messages across its screen. The LCD relies on a device called a contrast setting potentiometer. This is a variable resistor that is used to set the contrast of the LCD screen. The potentiometer has three electrical connections and a knob, allowing it to adjust the path that current has to flow, which is what allows it to vary its resistance. The analog knob allows for smooth varying of the resistance and hence, smooth variation in the contrast of the LCD.

The keypad is a device that we use to interface with the Arduino. It has buttons that are labeled 0-9 along with # and * in a 3x4 matrix. Each column and each row corresponds to a different pin connected to the Arduino, hence the 7 utilized (or 8 total) pins soldered to keypad. When a button

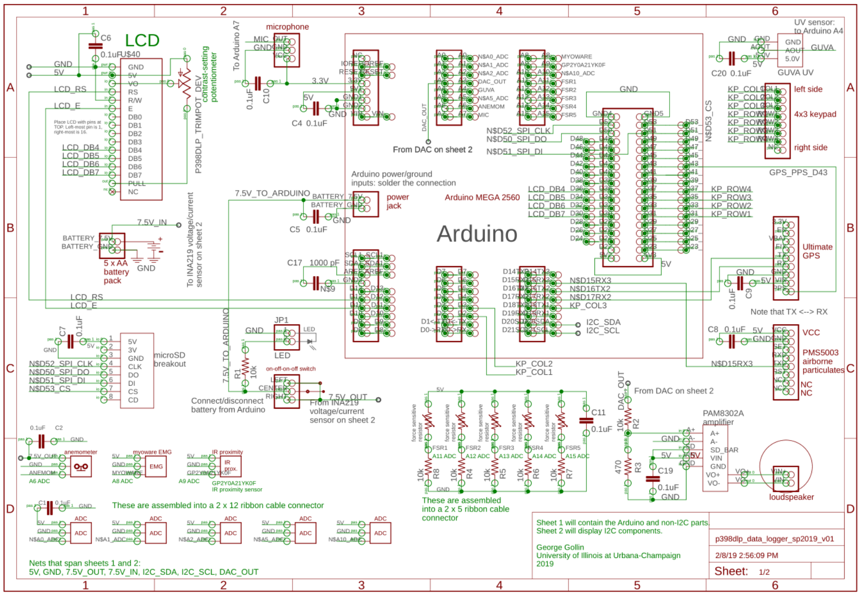


Figure 3: A schematic of the circuit design for both the breadboards and PCBs [5]. Not all components were implemented into the final design and the INA219 is not shown.

is pressed, two signals are sent through their respective pins, signalling the coordinate of which button was pressed, and subsequently released, on the Arduino. This allows for specific inputs from the user, such as a stop and start command used during data acquisition. Another important device used for data acquisition is the MicroSD Breakout Board. The MicroSD Breakout Board is used to interface with a micro SD card, which stores the samples for analysis. The breakout board has a series of pins that allow for data input and output to and from the micro SD card. This allows for information to be saved and analyzed later. This drastically improves the functionality of our device as it allows for the separation of data acquisition and data analysis into disconnected processes.

2.2 Microphone

The microphone is arguably the most important piece of hardware. It records with a 32kHz sampling rate[2] and is the component that this project relies on the most. The microphone used is an electret microphone, which is a type of electrostatic capacitor-based microphone. The electret in the microphone is a stable, permanently charged dielectric material. This electret is placed in parallel with a conductive backplate and does not require power to operate. Sound waves cause the electret to vibrate which produces a small electrical output. For the output to be useful, an amplifier is required, which does require power to operate.

---

[2] It should be noted, however, that the sampling rate used in the final WAV files was 44.1kHz. An explanation can be found in the Data Collection section.

2.3 Power Management

There are several important components that enable the use of all other components. The power management related components are the INA219 and the battery pack. The INA219 is a current and power monitor with an I2C interface. It allows us to monitor the voltage and current between the batteries and the Arduino. Its location in the circuit is between the power source, in this case five AA batteries, and the Arduino. The battery pack, however, is not a sensor but rather a necessary component of our setup that enables us to use the breadboards and PCBs without needing a connection to an external power source. The battery pack takes five AA batteries and allows for approximately 5-6 hours of constant use before replacing the batteries.

3 Data Collection

Initially, data was to be collected from volunteers within the music department inside soundproof rooms in the Krannert Performing Arts Center. However, due to complications that arose during the COVID-19 pandemic, methods for data collection for the project were modified. Preliminary testing samples were taken from YouTube and other online sources, initially using the microphone on the PCB to record them and then, in some cases, downloading samples directly. Due to the change in the data acquisition method, there was a noticeable increase in noise in samples when the microphone was used to re-record audio from the internet. That is to say that noise was compounded from the original recording setup, YouTube's compression algorithm, the speakers playing the sample, and from the PCB. To clarify, noise is defined as undesirable and extraneous sound. It comes from a source that is not the instrument or voice that is desired. Noise has the potential to detract from the experiment if it is apparent at an extreme enough level as it would reduce the quality and clarity of data being analyzed. Although the magnitude of the noise has increased in those samples, it has not increased to the point that would cause an issue when analyzing the data. Thus there does not appear to be any severe adverse effects to using data from the internet. Additionally, noise was naturally suppressed during analysis, cutting out the majority of random noise. The final samples used to characterize vibrato, however, were recorded using an iPhone and the microphone on the PCB for cello notes and vocals, respectively. The notes chosen to be analyzed were G2, D3, and E4 for both cello and vocals. All final data presented in this paper was collected via friends and family. The initial testing phase helped us determine which parameters are useful to characterize vibrato and which instruments' samples are too complex to analyze. The final recordings in particular helped us directly compare identical notes from a cello and human vocals as well as vibrato and non-vibrato notes from the same source.

3.1 Control Group

A non-vibrato control group was used to characterize and isolate differences between notes with and without vibrato. The non-vibrato control group consisted of single notes played or sung at particular frequencies with no modulation. The group of notes consisted of G2, D3, and E4, which are at a frequency of 98Hz, 146Hz, and 330Hz, respectively. For reference, a middle C has a frequency of 261Hz. This standard set of notes was then compared to recordings of the same instrument or singer repeating the corresponding note with vibrato. This allowed for direct comparison of notes with and without vibrato and aided in the characterization of the technique as a whole.

3.2 Data Acquisition Software

The data acquisition software (DAQ) was entirely contained to the Arduino in internal storage. The DAQ activated the microphone and started recording upon the keypad being pressed. The recording was then saved as a binary file to the microSD card after the keypad was pressed again. This file could then be extracted from the microSD card and analyzed remotely using our analysis software. During the recording process, the DAQ activates the microphone and enables 32kHz audio recording. The raw output was stored as a binary file of ADC counts. The code was a replication of an Arduino code that activates the microphone, LCD, and keypad simultaneously [4]. A delay was present with input from the keypad, adding extraneous runtime to edges of our samples, and was dealt with by clipping the recording appropriately during the analysis phase.

3.3 Software

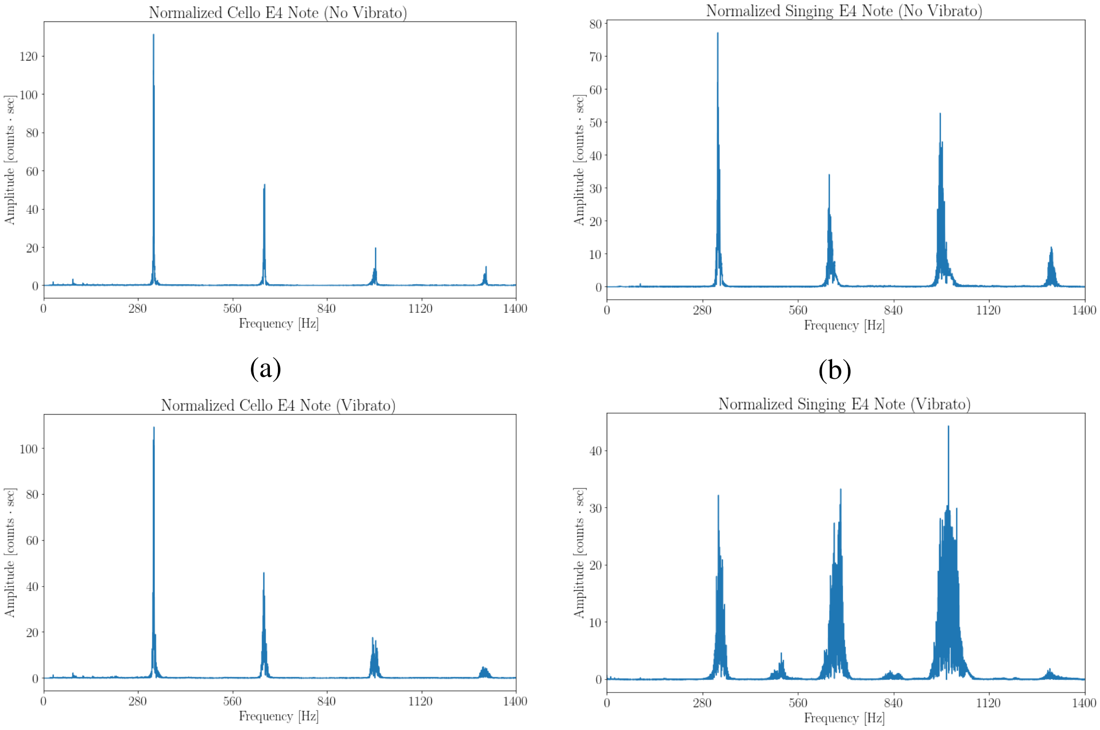
The analysis software was written in Python and utilized a variety of functionality from the Matplotlib, NumPy, and SciPy libraries. The first portion of the analysis software came from code that converted binary files to .WAV files [4]. This was used to convert the files on the microSD Card into usable data. The second portion of the code is responsible for the conversion of .WAV files into their respective FFT plots. The generation of the plots relied on the SciPy Fast Fourier Transform function and, as such, there was an option between a frequency based FFT and an amplitude, or intensity, based FFT. While the amplitude FFT was not used explicitly, the code for creating plots of amplitude and plots for spectrograms are still present for future analyses.

The plots were generated automatically when fed .WAV files from a local directory. The plots, some of which are shown in Figure 4, show the fundamental frequency, or the frequency of the note played, as the peak with the lowest frequency of all large peaks. The amplitude of the peaks, fundamental and overtones, is given by ADC counts $\cdot$ seconds. However, these plots have been normalized to allow for easier interpretation and comparison. The frequency is clearly in units of Hz, and the peaks correspond to a note's fundamental frequency and overtones in order of lower to higher frequencies.

Additionally, spectrograms were produced using pre-existing software known as Sonic Visualiser [17]. The spectrogram is a visual representation of frequency, in Hertz, versus time, in seconds. The plot utilizes a heat map where the intensity, in dB, of a signal appears as colors on a spectrum between cold and warm colors where warmer colors denote higher intensity.

4 Analysis

The backbones of the analysis are the Fourier Transform, which allows us to analyze the frequency spectra of samples, and the spectrogram, which allows for a thorough analysis of the waveform and its overtones by providing something similar to a heatmap. An in-depth explanation of the Fourier Transform can be found in Appendix A.



(a) (b)

(c) (d)

Figure 4: Figures a) and c) are the Fourier spectra for a cello playing an E4 note without vibrato and with vibrato, respectively. Figures b) and d) are the Fourier spectra for a vocalist singing an E4 note without vibrato and with vibrato, respectively.

The type of Fourier Transform that was utilized was known as a Discrete Fourier Transform (DFT). A DFT can be summarized as a Fourier Transform that is performed over a finite or discrete interval of time. The algorithmic implementation of the DFT is commonly referred to as a Fast Fourier Transform (FFT), however DFT and FFT are often used interchangeably. Since the Fourier spectra served mainly to bolster qualitative data, we decided to supplement our Fourier spectra with quantitative data from spectrograms to get a more complete understanding of the data.

To give a more thorough description of a spectrogram, a spectrogram is a visual representation of the frequencies of a signal as it varies with time. Spectrograms essentially take FFTs of small windows of time, the center of that window being the time coordinate of the data point. Spectrograms are sometimes otherwise referred to as sonographs, voiceprints, or voicegrams. Examples of the Fourier spectra and spectrograms used in analysis are present in Figures 4 and 5. While the FFTs are capable of depicting amplitude as a function of frequency, spectrograms depict frequency as a function of time instead. This variation in types of data representation provides many different ways to interpret our data, and as such, it was essential to our project to include both FFT analysis and Spectrogram analysis.

5 Results

5.1 Fourier Transforms

Analysis of the Fourier spectra plots for stringed instruments showed that when a note is played, even if there is no vibrato, there is still a small range of frequencies, other than the fundamental frequency and its overtones. This is represented by the Fourier spectra plots' peaks having bases that cover more then one frequency. When compared to the graphs for vocals, we can clearly see that for cases of both vibrato and no vibrato the bases for vocals are much wider on both overtones and the fundamental. Qualitatively speaking, a clear way to distinguish between vocals and instrumental waves is by the presence of higher overtones, as while instrumental sound waves can produce overtones that go far beyond even the 10th overtone, human vocals tend to be limited to only having established overtone presence up to around the 4th overtone. This is certainly due to the mechanical nature of a string. With experience and certain skeletal structures, there are singers that can create overtones higher than the 4th overtone, however, it is unlikely for there to be singers that can produce as many overtones as those produced in the sound of stringed instruments. The human body is not perfect and neither a string thus a string made to specifically produce a single note is obviously going to be more accurate than a person's voice.

Vibrato for instrumental samples was vastly more narrow and centered more tightly on frequencies that the stringed instrument notes occurred on (i.e. violin recordings showed tight peaks with evenly distributed sides around the frequencies of the notes G3, D4, A4, and E5). There is also a smooth slope that is found in stringed instrument vibrato, which is most likely due to the mechanical nature of stringed instruments. The vibrato in a stringed instrument can be characterized by the narrow frequency range that surrounds the notes playable by the stringed instrument, whereas a lack of vibrato can be characterized by an incredibly narrow frequency range that approximates to a straight vertical line on the Fourier spectra plots

Analysis of the Fourier spectra plots for human vocals showed that vibrato for vocals could be characterized by a vastly large range of frequencies around the fundamental and overtone frequencies of notes that are singable by humans. When singing with vibrato, the frequency can span an extremely large range with several peaks, which can easily be distinguished from the tight range of frequencies that can be found within stringed instruments.

Note	Non-Vibrato Amplitude (Hz)		Vibrato Amplitude (Hz)	
	Voice	Cello	Voice	Cello
G2	22.5	21.0	28.0	21.0
D3	26.5	18.5	28.0	24.0
E4	26.0	22.5	34.5	24.0

Table 1: Frequency amplitudes of notes rounded to the nearest half-integer frequency.

5.2 Spectrograms

While our Fourier spectra mainly provided qualitative analysis, using spectrograms for our samples provided a way to quantitatively analyze vibrato in our samples. That is, it is quite easy to identify the exact difference in variance between vibrato in spectrograms through the comparison of amplitudes, periods, and from the overall sinusoidal shape of the audio sample. For example when comparing the note E4, which has a fundamental frequency of 330Hz, from both the cello spectrogram and the singing spectrogram, we can see that the cello varies from about 310Hz to just below 355Hz while the singing varies from around 310Hz to 380Hz. We can also see that the frequency of the cello modulates with a period of 0.25 seconds, while the singer modulated with a period of 0.178 seconds. While the difference in frequency modulation is similar across all three notes and while the period of modulation was the same between the cello and voice for both G2 and D3, We see a difference in modulation period for E4.

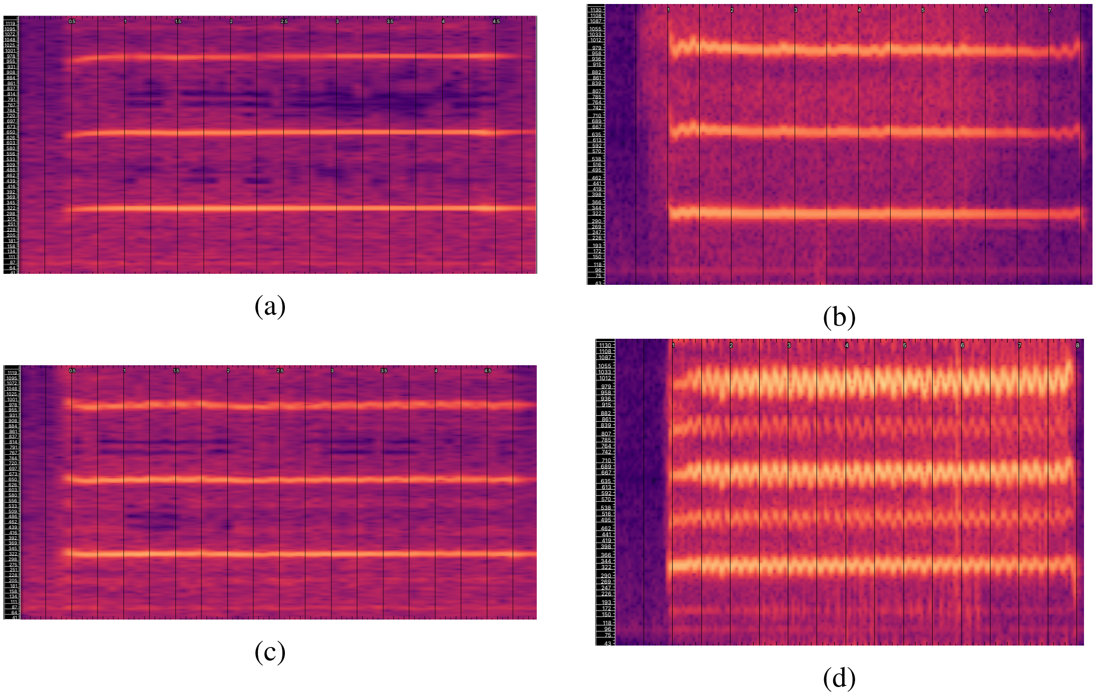

(a)
(b)
(c)
(d)

Figure 5: Figures a) and c) are spectrograms for a cello playing an E4 note without vibrato and with vibrato, respectively. Figures b) and d) are spectrograms for a vocalist singing an E4 note without vibrato and with vibrato, respectively. The y-axis is frequency in Hz and the x-axis is time in seconds, while brightness indicates intensity.

Further examination of the spectrograms for instrumental samples showed that vibrato was characterized as a varying in the frequency around each of the main notes and respective overtones (similarly shown by the Fourier spectra plots). Non-vibrato samples showed no variation in amplitude from the frequencies of the stringed instrument's playable notes.

The analysis of spectrograms in vocals provides a much more varied and chaotic heat map due to the large variation in frequency that human vocals produce in their vibrato. For vibrato in singing notes, there was a positive correlation between increasing overtones and increasing amplitude of the modulation. The fourth overtone is usually about double in amplitude of the fundamental frequency. This phenomena was not observed in non-vibrato samples of singing or any case of the stringed samples. While not exactly similar, non-vibrato in vocals is characterized similarly to non-vibrato in stringed instruments, as in vocals non-vibrato creates a narrow peak focused near a human singable note's frequency, but due to human imperfection there is often still a narrow but noticeable frequency range, or shift in frequency for any sung note.

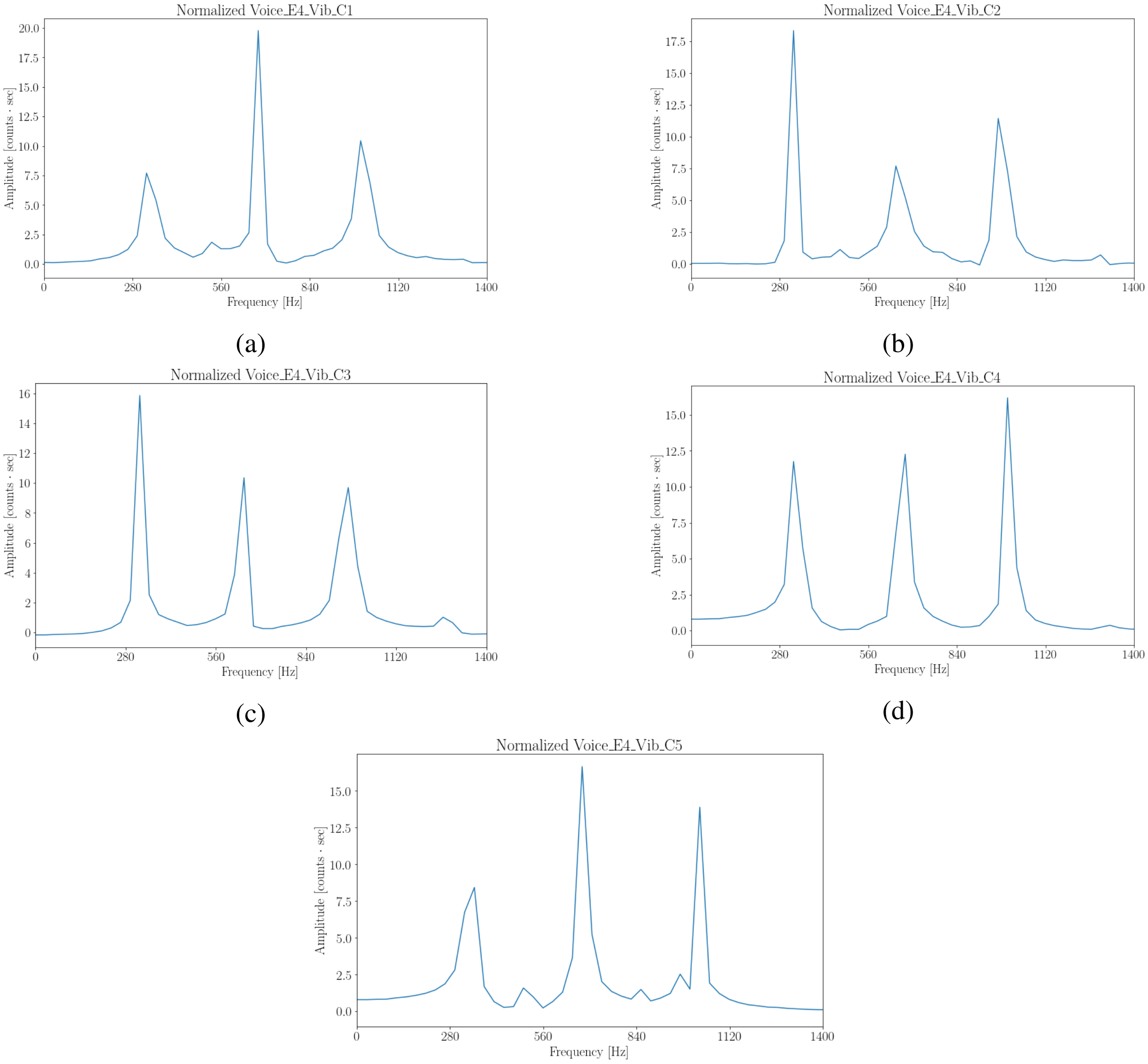
It is difficult to get a baseline from non-vibrato samples for a width using spectrograms. Spectrograms have a frequency and time resolution which were too large on our graphs to pick out useful information for non-vibrato samples. Frequency resolution is equal to Sample rate/Window size. We used a sample rate of 44,100hz and a window size of 4096, giving a frequency resolution of about 10.8Hz. While vocal non-vibrato seemed to have a wider width than stringed non-vibrato, it is difficult to determine exactly how much they varied due to our resolution.

