

# A Dynamical Systems Approach to Musical Tonicity

Edward W. Large

**Abstract.** Music is a form of communication that relies on highly structured temporal sequences comparable in complexity to language. Music is found among all human cultures, and musical languages vary across cultures with learning. Tonicity – a set of stability and attraction relationships perceived among musical frequencies – is a universal feature of music, found in virtually every musical culture. In this chapter, a new theory of central auditory processing and development is proposed, and its implications for tonal cognition and perception are explored. A simple model is put forward, based on knowledge of auditory organization and general neurodynamic principles. The model is simplified as compared to the organization and dynamics of the real auditory system, nevertheless it makes realistic predictions about neurodynamics. The analysis predicts the existence of natural resonances, the potential for tonal language learning, the perceptual categorization of intervals, and most importantly, relative stability and attraction relationships among musical tones. This approach suggests that high-level music cognition and perception may arise from the interaction of acoustic signals with the dynamics of the auditory system. Musical universals are predicted by intrinsic neurodynamics that provide a direct link to neurophysiology, and Hebbian synaptic modification could explain how different tonal languages are established.

## 1 Introduction

The music of almost every instrumental culture is tonal. In tonal music, one specific tone, called the tonic, provides a focus around which other tones are organized. Musical melodies typically involve discrete tones, organized in

---

Edward W. Large  
Center for Complex Systems and Brain Sciences  
Florida Atlantic University 777 Glades Road, Bldg 12, Rm 327,  
Boca Raton, FL 33431  
Ph.: 561.297.0106  
Fax: 561.297.3634  
e-mail: [large@ccs.fau.edu](mailto:large@ccs.fau.edu)

archetypal patterns that are characteristic of musical genres, styles, and cultures. These patterns may be related to a scale, an ordered collection of all the tones used in a given melody, summarizing the frequency ratios that govern the intervals between tones in a melody. The Western tuning system defines an inclusive scale, the chromatic scale, which divides the octave into 12 steps, called semitones. A subset of these 12 tones, called a diatonic scale, is typically used to create a melody. In modern Western tonality, the main diatonic scales are major and minor. When a melody is in a key, say C major, the notes of the C major scale are used to create that melody. One feature that the melodies of most musical systems share is that they give rise to tonal percepts. The tonic – C, in the key of C major – is said to be the most stable tone in that key. Stability means that the tone is perceived as a point of repose. For example, a melody in the key of C major will almost always end on the tonic, C. Among the other tones of the scale, there is a hierarchy of relative stability, such that some tones are perceived as more stable than others. Less stable tones provide points of dissonance or tension, more stable tones provide points of consonance or relaxation. Finally, less stable tones are heard relative to more stable ones, such that more stable tones are said to attract the less stable tones.

What processes and network architectures in the nervous system could give rise to such perceptions in music? This chapter argues that nonlinear neural resonance underlies the perception of tonality. Universal properties of nonlinear resonance predict the perception of stability and attraction in tonal music as well as preferences for small integer ratios and perceptual categorization. Such principles could provide a set of innate constraints that shape human musical behavior and enable the acquisition of musical knowledge.

## 2 Tonality

The oldest theory of musical consonance is that perceptions of consonance and dissonance are governed by ratios of whole numbers. Pythagoras is thought to have first articulated the principle that intervals of small integer ratios are pleasing because they are mathematically pure (Burns, 1999). He used this principle to explain the musical scale that was in use in the West at the time, and Pythagoras and his successors proposed small-integer-ratio systems for tuning musical instruments, such as Just Intonation (JI) (see Table 1). Because transposition on fixed tuning instruments, like the piano, is problematic for JI, modern Western equal temperament (ET), divides the octave into 12 intervals that are precisely equal on a log scale. ET approximates JI, and transposition in ET is perfect, because the frequency ratio of each interval is invariant. However, aside from octaves the intervals are not small integer ratios, they are irrational. The fact that equal tempered intervals sound approximately as consonant as neighboring small-integer-ratio intervals is generally considered *prima facie* evidence against the Pythagorean theory of musical consonance.

