

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at ScienceDirect

## Hearing Research

journal homepage: [www.elsevier.com/locate/heares](http://www.elsevier.com/locate/heares)

Research paper

## Mode-locking neurodynamics predict human auditory brainstem responses to musical intervals



Karl D. Lerud, Felix V. Almonte, Ji Chul Kim, Edward W. Large\*

University of Connecticut, Department of Psychology, 406 Babbidge Road, Storrs, CT 06269-1020, USA

## ARTICLE INFO

## Article history:

Received 22 January 2013

Received in revised form

13 September 2013

Accepted 17 September 2013

Available online 1 October 2013

## ABSTRACT

The auditory nervous system is highly nonlinear. Some nonlinear responses arise through active processes in the cochlea, while others may arise in neural populations of the cochlear nucleus, inferior colliculus and higher auditory areas. In humans, auditory brainstem recordings reveal nonlinear population responses to combinations of pure tones, and to musical intervals composed of complex tones. Yet the biophysical origin of central auditory nonlinearities, their signal processing properties, and their relationship to auditory perception remain largely unknown. Both stimulus components and nonlinear resonances are well represented in auditory brainstem nuclei due to neural phase-locking. Recently mode-locking, a generalization of phase-locking that implies an intrinsically nonlinear processing of sound, has been observed in mammalian auditory brainstem nuclei. Here we show that a canonical model of mode-locked neural oscillation predicts the complex nonlinear population responses to musical intervals that have been observed in the human brainstem. The model makes predictions about auditory signal processing and perception that are different from traditional delay-based models, and may provide insight into the nature of auditory population responses. We anticipate that the application of dynamical systems analysis will provide the starting point for generic models of auditory population dynamics, and lead to a deeper understanding of nonlinear auditory signal processing possibly arising in excitatory-inhibitory networks of the central auditory nervous system. This approach has the potential to link neural dynamics with the perception of pitch, music, and speech, and lead to dynamical models of auditory system development.

*This article is part of a Special Issue entitled <Music: A window into the hearing brain>.*

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Peripheral auditory nonlinearities have received a great deal of attention in the theoretical literature (Eguíluz et al., 2000; Kern and Stoop, 2003; Jülicher et al., 2001; Szalai et al., 2013), however, far less is known about the nature and the role of central auditory nonlinearities. Nonlinear receptive fields have been found in inferior colliculus of the cat (Escabi and Schreiner, 2002) and gerbil (Langner, 2007), and in auditory cortex of the cat (Sutter and Schreiner, 1991) and rhesus monkey (Kikuchi et al., 2011), for example. Interestingly, nonlinear responses to intervals composed of pure tones have been found in auditory evoked potentials of guinea pigs (Chertoff and Hecox, 1990), and nonlinear frequency-following responses (FFRs) have been observed in human brainstem using electroencephalography (EEG) (Galbraith, 1994; Pandya

and Krishnan, 2004; Lee et al., 2009), and in human auditory cortex using steady-state methods in magnetoencephalography (MEG) (Purcell et al., 2007). Together these facts pose questions about the auditory system: What are the generating mechanisms of these nonlinearities? Do brainstem responses represent only cochlear mechanics, or does brainstem signal processing play a significant role? Perhaps we can think of the central auditory system as performing a series of nonlinear transformations of sound. Significant theoretical advances in the study of neural signal processing will be necessary to understand how the brain harnesses nonlinear dynamics to analyze complex, temporally structured acoustic signals (Koepsell et al., 2010).

In central auditory circuits, action potentials phase-lock to both the fine time structure and the temporal envelope modulations of auditory stimuli at many different levels, including cochlear nucleus, superior olive, inferior colliculus, thalamus, and A1 (Langner, 1992). Traditionally, phase-locked spiking in the central auditory system is thought to represent an essentially passive transmission of synchronized basilar membrane motion, as illustrated

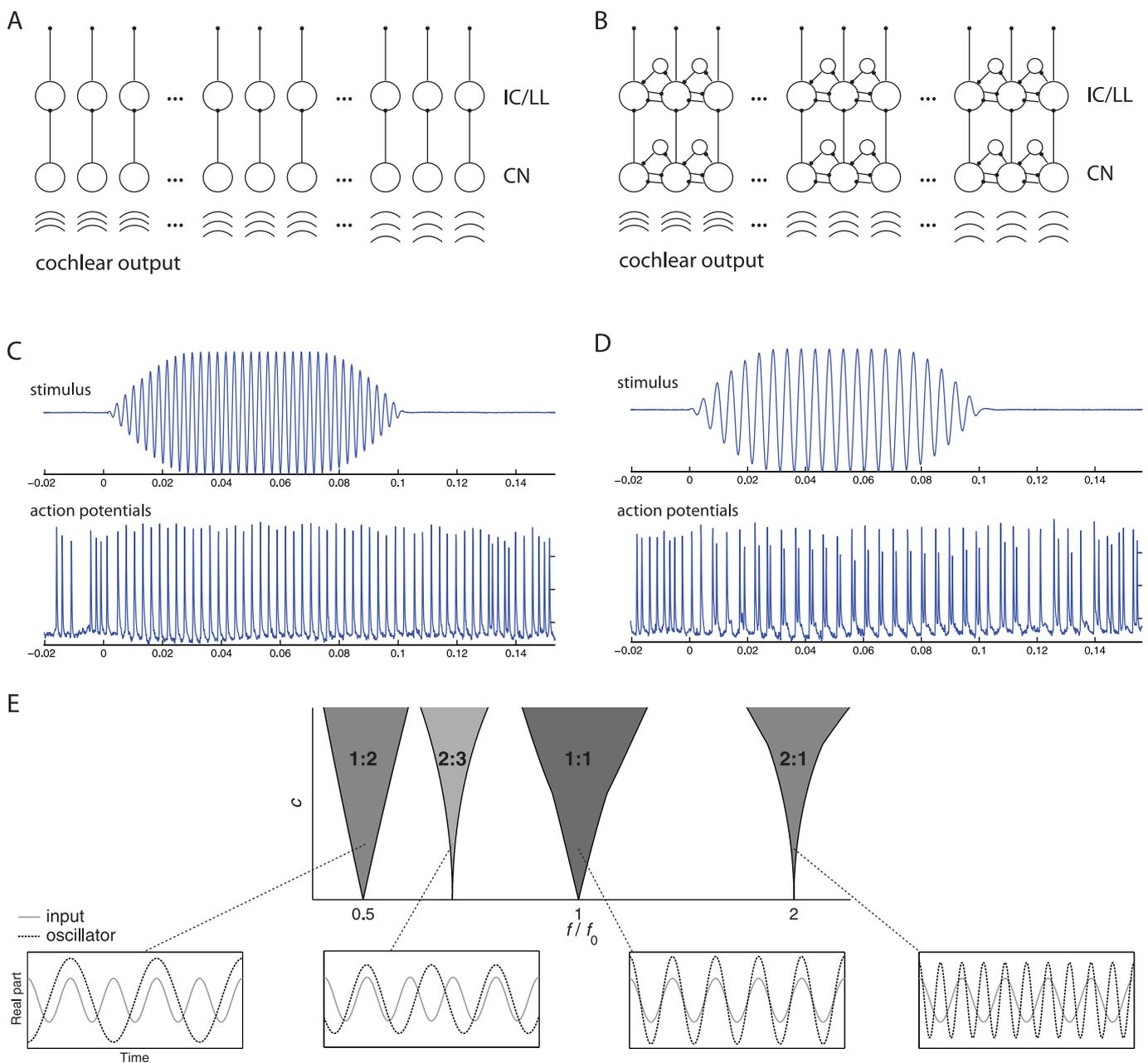
\* Corresponding author.

E-mail address: [edward.large@uconn.edu](mailto:edward.large@uconn.edu) (E.W. Large).

schematically in Fig. 1A (Koepsell et al., 2010). If this is the case, then nonlinear responses at the level of the brainstem would represent mainly distortion products that arise in the cochlea. An alternative possibility is that synchronized neural activity is carried forward by active circuits in the central auditory system, illustrated schematically in Fig. 1B. If this is the case, nonlinearities observed at the level of the brainstem might also arise due to mode-locking, a phenomenon that has been observed in the auditory brainstem (Large et al., 1998; Laudanski et al., 2010; Arnold and Burkard, 1998, 2000), and is physiologically distinct from the mechanical compression and half-wave rectification that occurs in the organ of

Corti. Mode-locking is a generalization of phase-locking in which a periodic stimulus interacts with an intrinsic oscillatory dynamics of a neuron or neural circuit, causing  $k$  cycles of an oscillation to lock to  $m$  cycles of the stimulus, where  $k$  and  $m$  are integers. In phase-locking  $k = m = 1$ .

