Comments on Banjo Head Tapping, Tuning, and Tension

David Politzer*

California Institute of Technology

(Dated: July 31, 2022)

Some aspects of banjo head tapping, tuning, and tension are reviewed and demonstrated. Most or all are familiar to experienced players. Measurements made from tap recordings fit a simple theory. The discussion of the physics is entirely qualitative but does shed some light on both the challenges and value of tap tuning in relation to head tension. The quantitative details and underlying physics bear a complicated relation to the perception of pitch.

* politzer@theory.caltech.edu; http://www.its.caltech.edu/~politzer; 452-48 Caltech, Pasadena CA 91125; *452-48 Caltech, Pasadena CA 91125*

Comments on Banjo Head Tapping, Tuning, and Tension

I. INTRODUCTION

Head tension is easy to adjust on many banjos. This produces a considerable range of how plucking converts into sound. Certainly, players should experiment to find the tension that sounds best to them. But often, someone may want to reproduce a successful set-up or to emulate a desired sound. Giving and receiving advice require some way to characterize a target tension that can be implemented directly.

In the world of standardized mylar heads and resonator backs, very fine distinctions are commonly made. Distinctions with open-back banjos are typically made with somewhat less precision, and going to skin heads necessarily blurs the issues even more. In all these cases, the static deflection produced by a localized downward force is a common tension measure. In ascending order of precision, it could be a thumb near the rim, the bridge itself (with strings tuned) relative to a straight edge set across the head, or a dial gauge (DrumDial being a commonly used commercial version).

A very different method involves identifying a particular pitch produced when tapping on the head. If it can be done at all, it can be done quickly and with considerable precision after a single round of tightening or loosening. In contrast, listening to some playing and comparing to the previous round requires considerably more time and expertise. Where and how to tap is a learned skill. Identifying a single pitch is also a learned skill — and it can be beyond a particular person's ability. The underlying physics connects all these issues.

The perceived pitch of a tuned string is simply related to the frequency of its lowest mode of vibration. They're equal. A given string is tuned by varying its tension. In contrast, the lowest vibrational frequency of a particular banjo head depends on several factors beside its tension. Pot geometry is crucial, and other aspects (e.g., string tuning, break angle, and bridge weight) contribute, too. However, a proper head tap tone depends on the head tension, diameter, and weight and not on the other aspects of the banjo.

So, the tap tone is an unambiguous measures of tension, but its relation to a particular banjo's sound will be tangled up with all the other parameters. At a minimum, the impact of a given head tension will depend not only on head diameter and density but also on string tuning and break angle, bridge weight, and pot design.

Banjos produce a huge range of many component frequencies simultaneously with each pluck or head tap. Physics can say something about the numbers, but brains turn all that into perception of pitch and timbre. Furthermore, for a given note, the banjo component frequencies are not all related as they are in “fine” musical instruments. Some people’s brains turn banjo sounds into music. Others resist. Distinguishing perception from physical measurement is crucial to making some sense of it all. In the 21st Century, the ubiquitous three-color screens provide visual examples of a somewhat analogous distinction between perception of color and light frequency.

II. OUTLINE

The relation of pitch to frequency can be quite complex. The phenomenon of the “missing fundamental” is a simple, illustrative example. It is demonstrated here with plucks on a banjo 4th string. For the lowest half-octave, the perceived pitch is not one of the frequency components of the sound — at least not in the sense of Fourier components and spectrum. Nevertheless, that sound is periodic with a period that matches the perceived pitch. The case of head taps is yet more complicated. The frequency components are not multiples of the perceived pitch frequency, as they are with the plucked string. Nevertheless, their combined effect has a periodic aspect that determines the pitch. Furthermore, brains often try to identify a unique pitch even when nothing is “mathematically” precise. One very clean example illustrates the sound and quantitative analysis of taps with a particular set-up on a resonator banjo.

It was suggested in ref.s [1] and [2] that a proper tap excites a whispering gallery mode[3] of the head, and the period of that motion is heard as the tap pitch. The name comes from a classic example in architecture, but the basic phenomenon underlies devices in laser optics and precision quantum measurement. The earliest physics accounts got into the details of the specific system, in particular, the sounds inside hemispherical domes. Analogously, in the present context, it helps to look at some of the physics specific to the ideal circular drum head. However, the phenomenon is far more general, and that general perspective offers a far simpler and less mathematical perspective.

For details specific to head tapping, the ideal circular drum head is a good place to begin. Then air, bridge, and strings can be added. All of those impact the vibrational resonant frequencies of the head and are reflected in the frequency and strength of Fourier components of the sound of a plucked string. However, the pitch of head vibrations produced by a proper tap depends only on the head itself, i.e., its tension, density, and diameter. This is magic and definitely makes tapping a brilliant measure of tension. These taps launch the whispering gallery modes. Their motion and sound are unlike anything produced by a plucked string.

Physics cannot determine that one sound is better than another. But it can describe how some parameters combine to produce a given aspect of the sound. For example, physics can address the relation of tapping on different size heads (e.g., flat-top, arch-top, and 12″).

III. THE MISSING FUNDAMENTAL

Air pressure that varies sinusoidally as a function of time produces a sensation known as a pure tone, and we identify the single pressure frequency with the pitch of the tone. It can be produced electronically, and some musical instruments come close to producing pure tones. Strings and pipes naturally generate sounds with a spectrum of frequencies that are very close to integer multiples of the lowest one, which is known as the "fundamental." The perceived pitch is the same as would be heard if only the fundamental frequency were present. The higher harmonics determine a timbre.

If the sound is missing its fundamental but has more than one higher harmonic (preferably several), then the pitch of the "missing" fundamental is still what is heard. This allows small speakers to produce some semblance of low notes. It also is a feature of most banjos and many guitars. In the physical examples, the fundamental is not completely missing, but it is inordinately weak compared to the higher harmonics.

The mass on a spring in high school physics can be driven by a sinusoidal force of any frequency. The long-term response of the mass is largest when the drive matches the mass-spring resonant frequency. However, there is always long-term motion at the driving frequency — but just smaller in amplitude when not at resonance. In the spring-mass system, the response is non-zero no matter how small the driving frequency. Fig. 1 displays the qualitative shape of the amplitude as a function of driving frequency of the steady-state mass motion arising from an oscillating force of a given magnitude. (The parameters and units used for Fig. 1 are random.)