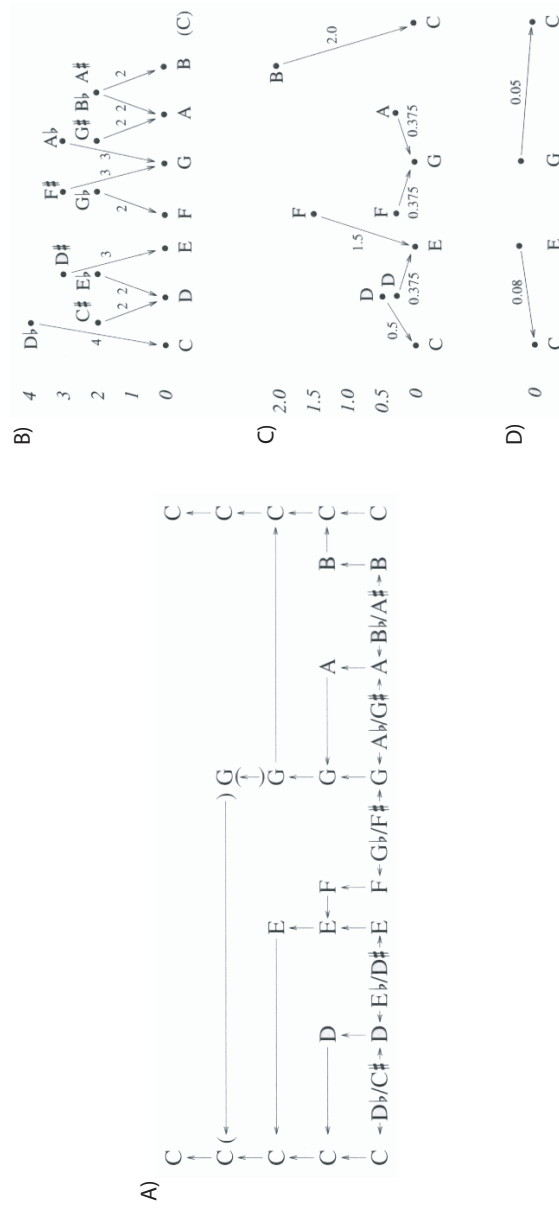
Helmholtz (1863) hypothesized that the dissonance of a pair of simultaneously sounding complex tones was due to the interference of its pure tone components, explaining dissonance as an unpleasant sensation of roughness produced by the

beating of sinusoids. This phenomenon, called sensory dissonance, is heard when tones interact within an auditory critical band (Plomp & Levelt, 1965), and the interaction of pure tone components correctly predicts ratings of consonance for pairs of complex tones (Kameoka & Kuriyagawa, 1969). However, sensory consonance does not fully explain the perception of musical consonance. For one thing, the sensory dissonance phenomenon applies to isolated clusters of simultaneously sounded tones, whereas musical consonance and dissonance are intrinsically dynamic: “... a dissonance is that which requires resolution to a consonance” (Dowling, 1978).

**Table 1** Tuning systems: Tone frequencies are chosen to divide the octave into (approximately) equal steps. Just intonation uses small integer ratios; equal temperament provides an approximation. The major scale (white notes) is one diatonic subset used to create melodies.

|   | Just Intonation | JI Decimal Equivalents | Equal Temperament |
|---|-----------------|------------------------|-------------------|
|   | 1:1             | 1.000                  | 1.000             |
| ■ | 16:15           | 1.067                  | 1.060             |
|   | 9:8             | 1.125                  | 1.123             |
| ■ | 6:5             | 1.200                  | 1.189             |
|   | 5:4             | 1.250                  | 1.260             |
|   | 4:3             | 1.333                  | 1.335             |
| ■ | 45:32           | 1.406                  | 1.414             |
|   | 3:2             | 1.500                  | 1.498             |
| ■ | 8:5             | 1.600                  | 1.587             |
|   | 5:3             | 1.667                  | 1.682             |
| ■ | 7:4             | 1.750                  | 1.782             |
|   | 15:8            | 1.875                  | 1.888             |
|   | 2:1             | 2.000                  | 2.000             |

In tonal music, some tones are perceived as more stable than others (Krumhansl, 1990; Lerdahl, 2001). More stable tones function as points of relative rest, whereas less stable tones tend to resolve to more stable ones. Theoretically, the stability of each pitch class relative to the other pitch classes is often described as a hierarchy (e.g., Lerdahl, 2001; see Figure 1A). Tones that are more stable occupy higher levels in the hierarchy; tones on a lower level are heard in relation to tones on the adjacent higher level. Krumhansl and Kessler (Krumhansl & Kessler, 1982) asked listeners to rate how well individual pitches fit within a tonal context. Such experiments provide profiles that quantify the stability of each tone within a musical key (see Figure 4, below). When applied to Western tonal contexts, the measured hierarchies are found to be consistent with music-theoretic accounts (Krumhansl, 1990; Krumhansl & Kessler, 1982). Moreover, stability measures correlate well with empirical frequencies of occurrence of tones in tonal



