Gabor domain analysis of a three spring damped oscillator

Robert J. Ferguson

ABSTRACT

Familiar aspects of exploration seismology are applied to the study of a class of membranophones that are well modelled as three spring damped oscillators. As a first attempt, damping is ignored in this treatment. Analysis using vibroseis-like source sweeps, cross-correlation, Fourier decomposition, and Gabor domain analysis provide insight into how this class is tuned. It is found that for a particularly good set of three examples, they are tuned to pitches such that they resonate in combinations that are particularly harmonious. Combinations of octave, perfect fourth, and perfect fifth are found and, in particular, the Gabor domain is most useful for discrimination of resonant tones from the ambient noise of the recording system and surroundings.

INTRODUCTION

Membranophones are a class of instruments that use stretched membranes to produce sound. Within this class of instrument are drums (number 21 on the Hornbostel-Sachs musical instrument taxonomy) that produce sound from two stretched membranes that enclose a column of air within a cylinder; the membranes or heads (made of Mylar) are clamped to each end, and the cylinder is often made of wood (Fletcher and Rossing, 1998, pg. 599). Typically, the cylinder is ported by a dime-sized hole (Fletcher and Rossing, 1998, pg. 601). The two heads and the column of air are well modelled as springs, and air rushes in and out through the port during sound production. The port and the load of the ambient air have the effect of damping the produced sound.

Vibration of such a system is governed by the acoustic wave equation, and so the study of the associated acoustics may be expected to be well facilitated using the familiar techniques of reflection seismology: seismic data processing, seismic imaging, and time-lapse seismology. One open question in the study of drum acoustics is what is nature of good drum sound? Drum tuning, for example, is not well documented, and most percussionists tune by feel rather than by a set of procedures or with some kind of tuning equipment. The exception to this is the timpani and concert toms - drums with only a single head. Single headed drums can in fact be tuned (Fletcher and Rossing, 1998, pg. 591).

Western music is based on the diatonic scale of sound Pierce (1994). Diatonic tone combinations are more pleasing musically, and other combinations are found to be not pleasing. These combinations give rise to the "perfect" musical relationships as in Table 1. For the dual headed drums found in jazz and popular music, I hypothesize that the pleasing drum sound that is created by very few tuners is the result of tuning to promote a low tone relative to higher frequency overtones. Further, I hypothesize that combinations of drums, as found in jazz music and pop music, are tuned to each other according to the the most pleasing harmonic intervals of classical music.

For this study, an example of pleasing drum sound had to be found. Fortunately Bob Everett, owner of renowned drum store "Beat it" and noted for quality of his drum tuning,
allowed me to record an example of a set of drums that he had tuned for sale. His strategy for tuning drums is entirely qualitative "You've got to sit on the kit and it's got to feel good" - this is a qualitative statement about how each drum in a set of drums sound in combination with the others. "To me, it's more about feel and torque that you can feel - sound and little vibrations - you notice those things and you get rid of them" - another qualitative statement this time about overtones and the need to reduce their presence in the sound field. Asked to describe his tuning process, Mr. Everett replies "I don't know what the process is ... it's a gut feeling ... there's so many things that can screw up tuning ... you've got a floor tom that sounds great, and then you move it, and then it sounds terrible." This last statement suggests that because there are so many important variables, it might not be possible to completely describe physically the nature of good drum sound. Rather, the most tractable approach is to determine what desirable outcome is achieved through the complex and rather arbitrary art of tuning.

To determine the nature of good sound, a number of recordings were made over two afternoons using the following procedure: 1) Rather than strike each drum and record the response, adopt the seismic Vibroseis procedure*: 1) broadcast Vibroseis-like sweep tones through loud speakers at the individual drums, and acquire reference recordings by close miking with a pillow on the drum head closest to the mike. 2) Remove the pillow and make a monitor recording. 3) Cross correlate the reference and monitor signals to reveal, approximately, the impulse response for each head of each drum. 4) Analyze the recordings in time, frequency, and in the Gabor domain (Margrave et al., 2005).