The lowest octave chromatic scale on an open-back banjo was plucked and recorded. Fig. 2

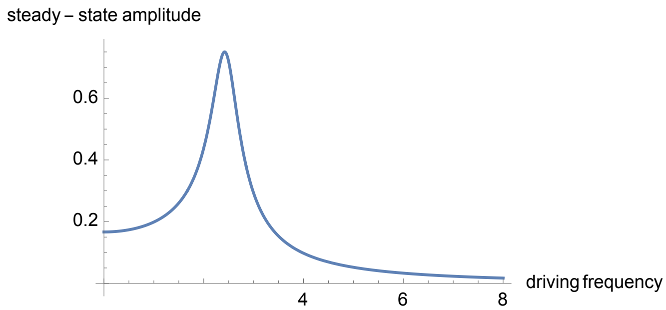


FIG. 1. the shape of amplitude *vs* frequency for a driven harmonic oscillator

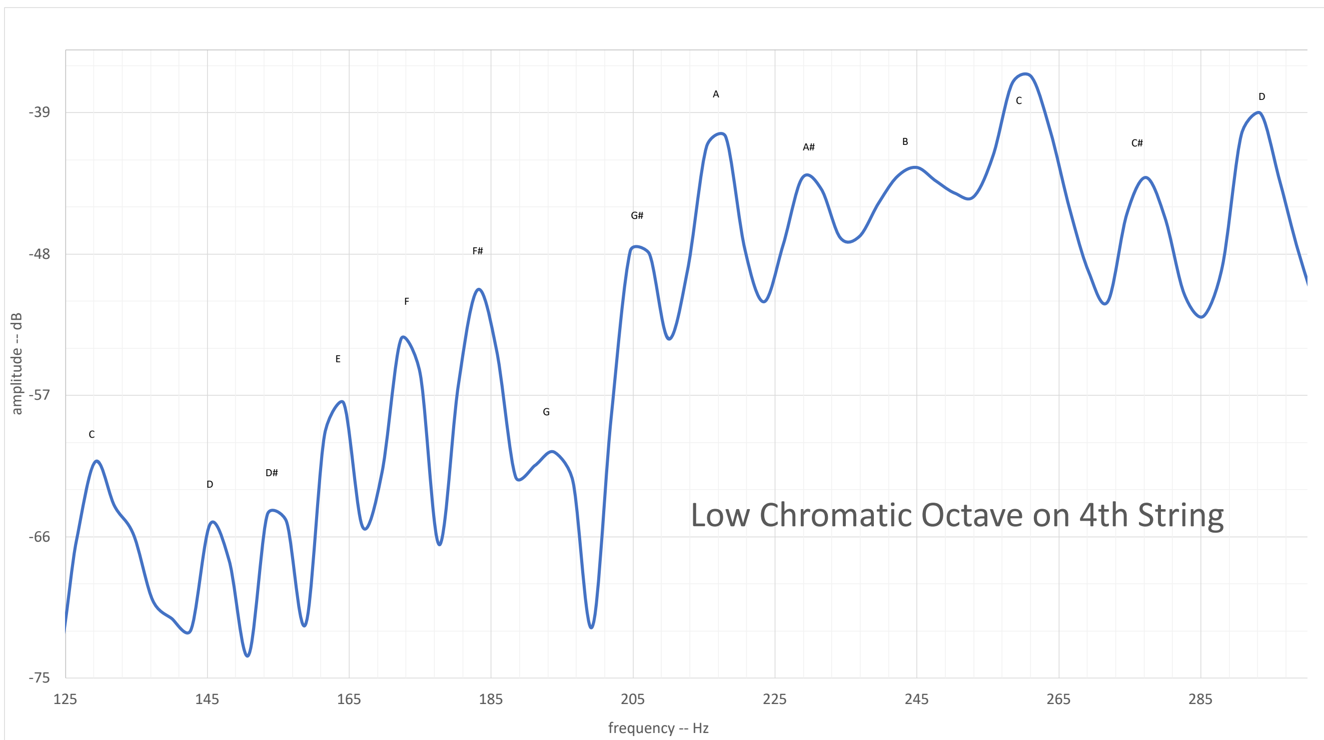


FIG. 2. spectrum of the 4th string lowest chromatic octave

is the spectrum of the entire twelve-pluck recording. There is nothing odd about the sound of the low notes, but the spectrum gets dramatically weaker below 200 Hz. Presumably, that is the lowest resonance of the head. Driving the head by the string at frequencies below that lowest head resonance produces weak head response at the fretted string fundamental frequency. The lowest C is heard clearly, but the spectrum amplitude there is down by about 25 decibels compared to the upper half of the octave.[4] Similarly, the low C# and D sound fine, curtesy of their very many strong, higher harmonics. (Those harmonics were produced by fretting at the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ frets and *not* the 13$^{\text{th}}$ and 14$^{\text{th}}$.)

IV. PROPER HEAD TAP SOUNDS

Reproduced here is some very successful head tapping on a Deering Sierra for an appendix of the *Pickers' Guide...* in April 2021.[1] The Sierra is an 11″ flat-top with a top-frosted mylar head and a resonator back.

This link is to the sound of a single tap: http://www.its.caltech.edu/~politzer/head-taps/whisper-1-just-one.mp3 .

Fig. 3 is a spectrogram of that tap sound. Fig. 4 is its waveform. And Fig. 5 is a spectrum computed for thirty essentially identical taps. In the spectrogram, the linear vertical scale is frequency in Hertz; time is displayed horizontally in fractions of a second; and spectral intensity is represented by a grey scale, with black as the highest. The waveform is pressure amplitude *versus* fractions of a second. The spectrum was computed from a recording of

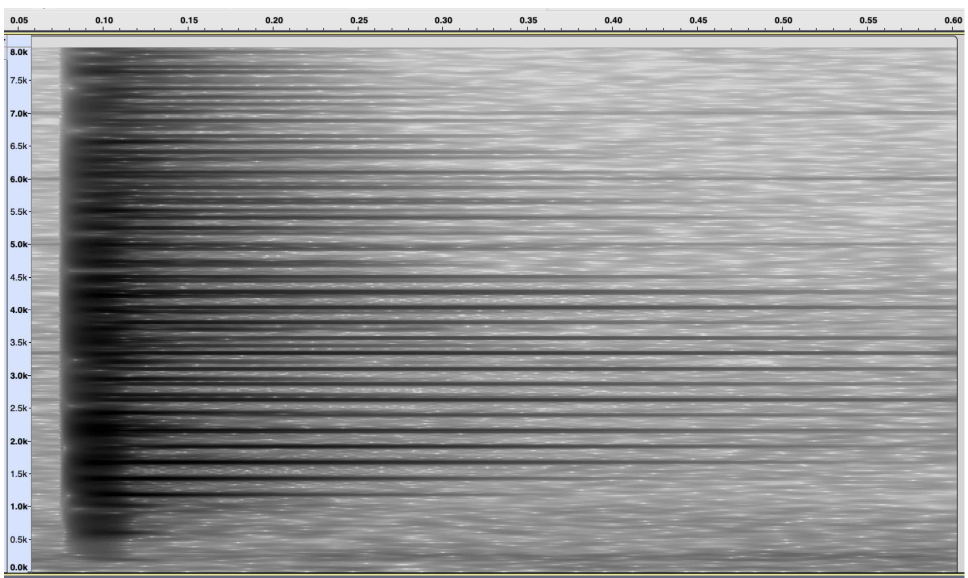


FIG. 3. single head tap spectrogram

thirty consecutive taps; hence, it averages over the individual tap slight variations. The spectrum and spectrogram are related in a simple way. The spectrum adds up the intensity for each tiny frequency interval over the duration of the recording being analyzed and plots that sum as a function of frequency.