**Fig. 1** Music theoretic depictions of tonal stability and attraction. A) The basic pitch class space for a tonic chord in the key of C Major. Attractions to superordinate neighbors are shown separately for each level of the basic space: chromatic (B), diatonic (C) and triadic (D). Arrows indicate goals and sizes of attractions. Reproduced from Lerdhal, 2001.

songs (e.g. Knopoff & Hutchinson, 1983). The strong expectancy for less stable tones to resolve to more stable ones is called *attraction*. Some theorists have described tonal attraction by analogy to physical forces, such as gravity and inertia (Larson, 2004); others link it to the resolution of musical dissonance (Bharucha, 1984). Transition probabilities in databases of folksongs (e.g., Eerola & Toiviainen, 2004), as well as listener's expectations about the completion of musical sequences (Bharucha & Stoeckig, 1986; Cuddy & Lunney, 1995; Larson, 2004), all show a strong influence of tonal attraction. Lerdahl's tonal pitch space summarizes such findings, quantifying net resultant attraction of pitches at one level of a stability hierarchy toward pitches at the next level of stability (Lerdahl, 2001) as shown in Figure 1, Panels B, C and D.

The tuning systems of the world's largest musical cultures, Western, Chinese, Indian, and Arab-Persian, are based on small integer-ratio relationships (Burns 1999). However, each tuning system is different, and this has led to the notion that frequency relationships do not matter in high level music cognition; rather, auditory transduction of musical tones results in abstract symbols, as in language. If this were true, stability and attraction relationships would have to be learned solely based on the frequency-of-occurrence statistics of tonal music (e.g. Krumhansl, 1990; Tillmann, Bharucha, & Bigand, 2000). However, this general approach does not explain why the statistics of tonal music would develop as they have; it assumes that statistical properties are given *a priori*. It also does not explain the significance of different tuning systems; it would make the same predictions given two different sets of tone frequencies with the same statistical relationships. The main hypothesis of this chapter is that nonlinear resonance in the central nervous system underlies the perception of tonality. Universal characteristics of nonlinear resonance predict that the perception of stability, tonal attraction, preference for small integer ratios, and perceptual categorization, are intrinsic to the dynamics of neural processing. Hebbian learning, which requires only passive exposure, accommodates the heterogeneity of tonal systems while the physics of nonlinear resonance predicts constraints on what can be learned.

### 3 A Dynamical Systems Approach

Interaction of excitatory and inhibitory neurons, illustrated schematically in Figure 2A, can give rise to neural oscillation. The current discussion follows the analysis of neural oscillation by Hoppenstadt and Izhikevich (Hoppensteadt & Izhikevich, 1996a, 1996b, 1997), extending the analysis to gradient frequency networks of neural oscillators driven by acoustic stimuli (Large, Almonte, & Velasco, 2010). Neural oscillation can be modeled theoretically using Equation 1 (Wilson & Cowan, 1973), which consists of two variables, describing the activity of excitatory ( $x$ ) and inhibitory ( $y$ ) neural populations:

$$\begin{aligned}\dot{x} &= -x + S(\rho_x + ax - by) \\ \dot{y} &= -y + S(\rho_y + cx - dy)\end{aligned}\tag{1}$$

The overdot represents differentiation with respect to time,  $\dot{x} = dx/dt$ , and  $S$  is a sigmoid function.  $\rho_x$  and  $\rho_y$  are bifurcation parameters. Gradient frequency neural oscillator networks – under the influence of external input – can be written as

$$\tau_i \dot{x}_i = f_i(x_i, \lambda) + \varepsilon g_i(x, \lambda, \rho, \varepsilon) \quad (2)$$

where  $x$  is the 2-dimensional state vector for an oscillator,  $\lambda$  is a vector of parameters,  $\rho$  represents external input, and  $\varepsilon$  is the strength of the coupling nonlinearity (Hoppenstadt & Izhikevich, 1996a). The parameters are chosen such that  $\tau_i = 1/f_i$ , where  $f_i$  is natural frequency in Hz. The behavior of an oscillator under the influence of external input (Figure 2A) can be understood in detail by rewriting it in normal form (Wiggins, 1990; see Equation 3). The analysis involves a coordinate transformation, followed by Taylor expansion of the nonlinearities, truncating at some point to eliminate high order terms (abbreviated as *h.o.t.* in Equation 3). This results in  $z$ , a new, complex valued state variable, resulting from the coordinate transformation, and complex-valued parameters  $a$  and  $b$  ( $a = \alpha + i\omega$ , where  $\omega$  is the radian frequency,  $2\pi$ , and  $b = \beta + i\delta$ ), which can be related to the parameters of the original system.  $x(t)$  represents external input, either from another oscillator or from an acoustic signal.