From our spectrograms we have determined that vocal samples always have a wider width or peak than stringed samples for similar vibrato or non-vibrato cases. This difference was hard to extrapolate for non-vibrato samples, coming out to about a 4Hz difference on average. We say this with low confidence due to the frequency resolution of our spectrograms. For vibrato samples, the difference in amplitude comes out to 10Hz on average. This 10Hz difference is visually easy to distinguish when looking at the spectrograms. We have also determined that the period of fluctuation is usually around .25 seconds, regardless of how the vibrato is produced. There was a deviation in our data for our stringed instrument vibrato for E4, however, with a period of .36 seconds.

5.3 Fourier Transform Chunks

For the purpose of determining how vibrato changes on a small timescale, we collected small data slices of audio waves that we will refer to as ”chunks.”

We see that in the case that the audio is sliced into these small chunks of vibrato that there is an amplitude ”shifting” that occurs between the fundamental frequency and overtone frequencies so to speak. When examining the Fourier spectra of these chunks of vibrato, the relative amplitude between the fundamental frequency and the first and second overtone frequencies change through the progression of the chunks (i.e. from the graphs with end labels C1 to C5, there were slices taken of the same audio file that divided a single period of the vibrato wave into 5 chunks). The shifting of largest relative amplitude peaks is a strong indicator of the presence of vibrato, as in the examination of non-vibrato waves this shifting is mainly non-existent. It is also important to note that while the shifting mainly occurs in the fundamental frequency and first and second overtone frequency peaks, in some samples there can be seen minuscule changes in higher overtones as well. While this ”shifting” is an important indicator of vibrato, it appears that when examining the chunks of human vocals and string instruments, there does not appear to be any distinguishing features between the shifting in the two sets of chunks.



(a) (b) (c) (d) (e)

Figure 6: All figures (a) through (e) above represent human vocals singing the note E4 using vibrato. From figure (a) to figure (e) there is a relative progression in time such that figure (a) was a chunk of a single period of a vibrato wave that occurred just before the chunk that occurred to create figure (b) and so on.

6 Conclusion

The goal of this project was to determine a method of characterizing vibrato in human vocals and string instruments. From analyzing the data, it can be determined that not only is the vibrato in human vocals and string instruments detectable, but via the use of Fourier Transforms and spectrograms, there is a clear and distinct difference that separate human vocals, string instruments, and their respective vibrato characteristics. The Fourier Transform graphs can, firstly, show that human vocals and string instruments can be distinguished by the presence of higher overtone frequencies that populate the string instrument Fourier spectra. Any overtones past the 4th overtone in human vocals tends to be deformed in the sense that they are non-existent, have minuscule amplitude relative to the fundamental, or cover such a wide range of frequencies that they mix together with other overtones. Secondly, the Fourier Transform graphs show that, in general, the bases of peaks for the fundamental and overtone frequencies are very narrow and precise for string instruments. This means that, compared to singing, a smaller range of frequencies are reached around the fundamental frequency and overtone frequencies. In the examination of the vibrato examples, it is important to note that while both human vocals and string instrumental peaks have wide bases (due to the fact that they are both fluctuating), string instruments still fluctuate in frequency around each fundamental and overtone significantly less than with human vocals.

If we take a recording with vibrato and cut it into small clips and perform Fourier Transforms on each of them, we can see that our frequency is indeed oscillating and not just randomly fluctuating. We can quickly do this en mass by using spectrograms. When we preform spectrogram analysis on our recordings, we also saw a very distinct difference between vocal vibrato and stringed instrument vibrato, with the difference in amplitudes being anywhere from 17% larger to 44% larger for vocal vibrato than stringed vibrato. Visually, the difference between stringed vibrato and vocal vibrato is obvious; vocal vibrato recording are easily identifiable as sinusoidal waves while the oscillations in stringed vibrato are more difficult to identify. Due to our frequency resolution, it was more difficult to distinguish between non-vibrato samples using the width of the frequency. However, there was a visual difference.

There are circumstances where human vocal vibrato and stringed instrument vibrato may appear to be similar, such as in the case where data for an extremely talented and experienced singer is examined alongside the performance of a stringed instrument. However, if provided with enough data over time, there will be a clear variation between the two performances, as well as in any vibrato techniques via the examination of their respective Fourier spectra's relative peak widths and the presence of very high frequency overtones.

The importance of music is known throughout society, however, the physics behind many of musics common techniques are not. Through the analysis of the technique known as vibrato, the difference in vibrato that can be found in both human vocals and stringed instruments can be distinguished by analysis of their Fourier Transform graphs and spectrograms quite easily, and the shape of the graphs that are formed can be further analyzed to provide insight on the mechanical smoothness of instrumental vibrato as well as the fine-tuned and rare smoothness that is found within human vocal vibrato.

References

[1] *Discrete Fourier Transform (DFT)*. [Online]. 2017. URL: `https://web.archive.org/web/20171010173422/http://en.dsplib.org/content/dft.html#r6`.

[2] Victor Epand. “Why the Tremolo Arm is Misnamed”. In: (2008). [Online]. URL: `https://ezinearticles.com/?Why-the-Tremolo-Arm-is-Misnamed&id=1308145`.

[4] George Gollin. *32 kHz audio recording to binary file code*. 2018. URL: `https://courses.physics.illinois.edu/phys398DLP/sp2020/code.asp`.

[7] J. D. Hunter. “Matplotlib: A 2D graphics environment”. In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: `10.1109/MCSE.2007.55`.

[8] Eric Jones, Travis Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. [Online]. 2001–. URL: `http://www.scipy.org/`.

[9] Danya Katok. “The versatile singer: a guide to vibrato & straight tone”. In: (2016).

[10] Karyn O’Connor. *Singwise - False Vibratos*. [Online]. URL: `http://singwise.com/cgi-bin/main.pl?section=vocalfaqs&doc=FakeVibratos`.

[11] Travis Oliphant. *NumPy: A guide to NumPy*. USA: Trelgol Publishing. [Online]. 2006–. URL: `http://www.numpy.org/`.

[13] L. Potter. *The Art of Cello Playing: A Complete Textbook-method for Private Or Class Instruction*. Art of. Summy-Birchard Music, 1980. ISBN: 9780874870718.

[14] Casey Reas and Ben Fry. *Processing: A Programming Handbook for Visual Designers and Artists*. 2nd ed. Extension 3. The MIT Press, 2014. ISBN: 9780262028288. URL: `https://processing.org/tutorials/sound/`.

[15] L. Regnier and G. Peeters. “Singing voice detection in music tracks using direct voice vibrato detection”. In: *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2009, pp. 1685–1688.

[16] Stanley Sadie and John Tyrrell. *The new Grove dictionary of music and musicians*. Grove, 2002.

[17] *Sonic Visualiser*. [Online]. URL: `https://www.sonicvisualiser.org/index.html`.

[18] *Vibrato*. [Online]. 2020. URL: `https://github.com/sergicastells9/Vibrato`.

External Figures

[3] *Fourier Transform of Square Function*. 2010. URL: `thefouriertransform.com`.

[5] George Gollin. *PCB Data Logger Schematic*. 2019. URL: `https://courses.physics.illinois.edu/phys398DLP/sp2020/code.asp`.

[6] Mark Huckvale. *Vocal Cords*. University College London, 2019. URL: `https://www.phon.ucl.ac.uk/courses/plin0065/week4.php`.

[12] *Piano Keys*. 2012. URL: `http://musictheoryfundamentals.com/MusicTheory/intervals_part1.php`.

[19] Audrey Williams. *Cello*. 2012. URL: `https://audreywilliamsmusic.wordpress.com/tag/cello-diagram/`.

A The Fourier Transform

The Fourier Transform of a function, $g(t)$, is given by:

$$\mathscr{F}\left\{g(t)\right\} = G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt \tag{1}$$

We can see that this transforms the function from the time domain to the frequency domain. $G(f)$ is called the spectrum of g. The inverse Fourier Transform allows us to obtain g from G.

$$\mathscr{F}^{-1}\left\{G(f)\right\} = \int_{-\infty}^{\infty} G(f)e^{2\pi i f t}df = g(t) \tag{2}$$

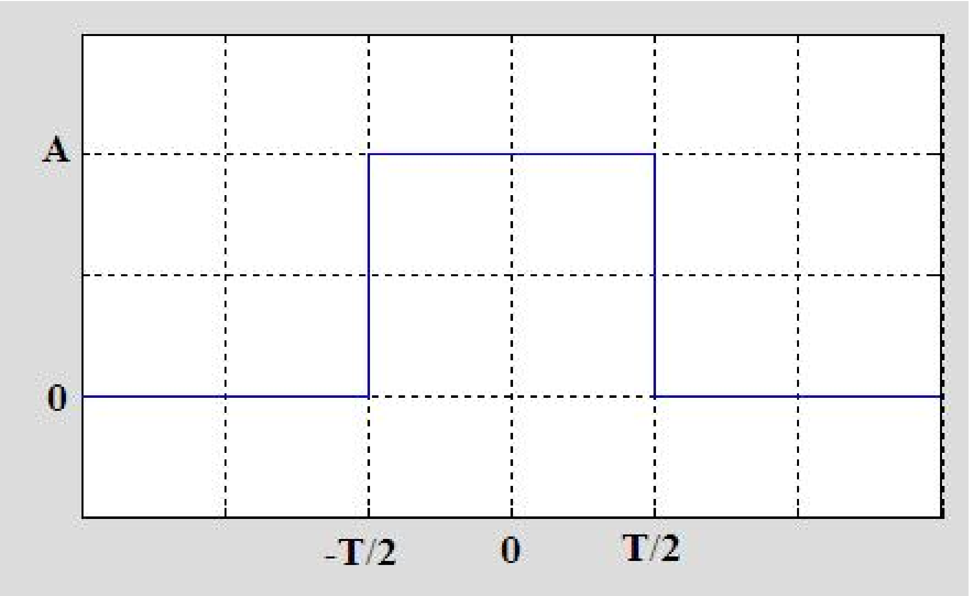


Figure A.1: A square wave function of amplitude A and width T [3].

If we analyze the square wave shown in Figure A.1 and evaluate the Fourier Transform we obtain:

$$\begin{aligned}\mathscr{F}\left\{g(t)\right\} = G(f) &= \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt \\ &= \int_{-T/2}^{T/2} Ae^{-2\pi i f t}dt \\ &= \frac{AT}{\pi f T}\left[\frac{e^{i\pi f T} - e^{-i\pi f T}}{2i}\right] \\ &= \frac{AT}{\pi f T}\sin(\pi f T) \\ &= AT\text{sinc}(fT)\end{aligned} \tag{3}$$

where we define the $sinc$ function as $sinc(t) = \frac{sin(\pi t)}{\pi t}$.

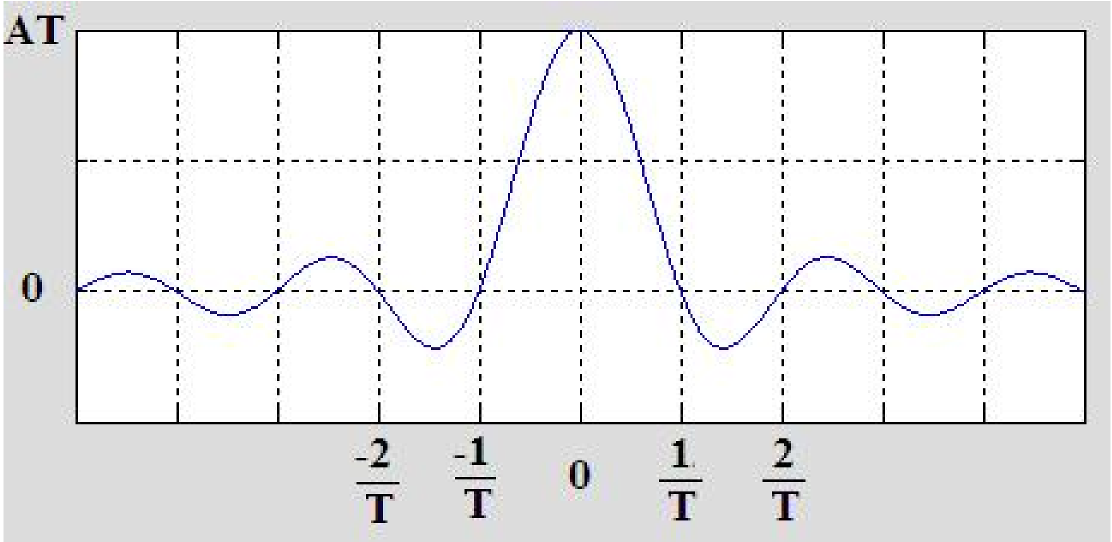


Figure A.2: The Fourier Transform of a square wave function of amplitude A and width T [3].

This leaves us with $G(f) = AT\mathrm{sinc}(\pi fT)$, the frequency domain equivalent of the square wave, which is plotted in Figure A.2.
If we compute $\mathscr{F}^{-1}\left\{G(f)\right\}$, we obtain

$$\begin{aligned}\mathscr{F}^{-1}\left\{g(t)\right\} &= \int_{-\infty}^{\infty} G(f)e^{2\pi ift}df \\ &= \int_{-\infty}^{\infty} AT\mathrm{sinc}(fT)e^{2\pi ift}df \\ &= A\,\mathrm{rect}\left(\frac{T}{2}\right)\end{aligned} \tag{4}$$

The inverse Fourier Transform of the $sinc$ function is defined as the $rect$ function. Thus we have precisely what we expect as an output. The result is the square wave from Figure A.1.

Now if we look at a periodic square wave, as seen in Figure A.3, we see the effect of a Fourier Transform on a periodic function.

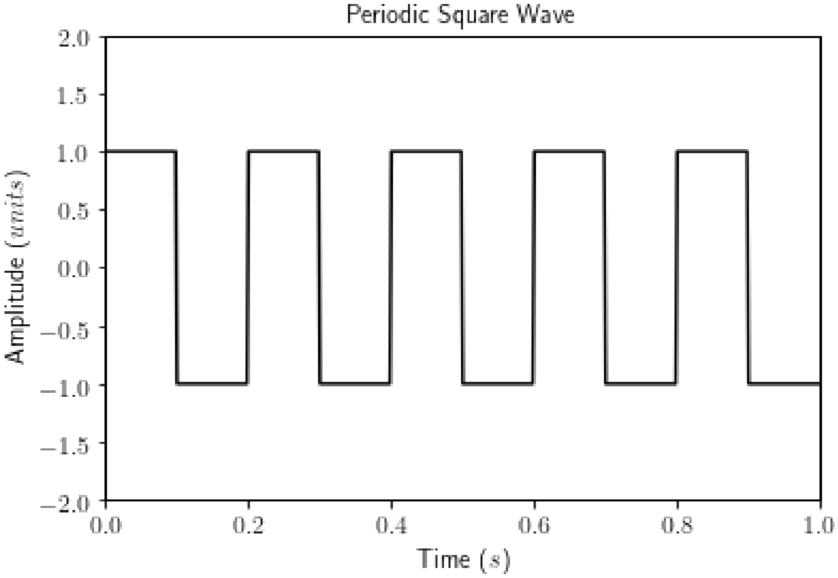


Figure A.3: A periodic square wave with an amplitude of 2 units and a width 0.1 seconds.

The Fourier series representation of the periodic square wave function is given by:

$$f(t) = \sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n\pi} \sin\left(\frac{n\pi t}{0.1}\right) \tag{5}$$

Taking the Fourier Transform of Eqn. 5,

$$\begin{aligned}
\mathscr{F}\left\{f(t)\right\} = F(f) &= \int_{-\infty}^{\infty} f(t) e^{-2\pi i f t} dt \\
&= \int_{-\infty}^{\infty} \left[\sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n\pi} \sin\left(\frac{n\pi t}{0.1}\right)\right] e^{-2\pi i f t} dt \\
&= \sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n\pi} \int_{-\infty}^{\infty} \sin\left(\frac{n\pi t}{0.1}\right) e^{-2\pi i f t} dt \\
&= \sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n\pi} \int_{-\infty}^{\infty} \sin\left(\frac{n\pi t}{0.1}\right) \big[\cos(2\pi f t) - \sin(2\pi f t)\big] dt \\
&= \sum_{n=1,3,5,\ldots}^{\infty} \left[\frac{4}{n\pi} \int_{-\infty}^{\infty} \sin\left(\frac{n\pi t}{0.1}\right) \cos(2\pi f t) - \frac{4}{n\pi} \int_{-\infty}^{\infty} \sin\left(\frac{n\pi t}{0.1}\right) \sin(2\pi f t)\big] dt\right]
\end{aligned} \tag{6}$$

While it is possible to solve this analytically, it is tedious. It requires a similar computation as in Eqn. 3 over an infinite sum. Finding a solution is aided by solving it numerically. The Fourier Transform is shown in Fig. A.4.

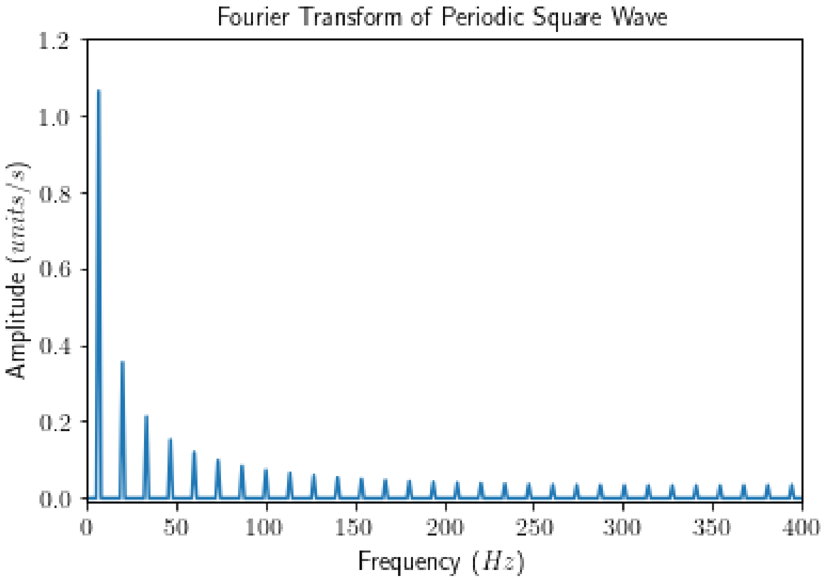


Figure A.4: The Fourier Transform of a periodic square wave function with an amplitude of 2 units and a width 0.1 seconds.

We see peaks in the Fourier spectrum of the periodic square wave function at multiples of the fundamental frequency. The fundamental frequency is defined by $T = \frac{1}{f_0} \implies f_0 = \frac{1}{0.1s} = 5Hz$. The subsequent harmonics are odd integer multiples of the fundamental frequency (the square wave only has odd harmonics). The first few harmonics are 5Hz, 15Hz, 25Hz, and so on.

B Data Tables

All data is presented below in tables. They are arranged by source and in consistent order of notes.

G2 Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	98	196	294	392	490
Vocal High (Hz)	133.3	240.6	343	448	547
Vocal Low (Hz)	77.4	168.5	266	352	452
Vocal Amplitude (Hz)	27.95	36.05	38.5	48	47.5
Vocal Freq (Hz)	105.35	204.55	304.5	400	499.5
Instrumental High (Hz)	123.81	226.1	323	420	509
Instrumental Low (Hz)	82.47	179.37	276.27	375	473
Instrumental Amplitude (Hz)	20.67	23.365	23.365	22.5	18
Instrumental Freq (Hz)	103.14	202.735	299.635	397.5	491
Amplitude Percent Diff	135.2201258	154.2906056	164.7763749	213.3333333	263.8888889

Table A.1: Cello & Voice – G2 Vibrato Data – Source 1

D3 Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	146.83	293.66	440.49	587.32	734.15
Vocal High (Hz)	182.6	341.1	503	654.6	813.1
Vocal Low (Hz)	127.5	268.7	415	546.1	687.3
Vocal Amplitude (Hz)	27.55	36.2	44	54.25	62.9
Vocal Freq (Hz)	155.05	304.9	459	600.35	750.2
Instrumental High (Hz)	176.9	328.8	473.9	617.2	770.2
Instrumental Low (Hz)	129.7	276.8	420.3	567.1	718.1
Instrumental Amplitude (Hz)	23.6	26	26.8	25.05	26.05
Instrumental Freq (Hz)	153.3	302.8	447.1	592.15	744.15
Amplitude Percent Diff	116.7372881	139.2307692	164.1791045	216.5668663	241.4587332

Table A.2: Cello & Voice – D3 Vibrato Data – Source 1

E4 Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	329.63	659.26	988.89	1318.52	1648.15
Vocal High (Hz)	377	722	1074	1405	1742
Vocal Low (Hz)	308	617	930.6	1268	1588.5
Vocal Amplitude (Hz)	34.5	52.5	71.7	68.5	76.75
Vocal Freq (Hz)	342.5	669.5	1002.3	1336.5	1665.25
Instrumental High (Hz)	356	692	1017	1340	1672
Instrumental Low (Hz)	308	632	954	1288	1609.5
Instrumental Amplitude (Hz)	24	30	31.5	26	31.25
Instrumental Freq (Hz)	332	662	985.5	1314	1640.75
Amplitude Percent Diff	143.75	175	227.6190476	263.4615385	245.6

Table A.3: Cello & Voice – E4 Vibrato Data – Source 1

G2 Non-Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	98	196	294	392	490
Vocal High (Hz)	123	219	327	422	519
Vocal Low (Hz)	78	176	274	373	472
Vocal Amplitude (Hz)	22.5	21.5	26.5	24.5	23.5
Vocal Freq (Hz)	100.5	197.5	300.5	397.5	495.5
Instrumental High (Hz)	123	219	318	407	509
Instrumental Low (Hz)	81	180	282	377	482
Instrumental Amplitude (Hz)	21	19.5	18	15	13.5
Instrumental Freq (Hz)	102	199.5	300	392	495.5
Amplitude Percent Diff	107.1428571	110.2564103	147.2222222	163.3333333	174.0740741

Table A.4: Cello & Voice – G2 Non-Vibrato Data – Source 1

D3 Non-Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	146.83	293.66	440.49	587.32	734.15
Vocal High (Hz)	172	316	459	603	746
Vocal Low (Hz)	119	266	410	558	698
Vocal Amplitude (Hz)	26.5	25	24.5	22.5	24
Vocal Freq (Hz)	145.5	291	434.5	580.5	722
Instrumental High (Hz)	172	314	463	600	755
Instrumental Low (Hz)	135	277	430	578	721
Instrumental Amplitude (Hz)	18.5	18.5	16.5	11	17
Instrumental Freq (Hz)	153.5	295.5	446.5	589	738
Amplitude Percent Diff	143.2432432	135.1351351	148.4848485	204.5454545	141.1764706