In the spectrogram, we see that much of the strongest sound occurs at roughly equally spaced frequencies. This is particularly true of the longest lasting sound. That spacing is approximately 237 Hz. In the wave form, we see that after the initial attack, the amplitude

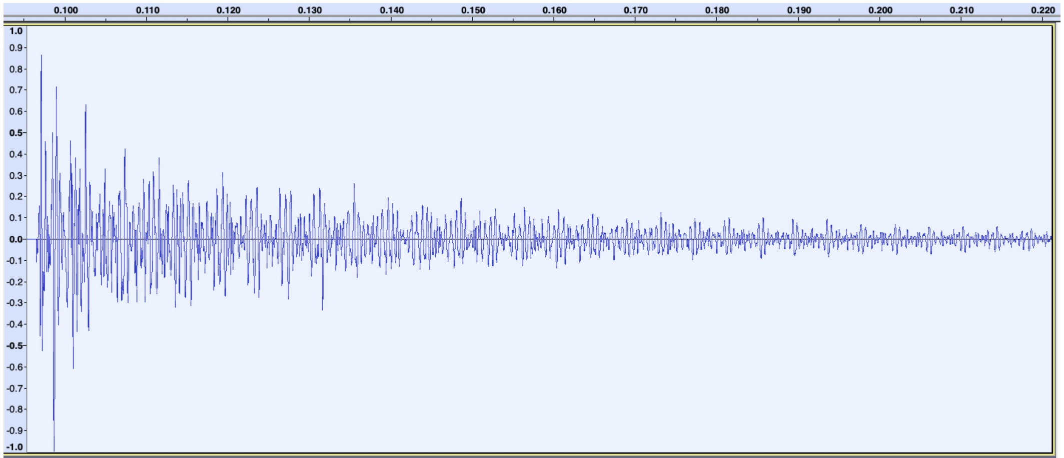


FIG. 4. single head tap waveform

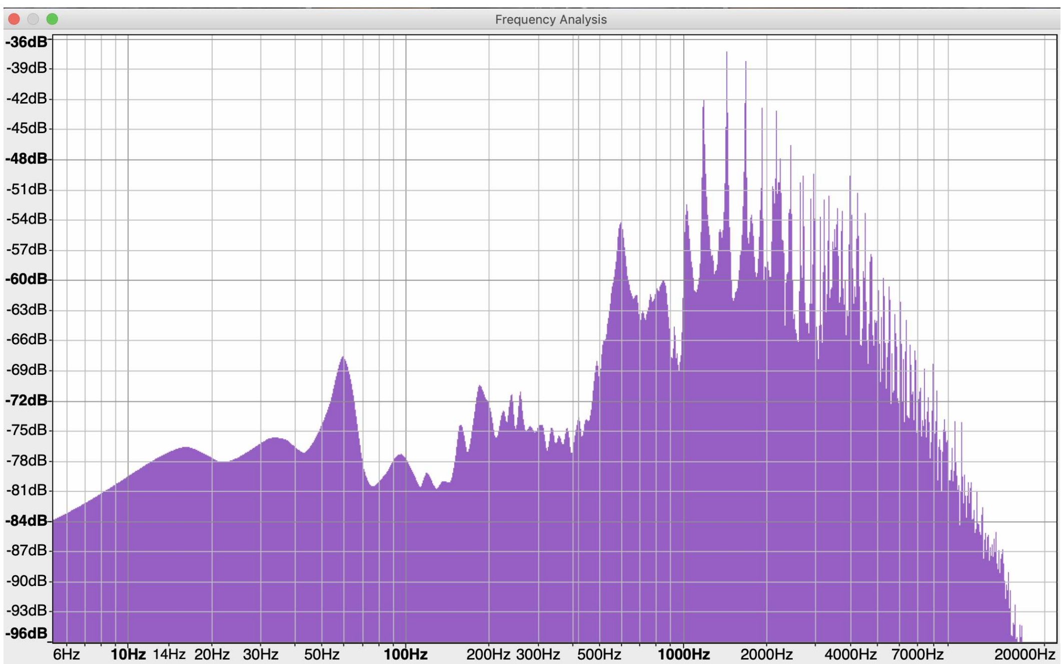


FIG. 5. spectrum computed from 30 consecutive taps

gets larger and softer with a fairly regular period while it decays on average. I count 29 cycles in 0.121 seconds. That corresponds to 240 cycles per second. However, the structure within cycles does not repeat from one to the next. So, the waveform is nothing like a periodic function of period 0.0042 seconds multiplied by a decaying envelop. The spectrum shows a fat peak at 60 cycles per second, curtesy of the electric company. The peak around 190 Hz is the lowest mode of the head, i.e., corresponding to an up-down motion over the entire head, vanishing only at the fixed rim. Significantly, there is no strong spectral peak anywhere near 240 or 480 Hz. Consistent with the spectrogram, the really strong peaks begin around 1200 Hz and continue to 6000 or 7000 Hz.

This is a sound that some people can identify as $A^{\#}$. (With $A_4 = 440$ Hz, $A_3^{\#} = 233$ Hz in equal temperament.) There is certainly something $A^{\#}$-like in the quantitative versions of the recording. Explaining why it sounds like $A^{\#}$ would require serious brain science.

It may well be that a sequence of strong, approximately equally spaced frequencies can produce a perception of a definite pitch. But there are obvious challenges to producing and hearing such a sequence. Good high frequency hearing is a prerequisite because frequencies much higher than the perceived pitch are necessary. When tapping, proper technique is required to avoid making lots of other sounds in addition. Some of those other sounds are inevitable. So, hearing the tap pitch requires some amount of "hearing out." The most common example given is the ability in a noisy space to focus on a single person's voice and hear what they're saying. Musicians can often hear out individual notes in a chord or even harmonics within a note. An orchestra conductor can choose to listen to individual instruments or to the whole ensemble. These are all skills that rest on some combination of talent and experience. Hearing a banjo head tap pitch is a similar challenge.[5]

An electronic tuner faces similar challenges. Its success at identifying the tap tone does not depend on its sensitivity or accuracy but rather on the programmed algorithms used to display a single pitch from a sound with many component frequencies.

V. WHISPERING GALLERY MODES

In its most general form, a whispering gallery mode is a vibrational motion of some medium that is confined by a concave surface. The motion is restricted to be very close to the surface by its relation to or interaction with the surface. For a traveling wave (or disturbance, pulse, or wave packet), the wave is guided along by the surface. The surface prevents the wave from spreading outward while its curvature prevents the wave from spreading inward. Some researchers have described the phenomenon as a continuous (or repeated) total internal reflection. That perspective explains (and quantifies) that the curvature cannot be too sharp relative to the tangent direction. If the curvature were too sharp, the wave would reflect inwards and, in cases of transparent media (such as a glass-air interface) also escape outward. If the surface is too flat or even convex, the wave motion would spread into the interior. For a three dimensional medium, the surface must be concave in both its directions. Otherwise, the motion would spread in the direction that is not concave. Note that a stationary wave, pulse, or disturbance can be represented as the sum of waves traveling in opposite directions.

The eponymous whispering gallery is in St. Paul's cathedral in London. Standing at the wall inside the huge dome, one can whisper and be heard halfway round the dome. Rayleigh was the first to give a correct physics account[3], and he explored the details using properties of Bessel functions, which are part of the solution to the wave equation with a spherical boundary.

The fixed edge of a banjo head can also guide a whispering gallery mode. Again, some of the details can be illuminated using the properties of Bessel functions. Actually, the general, qualitative description and the Bessel functions provide complementary understanding and details.

A. proper tap technique

A recommended method is to tap smartly near the head edge with the eraser end of a wooden pencil. (Some people tap with a metal finger pick; some scratch with a fingernail.) If too close to the edge, the rim limits the motion to an extent that there's not much sound. As the tap point moves inward, the sound becomes louder and richer. However, "richer" means that other types of motion are produced as well. These can confuse the brain that is trying to identify a single pitch. Multiple pitches may be heard. So, a practical compromise must be made regarding the location.

