Beat encoding at mistuned octaves within single sensory neurons

Alexandra Rudnaya[a,*], Ramona Schäfer[a], Jan Grewe[a], Jan Benda[a,b,c,*]

[a] *Neuroethology, Institute for Neurobiology, Eberhard Karls University, Tübingen, Germany*
[b] *Bernstein Center for Computational Neuroscience Tübingen, Tübingen, Germany*
[c] *Werner Reichardt Centre for Integrative Neuroscience, Tübingen, Germany*

Abstract

Beats are periodic amplitude modulations resulting from the superposition of two periodic signals of different frequencies. Psychophysical experiments by Georg Simon Ohm, Hermann Helmholtz, and others from the 19th century demonstrate that beats are not only perceived for two close-by frequencies but also at mistuned octaves. However, the physiological mechanisms of this percept are still debated. Motivated by a field study, in which we observed the behavioral relevance of beats at high difference frequencies, we here study the beat encoding over a wide range of difference frequencies in the electric fish *Apteronotus leptorhynchus*. The activity of P-unit electroreceptor afferents, that share many properties with mammalian auditory fibers, follows a repetitive pattern with slow modulations of their firing rate reoccurring around multiples of the frequency of the carrier signal. By mathematical reasoning and supported by simulations of modified integrate-and-fire models we conclude that neither Hilbert transform, squaring, harmonics of the carrier, nor a threshold operation are sufficient to extract slow beats around the octave. Raising the thresholded signal to a power of three, however, is sufficient to explain the repetitive P-unit responses. Since the threshold-nonlinearity of the afferent's spike generator could be ruled out, it is most likely the transfer function of the electroreceptor synapse that implements such a non-linearity. In the auditory system the hair-cell synapse is known to act as a smooth threshold operation. We thus conclude that this mechanism within each auditory fiber contributes to the perception of beats at mistuned octaves.

Keywords: difference frequency, nyquist frequency, aliasing, *A. leptorhynchus*, electrosensory, mistuned octave, beats of mistuned consonance

1. Introduction

Periodic signals are ubiquitous in the auditory system (Lewicki, 2002; Romani et al., 1982; Köppl, 1997) and the electrosensory system of wave-type electric fish (Hopkins, 1976). The superposition of two periodic signals that differ in frequency results in a periodic amplitude modulation (AM) known as "beat". Tuning curves of both auditory fibres and p-type electroreceptor afferents typically decline with increasing difference frequencies (Rhode and Greenberg, 1994; Nelson et al., 1997; Walz et al., 2014).

This is in contrast to experiments on human perception in the 19th century demonstrating beating amplitude modulations to be perceived not only for close by frequencies, but also for mistuned octaves where the second frequency is close to octaves of the carrier frequency (Roeber, 1834; König, 1876). These and related percepts have been formalized by Ohm (1839) and Helmholtz (2009) and discussed to result from aural harmonics or interactions with combination tones. These potential mechanisms, however, have largely been ruled out by masking experiments (Plomp, 1967).

Modern electrophysiological studies in auditory systems focused on the encoding of sinusoidal amplitude modulations (SAMs) with frequencies below half the octave of the carrier (Joris et al., 2004). In contrast to beats, SAMs result from an interaction of three tones, a carrier flanked by two stimulating sine waves at plus and minus the SAM frequency. Thus, the physiological mechanisms underlying beat perception at mistuned octaves remain an open issue.

Similarly to electrophysiological research in auditory systems, the encoding of beats in electrosensory systems has also been studied at low difference frequencies only (Walz et al., 2014). Recent field studies, however, demonstrated behaviorally relevant difference frequencies above half of the octave of the carrier in the context of courtship and synchronization of spawning (Henninger et al., 2018) and inter-species detection (beyond 400 Hz, Henninger et al., 2020).

Wave-type Gymnotiform electric fish generate a sinusoidal electric organ discharge (EOD) of a species and individual specific frequency (Knudsen, 1975; Henninger et al., 2020). The EODs of two nearby fish superimpose and thus produce a beat. Beats and their modulations are

*Corresponding author
Email address: jan.benda@uni-tuebingen.de (Jan Benda)
[1] Funding:
A.R. is funded by DFG grant *BE3699* which is part of the priority program *SPP 2205: "Evolutionary optimization of neural processing"*.

the fundamental signals in electrocommunication. Beat amplitude declines with distance between the two fish and beat frequency is modulated by various electrocommunication signals (Benda, 2020). Beats are encoded by tuberous electroreceptors distributed all over the body (Carr et al., 1982). T-units fire phase locked to each EOD cycle and their spike timing encodes phase modulations (Scheich et al., 1973). P-units respond with changes in their firing rate to AMs (Bastian, 1981; Nelson et al., 1997; Benda et al., 2005), similar to auditory fibers. So far, P-unit tuning to beat frequencies has been analyzed in a range up to 300 Hz. Beyond the strongest firing rate modulations in response to beat frequencies of 60 – 100 Hz the response declines down to baseline at about 250 Hz (Bastian, 1981; Benda et al., 2006; Walz et al., 2014).

Here we address the apparent mismatch between P-unit tuning and behavioral relevant beat frequencies. We record P-unit activities of *Apteronotus lepthorynchus* in response to a much wider range of difference frequencies than was used before ($-750 \leq \Delta f \leq 2500$ Hz), also exceeding the ranges observed for inter- and intra species encounters in the field (Henninger et al., 2018, 2020). Slow beats repetitively reoccur at octaves of the carrier EOD and P-unit firing rates follow this pattern known from early psychophysical experiments (Ohm, 1839; Helmholtz, 2009). By mathematical reasoning and by modelling the signal transduction pathway (Chacron et al., 2001; Savard et al., 2011; Sinz et al., 2020) we then explain how single sensory cells extract beat frequencies at multiple octaves. Finally, we test our model predictions in behavioral experiments based on the jamming avoidance response (Watanabe and Takeda, 1963).

2. Results

We studied the encoding of beats over a wide range of frequencies by presenting sinusoidal electrical stimuli with absolute frequencies ranging from 20 up to 3200 Hz to electric fish of the species *A. leptorhynchus*. The stimuli superimpose with the fish's own electric field and result in beating amplitude modulations. We recorded the spiking activity of $n = 42$ P-unit electroreceptors in response to these beats.

Responses to low difference frequencies P-units are known to respond to low-frequency beats by modulating their firing rate (Bastian, 1981; Nelson et al., 1997; Benda et al., 2005; Walz et al., 2014). For a fish with EOD frequency $f_{EOD} = 664$ Hz and a stimulus with frequency $f_{stim} = 730$ Hz, mimicking a fish with higher EOD frequency, we get a beat at the difference frequency $\Delta f = f_{stim} - f_{EOD} = 66$ Hz (Fig. 1 C, top). A P-unit responds to this amplitude modulation as demonstrated by the spike raster and the corresponding firing rate (Fig. 1 C, center). The strongest peaks in the power spectrum of the spike response are at the beat frequency (red) and the receiver's EOD frequency (gray circle, Fig. 1 C bottom).

When f_{stim} is 66 Hz below f_{EOD}, the resulting beat is similar to the one discussed above, although at a negative difference frequency. The P-unit response has the same features as for the positive difference frequency and the spectrum has prominent peaks at the same locations (compare Figs. 1 B and C). Signals generated by positive or negative Δf differ only in small phase shifts in the carrier, encoded by another population of electroreceptors, the T-units (Scheich et al., 1973).

As f_{stim} increases, the beat frequency increases accordingly and the response strength, characterized by the modulation depth of the P-unit's firing rate, declines (Fig. 1 D), confirming previous results (Bastian, 1981; Nelson et al., 1997; Savard et al., 2011; Walz et al., 2014).

Beat frequency does not match difference frequency for high stimulus frequencies In the examples discussed so far, $|\Delta f|$ and the frequency of the beating pattern induced in the receiver's EOD are identical. Increasing (or decreasing) f_{stim} beyond $f_{EOD} \pm f_{EOD}/2$, however, breaks this relation. Instead, at $f_{stim} = 0.1 f_{EOD}$, $2.1 f_{EOD}$, or $3.1 f_{EOD}$ the resulting beats have the same frequency as for $f_{stim} = 1.1 f_{EOD}$, a stimulus close to the receiver's EOD frequency (Fig. 1 A, E, F, top). Even for extremely high difference frequencies the resulting beat can be rather slow.

P-units respond to an extremely wide range of stimulus frequencies These slow beats modulate the P-unit's spike responses for difference frequencies up to roughly 1400 Hz, far beyond the known response range up to $\Delta f \approx 300$ Hz (Fig. 1 C – G). Also, at large negative difference frequencies corresponding to absolute stimulus frequencies down to 0 Hz, a range commonly assumed to be primarily driving the ampullary electrosensory system (Kalmijn, 1974; Engelmann et al., 2010; Grewe et al., 2017), we observed clear responses (Fig. 1 A). The multiple ranges of stimulus frequencies clearly driving the P-units are centered on harmonics of f_{EOD}.

Aliasing structure of beat responses Close to integer multiples of f_{EOD} (i.e. its harmonics) we see slow beats and strongly modulated spike responses. Towards odd multiples of $f_{EOD}/2$, beats get faster and spike responses get weaker. The position of the strongest peak in the response power spectrum below $f_{EOD}/2$ is the frequency of the P-unit's firing modulation. Plotting this frequency as a function of f_{stim} reveals a repetitive pattern shaped like a "Toblerone" (Fig. 1 H). This zig-zag pattern resembles the phenomenon of aliasing known from the sampling theorem, with $f_{EOD}/2$ being the Nyquist frequency (Fig. S2). This analogy suggests that the carrier, the receiver's EOD, samples the stimulus with its peaks, as has been suggested previously (Sinz et al., 2020). Thus, f_{EOD} defines the scale in which f_{stim} has to be interpreted. Accordingly, from now on we express f_{stim} relative

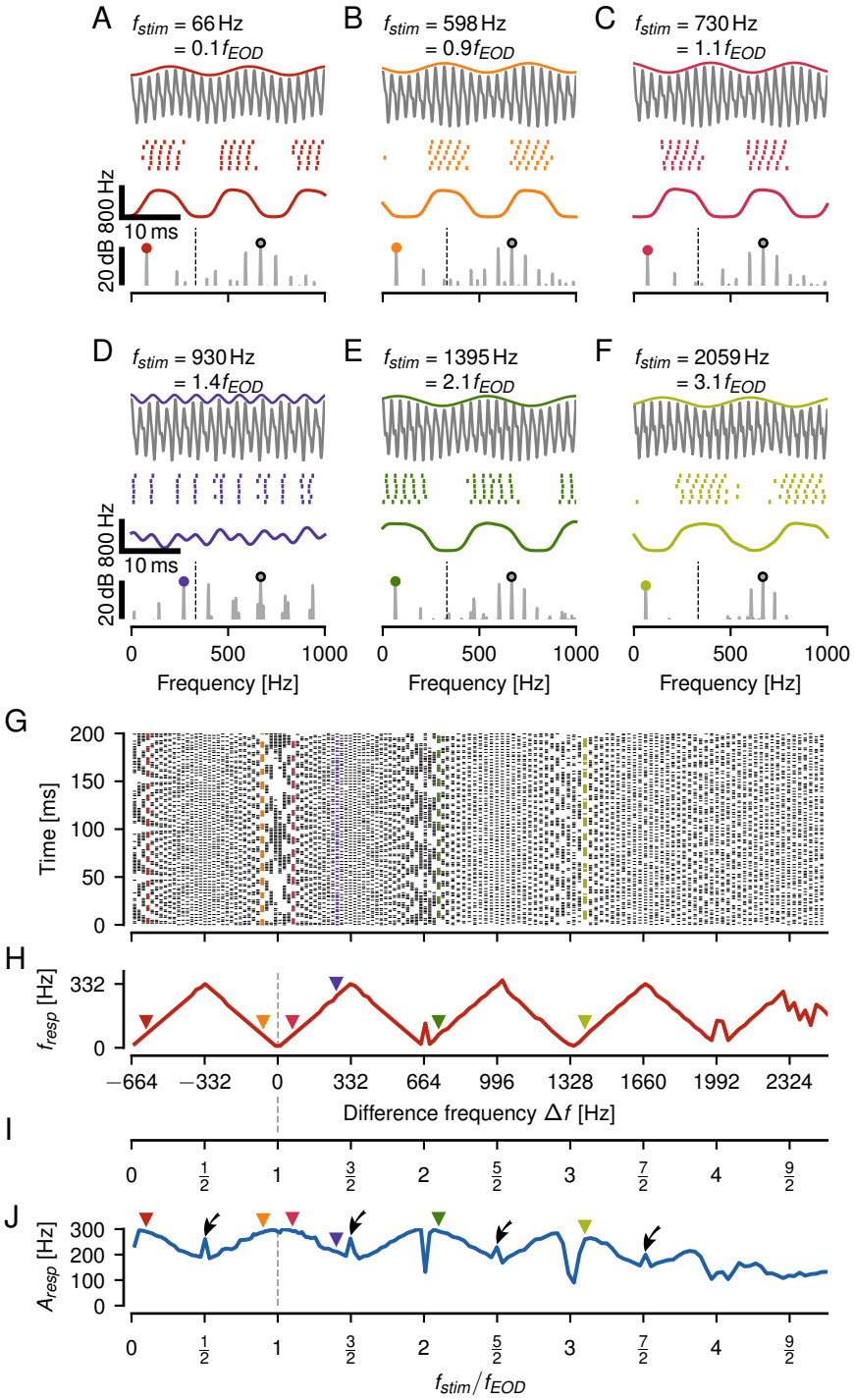

Figure 1: Encoding of a wide range of difference frequencies in an example P-unit. **A–F** Beating amplitude modulation (colored line) of the carrier EOD (gray, top), spike raster and instantaneous firing rate of evoked P-unit response (center), and corresponding power spectrum (bottom, gray circle: f_{EOD}, dashed line: $f_{EOD}/2$, colored circle: strongest frequency below $f_{EOD}/2$) of a few selected stimulus frequencies. **G** Vertical raster plots to a wide range of stimulus frequencies (one trial per frequency) indicate a repetitive structure of P-unit responses. Colored rasters mark examples shown in panels A–F. **H** Frequency tuning, f_{resp}, of the P-unit response, i.e. frequency of its firing rate modulation, retrieved as the strongest frequency below $f_{EOD}/2$, grows with stimulus frequency and repeats every integer multiple of f_{EOD}. Colored triangles mark examples shown in panels A–F. **I** Alternate stimulus frequency axis in multiples of f_{EOD}, i.e. $\Delta f/f_{EOD} + 1$. **J** Amplitude tuning curve, quantified as the amplitude of the peak at f_{resp} (square-root of the integral below the peak at f_{resp} of the power spectrum of the spike response convolved with a Gaussian kernel with $\sigma = 0.5$ ms), also repeats at harmonics of f_{EOD}. Strongest responses are close to multiples of f_{EOD}. Right at odd multiples of $f_{EOD}/2$ responses are enhanced (arrows).

to f_{EOD}, which also allows for comparisons across animals with distinct EOD frequencies (Fig. 1 I).