Fig. 1C and D illustrate phase- and mode-locking to acoustic signals in the auditory midbrain of the fish *Pollimyrus* (Large et al., 1998). On the left, a neuron produces one action potential for every cycle of the sinusoidal stimulus (1:1 phase-locking), on the right, the same neuron produces two action potentials for every cycle of the sinusoidal stimulus (2:1 mode-locking). Mode-locking to



**Fig. 1.** Phase & mode locking. (A) Phase-locked spiking in the central auditory system may represent passive transmission of synchronized basilar membrane motion via a labeled line code. In such a passive model, spiking patterns are transferred directly from layer to layer of the auditory system. (B) Alternatively, synchronized neural activity may be carried forward by active oscillatory circuits in the central auditory system, which rely on interaction of excitatory and inhibitory subprocesses. (C) Phase-locking (1:1) and (D) mode-locking (2:1) of an auditory medulla neuron in the fish *Pollimyrus* to different stimulus frequencies (from Large et al., 1998). (E) Mode locking in a canonical model. Within each resonance region (shaded), the canonical model mode-locks to external input at the ratio shown in figure ( $c$ : coupling strength,  $f$ : oscillator's intrinsic frequency,  $f_0$ : input frequency). Insets show the inputs and traces produced by a canonical model.

acoustic signals has been observed in guinea pig cochlear nucleus chopper and onset neurons (Laudanski et al., 2010), and mode-locking to the difference tone of two dichotically presented stimulus frequencies has been observed in vivo and isolated to the inferior colliculus of the chinchilla (Arnold and Burkard, 1998, 2000). Mode-locked spiking patterns are often observed in vitro under DC injection (Brumberg and Gutkin, 2007), and active oscillations have been observed in vivo in the inferior colliculus of the chicken (Schwarz et al., 1993). Such observations lead to the possibility that the nonlinear responses observed in the human auditory brainstem may arise, in part, due to mode-locking neurodynamics.

Here, we consider nonlinear responses to musical intervals that have been measured in the human auditory brainstem response (Lee et al., 2009, see Fig. 2). In that study, the brainstem representation of the musical intervals comprised not only stimulus frequencies, but also numerous resonances at frequencies that were not physically present in the stimulus. How did these frequencies arise? The stimuli were the intervals major sixth (G and E, “consonant”) and minor seventh (F# and E, “dissonant”) which have fundamental frequency ratios of 1.6 (166 Hz/99 Hz) and 1.7 (166 Hz/93 Hz), making it unlikely that interaction of the fundamental frequencies created strong distortion products in the cochlea (Dhar et al., 2009, 2005; Knight and Kemp, 2001). Moreover, the responses of trained musicians were significantly enhanced compared with those of novice listeners, implying experience-based differences that would not have arisen at the level of the cochlea or auditory nerve (Lee et al., 2009). However, stimuli were recorded from an electric piano and thus also included several harmonics of the fundamentals which could have generated cochlear distortions through their interaction. To guard against the possibility of seeing cochlear responses in the FFR, Lee et al. (2009) used the polarity reversal technique, averaging brainstem responses to stimuli with alternating polarities (Skoe and Kraus, 2010).

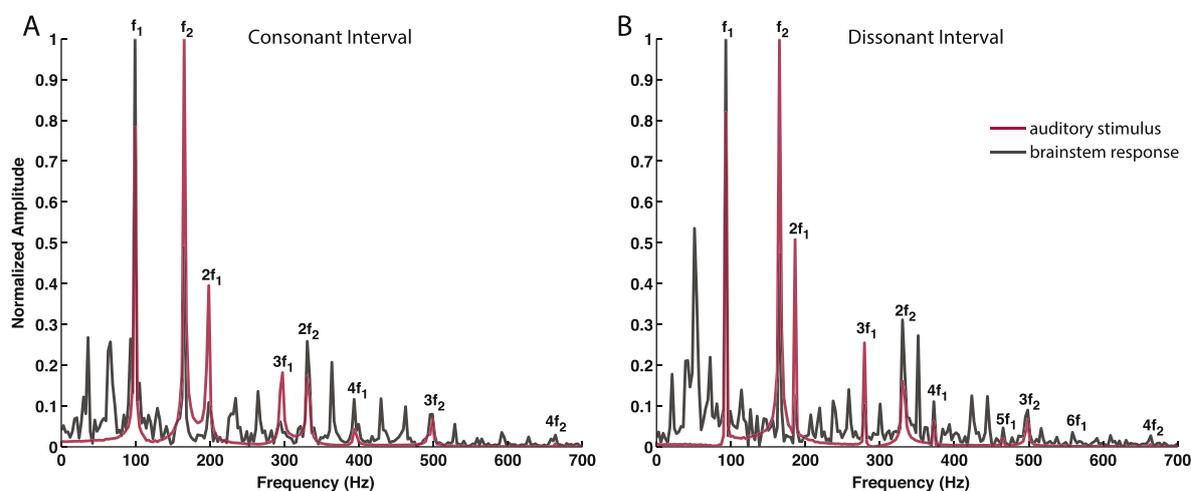
The polarity reversal technique is commonly taken to eliminate the cochlear microphonic from the scalp-recorded potential (Skoe and Kraus, 2010), and it is also said to eliminate the fine-structure-following response and amplify the envelope-following response in the scalp-recorded potential, irrespective of the particular place of generation (Wile and Balaban, 2007). While there is truth to both of these interpretations, it is most accurate to say that the technique of averaging both polarity conditions

minimizes odd-order nonlinearities, and maximizes even-order nonlinearities (Rickman et al., 1991, see also Appendix A). An understanding of this technique is critical for correctly interpreting brainstem EEG results and attempting to model them. Odd-order responses include those to the primaries themselves (first order) and, for instance, a cubic difference tone,  $2f_1 - f_2$ , arising from any pair of primaries (third order). Even-order responses include the simple difference tone,  $f_2 - f_1$ , and summation tone,  $f_2 + f_1$  (both second order). Due to the nature of cochlear compression (Ruggero et al., 1997; Robles and Ruggero, 2001) cochlear distortion products are mostly odd-order (Jülicher et al., 2001; Szalai et al., 2013), thus this technique minimizes most nonlinearities that would have arisen via the cochlea.

We explore the implications of mode-locking using a nonlinear cochlear model (e.g., Jülicher et al., 2001), coupled to a generic model of mode-locking, which we take to represent cochlear nucleus and inferior colliculus/lateral lemniscus (Fig. 1B). Mode-locked dynamics can be observed in generic neuron models (Lee and Kim, 2006), and in models of oscillatory neural populations (Hoppensteadt and Izhikevich, 1996a; Large et al., 2010, see Fig. 1E). Here, we employ a canonical model to make predictions about the responses of neural populations, as observed in brainstem recordings. The canonical model (Large et al., 2010) was derived from a model of neural dynamics in oscillatory populations (Hoppensteadt and Izhikevich, 1996a; Wilson and Cowan, 1973), using normal form theory (Wiggins, 1990), and it makes generic predictions that do not depend on the specific anatomical/physiological mechanisms at the neuronal level that produce mode-locked responses. This model predicts mode-locked population responses to periodic stimuli, as illustrated in Fig. 1D, and is in principle capable of explaining many kinds of nonlinearities that have been reported in the central auditory system (Chertoff and Hecox, 1990; Escabi and Schreiner, 2002; Joris et al., 2004; Langner, 1992, 2007; Large and Crawford, 2002; Lee et al., 2009; Pandya and Krishnan, 2004; Sutter and Schreiner, 1991; Kikuchi et al., 2011). Here we ask whether mode-locking could explain the specific nonlinear population responses to musical intervals observed in human auditory brainstem (Lee et al., 2009).