Several taps in rapid succession help overcome the fact the single tap sounds are very short.

If there is a resonator, it should be on the banjo. The wood back does not absorb as much sound as a player's belly, especially above 4 kHz.[6]. On this issue, an open-back on a stand with the back open to the air does worst. It could get a hard back simply by placing it on a hard surface.

The strings are usually damped to avoid extraneous sound (but see ref. [5]). If the strings are damped, it helps considerably to also damp the bridge with your hand. The desired head motion is near the edge. Damping the bridge can reduce the motion towards the middle, thus making the pitch less ambiguous.

B. simple, general picture

The tap creates a local distortion of the head — a lump or a pulse. It has no preferred direction of motion along the head surface. Just like a pluck of a string, the distortion does not actually spread out. Rather, it splits into two identical parts going in opposite directions, each with half the amplitude. These pulses travel at the speed of transverse waves in the head. That wave speed c is independent of frequency and geometry for an ideal drum head and is given by $c = \sqrt{\frac{\tau}{\rho}}$ where τ is the tension (a force per unit length) and ρ is the mass per unit area. For head diameter D, each half of the pulse goes around the head, i.e., a distance equal to πD, at speed c. Hence the frequency of the motion is $c/(\pi D)$.

C. Bessel functions, normal modes, & a more detailed and accurate picture

The possible motions of an ideal circular drum head, vibrating on its own, i.e., without additional extra forces, can be expressed as sums or superpositions of "normal modes." For these motions, each part moves at the same frequency. Each normal mode has its own characteristic frequency. The normal mode motions have different amplitudes at different points. One characteristic feature is the node lines: fixed lines on the head at which the amplitude is zero. Normal modes are the motions that arise after the drum is disturbed and then left free to vibrate while it dies down. When a driving force is applied at the natural frequency of one of the normal modes, the resulting motion will include that mode with a large amplitude. The extra mass and forces of bridge and strings alter the shapes and frequencies of the normal modes. Those effects are largest for the lowest frequencies and in the immediate region of the bridge feet. Whispering gallery modes live near the edge, away from the bridge and, as suggested by Fig. 3, involve normal modes with high frequencies, with little contribution from the lower ones. So, the drum without bridge and strings is a simple place to start.

The normal modes of the ideal circular drum head are labeled by two integers: the number of circular node lines and the number of straight-line nodes that are diameters equally spaced in angle. There is always at least one circular node line at the edge. The radii of the other circular node lines can be expressed in terms of the famous Bessel functions.[7] And all the resulting resonant frequencies can be calculated in terms of Bessel functions.

The drum head normal modes are often listed and pictured by ascending frequency. For many applications, that is exactly what is needed. Of interest are all motions up to some highest frequency determined by the practical situation. For the whispering gallery discussion, it is appropriate to group the normal modes by the number of circular node lines. For each circle number, there is an infinite list of modes with 0, 1, 2, 3, 4,... diameter node lines.

The shapes of drum normal modes are products of a function of distance from the center times a function of the angle going around. The angular functions are simply sines and cosines. The radial functions are again Bessel functions.

The drum normal frequencies, listed in ascending order, show no regular patterns. That agrees with the notion that the ideal drum has indefinite pitch. At high frequencies, the drumhead is divided by the many circular and diameter node lines into lots of patches that move in opposite direction to their immediate neighbors.

However, a very distinct picture emerges when we group the modes by number of circular node lines. Consider the case of no circles besides the one around the edge. The diameter lines are equally spaced in angle, and the distance between them goes to zero towards the center. Basic physics implies that the force on a patch of head is proportional to its curvature. Hence, for a given mode, the amplitude must decrease as the node lines get closer together and vanish in the center. This happens more and more dramatically with increasing number of diameter node lines.

Fig. 6 shows the normal mode amplitudes as a function of radius, i.e., distance from the center, for n diameter lines, and 0 and 1 circular nodes besides the rim.

Clearly, high values of n have their motion restricted to near the edge. An arbitrary pulse shape near the edge would require superposing various combinations of numbers of circle nodes. However, if the shape is roughly just a lump, i.e., goes up and comes back down, it will be almost entirely made up of the no-extra-circles modes.

The frequencies of the sets of modes organized by circle line number show a clear pattern and simple behavior for large numbers of diameters and high frequencies. This is illustrated in Fig. 7.

The top graph of Fig. 7 shows that, for a given number of circles, the frequency as a function of the number of diameters, n, becomes very linear. That means that the spacing between successive values of n approaches a constant. The lower graph highlights this

