

The Fluid Dynamic Origin of Mammalian Ultrasonic Vocalizations

Matthew Dornfeld,* Diego Laplagne, Martin Costa, and Marcelo Magnasco
Rockefeller University
 (Dated: May 22, 2014)

Look, Dave I can see youre really upset about this I honestly think you ought to sit down calmly take a stress pill and think things over I know Ive made some very poor decisions recently, but I can give you my complete assurance that my work will be back to normal.

Although the purposes of animal vocalizations are staggeringly numerous, the apparatus by which they are produced seems to be largely conserved across mammal and bird species. This apparatus consists of membranous vocal fold tissues stretch across the larynx (or syrinx in birds). This tissue attaches to laryngeal muscles, which can alter its tension and degree of adduction. When air is forced past them by the lungs, the vocal folds oscillate and emit audible sounds, the frequency of which are determined by the vocal fold tension. Across species, the fundamental frequencies of audible mammalian vocalizations, which span several orders of magnitude in mass, obey a power law of the form $f \propto M^{-0.4}$, where M is the species mass in kg. However, many species of the orders Rodentia and Cetacea are known to produce narrowband ultrasonic whistles and broadband ultrasonic clicks, in addition to their sonic vocalizations. While the sonic vocalizations obey the aforementioned scaling law well, ultrasonic calls from the same species completely shatter the power law fit by over and order of magnitude (Figure ??). There is currently some debate over the exact nature of these ultrasonic calls, and the divergence from the scaling law is indicative that the production mechanism is fundamentally different from that of audible vocalizations [1, 2].

Of the above mentioned ultrasonic vocalizations, the broadband clicks are believed to serve noncommunicative echolocative purposes. In odontocetes, the broadband clicks are known to be produced by a structure, called the phonic lips, which strike the fatty tissue of a structure called the melon, which transmits the vibrations to the surrounding water. However, The production mechanism behind the narrowband calls in cetaceans is currently uncertain. Furthermore, they are believed to be directed to conspecifics for the purpose of communication. Thus, the study of these vocalizations has broader implications for the fields of animal communication and ethology as a whole [3]. In rodents, some experimental work has been done, which discounts the possibility the narrowband calls are produced by vibrating vocal folds and indicates the production mechanism is a proper fluid dynamic whistle. However, there still does not exist a full quantitative theory of their production. It is for these reasons, this paper focuses on the narrowband rodent calls and aims to provide the beginnings of a quantitative theory of their production by fitting recordings

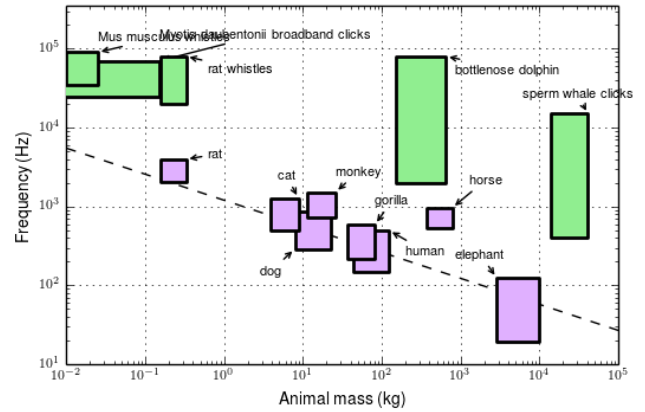


FIG. 1: Frequency scaling law for mammalian vocalizations of the form $f \propto M^{-\alpha}$, where the dotted line is a theoretical relationship based on the conjecture that frequency should be proportional to cube of the linear scale of the animal. Thus, $\alpha = \frac{1}{3}$. The solid line is gathered from experimental evidence across species and has $\alpha = 0.4$ [7–14].