## 2. Theoretical framework

In general, oscillatory neurodynamics arise through the interaction of excitatory and inhibitory subprocesses with implications



**Fig. 2.** Auditory stimuli and brainstem response. Auditory stimuli (red) and brainstem response (gray) observed by Lee et al. (2009) for (A) a consonant interval and (B) a dissonant interval. The consonant interval was a major sixth, G2 ( $f_1 = 99$  Hz) and E3 ( $f_2 = 166$  Hz); the dissonant interval was a minor seventh, F#2 ( $f_1 = 93$  Hz) and E3 ( $f_2 = 166$  Hz). The brainstem response contains spectral information not present in the stimulus.

for many different aspects of neural information processing (Wilson and Cowan, 1973; Hoppensteadt and Izhikevich, 1996a; Koepsell et al., 2010; Varela et al., 2001; Ainsworth et al., 2012). The following generic model captures the dynamics of weakly coupled networks of neural oscillators with different natural frequencies responding to external stimulation (Hoppensteadt and Izhikevich, 1996a; Large et al., 2010):

$$\begin{aligned} \frac{dx_i}{dt} &= f_i(x_i, y_i, \lambda) + \varepsilon p_i(x_1, y_1, \dots, x_n, y_n, s(t), \lambda, \varepsilon) \\ \frac{dy_i}{dt} &= g_i(x_i, y_i, \lambda) + \varepsilon q_i(x_1, y_1, \dots, x_n, y_n, s(t), \lambda, \varepsilon) \end{aligned} \quad (1)$$

In this general formulation, the variables  $x_i$  and  $y_i$  represent excitatory and inhibitory activities in the  $i$ th neuronal population (or neural oscillator) respectively. The nonlinear functions  $f_i$  and  $g_i$  describe the intrinsic (uncoupled) dynamics of the excitatory and inhibitory subpopulations, and  $p_i$  and  $q_i$  describe interactions between neuronal subpopulations with external input  $s(t)$ .  $\lambda$  is a set of model parameters, and  $\varepsilon$  is a small number that represents weak interaction (Hoppensteadt and Izhikevich, 1996a).

To understand the implications of mode-locking for brainstem responses, Eq. (1) is analyzed using normal form theory assuming the system is near an oscillatory instability called a Hopf bifurcation (Hoppensteadt and Izhikevich, 1996a; Large et al., 2010). This analysis produces a canonical model, the simplest model in a mathematical sense (see Hoppensteadt and Izhikevich, 1997; Wiggins, 1990; Perko, 1996; Murdock, 2003) that captures the dynamics of the system. The analysis does not depend on the details of Eq. (1), only on the existence of a Hopf bifurcation. The canonical model makes important relationships explicit, including the relative strength of mode-locked responses. The canonical network corresponding to Eq. (1) can be written as (Large et al., 2010):

$$\begin{aligned} \tau_i \frac{dz_i}{dt} &= z_i \left( \alpha + i2\pi + \beta_1 |z_i|^2 + \varepsilon \beta_2 |z_i|^4 + \dots \right) \\ &+ c \mathcal{P}(\varepsilon, s(t)) \mathcal{A}(\varepsilon, z_i) \end{aligned} \quad (2)$$

where the roman  $i$  denotes the imaginary unit and  $z_i$  is the complex-valued state variable for the  $i$ th neural oscillator whose real and imaginary parts can be thought of as the activation of the excitatory and inhibitory subpopulations respectively. The parameter  $\alpha$  controls the bifurcation of the system's intrinsic behavior:  $\alpha = 0$  is the critical point, above which the system exhibits spontaneous oscillation and below which damped oscillation. The  $i$ th oscillator's natural frequency is given by  $f_i = 1/\tau_i$ , and  $\tau_i$  is varied to create a tonotopic arrangement as is found in the cochlea and brainstem (Langner, 1992). The  $\beta$ 's are nonlinear damping parameters, and the complete expansion of intrinsic terms describes a fully saturating nonlinearity (Large et al., 2010; Murdock, 2003).

Coupling to the stimulus  $s(t)$  is nonlinear, and consists of a series of terms called resonant monomials. Each resonant monomial accounts for a different mode-locked response of the neural oscillator to the stimulus (see Appendix A for a further mathematical description of resonant monomials). The canonical model includes a full expansion of resonant monomials which can be expressed as a 'passive' nonlinear function  $\mathcal{P}(\varepsilon, s(t))$  of the external stimulus,  $s(t)$ , multiplied by an 'active' nonlinear function  $\mathcal{A}(\varepsilon, z_i)$  of the current state,  $z_i$ . The full expansion of resonant monomials captures mode-locked responses at stimulus frequencies as well as at the harmonics, subharmonics, and any combination frequencies (see Appendix A). Thus, the nonlinear coupling predicts the nature and relative strength of mode-locked responses.

The ability to predict the full range of mode-locked responses to a complex, multi-frequency stimulus allows us to use the canonical model to extrapolate from observations of mode-locking in

individual auditory neurons (Arnold and Burkard, 1998, 2000; Large et al., 1998; Laudanski et al., 2010) to the responses of mode-locking auditory populations. This, in turn, enables prediction of the responses that would be observed in the human auditory brainstem. The parameter  $\varepsilon$  can be varied between 0 and 1 to control the contribution of each monomial to the overall dynamics of an oscillator. When  $\varepsilon$  is small only low order mode-locks are observed, as  $\varepsilon$  is increased, higher order mode-locks begin to appear. This provides a means to optimize model fits using a single degree of freedom, yielding a strong constraint on the relative amplitude of nonlinear resonances predicted for the brainstem response to a complex input signal.