$$\dot{z} = z(a + b|z|^2) + c x(t) + \text{h.o.t.} \quad (3)$$

The parameter,  $c$ , determines input connectivity and here we assume it to be real, although in general it could be complex (Hoppensteadt & Izhikevich, 1996a). The external stimulus  $x(t) = A(t)e^{i\theta(t)}$  has frequency  $\omega_0$ .

The behavior of this system can be understood by transforming it to polar coordinates using the relation  $z = r(t)e^{i\phi(t)}$ . This allows the independent study of amplitude,  $r$ , and phase,  $\phi$ , dynamics of the oscillator.

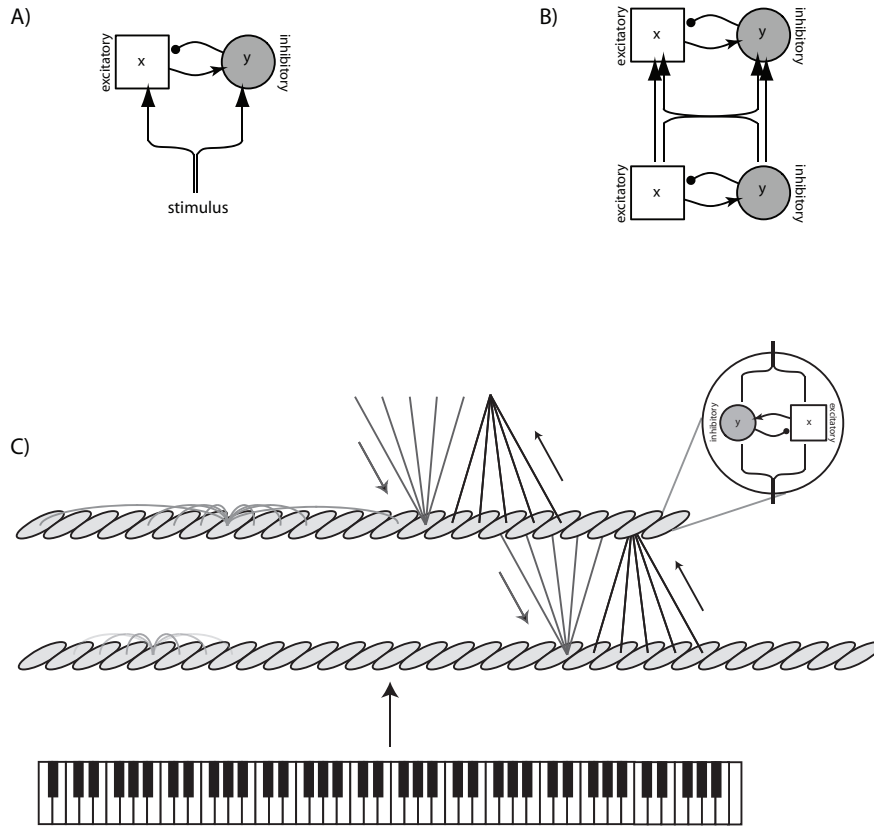
$$\begin{aligned} \dot{r} &= r(\alpha + \beta r^2) + cA \cos(\theta - \phi) + \text{h.o.t.} \\ \dot{\phi} &= \omega + \delta r^2 + c \frac{A}{r} \sin(\theta - \phi) + \text{h.o.t.} \end{aligned} \quad (4)$$

Interactions between oscillators of different frequencies are found in the higher order terms. To understand these interactions, we expand out the higher order terms, keeping in mind that any resonant relationship among oscillator frequencies of the form

$$m_1 \omega_1 + \dots + m_n \omega_n + m_R \omega_R = 0 \quad (5)$$

is a resonance among eigenvalues of the uncoupled system and thus cannot be eliminated (Hoppenstadt & Izhikevich, 1997). This includes harmonics, subharmonics, and summation and difference tones of various orders. For example, for a pair of oscillators with  $\omega_1 = 2\omega_2$

$$\begin{aligned} \dot{z}_1 &= z_1(a + b|z_1|^2) + \sqrt{\varepsilon} c_{12} z_2^2 + O(\varepsilon) \\ \dot{z}_2 &= z_2(a + b|z_2|^2) + \sqrt{\varepsilon} c_{21} z_1 \bar{z}_2 + O(\varepsilon) \end{aligned} \quad (6)$$