Table A.5: Cello & Voice – D3 Non-Vibrato Data – Source 1

E4 Non-Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	329.63	659.26	988.89	1318.52	1648.15
Vocal High (Hz)	352	682	1002	1336	1653
Vocal Low (Hz)	300	617	930	1268	1598
Vocal Amplitude (Hz)	26	32.5	36	34	27.5
Vocal Freq (Hz)	326	649.5	966	1302	1625.5
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.6: Cello & Voice – E4 Non-Vibrato Data – Source 1

G2 Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	98	196	294	392	490
Instrumental High (Hz)	121.2	227	320	423	520
Instrumental Low (Hz)	84.7	177	277	379	470
Instrumental Amplitude (Hz)	18.25	25	21.5	22	25
Instrumental Freq (Hz)	102.95	202	298.5	401	495
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.7: Cello – G2 Vibrato Data – Source 2

D3 Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	146.83	293.66	440.49	587.32	734.15
Instrumental High (Hz)	175	326	470	618	765
Instrumental Low (Hz)	129	274	416	560	711
Instrumental Amplitude (Hz)	23	26	27	29	27
Instrumental Freq (Hz)	152	300	443	589	738
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.8: Cello – D3 Vibrato Data – Source 2

E4 Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	329.63	659.26	988.89	1318.52	1648.15
Instrumental High (Hz)	360	688	1007	1355	1674
Instrumental Low (Hz)	303	628	953	1278	1603
Instrumental Amplitude (Hz)	28.5	30	27	38.5	35.5
Instrumental Freq (Hz)	331.5	658	980	1316.5	1638.5
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.9: Cello – E4 Vibrato Data – Source 2

G2 Non-Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	98	196	294	392	490
Instrumental High (Hz)	122	224	323	419	512
Instrumental Low (Hz)	82	180	277	373	477
Instrumental Amplitude (Hz)	20	22	23	23	17.5
Instrumental Freq (Hz)	102	202	300	396	494.5
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.10: Cello – G2 Non-Vibrato Data – Source 2

D3 Non-Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	146.83	293.66	440.49	587.32	734.15
Instrumental High (Hz)	171	313	460	611	754
Instrumental Low (Hz)	135	276	432	571	717
Instrumental Amplitude (Hz)	18	18.5	14	20	18.5
Instrumental Freq (Hz)	153	294.5	446	591	735.5
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.11: Cello – D3 Non-Vibrato Data – Source 2

E4 Non-Vibrato	Fundamental	First Overtone	Second Overtone	Third Overtone	Fourth Overtone
Frequency (Hz)	329.63	659.26	988.89	1318.52	1648.15
Instrumental High (Hz)	356	677	1008	1333	1657
Instrumental Low (Hz)	313	641	971	1299	1624
Instrumental Amplitude (Hz)	21.5	18	18.5	17	16.5
Instrumental Freq (Hz)	334.5	659	989.5	1316	1640.5
Instrumental High (Hz)	350	679	1011	1336	1661
Instrumental Low (Hz)	305	641	963	1298	1627
Instrumental Amplitude (Hz)	22.5	19	24	19	17
Instrumental Freq (Hz)	327.5	660	987	1317	1644
Amplitude Percent Diff	115.5555556	171.0526316	150	178.9473684	161.7647059

