A generalized model for communicating individuality through teleost swim bladder modulation

Cameron A. Matthews¹,* and Pierre-Philippe J. Beaujean²

ABSTRACT
Arguments have been made for and against the traditional swim bladder model as a primary component of fish vocalization. This paper presents arguments for decoupled forced and resonant responses being extractable features within a variable air volume. As such, a mechanical analog is used to show how envelope modulation may be used by some species to identify air volume and consequently size in conspecifics. These arguments consider how an arbitrary fish may apply a genetic strategy of forcing vocalization through slow, fast, or both slow and fast sonic musculature while amplitude modulating via swim bladder. The classic resonant bubble model is revised to account for a hypothetical carrier signal resonance associated with static or varying volume. In the absence of live specimens, a test is conducted in different cylindrical structures with equally sized air volumes. First, a proposed method for extraction of swim bladder volume features through blind amplitude demodulated signals in the time and frequency domain is applied. Second, a proposed method for extraction of swim bladder volume features through cyclostationary analysis of the cross-spectral coherent spectra of the modulated and demodulated signal is applied. Both methods take average frequency content as derived by the prescribed signal processing techniques as the input to the correlator functions used to identify air volumes. Vocalizations of Epinephelus guttatus, or more commonly known as the red hind grouper, are used as test signals.

KEY WORDS: Biomodulation, Bioresonance, Bioacoustics

INTRODUCTION
The swim bladder model for resonance as defined in Ladich (2015) is understood to be one of the most widely accepted approaches to characterizing vocalization from fish who apply muscle contractions across the swim bladder to generate sound. Normal mode analysis can yield varied outputs for any particular geometry and are generally considered as a primary component of fish vocalization, specifically territorial assertion and mating ritual, some fish have indeed exhibited correlations in mass and swim bladder size (Fine et al., 1962), but not with enough detail to differentiate size between referenced power sources. With respect to the pivotal behaviors typically seen as associated with vocalization, specifically territorial assertion and mating ritual, some fish have exhibited correlations in mass and swim bladder size (Fine et al., 2001; Suthers et al., 2016; Ali et al., 2016). If the frequency is indeed coupled with the radius of the swim bladder, then the frequency drop indicated in Suthers et al. (2016) could potentially be indicative of an attempt to maintain a particular modulation effect about an increasing swim bladder volume.

DISCUSSION
This paper presents a methodology for isolating envelope modulation characteristics associated with resonance of air cavities in water being impulsively excited by an external source. As a means of controlling the signal output, an identical test reference vocalization with no variation in power was played through four varying air volumes, and each air volume was correctly assessed by both methods considered. The intent of showing two methods is to corroborate each other. Further, the air volumes were static, implying the most consistent possible modulation scheme as described in Eqn 2 with respect to radius of the air volume. It is important to consider some of the assumptions in developing a successful strategy for accurately estimating the air volumes used as test targets.

With respect to static power, this was intentional with respect to the candidate fish species selected for vocalization; no references on signal power (P) characteristics of the red hind were discovered in the literature review. Relatives of the red hind have been analyzed from a transmission power standpoint (Hazlett and Winn, 1962), but not with enough detail to differentiate size between referenced power sources. With respect to the pivotal behaviors typically seen as associated with vocalization, specifically territorial assertion and mating ritual, some fish have exhibited correlations in mass and swim bladder size (Fine et al., 2001; Suthers et al., 2016; Ali et al., 2016). If the frequency is indeed coupled with the radius of the swim bladder, then the frequency drop indicated in Suthers et al. (2016) could potentially be indicative of an attempt to maintain a particular modulation effect about an increasing swim bladder volume.

Multipath effects of the pipes were completely neglected in consideration. The pipes may be expected to act as nearly perfect Von Neumann boundaries for the frequencies of interest and have relatively long periods of bounce along the pipe walls affecting linear summing at the receiver. This does not preclude multipath interference from contributing to variations in spectral profile, though none were visibly observed in the collected records.