The result of the analysis strongly suggests that the drumset is tuned approximately to three of the most pleasing classical musical intervals octave (tom tom / bass drum), perfect fifth (tom tom / floor tom), and perfect fourth (floor tom / bass drum). The Gabor domain was essential in the determination of the resonant mode of the tom tom and the decay of overtones. The use of reference recordings is essential in determining the drum effect by visual comparison. Though the spectrum of the sweep signal was not flat, analysis of cross-correlation spectra was more revealing of the fundamental drum modes than analysis of the uncorrelated recordings.

---

*A struck drum sets up complex overtones where the goal here is to determine to what fundamental frequency each drum is tuned to.
The purpose of this section is to provide a very simple, idealized physical description of a vibrating drum. To facilitate this description, the drum is reduced from a three spring, damped oscillating system to a single circular membrane whose vibration is excited by striking in the centre of the head. The result is that, ideally, a drum has a low, fundamental frequency that is determined by head diameter, density, and head tension, and the root of the lowest order Bessel function of the first kind.

In its most basic form, then, a drum consists of a circular clamped membrane governed by the wave equation in spherical coordinates according to

$$\left[ \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right] u(r, t) = \frac{1}{c^2} \ddot{u}(r, t), \quad 0 < r < a, \quad t > 0,$$

where \(a\) is the speed of waves travelling on the drumhead and

$$u(a, t) = 0, \quad u(r, 0) = \xi(r), \quad \dot{u}(r, 0) = 0,$$

are boundary conditions that impose no motion of the drumhead at the boundary \(r = a\), a membrane configuration at \(t = 0\) of \(\xi(r)\), with no motion of the membrane at \(t = 0\) (Boyce and DiPrima, 1987, pg. 641), and with no air loading or membrane stiffness (Worland, 2010).

The solution \(u(r, t)\) is obtained by separation of variables according to the boundary conditions and the expectation that the solution be bounded and oscillatory with the result

$$u(r, t) = \sum_{n=1}^{\infty} c_n J_0 \left( \lambda_n \frac{r}{a} \right) \cos \left( \lambda_n \frac{ct}{a} \right),$$

where Bessel functions \(J_0(\lambda_n \frac{r}{a})\) are functions of an infinite number of roots \(\lambda_n, 1 \leq n \leq \infty\) (see Table 2 for numerical values for \(1 \leq n \leq 4\)) (Boyce and DiPrima, 1987, pg. 642). Coefficients \(c_n\) depend on the initial configuration \(\xi(r)\) of the membrane according to:

$$c_n = \frac{\int_0^a r \xi(r) J_0 \left( \lambda_n \frac{r}{a} \right) dr}{\int_0^a r J_0^2 \left( \lambda_n \frac{r}{a} \right) dr}$$

(Boyce and DiPrima, 1987, pg. 642).

In Figure 1 is the graph of a few snapshots of a theoretical vibrating drumhead according to equations 3, 4, and the values for \(\lambda_n\) found in Table 2. Each line is a snapshot of the vibrating head at approximately 1 ms intervals. The starting function \(\xi(r)\) at \(t = 0\) is a pyramid shape with the apex pointing down - much like the response of a drum head to being pressed in its center. Under the assumptions of no air loading and a symmetric

<table>
<thead>
<tr>
<th>(n)</th>
<th>(\lambda_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.04048</td>
</tr>
<tr>
<td>2</td>
<td>5.5201</td>
</tr>
<tr>
<td>3</td>
<td>8.6537</td>
</tr>
<tr>
<td>4</td>
<td>11.792</td>
</tr>
</tbody>
</table>

Table 2. Roots \(\lambda_n\) for \(1 \leq n \leq 4\) Bessel functions of the first kind (Powers, 1987, pg. 256).
FIG. 1. Cross section of a vibrating drum head. The drum head is 80 cm in diameter. Each line is a snapshot of the drum head at approximately 1 ms intervals.

FIG. 2. Bessel functions $J_1$, $J_2$, $J_3$, and $J_4$ for the drumhead in Figure 1.

initial configuration of the membrane, the graphs show that during vibration the membrane tends to retain its shape. That is, the simplistic, low-order model of vibration suggests that there is a fundamental vibration mode for a drum head.