Table A.12: Cello – E4 Non-Vibrato Data – Source 2

C Figures

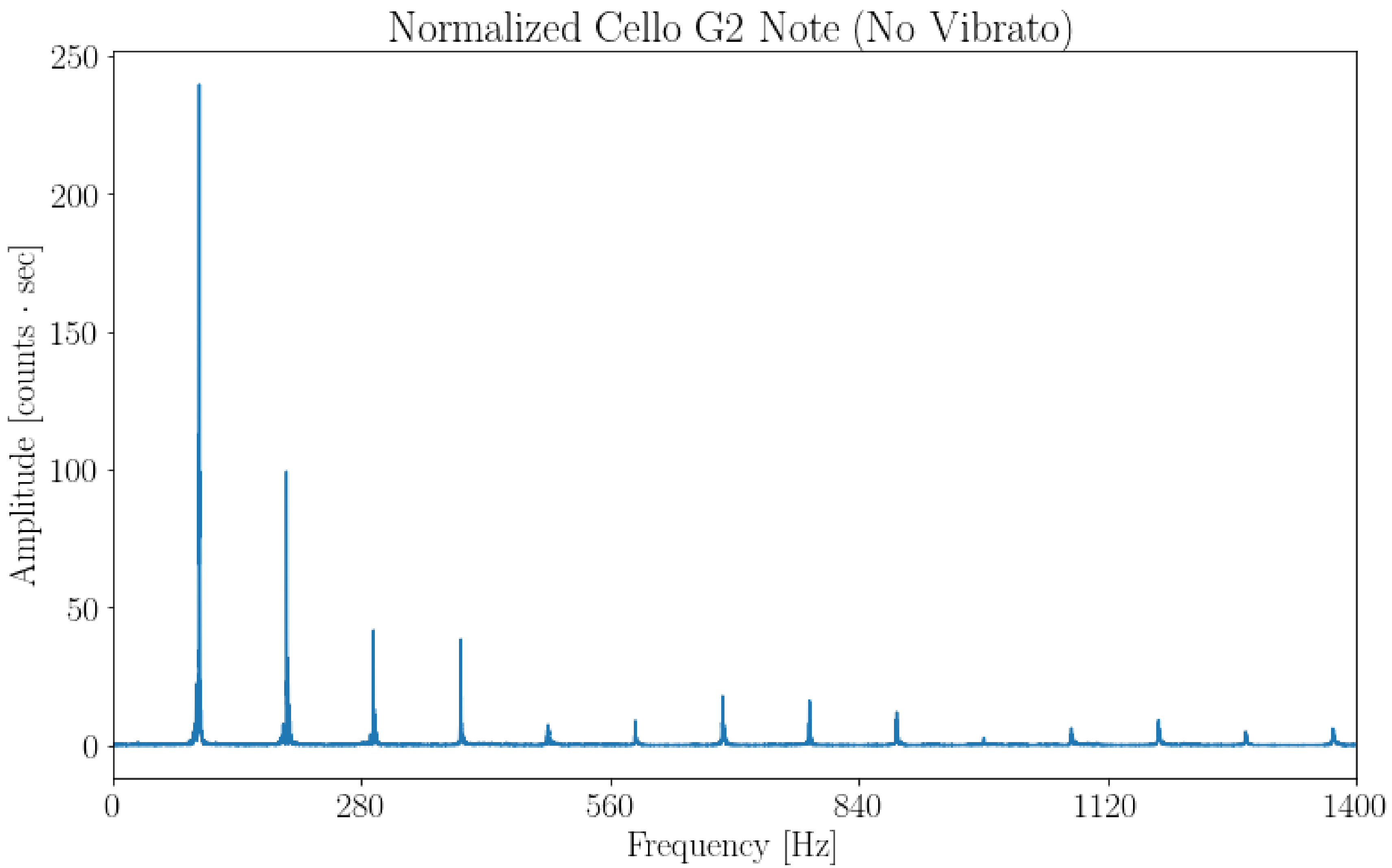


Figure A.1: Cello – G2 No Vibrato – Source 1

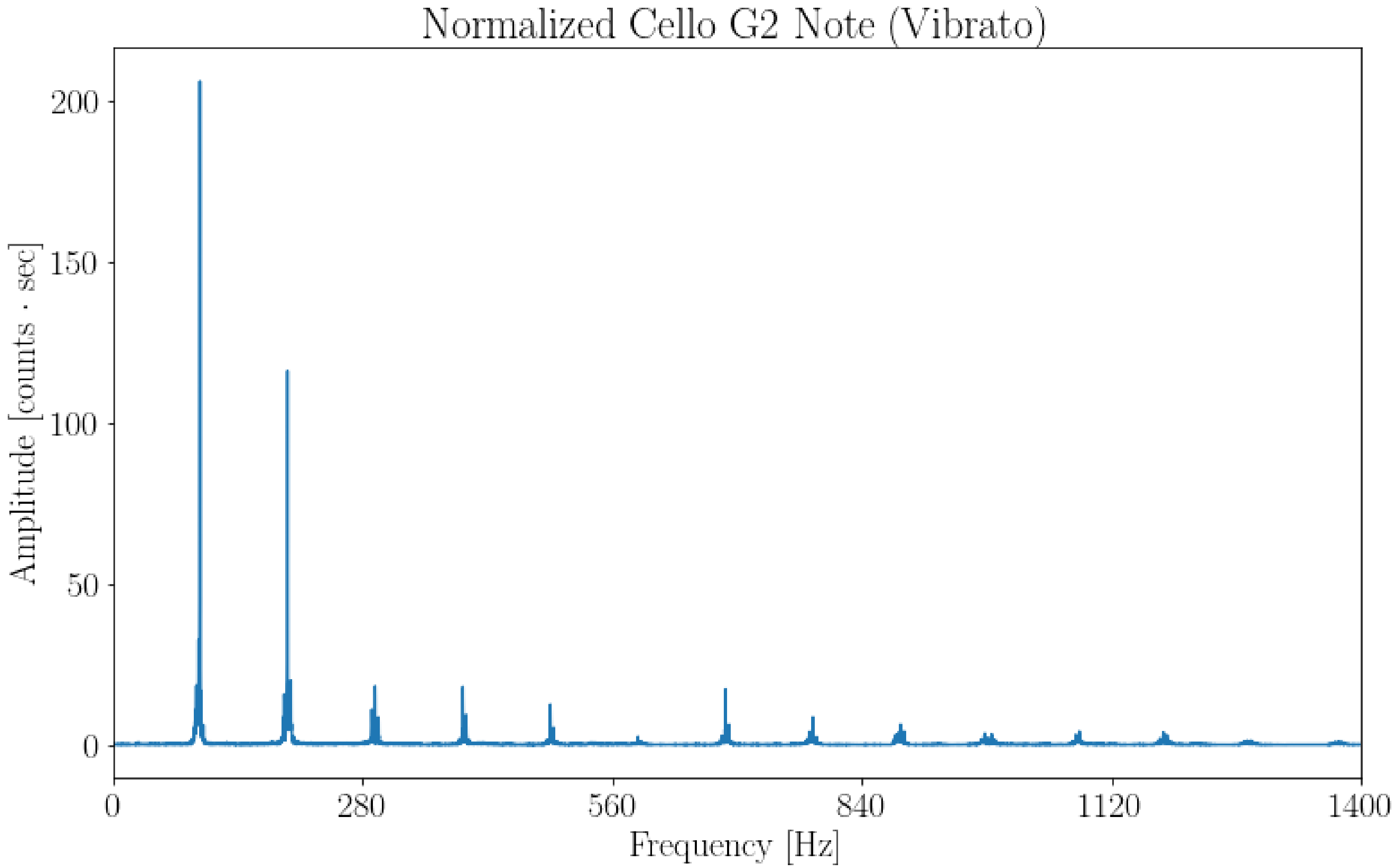


Figure A.2: Cello – G2 Vibrato – Source 1

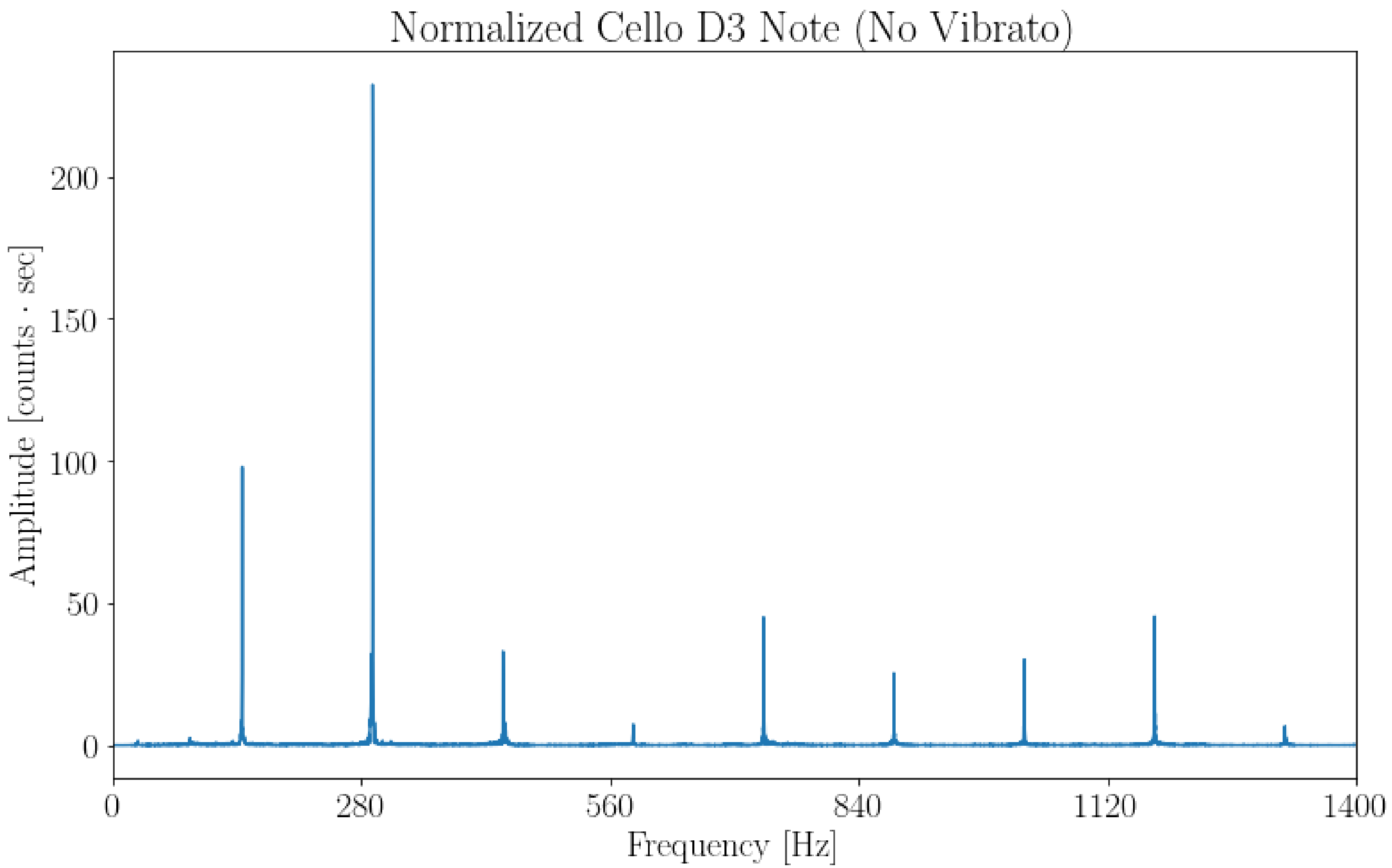


Figure A.3: Cello – D3 No Vibrato – Source 1

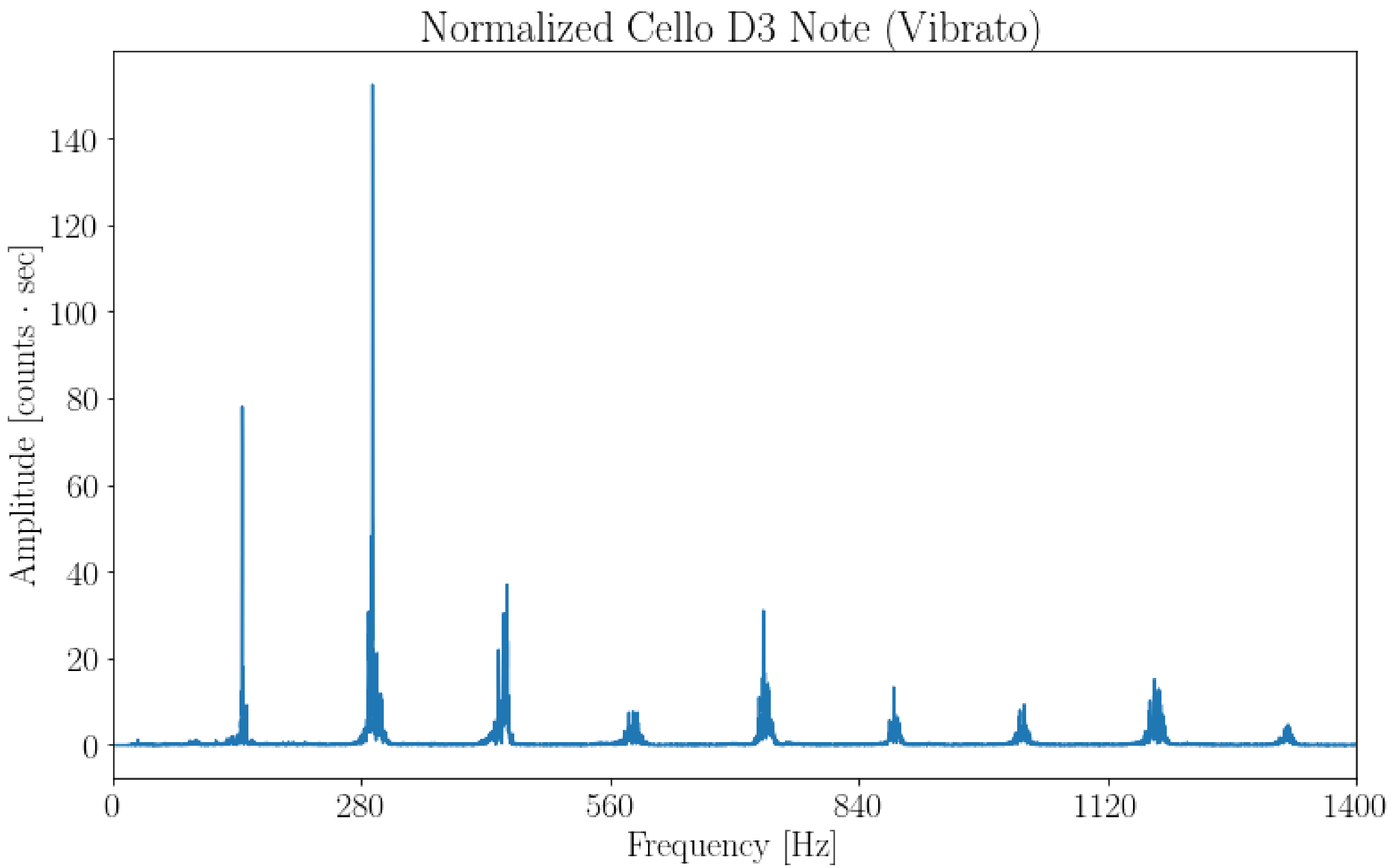


Figure A.4: Cello – D3 Vibrato – Source 1

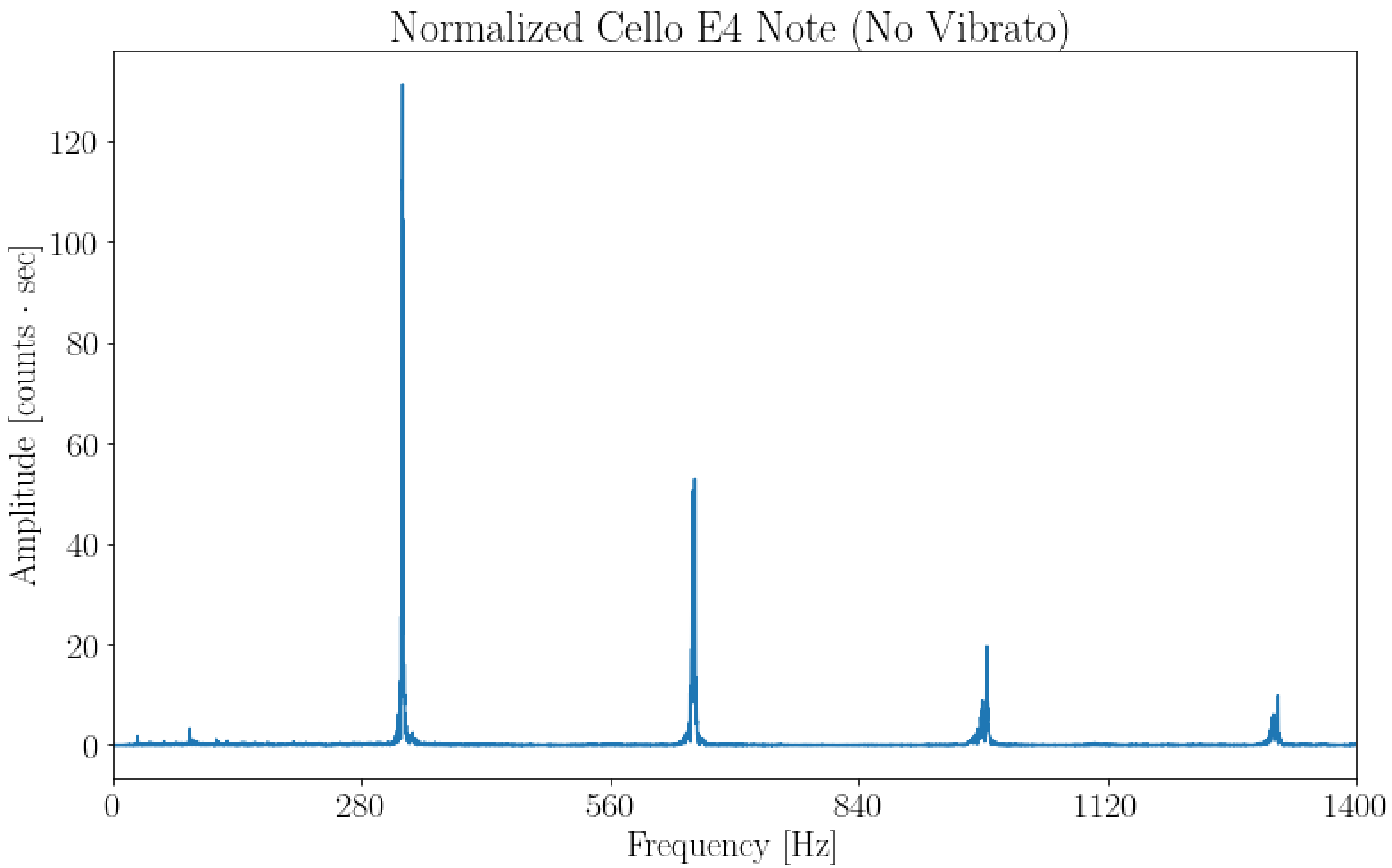


Figure A.5: Cello – E4 No Vibrato – Source 1

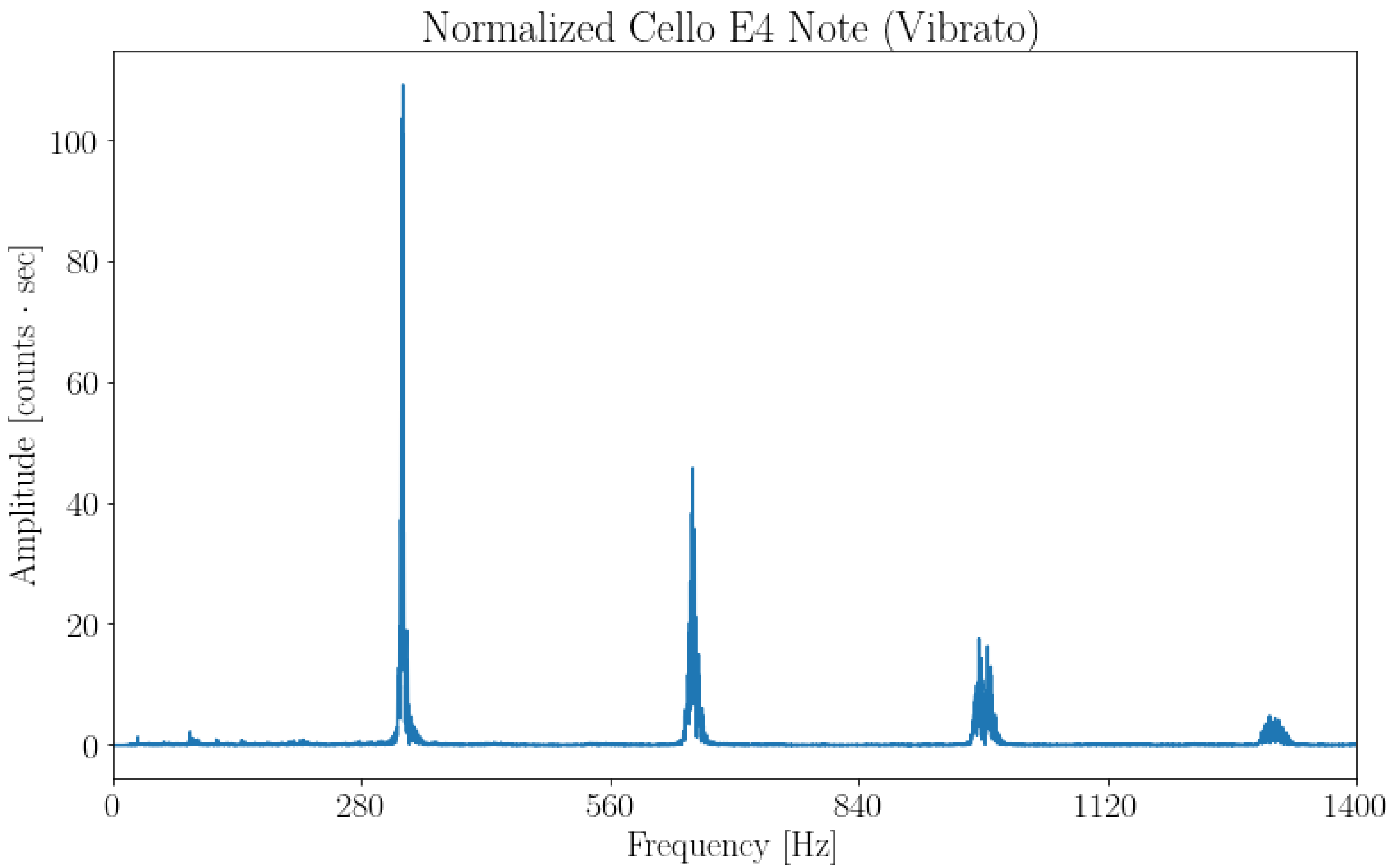


Figure A.6: Cello – E4 Vibrato – Source 1

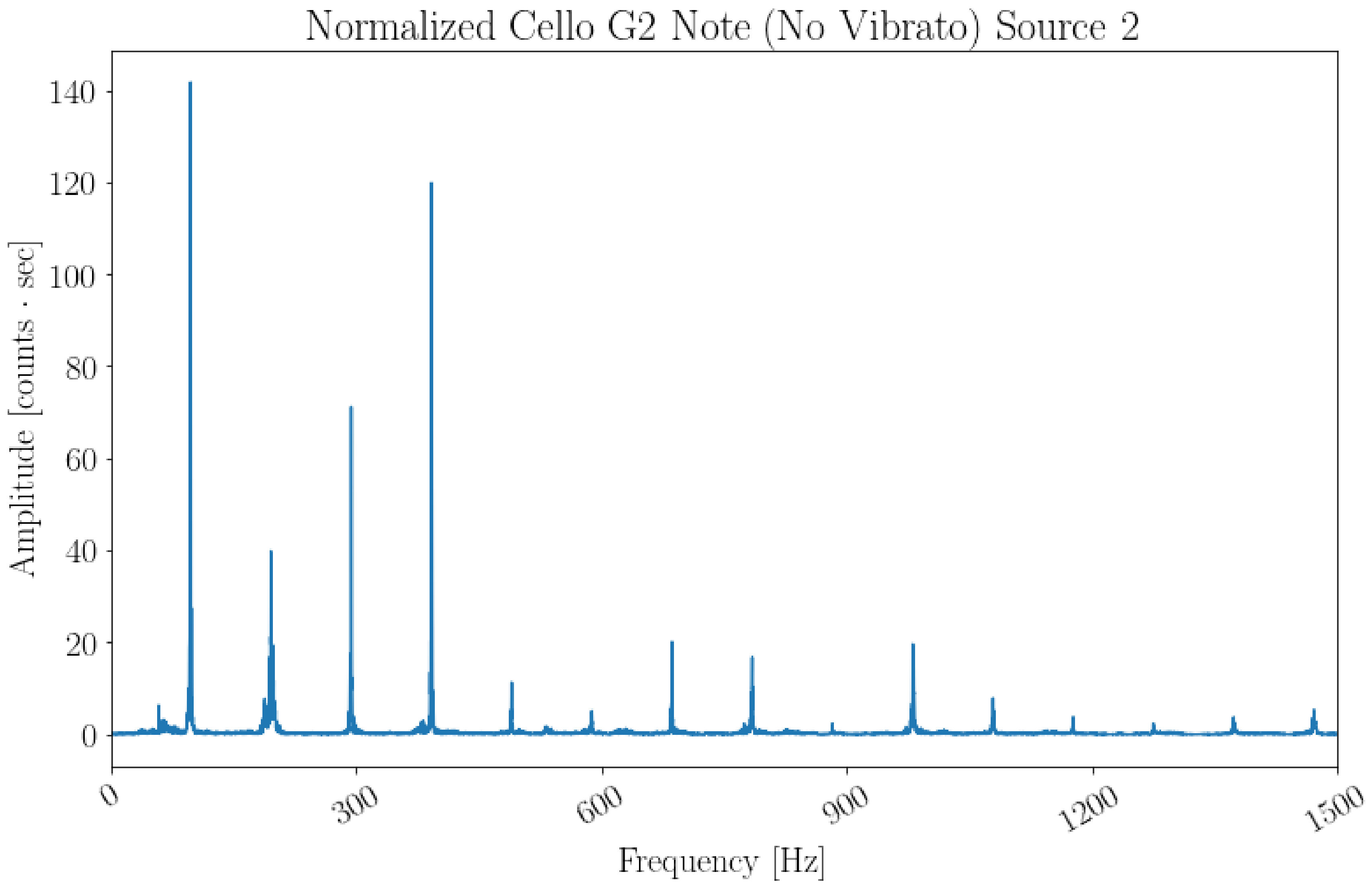


Figure A.7: Cello – G2 No Vibrato – Source 2