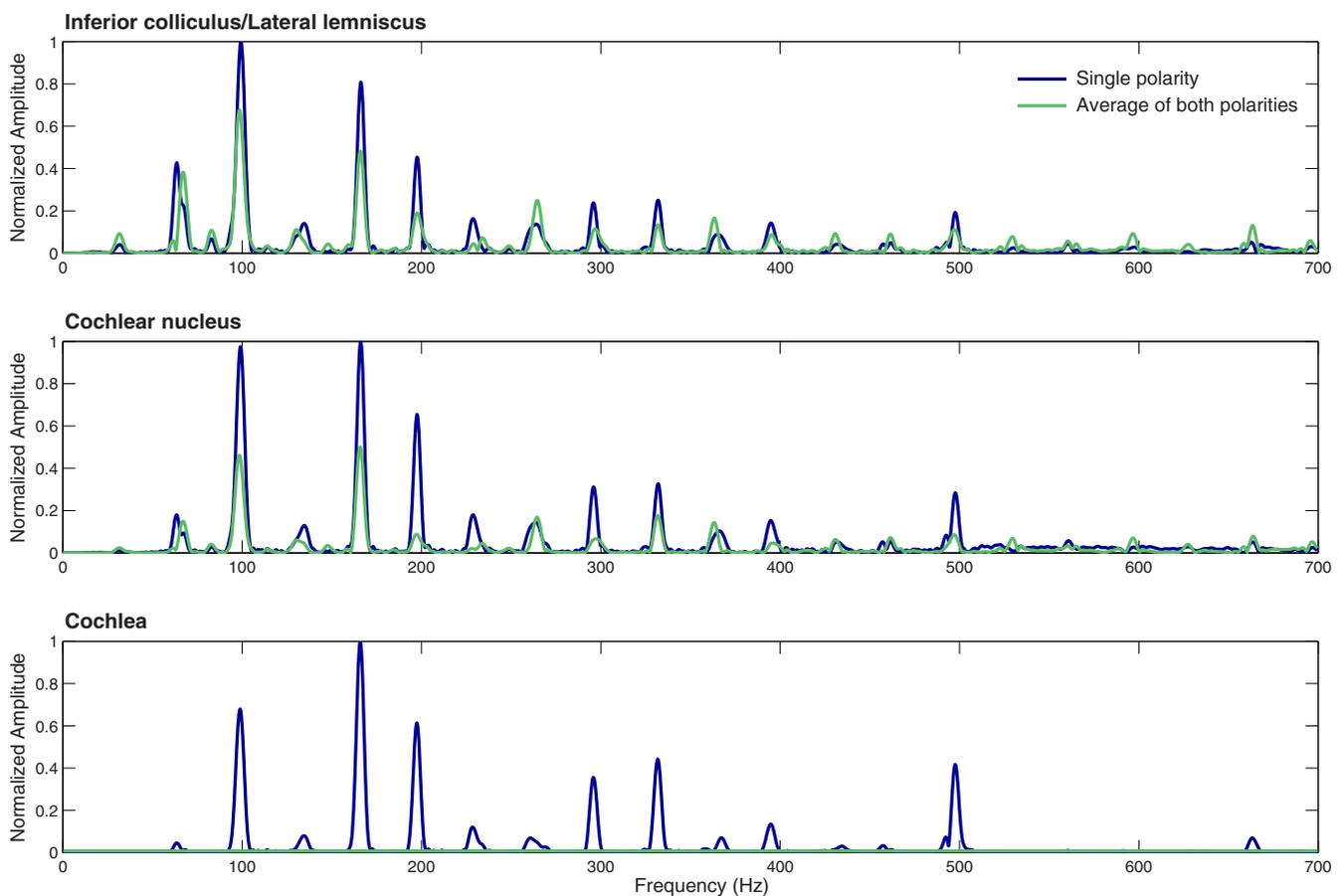
### 3. Methods

The stimuli from the Lee et al. (2009) study were used as input to a cochlear model, which in turn provided input to two brainstem network layers. The characteristic frequencies of the cochlear layer and both brainstem layers spanned four octaves with 99 oscillators per octave. Thus, each layer included 397 oscillators, with characteristic frequencies ranging from 64 Hz to 1024 Hz, encompassing the range of frequencies for which time-locked responses have been observed in midbrain physiology (Langner, 1992). The cochlear model includes a middle ear filter and simulates the basilar membrane and the organ of Corti (cf. Jülicher et al., 2001, see Appendix B). The cochlea is connected to the first brainstem layer, representing the cochlear nucleus (CN), and the CN is connected to the second brainstem layer, representing the inferior colliculus/lateral lemniscus (IC/LL). The oscillator equations for the brainstem layers utilize the complete canonical model, including the full expansion of resonant monomials. In these layers, the parameters were  $\alpha = 0$ ,  $\beta_1 = 0$ ,  $\beta_2 = -1$ .

We modeled four FFRs from Lee et al., musicians' and non-musicians' responses to both the consonant and dissonant intervals. Lee et al. used stimuli with alternating polarity and averaged the responses to these conditions, and our simulations mimicked this situation. For each model fit, a simulation was run for each polarity condition. The stimulus was fed to the cochlea, all oscillator equations were numerically integrated for the length of the stimulus, and the responses in all layers were stored. To compute the model brainstem FFR, the responses of all oscillators in each layer were averaged, leaving a single time series for each layer (see Fig. 3). There is evidence for the relative contributions of the various generators of the brainstem FFR from lesion studies in cats (Gardi et al., 1979), which indicates that the cochlear microphonic is responsible for 25%, the cochlear nuclei 50%, and the lateral lemniscus and inferior colliculus the remaining 25%. Following this scheme, the model FFR was a weighted average of the cochlea, CN, and IC/LL layers. This weighted average was filtered (3rd-order digital Butterworth low-pass, 450 Hz cutoff) to account for the low-pass effect of the skull, meninges, and scalp on the FFR. Finally, the resultant model time series for each polarity condition were averaged and fast Fourier transformed to produce a model fit. The four model fits, for the consonant and dissonant intervals, and for musician and nonmusician subjects, were optimized with a single degree of freedom through a series of simulations. In these simulations, the parameter  $\varepsilon$  was systematically varied between zero and one to yield the highest correlation for each fit.

### 4. Results

Fig. 3 shows the predicted responses of nonmusicians to the consonant interval at each model layer. For each layer, linear amplitude spectra for a single polarity condition and for the average of both polarity conditions are plotted. In addition to stimulus



**Fig. 3.** Mode-locking and population responses in a canonical model. Canonical neural oscillators can be organized into tonotopic networks and used to predict population responses. Shown are the predicted responses (linear amplitude spectra) of each layer for nonmusicians to the consonant stimulus. Both the averaged responses for both polarity conditions (green) and the response for a single polarity condition (blue) are shown. Notice that averaging over both polarity conditions cancels out any contribution of the cochlea to the model response.

frequencies, oscillations at many other frequencies achieve noticeable amplitude. It is important to note that the two-polarity average zeros out all cochlear responses; there is only odd-order activity, due to the nature of compression in the model cochlea. Because our predicted response is an average of both polarity conditions, it can only contain even-order nonlinearities. Odd-order distortions in the cochlear model do propagate up to the neural layers and there contribute to the formation of even-order nonlinearities. However, the central point is that the only mechanism for even-order nonlinearities in this model is mode-locking in the neural layers due to the full expansion of resonant monomials.

Fig. 4 shows the predicted brainstem responses, based on the two-polarity average. The canonical population response predicts each peak in the nonmusician brainstem response with remarkable accuracy for both intervals (consonant:  $R^2 = 0.77$ ,  $p < 0.0001$ ; dissonant:  $R^2 = 0.67$ ,  $p < 0.0001$ ). Musicians' responses did not correlate as highly, but were also predicted with good accuracy (consonant:  $R^2 = 0.74$ ,  $p < 0.0001$ ; dissonant:  $R^2 = 0.52$ ,  $p < 0.0001$ ).

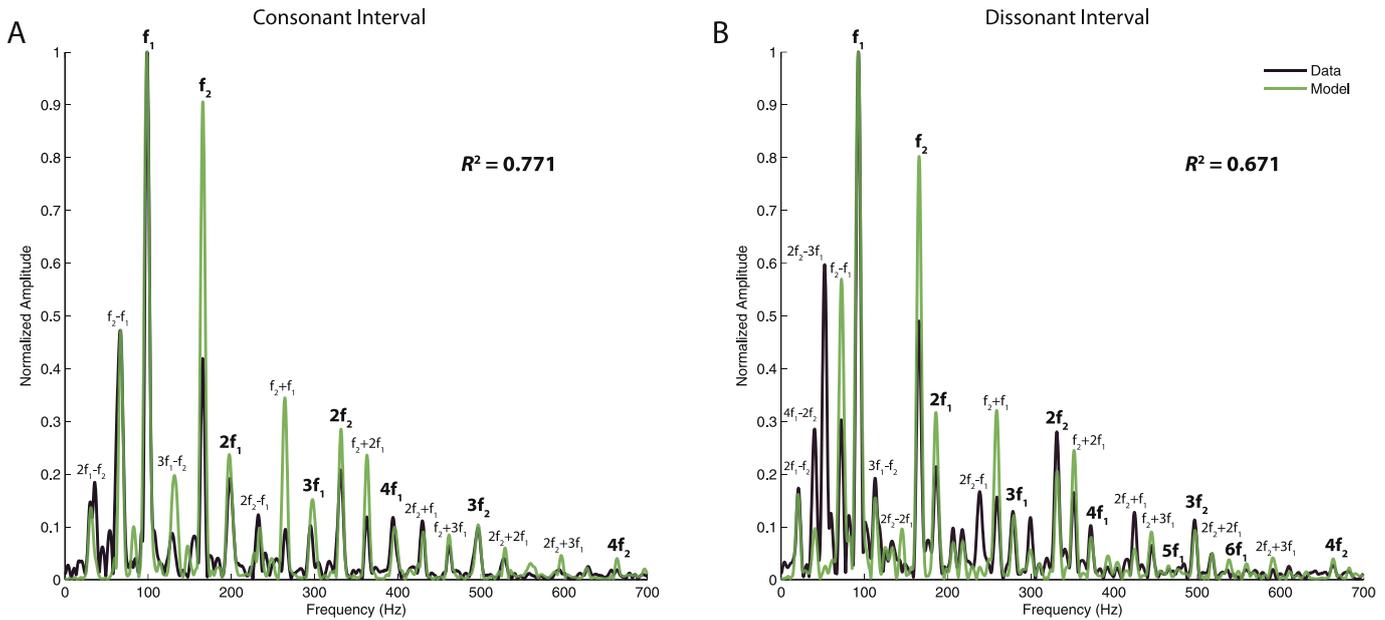
Best values of  $\epsilon$  were greater for musicians (consonant:  $\epsilon = 0.48$ ; dissonant:  $\epsilon = 0.41$ ) than for nonmusicians (consonant:  $\epsilon = 0.28$ ; dissonant:  $\epsilon = 0.07$ ). This suggests that the difference between musician and nonmusician responses occurs partly because musicians' processing of musical intervals is more strongly nonlinear. Stronger nonlinearities are observed in such systems as a result of stronger gain, therefore this could point to overall synaptic efficacy.