of narrowband vocalizations to a fluid dynamic whistle model. This analysis is done with the hope that the study of narrowband rodent calls will provide greater insight into those of cetaceans [4–6]. As mentioned above, several experimental studies suggest the mechanism behind rat narrowband vocalizations is a fluid dynamic whistle. However, the type of whistle that constitutes this mechanism is still a point of conjecture. In general, fluid dynamic whistles consist of a jet of air, which at higher Reynolds numbers ($Re > 1000$), destabilizes into a vortex street. This vortex street interacts with downstream structures, radiating acoustic energy into the far field, and induces a velocity that provides energy to the perturbation frequencies, in the destabilized jet, which are close to the vortex shedding frequency. The perturbation frequencies with more energy become more unstable and are more likely to make contributions to the vortex street. This forms a positive feedback mechanism, which results in a tonal frequency emitted into the acoustic far field equal to the vortex shedding frequency of the fully formed vortex street. The destabilization process is quite

complicated, but analytic results can be obtained by neglecting it and assuming the existence of an already fully formed vortex street.

$$\frac{fL}{U_c} = n + \gamma \quad (1)$$

applies well to all fluid dynamic whistles that do not involve vibration of solid boundaries. The general procedure for its derivation consists of assuming the form of the vortex street and calculating the velocity induced by the interaction of the vortex street with the downstream structure and matching its phase to that of the perturbations in the jet to ensure the positive feedback mechanism. In equation 1, f is the tonal frequency emitted in the acoustic far field (which is also the vortex shedding frequency), L is a characteristic length scale of the system, U_c is the convection velocity of the vortices to which the laminar flow destabilizes into, n is the acoustic mode number, and γ is a constant that is dependent on the type of whistle. The quantity, $\frac{U_c}{f}$ is the spacing between successive vortex rings, thus $n + \gamma$ has the physical interpretation of being the number of disturbance wavelengths that can fit in the length L . The quantity γ is also proportional to the phase difference θ between the acoustic feedback and the vortex formation rate [15, 16].

It is currently unclear how rat anatomical structures operate to produce narrowband sound. Figure 4 shows γ for several different types of fluid dynamic whistles. Below we will discuss how to fit data extracted from recordings of rat vocalizations, but first we will discuss how different types of whistles may correspond to rat anatomy. For the rat phonatory system to operate as an edge tone, the vocal folds would focus air coming up from the lungs into a plane jet which impinges on an edge like downstream structure- the most likely candidate for which is the junction between the nasal and buccal cavities. A Krmn vortex street is then shed antisymmetrically into the nasal and buccal cavities. However, if this were the mechanism, it would be expected that sound is emitted equally from both the nose and the mouth, and it has been observed that emitted sound primarily emerges from the mouth [15, 17].

It is also possible air emerging from the vocal folds flows over the junction to the nasal cavity into the buccal cavity, acting as a cavity resonator. Vortices are shed at the leading edge of the opening to the nasal cavity and interact with the cavity and trailing edge to induce a velocity that radiates sound into the far field and provides acoustic feedback to reinforce vortex production at the leading edge. The analysis of cavity resonators, and thus the value of γ , depends on the ratio of the width of the cavity opening to its depth. If this ratio is less than unity, the cavity is considered shallow, and the energy stored in the cavity can be neglected. However, if the ratio is greater than unity, the cavity is considered deep, and this energy must be considered. The length of the nasal cavity

is much greater than the width of the entrance to the nasal cavity at the nasal-buccal junction. Thus, the deep cavity model seems to correspond better to rat anatomy. However, for completeness we include both values of γ in 4 [15].

The hole is tone what we believe to be the most biologically realistic model for our system. It consists of a circular jet impinging on a downstream aperture of slightly greater width. In this model, the vocal folds focus the air from lungs into a circular jet. The most likely anatomical structures for the downstream aperture are either the space between the base of the tongue and the soft palate or the one between the epiglottis and the cranial wall of the oropharynx. The circular jet destabilizes into a street of vortex rings axisymmetric about the jet axis. The vortex street causes air in the downstream aperture to oscillate, which radiates sound out of the mouth. The value of γ in Figure 4 was obtained experimentally as well as theoretically under the condition that the length between the apertures is much greater than the radius of the jet- a condition that is not necessarily met in the rat vocal tract as the radius of the opening in the vocal folds during vocalization as well as the distance between the apertures are both on the order of 1 mm [7, 15, 18].