In fact, a drumhead with density $\rho$ and under uniform tension $T$, the fundamental (often called the 01 mode) is given by

$$f_n = \frac{\lambda_n}{2\pi a} \sqrt{\frac{T}{\rho}},$$

for $n = 1$, where $\lambda_1 = 2.04048$ from Table 2 and $a$ is the radius of the drum head (Fletcher and Bassett, 1978; Worland, 2010). The lowest (fundamental) vibration mode of a drumhead, then, is directly proportional to $\lambda_1$. This is not surprising when one considers a number of the Bessel functions, and these are graphed for the theoretical drum head of Figure 1 in Figure 2. From Table 2, $\Lambda_1 < \lambda_2 < \lambda_3 < \cdots$, so according to equation 5, the higher Bessel functions in the model contribute frequencies higher than the fundamental. Moreover, because the maximum amplitudes of the Bessel functions decrease from $J_1$, the higher frequencies contributed by $J_2, J_3, \cdots$ are of lower amplitude.
Vibroseis

The recording of a struck drum results in a complex system of overtones that lie on top of the fundamental tone that is of interest in this study. Rather than strike the drum, then, a linear sweep of frequencies from low to high frequency is broadcast through loudspeakers at the drum where the fundamental tone is excited. To determine the fundamental tone, then, the sweep plus the response of the surrounding room response to the sweep must be removed. This is done by recording two sweep signals. The first with the experimental set up in place with an individual drum, but with the drum muffled by a pillow; this is the reference recording. The pillow is removed for the second recording, the monitor recording, is made. Following the Vibroseis method, where a linear sweep is broadcast into the ground through a baseplate coupled to the ground, the convolutional model of source and system response is adopted.

In the convolutional model, the source sweep $s(t)$ is convolved with the impulse response $g(t)$ to give the recording $x(t)$ according to

$$x(t) = s(t) * g(t),$$

where $*$ indicates convolution (Baeten, 1989). Convolution in the Fourier transform domain is

$$X(\omega) = S(\omega) \cdot G(\omega),$$

where, $\omega$ is angular frequency, $X$, $S$, and $G$ are the Fourier transform spectra of $s$, $s$, and $g$ respectively. Cross-correlation $U$ of spectrum of the recording $X$ and the sweep $S$ is

$$U(\omega) = X(\omega) \cdot S^\dagger(\omega) = |S(\omega)|^2 \cdot G(\omega),$$

where $\dagger$ is complex conjugate. Equation 8 implies that, so long as the spectrum of the sweep is flat over the range of $\omega$, cross-correlation $C$ is simply a scaled version of the spectrum of the impulse response of the drum. Inverse Fourier transform returns the spectrum to the time domain.

The Gabor transform domain

Gabor transforms are used in seismic imaging and they are a special case of the generalized S-transform. Given a 1D signal, the Gabor transform returns a 2D spectrum (complex valued) in time and frequency. Analysis in this domain allows the interpreter to study the frequency content of a signal as it changes through the length of the recording. Given the recording $u(t)$ of the vibrations of a drum, then, the Gabor transform $U(t, f)$ is given by

$$U(t_k, f) = \int_{-\infty}^{\infty} u(t) \cdot g_k(t) \cdot e^{-2\pi ift} dt, \quad 1 \leq k \leq N,$$

where $t_k$ is the $k$th discrete time (time sample), $N$ is the number of time samples in the signal, and $g_k(t)$ is a window operator centred on the $k$th time sample (Margrave et al., 2005). The window function is often chosen as a Gaussian shape (Margrave et al., 2005) and it’s job is to set to zero signal amplitudes that lie outside of the window. A Fourier transform is then applied to the windowed signal, and the output spectrum is written to the
Table 3. Summary of acquisition hardware and software.

<table>
<thead>
<tr>
<th>Component</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital recorder (16 bit / 48 kHz software)</td>
<td>Apple Logic Express 8</td>
</tr>
<tr>
<td>Digital recorder (hardware)</td>
<td>Apple macbook air</td>
</tr>
<tr>
<td>Microphone</td>
<td>Shure SM27-SC Cardioid Condenser on a suspension shock mount</td>
</tr>
<tr>
<td>Function generator (software)</td>
<td>SignalScope by Faber Acoustical</td>
</tr>
<tr>
<td>Function generator (hardware)</td>
<td>Apple iPad2</td>
</tr>
<tr>
<td>Function generator (24 bit A to D)</td>
<td>iMic by Griffin Technology</td>
</tr>
<tr>
<td>Preamp</td>
<td>Eurorack UB1202FX by Behringer</td>
</tr>
<tr>
<td>Output</td>
<td>16 bit aip format</td>
</tr>
</tbody>
</table>

In this way, a spectrum for each of the $N$ time samples is written to each row of the data matrix, where the frequency content is approximately local to each time sample.