Periodic amplitude tuning curve The amplitude A_{resp} of the largest peak in the response spectrum below $f_{EOD}/2$ reflects the strength of the P-unit response, i.e. the modulation depth of its firing rate. The resulting tuning curve shows the same repetitive structure as the frequency of the spike response (Fig. 1 J). Close to harmonics of f_{EOD} the response is strongest. These maxima, however, become smaller the higher f_{stim}. Right at the harmonics we observe dips in the response amplitudes which can be attributed to the P-unit's spike-frequency adaptation (Benda et al., 2005). Response amplitudes decline as the stimulus frequency approaches odd multiples of $f_{EOD}/2$. Exactly at odd multiples of $f_{EOD}/2$, response amplitudes often are markedly elevated (arrows in Fig. 1 J). For higher stimulus frequencies response amplitude increases again towards the next harmonic. This increase of the response amplitude beyond $f_{stim} = 1.5 f_{EOD}$ was not expected given previous data on the encoding of amplitude modulations in P-units (Bastian, 1981; Nelson et al., 1997; Savard et al., 2011; Walz et al., 2014).

Amplitude tuning depends on synaptic filtering The synapse between P-units and their target neurons in the electrosensory lateral line lobe acts as a low-pass filter (Berman and Maler, 1998). The shape of the P-unit's amplitude tuning curve strongly depends on this low-pass filtering. It is relatively flat when computed directly from the spike trains but becomes more modulated when the spike train is low-pass filtered (Figs. 2 and S1). Low-pass filtering the spike train with a physiologically plausible Gaussian kernel with $\sigma = 0.5\,\text{ms}$ attenuates high beat frequencies while keeping low beat frequencies untouched, resulting in a periodically modulated amplitude tuning curve.

Sensitive cells respond to a larger frequency range P-unit responses scale with stimulus amplitude (Bastian, 1981; Nelson et al., 1997) and different P-units differ in their sensitivities to a global stimulus (Grewe et al., 2017). To account for both, sensitivity and stimulus intensity, we quantified the P-unit's response amplitude to a standard stimulus, a $\Delta f = 50\,\text{Hz}$ beat (Fig. 2 D). This quantifies the modulation depth of a P-unit's firing rate evoked by the 50 Hz beat and ranged from 70 to 360 Hz.

Based on these amplitudes we divided the P-unit recordings into three sensitivity categories (Fig. 2 D). The responses of the most sensitive cells show the repetitive aliasing structure of the frequency and amplitude tuning curves up to about 3.5 harmonics (Fig. 2 A). Low-pass filtering with a postsynaptic kernel of 0.5 ms standard deviation enhances the repetitive structure of the amplitude tuning curve. The less sensitive a P-unit, the less far its response follows the expected aliasing frequencies and the weaker its amplitude tuning curve is modulated (Fig. 2 B, C).

In the following sections we explore prerequisites necessary for neurons to extract the aliased frequency of the amplitude modulations of beats. The resulting theory fully explains the experimental observations, including the enhanced responses at odd multiples of $f_{EOD}/2$ (arrows in Figs. 1 J).

Neither the analytic signal nor squaring explains the aliasing structure of the beat As suggested by Figs. 1 and 2 above, the aliasing structure of the P-unit response is inherited from the amplitude modulation resulting from the superposition of the EODs of the two fish. To understand the P-unit responses we need to understand how the amplitude modulation of a beat can be retrieved from the superposition of the two EODs, i.e. how a peak at the beat frequency can be generated. This is a generic problem independent of the electrosensory system and thus we express the problem in terms of two cosine waves: a carrier, the EOD of the receiving fish, with frequency $\omega_1 = 2\pi f_1$ and amplitude one and a stimulus, the EOD of another fish, with frequency $\omega_2 = 2\pi f_2$ and an amplitude α measured relative to the amplitude of the carrier (stimulus contrast). Both signals superimpose:

$$x(t) = \cos(\omega_1 t) + \alpha \cos(\omega_2 t) \tag{1}$$

The resulting signal $x(t)$ shows the characteristic beating amplitude modulation, but the spectrum has peaks only at the original stimulus frequencies $\pm\omega_1$ and $\pm\omega_2$. There are no peaks at the beat frequency (Fig. 3 A). A non-linear operation needs to be applied to the signal to generate spectral peaks at the observed beat frequency. Commonly used non-linearities to retrieve the amplitude modulation of a beat are the absolute value of the analytic signal obtained by means of a Hilbert transformation, squaring, or thresholding (Middleton et al., 2006).

Both the analytic signal and squaring predict the beat frequency to be identical to the difference frequency. This is exactly what we expect for low difference frequencies, i.e. for stimulus frequencies ω_2 close to ω_1. However, for higher difference frequencies the beat frequency is predicted to keep growing (Fig. S3). These non-linearities do not explain the aliasing structure we observe in the beats and the P-unit responses (Fig. 1 H).

Thresholding explains aliasing at odd multiples of the carrier frequency A threshold non-linearity

$$x_{th}(t) = \lfloor x(t) \rfloor_0 = \begin{cases} x(t) & ; \quad x(t) \geq 0 \\ 0 & ; \quad x(t) < 0 \end{cases} \tag{2}$$

sets all negative values of a signal to zero. Only the positive half-waves of the beating signal are passed through (Fig. 3 C, left).

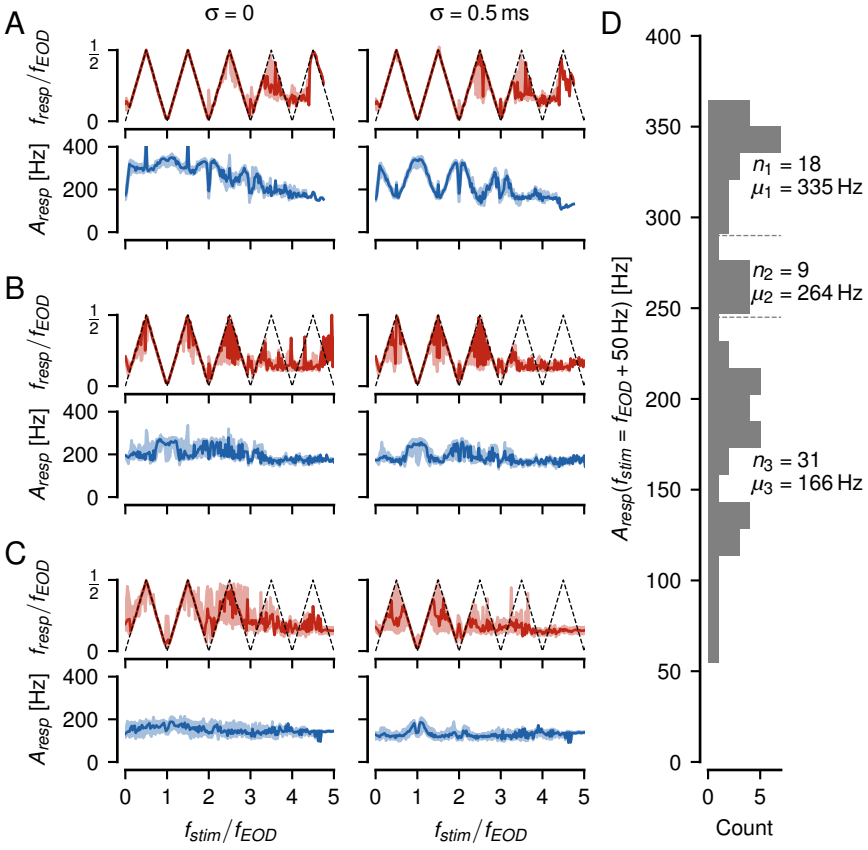

Figure 2: Aliasing structure of population averaged responses. P-unit recordings were classified according to their sensitivity in D. **A** The frequency tuning curve of good encoders ($n_1 = 18$) follows the aliased beat frequency up to almost $3.5 f_{EOD}$. **B** Medium encoders ($n_2 = 9$) respond up to almost $3 f_{EOD}$. **C** Bad encoders ($n = 31$) respond up to about $2.5 f_{EOD}$. Plots show median (solid line) and interquartile range (shading) of the frequency f_{resp} (top) and amplitude A_{resp} (bottom) of the strongest peak in the response spectrum below $f_{EOD}/2$. Response spectra were computed directly from binary spike times (resolution of 40 kHz, left column) or from spike trains convolved with a Gaussian kernel with $\sigma = 0.5$ ms (right column). **D** Distribution of each recording's sensitivity quantified as the amplitude A_{resp} of the response evoked by a $\Delta f = +50$ Hz beat with a contrast of 10 % or 20 %. The recordings were categorized as indicated by the dashed lines. Note, that cells that have been measured with both 10 and 20 % contrast contribute two recordings to the histogram.

To make this operation analytically tractable, we approximate the thresholding operation, Eq. (2), by a multiplication with a pulse train of the same frequency ω_1 as the carrier signal (Fig. 3 B left):

$$x_{th}(t) \approx x(t) \cdot p(\omega_1 t) \tag{3}$$

The pulse train $p(\omega_1 t)$, Eq. (S8), multiplies positive halfwaves with one and negative halfwaves with zero (Fig. 3 C left). Note that this approximation is valid only in the limit $\alpha \to 0$. For larger amplitudes the stimulus distorts the carrier, shifting also its zero crossings. The spectrum of the pulse train has a peak at zero frequency, Eq. (S10), and peaks at all odd multiples of the carrier frequency ω_1, Eq. (S9), (Fig. 3 B, right).

According to the convolution theorem a multiplication in time equals a convolution in the Fourier domain. Thus, the Fourier spectrum of a thresholded beat (Fig. 3 C) is given by the convolution of the spectrum of the beat (Fig. 3 A) and that of the pulse train (Fig. 3 B). A first component of the resulting spectrum is the convolution of the carrier (ω_1) with the pulse train. This results in peaks at even multiples of ω_1, Eq. (S12), and at $\pm\omega_1$, Eq. (S13) (horizontal lines in Fig. 3 D). These frequency components do not make up the amplitude modulation, because they do not depend on ω_2.

The second component of the spectrum, the convolution of the stimulus (ω_2) with the pulse train, provides side peaks at $\pm\omega_2$ to all the peaks of the pulse train. These peaks can explain the aliasing structure of the beat around odd multiples of ω_1, Eq. (S15), and around zero frequency, Eq. (S16). Since they depend on ω_2, they form the diagonal lines in Fig. 3 D. These peaks fall into the range below the Nyquist frequency ($\omega_1/2$, gray area in Fig. 3 C, D, F) around odd multiples of ω_1, indicating slow beat modulations.

A few example spectra of the thresholded beat are shown in Fig. 3 F. While there is a peak below the Nyquist frequency for ω_2 at 0.15, 1.15, and 3.15 times ω_1 (Fig. 3 F_i, C, F_{iii}), it is missing around even multiples of ω_1 (2.15 and 4.15 times ω_1 in Fig. 3 F_{ii}, F_{iv}). In contrast to the am-

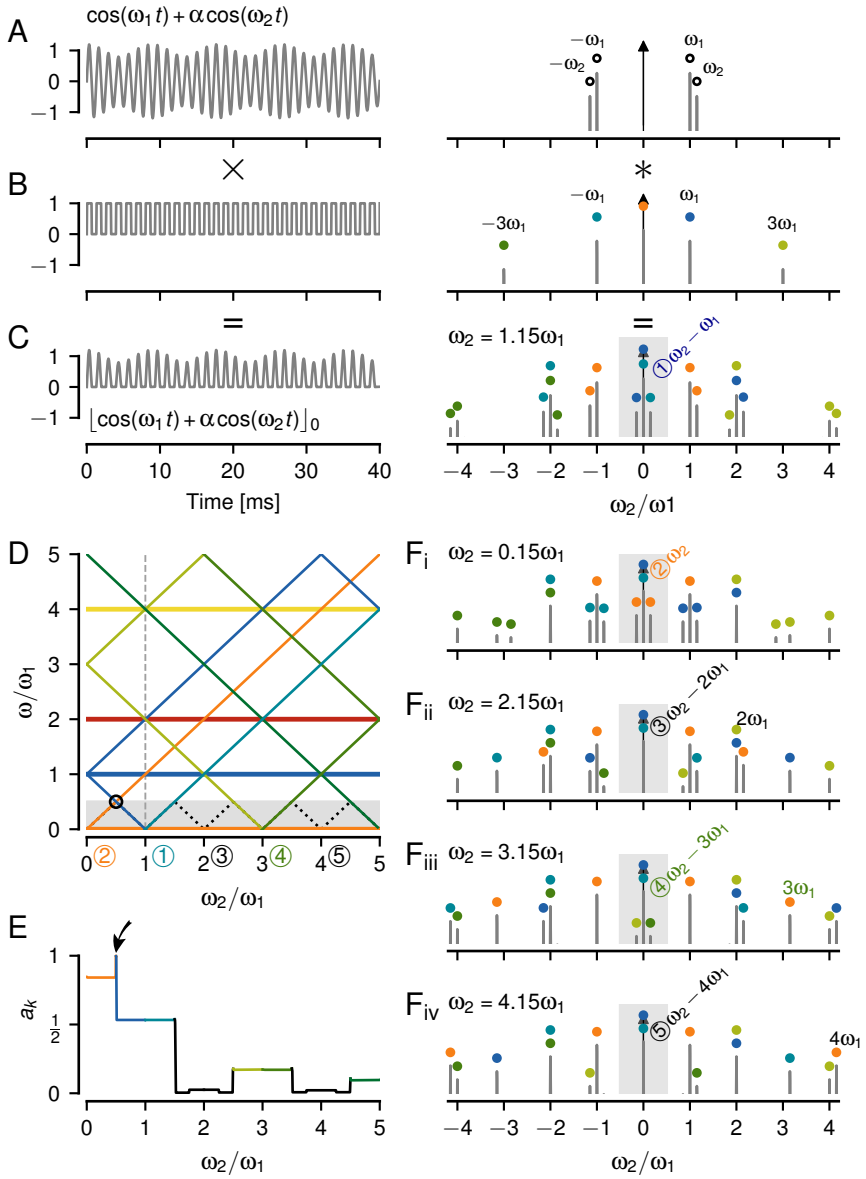

Figure 3: Thresholding is not sufficient to explain full aliasing structure of beats. **A** A beat generated by summing up two frequencies ω_1 and ω_2 (left) only contains these two frequencies in its Fourier spectrum (right). **B** A pulse train, Eq. (S8), used to approximate a threshold operation, Eq. (2), has peaks at odd multiples of ω_1 and at zero. **C** The thresholded beat, approximated by multiplying the beat with the pulse train, has a rich spectrum that can be computed by convolving the spectrum of the beat with the one of the pulse train (colored circles). Here, for $\omega_2 = 1.15\omega_1$ a peak appears at the difference frequency $\omega_2 - \omega_1$ below the Nyquist frequency at $\omega_1/2$ (gray area). This peak describes the slow amplitude modulation of the beat in **A**. **D** The position of all peaks of the thresholded beat, Eqs. (S12), (S13), (S15), and (S16) (solid lines), as a function of stimulus frequency ω_2. The black dashed line marks the expected aliased frequencies, and the circle the only crossing of peaks below the Nyquist frequency (gray area). **E** Amplitude of the peaks in **D** below the Nyquist frequency, Eqs. (S15) and (S16), that basically reflect the the respective Fourier coefficient of the pulse train, Eqs. (S9) and (S10). At the crossing of peaks (circle in **D**) amplitudes sum up (arrow). **F** Further examples of power spectra for various stimulus frequencies as marked by circled numbers in **D**. Same color scheme as in **C**.

plitude of the analytic signal and to squaring, the thresholding operation introduces many additional peaks in the spectrum, necessary for explaining some but not all of the aliasing structure of beats.