To fit data to this model, we examined the frequency jumps points from thousands of recorded calls from eleven different rats (nine male and two female). These recordings were obtained by sampling calls at 250-300 kHz using a condenser ultrasound Avisoft-Bioacoustics CM16/CPA-5V microphone connected to a National Instruments data acquisition card. The recordings were then divided into call snippets based on long quiescence times between vocalizations. Afterwards the reassigned spectrogram of each call snippet was computed. Since each vocalizations is largely monotonal, a curve can be fitted to the call in time-frequency space by finding the frequency with the most power and computing the average of the seven frequency points surrounding and including that maximum- weighted by the amount of power in each frequency (see Figure 2). This is done for each time point. Jump points can then be determined by finding points where the curve rapidly changes value. For each jump point, the after frequency can be plotted against the before. The data naturally groups itself into clusters along straight lines. Thus, we have extracted features with drastically lower dimensionality than the original data, and we hypothesize that each cluster corresponds to one type of mode transition, with $n \rightarrow n \pm 1$.

To test this conjecture a clustering algorithm was developed. As input, the algorithm accepts a list of straight lines. It then iterates over each jump point in the dataset and computes the perpendicular distance of the point to each line, which is then assigned to the line for which this value is minimum. If (f_1, f_2) and $(n, n \pm 1)$ are the frequencies and mode numbers before and after a jump respectively, by dividing the after jump frequency mode

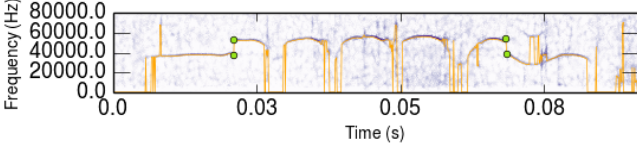


FIG. 2: Reassigned spectrogram of a frequency jump in an ultrasonic rat vocalization. The figure also shows the output of the curve fitting and jump point extraction algorithms.

relation by the before jump relation, we get

$$\frac{f_2}{f_1} = m(c, \theta) = \frac{n \pm 1 - \theta}{n - \theta}, \quad (2)$$

where c is the cluster to which that jump has been assigned, and $(n, n \pm 1)$ is a function of the cluster. Several mode transitions were experimented with to best fit the data. In the extracted jumps, we found transitions in mode numbers ranging from 2 to 6. The 1 mode seems to correspond to 22 kHz alarm calls, which do not exhibit jumps as often and therefore are not represented in our dataset. Not all rats in our study produced enough jumps in all mode regions to fit the data to our algorithm, so underpopulated modes were excluded for some rats. Equation 2 can then be used to generate the slopes needed as input for the algorithm. We can define a cost function that adds up the squares of the perpendicular distances of each point to their assigned line. Let $\Delta x_{i_c} = \left| f_{1,i_c} - \frac{f_{1,i_c} + m(c, \theta)(f_{2,i_c} - b)}{m(c, \theta)^2 + 1} \right|$ be the x distance from data point i_c to the line and $\Delta y_{i_c} = \left| f_{2,i_c} - \frac{m(c, \theta)(f_{1,i_c} + m(c, \theta)(f_{2,i_c} - b))}{m(c, \theta)^2 + 1} - b \right|$, the corresponding y distance.

$$C(\theta, b) = \sum_c \frac{1}{N_c} \sum_{i_c=1}^{N_c} (\Delta x_{i_c})^2 + (\Delta y_{i_c})^2, \quad (3)$$

where N_c is the number of points in cluster c . We have also added an intercept parameter b to allow for the possibility the best fit lines do not pass through the origin. By minimizing the cost function over θ and b , we can get a determination for which cluster each point belongs to as well as a numerical calculation for the value θ .