**REAL DATA EXAMPLES**

The data acquisition system through which data were obtained in this study consists of a function generator, a fairly high-fidelity microphone, and a digital recorder (Table 3). Recordings are made in a relatively quiet room, and the drums are close miked. Proceedurally, each drum is miked first on either the batter or resonant side with a pillow placed on the drum head. A repeating sweep is generated on the function generator, amplified, and then broadcast through loudspeakers at the drum. The corresponding recording is the reference recording for that head for that drum. The pillow is then removed and a second recording is made. That recording is the monitor recording for that head for that drum. The drum is then flipped over and the process is repeated. The bass drum is naturally played on it’s side, so the recording procedure is modified accordingly.

The recorded format is based on an analogue to digital internal conversion first to 24bit and then desampled to 16bit / 48kHz for mac OS. Data are stored in a big endian format like the SEG-Y format but with much less header information. Data are organized into a header and two chunks. The header declares what kind of file (picture, audio, or video) and how much data to expect. The common chunk declares the signal length and sample rate, and the Sound Data Chunk contains the waveform data.

**Time domain**

The recorded drum signals are given in Figures 3 through 5. Beginning with the tom tom, we see that that the reference recordings for the batter and resonant heads (Figures 3(a) and 3(b)) are quite similar in amplitude with the resonant head peak amplitude about 1.4 dB greater than the batter head. This difference is probably due to a number of factors the most important of which is microphone placement. Because the batter and resonant measurements are not critical in an absolute sense for this experiment, no special effort was made to ensure identical placement head to head or drum to drum. Similarly, the monitor recordings (Figures 3(c) and 3(d)) are close, but here the peak amplitude of the
batter head is about 0.3 dB greater than the resonant head.

Reference to monitor comparisons of these data indicate that the presence of the drum unmuffled increases the recorded signal (care was taken to not alter mike placement between reference and monitor recordings). For the batter head the increase is 2.4 dB and it is 1.5 dB for the resonant side. This difference, though to a small extent due to subtle changes in the ambient noise in the recording studio, is due mostly to the response of the drum to the driving signal. The floor tom sweeps (Figure 4) indicate a change in the driving signal from a long sweep to a short sweep.† As for the tom tom, the presence of the floor tom causes an increase in amplitude of the recorded monitor signal relative the reference signal. The largest increase on the batter head with nearly 3 dB gained due to the presence of the drum, with a relatively small increase of 0.4 dB on the resonant side. Recordings for the bass drum (Figure 5) have the lowest amplitudes overall for the three drums owing, probably, to the location of the drum and microphone combination. Where

†This change, unfortunately, is an unaccountable function of the function generator, and it has prompted a search for a more controllable source.
FIG. 4. Reference and monitor recordings for the floor tom recordings. Amplitudes for the reference recordings for the floor tom differ overall (4(a) and 4(b)) with the resonant head 2 dB down overall from the batter head. Batter head monitor amplitudes (4(c)) are 3 dB larger overall than the reference batter (4(a)), and the reference and monitor amplitudes for the resonant head (4(b) and 4(d)) are similar overall.
the floor tom and tom tom were recorded in an upright position with the microphone overhead in a suspension mount on a stand, the bass drum was recorded on its side with the microphone suspended beside each head. Presumably, the orientation of the microphone/drum combination reduced the amplitude of the overall bass drum recordings.

Relative to the reference recording (Figure 5(a)), the batter head of the bass drum caused an increase of 3.4 dB (Figure 5(c)), and the resonant head caused an increase of 1 dB (Figures 5(b) and 5(d)).