Threshold cubed fills in additional beat frequencies How can we fill in the missing components in the spectrum around even multiples of ω_1? Let's try to make the threshold operation more non-linear by raising its output to a power of three:

$$x_c(t) = \lfloor x(t) \rfloor_0^3 \tag{4}$$

Again, this can be approximated by taking the beat signal to the power of three and multiplying the result with a pulse train (Fig. 4 A–C). The spectrum of the cubed beat signal has $2^3 = 8$ peaks, Eq. (S18) (two times convolution of two peaks with themselves). Of those the peaks at $|2\omega_1 - \omega_2| = |\omega_1 - \Delta\omega|$ (Fig. 4 D, red and purple) are the only relevant additions in comparison to the threshold without exponent. The convolution of these peaks with the zero-frequency peak of the pulse train fills in the missing beat frequencies around twice the carrier frequency (Eq. (S24), Fig. 4 E, red and purple). Threshold cubed can explain beat frequencies up to inclusively three multiples of ω_1.

Higher powers would introduce more peaks in the spectrum and fill up towards higher multiples of ω_1. In the limit of infinitely high exponents this would approach the situation of the sampling theorem — the half-waves of the carrier would get infinitesimally narrow and sample the stimulus at integer multiples of the carrier period $2\pi/\omega_1$. But the higher the power, the more the amplitude modulation is distorted.

The amplitudes of the peaks below the Nyquist frequency decline in a step-wise manner for each harmonic, Eqs. (S22), (S23), and (S24). At $\frac{1}{2}\omega_1$, $\frac{3}{2}\omega_1$, and $\frac{5}{2}\omega_1$ spectral peaks cross each other (circles in Fig. 4 E). Their respective amplitudes add up and result in elevated amplitudes exactly at these frequencies (Fig. 4 F, arrows). These peak crossings explain the elevated P-unit responses at these frequencies (arrows in Fig. 1 J).

Non-linear spiking dynamics cannot explain responses to higher beat frequencies Eventually, a spike generator, here modeled as a leaky integrate-and-fire neuron (LIF) with adaptation, Eqs. (A.1), (A.3), (A.4), encodes the extracted amplitude modulation in a train of action potentials (Fig. 5 D). Neither the hard threshold non-linearity the LIF applies on the membrane voltage nor the smooth voltage-threshold, Eq. (A.2), of the exponential integrate-and-fire neuron (EIF, Fig. 5 E, Fourcaud-Trocmé et al., 2003) can replace the power of three to extract the amplitude modulation of a beat resulting from a stimulus at around twice the carrier frequency. The non-linear dynamics of a spike generator can not generate the aliasing structure of the P-unit responses to beats. Rather, a sufficiently strong static non-linearity (Fig. 5 B) has to be applied to the beat, such that the necessary low-frequency peak in the spectrum is generated. Subsequent low-pass filtering then isolates this peak (Fig. 5 C) and this is what the spike generator responds to.

A power of three describes P-unit responses best For a more systematic evaluation which exponent on the threshold operation describes the P-unit responses best, we simulated LIF models using exponents at the threshold non-linearity ranging from $p = 0.2$ to 5. The resulting frequency and amplitude tuning curves were compared to the experimentally measured ones (Fig. 6). LIF models with a threshold non-linearity without exponent (power of one), Eqs. (A.1), (A.3), (A.4) (Sinz et al., 2020), were individually fitted to response characteristics of $n = 9$ cells.

As predicted, a pure threshold without exponent shows responses at the zeroth, first and third harmonic (Fig. 3), but diverges from the measured activity (Fig. 6 A) around the second harmonic (Fig. 6 C). Exponents both higher (Fig. 6 D, E, F) and lower than one (Fig. 6 B) fill in the response at the second EOD multiple. Models with powers of 1.5 and 2.5, and in particular of 0.5, additionally respond to the forth EOD multiple.

To quantify the model performance we computed mean squared errors (MSE) between 18 experimentally measured cells with the 9 cells of similar sensitivity in the model population. The MSEs between frequency tuning curves were minimal at powers of about 0.5, 1.5, and 2.5–3 (Fig. 6 G, left). The MSEs for the corresponding amplitude tuning curves showed similar minima but with the smallest MSE at a power of 3 (Fig. 6 H, right). A power of three indeed describes both the frequency and amplitude tuning curves of P-unit responses best.

Harmonics of the carrier are not sufficient to explain aliasing So far our reasoning was based on pure sine waves. In the electrophysiological recordings, however, the carrier was a real EOD waveform of *A. leptorhynchus*. Using these EOD waveforms instead of sine waves for the models tilted the MSE curves slightly towards higher exponents. Exponents smaller than one increased the MSE, whereas exponents larger than two further decreased the MSE with a minimum at $p = 3$ (Fig. 6 H). In a model with a pure threshold ($p = 1$) the harmonics of the carrier do not contribute to shift the stimulus frequency to all multiples of f_{EOD}. A wider or narrower EOD waveform, however, modifies the aliasing structure introduced by the threshold operation in a way we do not observe in the data. Adding a power of three to the threshold makes the P-unit responses more robust against changes in the EOD waveform (see Fig. S4 and supplement for detailed explanation).

Ambiguous beats evoke similar behavioral responses Our results imply that P-unit responses are am-

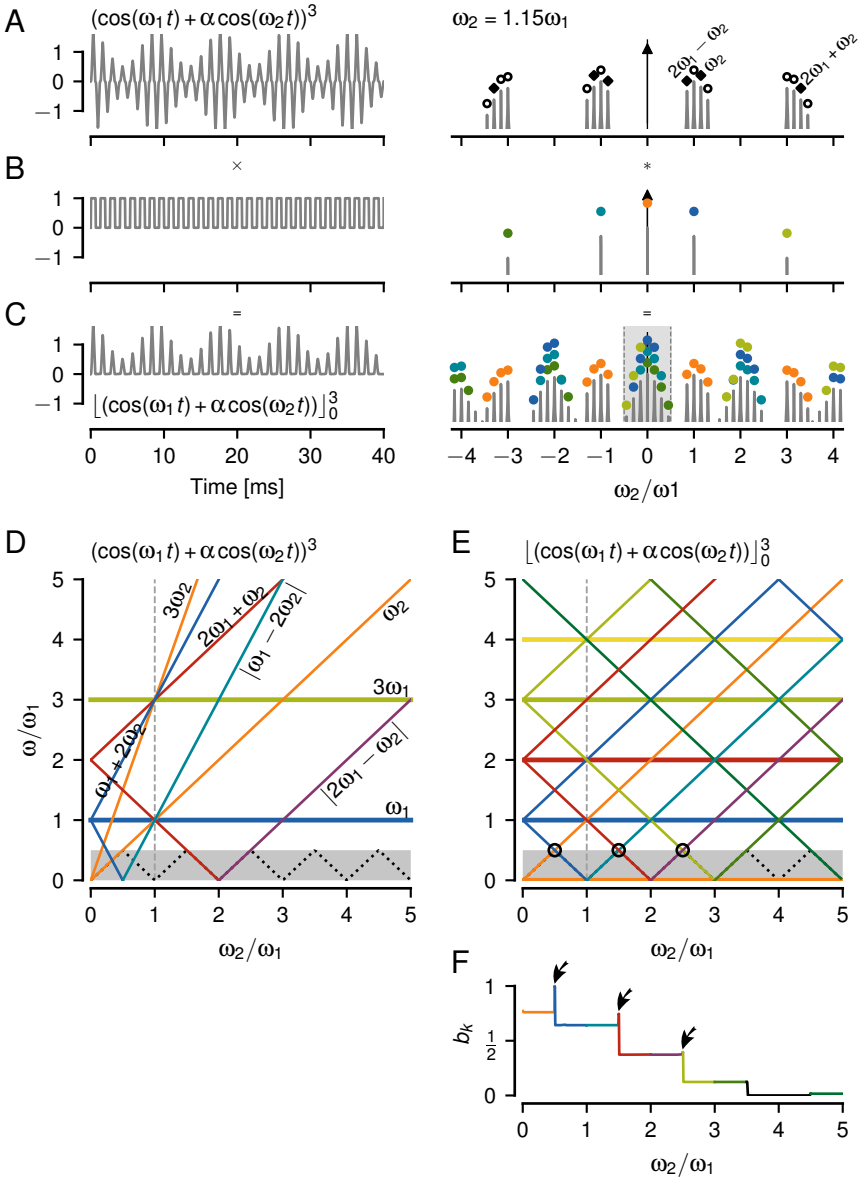

Figure 4: Threshold cubed fills in beat frequencies around the second harmonic of the carrier. **A** The beat cubed (left) has a more pointed carrier and eight peaks in the positive half of the Fourier spectrum. **B** Again we approximate the thresholding operation by a multiplication with a pulse train. **C** The threshold-cubed beat has a much richer power spectrum compared to the thresholded beat without exponent (Fig. 3 C). The colored markers illustrate the convolution process and the gray shading frequencies below the Nyquist frequency. **D** Position of all peaks of the cubed beat as a function of stimulus frequency. Dashed line indicates the expected beat frequencies. **E** Position of the major peaks of the threshold-cubed beat. Beat-frequencies are faithfully extracted up to $\omega_2 = 3.5\omega_1$. **F** Amplitudes of the peaks in E describing amplitude modulation of beats, Eqs. (S22), (S23), and (S24). Crossing peaks in E (circles) add up and result in elevated amplitudes (arrows in F).

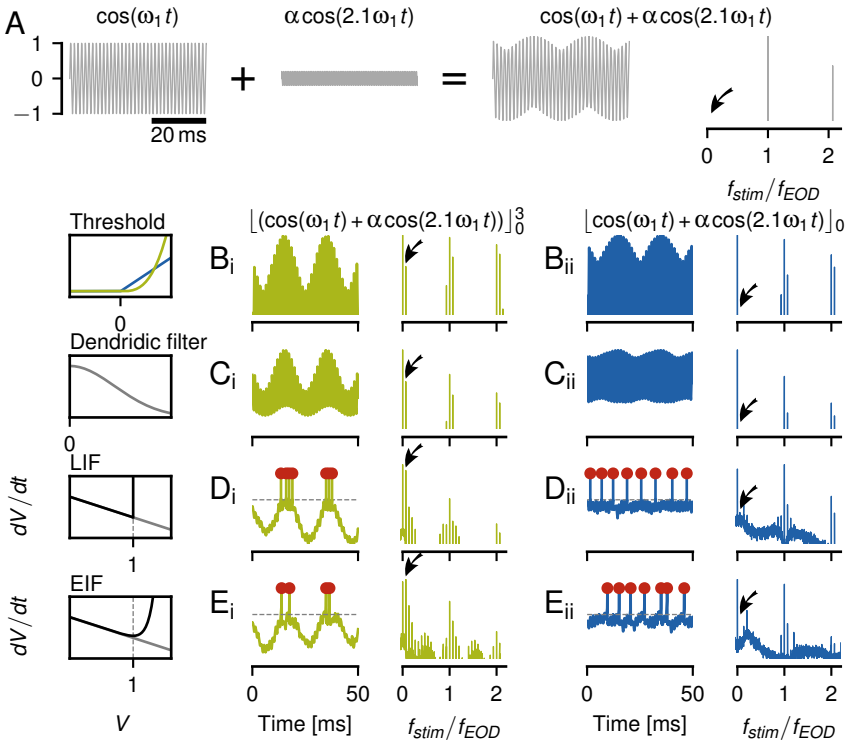

Figure 5: Integrate-and-fire models of P-unit spiking activity. **A** Superposition of two cosine waves with $\omega_2 = 2.1\omega_1$ results in a beat with an amplitude modulation oscillating at $0.1\omega_1$. **B** This signal is thresholded, Eq. (2), (blue) or thresholded and cubed, Eq. (4), (green), potentially at the synapse between the receptor cells and the P-unit afference. **C** The dendritic cable of the afferent acts as a low pass filter Eq. (A.4). **D** The spike generator is modeled by a leaky integrate-and-fire neuron (LIF) with adaptation, Eqs. (A.1) & (A.3). **E** Alternatively, we also tested an adapting exponential integrate-and-fire model (EIF), Eq. (A.2). A threshold raised to a power of three generates a modulated sub-threshold potential that reflects the amplitude modulation of the beat. Both the LIF and EIF respond to this and this way encode the beat (left columns). Without the exponent of three neither the LIF nor the EIF are able to respond to the beat, because the amplitude modulation signal is not contained in the sub-threshold potential any more (right columns).

biguous with respect to the absolute stimulus frequency. Whether the stimulus frequency was close to one, two, or three times f_{EOD} should not be distinguishable by the fish based on P-unit responses.

We tested this hypothesis behaviorally by means of the jamming avoidance response (JAR, Watanabe and Takeda, 1963). When a receiving fish is stimulated with a sinusoidal mimic that is close to but below the own EOD frequency, it will raise its EOD frequency by a few Hertz on a time scale of about 10 seconds (Fig. 7 A). We repeated the experiment with stimulus frequencies 5 Hz below one to five times f_{EOD}. Indeed, all the fish tested ($n = 5$) responded with a JAR to one, two, and three multiples of the EOD frequency (Fig. 7 B). None of the fish responded to four times f_{EOD}, but some fish did respond even at five times f_{EOD} with a very small JAR of less than 1 Hz. The fish also showed a weak JAR to a stimulus with an absolute frequency of 5 Hz. The JARs to non-zero multiples of f_{EOD} closely follow the beat amplitudes predicted by a cubed threshold, Eqs. (S22), (S23), and (S24) (Fig. 7 C). Only the response to the zeroth EODf multiple was much smaller than expected from the beat amplitude.

3. Discussion

We observed that P-units encode a wide range of stimulus frequencies up to about 3000 Hz (Fig. 1). Superposition of the EOD carrier and the stimulus led to beating amplitude modulations. At every integer multiple of the carrier frequency these AMs are slow, between the integer multiples AM frequencies were fast. The repeating pattern resembles aliasing known from the sampling theorem (Fig. S2). In accordance with the beat pattern, the P-unit tuning curves were also periodic in harmonics of the carrier (Fig. 2), but the amplitude of the responses declined with higher stimulus frequencies.