This cannot be the entire explanation, however, because the model also explained a smaller overall proportion of the variance due to enhanced musician responses at certain frequencies. These more subtle differences could not be accounted for by manipulation of  $\epsilon$  alone, implying that incorporation of other network properties into the model, i.e., synaptic coupling, will be necessary to explain the responses of trained listeners. This is consistent with the interpretation that the refinement of auditory sensory encoding in musicians is driven by synaptic plasticity that links learned representations to the neural encoding of acoustic features (Lee et al., 2009).

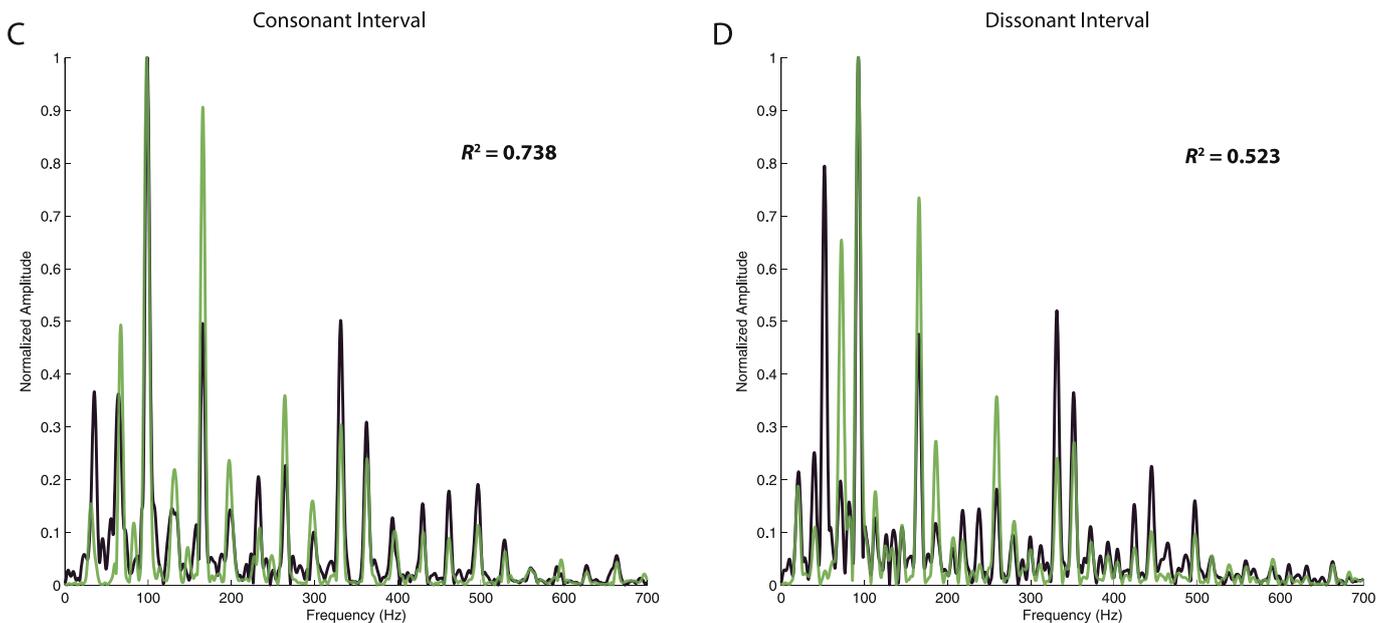
## 5. Discussion

Overall, the model explained 68% of the variance in human brainstem data with a single degree of freedom fit, while holding intrinsic and network parameters constant to evaluate the basic approach. The remaining discrepancies could likely be accounted for if these parameters were allowed to vary, modulating the relative amplitude of response at different frequencies, as no doubt would occur in the central auditory pathway. Therefore, the hypothesis of active mode-locking neurodynamics in the central auditory pathway appears sufficient to explain most of the nonlinear responses observed in the human brainstem. Both the CN and IC/LL layers contributed to the relative strengths of model components, and from a signal processing point of view, processing

Nonmusicians



Musicians



**Fig. 4.** Model & data comparison. Comparisons of model predictions and auditory brainstem responses of nonmusicians to (A) the consonant interval (99 Hz, 166 Hz) and (B) the dissonant interval (93 Hz, 166 Hz), and of musicians to (C) the consonant interval and (D) the dissonant interval. The labels above each spectral component refer only to their specific frequencies as functions of the primaries, and do not necessarily reflect the generating processes of those components (see Discussion).

by multiple nonlinear transformations is not redundant; information is added at every processing stage, and each stage would contribute to perceptual processing in important, and possibly unique ways.

While odd-order nonlinearities in the FFR are typically attributed to the cochlea, the generation site of even-order nonlinearities is far less clear (Bhagat and Champlin, 2004). In our model of the FFR, the absence of contribution of the cochlea and the strictly even-order nature of the response are true by definition. However, it is easy to see how this could be the case in the human auditory system as well: Every prominent frequency in all responses can be generated by even-order interaction between

stimulus components (see Fig. 4). The response at  $2f_1 - f_2$ , for instance, is a typical example of a cubic distortion product; however in this case the frequency  $2f_1$  actually occurs in the stimuli as the second harmonic of  $f_1$ . Thus the component at  $2f_1 - f_2$  could be a simple difference tone. If one compares all the nonlinearities in the response to the frequencies in the stimuli, the potential quadratic nature of all response frequencies is clear. As the generation of even-order nonlinearities in scalp-recorded potentials is much more likely neural, possibly as a result of envelope-following, than that of odd-order nonlinearities, the similarity of our model FFR to the actual FFR is perhaps further explained.

Dichotic stimulation is commonly held to be a good potential indicator of the neural generation of nonlinear resonances. Yet there is little evidence of any nonlinearity in scalp-recorded potentials under dichotic stimulation (Bidelman and Krishnan, 2009; Gockel et al., 2011). It is important to note, however, that dichotic stimulation prevents the interaction of stimulus components in the cochlear nuclei, sites of activity at which mode-locking has been directly observed (Laudanski et al., 2010). Thus, the lack of scalp-recorded nonlinear resonances under dichotic stimulation does not constitute strong counter-evidence for this model, or for the claim that the nonlinear resonances observed by Lee et al. (2009) were primarily generated neurally.

Traditional models of central auditory processing rely upon synaptic delays to enable temporal computations (e.g., autocorrelation), thought to be necessary for auditory processes such as pitch perception (Licklider, 1951). However, the stark lack of evidence for neural delays over the time-scale required for pitch perception presents a significant puzzle for this theoretical approach (de Cheveigné and Pressnitzer, 2006). Moreover, key results in pitch perception, including the pitch of sinusoidally amplitude-modulated noise (Burns and Viemeister, 1976, 1981), Stevens' rule (e.g., Hartmann, 1996), and the lack of pitch in certain sounds containing autocorrelation peaks (Kaernbach and Demany, 1998), cannot be explained by delay-based models. These problems in pitch perception, as well as more traditional ones such as the pitch of the missing fundamental and pitch shift of the residue (Schouten et al., 1962), could be explained as higher-order resonance in nonlinear oscillators (e.g., Cartwright et al., 1999).

Oscillatory neurodynamics could also explain key aspects of music perception and cognition. Recent empirical studies have shown that harmonicity (integer ratio frequency relationships) provides a better explanation for consonance and dissonance than does beating (McDermott et al., 2010; Cousineau et al., 2012). Thus, the explanation for consonance may lie in low-order integer frequency ratios, which is a fundamental principle of stability in oscillatory dynamical systems (Shapira Lots and Stone, 2008). It has also been shown that the relative strength of mode-locked neural responses could account for tonal stability, predicting a higher proportion of variance in listener judgments of tonal stability in Western major and minor modes than psychoacoustic measurements of sensory consonance (Large, 2011a; Krumhansl and Kessler, 1982; Krumhansl, 1990). Tonal stability and attraction are important because these are the properties that are thought to differentiate musical sound sequences from arbitrary sound sequences, and enable non-referential sound patterns to carry meaning (Zuckermandl, 1956). Moreover, neurodynamic predictions matched the perception of tonal stability in North Indian raga by familiar Indian listeners, as well as by unfamiliar Western listeners (Large et al., submitted; Castellano et al., 1984). Thus, the theoretical stability of mode-locking interactions in oscillatory neural populations predicts empirically measured stability cross-culturally, consistent with a neurodynamic explanation.