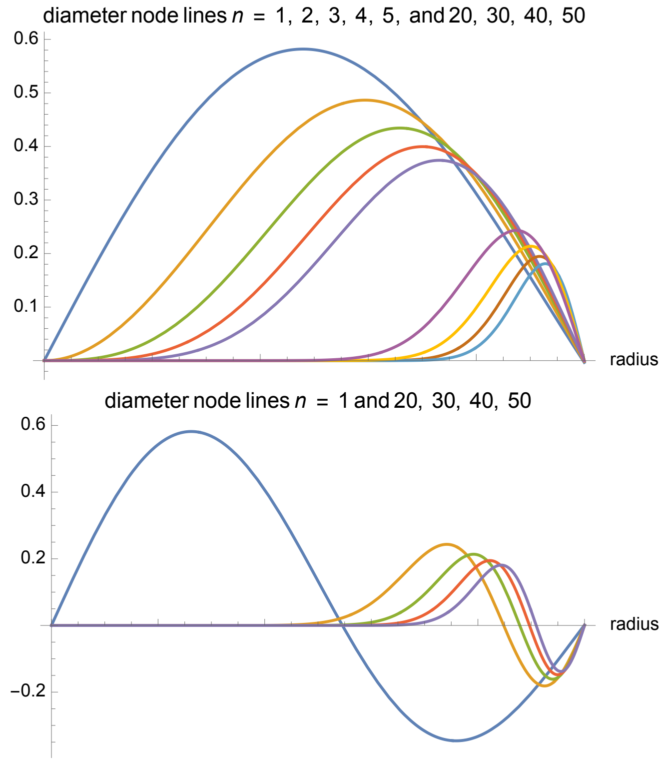


FIG. 6. radial amplitude for 1 and 2 circular nodes and n diameter nodes

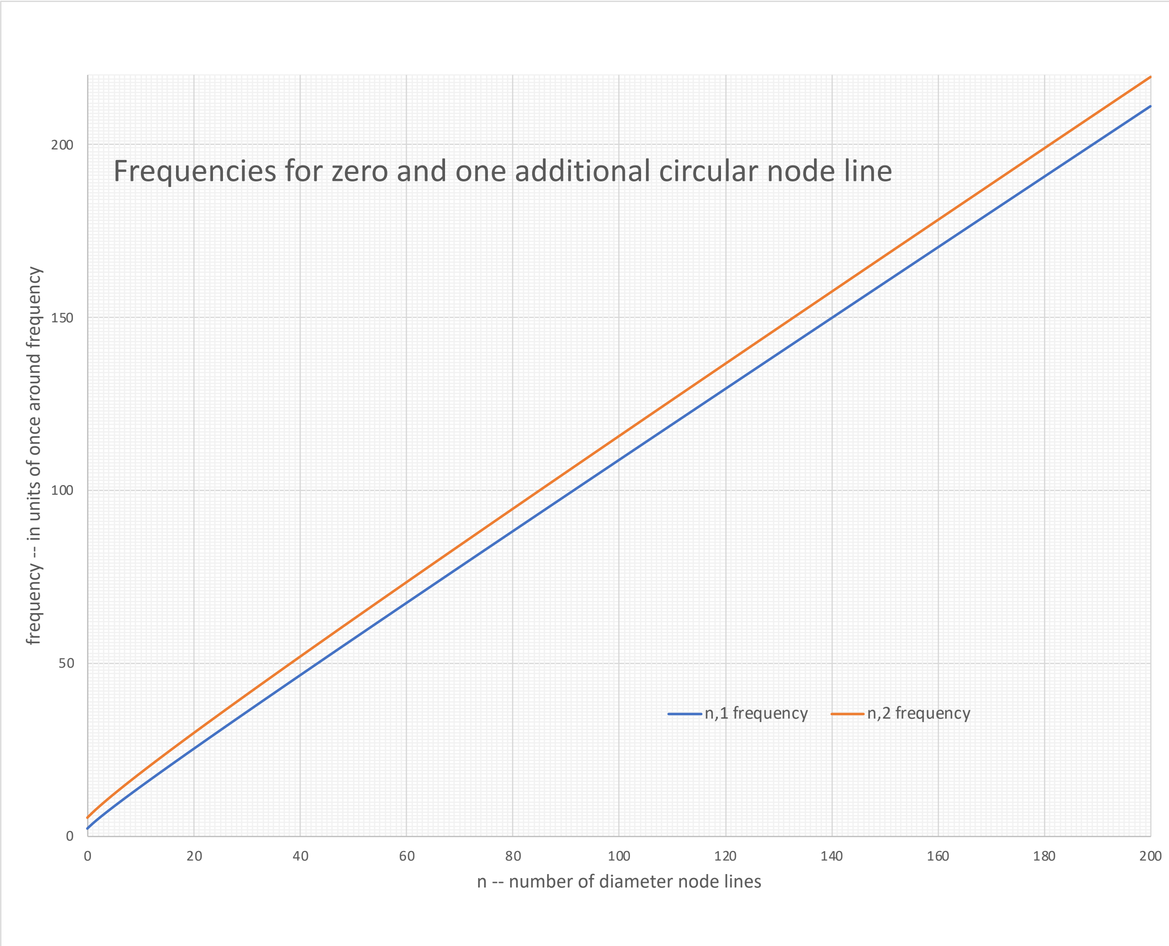
feature. The asymptotic frequency spacing is 1 in units of $c/(2\pi R)$, where c is the wave speed and R is the outer radius. (Note that the vertical axis goes from 1.0 to 1.5.)

For a pulse to propagate at a well-defined speed and to retain its shape, it is essential that wave speed be independent of frequency. That is true for the ideal drum head — and also true for sound in air (at acoustic frequencies).

D. frequency of the whispering gallery mode

The Bessel function calculation for the ideal circular drum gives the same frequency as the general argument for waves scooting around the inside of a concave surface at the wave speed because the one-cycle travel distance is the circumference, $2\pi R$.

And the Bessel function evaluation offers a further insight. In practice, we are not dealing with $n \to \infty$. From Fig. 3, we might imagine that the values of interest are something like $5 \leq n \leq 30$. So, we expect the spacings in the spectrogram and spectrum to not be exactly equal and to decrease with increasing frequency. If we make an estimate by counting the spacing for several successive ones and dividing by the number, the result will be larger than



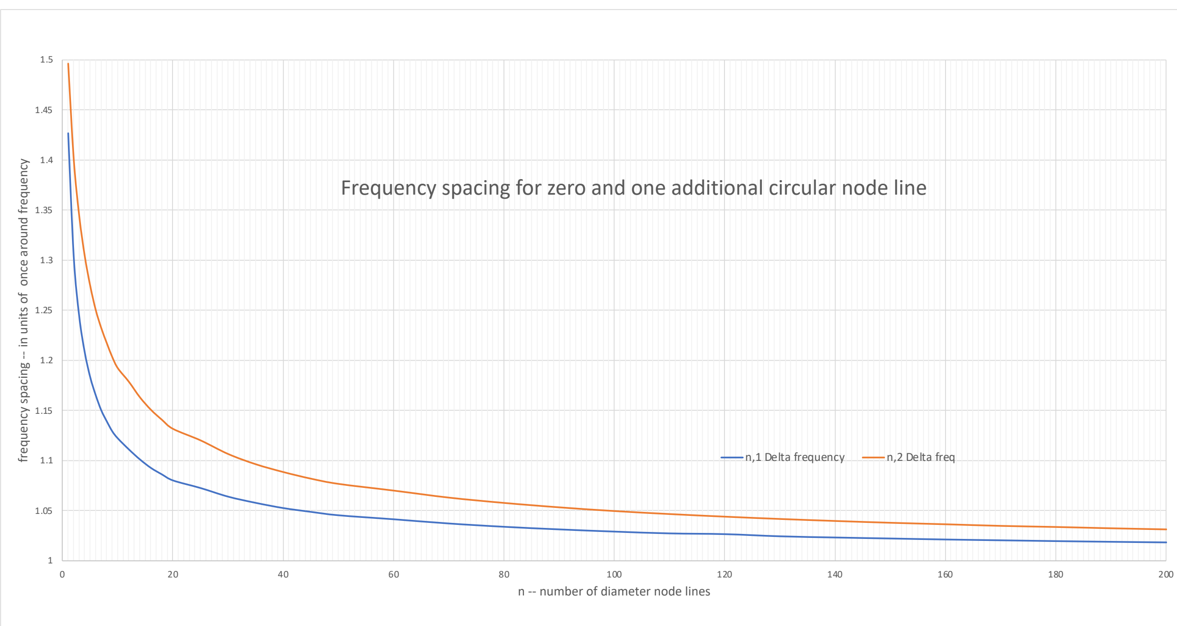


FIG. 7. frequencies and successive frequency spacing for 1 and 2 circles and n diameters

the $n \to \infty$ limit and will depend slightly on how they are chosen. What the brain does with this pattern is up to the brain and not mathematics [8], but it usually attempts to identify a particular pitch. Fig. 6 also gives a sense of where the superposition of the observed strong frequencies resides relative to the edge of the head. Both of these effects of $n \not\to \infty$ impact the relation of tension to measured frequency spacings. Examples are seen in section §VIII.

VI. THE PHYSICS OF AIR LOADING

The air that surrounds the head must move when the head moves. Most of the effect of that air on head motion can be expressed as an added inertia when the head moves in air compared to moving in vacuum, e.g., under the forces of the strings and its own tension.[9] The relevant banjo modeling, calculations, and references are provided in the second paper of ref. [2], section §2.3. Precise characterization of air loading is best expressed on a mode-by-mode basis. It is found to be a small fraction, e.g., typically around 0.14, of the mass of the air above the banjo within a distance of one air wavelength at that frequency. The actual fraction depends weakly on the actual mode number but quickly approaches a limiting value for high frequencies. The limiting value can be expressed in terms of the speed of sound in air and the rather slower speed of transverse waves across the head, a result that can be derived analytically for idealized geometries. For the lowest few modes, the air inertia is comparable to the mass of the head itself.