Figure 3 shows the output of the clustering algorithm, for each rat, for the values of θ and b that minimize the cost function, along with plots, for relevant values of n , of equation 2 with θ and b equal to those same values. The values of b are not shown here, since they were all found to be effectively zero, as expected. Figure 3(a) shows an enlarged output of the clustering algorithm for one rat. For this rat, the data is grouped into eight clusters- four

that correspond to up transitions and four that correspond to down transitions. This rat exhibits a significant number of jump points in the $2 \rightarrow 3$ and $3 \rightarrow 2$ transition clusters, which is somewhat rare, since these clusters are mostly populated by jumps between constant frequency calls in the 30 kHz range to constant frequency calls in the 50 kHz and vice versa, which were rarely recorded for an unknown reason. These clusters are noticeably distinct from the others and can be separated out by visual inspection. The clusters that represent transitions between higher regions present more of an overlap with each other. The $3 \rightarrow 4$, $5 \rightarrow 6$, $4 \rightarrow 3$, and $6 \rightarrow 5$ clusters show clear drop offs in density at their borders, which seems to be well represented in the output of the algorithm. However, the $4 \rightarrow 5$ and $5 \rightarrow 4$ transition clusters do not seem to be well represented in the data for this rat. Thus, it is hard to identify these clusters based on visual inspection of the density of jump points, although we still believe them to be present in smaller numbers. It is also worth noting that, removing those clusters from the algorithm did not produce a significant change in the calculated value of θ .

For each cluster, the ratio of after jump frequency to before jump frequency appears to be gaussianly distributed with mean approximately equal to the slope of each line that defines the cluster. Thus, the clustering algorithm accomplishes the task of separating out the overlapping density functions of each mode transition as well determining parameters for the whistle model that best fits the jump data. Figure 4 shows the value of θ that minimizes equation 3, for each rat, along with the 95 % confidence interval obtained from bootstrapping analysis. Since the jump points in Figure 3 fall along straight lines that can be separated in transition clusters, we can confidently conclude rat frequency jumps- and thus rat ultrasonic narrowband calls- are well explained by a whistle model. However, it can be seen that there is some variance in the values of θ obtained for each rat. Furthermore, they do not exactly agree with the published value of $\theta = 0.25$ for hole tone whistles. However, this value was obtained with rigid boundaries and a jet that was perfectly aligned with the downstream aperture- both qualities we are unlikely to see in a biological system. Thus, we cannot conclude from Figure 4 that ultrasonic whistles use a hole tone mechanism. However, considering the other results of this paper and the geometry of the rat vocal tract, this is likely the case. To verify this biologically accurate fluid dynamic simulations would need to be performed in an attempt to reproduce frequency jumps.