## 6. Limitations and future work

A single parameter of a model brainstem network was manipulated to fit Lee et al.'s (2009) brainstem FFR data. The parsimony of the model, its basis in neurophysiological observations of mode-locking, and the quality of the fits all speak to the potential of this theoretical approach. It is clear that no model can provide direct evidence for the physiological generators of auditory nonlinearities. However, to the extent that such modeling efforts clarify the nature of nonlinear resonance in the auditory system, predictions and implications for empirical research may ultimately lead to a fuller understanding of the mechanisms generating the

FFR. If successful, dynamical systems theory applied to nonlinear auditory signal processing may have further application in understanding other dynamic aspects of neural processing.

Our simplified cochlear model only generates odd-order nonlinearities, however even-order nonlinearities, mainly the simple difference tone (DT), have also been observed in mammalian cochleae. A more accurate cochlear model that performed this way would no doubt be more biologically plausible, and would likely alter the modeled brainstem response, if only slightly. It is important to note, however, that many studies show substantial differences in the behavior of the difference tone and the odd-order cubic difference tone (CDT) under two-tone stimulation (Cooper and Rhode, 1997; Bian and Chen, 2008). These differences are evident, for instance, in their sensitivities to  $f_2/f_1$  ratio, absolute frequencies of stimulation, ratio of primary levels, and absolute primary levels. While the CDT, along with the rest of the odd-order nonlinearities commonly seen in both real and model cochleae, can be fully explained through compressive nonlinearity, the explanation of the DT is far less clear. These facts together suggest different generating mechanisms for the components (Hall, 1974; Kujawa et al., 1995; Chang and Norton, 1997), and it cannot be ruled out that the DT is partially generated neurally in the brainstem and is propagated back to the basilar membrane via the efferent olivocochlear system. Whether or not this is the case, it is well known that this efferent connectivity is ubiquitous and important in the auditory system; thus the addition of efferent connectivity to this model is in order.

In general, improvements will be necessary to enable more comprehensive simulations of the early auditory system that include all relevant aspects of cochlear dynamics, as well as parameterization of CN and IC/LL dynamics (Large and Almonte, 2012). In comprehensive models, parameter fitting will become a significant issue. Finally, it will be necessary to explore models of synaptic plasticity for neural oscillator networks (e.g., Hoppensteadt and Izhikevich, 1996b; Large, 2011b) in order to better explain the responses of trained listeners.

## 7. Conclusions

Neural models explain the perception of sound patterns as the interaction of sound stimuli with neurophysiological processes. Canonical models use the tools of nonlinear dynamical systems to make predictions about perception that are consistent with our understanding of the neurophysiology (e.g., mode-locking), but are not overly dependent on the neuronal-level details. This approach may lead to an understanding of general neural signal processing principles underlying music and pitch perception. Moreover, canonical analysis of plasticity in neural oscillator networks may help us to understand the role of learning in modulating these responses. Thus, future modeling efforts based on canonical dynamical systems could bring us closer to understanding fundamental mechanisms of hearing, communication, and auditory system development.

## Acknowledgments

This work was supported by AFOSR FA9550-07-C0095 and NSF BCS-1027761. Thanks to Kimi Lee, Erika Skoe, Nina Kraus and Ric Ashley for providing the brainstem data, and for valuable discussions and comments regarding this work.

## Appendix A. Resonant monomials in the canonical model

The derivation of resonant monomials makes use of Poincaré-Dulac normal form theory (Hoppensteadt and Izhikevich, 1997;

Murdock, 2003; Wiggins, 1990), and yields resonant monomials of the form

$$\varepsilon^{((\sum_j p_j + q_j) + r_i - 2)} / 2 x_1^{p_1} \dots x_n^{p_n} \bar{x}_1^{q_1} \dots \bar{x}_n^{q_n} \bar{z}_i^{r_i - 1} \quad (\text{A.1})$$

where each monomial leads to a specific resonance in the network. Here,  $x_1, \dots, x_n$  and  $\bar{x}_1, \dots, \bar{x}_n$  represent the individual frequency components of the input signal and their complex conjugates, and  $p_1, \dots, p_n$ ,  $q_1, \dots, q_n$ , and  $r_i$  are natural numbers where at least one of them is not zero. This analysis predicts responses at harmonics (e.g., the monomial  $\sqrt{\varepsilon} x_1^2$  implies a resonant frequency equal to  $2f_1$ ), subharmonics (e.g.,  $\sqrt{\varepsilon} x_1 \bar{z}_i$  corresponds to  $f_1/2$ ), integer ratios (e.g.,  $\varepsilon \sqrt{\varepsilon} x_1^3 \bar{z}_i$  corresponds to  $3f_1/2$ ), combination frequencies (e.g.,  $\varepsilon x_1^2 \bar{x}_2$  corresponds to  $2f_1 - f_2$ ) and so forth. The  $\varepsilon$ -based coefficient predicts the relative strength of the nonlinear resonances.

The canonical model includes a full expansion of resonant monomials, one for every combination of input frequencies that could lead to a mode-locked resonance, so that each oscillator responds appropriately to a complex stimulus that contains multiple frequency components (Large et al., 2010). The analysis leads to one form of full expansion that includes all possible resonant monomials of the form shown in (A.1). This infinite sum of resonant monomials can be factored into a passive nonlinear function  $\mathcal{P}(\varepsilon, s(t))$  and an active nonlinear function  $\mathcal{A}(\varepsilon, z_i)$ :

$$\begin{aligned} \mathcal{P}(\varepsilon, s(t)) &= (s + \sqrt{\varepsilon} s^2 + \varepsilon s^3 + \dots) \left( 1 + \sqrt{\varepsilon} \bar{s} + \varepsilon \bar{s}^2 + \dots \right) \\ &= \left( \frac{s}{1 - \sqrt{\varepsilon} s} \right) \left( \frac{1}{1 - \sqrt{\varepsilon} \bar{s}} \right), \\ \mathcal{A}(\varepsilon, z_i) &= 1 + \sqrt{\varepsilon} \bar{z}_i + \varepsilon \bar{z}_i^2 + \dots = \frac{1}{1 - \sqrt{\varepsilon} \bar{z}_i}, \end{aligned} \quad (\text{A.2})$$

where  $s(t) = x_1 + \dots + x_n$  is the input signal.

The stimulus polarity reversal technique is commonly used to separate out different types of nonlinear responses. The summation of responses obtained using alternating stimulus polarity contains only even-order nonlinearities whereas the subtraction of responses to different polarities contains only odd-order nonlinearities. This relationship between stimulus polarity and nonlinear response can be explained succinctly in the context of the canonical model presented in this paper.

Let us consider, for instance, a stimulus including two frequency components,  $s(t) = x_1(t) + x_2(t)$ . Reversing the polarity of the stimulus (i.e., using  $-s(t)$  instead of  $s(t)$  as stimulus) reverses the polarity of the third-order passive nonlinearity  $x_1^2 \bar{x}_2$  (corresponding to the cubic difference tone  $2f_1 - f_2$ ) since  $(-x_1)^2 (-\bar{x}_2) = -x_1^2 \bar{x}_2$ . On the other hand, the second-order  $x_2 \bar{x}_1$  (the quadratic difference tone  $f_2 - f_1$ ) does not change polarity under stimulus polarity reversal since  $(-x_2)(-\bar{x}_1) = x_2 \bar{x}_1$ . To generalize, odd-order nonlinear responses change polarity under stimulus polarity reversal whereas even-order ones do not. Thus, adding responses to alternating stimulus polarity leads to doubling even-order nonlinear components while eliminating odd-order ones. Subtracting the two responses, on the other hand, doubles odd-order nonlinearities and eliminates even-order ones.