P-unit tuning to high difference frequencies P-unit tuning curves for beat stimuli have so far been only measured for difference frequencies up to about 300 Hz (Bastian, 1981; Nelson et al., 1997; Benda et al., 2006; Walz et al., 2014). The amplitude of the response modulation induced by beating amplitude modulations resembled a band-pass tuning. The reduced response to difference frequencies close to zero (Nelson et al., 1997) reflects the high-pass filter induced by the P-units' fast spike-frequency adaptation (Benda et al., 2005). Towards higher

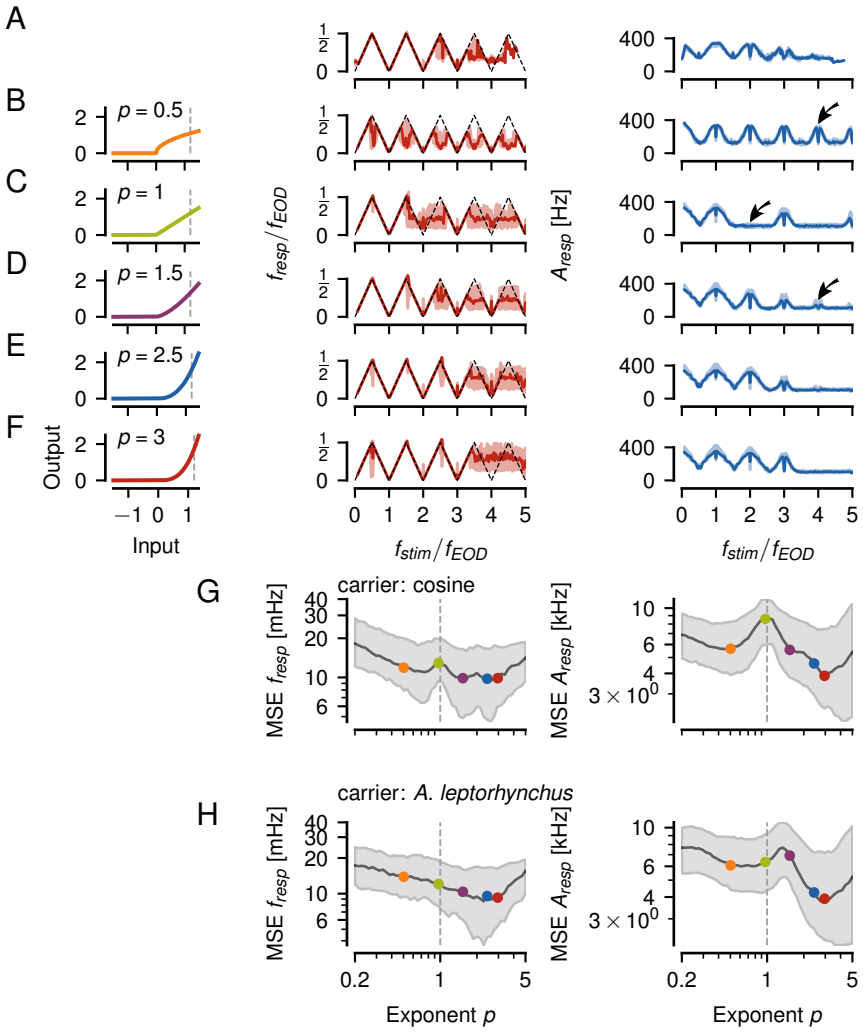

Figure 6: Dependence of LIF model performance on threshold exponent. **A** Frequency (left) and amplitude tuning curves (right, solid line indicates population median and shaded area the interquartile range) of $n = 18$ experimentally measured P-units of highest sensitivity (same as Fig. 2 A). **B–F** Tuning curves simulated from a population of $n = 9$ LIF models (Fig. 5) that have been fitted to individual P-units. Exponent p of the threshold non-linearity applied to the beat as indicated and illustrated in the left column. Arrows indicate missing or additional responses. **G** Mean squared error (MSE, median with interquartile range) between the tuning curves of each experimentally measured P-unit shown in A and each LIF model in dependence on the threshold exponent used in the models. Examples from B–F are indicated by the correspondingly colored circles. Here, both the carrier and the stimulus are pure sine waves. **H** Same as in G but with an EOD waveform of *A. leptorhynchus* as carrier and a pure sine wave as stimulus, resembling the situation in the electrophysiological experiments (see Fig. S4).

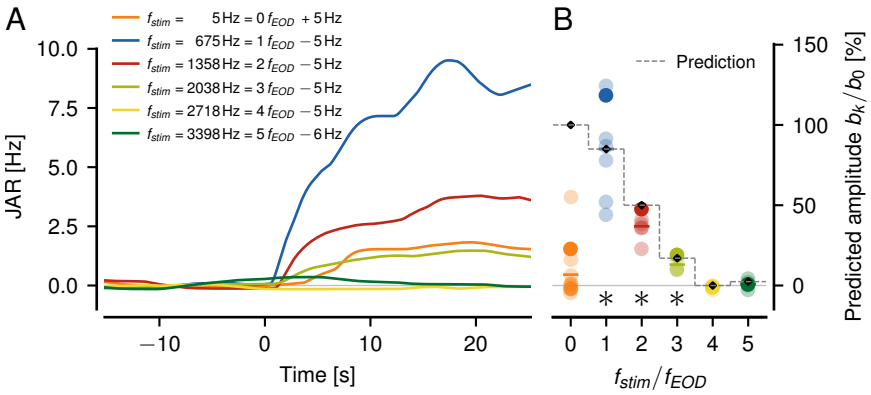

Figure 7: Jamming Avoidance Response (JAR) to higher harmonics of f_{EOD}. **A** JARs of one fish to a 5 Hz beat at 0 to 5 times f_{EOD}. Stimulus onset at 0 s. The fish responds to the sinusoidal stimulus up to at least the third harmonic of its EOD frequency. **B** Final JAR amplitudes of five fish. JARs of the fish shown in **A** are highlighted by more saturated colors. The average JARs (horizontal lines) closely follow the beat amplitudes predicted by a cubed threshold, Eqs. (S22), (S23), and (S24) (dashed line) for non-zero multiples of f_{EOD}. Asterisks indicate significant deviations of the JAR from zero ($\alpha < 5\,\%$).

beat frequencies the response was found to steadily decline towards zero as expected for a spiking neuron (Fourcaud-Trocmé et al., 2003) and as set by the neuron's baseline firing rate (Knight, 1972).

The experimental findings reported here clearly demonstrate that P-units do show responses to difference frequencies beyond 300 Hz. Instead of the expected steady decrease, the tuning repeats at harmonics of the EOD carrier frequency which has profound consequences for the encoding of chirps, electrocommunication signals that transiently increase the difference frequency (Benda et al., 2005). The P-unit response to chirps could be explained by transient firing rate modulations mediated by the P-units' beat tuning curve (Walz et al., 2014). In a field study the behavioral relevance of chirps at difference frequencies beyond 400 Hz and thus beyond the Nyquist frequency of the female has been shown (Henninger et al., 2018). A monotonously declining tuning curve would not suffice to explain how chirps could be encoded at such high beat frequencies. The repetitive tuning we describe here increases again beyond the Nyquist frequency and thus leads to changes in the firing modulation induced by chirps. The small peaks in the tuning curve at exactly half f_{EOD}, that can be explained as a crossing of two peaks in the spectrum (Fig. 4), might add an interesting aspect to the encoding of chirps.

In this context, the low-pass filtering of a post-synaptic kernel plays an important role (Fig. S1). If the kernel is too narrow, then the tuning curve of the P-units is almost flat and firing rate modulations caused by chirps would be quite small. If the kernel is too wide, the tuning curve is only modulated within a narrow range of stimulus frequencies around the harmonics of the carrier. Only for kernels resembling the experimentally measured postsynaptic potentials ($\sigma = 0.5$ ms, Berman and Maler, 1998) is the tuning curve fully modulated without being flat between the harmonics (Fig. 2). We therefore hypothesize electric fish with lower EOD frequencies to have correspondingly wider postsynaptic potentials.

Beat extraction at high difference frequencies
The repetitive tuning curve of P-units is in a way trivial in that it simply follows the amplitude modulations of beats. But how is this apparent amplitude modulation extracted from the original signal, which does not contain the beat frequency in its power spectrum?

Usually, beat frequency is considered as equalling the difference frequency (Walz et al., 2014; Joris et al., 2004). A commonly used method to extract amplitude modulations is the magnitude of the analytic signal obtained via Hilbert transform (Middleton et al., 2006; Longtin et al., 2008; Carriot et al., 2017; Stamper et al., 2012), which also predicts the beat frequency to be given by the difference frequency (Fig. S3 A). Similarly, squaring the signal (Fig. S3 B) generates a peak in the spectrum at the difference frequency. However, the observed beat frequencies equal difference frequencies only for difference frequencies below the Nyquist frequency and thus are not sufficient to explain our results.

Yet another common method for computing amplitude modulations of, for example, EMG or EEG signals (Myers et al., 2003) or acoustic signals (Khanna and Teich, 1989), is thresholding. A threshold, also known as a Rectifying Linear Unit (ReLU) in deep learning, extracts beat frequencies only at odd harmonics of the carrier (Fig. 3). Rather, a threshold operation followed by exponentiation is required to generate peaks in the power spectrum at integer multiples of the carrier frequency (Fig. 4). Neither the non-linearity of action-potential generation (Fig. 5), nor higher harmonics of the carrier and the signal are sufficient to substitute an exponentiated threshold to explain our data (Fig. S4).

In deep learning it is common to apply smooth threshold functions such as ELU (Clevert et al., 2015) or Softplus (Glorot et al., 2011). All of these are potential alternatives for the threshold raised to a power of three we suggest to be

at work in P-units, because they all could be approximated by a ReLU raised to a power of three in the vicinity of their threshold. The same holds true for sigmoidal activation functions. As they are discussed for the transformation of a hair-cells membrane voltage by their ribbon synapse (Peterson and Heil, 2019). Only for larger inputs their saturation will lead to noticeable deviations.

Sinusoidal amplitude modulations (SAMs) versus beats For characterizing signal processing of the electrosensory system (Bastian, 1981) and in particular of the mammalian auditory system (Joris et al., 2004), SAM stimuli of various frequencies and amplitudes have been used. SAM stimuli multiply a carrier with a periodic amplitude modulation Eq. (S6) and differ from beat stimuli by having two peaks in the spectrum instead of a single one which are located at $\pm\Delta f$ flanking the carrier frequency (Fig. S5). The additional peak induced by a SAM stimulus already fills in responses at the second harmonic of the carrier when used in conjunction with a power-of-one threshold Fig. S5 E. An effect that would have obscured the underlying mechanism, i.e. the power of three threshold, if we had applied SAMs instead of realistic beats.

Relation to the sampling theorem In the limit to infinite power the thresholded and exponentiated carrier approaches Dirac delta functions positioned at multiples of the carrier's period. This pulse train can be thought to sample the stimulus waveform with the carrier frequency exactly like in the setting of the sampling theorem (Fig. S2). However, the stimulus waveform also gets transformed by the threshold and the power operation. The higher the exponent the larger the distortion of the extracted amplitude modulation. Thus, the exponent should be not too large in order to maintain an accurate representation of the amplitude modulation. In this sense, a sharp threshold without exponent would be ideal.

Physiological mechanisms for beat extraction The most likely site for the threshold operation are the ribbon synapses of the electroreceptor cells onto the afference (Northcutt, 1986) as has been suggested previously (Chacron et al., 2001). It is quite unlikely that these synapses act like a perfectly sharp threshold. Rather there will be a smooth transition in a way that is best described by a cubic power applied to the threshold (Fig. 5 B). Indeed, cooperativity of calcium channels in the presynapse has been discussed for hair cells in the auditory system to result in powers of three or higher (Roux et al., 2006; Michalski et al., 2017; Peterson and Heil, 2019).

Before the postsynaptic potential reaches the spike-generation site it is low-pass filtered by the dendrite of the afferent (Sinz et al., 2020). With the right time-constant this isolates the low-frequency amplitude modulation but does not entirely remove the EOD (Fig. 5 C).

Sensitivity curves measured for a range of EOD frequencies demonstrate that P-units are tuned to a fish's EOD frequency (Hopkins, 1976) and to limit P-unit responses to stimulus frequencies close to the EOD frequency. The corresponding band-pass filter is probably caused by electric resonance in the electroreceptor cells (Viancour, 1979). Adding a damped oscillator, Eq. (S28), to our P-unit models, Eqs. (A.1) – (A.4), reproduces P-unit tuning to EOD frequency (Fig. S6 A–B) but does not impair responses to beats at high difference frequencies (Fig. S6 C–D).

Ambiguity in beat perception The decline in response amplitude with higher harmonics of the carrier (Fig. 2) is a consequence of the beat amplitude that gets smaller with higher harmonics (Fig. 4 F). Since the beat amplitude also depends on the distance between two fish — the larger the distance the smaller the beat amplitude, the P-units' response is ambiguous and the fish should not be able to resolve this based on the firing rate of the P-units.

The jamming avoidance response (JAR) is evoked by stimulus frequencies close to the receivers's own EOD frequency and also by stimulus frequencies close to harmonics of the carrier (Watanabe and Takeda, 1963). The amplitude of the JAR closely follows the beat amplitude extracted by a cubed threshold (Fig. 7). Also the behavioral threshold for detecting a beat at least qualitatively follows the tuning curve of the P-units (Knudsen, 1974). This suggests that wave-type electric fish indeed can not disambiguate stimuli at different multiples of their EOD frequency.

In contrast, the JAR at stimulus frequencies close to the zeroth multiple of f_{EOD} is much smaller than the amplitude of the corresponding beat. Such low-frequency stimuli also evoke responses in ampullary cells of the passive electrosensory system (Kalmijn, 1974; Engelmann et al., 2010; Grewe et al., 2017). Combining ampullary and P-unit information would allow the fish to disambiguate the P-unit response and to inhibit the JAR response.

Perception of other wave-type species The wide range of difference frequencies covered by P-unit responses (Fig. 2) extends far beyond the range of EOD frequencies covered by conspecific wave-type electric fish (usually about one octave). Consequently, the fish should be able to detect the presence of sympatric species covering higher or lower frequency ranges than conspecifics (Steinbach, 1970; Hopkins, 1974; Kramer et al., 1981; Stamper et al., 2010; Henninger et al., 2020). Whether and how the different species interact or communicate is an open issue that could be resolved by analyzing electrode-array data recorded in the field (Henninger et al., 2020).

Beat perception in the auditory system Early psychophysical experiments with interacting pure tones

demonstrated that beats are not only perceived at low difference frequencies but also for mistuned octaves, when the second tone is close to octaves of the first tone (Roeber, 1834; König, 1876). These experimental findings were formalized in the 19th century by Ohm (1839) and Helmholtz (2009). Underlying physiological mechanisms, i.e. non-linear mechanisms necessary for extracting beat frequencies, are still debated, in part because the focus of auditory neuroscience has been on the encoding of SAMs (Joris et al., 2004).

Distortion-product otoacoustic emmissions (DPOEs) are hallmark non-linear phenomena of the ear that have been attributed to the mechanic properties of the cochlea and in particular to the active amplification of outer hair cells (Brownell, 1990). The most prominent DPOEs are the cubic distortion, $2\omega_1 - \omega_2$, and the quadratic distortion, $\omega_2 - \omega_1$ (Kujawa et al., 1995), which is the difference frequency. Both frequencies grow linearly in the stimulus frequency ω_2 and thus can not explain the aliasing pattern of the beat frequency. Further, aural harmonics and interactions with combination tones have largely been ruled out by masking experiments (Plomp, 1967).

Our results suggest the non-linear transformation at the ribbon synapse (Michalski et al., 2017) to generate distortion products that extract the aliasing structure of beats. The two tones making up a beat are interacting within a single hair cell. The lower the characteristic frequency of an auditory fibre and the louder the two tones, the wider its effective tuning (Sumner and Palmer, 2012; Evans, 1972), allowing for beats at multiple octaves of the carrier tone exciting the auditory fibre. This is in line with beats being better perceived with louder tones and at higher harmonics the lower the frequency of the carrier signal (Plomp, 1967) and with a second peak at the octave of the carrier stimulus observed in the response gain of auditory fibres to SAMs (Palmer, 1982).