**FIG. 5.** Reference and monitor recordings for the floor tom. Like the reference amplitudes of the tom tom (Figure 3), the reference amplitudes for the batter and resonant heads of the bass drum (5(a) and 5(b)) are similar overall but like the floor tom, monitor amplitudes are greater for the batter head than for the resonant head (5(c) and 5(d), approximately 3 dB greater).

**Frequency domain**

The reference and monitor recordings, and their cross-correlations, for the three drums are analyzed in the Fourier transform domain. This is done prior to analysis in the Gabor domain as a check on the overall spectral content of the recordings, whether there is significant 60 Hz line noise, and at approximately what frequencies are the fundamental resonances of each drum might be expected in the Gabor domain.

Spectra for the reference and monitor recordings of the tom tom (Figure 6) indicate that
the presence of the drum causes a significant overall increase in amplitude (as expected from the time domain analysis (Figure 3), with enhanced high frequency content.

The peak amplitude in the reference recording (Figure 6(a)) occurs at 74 Hz for both the batter and resonant heads (both muffled) and this is found to be consistent with the other the reference recordings for the floor tom and the bass drum; both have at 76 Hz (batter and resonant (Figures 7(a) and 8(a)) indicating a room resonance between 70 and 80 Hz. Note, the two Hz difference between the tom tom and the other drums is due probably to the different driving signal applied to the tom tom exciting a slightly different room resonance.

The peak amplitude for the monitor recordings differ significantly between the batter and resonant heads (Figure 6(b)) at 84 Hz and 116.3 Hz respectively as do frequency content above about 350 Hz. The presence of the floor tom causes a significant alteration in the frequency content of the recorded signal (Figure 7. The peak amplitudes show a greater increase relative to the increase for the tom tom, and the corresponding peak frequencies (76 Hz and 80 Hz for the batter and resonant heads respectively. Figure 7(b)) are close to the room resonance of 76 Hz (Figure 7(a)), with an increase in amplitude around 200 Hz. Similar to the floor tom, the presence of the bass drum significantly alters the spectrum of
the recorded signal. Peak frequencies of 60 Hz and 64 Hz for the batter and resonant heads respectively (Figure 8(b)) are found below the room resonance of 76 Hz (Figure 8(a)).

Cross correlation of the monitor recordings with their respective reference recordings appears to have the desired effect of revealing the impulse response of each drum (Figure 9) though it is clear from the reference spectra that the sweep spectra are not flat. Compared with the spectra of the monitor recordings (Figures 8(b), 7(b), and 6(b)), spectra of the cross correlations (Figure 9) - in particular the cross-correlation spectra of the floor tom (Figure 9(b)) and the bass drum (Figure 9(c)). Though frequencies of around 60 Hz for the bass drum and 80 Hz for the floor tom are reasonable, a tom tom resonant frequency of around 70 Hz is unexpected in that the tom tom is of smaller diameter, so according to equation 5, so long as the head densities and tensions are similar, the tom tom should have a higher resonant frequency. Also, it is customary for the tom tom to have the highest pitch - in the studio, the tom tom did sound higher in pitch than the floor tom and bass drum.

### Gabor domain

Analysis of the cross correlation recordings in the Gabor domain reveals fundamental modes for all three drums (Figures 10 through 13). Fundamentals for the floor tom (Figures 12(a) and 12(b)) and bass drum (Figures 13(a) and 13(b)) are quite obvious, and they are tabulated along with the interpreted tom tom fundamentals in Table 4. The tom tom fundamentals of 120 Hz and 118.4 Hz for the batter and resonant heads respectively were determined through analysis of Figures 10 and 11. Gabor spectra for both heads on all three drums were annotated with a line that indicates the frequency of the peak amplitude a given time in the decay of the signal. For the floor tom and bass drum (Figures 12 and 13, these lines are straight and correspond to the interpreted fundamentals. For the tom tom, the

<table>
<thead>
<tr>
<th>Drum</th>
<th>Batter (Hz)</th>
<th>Resonant (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom tom</td>
<td>120</td>
<td>118.4</td>
</tr>
<tr>
<td>Floor tom</td>
<td>80.2</td>
<td>75.9</td>
</tr>
<tr>
<td>Bass drum</td>
<td>59.8</td>
<td>75.9</td>
</tr>
</tbody>
</table>