This air loading can be pictured qualitatively. The air at the head surface acts like a blanket that must go up and down as the head vibrates, and that blanket effectively adds inertia to the head motion. But how thick is the blanket? Because the vibration is characterized by a frequency, the air motion has its own characteristic wavelength at that frequency. The lower velocity head waves have a shorter wavelength. As an elastic medium, the air really cannot support variations over those head lengths that are shorter than the air natural wavelength. Hence, the air motion that directly follows the head motion has to die out away from the head, and that happens on the scale of the air wavelength.

The crucial take-away for head taps near the rim is that air loading is negligible. That is because the wavelength in air decreases with frequency f like $\frac{v_{\rm sound}}{f}$, and the added inertia is effectively that of a blanket of air whose thickness is $\sim 0.14 \times \frac{v_{\rm sound}}{f}$.

A. bridge or no-bridge?

When bridge and strings are mounted on a head without altering the tension hooks, the shapes and frequencies of the low frequency modes certainly are altered. It stands to reason that the overall tension increases as well. However, in the context of the measurements I did for the present effort, that increase for almost all of the way around near the rim was no greater than the ∼2% tension resolution of a DrumDial for a typical set-up on the 12″ pot. On the 8″ pot, the increase seemed to be about equal to that resolution value. Since DrumDial measurements are typically taken around the edge and that is where the whispering gallery modes live, it seems that bridge-on or bridge-off makes no practical difference in the tap pitch. As mentioned above, bridge-on with the bridge damped by hand (in addition to damping the strings) suppresses some of the head vibrations that can confuse tap pitch recognition.

VII. EXAMPLES FROM THE 12″ HEAD AT 92 DD

Taps on the 12″ head set to 92 on the DrumDial provide examples of some of what's been described so far. In particular, the comparisons include 1) without strings with the back completely open to the air, 2) without strings and backed by wood with an air gap producing a "sound hole" air comparable to the back and flange of a resonator banjo, and 3) the backed rim with strings tuned to pitch and damped.

The frequencies of the lowest head vibrational mode in each of these three cases are easy to identify from the spectrum of the sound of any sort of tapping. The head motion in those lowest modes is certainly just up and down in unison over the whole head. But head tension, diameter, and mass are not the only relevant factors. Air adds to the inertia of the motion in a way that depends on the resulting frequency and the details of the shape across the head. The contribution from the air under the head also depends on details of the pot's geometry. Head tension is, indeed, one of the primary determinants of a banjo's sound, but this lowest frequency is not a very useful way to characterize the banjo's overall sound or to determine the head tension. In fact, the banjo's sound spectrum is so rich and goes up to such very high harmonics that the exact details of the low frequency sound production is of minor importance.

The observed values of those lowest frequencies for the three situations were 313 Hz for the open, bare head; 218 Hz for the bare head backed by wood; and 428 Hz for the strung up and tuned banjo with the wood back.

However, tap pitches can be estimated from the spectrograms to be 217 Hz for the open, bare head; 219 Hz for the bare head backed by wood; and 219 Hz for the strung up and tuned banjo with the wood back.

The tap tone theory explained above, the observed value of the tap pitch for these setups, and the standard theory for the ideal drum's lowest mode can be combined to estimate the amount of air loading on these particular modes. For example, for the open back, bare head, the extra inertia due to the air is an additional 67% of the bare mylar inertia.

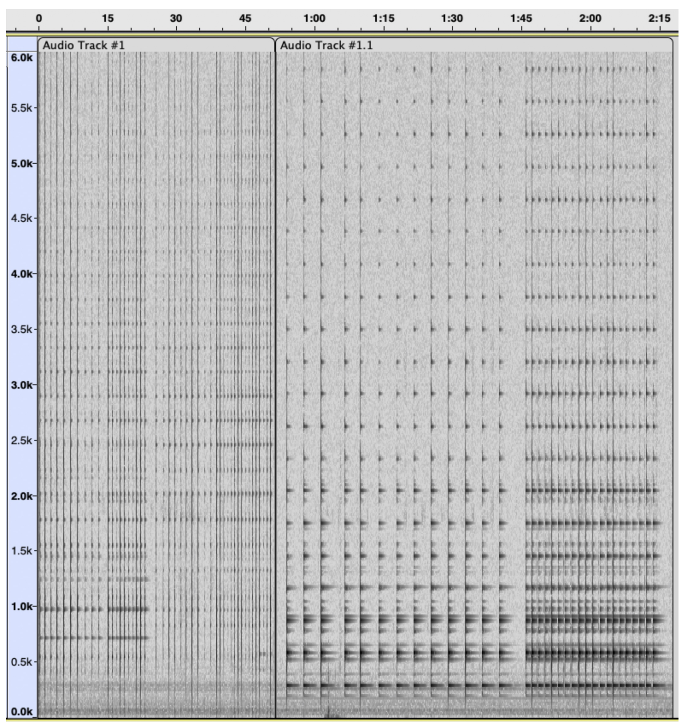


FIG. 8. 12″ 92DD damped string taps (on left) & 1^{st} string plucks (on right)

To put the tap spectrograms and spectra in perspective, Fig. 8 is a spectrogram and Fig. 9 the associated spectrum of head taps with damped strings and a pluck of the 1^{st} string with all strings open. From the average spacing of prominent tap frequencies between ∼1800 Hz and ∼5000 Hz, the tap pitch is around 219 Hz. The plucked 1^{st} string was tuned to D_4, i.e., approximately 294 Hz. There are actually two distinct sets within the tap series. The first one, shown on the left in Fig. 8, had strings damped. The second one that follows also has damping by hand on the bridge. That bridge damping muffles much of the sound that is heard, but close inspection confirms that the strong sound without bridge damping is not at the frequencies of the whispering gallery mode. In fact, bridge damping enhances the frequency content of the whisper. (The two string pluck sets differ only by the pluck tempo.) Note that 12dB was added to the tap spectrum in Fig. 9 to offset it from the pluck for better visibility.