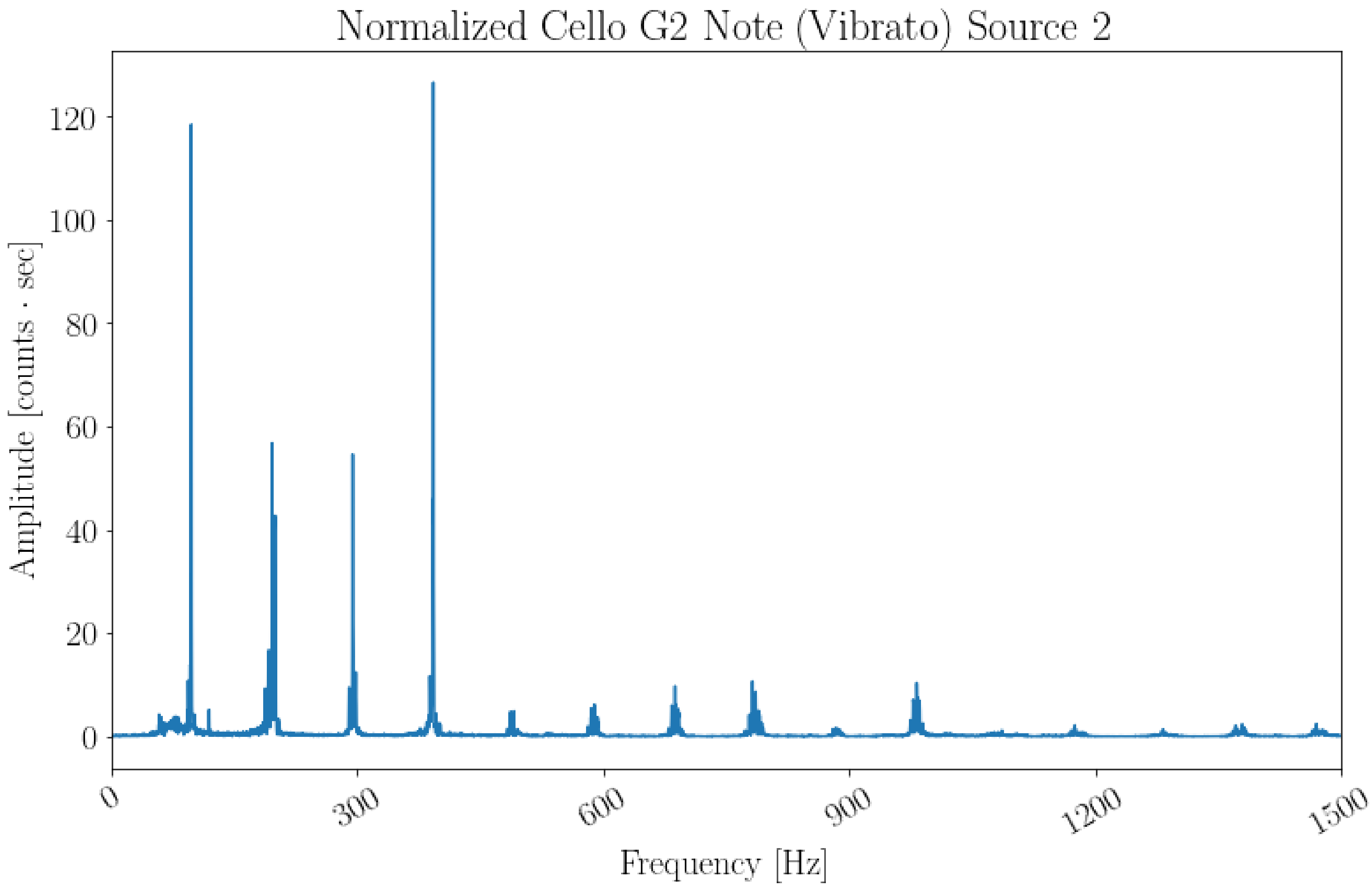


Figure A.8: Cello – G2 Vibrato – Source 2

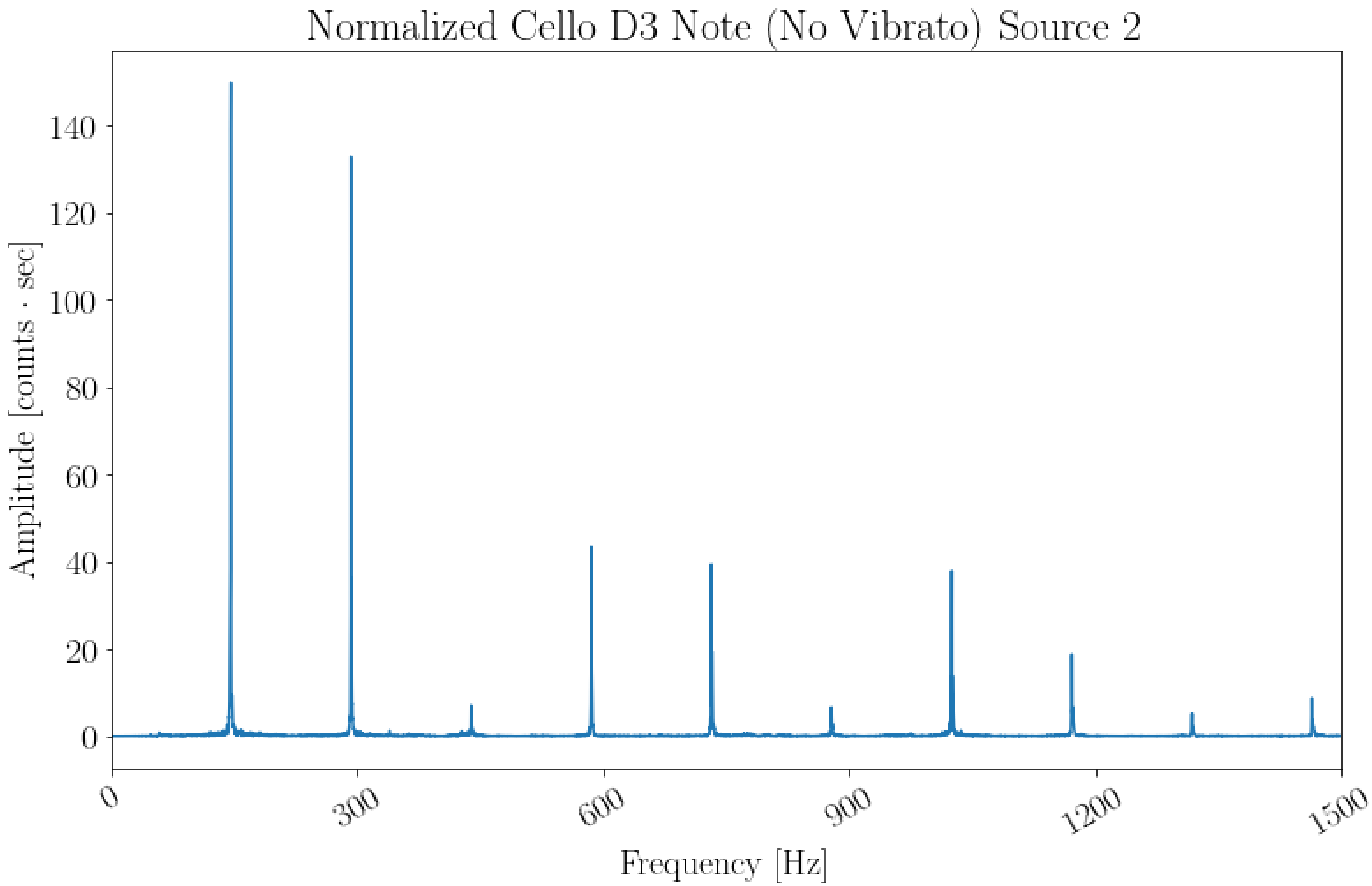


Figure A.9: Cello – D3 No Vibrato – Source 2

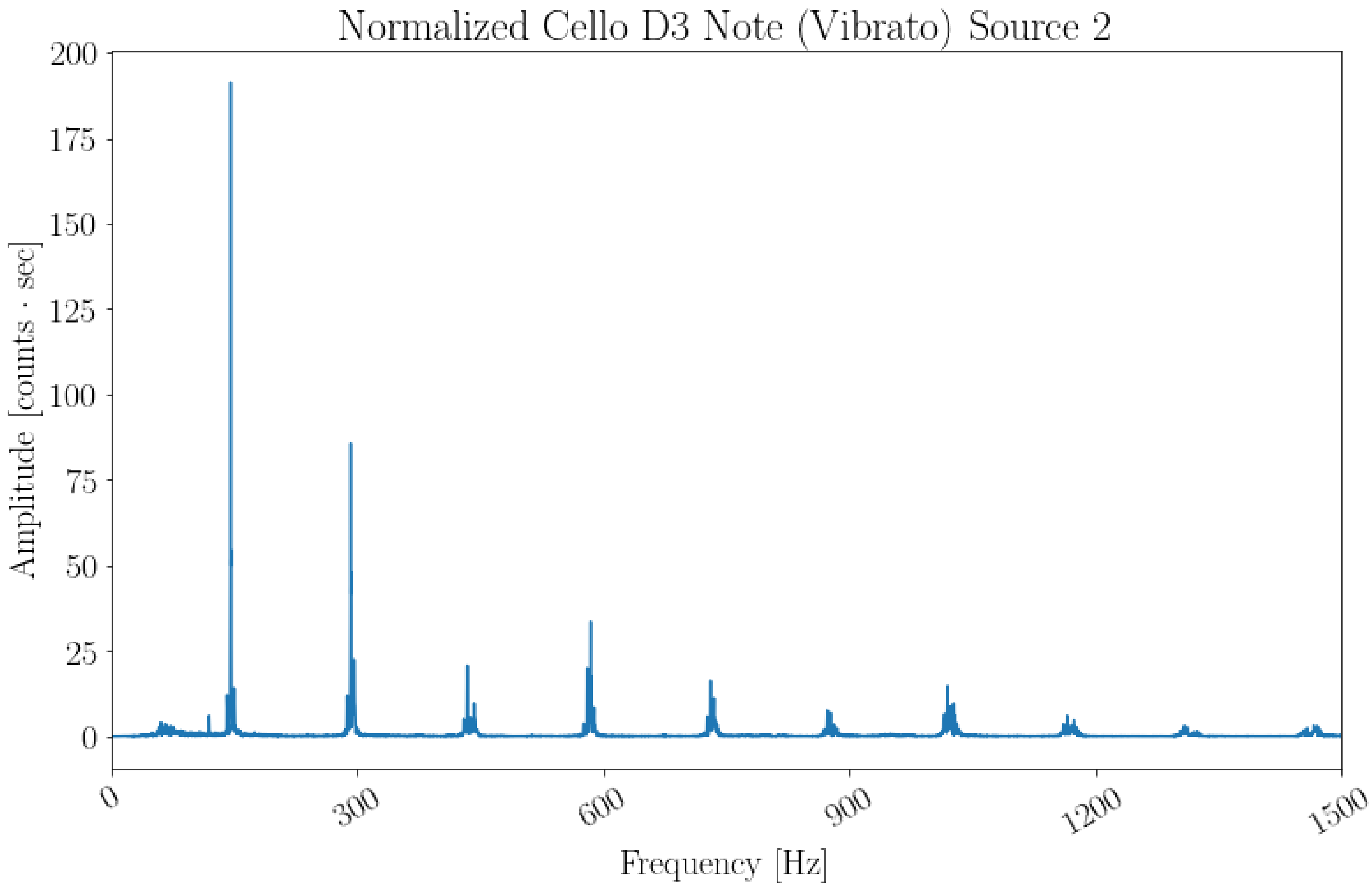


Figure A.10: Cello – D3 Vibrato – Source 2

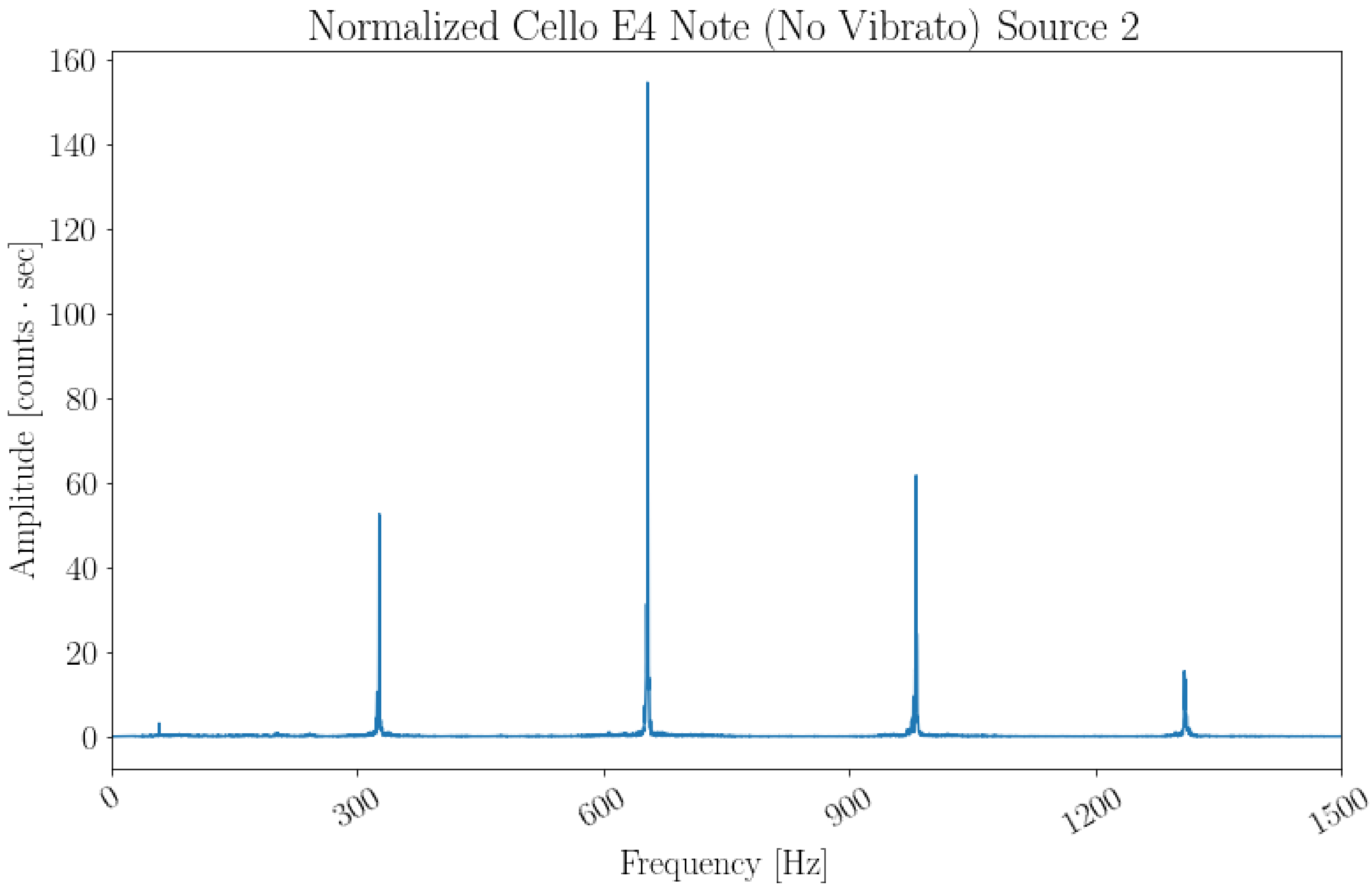


Figure A.11: Cello – E4 No Vibrato – Source 2

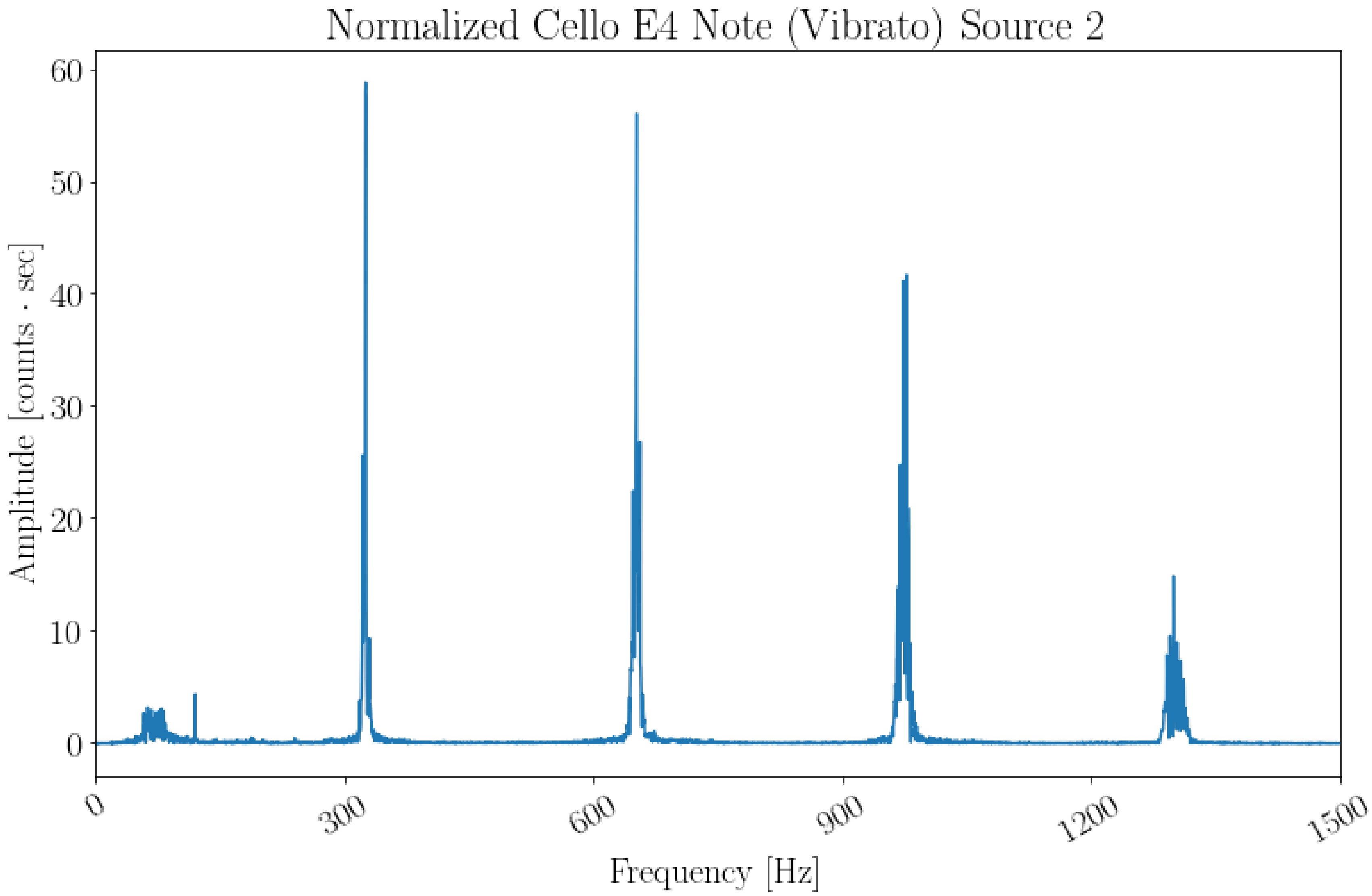


Figure A.12: Cello – E4 Vibrato – Source 2

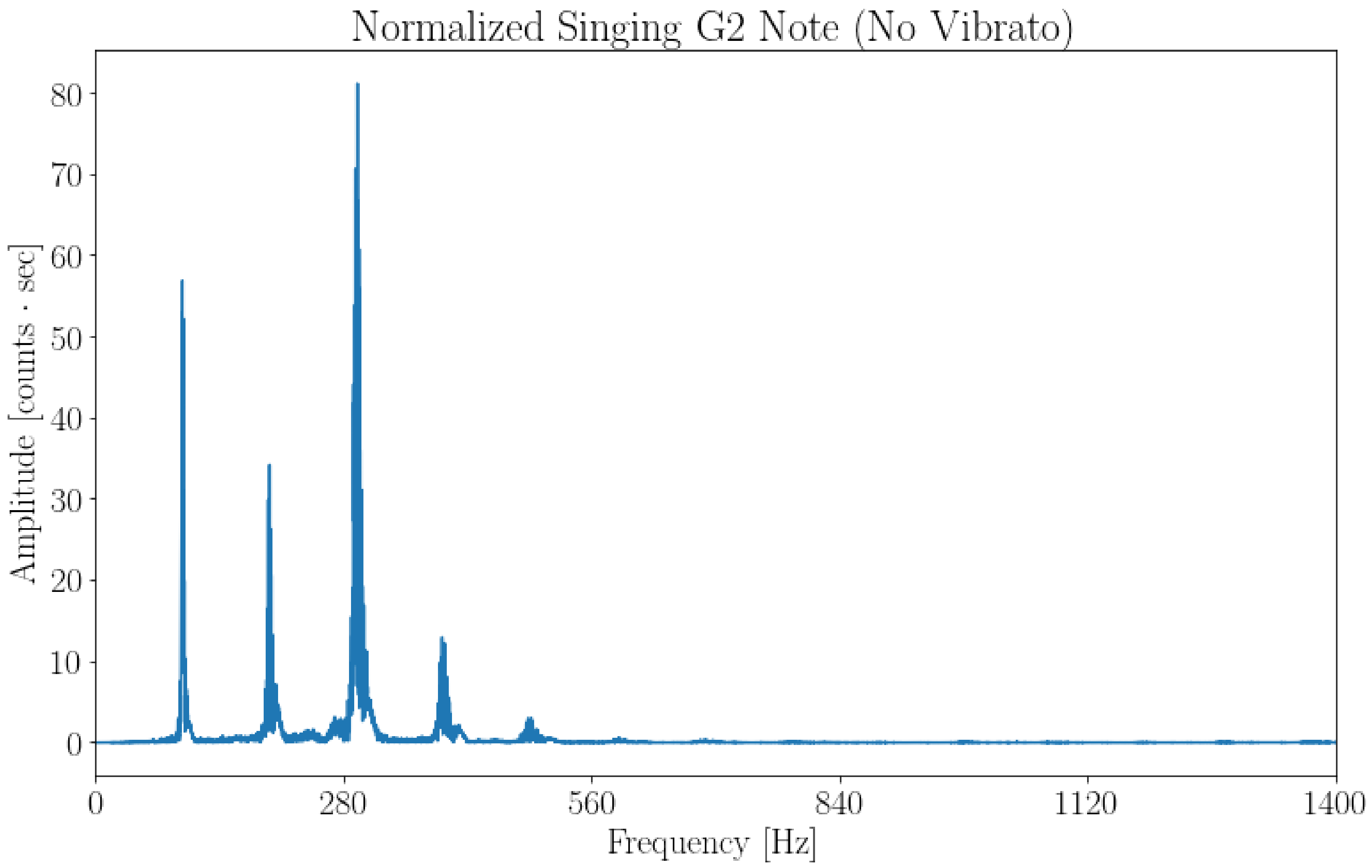


Figure A.13: Voice – G2 No Vibrato

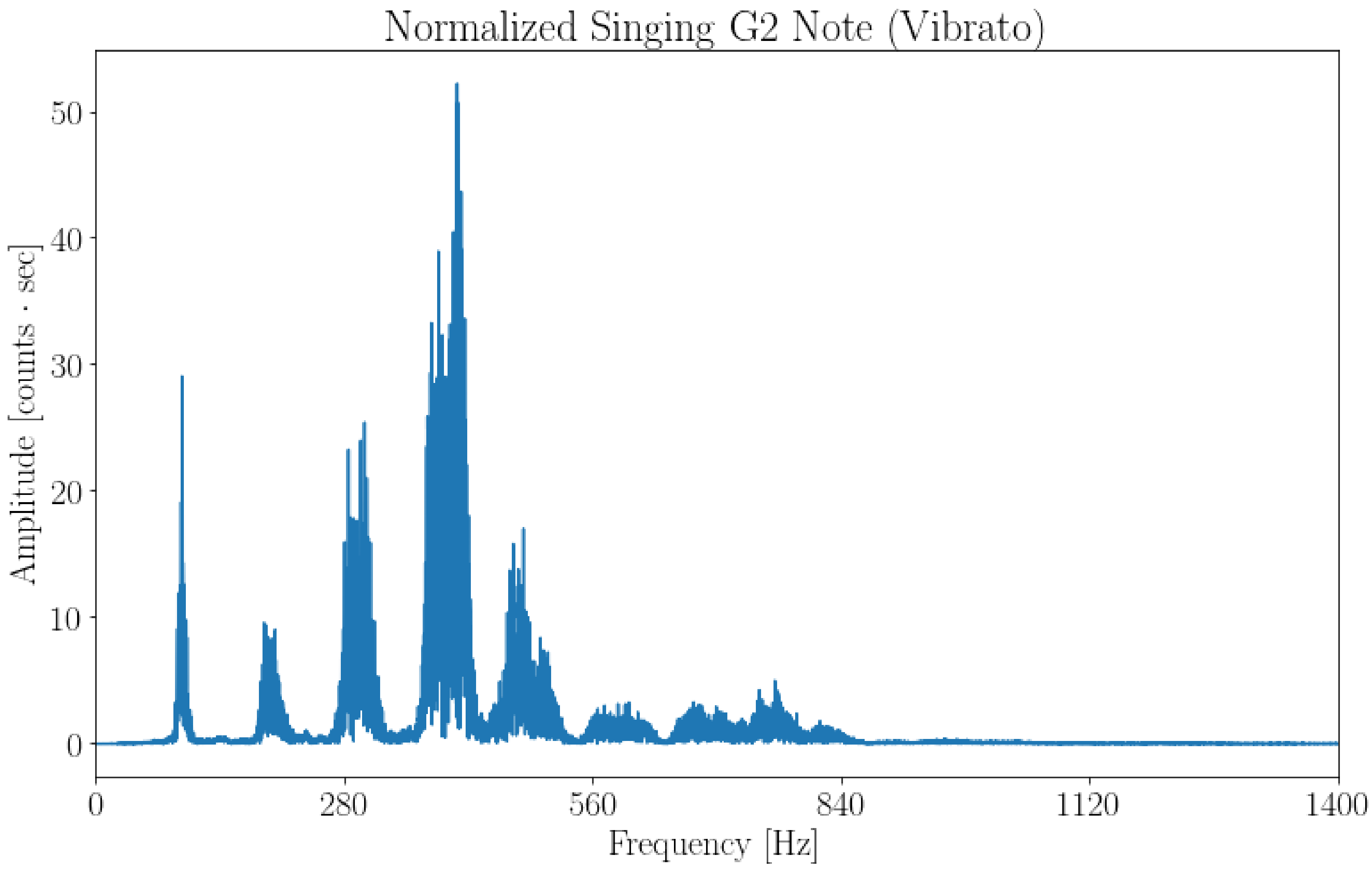


Figure A.14: Voice – G2 Vibrato

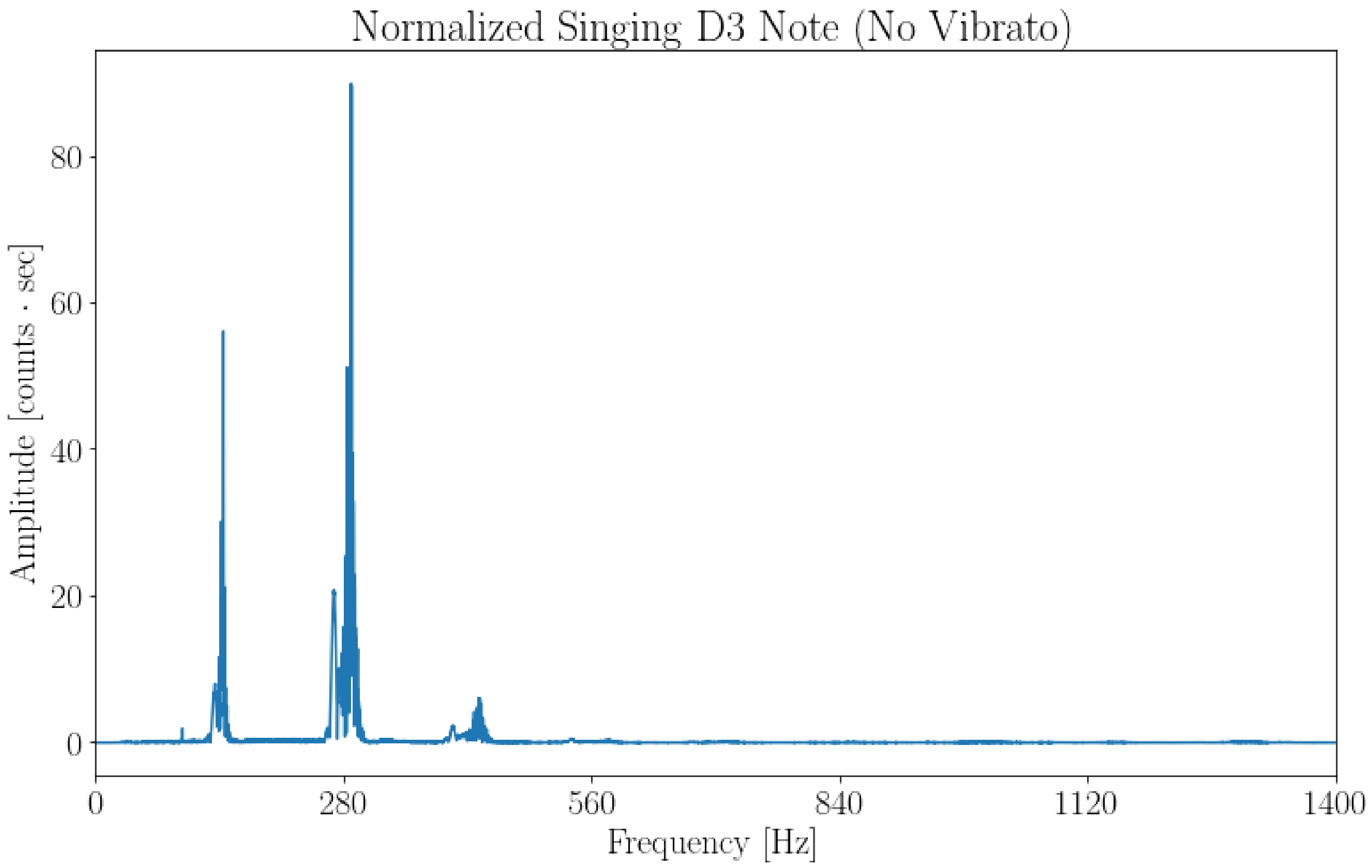