The resonant bubble model does not strictly require a spherical reverberation, though in the classical form in Ladich (2015) it does not consider the complete effects of normal modes transferring in the wall as lobed or otherwise geometrically defined waves. All of these effects, should they be present, could contribute to variations in vocalizing. When considering the possibility of a time and envelope modulation dependent strategy in terms of the findings in McKibben and Bass (2001), it becomes of interest to consider possible biological rationale for developing such proclivities.

RESULTS
Observations were made on a controlled dataset, where multiple records were collected using 4 inch long standard PVC pipe sections of 1.5, 2, 3 and 4 inch diameter.

© 2018. Published by The Company of Biologists Ltd| Biology Open (2018) 7, bio023515. doi:10.1242/bio.023515

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Received 6 December 2016; Accepted 24 April 2018
in the signal envelope in either time or time-frequency analysis. The air bubble effect of a rubber balloon in a pipe does not provide the same biological controls as a fish in terms of damping, movement, etc.

**MATERIALS AND METHODS**

To analyze the effect of temporally varying modulation, we first consider the basic problem to be one of impulse (generated in the musculature of the fish) and impulse response (generated in the swim bladder). A classical analysis tool can be found in convolution for separation of such a signal using the definitions found in Ifeachor and Jervis (2002):

$$ y(f) = x(f) * h(f, t) $$

In (1), an impulse signal is presumed to be convolved in the time domain by an impulse response signal and thus multiplied in the frequency domain. This simple model helps describe the effects seen in McKibben and Bass (2001). An altered model can be generated by introducing an ideal filter model as described in Bianchi and Sorrentino (2007). This model implies that the output signal will have a form of carrier signal or a ‘carrier ID’ for a receiver set capable of demodulating both the carrier impulse response, the information signal, and the carrier ID signal. The results in McKibben and Bass (2001) show the capability of a fish being capable of recognizing modulation effects, and at least suggest the possibility of a listening fish being able to infer information from such a modulation such as bladder size and in turn the size of the vocalizing fish.

Based on this possibility, a model is first constructed about the conventional resonant bubble model described in Ladich (2015) for estimating frequency as a function of swim bladder radius and environmental factors. As a variation to this model, we consider an impulse and impulse response model, where the musculature acts as an impulsive force while the bladder acts as an impulse response. For the purposes of this paper, which are to isolate a general estimate of the impulse response, the model is not concerned with features associated with the forcing function of the musculature. The model incorporates time as a variable to indicate the effects as a function of the period of the vocalization. The vocalization over time $t$ is defined over a vocalizing record length $r$, so that $t$ varies between $-r/2$ and $r/2$. The vocalization is centered at an epoch of time $t=0$:

$$ M(F, R, P, t) = 1 + \frac{\cos(2\pi R t)}{2\pi FR} \sqrt{\frac{3\gamma P}{\rho}} \text{ for } -\frac{r}{2} \leq t \leq \frac{r}{2} $$

As shown in Fig. 1, the variable $R$ represents the radial distance in cm, the ratio of specific heats is $\gamma=1.4$, $P$ is the combined atmospheric and hydrostatic pressure in Pascals (Pa) and $\rho$ for the mass density of water in atmosphere. Specific heat, hydrostatic pressure and water mass density remain unchanged in this model, with forcing function as related to the musculature having a frequency $F$. This revision to the model as described in Ladich (2015) describes a generalized modulation signal...
An implication of this model is that the peak modulation amplitude drops off significantly as frequency $F$ of the forcing function increases. The next pair of figures illustrates these decay models over a single period of unit vocalization time for a hypothetical raised cosine distribution strategy, where a magnitude operator is applied to (1) for a radial variation of 1–10 cm.