Table 4. Fundamental frequencies interpreted from Gabor domain analysis.
FIG. 9. Modulus of cross-correlation spectra. Peaks in the spectra of the tom tom occur at 74 Hz (batter) and 72.8 Hz (resonant) 9(a). For the floor tom 9(b) and bass drum 9(c) they are 80 Hz (batter) and 76 Hz (resonant), and 60 Hz (batter) and 76 Hz (resonant) respectively.
lines are crooked and indicate a confusion of various peak frequencies. Looking closely, however, reveals constant tones that persist on both the batter and resonant heads. A zoom in of the spectra reveals that indeed there are persistent tones at 120 Hz (batter) and 118.4 (resonant) that support the fact that the tom tom sounds high in pitch. Note, the relative low amplitude of the interpreted fundamental for the tom tom is possibly due in part to the age of the heads; both the batter and resonant heads are known to be considerably older than those for the floor tom and bass.

![Figure 10](image.png)

**FIG. 10.** Tom tom Gabor transforms. Blue lines indicate peak amplitude as a function of time and frequency. For the tom tom batter 10(a), the median frequency over time for the peak amplitude is 74.3 Hz, and for the resonant side 10(b) it is 101.7 Hz. Residual reference energy and very-strong overtones are present with nearly equal strength on both the batter side and the resonant side.

![Figure 11](image.png)

**FIG. 11.** Zoomed in tom tom Gabor transforms. In this zoomed in version of Figure 10, the median frequency for the batter side is 120 Hz 11(a), and the resonant side 11(b) it is 118.4 Hz.

**DISCUSSION**

It is curious to consider how high the resonant side head is tuned on the bass drum (76 Hz) relative to the batter side (60 Hz) where tuning for the tom tom and floor tom is quite consistent with the resonant sides generally being lower. If we assume that, because the player/tuner is closest to the batter heads, batter heads carry the most sound to the person tuning the drums, then it is reasonable to form musical ratios based on the batter heads alone. From this argument, then, it is found that this drumset is tuned to three of the most pleasing musical intervals as tabulated in Table 5.
FIG. 12. Floor tom Gabor transforms. Blue lines indicate peak amplitude as a function of time and frequency. For the floor tom batter 12(a), the median frequency over time for the peak amplitude is 80.2 Hz, and for the resonant side 12(b) it is 75.9 Hz. Overtones are strongest on the batter side.

(a) Floor tom Gabor transform (batter).  (b) Floor tom Gabor transform (resonant).

FIG. 13. Bass drum Gabor transforms. Blue lines indicate peak amplitude as a function of time and frequency. For the bass drum batter 13(a), the median frequency over time for the peak amplitude is 59.8 Hz, and for the resonant side 12(b) it is 75.9 Hz. Overtones are strongest on the resonant side.

(a) Bass Gabor transform (batter).  (b) Bass Gabor transform (resonant).

Table 5. Table of ratios of drum fundamentals and corresponding classical musical intervals.

<table>
<thead>
<tr>
<th>Drum pair</th>
<th>Ratio</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom tom / bass drum</td>
<td>120 / 60</td>
<td>2:1, Octave</td>
</tr>
<tr>
<td>Tom tom / floor tom</td>
<td>120 / 80</td>
<td>3:2, Perfect fifth</td>
</tr>
<tr>
<td>Floor tom / bass drum</td>
<td>80 / 60</td>
<td>4:3, Perfect fourth</td>
</tr>
</tbody>
</table>
CONCLUSIONS

The development above reveals that the Gabor domain has great utility in acoustic analysis beyond the familiar seismic applications. In particular, it was central to the determination of the resonant frequency of the tom tom as that frequency was lost in the room resonance in the conventional Fourier domain. Overall, the tones and tonal relationships of a particularly well sounding set are found to correspond to three of the most pleasing intervals of classical music: octave, perfect fourth, and perfect fifth.

ACKNOWLEDGEMENTS

I wish to thank the staff and sponsors of CREWES for their support. I also thank NSERC for their support of this work through CRD grant CRDPJ 379744-08.

REFERENCES

Pierce, A. D., 1994, Acoustics: an introduction to its physical principles and applications: Acoustical Society of America through the American Institute of Physics.