Figure A.15: Voice – D3 No Vibrato

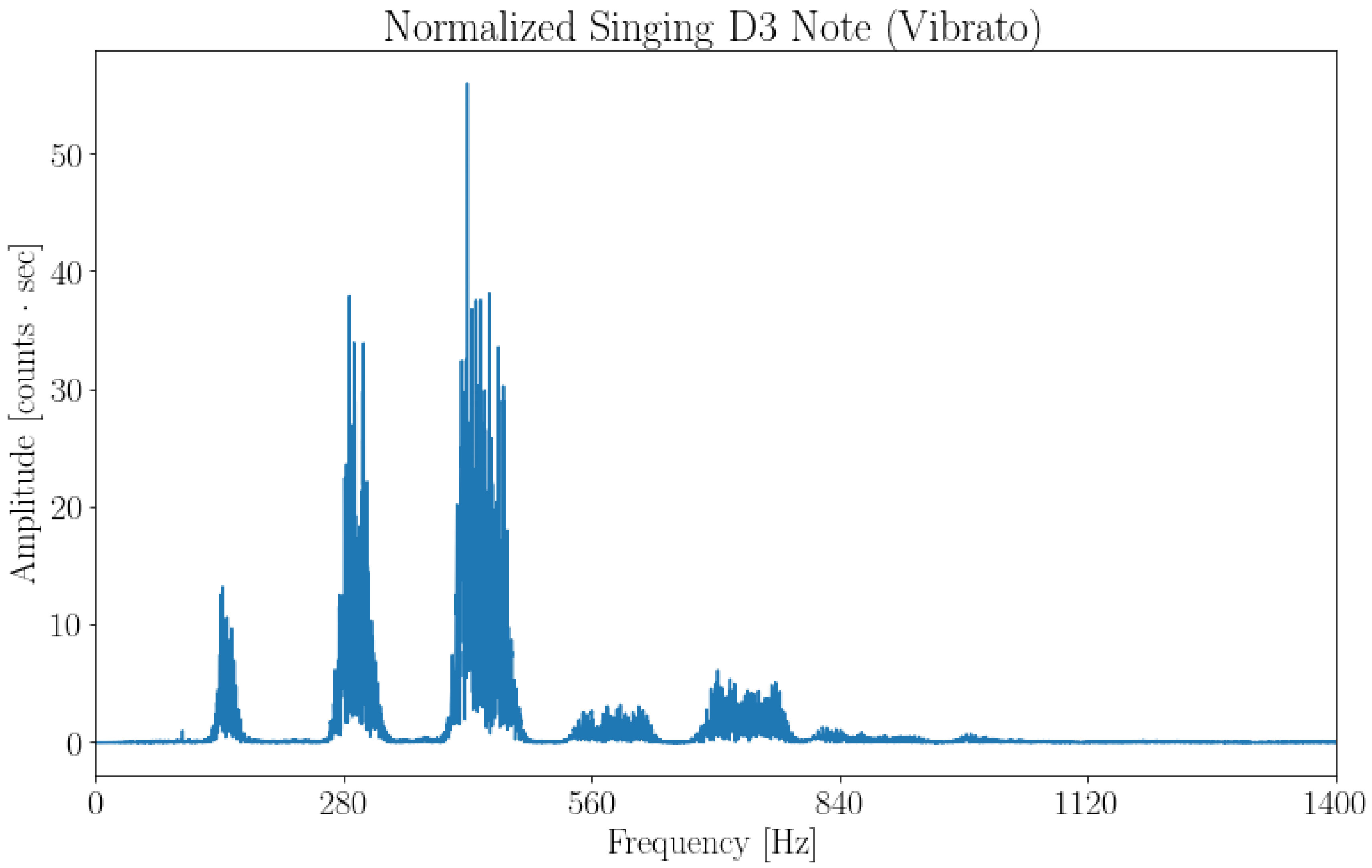


Figure A.16: Voice – D3 Vibrato

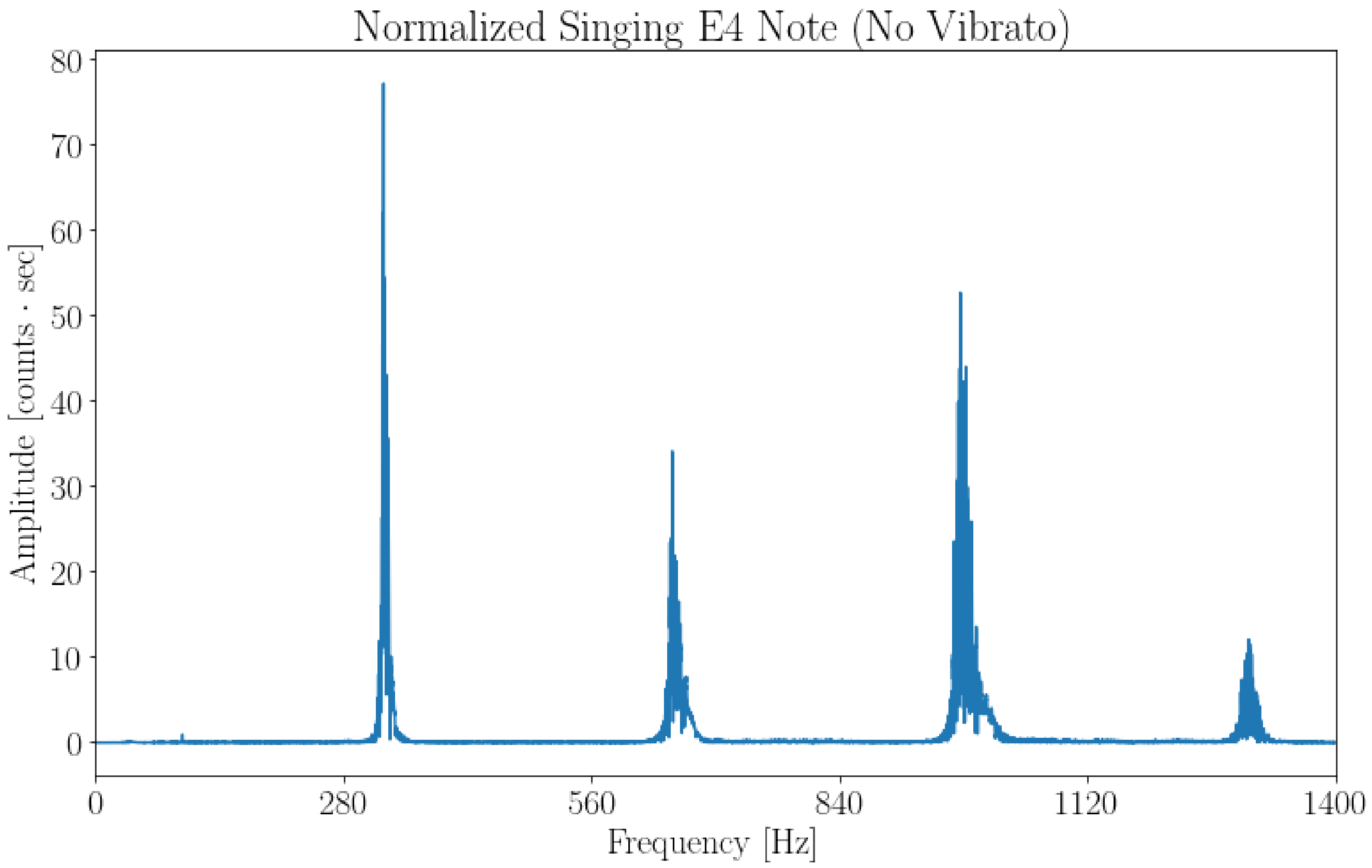


Figure A.17: Voice – E4 No Vibrato

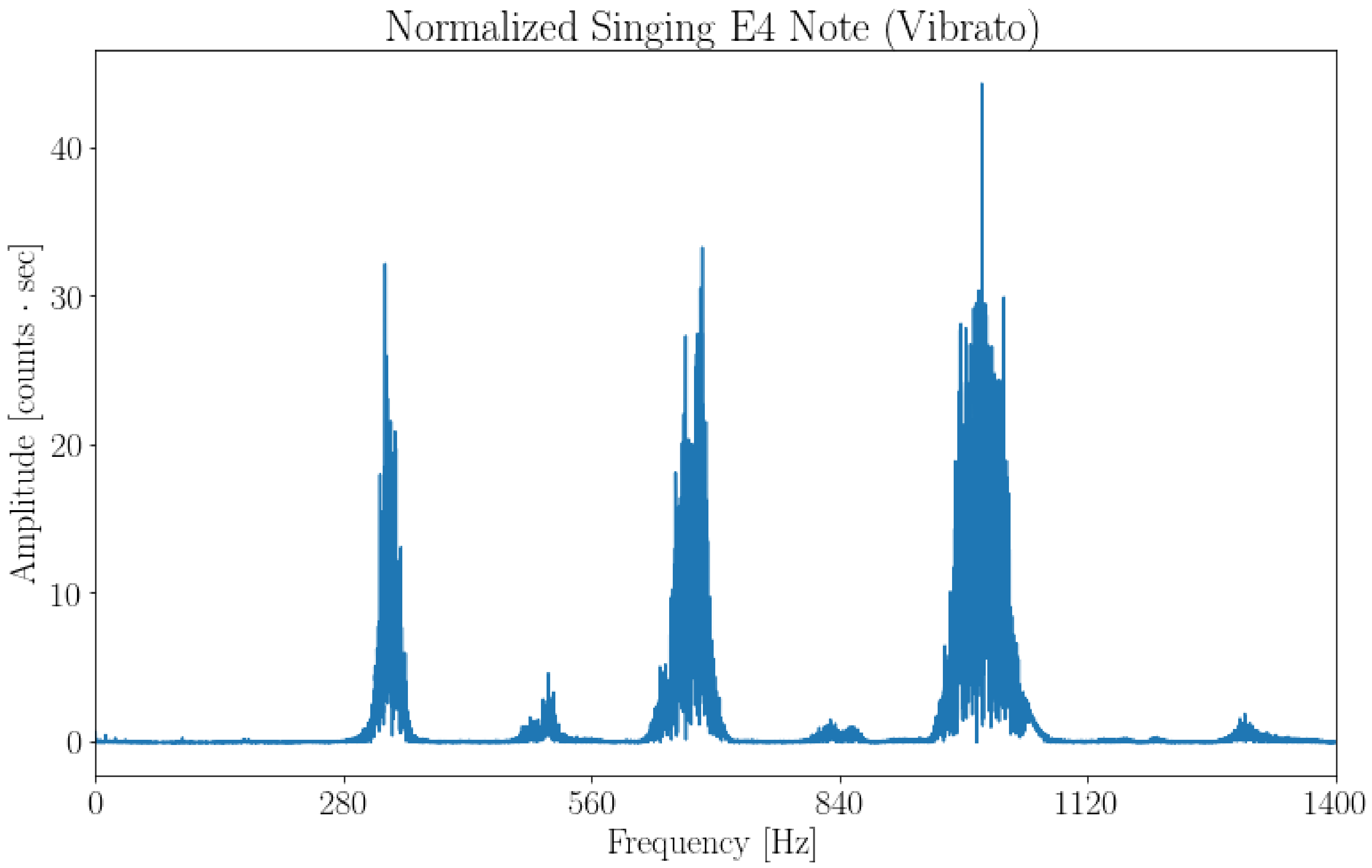


Figure A.18: Voice – E4 Vibrato

Figure A.19: Cello – G2 No Vibrato – Spectrogram – Source 1

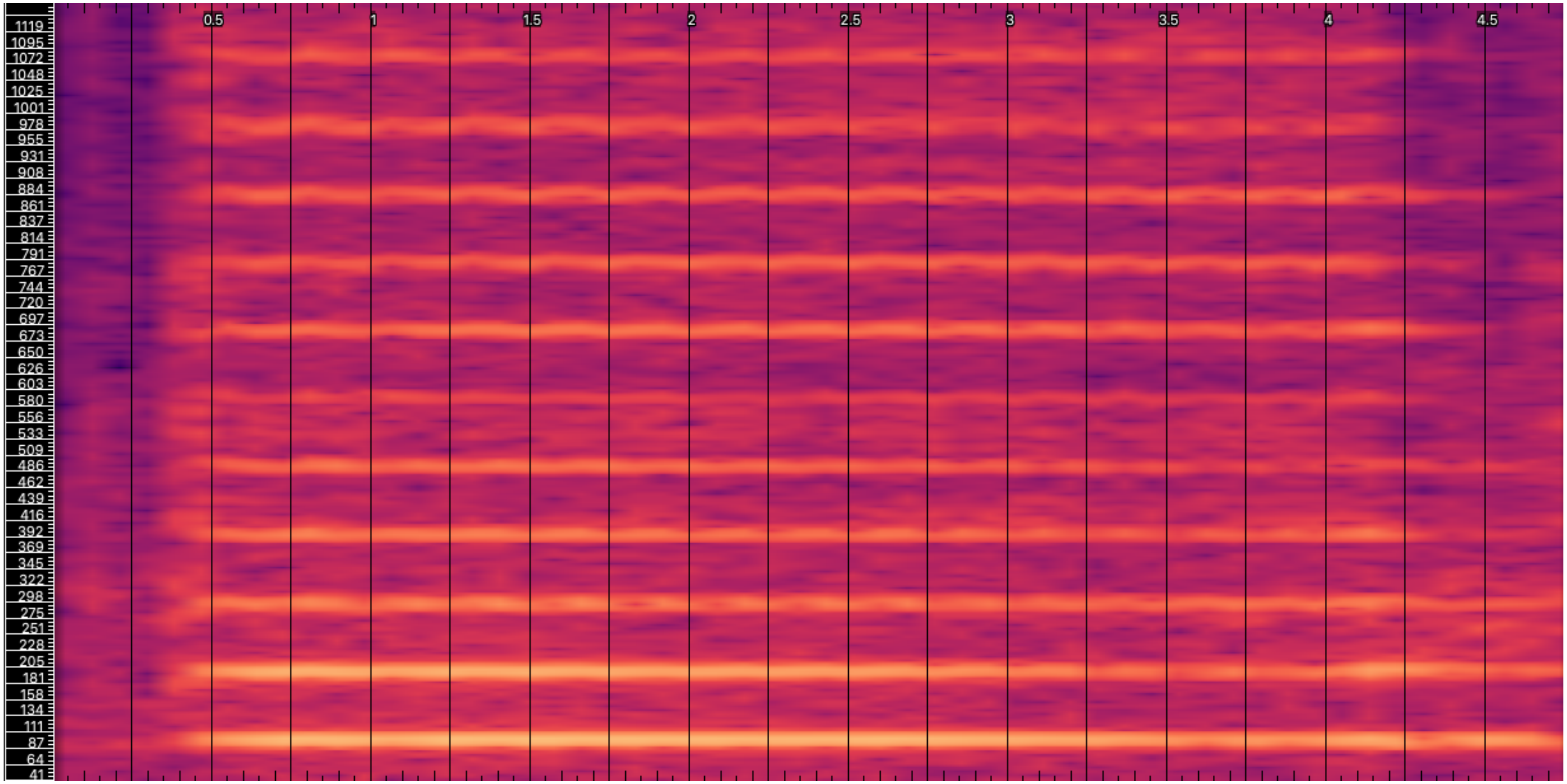


Figure A.20: Cello – G2 Vibrato – Spectrogram – Source 1

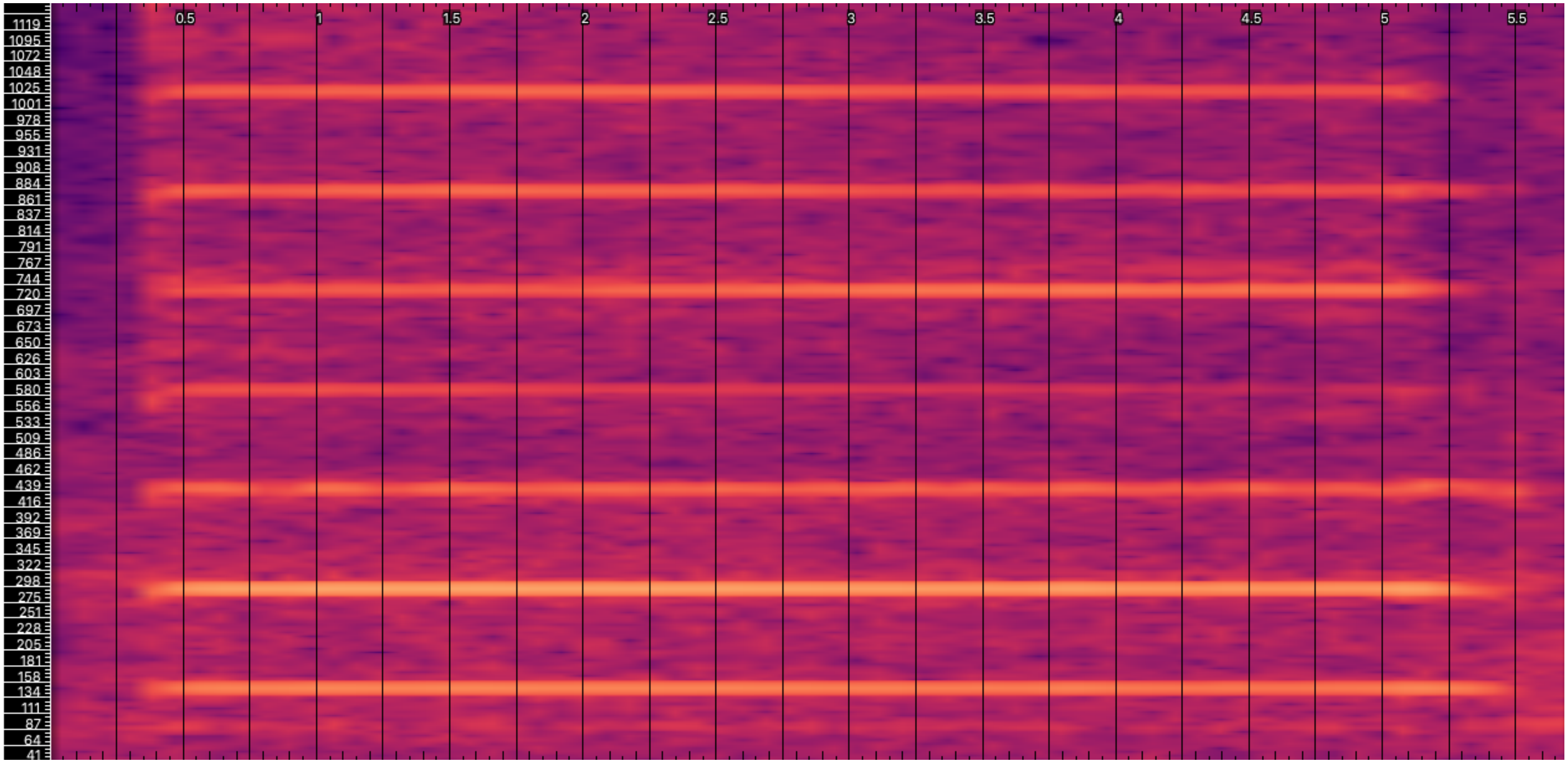


Figure A.21: Cello – D3 No Vibrato – Spectrogram – Source 1

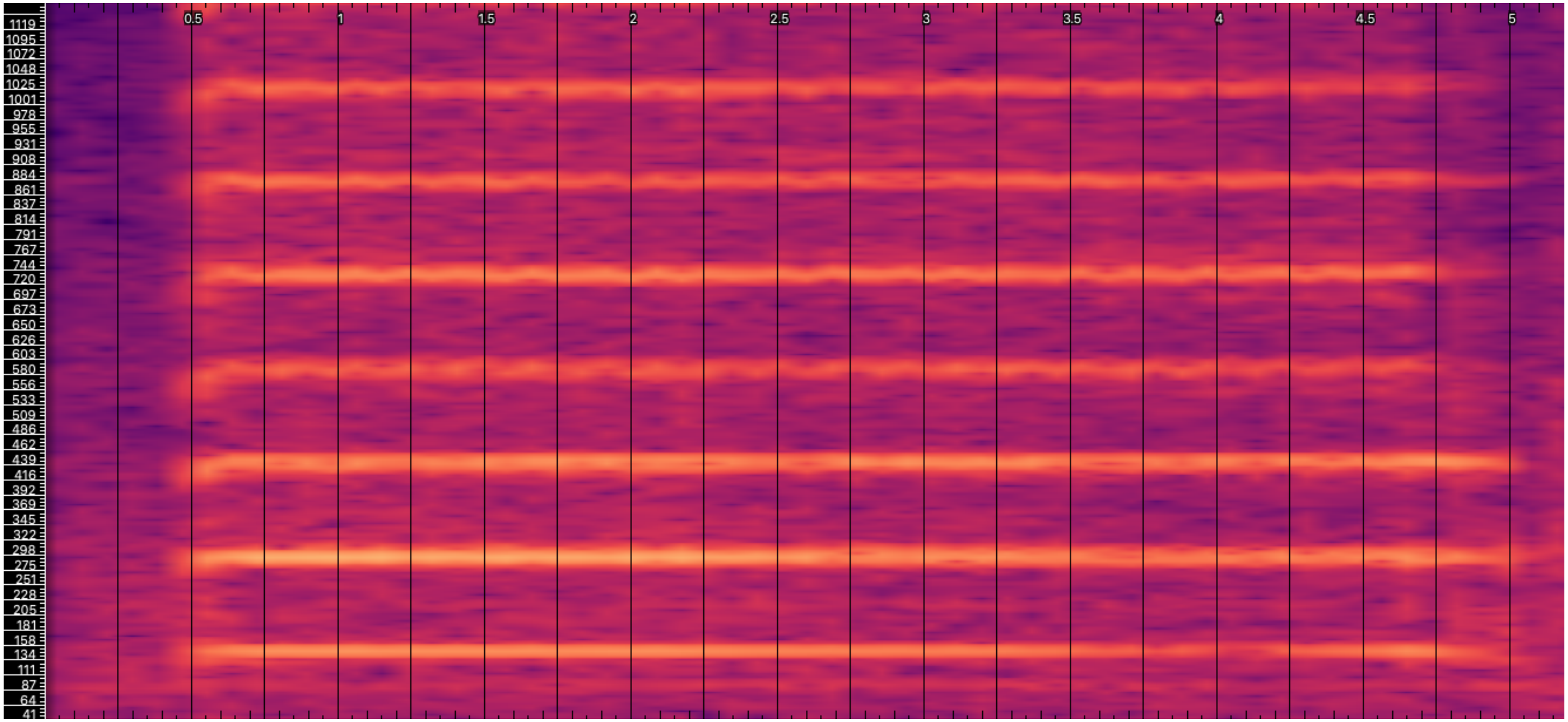


Figure A.22: Cello – D3 Vibrato – Spectrogram – Source 1

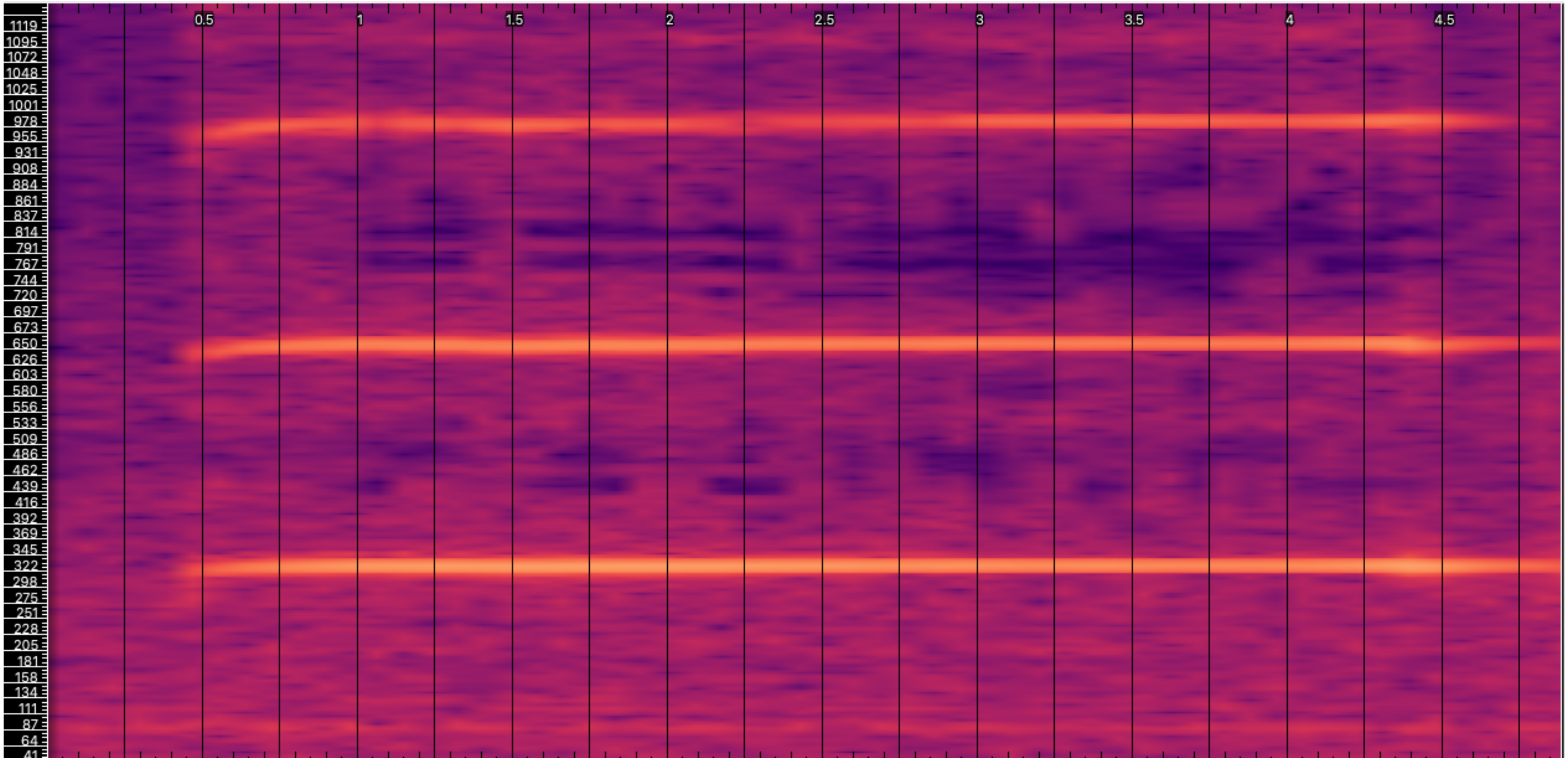