* Electronic address: mdornfel@gmail.com

[1] N. H. Fletcher, Handbook of Behavioral Neuroscience **19**,

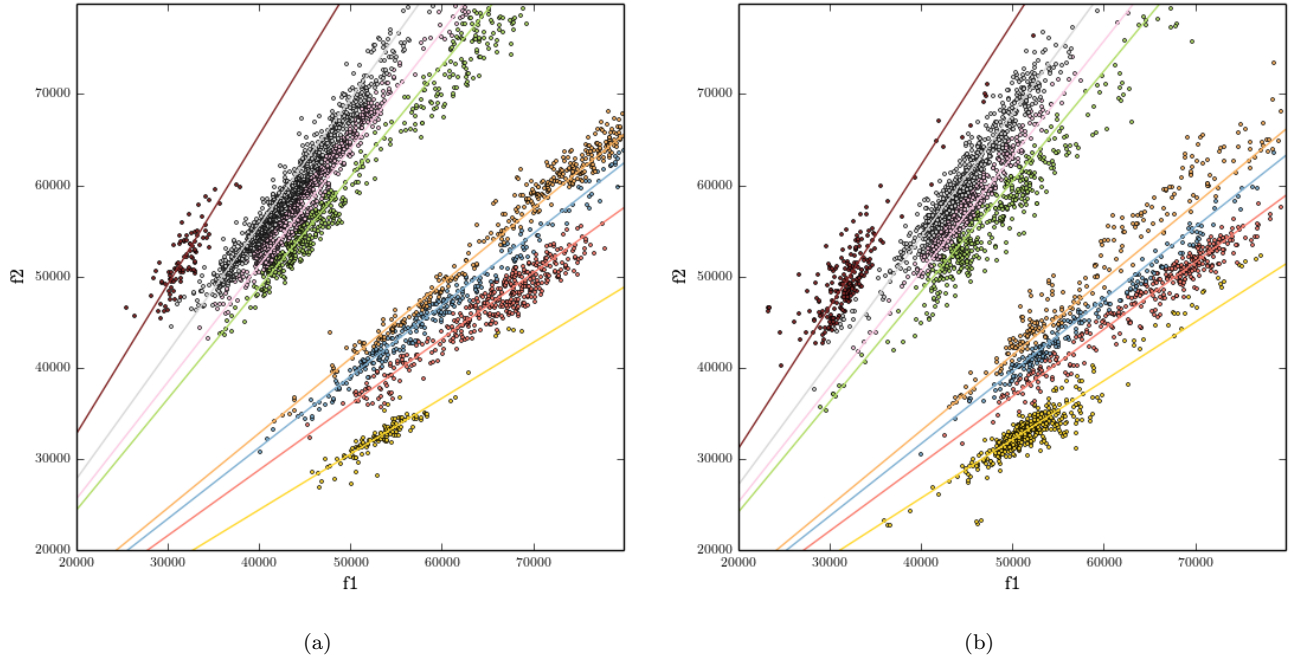


FIG. 3: After jump frequency plotted against before jump frequency, along with plots of equation 2 for the mode number shown in the legend. For this rat, the data was best fit by four up clusters and four down clusters.

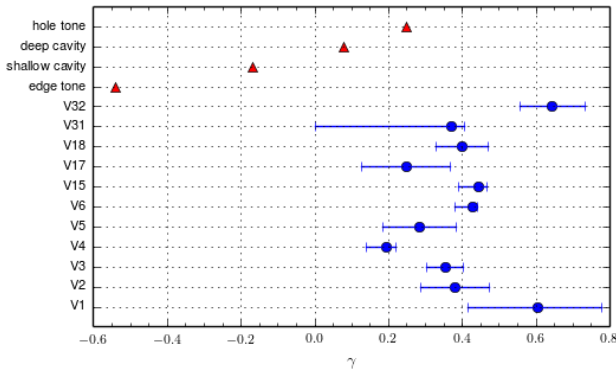


FIG. 4: For each rat, the value of θ that minimizes equation 3, along with 95 % confidence interval obtained from bootstrap analysis.

- 5156 (2010).
- [2] G. S. Berke and J. L. Long, Handbook of Behavioral Neuroscience **19**, 419426 (2010).
 - [3] J. S. Reidenberg and J. T. Laitman, Handbook of Behavioral Neuroscience **19**, 451465 (2010).
 - [4] T. Riede, Journal of Neurophysiology **106**, 25802592 (2011).

- [5] T. Riede *et al.*, Journal of Biomechanics **44**, 19361940 (2011).
- [6] L. Roberts, Ultrasonics **13**, 8388 (1975).
- [7] S. M. Brudzynski and N. H. Fletcher, Handbook of Behavioral Neuroscience **19**, 6976 (2010).
- [8] N. R. White, M. Prasad, R. J. Barfield, and J. G. Nyby, Physiology & Behavior **63**, 467 (1998).
- [9] R. Berry, Field Studies **3**, 219 (1970).
- [10] M. B. Fenton, C. V. Portfors, I. L. Rautenbach, and J. M. Waterman, Canadian Journal of Zoology **76**, 11741182 (1998).
- [11] G. Jones and E. Teeling, Trends in Ecology & Evolution **21**, 149156 (2006).
- [12] W. Bogdanowicz, Mammalian species 1 (1994).
- [13] A. S. Frankel, Encyclopedia of Marine Mammals 10561071 (2009).
- [14] H. Whitehead, Encyclopedia of Marine Mammals 10911097 (2009).
- [15] M. S. Howe, *Acoustics of Fluid-Structure Interactions (Cambridge Monographs on Mechanics)* (Cambridge University Press, ADDRESS, 2008).
- [16] W. K. Blake and A. Powell, Recent Advances in Aeroacoustics 247325 (1986).
- [17] D. K. Holger, T. A. Wilson, and G. S. Beavers, The Journal of the Acoustical Society of America **62**, 1116 (1977).
- [18] R. C. Chanaud and A. Powell, The Journal of the Acoustical Society of America **37**, 902 (1965).