Within the model described, it is possible to estimate any number of resonant modulation strategies on the part of the fish, including variable compression of the air bladder to vary the output of $M(t)$. The swim bladder modulation effect is essentially an all-pass filter for the vocalization, with passband ripple and natural curvature over the ideal 1-D filter envelope [it is important to note that $M(t)$ is a 1-D function, the previous figure would a line drawn through it somewhere to define $M(t)$]. As the frequency increases, peak cyclical effects associated with the swim bladder approach unity, indicating that the function described in (2) will be driven to unity – essentially an ideal all-pass filter. This supports the evolutionary strategies described in Ladich (2015) regarding low frequency acoustic communication and shallow running teleost fishes if swim bladder modulation features do in fact help identify attractive mates.

The proposed modulation is expected to be, relatively speaking, highly overshadowed by the information signal. As such, in the following section, we consider a well-defined and strong red hind grouper vocalization played through several increasing diameters of pipe filled with a roughly spherical and equally increasing air volume in a typical rubber balloon. It should be noted that any acoustic signal could be applied as the input to the model; the intent is to measure effects of an increasing air volume on the sound passing through it.

In order to measure effects in a controlled setting, a physical analog was constructed to allow variations in the air volume for a static test audio file to be played through. Fig. 3 shows a diagram of the test set-up.

With a general framework for analysis and an absence of live specimens to work with, we consider the model thus far in terms of a static volume with fish vocalizations played back through the static volume increased sequentially. The original file is shown below in Short Time Fourier Transform (STFT) – often referred to as spectrogram – format. The reference signals’ spectral content is shown in Fig. 4.

In analyzing the signal, methods as described in work by E. C. Like (Non-Cooperative Modulation Recognition Via Exploitation of Cyclic...
As an assumption in developing the demodulation strategy, the same methods in McKibben and Bass (2001) are presumed upon regarding amplitude modulation. The blind demodulation estimate also presumes on the fish applying Single Side Band (SSB) amplitude demodulation. The Amplitude Modulation SSB (AM-SSB) model as shown in Fig. 5 is attractive not only as a possible physiological effect of the damping swim bladder wall but also as a carrier suppression tool. The intent of this operation is to decorrelate impulse and impulse response to extract a hypothetical carrier identification signal. The demodulation routine then becomes the signal multiplied by a cosine function using the fundamental of the presumed vocalization [this is readily achieved for red hind vocalization – potentially any communication signal – with peak-finding in energy detectors following a front-end speech detection process, see Matthews and Beaujean (Edge Detection of Red Hind Grouper Vocalizations in the Littorals, Society for Photonics and Imaging Electronics, Buried and Obscured Objects Detection Session, Defense Security Systems, 2016) and C. A. Matthews (Acoustic Tonal And Vector Properties Of Red Hind Grouper Vocalizations, Doctoral Thesis, Florida Atlantic University, 2017)]. The following flow diagram describes the process in detail, which consists in low-pass filtering the modulated signal through a 5th order Butterworth filter as described in Bianchi and Sorrentino (2007).

When the input signal \( x \) is demodulated, it yields the information signal, termed \( y \). When these signals are decorrelated, the remainder is expected to yield carrier signal information, which can then be cross-correlated as a template to an input signal \( z \). First, we consider an arbitrary overlap and frequency precision for conversion to time–frequency, where all frequencies \( k \) being considered, an approximation of the frequency content over time \( t \) is estimated for \( x, y \) and \( z \), respectively.

\[
G_k(t) = F\{x(t)\} \quad T_k(t) = F\{y(t)\} \quad R_k(t) = F\{z(t)\} \quad (3)
\]

The expectation, or mean, denoted by \( E\{\} \), is applied for the deconvolved frequency content over time for Regions of Interest (ROI) that are germane to the feature set as a means of smoothing information content in frequency \( f \) about a region of frequencies \( k(m:n) \) of interest contained within \( f \). In the case of the red hind grouper for example, the feature set is expected to be contained within the region about 185 Hz as previously considered, such that \( 150 \leq f \leq 200 \) Hz. These signals are then integrated...
for lag $\tau$.