Figure A.23: Cello – E4 No Vibrato – Spectrogram – Source 1

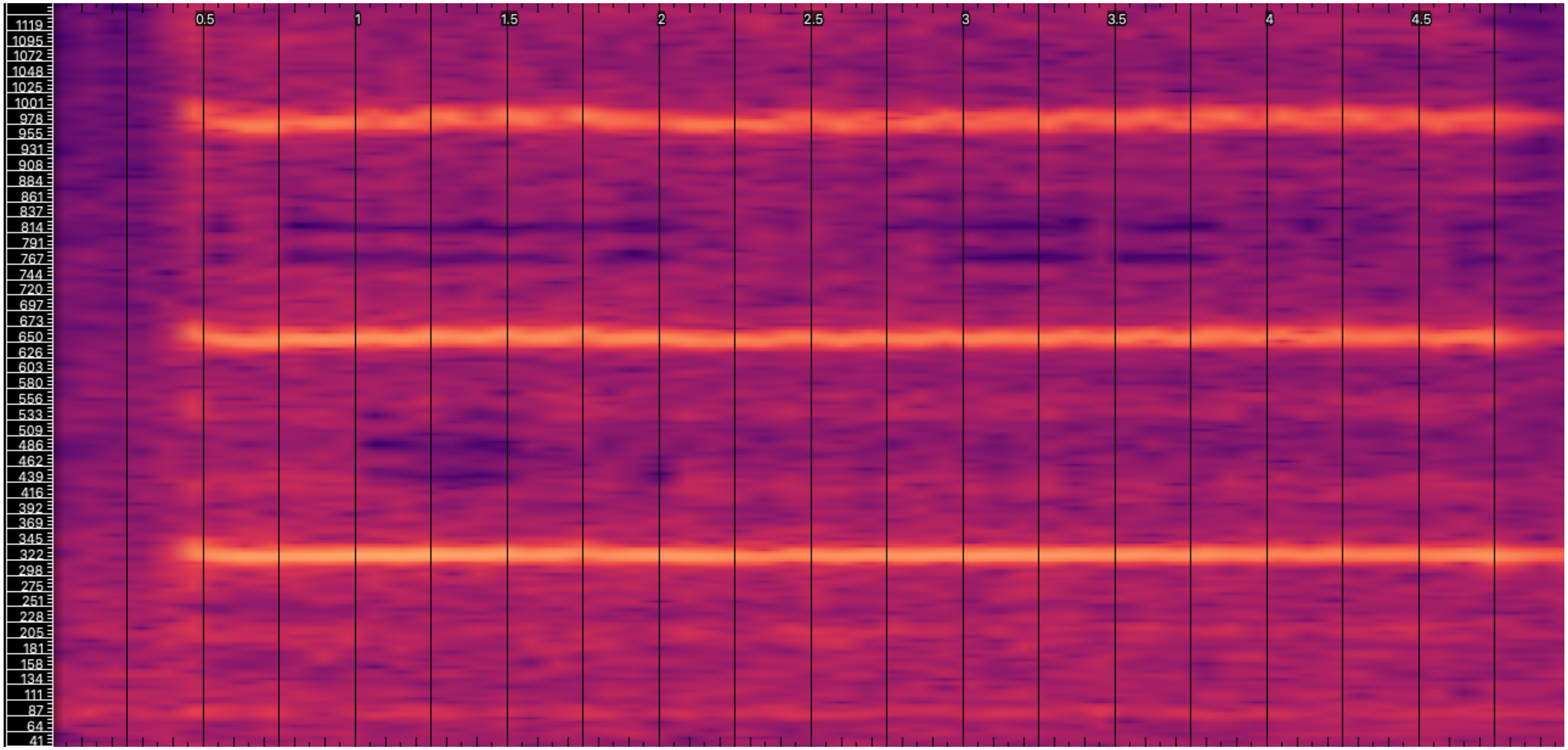


Figure A.24: Cello – E4 Vibrato – Spectrogram – Source 1

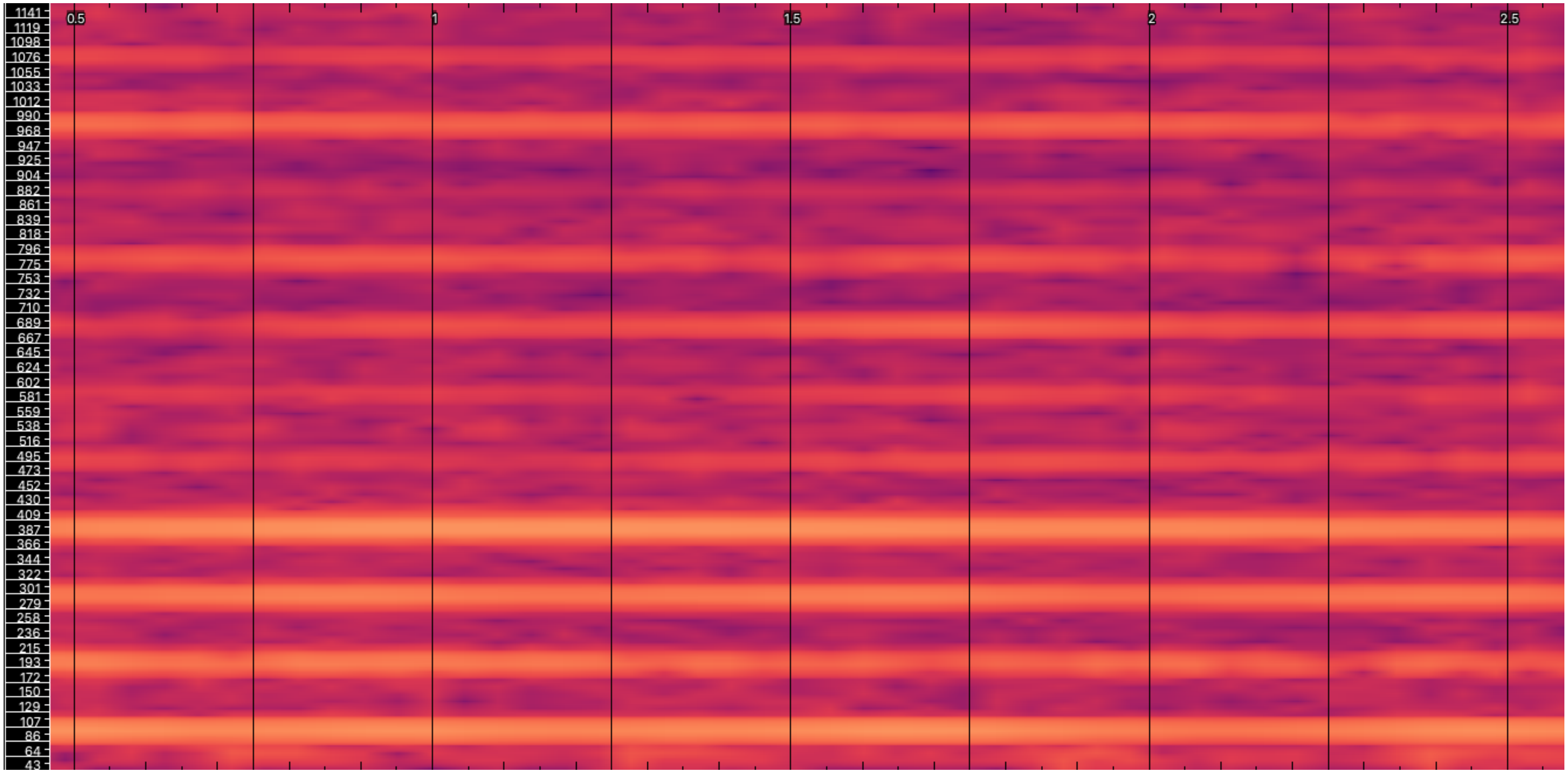


Figure A.25: Cello – G2 No Vibrato – Spectrogram – Source 2

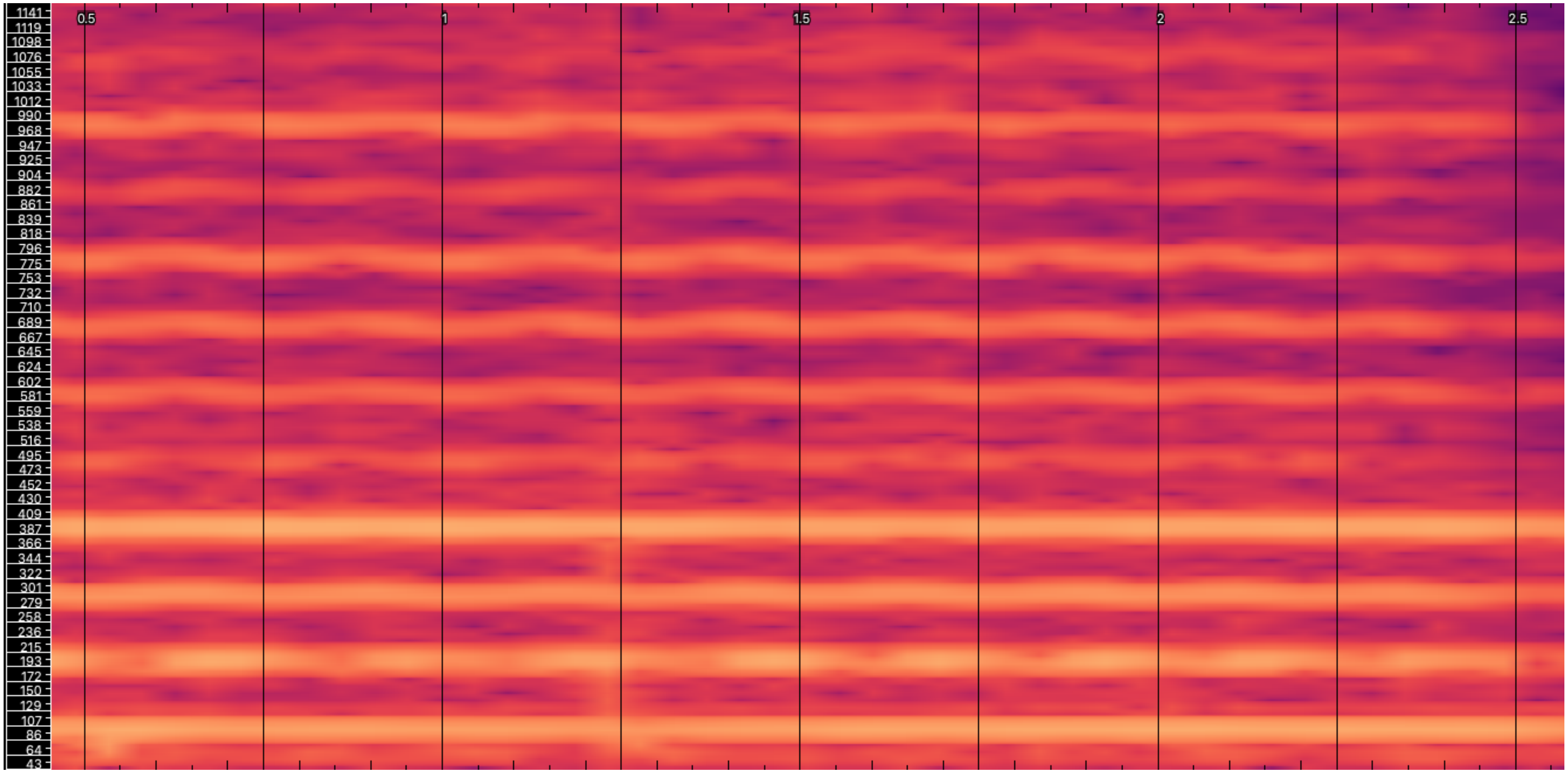


Figure A.26: Cello – G2 Vibrato – Spectrogram – Source 2

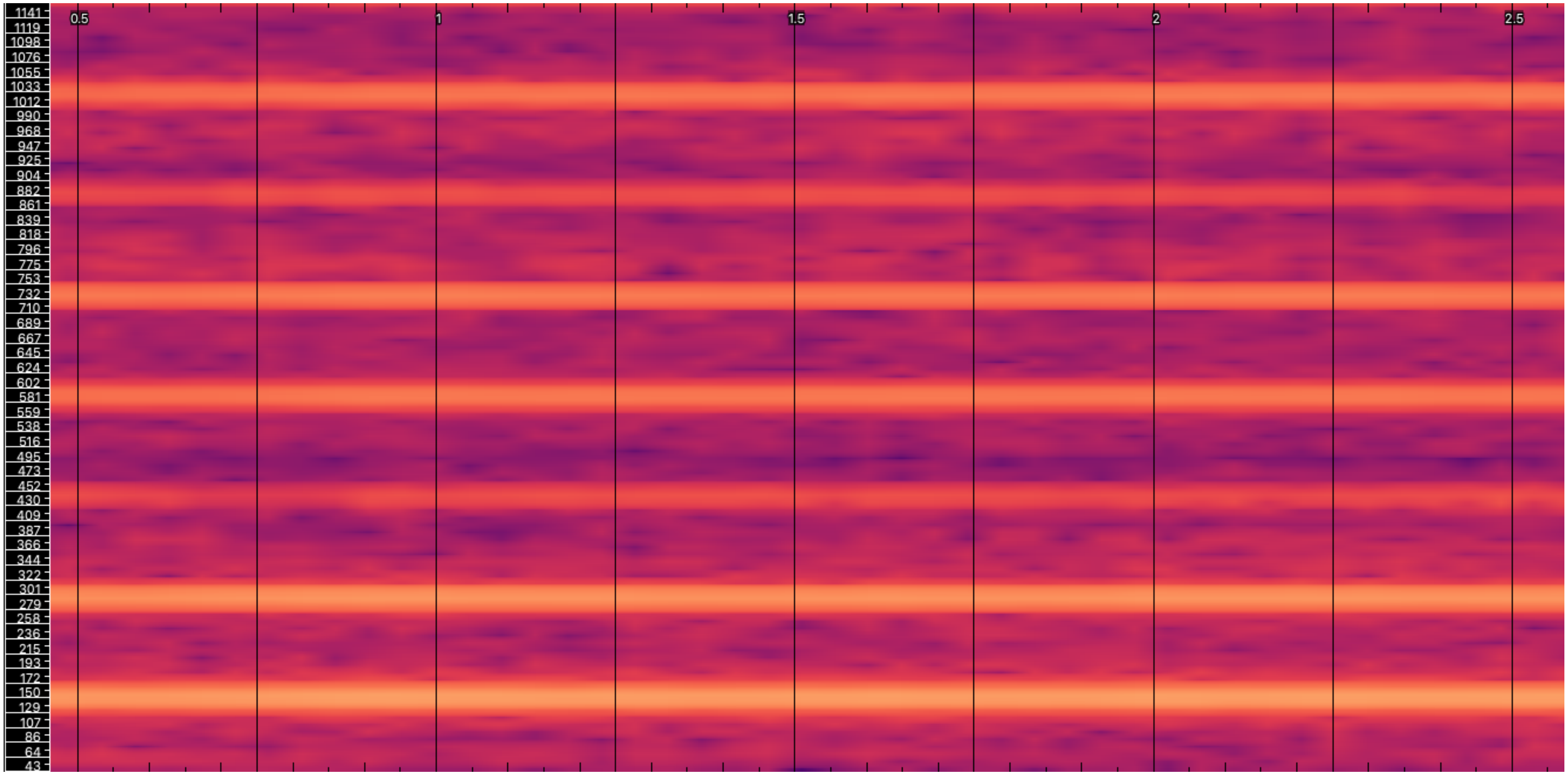


Figure A.27: Cello – D3 No Vibrato – Spectrogram – Source 2

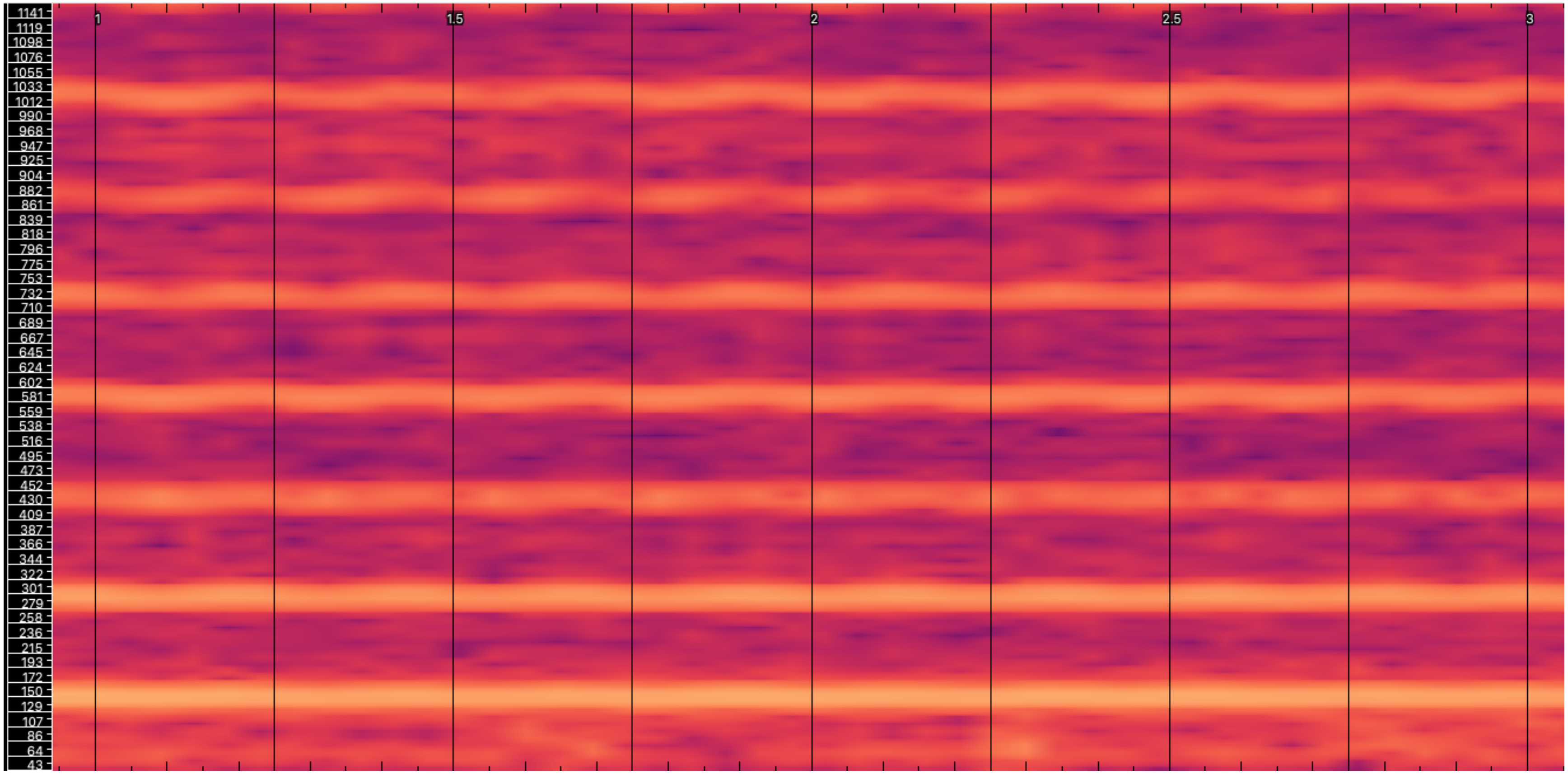


Figure A.28: Cello – D3 Vibrato – Spectrogram – Source 2

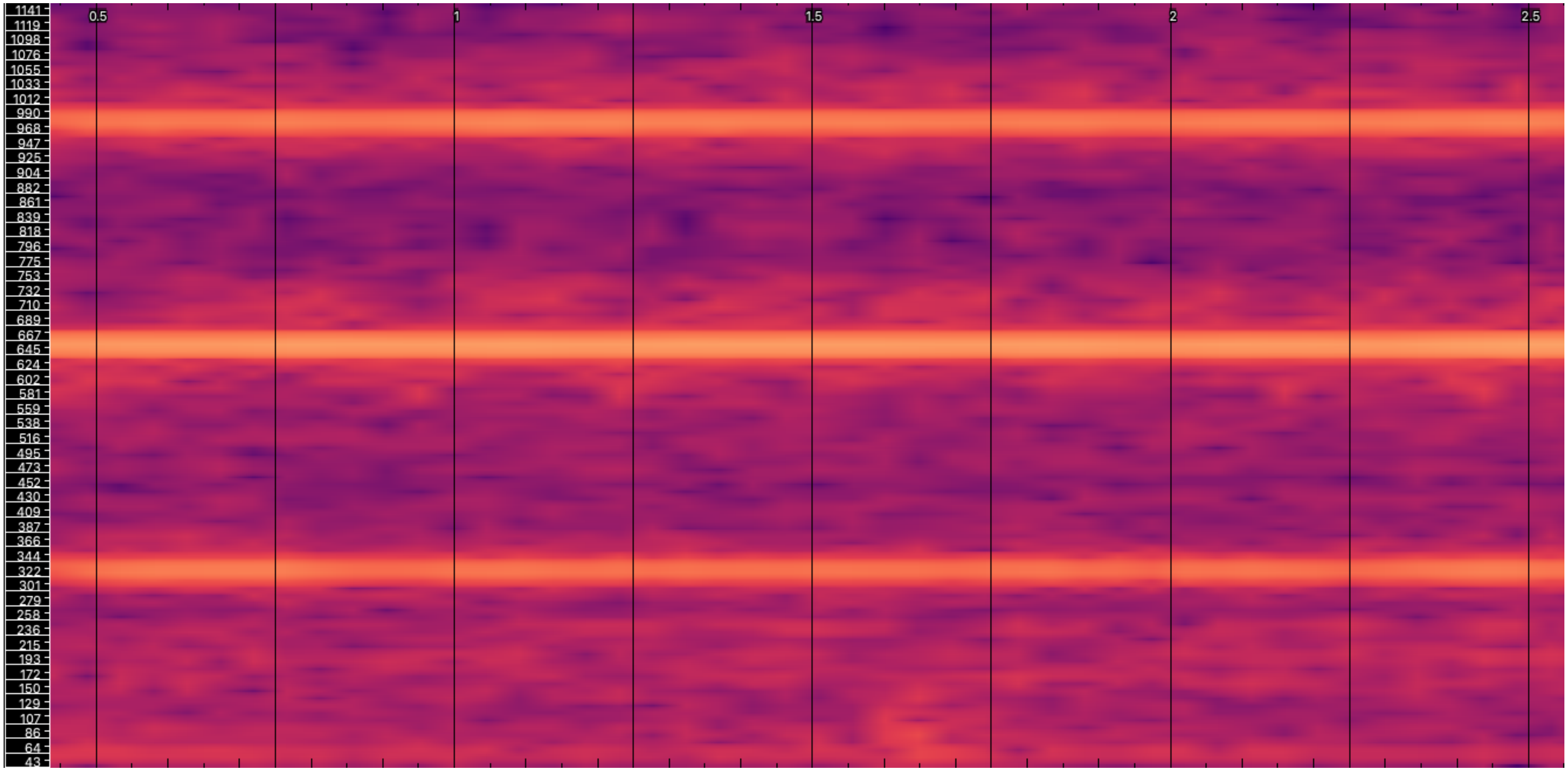


Figure A.29: Cello – E4 No Vibrato – Spectrogram – Source 2