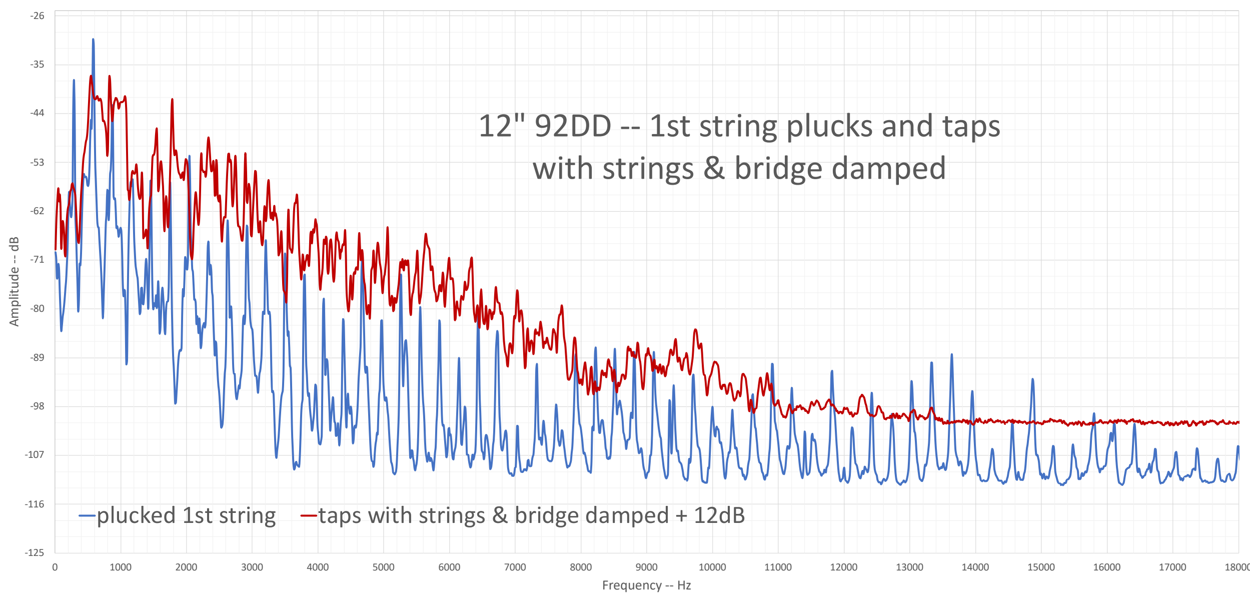


FIG. 9. 1st string pluck spectrum and tap spectrum (shifted up 12 dB) on same head

Some features are easier to see in the spectrograms, and others are clearer in the spectra. Both are computed from the same recorded data.[10] It is possible that the brain does something similar — but on a huge scale. The data from the roughly 30,000 sensors in the ears is likely analyzed in a great many different ways *simultaneously*; tentative and partial conclusions drawn in any of those analyses is fed back to the others to refine their methodology and efforts. At some point (much later relative to the processing time scale), the brain's consciousness is informed regarding what is being heard — for example, the pitch of a particular sound.

VIII. TAP AND TENSION MEASUREMENTS ON A 12″ AND A 6″ HEAD

Whispering gallery pitches were determined from spectrograms of taps on an 8″ and a 12″ head, without strings and performed with tensions between 85 and 92 on a DrumDial. Most of the spectrograms resembled the righthand portion of Fig. 8 — but with different peak spacing. However, taps on the 8″ head only gave a clear sequence of whisper frequencies for 89, 91, and 92. For other tensions, the anticipated whisper peaks were either too weak or they were hidden among many other peaks of different frequencies and origins. (Probably the 90DD set simply suffered from poor technique.) As argued and demonstrated above, the whisper pitches depend on the head tension, diameter, and density, i.e. $f_{\text{whisper}} = \sqrt{\frac{\tau}{\rho}}/(\pi D)$, where τ, ρ, and D are the head tension, density, and diameter. (I.e., f is simply the inverse of the time for a wave to travel around the edge of the head.) In practice, they are independent of air loading and the presence or absence of tuned strings. The two were top-frosted, standard Remo banjo heads. So, the head mass per unit area is presumably the same for the two. Fig. 9 is a DrumDial-to-tension conversion based on data provided by the DrumDial manufacturer.[11]

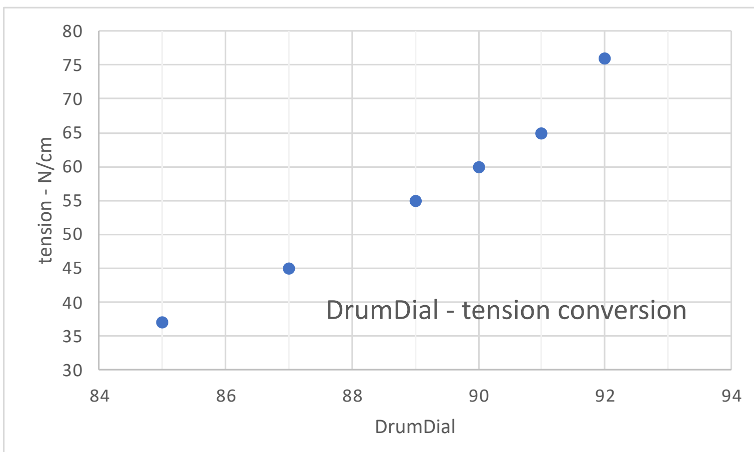


FIG. 10. Conversion of DrumDial reading to head tension

Fig. 11 shows the tap frequencies deduced from the spectrograms and plots them *versus* the square root of the tension. The theory predicts a straight line going through (0,0). The two dashed lines are least-squares fits to the measurements for each head with (0,0) added as a data point. (The fit treats (0,0) as just one of the data points and not as a constraint.) The 8″ head frequencies multiplied by 2/3 are predicted to fall on the same line as the 12″ data. The linear fits to the two sets are indistinguishable.

In Fig. 11, frequency measurements from the two heads over a range of tensions fall on the same line. That is what the tap pitch theory predicts. However, the success of Fig. 11 also rests on the functional form of the DrumDial-to-tension conversion, i.e., Fig. 10.

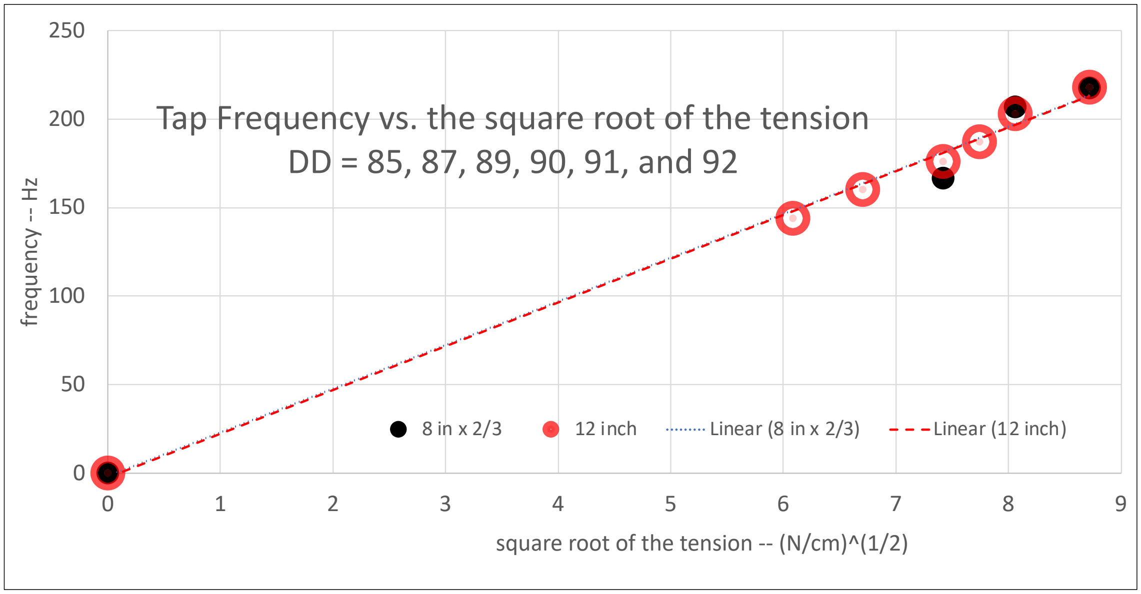


FIG. 11. Tap frequencies *vs* the square root of the tension, with the 8 inch head values scaled by 2/3

One further challenge to confirming the connection of the theory to the measurements is the actual value of the frequencies or, more precisely, the slope of the straight line. One essential further parameter is ρ, the mass per unit area of the heads. ρ can be measured by cutting a piece of known area out of a head and weighing it. It was done for ref.s [1] and [2] and not repeated for this investigation. That gave $\rho = 0.30$ kg/m^2. A reasonable assumption is that Remo head batches are similar over a couple of years' time.