Conclusion The sharp kink introduced by a simple threshold is a mathematical abstraction. Any physiological mechanism implementing this non-linearity, like for example the activation curve of voltage-gated calcium currents, has a rather smooth transition. The cubed threshold we derive from our recordings is a mathematically simple way for modeling such a physiologically realistic smooth threshold. In this sense, the ability of the P-units to extract beats at multiples of the carrier frequency is a by-product of their physiology. Natural and sexual selection need to work on the resulting ambigous code for beat frequencies. For the same reason, mammalian auditory fibers are bound to respond to mistuned octaves and thus should contribute to the percept of beats at higher difference frequencies.

Appendix A. Methods

Electrophysiology 42 P-units were recorded from six weakly electric fish of the species *Apteronotus leptorhynchus* obtained from a commercial tropical fish supplier (Aquarium Glaser GmbH, Rodgau, Germany). The fish were kept in tanks with a water temperature of 25 °C and a conductivity of around 300 µS/cm under a $12\,h : 12\,h$ light-dark cycle. Body sizes oft the fish were between 15 and 17.5 cm and 11.1 and 13.2 g. EODf varied between 558 and 787 Hz. All experimental protocols complied with national and European law and were approved by the Ethics Committee of the Regierungspräsidium Tübingen (permit no: ZP1-16).

Surgery Prior to surgery, anesthesia was provided via bath application of a solution of MS222 (120 mg/l, PharmaQ, Fordingbridge, UK) buffered with Sodium Bicarbonate (120 mg/l). For the surgery the fish was fixed on a stage via a metallic rod glued to the skull. The posterior anterior lateral line nerve (pALLN) above the gills, before its descent towards the anterior lateral line ganglion (ALLNG) was disclosed for subsequent P-unit recordings. During the surgery water supply was ensured by a mouthpiece, sustaining anaesthesia with a solution of MS222 (100 mg/l) buffered with Sodium Bicarbonate (100 mg/l).

Experimental setup Fish were immobilized by an initial intramuscular injection of Tubocurarine (Sigma-Aldrich, Steinheim, Germany; 25–50 µl of 5 mg/ml solution). For the recordings fish were fixated on a stage in a tank, with a major part of the body immersed in water. Analgesia was refreshed in intervals of two hours by cutaneous Lidocaine application (2 %; belapharm, Vechta, Germany) around the operation wound and the head mounting rod. Electrodes (borosilicate; 1.5 mm outer diameter; GB150F-8P; Science Products, Hofheim, Germany) were pulled to a resistance of 50–100 MΩ (model P-97; Sutter Instrument, Novato, CA) and filled with 1 M KCl solution. Electrodes were fixed in a microdrive (Luigs-Neumann, Ratingen, Germany) and lowered into the nerve. Recordings of electroreceptor afferents were amplified (SEC-05, npi-electronics, Tamm, Germany, operated in bridge mode) and digitized at 40 kHz (PCI-6229, National Instruments, Austin, TX). RELACS (`www.relacs.net`) running on a Linux computer was used for online spike and EOD detection, stimulus generation, and calibration.

P-unit identification P-units were identified based on their firing properties with a baseline firing rate between 64–470 Hz (Grewe et al., 2017; Gussin et al., 2007; Ratnam and Nelson, 2000), phase locking to the EOD, indicated by multimodal interspike-interval (ISI) histograms, and by responses to amplitude modulations of the EOD.

Electric field recordings Global EOD for monitoring EODf was measured with two vertical carbon rods (11 cm long, 8 mm diameter) in a head-tail configuration. The signal was amplified 200–500 times and band-pass filtered (3 to 1,500 Hz passband, DPA2-FX; npi electronics, Tamm, Germany). A local EOD including the stimulus was measured between two, 1 cm-spaces silver wires located next to the left gill orthogonal to its longitudinal body axis (amplification 200–500 times, band-pass filtered with 3 to 1,500 Hz passband, DPA2-FX; npi-electronics, Tamm, Germany).

Stimulation Sine wave stimuli (10–3300 Hz) imitating another fish were isolated (ISO-02V, npi-electronics, Tamm, Germany) and delivered via two horizontal carbon rods located 15 cm laterally to the fish. Depending on f_{EOD} of the fish, the stimuli resulted in difference frequencies between −750 and 2495 Hz. Each stimulus was repeated twice either for 0.5 s (20% of the trials) or 1 s (80% of the trials). Stimulus amplitude was 10 % or 20 % of the fish's local EOD amplitude (contrast) prior to each stimulation. Cells measured with two different contrasts contribute two recordings to the subsequent analysis.

Data analysis Data analysis was performed with Python 3 using the packages matplotlib, numpy, scipy, sklearn, pandas, nixio (Stoewer et al., 2014), and thunderfish (`https://github.com/bendalab/thunderfish`).

In binary spike trains with a time step of 0.025 ms each spike was indicated by a value of 40 kHz and all other time bins were set to zero. Instantaneous firing rates were computed by convolving the spike trains with a Gaussian kernel. The standard deviation of the kernels was set to $\sigma = 0.5$ ms or $\sigma = 2$ ms. In the frequency domain, are also Gaussians centered at zero frequency and with a standard deviation of $\sigma_f = (2\pi\sigma)^{-1} = 318$ Hz or $\sigma_f = 80$ Hz, respectively. Cells with less than 50 different beat frequencies and cells with no beat frequencies higher than 2.6 EODf were excluded from the analysis.

Power spectra of the binary spike trains or the instantaneous firing rates in response to beat stimuli were computed from fast Fourier transforms on $n_{fft} = 4096$ long data segments that overlapped by 50 %. The initial and last 5 ms of each spike train were excluded from the analysis.

Modulation depth was estimated as the square root of the integral of the power spectrum over the 4 frequencies closest to the beat frequency. Baseline firing rate was calculated as the number of spikes divided by the duration of the baseline recording (on average 18 s).

Leaky integrate-and-fire models We constructed leaky integrate-and-fire (LIF) models to reproduce the specific firing properties of P-units (Chacron et al., 2001; Sinz

et al., 2020):

$$\tau_m \frac{dV_m}{dt} = -V_m + f(V_m) + \mu + \beta V_d - A + \sqrt{2D}\xi \quad \text{(A.1)}$$

where τ_m is the membrane time constant, μ a bias current, and D is the strength of Gaussian white noise ξ. Whenever the unitless membrane voltage V_m crosses the threshold of $\theta = 1$, a spike is generated and the voltage is reset to $V_m = 0$.

The static non-linearity $f(V_m)$ equals zero for the LIF. In case of an exponential integrate-and-fire model (EIF, Fig. 5), this function was set to

$$f(V_m) = \Delta_V e^{\frac{V_m - 1}{\Delta_V}} \quad \text{(A.2)}$$

(Fourcaud-Trocmé et al., 2003), where we set Δ_V to 0.1. Varying Δ_V from 0.001 to 0.1 did not change the results.

The prominent spike-frequency adaptation of P-units (Benda et al., 2005) is modeled by an adaptation current A with dynamics

$$\tau_A \frac{dA}{dt} = -A \quad \text{(A.3)}$$

and adaptation time-constant τ_A. Whenever a spike is generated, the adaptation current is incremented by Δ_A (Benda et al., 2010).

The input to the LIF is the membrane voltage V_d of a dendritic compartment scaled by β, that low-pass filters the rectified, Eq. (2), electrosensory stimulus $x(t)$ with a time constant of τ_d:

$$\tau_d \frac{dV_d}{dt} = -V_d + \lfloor x \rfloor_0^p \quad \text{(A.4)}$$

This dendritic low-pass filtering was needed for reproducing the loose coupling of P-unit spikes to the EOD, while maintaining high sensitivity to small amplitude modulations. The rectified stimulus was optionally taken to a power of p.

The stimulus is the EOD of the receiving fish normalized to an amplitude of one plus the EOD of a second fish. If not stated otherwise, a superposition of cosine waves, Eq. (1), was used to mimic the EODs. Realistic EODs (Fig. 6 H) were generated by summing up the first 10 harmonics whose relative amplitudes and phases have been extracted from head-tail recordings obtained during measurements of P-unit baseline activity using our thunderfish software, `https://github.com/bendalab/thunderfish`.

The 8 free parameters of the P-unit model, τ_m, μ, β, D, τ_A, Δ_A, τ_d, and t_{ref}, were fitted to both baseline activity (baseline firing rate, CV of inter-spike intervals (ISI), serial correlation of ISIs at lag one, and vector strength of spike coupling to EOD) and responses to step in- and decreases in EOD amplitude (onset- and steady-state responses, effective adaptation time constant) of 9 specific P-units (table A.1). When modifying the model (e.g. varying the powers p), we adapted the bias current to restore the original baseline firing rate.

References

M. S. Lewicki, Efficient coding of natural sounds., Nat Neurosci 5 (2002) 356–363.

G. L. Romani, S. J. Williamson, L. Kaufman, Tonotopic organization of the human auditory cortex, Science 216 (1982) 1339–1340.

C. Köppl, Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, tyto alba, Journal of Neuroscience 17 (1997) 3312–3321.

C. D. Hopkins, Stimulus filtering and electroreception: tuberous electroreceptors in three species of gymnotoid fish., J Comp Physiol 111 (1976) 171–207.

W. S. Rhode, S. Greenberg, Encoding of amplitude modulation in the cochlear nucleus of the cat, Journal of neurophysiology 71 (1994) 1797–1825.

M. E. Nelson, Z. Xu, J. R. Payne, Characterization and modeling of p-type electrosensory afferent responses to amplitude modulations in a wave-type electric fish, Journal of Comparative Physiology A 181 (1997) 532–544. URL: http://dx.doi.org/10.1007/s003590050137. doi:10.1007/s003590050137.

H. Walz, J. Grewe, J. Benda, Static frequency tuning accounts for changes in neural synchrony evoked by transient communication signals., J Neurophysiol 112 (2014) 752–765.

A. Roeber, Untersuchungen des hrn. scheibler in crefeld über die sogenannten schläge, schwebungen oder stösse, Annalen der Physik 108 (1834) 492–520.

R. König, Li. on the simultaneous sounding of two notes, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 1 (1876) 417–446.

G. S. Ohm, Bemerkungen über combinationstöne und stösse, Annalen der Physik 123 (1839) 463–466.

H. L. Helmholtz, On the Sensations of Tone as a Physiological Basis for the Theory of Music, Cambridge University Press, 2009.

R. Plomp, Beats of mistuned consonances, The Journal of the Acoustical Society of America 42 (1967) 462–474.

P. Joris, C. Schreiner, A. Rees, Neural processing of amplitude-modulated sounds, Physiological reviews 84 (2004) 541–577.

J. Henninger, R. Krahe, F. Kirschbaum, J. Grewe, J. Benda, Statistics of natural communication signals observed in the wild identify important yet neglected stimulus regimes in weakly electric fish., J. Neurosci. 38 (2018) 5456–5465.

J. Henninger, R. Krahe, F. Sinz, J. Benda, Tracking activity patterns of a multispecies community of gymnotiform weakly electric fish in their neotropical habitat without tagging, J Exp Biol 223 (2020) jeb206342. URL: https://jeb.biologists.org/content/223/3/jeb206342. doi:10.1242/jeb.206342.

E. Knudsen, Spatial aspects of the electric fields generated by weakly electric fish, J Comp Physiol A 99 (1975) 103–118.

J. Benda, The physics of electrosensory worlds., volume 7, Elsevier, Academic Press, 2020, pp. 228–254.

C. E. Carr, L. Maler, E. Sas, Peripheral organization and central projections of the electrosensory nerves in gymnotiform fish., J Comp Neurol 211 (1982) 139–153.

H. Scheich, T. H. Bullock, R. Hamstra Jr, Coding properties of two classes of afferent nerve fibers: high-frequency electroreceptors in the electric fish, eigenmannia., Journal of Neurophysiology 36 (1973) 39–60.

J. Bastian, Electrolocation I. How electroreceptors of *Apteronotus albifrons* code for moving objects and other electrical stimuli., J Comp Physiol A 144 (1981) 465–479.

J. Benda, A. Longtin, L. Maler, Spike-frequency adaptation separates transient communication signals from background oscillations., J Neurosci 25 (2005) 2312–2321.

J. Bastian, Electrolocation, Journal of comparative physiology 144 (1981) 465–479.

J. Benda, A. Longtin, L. Maler, A synchronization-desynchronization code for natural communication signals., Neuron 52 (2006) 347–358.

M. J. Chacron, A. Longtin, L. Maler, Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli, Journal of Neuroscience 21 (2001) 5328–5343.

Table A.1: Model parameters, fitted to nine specific P-units.

cell	β	τ_m/s	μ	D	τ_A/s	Δ_A	τ_d/s	t_{ref}/s
2012-04-20-ak	373.7	0.0017	21.88	0.000130	0.2950	142.7	0.0043	0.0009
2011-10-25-ad	301.4	0.0029	−34.38	0.000440	0.1161	60.7	0.0041	0.0007
2012-12-20-ab	46.6	0.0013	−4.79	0.000029	0.0480	9.3	0.0012	0.0011
2012-12-20-ad	124.2	0.0011	−16.21	0.000056	0.0932	23.0	0.0045	0.0011
2012-12-20-ae	190.1	0.0016	−31.84	0.000071	0.0612	28.8	0.0054	0.0009
2012-12-21-ai	291.2	0.0021	−54.69	0.000552	0.1274	36.8	0.0031	0.0012
2014-06-06-ac	382.9	0.0050	−70.70	0.004005	0.1120	55.1	0.0024	0.0006
2018-05-08-af	266.9	0.0012	−35.16	0.000168	0.0620	50.7	0.0089	0.0011
2018-06-25-ad	286.1	0.0024	−29.69	0.002437	0.0963	62.5	0.0019	0.0011
median	286.1	0.0017	−31.84	0.000168	0.0963	50.7	0.0041	0.0011

M. Savard, R. Krahe, M. Chacron, Neural heterogeneities influence envelope and temporal coding at the sensory periphery, Neuroscience 172 (2011) 270–284.

F. H. Sinz, C. Sachgau, J. Henninger, J. Benda, J. Grewe, Simultaneous spike-time locking to multiple frequencies, J Neurophysiol accepted (2020).

A. Watanabe, K. Takeda, The change of discharge frequency by A.C. stimulus in a weak electric fish., J Exp Biol 40 (1963) 57–66.

A. J. Kalmijn, The detection of electric fields from inanimate and animate sources other than electric organs., in: A. Fessard (Ed.), Electroreceptors and other specialized receptors in lower vertrebrates, Springer, Heidelberg, 1974, pp. 148–194.

J. Engelmann, S. Gertz, J. Goulet, A. Schuh, G. von der Emde, Coding of stimuli by ampullary afferents in gnathonemus petersii., J Neurophysiol 104 (2010) 1955–1968. URL: http://dx.doi.org/10.1152/jn.00503.2009. doi:10.1152/jn.00503.2009.

J. Grewe, A. Kruscha, B. Lindner, J. Benda, Synchronous spikes are necessary but not sufficient for a synchrony code, PNAS 114 (2017) E1977.