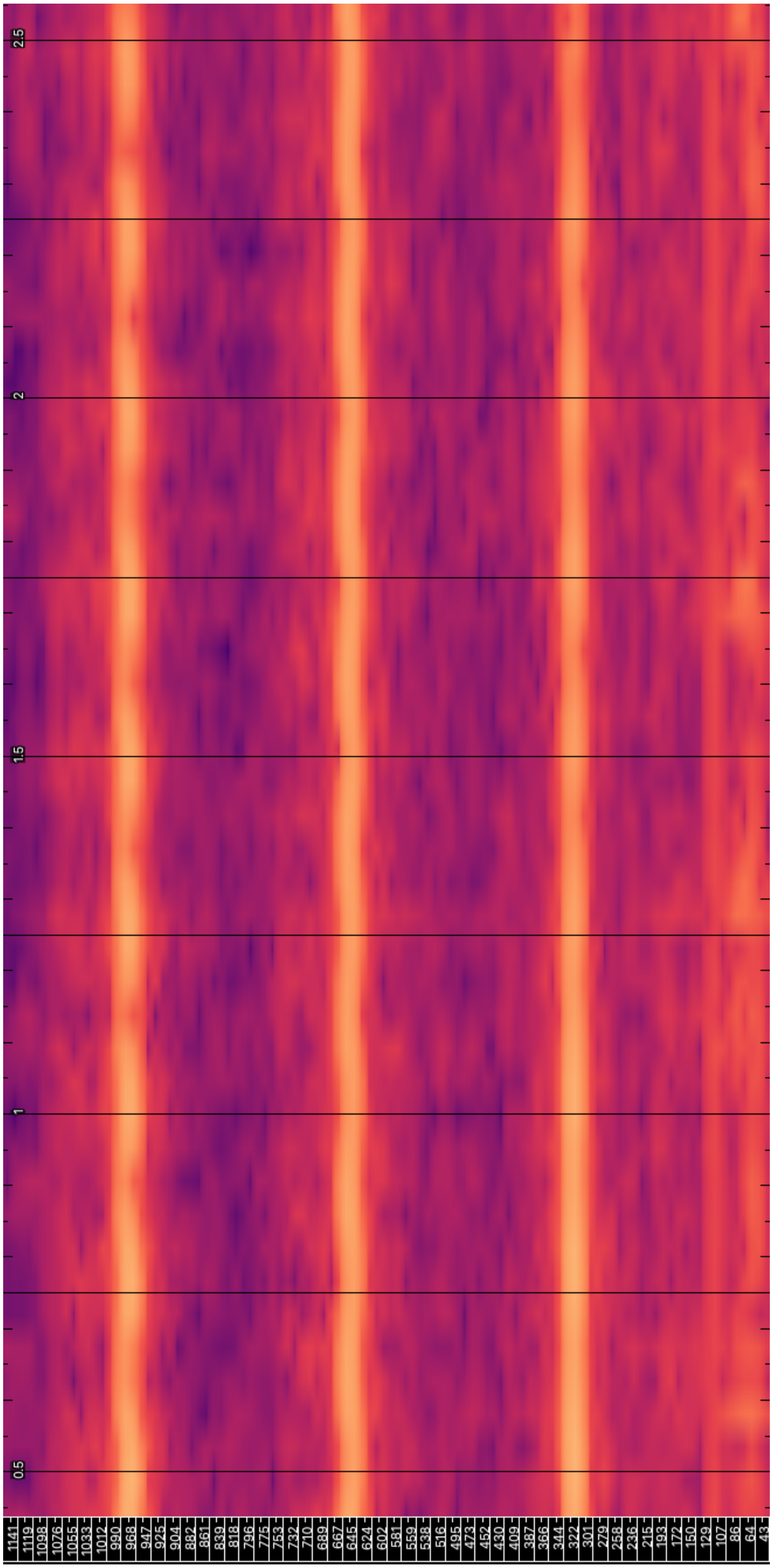


Figure A.30: Cello – E4 Vibrato – Spectrogram – Source 2

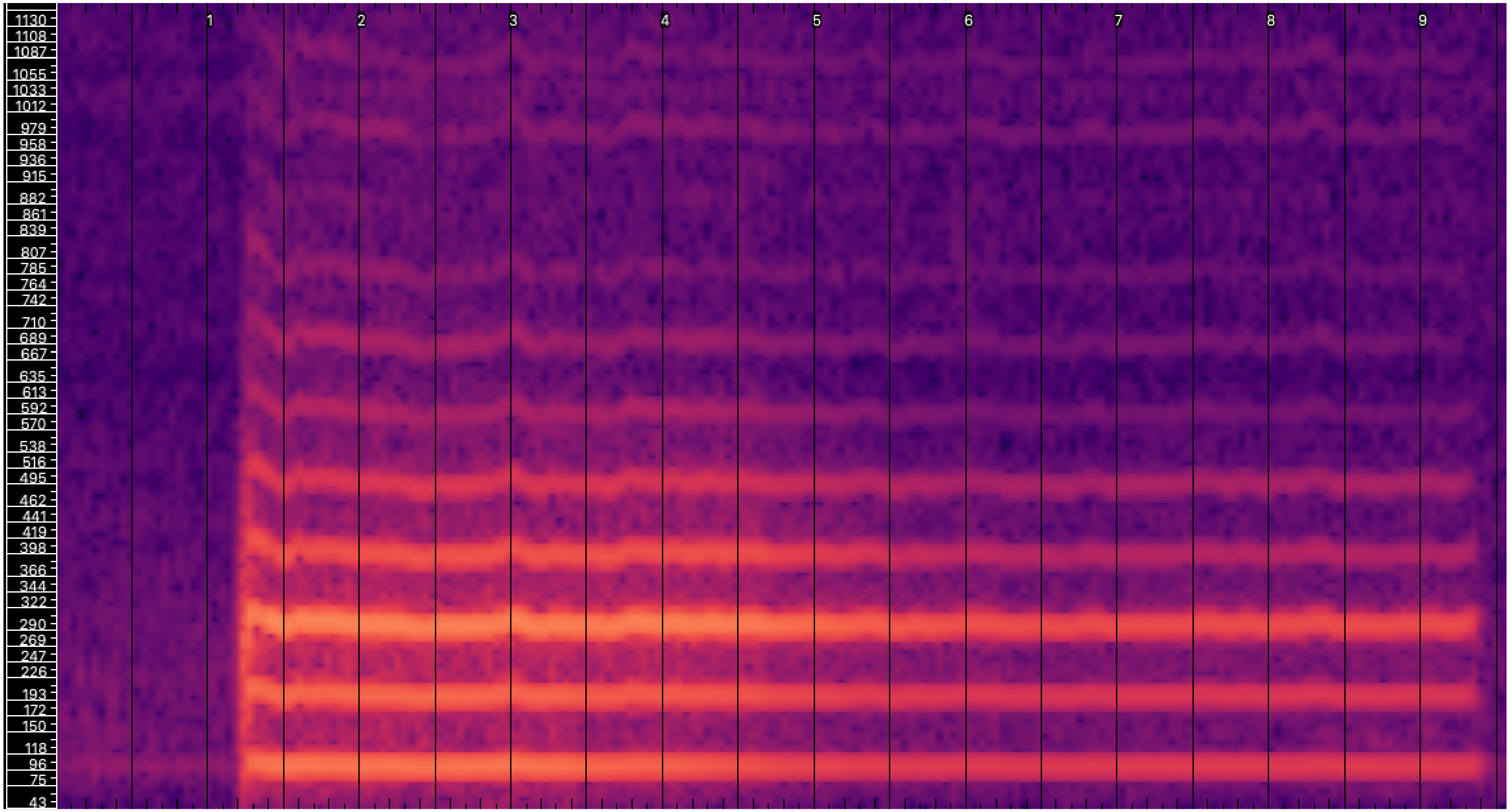


Figure A.31: Voice – G2 No Vibrato – Spectrogram

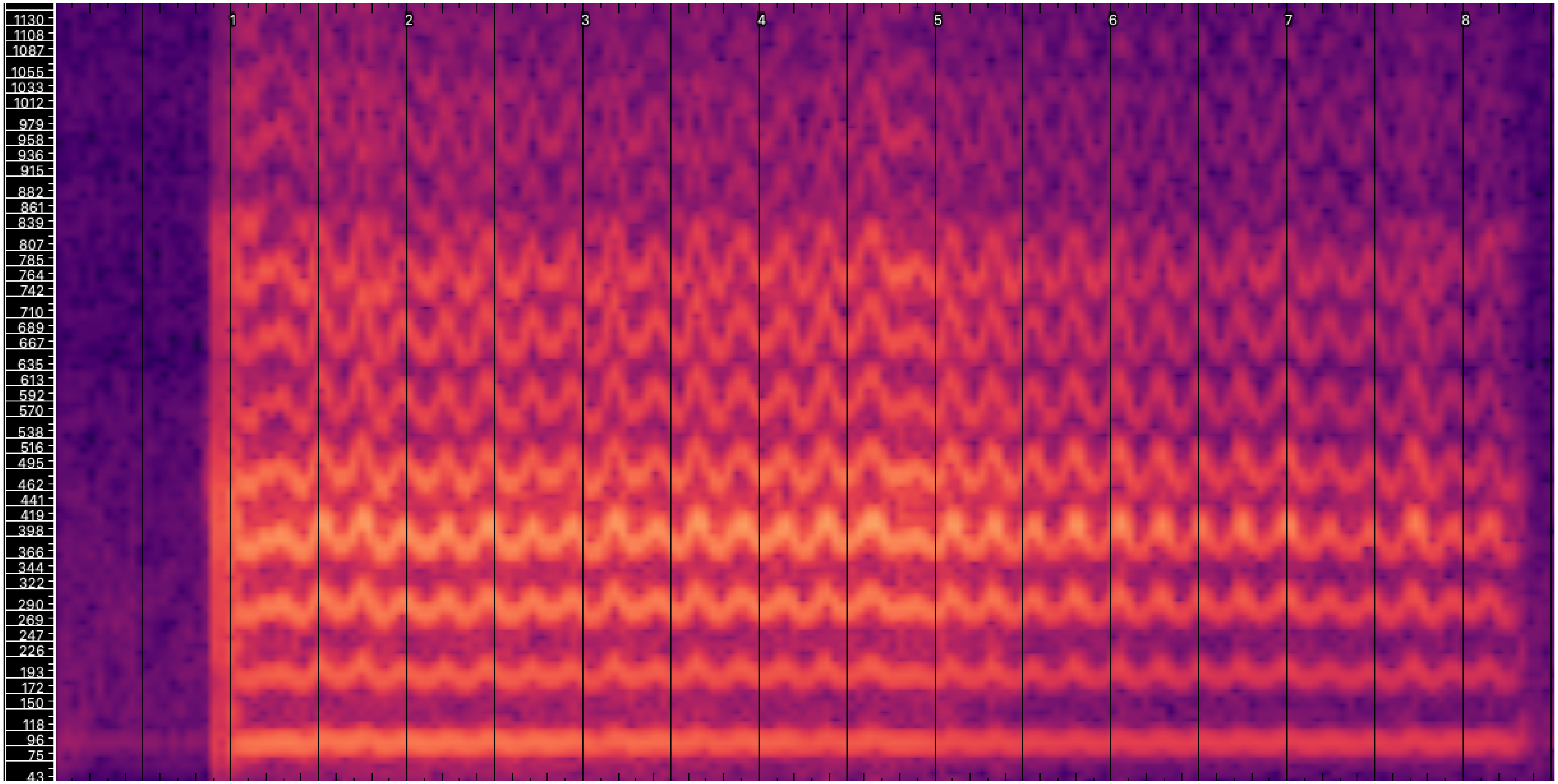


Figure A.32: Voice – G2 Vibrato – Spectrogram

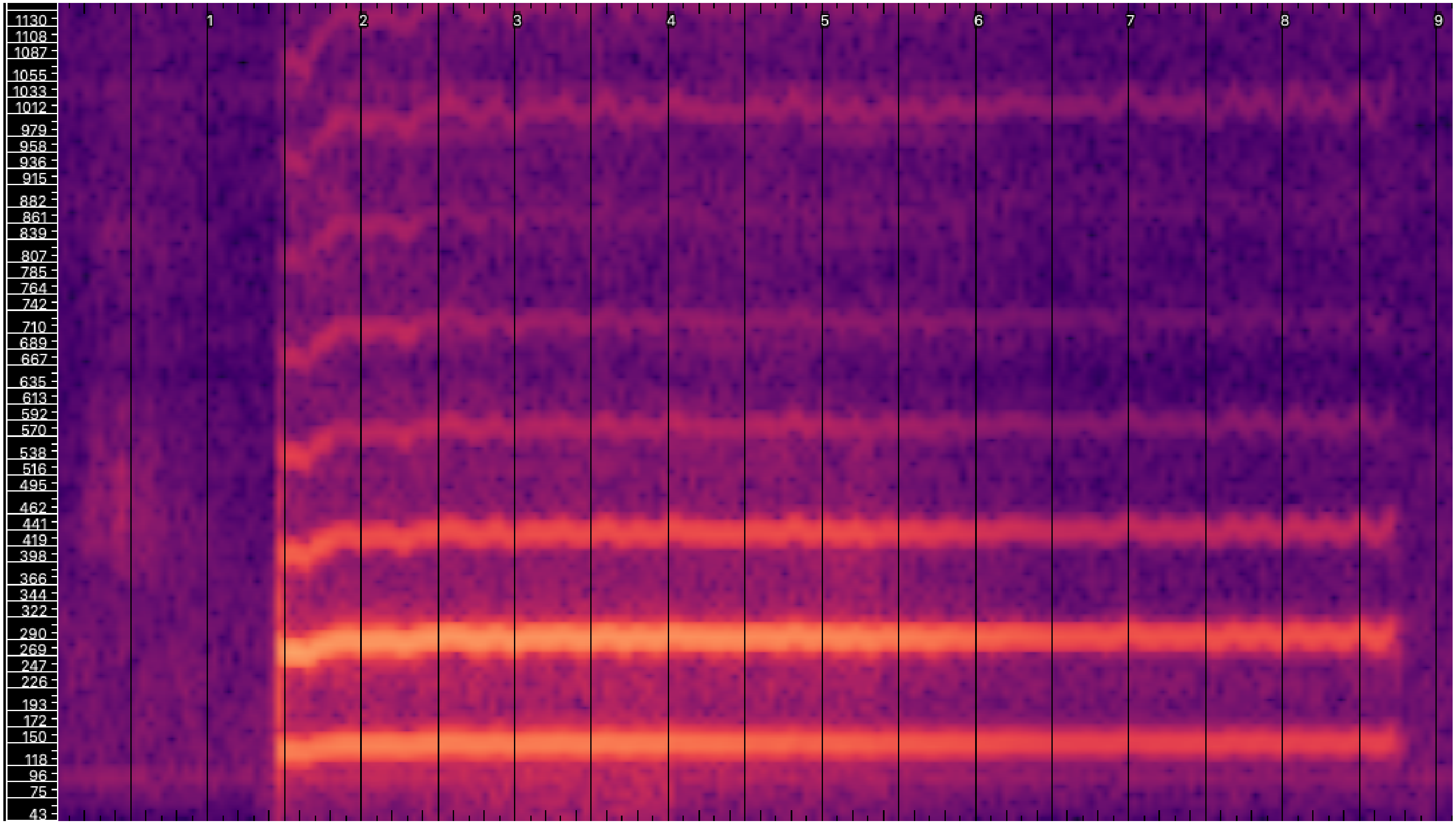


Figure A.33: Voice – D3 No Vibrato – Spectrogram

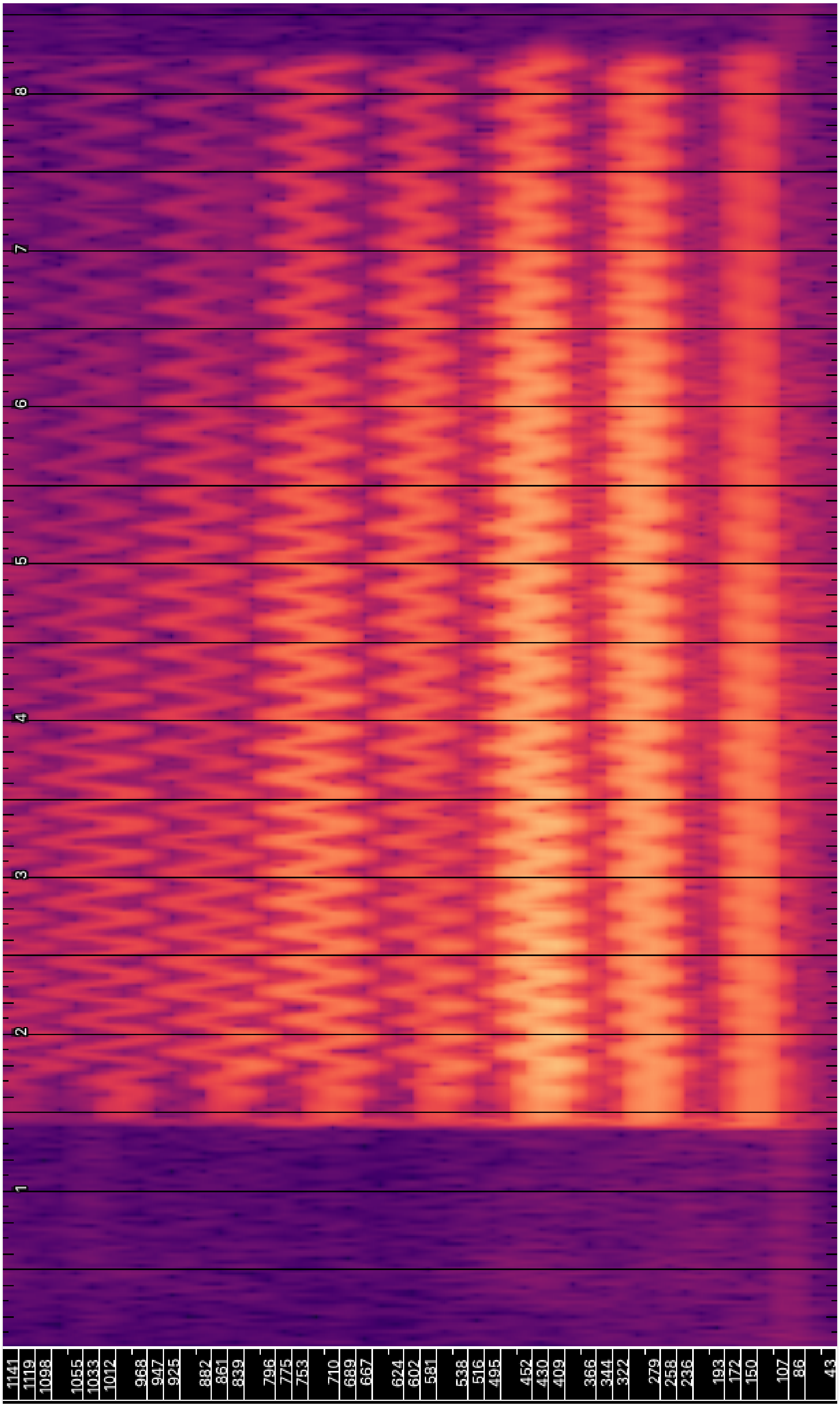


Figure A.34: Voice – D3 Vibrato – Spectrogram

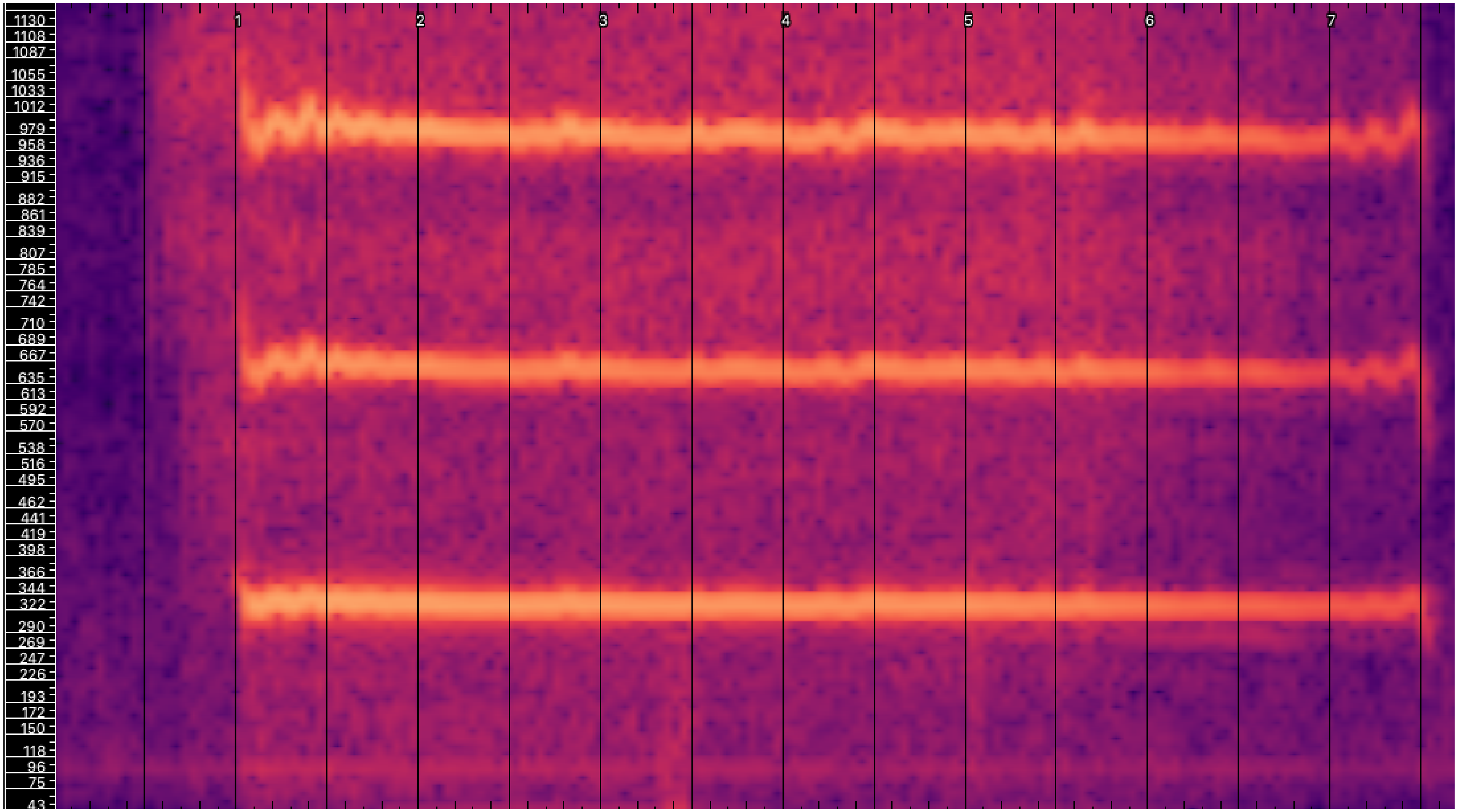


Figure A.35: Voice – E4 No Vibrato – Spectrogram

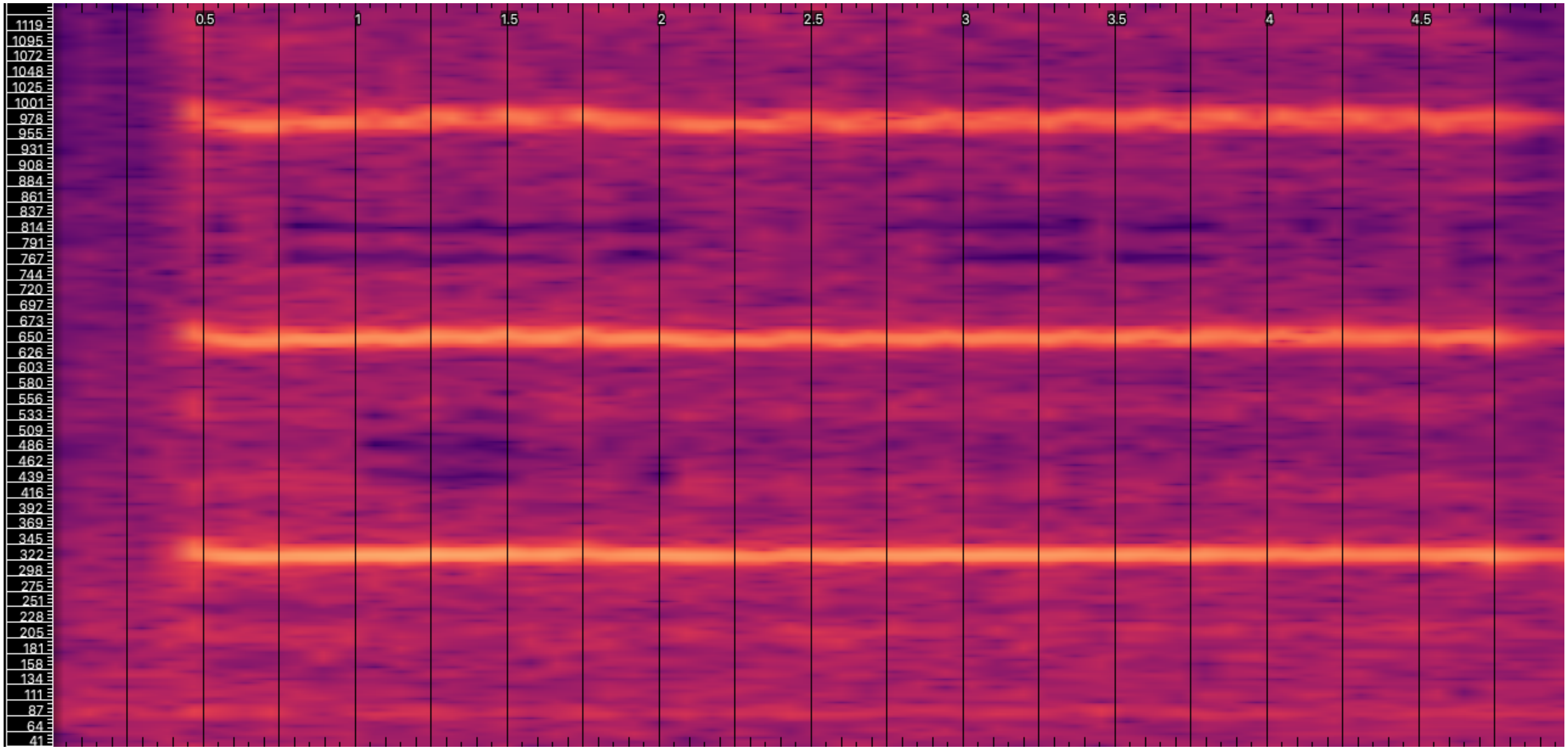


Figure A.36: Voice – E4 Vibrato – Spectrogram