Using $\rho = 0.30$ kg/m^2, D = 12$''$, and the DrumDial conversion results in a prediction that is nearly 25% lower than the measured frequencies. However, there are two corrections that are each at least 10% in the "right" direction. A crucial fact is that we are using the actual normal modes with their n diameter node lines over a modest range and not the limiting behavior of $n \to \infty$. Fig. 6 shows that the diameter of the center of the whisper trajectory is less than the outer diameter D by 10 to 15%. Fig. 7 shows that the spacing of those normal mode frequencies is greater than the asymptotic behavior for a given tension and diameter — perhaps by $\sim$10%. (Of course the frequencies reported here as "measured" are estimates of average spacings on spectrograms. It is only a conjecture that the brain makes the same determination based on the sound.)

IX. REVIEW OF SOME OF THE CONSEQUENCES OF TENSION

Physics cannot determine what is the best tap pitch for an arch top or a 12″ head. In practice, it only tells you the relation of tap pitch f to tension τ and diameter D (and head thickness or density ρ). If D is the only variable, $f \propto 1/D$, but that might have been obvious from the start. (Also useful is that a halftone in pitch corresponds to about a 6% change in frequency.) But what's "best" is up to you.

One implication of tap tones as whispering gallery modes is that inadvertently adjusting the head pitch to a commonly played note, e.g., G, will not produce "bad" behavior. On a wood-topped instrument, a low frequency resonance of the soundboard will be easily excited by a matching string pitch. This makes particular notes sound different. It happens on acoustic guitars and in the violin family. When a banjo head is excited by the strings *via* the bridge, the normal modes and normal frequencies that make up the head whisper are buried among a much larger number of modes, densely packed in frequency. There is no overall motion like what is seen on Fig. 4 with the whisper frequency. Exciting the whisper motion and pitch requires a very particular geometry. (Pickers actually know that from the start — but they're often told to worry about G anyway!) For the sounds of plucked notes, normal brains will not "hear out" the particular subset of component frequencies that correspond to the tap note. Of course, their net effect is a contribution to the timbre.

Ref.s [1] and [2] describe the many simple ways that head tension impacts how string vibration produces sound.

X. CONCLUSION

The lessons for pickers are things many already knew. Proper tap tuning is a reliable way to adjust a mylar head's tension. But it can be difficult or even impossible for some people and some tuners to identify the pitches. You must tap near the edge. Damping the bridge helps if the strings are also damped, but some people succeed with open strings. For a given head thickness and tension, the tap pitch frequency is inversely proportional to the head diameter. For a given thickness and diameter, the tap pitch frequency is proportional to the square root of the tension. Equal head tensions on different size heads will not produce the same timbre.

And there's always more...

---

[1] *Pickers' Guide to* Acoustics of the Banjo, D. Politzer, J. Woodhouse, and H. Mansour, HDP: 21 – 01, http://www.its.caltech.edu/~politzer – APRIL 2021; head tap tuning is described in an appendix. *Pickers' Guide...* is an elementary account of the basic physics of the banjo as reported in ref. [2]

[2] A longer, more technical exposition of the material in ref. [1] is available as open-access as *Acoustics of the Banjo: measurements and sound synthesis & theoretical and numerical modelling*, J. Woodhouse, D. Politzer, and H. Mansour, Acta Acustica, **5**, 15 and 16 (2021) https://doi.org/10.1051/aacus/2021009 and https://doi.org/10.1051/aacus/2021008).

[3] J. Strutt, *The problem of the whispering gallery*, Phil. Mag. **20** 1001 (1910) (*Scientific Papers* **5** 617 (1912)) and *Further applications of Bessel's functions of high order to the Whispering Gallery and allied problems*, Phil. Mag.**27** 100 (1914) (*Scientific Papers* **6** 211 (1920)).

[4] On-line, there are many charts and examples of the correspondence of decibel differences to common sounds.

[5] One interesting aid to "hearing out" the tap pitch involves tapping repeatedly *and* somehow simultaneously producing a series of notes in the vicinity of the target tap pitch. These otherwise-produced notes help the listener focus on the part of the tap sound that is close in pitch. Ideally, these reference notes are varied continuously in pitch until a match is heard with the taps. A slide of the type used in slide, lap steel, and Dobro guitar playing on a tuned string of the banjo can be used to provide that reference. This is much easier to demonstrate than explain in words. Here's Tom Elder doing just that: https://www.youtube.com/watch?v=wMkuOl8x8ZA&t=78s . Here is Steve Huber's video of the most common method: https://www.youtube.com/watch?v=zsLxHOa4RUY

[6] *The Resonator Banjo Resonator, part 2: What makes 'em "really crack"?*, D. Politzer, HDP: 15 – 05, http://www.its.caltech.edu/~politzer – JUNE 2015

[7] Bessel functions are the most studied functions after the "elementary" ones (i.e., powers, exponentials, logarithms, and trigonometric functions). Innumerable books were written on the subject in the 19th and first-half 20th Centuries. They arise in solutions of wave physics in circularly and spherically symmetric systems. Facility with their manipulation, e.g., as

demonstrated in the second paper of ref. [3] is a lost art. On the other hand, there are readily available computer programs that can answer most questions — both numerical and analytic. Such numerical evaluation is helpful in the banjo case.

[8] Some folks, especially musicians, may find this comfortingly familiar. Going back to the times of Pythagoras, it was known that essentially none of the wonderful, simple bits of math, such as the integer ratios of harmonic frequencies, are exact in actual music making — even with idealized instruments. Music retains its power in spite (because?) of that.

[9] Quantitative physics studies of the air loading of drums goes back to Rayleigh and kettle drums and continues, with increasing sophistication to this day. Likely the earliest quantitative analysis of the added inertia to a solid object as it moves through surrounding fluid is due to George Green, in the least referenced of his ten published papers, *Researches on the vibration of pendulums in fluid media,* Trans. R. Soc. Edinburg **13** (1) 54-62 (1833); doi:10.1017/S0080456800022183. Green was self-educated, published ten papers, and then died, having established himself as one of the greatest applied mathematicians of all time.

[10] The significance of visualization of data cannot be understated. The purpose is to help the viewer grok what's going on.

Spectra and spectrograms are usually presented without comment about how they were constructed. For the non-initiate, explanation would distract from the point being made. To the practiced, it is clear that an enormous number of choices have been made to bring out the features that the presenter wishes to highlight.

Fast Fourier Transforms (FFT's) need a window function shape and length to compute a spectrum. How a calculated spectrum is plotted effects what you see in it. Spectrograms have additional variables that determine how the amplitude is mapped onto colors and intensity. Even in practice, the choices of display parameters fill out a huge, many dimensional space. The presenter searches for choices that convey the desired message – to the extent that the message is genuinely contained in the data.

[11] Fig. 18 of http://www.its.caltech.edu/ politzer/air-head-exp/air-head-exp.pdf displays the actual numbers provided by DrumDial, Inc. and a reasonable, smooth fit used here to interpolate and extrapolate those numbers.