$$U_{k(m) \lambda(n)}(t) = E\left\{ \frac{G_{k(m) \lambda(n)}(t)}{T_{k(m) \lambda(n)}} \right\} J_u = \int_{-\infty}^{\infty} |U^2_{k(m) \lambda(n)}(\tau)| d\tau$$

$$V_{k(m) \lambda(n)}(t) = E\left\{ \frac{R_{k(m) \lambda(n)}(t)}{T_{k(m) \lambda(n)}} \right\} J_v = \int_{-\infty}^{\infty} |V^2_{k(m) \lambda(n)}(\tau)| d\tau$$

(4)

The resulting normalized cross correlation is then derived and labelled $Z_1$:

$$Z_1(\tau) = \frac{U_{k(m) \lambda(n)}(\tau) \times V_{k(m) \lambda(n)}(\tau)}{\sqrt{J_u \times J_v}}$$

(5)

As noted, average frequency content is used since the true modulating signal is unknown, but presumed to reside within the available communication band of the information signal.

Using the same variable notation where appropriate for continuity, the cross-spectral density and coherence models for cyclostationarity are applied. In particular, we consider the spectral density function as defined in Like (2007):

$$S_{xy}(\alpha, f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f \tau} d\tau$$

(6)

In this notation, the spectral content in $f$ is differentiated from the cyclic content in $\alpha$ as defined in Napolitano (2012). This is important as we are now dealing in the cyclic-spectral perspective rather than time-spectral perspective, though the datasets remain the same. The implication is that we are now analyzing frequency content that presents with periodic behavior over time. When taking the autocorrelation of a signal, a real signal in time $x(t)$ will be correlated with itself to produce

$$R_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{0}^{T} x(t + \tau)x(t) dt$$

(7)

Given two input signals, it is possible to derive the cross-spectral coherence function for a given $\alpha$ band:

$$C_{xy}^m(\alpha, f) = \frac{|S_{xy}^m(\alpha, f)|^2}{S_{xx}^m(\alpha, f) S_{yy}^m(\alpha, f)}$$

(8)

To achieve a template structure, a variation on (8) is applied in an effort to isolate cyclostationary elements associated with the air volume being actuated. In doing so, we derive cross-spectral density functions for both the modulated information signal $x$ and the hypothesized demodulated information signal described $y$ [see (3)]. The intent will be to derive the modulation signal estimated in (2). We then derive the spectral autocorrelation of a candidate modulated information signal $z$. The resulting coherent spectrum – not to be confused with the cross-spectral coherence in (8) – is used as a cross-correlating metric against known quantities of previously cross-spectrum deconvolved signals of known volume.

$$T_{xx}(\alpha, f) = \frac{S_{xx}^m(\alpha, f)}{S_{xy}^m(\alpha, f)}$$

(9)

In this equation, $T$ represents the coherent spectrum of a template signal $x$ and its information deconvolved signal $y$. The intent of this exercise is to take the assumption that signals being shaped in the time domain are thus correlated in the frequency domain; deconvolving the input impulse (muscle input) will yield the desired characteristic impulse response modulation (bladder resonance). Ideally, the only elements left will be the residual modulation from the swim bladder as sought in (2). A template function for each specific volume to be tested must be created, forming a set of matched filters. When an input is applied to this model, it is possible to derive a

Fig. 7. Cyclic spectrum input signal and reference cyclic spectrum signal peak correlations in mean frequency content as described in Eqn (10) are shown for all pipe diameters. A red marker is used to denote the highest correlation in each template to the reference signal.
coherent spectrum of signals which are expected to represent the residual carrier as sought in (2):

\[ G_{\alpha, f} = \frac{S_{\alpha}(\alpha, f)}{S_0(\alpha, f)} \]  \hspace{1cm} (10)