N. J. Berman, L. Maler, Inhibition evoked from primary afferents in the electrosensory lateral line lobe of the weakly electric fish (apteronotus leptorhynchus), Journal of Neurophysiology 80 (1998) 3173–3196.

J. W. Middleton, A. Longtin, J. Benda, L. Maler, The cellular basis for parallel neural transmission of a high-frequency stimulus and its low-frequency envelope., PNAS 103 (2006) 14596–14601.

N. Fourcaud-Trocmé, D. Hansel, C. van Vreeswijk, N. Brunel, How spike generation mechanisms determine the neuronal response to fluctuating inputs, Journal of Neuroscience 23 (2003) 11628–11640. URL: https://www.jneurosci.org/content/23/37/11628. doi:10.1523/JNEUROSCI.23-37-11628.2003. arXiv:https://www.jneurosci.org/content/23/37/11628.full.pdf.

B. W. Knight, Dynamics of encoding in a population of neurons., J Gen Physiol 59 (1972) 734–766.

A. Longtin, J. W. Middleton, J. Cieniak, L. Maler, Neural dynamics of envelope coding, Mathematical Biosciences 214 (2008) 87–99.

J. Carriot, M. Jamali, K. E. Cullen, M. J. Chacron, Envelope statistics of self-motion signals experienced by human subjects during everyday activities: Implications for vestibular processing, PLoS One 12 (2017) e0178664.

S. A. Stamper, M. S. Madhav, N. J. Cowan, E. S. Fortune, Beyond the Jamming Avoidance Response: weakly electric fish respond to the envelope of social electrosensory signals., J Exp Biol 215 (2012) 4196–4207.

L. Myers, M. Lowery, M. O'malley, C. Vaughan, C. Heneghan, A. S. C. Gibson, Y. Harley, R. Sreenivasan, Rectification and non-linear pre-processing of emg signals for cortico-muscular analysis, Journal of neuroscience methods 124 (2003) 157–165.

S. Khanna, M. Teich, Spectral characteristics of the responses of primary auditory-nerve fibers to amplitude-modulated signals, Hearing research 39 (1989) 143–157.

D.-A. Clevert, T. Unterthiner, S. Hochreiter, Fast and accurate deep network learning by exponential linear units (elus), arXiv preprint arXiv:1511.07289 (2015).

X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in: Proceedings of the fourteenth international conference on artificial intelligence and statistics, 2011, pp. 315–323.

A. J. Peterson, P. Heil, Phase locking of auditory-nerve fibers reveals stereotyped distortions and an exponential transfer function with a level-dependent slope, Journal of Neuroscience 39 (2019) 4077–4099.

R. G. Northcutt, Electroreception in non-teleost bony fishes, in: H. W. Bullock T.H. (Ed.), Electroreception, Wiley, 1986, pp. 247 – 285.

I. Roux, S. Safieddine, R. Nouvian, M. Grati, M.-C. Simmler, A. Bahloul, I. Perfettini, M. Le Gall, P. Rostaing, G. Hamard, et al., Otoferlin, defective in a human deafness form, is essential for exocytosis at the auditory ribbon synapse, Cell 127 (2006) 277–289.

N. Michalski, J. D. Goutman, S. M. Auclair, J. B. de Monvel, M. Tertrais, A. Emptoz, A. Parrin, S. Nouaille, M. Guillon, M. Sachse, et al., Otoferlin acts as a ca2+ sensor for vesicle fusion and vesicle pool replenishment at auditory hair cell ribbon synapses, Elife 6 (2017) e31013.

T. A. Viancour, Electroreceptors of a weakly electric fish, Journal of comparative physiology 133 (1979) 327–338.

E. Knudsen, Behavioral thresholds to electric signals in high frequency electric fish, J Comp Physiol A 91 (1974) 333–353.

A. B. Steinbach, Diurnal movements and discharge characteristics of electric gymnotid fishes in the Rio Negro, Brazil, Biol Bull 138 (1970) 200–210.

C. D. Hopkins, Electric communication in the reproductive behavior of *Sternopygus macrurus* (gymnotoidei)., Z Naturf 35 (1974) 518–535.

B. Kramer, F. Kirschbaum, H. Markl, Species specificity of electric organ discharges in a sympatric group of gymnotoid fish from Manaus (Amazonas)., Adv Physiol Sci 31 (1981) 195–219.

S. A. Stamper, E. Carrera-G, E. W. Tan, V. Fugère, R. Krahe, E. S. Fortune, Species differences in group size and electrosensory interference in weakly electric fishes: Implications for electrosensory processing, Behav Brain Res 207 (2010) 368 – 376. URL: http://www.sciencedirect.com/science/article/pii/S0166432809006317. doi:https://doi.org/10.1016/j.bbr.2009.10.023.

W. E. Brownell, Outer hair cell electromotility and otoacoustic emissions, Ear and hearing 11 (1990) 82.

S. Kujawa, M. Fallon, R. Bobbin, Time-varying alterations in the f2-f1 dpoae response to continuous primary stimulation i: Response characterization and contribution of the olivocochlear efferents, Hearing research 85 (1995) 142–154.

C. J. Sumner, A. R. Palmer, Auditory nerve fibre responses in the ferret, European Journal of Neuroscience 36 (2012) 2428–2439.

E. Evans, The frequency response and other properties of single fibres in the guinea-pig cochlear nerve, The Journal of physiology

226 (1972) 263–287.

A. R. Palmer, Encoding of rapid amplitude fluctuations by cochlear-nerve fibres in the guinea-pig, Archives of oto-rhino-laryngology 236 (1982) 197–202.

D. Gussin, J. Benda, L. Maler, Limits of linear rate coding of dynamic stimuli by electroreceptor afferents, Journal of neurophysiology 97 (2007) 2917–2929.

R. Ratnam, M. E. Nelson, Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals, Journal of Neuroscience 20 (2000) 6672–6683.

A. Stoewer, C. J. Kellner, J. Benda, T. Wachtler, J. Grewe, File format and library for neuroscience data and metadata, Front. Neuroinform 8 (2014) 10–3389.

J. Benda, L. Maler, A. Longtin, Linear versus nonlinear signal transmission in neuron models with adaptation-currents or dynamic thresholds., J Neurophysiol 104 (2010) 2806–2820.

L. Maler, The posterior lateral line lobe of certain gymnotoid fish: Quantitative light microscopy., J Comp Neurol 183 (1979) 323–364.

Appendix B. Supplement

Appendix B.1. Tuning curve depends on time-scale of read-out

The strongest peak in the power spectrum of the P-unit response is at EODf (grey circle in Fig. S1 A). The relevant peak explaining the P-unit response, i.e. the modulation of its firing rate, appears below EOD$f/2$ (orange circle). The frequency of this peak shows the aliasing phenomenon and the corresponding amplitude is only slightly modulated as a function of stimulus frequency, if the power spectrum is computed from the binary spike train directly (time step of 0.025 ms, Fig. S1 A).

P-units project onto pyramidal cells in the ELL (Maler, 1979). The postsynaptic potentials low-pass filter the incoming P-unit spike trains (Berman and Maler, 1998). This filtering affects the shape of the P-unit's tuning curve (Fig. S1 B,C). In particular the EODf component is strongly attenuated, as well as other peaks at higher frequencies, like for example the one of the stimulus (blue circle). For a physiological plausible kernel width of about 1 ms ($\sigma = 0.5$ ms for the Gaussian kernel used here, Berman and Maler, 1998), the corresponding filter only slightly attenuates the peak at the beat frequency below EOD$f/2$ while strongly suppressing power at higher frequencies. This filter retains the aliased frequencies but in addition results in a more strongly modulated tuning curve. Response amplitudes at high beat frequencies are more strongly attenuated, resulting in reduced response amplitudes around odd multiples of EOD$f/2$ (arrows in Fig. S1 B). Using a wider filter attenuates even lower beat frequencies such that only stimulus frequencies really close to integer EODf multiples are transmitted (Fig. S1 C). The tuning curve is strongly modulated with large values at low beat frequencies only and wide regions without responses in between.

Appendix B.2. Sampling theorem

The zig-zag pattern of the frequency of the P-unit response as a function of stimulus frequency (Fig. 1 J and Fig. 2) resembles the phenomenon of aliasing known from the sampling theorem. Here, a signal is sampled at specific time points separated by the sampling interval (Fig. S2). In case of low stimulus frequencies a single cycle is sampled at many time points and the stimulus frequency appears as a peak in the power spectrum as expected (Fig. S2 A,B). Once the stimulus frequency overtakes half of the sampling rate, however, less than two samples cover a single cycle of the stimulus and the peak in the power spectrum no longer is at the stimulus frequency (Fig. S2 C–F). Instead, the peak first moves back towards zero as the stimulus frequency approaches the sampling rate (Fig. S2 C,D). For even higher stimulus frequencies the peak in the power spectrum first moves up again towards half the sampling rate (Fig. S2 E,F) and keeps oscillating below this frequency. Half the sampling frequency is also known as the Nyquist frequency. Stimulus frequencies above the Nyquist frequency can not be uniquely retrieved from the spectrum. They are mirrored into the frequency range below the Nyquist frequency, causing aliasing (Fig. S2 C–F).

Appendix B.3. Analytic signal does not describe the aliasing structure of the beat

The analytic signal corresponding to the original signal is constructed by means of the Hilbert transform. With this method any signal can be expressed as a product

$$x(t) = A(t)\cos(\varphi(t)) \tag{S1}$$

where the amplitude modulation $A(t)$ is the absolute value of the analytic signal and $\varphi(t)$ is the phase of the analytic signal. The amplitude of the carrier $\cos(\varphi(t))$ is modulated by $A(t)$. Whereas the Hilbert transform itself is linear, taking the absolute value is a non-linear operation.

For the beat (1) we get for the amplitude modulation

$$A(t) = |x(t)| = \sqrt{1+\alpha^2+2\alpha\cos((\omega_2-\omega_1)t)} \tag{S2}$$

and for the phase

$$\varphi(t) = \frac{\omega_1+\omega_2}{2}\,t + \arctan\left(\frac{1-\alpha}{1+\alpha}\cdot\tan\left(\frac{\omega_1-\omega_2}{2}\,t\right)\right) \tag{S3}$$

(Stamper et al., 2012). This is an exact identity. The Hilbert transform is just a mathematical trick to transform any signal into such a product of an amplitude modulation and a cosine carrier.

For $\alpha = 1$ (both cosine waves have the same amplitude) this reduces to the well known identity

$$x(t) = 2\cos\left(\frac{\omega_2-\omega_1}{2}t\right)\cos\left(\frac{\omega_1+\omega_2}{2}t\right) \tag{S4}$$

A carrier signal of frequency $(\omega_1+\omega_2)/2$ is multiplied with an amplitude modulation with frequency $(\omega_1-\omega_2)/2$. The latter frequency is half the frequency of the beating amplitude modulation.

For small amplitudes $\alpha \to 0$ the expansion of the amplitude modulation to first order results in

$$A(t) \approx 1+\alpha\cos(\Delta\omega t) \tag{S5}$$

This amplitude modulation has a constant zero-frequency component in the Fourier spectrum, and one at the difference frequency $\Delta\omega = \omega_2-\omega_1$. For larger amplitudes more and more harmonics of this peak appear.

This is exactly what we expect for low difference frequencies, i.e. for stimulus frequencies ω_2 close to ω_1. However, for higher difference frequencies, the amplitude of the analytic signal Eq. (S5) suggests that the beat frequency keeps increasing with increasing difference frequency, no matter how large the difference frequency (Fig. S3 A). It does not explain the aliasing structure we observe in the signals and the P-unit responses. This does not imply that the analytic signal is wrong. Rather the amplitude term

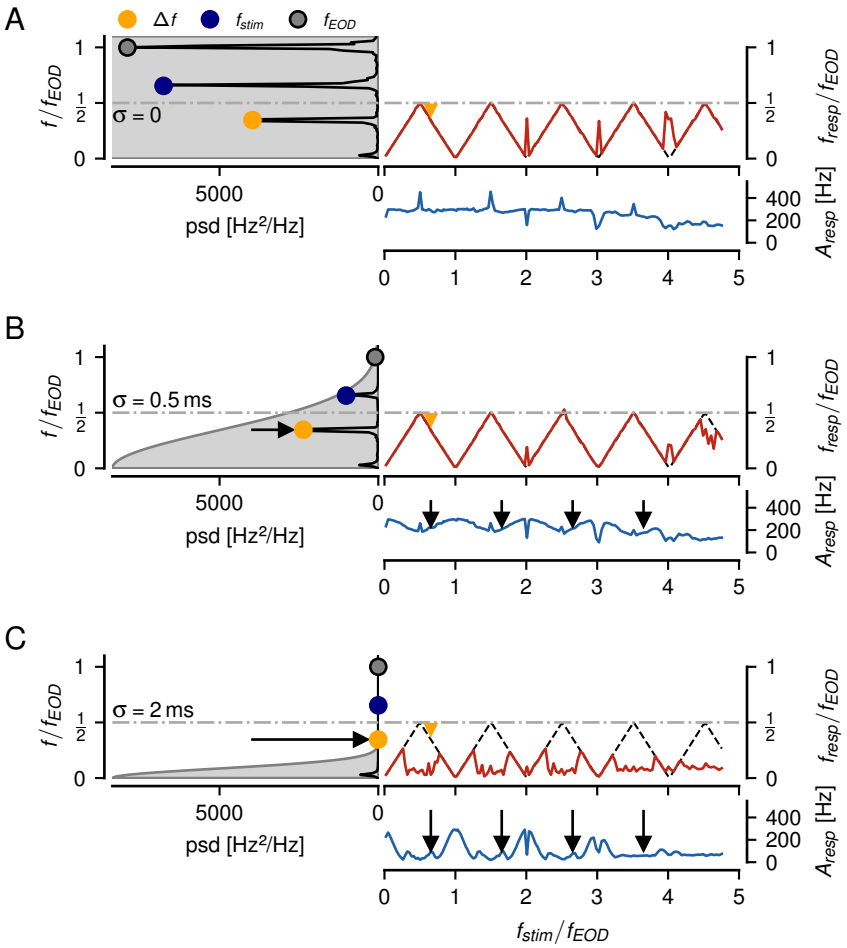

Figure S1: Influence of post-synaptic low-pass filtering on P-unit responses to beats. Left: Power spectra of a P-unit response to a $\Delta f = -220\,\text{Hz}$ beat. Convolution kernel mimicking post-synaptic filtering is indicated by gray area. Nyquist frequency at $f_{EOD}/2$ is indicated by the horizontal dash-dotted line. Right: Frequency (top) together with the expected aliasing frequencies (black dashed line) and amplitude (bottom) of the strongest peak of a P-unit's response below $f_{EOD}/2$. **A** Spectrum and tuning curves of the raw, binary spike trains recorded with a resolution of 40 kHz. Frequency tuning follows the aliasing frequencies over the whole measured range up to almost $5f_{EOD}$. The amplitude tuning curve is mostly flat with pronounced peaks at odd multiples of $f_{EOD}/2$. **B** A biological plausible post-synaptic filter, modeled by convolving the spike trains with a Gaussian kernel ($\sigma = 0.5\,\text{ms}$), keeps the frequency tuning, but reduces the amplitude of the P-unit's response for stimulus frequencies close to odd multiples of $f_{EOD}/2$ (arrows). **C** A wider PSP, modeled by a Gaussian with $\sigma = 2\,\text{ms}$, degrades the frequency tuning curves and strongly modulates amplitude tuning.