## Appendix B. Cochlear model

The cochlear model is motivated by recent nonlinear models of the cochlea (Jülicher et al., 2001; Szalai et al., 2013). The auditory stimulus provided to the cochlear model was Hilbert transformed and pre-filtered using the human middle-ear filter (MEF) model of Zilany and Bruce (2006) (see also Bruce et al., 2003). The resulting complex waveform, denoted by  $x(t)$ , was provided as input to the

cochlear network. We modeled cochlear dynamics using a tonotopically organized network of Hopf oscillators, each tuned to a distinct natural frequency. The coupled system given by Eq. (B.1) represents a cochlear section which encompasses the action of basilar membrane (BM) in conjunction with the dynamics of the outer hair cells and related organ of Corti (OC) components. The first equation below accounts for passive BM dynamics, similar to linear pre-filtering (Jülicher et al., 2001), and is formally related to a gammatone filter (Lyon et al., 2010). The second equation accounts for nonlinear OC dynamics. Nonlinearities in the cochlear response are generated due to amplitude compression and controlled by the  $\beta$  parameters.

$$\begin{aligned} \tau \frac{dz_{\text{bm}}}{dt} &= z_{\text{bm}} (\alpha_{\text{bm}} + i2\pi) + x(t) \\ \tau \frac{dz_{\text{oc}}}{dt} &= z_{\text{oc}} \left( \alpha_{\text{oc}} + i2\pi + \beta_1 |z_{\text{oc}}|^2 + \varepsilon \beta_2 |z_{\text{oc}}|^4 + \dots \right) + z_{\text{bm}} \end{aligned} \quad (\text{B.1})$$

The time dependent variables  $z_{\text{bm}}(t)$  and  $z_{\text{oc}}(t)$  (which, for notational simplicity, are written as  $z_{\text{bm}}$  and  $z_{\text{oc}}$ ) are complex-valued state variables representing the response of the BM and OC respectively. Each oscillator's natural frequency is  $f = 1/\tau$  Hz, and  $\alpha_{\text{bm}} < 0$ ,  $\alpha_{\text{oc}} = 0$  are linear damping and bifurcation parameters, respectively, and  $x(t)$  denotes linear forcing by a time-varying external signal. The only mechanism for generation of nonlinear distortion in this model is compression. Thus, all distortions produced by the cochlear model are odd-order (Jülicher et al., 2001; Stoop et al., 2005).

This kind of cochlear model can account for a nontrivial subset of cochlear dynamics such as sharp mechanical frequency tuning, exquisite sensitivity, and a large dynamic range (Eguíluz et al., 2000; Magnasco, 2003; Mora and Bialek, 2011; Ospeck et al., 2001). To tune the model's parameters, macaque monkey tuning curve data (Joris et al., 2011) was used to determine reasonable values. For the simulations, the parameter values  $\alpha_{\text{bm}} = -1$ ,  $\alpha_{\text{oc}} = 0$ ,  $\beta_1 = -10,000$ ,  $\beta_2 = -1$ ,  $\varepsilon = .0025$  were used to define the cochlear network spanning a frequency range within the bounds of the human cochlea. The natural frequencies of the oscillators were spaced logarithmically with a center frequency CF of 256 Hz, 99 oscillatory complexes as defined by Eq. (B.1) per octave, and 2 octaves on each side of CF.

## References

- Ainsworth, M., Lee, S., Cunningham, M.O., Traub, R.D., Kopell, N.J., Whitington, M.A., 2012. Rates and rhythms: a synergistic view of frequency and temporal coding in neuronal networks. *Neuron* 75, 572–583.
- Arnold, S., Burkard, R., 1998. The auditory evoked potential difference tone and cubic difference tone measured from the inferior colliculus of the chinchilla. *J. Acoust. Soc. Am.* 104, 1565–1573.
- Arnold, S., Burkard, R., 2000. Studies of interaural attenuation to investigate the validity of a dichotic difference tone response recorded from the inferior colliculus in the chinchilla. *J. Acoust. Soc. Am.* 107, 1541–1547.
- Bhagat, S.P., Champlin, C.A., 2004. Evaluation of distortion products produced by the human auditory system. *Hear. Res.* 193, 51–67.
- Bian, L., Chen, S., 2008. Comparing the optimal signal conditions for recording cubic and quadratic distortion product otoacoustic emissions. *J. Acoust. Soc. Am.* 124, 3739–3750.
- Bidelman, G.M., Krishnan, A., 2009. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J. Neurosci.* 29, 13165–13171.
- Bruce, I.C., Sachs, M.B., Young, E.D., 2003. An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses. *J. Acoust. Soc. Am.* 113, 369–388.
- Brumberg, J.C., Gutkin, B.S., 2007. Cortical pyramidal cells as non-linear oscillators: experiment and spike-generation theory. *Brain Res.* 1171, 122–137.
- Burns, E.M., Viemeister, N.F., 1976. Nonspectral pitch. *J. Acoust. Soc. Am.* 60, 863–869.
- Burns, E.M., Viemeister, N.F., 1981. Played-again sam: further observations on the pitch of amplitude-modulated noise. *J. Acoust. Soc. Am.* 70, 1655–1660.