Once the decorrelated frequency content of \( G \) is derived about the unit regions of interest, it is averaged and cross-correlated with the same for each template volume \( T \) to derive match criteria in the same fashion as the preceding section, where the cyclic content is operated on in the place of time content.

\[
V_{k(n),k(n)}(\alpha) = \mathbb{E}\{G_{\alpha, k(n)} : k(n)\} \quad J_F = \int_{-\infty}^{\infty} |V_{k(n),k(n)}(\tau)| \, d\tau
\]

\[
U_{k(n),k(n)}(\alpha) = \mathbb{E}\{T_{\alpha, k(n)} : k(n)\} \quad J_F = \int_{-\infty}^{\infty} |U_{k(n),k(n)}(\tau)| \, d\tau
\]  \hspace{1cm} (11)

Operating on the cyclic frequencies, a normalized cross-correlation \( Z_2 \) is derived for the average cyclic content in the spectral ROI of 150–200 Hz:

\[
Z_2(\tau) = \frac{U \times V}{\sqrt{J_F \times J_U}}
\]  \hspace{1cm} (12)

The resulting peak correlations of \( Z_1 \) and \( Z_2 \) are shown in Figs 6 and 7. In each dataset, the correct air volume can be ascertained from the peak correlation using either method. This suggests correlation in the modulation of each signal by the associated air volume.

**Acknowledgements**

The authors would like to thank Dr Michael Fine (Virginia Commonwealth University) and Dr Andrew Bass (Cornell University) for their valuable insights and multiple papers on fish vocalizations. This paper was prepared in conjunction and partial fulfillment of a Doctorate in Ocean Engineering at FAU, and was developed under training provided by the United States Navy (USN).

**Competing interests**

The authors declare no competing or financial interests.

**Author contributions**

Conceptualization: C.A.M.; Methodology: C.A.M.; Software: C.A.M.; Formal analysis: C.A.M.; Investigation: C.A.M.; Writing - original draft: C.A.M.; Writing - review & editing: C.A.M., P.-P.J.B.; Supervision: P.-P.J.B.

**Funding**

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**References**

Ali, H. A., Mok, H. and Fine, M. L. (2016). Development and sexual dimorphism of the sonic system in deep sea neobythitine fishes: the upper continental slope. *Deep-Sea Res. 115*, 293-308.

Antoni, J. (2007). Cyclic spectral analysis in practice. *Mech. Syst. Signal Process. 21*, 597-630.

Bianchi, G. and Sorrentino, R. (2007). *Electronic Filter Simulation & Design*, pp. 17-20. New York: McGraw-Hill Professional.

Fine, M. L. (2012). Swimbladder sound production: the forced response versus the resonant bubble. *Bioacoustics 21*, 5-7.

Fine, M. L., Malloy, K. L., King, C. B., Mitchell, S. L. and Cameron, T. M. (2001). Movement and sound generation by the toadfish swimbladder. *J. Comp. Physiol. A 187*, 371-379.

Frisch, G. (1994). *Ocean and Seabed Acoustics*. Upper Sadler River: Prentice Hall.

Hazlett, B. and Winn, H. E. (1962). Sound producing mechanism of the nassau grouper, epinephalus striatus, copeia. *ASIH 1962*, 447-449.

Ifeachor, E. C. and Jervis, B. W. (2002). *Digital Signal Processing: A practical Approach*. Essex: England Prentice Hall.

Ladich, F. (2015). *Sound Communication in Fishes*. New York: Springer.

McKibben, J. R. and Bass, A. H. (2001). Modulation and signal recognition in fish. *J. Acoust. Soc. Am. 109*, 2938.

Napolitano, A. (2012). *Generalizations of Cyclostationary Signal Processing*. Rochester: Wiley & Sons.

Suthers, R. A., Fitch, W. T., Fay, R. R. and Popper, A. N. (Eds.). (2016). *Vertebrate Sound Production and Acoustic Communication*, *Springer Handbook of Auditory Research* 53. New York: Springer.