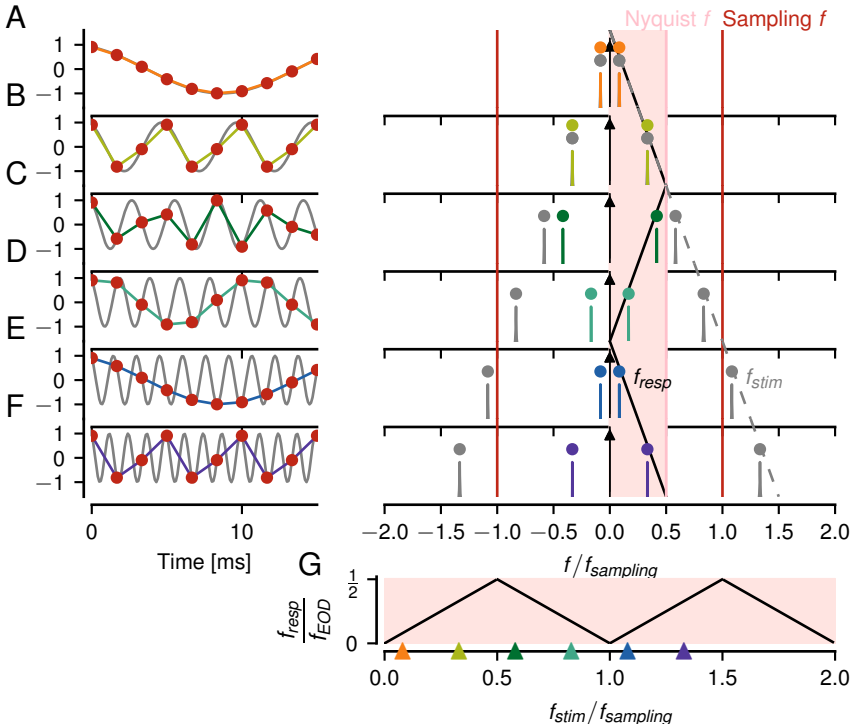

Figure S2: Sampling theorem visualization: aliasing induced at frequencies higher as the Nyquist f (sampling $f/2$). Left: Time series. Initial signal - grey. Sampling points - red. Sampled signal - colored lines.The frequency of the initial signal (grey) increases form **A** to **F**. Right: Power spectrum of the initial signal. Power spectrum of the sampled signal - colored. **A, B** Initial signal frequency below the Nyquist f - the frequency of the initial signal and the sampled signal are equal. **C, D** Frequency of the initial signal above the Nyquist f and below the sampling f - the sampled frequency moves backwards towards zero. **E, F** Frequency of the initial signal above the sampling f - the sampled frequency again moves form zero towards the Nyquist f. **G** A repetitive curve emerges for higher becoming frequencies of the initial signal. This curve resembles the one measured here in the data Fig. 1.)

(S5) simply does not capture the obvious aliasing structure of the beats. It is hidden in the phase term Eq. (S3).

On a first glance, the phase of the carrier simplifies to $\varphi(t) = \omega_1 t$ for small amplitudes. However, this is valid only for $\alpha = 0$, because only then $\frac{1-\alpha}{1+\alpha} = 1$ in Eq. (S3). The resulting small-amplitude approximation

$$x(t) \approx (1 + \alpha \cos(\Delta\omega t)) \cos(\omega_1 t) \tag{S6}$$

is in fact not a good approximation. In the Fourier spectrum it has two peaks at $\omega_1 \pm \Delta\omega$ instead of only one at $\omega_1 + \Delta\omega = \omega_2$ flanking the carrier at ω_1. Eq. (S6) no longer is a beat resulting from the superposition of two cosine waves, but is what is known in the auditory and electric fish literature as the sinusoidal amplitude modulation (SAM). The approximation fails, because the $\frac{1-\alpha}{1+\alpha}$-term in Eq. (S3) quickly deviates from one with slope -2 as amplitude increases.

Appendix B.4. Squaring does not retrieve the aliasing structure of the beat

An alternative method to retrieve amplitude modulations is to square the signal and then low-pass filter it. Squaring the beat (1), using the binomial theorem and the trigonometric power reduction formula results in

$$\begin{aligned} x^2(t) =& \tfrac{1}{2}(1+\alpha^2) + \tfrac{1}{2}\cos(2\omega_1 t) + \tfrac{1}{2}\alpha^2 \cos(2\omega_2 t) \\ &+ \alpha\cos((\omega_2 - \omega_1)t) + \alpha\cos((\omega_1 + \omega_2)t) \end{aligned} \tag{S7}$$

While the original signal (1) has two peaks in the power spectrum at ω_1 and ω_2 and no peak at the beat frequency, the power spectrum of the squared signal (S7) has five peaks, one for each term (Fig. S3 A). Shifting and generating new peaks in the spectrum is a hallmark of non-linear operations. The squaring operation doubles the two original frequencies and creates a new high-frequency peak at the sum of the two frequencies. In addition, a new peak occurs at zero, representing the non-zero mean of the squared signal. Another peak appears at the difference frequency $\omega_2 - \omega_1$. This is the amplitude modulation. By subsequent low-pass filtering this peak can be isolated and that way the amplitude modulation can be retrieved. However, as for the analytic signal, none of the five terms explain the aliasing structure of the beat.

Appendix B.5. Thresholding the beat

Let's define a pulse train with the same frequency ω_1 as the cosine signal of the carrier:

$$p(\omega_1 t) = \begin{cases} 1 & ; \quad \cos(\omega_1 t) \geq 0 \\ 0 & ; \quad \cos(\omega_1 t) < 0 \end{cases} \tag{S8}$$

It assumes one where the cosine is positive and zero where the cosine is negative. Multiplying a cosine with this pulse train thus sets the negative half waves to zero.

The Fourier spectrum of the pulse train turns out to have peaks at odd multiples of ω_1 with amplitudes

$$\begin{aligned} c_k =& \frac{\omega_1}{2\pi} \int_{-\frac{\pi}{2\omega_1}}^{+\frac{\pi}{2\omega_1}} e^{-i\omega_1 k t}\, dt = \frac{1}{\pi k} \sin\left(\tfrac{\pi}{2} k\right) \\ =& \frac{1}{\pi k} (-1)^{\frac{k-1}{2}} \quad , \quad k \text{ odd} \end{aligned} \tag{S9}$$

and an additional peak at zero frequency with amplitude

$$c_0 = \frac{\omega_1}{2\pi} [t]_{-\frac{\pi}{2\omega_1}}^{+\frac{\pi}{2\omega_1}} = \frac{1}{2} \tag{S10}$$

Thresholding a cosine with the same frequency ω_1 equals multiplication of the cosine with the pulse train Eq. (S8):

$$\lfloor \cos(\omega_1 t) \rfloor_0 = \cos(\omega_1 t) \cdot p(\omega_1 t) \tag{S11}$$

The corresponding Fourier spectrum is the convolution of the spectrum of the cosine with peaks of amplitude $1/2$ at $\pm\omega_1$ with the spectrum of the pulse train. The two peaks of the cosine are shifted to the positions of all the peaks of the pulse train and multiplied with their amplitude. Always two neighbouring peaks of the pulse train at odd multiples of ω_1 contribute to a peak at even multiples of ω_1 with amplitude

$$\tilde{a}_k = \frac{1}{2} c_{k+1} + \frac{1}{2} c_{k-1} = \frac{1}{\pi} (-1)^{\frac{k}{2}} \frac{1}{1-k^2} \quad , \quad k \text{ even} \tag{S12}$$

The zero-frequency peak of the pulse train gives rise to peaks at $\pm\omega_1$ with amplitude

$$\tilde{a}_{\pm 1} = \frac{1}{2} c_0 = \frac{1}{4} \tag{S13}$$

Because

$$\begin{aligned} &\lfloor \cos(\omega_1 t) + \alpha \cos(\omega_2 t) \rfloor_0 \approx \\ &\qquad \cos(\omega_1 t) \cdot p(\omega_1 t) + \alpha \cos(\omega_2 t) \cdot p(\omega_1 t) \end{aligned} \tag{S14}$$

for small contrasts a, the spectrum of the thresholded beat can be approximated by the spectrum of the thresholded carrier, Eqs. (S12) and (S13), and the one of the thresholded stimulus. The latter is the convolution of the spectrum of the pulse train, Eqs. (S9) and (S10), with the spectrum of the cosine with peaks of amplitude $\alpha/2$ at frequencies $\pm\omega_2$. Each peak of the pulse train at odd multiples of ω_1 is replaced by a pair of peaks at frequencies $k\omega_1 \pm \omega_2$ with amplitudes

$$a_k = \frac{\alpha}{2} c_k = \frac{\alpha}{2} \frac{1}{\pi k} (-1)^{\frac{k-1}{2}} \quad , \quad k \text{ odd} \tag{S15}$$

The zero-frequency peak of the pulse train, Eq. (S10), gives rise to two peaks at $\pm\omega_2$ with amplitude

$$a_0 = \frac{\alpha}{2} c_0 = \frac{\alpha}{4} \tag{S16}$$

The relative amplitudes $\bar{a}_k = a_k/a_0$ up to $k = 5$ multiples of ω_1 of the beat frequencies introduced by thresholding are $\bar{a}_0 = 100\,\%$, $\bar{a}_1 = \frac{2}{\pi} \approx 64\,\%$, $\bar{a}_2 = 0$, $\bar{a}_3 \approx \frac{2}{3\pi} = 21\,\%$, $\bar{a}_4 = 0$, and $\bar{a}_5 \approx \frac{2}{5\pi} = 13\,\%$.

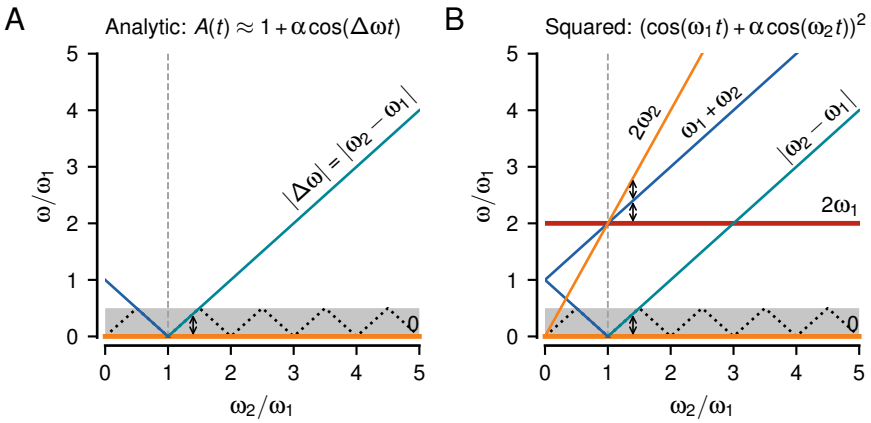

Figure S3: Amplitude modulations in analytic and squared signals of beats. Vertical arrows highlight the difference frequency $\Delta\omega = \omega_2 - \omega_1$. Frequencies below the Nyquist frequency at $\omega_1/2$ are marked by the gray band at the bottom. The black dotted line in this frequency band indicates the folded frequencies at $\omega_f = |\omega_2 - \omega_1\lfloor\omega_2/\omega_1\rceil|$. **A** The amplitude modulation Eq. (S5) of a beat computed from its analytic signal by means of a Hilbert transform has peaks in the spectrum only at 0 and at the absolute difference frequency $|\Delta\omega|$. **B** Squaring a beat retrieves, like the analytic signal, the difference frequency and an offset at zero frequency. In addition, three more peaks appear at $2\omega_1$, $2\omega_2$, and $\omega_1 + \omega_2$. Neither the analytic signal nor squaring explain the aliasing structure of the beat (dotted line).

Appendix B.6. Threshold cubed

Taking the beat to the power of three results in

$$x^3(t) = (\cos(\omega_1 t) + \alpha\cos(\omega_2 t))^3 \tag{S17}$$

$$\begin{aligned} &= \tfrac{3}{4}(1 + 2\alpha^2)\cos(\omega_1 t) + \underline{\tfrac{3}{4}(2\alpha + \alpha^3)\cos(\omega_2 t)} \\ &\quad + \tfrac{1}{4}\cos(3\omega_1 t) + \tfrac{1}{4}\alpha\cos(3\omega_2 t) \\ &\quad + \underline{\tfrac{3}{4}\alpha\cos((2\omega_1 + \omega_2)t)} + \underline{\tfrac{3}{4}\alpha\cos((2\omega_1 - \omega_2)t)} \\ &\quad + \tfrac{3}{4}\alpha^2\cos((\omega_1 + 2\omega_2)t) + \tfrac{3}{4}\alpha^2\cos((\omega_1 - 2\omega_2)t) \end{aligned} \tag{S18}$$

The dominant peaks depending on ω_2 are at ω_2 and $|2\omega_1 \pm \omega_2|$ (underlined).

Convolving the spectrum of the cubed beat, Eq. (S18), with the one of the pulse train, Eqs. (S9) and (S10), approximating the threshold operation, Eq. (2), boils down to replace all the peaks in the spectrum of the pulse train with the ones of the cubed beat shifted to the respective positions (Fig. 4 A–C). In the following calculations we ignore all terms of higher order in α.

The two purely ω_1-dependent terms with peaks at $\pm\omega_1$ and $\pm 3\omega_1$ result in peaks at even multiples of ω_1 with amplitudes

$$\begin{aligned} \tilde{b}_k &= \tfrac{1}{2}\tfrac{3}{4}(c_{k+1} + c_{k-1}) + \tfrac{1}{2}\tfrac{1}{4}(c_{k+3} + c_{k-3}) \\ &= \frac{3}{4\pi}(-1)^{\frac{k}{2}}\left(\frac{1}{1-k^2} - \frac{1}{9-k^2}\right), \quad k \text{ even} \end{aligned} \tag{S19}$$

and in addition in peaks directly at $\pm\omega_1$ and $\pm 3\omega_1$ with amplitudes

$$\tilde{b}_{\pm 1} = \tfrac{1}{2}\tfrac{3}{4}c_0 = \frac{3}{16} \tag{S20}$$

$$\tilde{b}_{\pm 3} = \tfrac{1}{2}\tfrac{1}{4}c_0 = \frac{1}{16} \tag{S21}$$

(horizontal lines in Fig. 4 E). The latter at the third harmonics of ω_1 is a new peak that the threshold without exponent does not generate.

Convolving the dominant ω_2 dependent terms in Eq. (S17) with the peaks at odd multiples of ω_1 of the pulse train, Eq. (S9), we get peaks at $k\omega_1 \pm \omega_2$ for odd k with amplitudes

$$\begin{aligned} b_k &= \tfrac{1}{2}\tfrac{3}{4}2\alpha c_k + \tfrac{1}{2}\tfrac{3}{4}\alpha(c_{k+2} + c_{k-2}) \\ &= \frac{3}{\pi}\alpha(-1)^{\frac{k+1}{2}}\frac{1}{k(k^2-4)}, \quad k \text{ odd} \end{aligned} \tag{S22}$$

These are peaks at the same frequencies as for the threshold without exponent, but with different amplitudes.