- Cartwright, J.H.E., González, D.L., Piro, O., 1999. Nonlinear dynamics of the perceived pitch of complex sounds. *Phys. Rev. Lett.* 82, 5389–5392.
- Castellano, M.A., Bharucha, J.J., Krumhansl, C.L., 1984. Tonal hierarchies in the music of north india. *J. Exp. Psychol. Gen.* 113, 394–412.
- Chang, K.W., Norton, S.J., 1997. Efferently mediated changes in the quadratic distortion product ( $f_2-f_1$ ). *J. Acoust. Soc. Am.* 102, 1719–1733.
- Chertoff, M.E., Hecox, K.E., 1990. Auditory nonlinearities measured with auditory-evoked potentials. *J. Acoust. Soc. Am.* 87, 1248–1254.
- de Cheveigné, A., Pressnitzer, D., 2006. The case of the missing delay lines: synthetic delays obtained by cross-channel phase interaction. *J. Acoust. Soc. Am.* 119, 3908–3918.
- Cooper, N.P., Rhode, W.S., 1997. Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea. *J. Neurophysiol.* 78, 261–270.
- Cousineau, M., McDermott, J.H., Peretz, I., 2012. The basis of musical consonance as revealed by congenital amusia. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19858–19863.
- Dhar, S., Abel, R., Hornickel, J., Nicol, T., Skoe, E., Zhao, W., Kraus, N., 2009. Exploring the relationship between physiological measures of cochlear and brainstem function. *Clin. Neurophysiol.* 120, 959–966.
- Dhar, S., Long, G.R., Talmadge, C.L., Tubis, A., 2005. The effect of stimulus-frequency ratio on distortion product otoacoustic emission components. *J. Acoust. Soc. Am.* 117, 3766–3776.
- Eguíluz, V.M., Ospeck, M., Choe, Y., Hudspeth, A.J., Magnasco, M.O., 2000. Essential nonlinearities in hearing. *Phys. Rev. Lett.* 84, 5232–5235.
- Escabi, M.A., Schreiner, C.E., 2002. Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. *J. Neurosci.* 22, 4114–4131.
- Galbraith, G.C., 1994. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalogr. Clin. Neurophysiol.* 92, 321–330.
- Gardi, J., Salamy, A., Mendelson, T., 1979. Scalp-recorded frequency-following responses in neonates. *Audiology* 18, 494–506.
- Gockel, H.E., Carlyon, R.P., Mehta, A., Plack, C.J., 2011. The frequency following response (ffr) may reflect pitch-bearing information but is not a direct representation of pitch. *J. Assoc. Res. Otolaryngol.* 12, 767–782.
- Hall, J.L., 1974. Two-tone distortion products in a nonlinear model of the basilar membrane. *J. Acoust. Soc. Am.* 56, 1818–1828.
- Hartmann, W.M., 1996. Pitch, periodicity, and auditory organization. *J. Acoust. Soc. Am.* 100, 3491–3502.
- Hoppensteadt, F.C., Izhikevich, E.M., 1996a. Synaptic organizations and dynamical properties of weakly connected neural oscillators. i. analysis of a canonical model. *Biol. Cybern.* 75, 117–127.
- Hoppensteadt, F.C., Izhikevich, E.M., 1996b. Synaptic organizations and dynamical properties of weakly connected neural oscillators. ii. learning phase information. *Biol. Cybern.* 75, 129–135.
- Hoppensteadt, F.C., Izhikevich, E.M., 1997. *Weakly Connected Neural Networks*. Springer, ISBN 0387949488.
- Joris, P.X., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., Shera, C.A., 2011. Frequency selectivity in old-world monkeys corroborates sharp cochlear tuning in humans. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17516–17520.
- Joris, P.X., Schreiner, C.E., Rees, A., 2004. Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577.
- Jülicher, F., Andor, D., Duke, T., 2001. Physical basis of two-tone interference in hearing. *Proc. Natl. Acad. Sci.* 98, 9080–9085.
- Kaernbach, C., Demany, L., 1998. Psychophysical evidence against the autocorrelation theory of auditory temporal processing. *J. Acoust. Soc. Am.* 104, 2298–2306.
- Kern, A., Stoop, R., 2003. Essential role of couplings between hearing nonlinearities. *Phys. Rev. Lett.* 91, 128101.
- Kikuchi, Y., Horwitz, B., Mishkin, M., Rauschecker, J., 2011. Harmonic Preference in the Lateral Belt of Rhesus Monkey Auditory Cortex. Society for Neuroscience. Poster No. 480.20/LL14.
- Knight, R.D., Kemp, D.T., 2001. Wave and place fixed dpoae maps of the human ear. *J. Acoust. Soc. Am.* 109, 1513–1525.
- Koepsell, K., Wang, X., Hirsch, J.A., Sommer, F.T., 2010. Exploring the function of neural oscillations in early sensory systems. *Front Neurosci.* 4, 53–61.
- Krumhansl, C.L., Kessler, E.J., 1982. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* 89, 334–368.
- Krumhansl, C.L., 1990. *Cognitive Foundations of Musical Pitch*. Oxford University Press.
- Kujawa, S.G., Fallon, M., Bobbin, R.P., 1995. Time-varying alterations in the  $f_2-f_1$  dpoae response to continuous primary stimulation. i: response characterization and contribution of the olivocochlear efferents. *Hear. Res.* 85, 142–154.
- Langner, G., 1992. Periodicity coding in the auditory system. *Hear. Res.* 60, 115–142.
- Langner, G., 2007. Temporal processing of periodic signals in the auditory system: neuronal representation of pitch, timbre, and harmonicity. *Z. Audiol.* 46, 8–21.
- Large, E.W., 2011a. A dynamical systems approach to musical tonality. In: Huys, R., Jirsa, V.K. (Eds.), *Nonlinear Dynamics in Human Behavior, Studies in Computational Intelligence*, vol. 328. Springer, Berlin Heidelberg, pp. 193–211.
- Large, E.W., 2011b. Musical tonality, neural resonance and hebbian learning. In: Agon, C., Andreatta, M., Assayag, G., Amiot, E., Bresson, J., Mandereau, J. (Eds.), *Mathematics and Computation in Music, Lecture Notes in Computer Science*, vol. 6726. Springer, Berlin Heidelberg, pp. 115–125.
- Large, E.W., Almonte, F.V., 2012. Neurodynamics, tonality, and the auditory brainstem response. *Ann. NY Acad. Sci.* 1252, E1–E7.
- Large, E.W., Almonte, F.V., Velasco, M.J., 2010. A canonical model for gradient frequency neural networks. *Physica D: Nonlinear Phenom.* 239, 905–911.
- Large, E.W., Crawford, J.D., 2002. Auditory temporal computation: Interval selectivity based on post-inhibitory rebound. *J. Comput. Neurosci.* 13, 125–142.
- Large E.W., Kim J.C., Flaig N.K., Bharucha J.J. and Krumhansl C.L., *Neurodynamics of music perception, (submitted for publication)*.
- Large, E.W., Kozloski, J., Crawford, J.D., 1998. A dynamical model of temporal processing in the fish auditory system. In: *Proceedings of the Association for Research in Otolaryngology*. Abstr., 21, p. 717.
- Laudanski, J., Coombes, S., Palmer, A.R., Sumner, C.J., 2010. Mode-locked spike trains in responses of ventral cochlear nucleus chopper and onset neurons to periodic stimuli. *J. Neurophysiol.* 103, 1226–1237.
- Lee, K.M., Skoe, E., Kraus, N., Ashley, R., 2009. Selective subcortical enhancement of musical intervals in musicians. *J. Neurosci.* 29, 5832–5840.
- Lee, S.G., Kim, S., 2006. Bifurcation analysis of mode-locking structure in a hodgkin-huxley neuron under sinusoidal current. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* 73, 041924.
- Licklider, J.C.R., 1951. A duplex theory of pitch perception. *Experientia* 7, 128–134.
- Lyon, R., Katsiamis, A., Drakakis, E., 2010. History and future of auditory filter models. In: *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pp. 3809–3812.
- Magnasco, M.O., 2003. A wave traveling over a hopf instability shapes the cochlear tuning curve. *Phys. Rev. Lett.* 90, 058101.
- McDermott, J.H., Lehr, A.J., Oxenham, A.J., 2010. Individual differences reveal the basis of consonance. *Curr. Biol.* 20, 1035–1041.
- Mora, T., Bialek, W., 2011. Are biological systems poised at criticality? *J. Stat. Phys.* 144, 268–302.
- Murdock, J.A., 2003. *Normal Forms and Unfoldings for Local Dynamical Systems*. Springer, ISBN 978-0387954646.
- Ospeck, M., Eguíluz, V.M., Magnasco, M.O., 2001. Evidence of a hopf bifurcation in frog hair cells. *Biophys. J.* 80, 2597–2607.
- Pandya, P.K., Krishnan, A., 2004. Human frequency-following response correlates of the distortion product at  $2f_1-f_2$ . *J. Am. Acad. Audiol.* 15, 184–197.
- Perko, L., 1996. *Differential Equations and Dynamical Systems*. In: *Texts in Applied Mathematics*, vol. 7. Springer, US.
- Purcell, D.W., Ross, B., Picton, T.W., Pantev, C., 2007. Cortical responses to the  $2f_1-f_2$  combination tone measured indirectly using magnetoencephalography. *J. Acoust. Soc. Am.* 122, 992–1003.
- Rickman, M.D., Chertoff, M.E., Hecox, K.E., 1991. Electrophysiological evidence of nonlinear distortion products to two-tone stimuli. *J. Acoust. Soc. Am.* 89, 2818–2826.
- Robles, L., Ruggero, M.A., 2001. Mechanics of the mammalian cochlea. *Physiol. Rev.* 81, 1305–1352.
- Ruggero, M.A., Rich, N.C., Recio, A., Narayan, S.S., Robles, L., 1997. Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J. Acoust. Soc. Am.* 101, 2151–2163.
- Schouten, J.F., Ritsma, R.J., Cardozo, B.L., 1962. Pitch of the residue. *J. Acoust. Soc. Am.* 34, 1418–1424.
- Schwarz, D.W., Dezs, A., Neufeld, P.R., 1993. Frequency selectivity of central auditory neurons without inner ear. *Acta Otolaryngol.* 113, 266–270.
- Shapira Lots, I., Stone, L., 2008. Perception of musical consonance and dissonance: an outcome of neural synchronization. *J. R Soc. Interface* 5, 1429–1434.
- Skoe, E., Kraus, N., 2010. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 31, 302–324.
- Stoop, R., Steeb, W.H., Gallas, J., Kern, A., 2005. Auditory two-tone suppression from a subcritical hopf cochlea. *Phys. A: Stat. Mechanics Appl.* 351, 175–183.
- Sutter, M.L., Schreiner, C.E., 1991. Physiology and topography of neurons with multi-peaked tuning curves in cat primary auditory cortex. *J. Neurophysiol.* 65, 1207–1226.
- Szalai, R., Champneys, A., Homer, M., Ó Maoiléidigh, D., Kennedy, H., Cooper, N., 2013. Comparison of nonlinear mammalian cochlear-partition models. *J. Acoust. Soc. Am.* 133, 323–336.
- Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239.
- Wiggins, S., 1990. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer, ISBN 0387001778.
- Wile, D., Balaban, E., 2007. An auditory neural correlate suggests a mechanism underlying holistic pitch perception. *PLoS One* 2, e369.
- Wilson, H.R., Cowan, J.D., 1973. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80.
- Zilany, M.S.A., Bruce, I.C., 2006. Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery. *J. Acoust. Soc. Am.* 120, 1446–1466.
- Zuckerandl, V., 1956. *Sound and Symbol: Music and the External World*. Pantheon Books.