However, from the convolution with the zero-frequency term of the pulse train, Eq. (S10), we get additional peaks at $\pm\omega_2$ and $\pm(2\omega_1 \pm \omega_2)$ with amplitudes

$$b_0 = \tfrac{1}{2}\tfrac{3}{4}2\alpha c_0 = \frac{3}{8}\alpha \tag{S23}$$

$$b_{\pm 2} = \tfrac{1}{2}\tfrac{3}{4}\alpha c_0 = \frac{3}{16}\alpha \tag{S24}$$

The latter is the one the power of three adds to the folding frequencies around the second harmonics of ω_1.

The relative amplitudes $\bar{b}_k = b_k/b_0$ of the beat frequencies introduced by a cubed threshold are $\bar{b}_0 = 100\,\%$, $\bar{b}_1 = \frac{8}{3\pi} \approx 85\,\%$, $\bar{b}_2 = \frac{1}{2} = 50\,\%$, $\bar{b}_3 \approx \frac{8}{15\pi} = 17\,\%$, $\bar{b}_4 = 0$, and $\bar{b}_5 \approx \frac{8}{105\pi} = 2.4\,\%$.

Appendix B.7. Harmonics of the carrier are not sufficient to explain aliasing

In reality the carrier EOD is a complex periodic wave and thus already provides harmonics at multiples of the carrier frequency. Wouldn't that be enough to explain the aliasing structure of the beat without non-linearities?

At least a threshold is needed. Without any-nonlinearity the only frequency component depending on the stimulus frequency would be the stimulus itself. With a threshold, Eq. (2), approximated by multiplication with a pulse train,

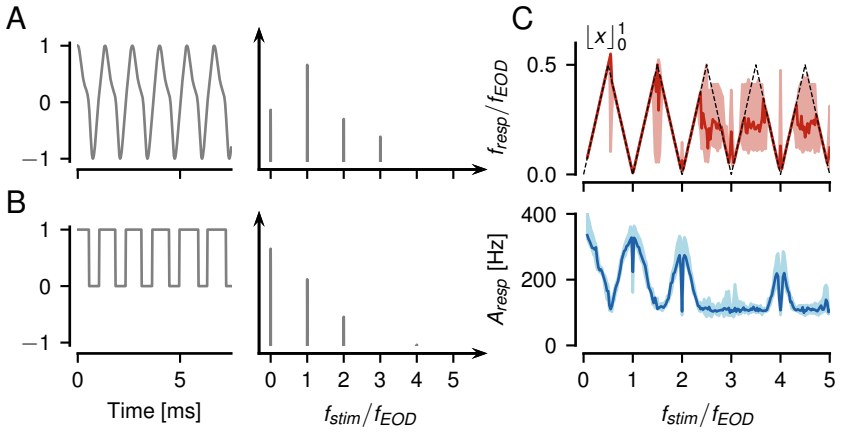

Figure S4: Influence of higher harmonics of the carrier on aliasing. **A** An EOD waveform and the corresponding power spectrum of an *A. leptorhynchus* used as a carrier signal of frequency ω_1 for simulating P-unit responses with the LIF models, Eqs. (A.1), (A.3), (A.4). **B** The corresponding pulse train, Eq. (S25), needed to approximate a threshold, Eq. (2), has a duty cycle larger than 50 %. **C** The enlarged duty cycle modifies the aliasing pattern of the resulting frequency (top) and amplitude (bottom) tuning curves, Eq. (S26), when using a threshold without exponent, Eq. (2). For the example EOD waveform shown, the third harmonics and not the second as for a pure sine wave is missing. Same models as in Fig. 6.

Eq. (S8), the higher harmonics of the carrier EOD would only add peaks to the resulting spectrum at multiples of the carrier frequency. The stimulus frequency still would be just convolved with the spectrum of the pulse train. As for the sine-wave carrier, the stimulus frequency would appear as beat frequencies around odd multiples of the carrier frequency and around zero frequency, Eqs. (S15) and (S16), but not at even multiples.

However, the harmonics of the carrier EOD modify the waveform. It is not a sine wave any more that stays positive for exactly half of the time and negative for the other half. Instead, the harmonics might distort the waveform such that we would need a pulse train with a duty cycle other than 50 % to emulate a threshold. For example, some *A. leptorhynchus* have a waveform that is wider than a sine wave at its zero crossings (Fig. S4 A). A matching pulse train would need a higher duty cycle (Fig. S4 B). We parameterize the pulse train by its duty cycle δ to account for this effect:

$$p(\omega_1 t;\delta) = \begin{cases} 1 & ; \quad -\delta\frac{\pi}{\omega_1} < t \bmod \frac{2\pi}{\omega_1} < \delta\frac{\pi}{\omega_1} \\ 0 & ; \quad \text{else} \end{cases} \tag{S25}$$

Changing the duty cycle modifies the spectrum of the pulse train:

$$c_k(\delta) = \frac{1}{\pi k}\sin(\pi k\delta)\ , \quad k \neq 0 \tag{S26}$$

$$c_0(\delta) = \delta\ , \quad k = 0 \tag{S27}$$

(Fig. S4 B). In particular, a peak at the second harmonics of the carrier appears with amplitude $c_2(\delta) = \frac{1}{2\pi}\sin(2\pi\delta)$. This peak is then convolved with the stimulus and fills in beat frequencies around the second harmonics (Fig. S4 C). The amplitude of the second harmonic of the pulse train equals zero for $\delta = \frac{1}{2}$, grows linearly in δ according to $c_2(\delta) \approx \frac{1}{2} - \delta$ as the duty cycle deviates from $\frac{1}{2}$, and assumes its maximum value of $\frac{1}{2\pi}$ for $\delta = \frac{1}{4}$ and $\delta = \frac{3}{4}$. It can get as large as the one of the fundamental, if, according to $c_2/c_1 = \cos(\pi\delta)$, the duty cycle approaches zero or one. However, the beat frequencies at the third harmonics are missing now, because the third harmonics of the pulse train is reduced by increasing the duty cycle.

To summarize, the harmonics of the carrier themselves do not contribute to extracting beat frequencies. However, the threshold operation applied to the non-sinusoidal carrier waveform modify the peaks relevant for the beat frequencies. Depending on the duty cycle of the carrier some harmonics are enhanced whereas others are suppressed. However, only a threshold cubed reliably reproduces the measured P-unit responses.

Appendix B.8. Harmonics of the stimulus are also not sufficient to explain aliasing

Alternatively, we could keep the carrier as a sine wave and use a realistic EOD for the stimulus. The components of the beat spectrum relevant for explaining beat frequencies result from the convolution of the spectrum of a pulse-train with a 50 % duty cycle, Eqs. (S9) and (S10), matching the sinusoidal carrier, with all the harmonics of the stimulus. For extracting the aliasing structure of the beat, however, only the fundamental of the stimulus is relevant. The higher harmonics introduce frequencies depending on multiples of the stimulus frequency and thus can not explain the beat frequencies that grow directly proportionally with stimulus frequency.

Appendix B.9. Tuning of P-units to EOD frequency

Silencing the fish's EOD and measuring the minimum amplitude of an artificial replacement EOD to make a P-unit fire action potentials results in V-shaped threshold curves centered at the fish's EOD frequency (Fig. S6 A, Hopkins, 1976). The corresponding band-pass filter is

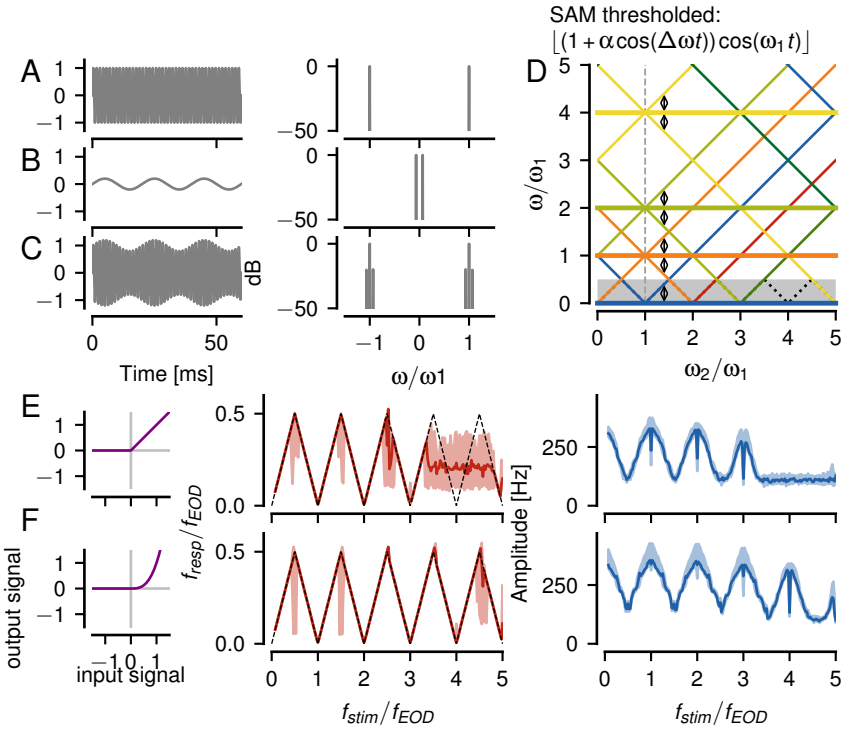

Figure S5: Model responses to SAM stimuli do not require an exponentiated threshold. **A-C** Left column: signals in the time domain, right column: the respective Fourier spectra. A: the receiver's EOD, B: the sinusoidal modulator with the desired difference frequency $\Delta\omega$, C: The receiver EOD is combined with the modulator according to Eq. (S6). Note: The modulator in B has two peaks at $\pm\Delta\omega$ which leads to additional peaks in the combined signal. **D** Positions of the major peaks in the combined signal as a function of the stimulus frequency. Grey band denotes the frequency range below the Nyquist frequency and the dashed line the shows the expected repetitive function. Colored lines the theoretical prediction of the relevant peaks. **E** Characteristics of the model response with a power-of-one threshold (left column). Center: Position of the response frequency normalized to the EOD frequency as a function of stimulus frequency. Dashed line is the repetitive expectation, solid red line the average and shaded area depicts the interquartile range. Right: response modulation as a function of stimulus frequency. Solid line is the average, shaded area the interquartile range. In contrast to Fig. 6 C, the additional peak introduced by the SAM stimulus fills in the responses at the 2nd harmonic of the receiver signal. **F** Same as E but applying a power-of-three threshold.

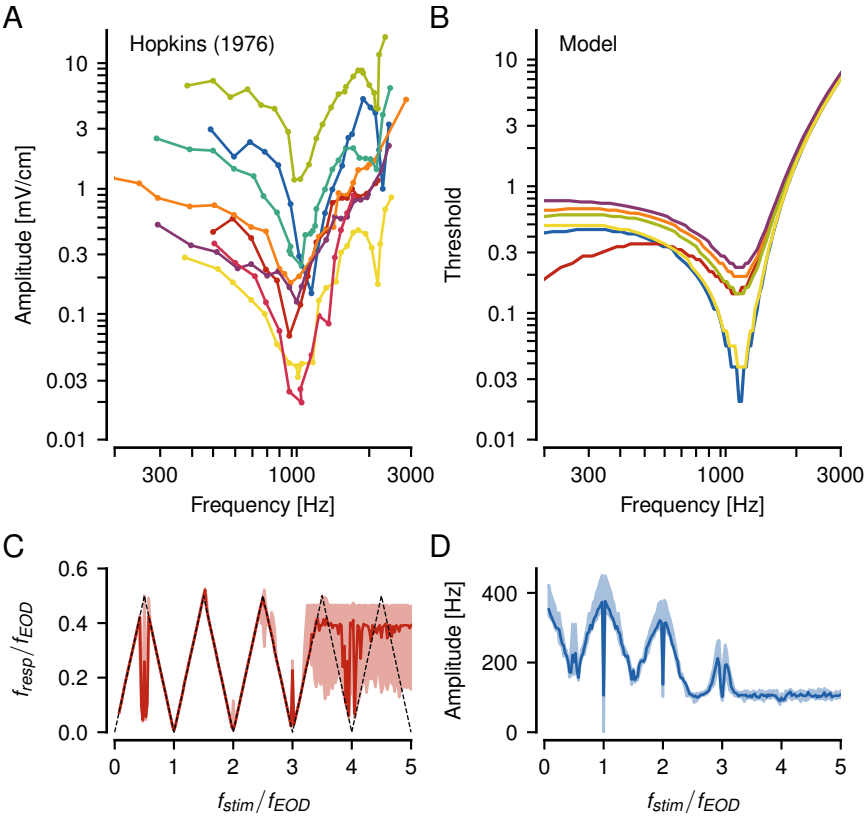

Figure S6: . Effects of the P-unit's EOD frequency filter. **A** Sensitivity of P-unit afferents to EOD frequency as reported by Hopkins (1976). The measured stimulus amplitudes were the minimum amplitude required to elicit just noticeable difference in firing rates of the P-units. **B** Corresponding sensitivities of our LIF models supplemented by a harmonic oscillator Eq. (S28). The amplitudes elicited an increase in firing rate of 10 % compared to baseline rate without stimulus. **C & D** The frequency f_{resp} and corresponding response amplitudes of these P-unit models to beats still reproduce our observed P-unit tuning to beats despite the EOD filter.

probably caused by electric resonance in the electroreceptor cells (Viancour, 1979). This could be modeled by a damped harmonic oscillator filtering the input signal before it is thresholded at the receptor synapse (Sinz et al., 2020).

To model this resonance filter we replaced the stimulus $x(t)$ in the P-unit models, Eqs. (A.1) – (A.4), by the output $y(t)$ of a harmonic oscillator

$$\frac{d^2y(t)}{dt^2} + 2\zeta w_0 \frac{dy(t)}{dt} + \zeta w_0^2 y(t) = x(t) \tag{S28}$$

multiplied with a normalization factor β. In Eq. (S28) the external force to the oscillator is the stimulus $x(t)$, ζ is the damping ratio of the harmonic oscillator, and

$$w_0 = \frac{w_R}{\sqrt{1-2\zeta^2}} \tag{S29}$$

is the eigen-frequency, where $w_R = 2\pi f_R$ is the resonance frequency that was set to the measured f_{EOD} of each fish. The normalization factor

$$\beta = 70 w_R \sqrt{(2w_0\zeta)^2 + (w_R^2 - w_0^2)^2/w_R^2} \tag{S30}$$

ensures that the fish's EOD is transmitted with a gain of one through the damped oscillator. The harmonic oscillator was solved using the differential equation solver from SciPy.

We varied ζ from 0.7 (almost no damping) to 0.1 (highest damping). A stronger damping factor of high frequencies can be compensated with a higher power modification of the input, with several combinations yielding similar results.

A mild damping coefficient of $\zeta = 0.45$ and an exponent $p = 5$ were sufficient to reproduce both the tuning of P-unit responses to f_{EOD} as reported by Hopkins (1976) (Fig. S6 B). This model still reproduces the responses to beats up to three multiples of the f_{EOD} (Fig. S6 C, D), suggesting that the EOD filter does not impede P-unit responses to high difference frequencies.