**Fig. 2** Nonlinear resonance. A) A neural oscillator consists of interacting excitatory and inhibitory neural populations. B) Four synapses are possible from one neural oscillator to another. Changes in synaptic efficacy affect both the strength and the phase of oscillators’ interaction and can be modified via Hebbian learning. C) A multi-layered, gradient frequency nonlinear oscillator network for responding to auditory stimulation.

and for a pair of oscillators with  $\omega_1 = 3\omega_2$

$$\begin{aligned} \dot{z}_1 &= z_1(a_1 + b_1|z_1|^2 + \epsilon d_1|z_1|^4) + \epsilon c_{12}z_2^3 + O(\epsilon\sqrt{\epsilon}) \\ \dot{z}_2 &= z_2(a_2 + b_2|z_2|^2 + \epsilon d_2|z_2|^4) + \epsilon c_{21}z_1^2\bar{z}_2 + O(\epsilon\sqrt{\epsilon}) \end{aligned} \quad (7)$$

Carrying the analysis out further leads to a canonical model for gradient-frequency networks of nonlinear neural oscillators (Large, et al., 2010):

$$\tau_i \dot{z}_i = z_i(a + b_1|z_i|^2 + \epsilon b_2|z_i|^4 + \dots) + (x + \sqrt{\epsilon}x^2 + \epsilon x^3 + \dots) \cdot (1 + \sqrt{\epsilon}\bar{z}_i + \epsilon\bar{z}_i^2 + \dots) \quad (8)$$

The simulations reported in the paper were based on numerical solution of this differential equation (see Large, Almonte & Velasco, 2010 for further details).

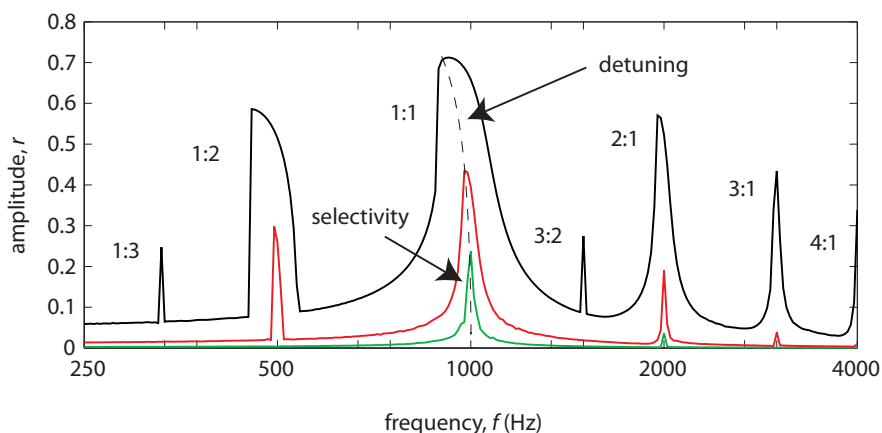
Next, consider analysis of sound by the auditory system. Acoustic signals stimulate the cochlea, which performs a nonlinear time-frequency transformation (e.g., Camalet, Duke, Julicher, & Prost, 1999; Ruggero, 1992). Central auditory networks in cochlear nucleus, inferior colliculus, thalamus, and primary auditory cortex phase-lock action potentials to both sinusoidal and amplitude modulated (AM) signal features, further transforming the stimulus (Langner, 1992). Phase-locking deteriorates at higher-frequencies as the auditory pathway is ascended. The role of neural inhibition in the central auditory system is not yet fully understood. However, phase-locked inhibition exists in many auditory nuclei and plays a role in the temporal properties of neural responses (Grothe, 2003; Grothe & Klump, 2000) that could be consistent with nonlinear resonance. A simple model consistent with the known facts and the hypothesis of nonlinear resonance in the auditory system is illustrated in Figure 2C. It is based on networks of neural oscillators, in which each is tuned to a distinct natural frequency, or *eigenfrequency*, following a frequency gradient, similar in concept to a bank of bandpass filters. Within this framework, the input,  $x$ , to a gradient-frequency network of neural oscillators, would consist of afferent, internal and efferent input. For a network responding directly to an auditory stimulus, the afferent input would correspond to a sound. Despite the fact that the physiology of neural oscillation of oscillators can vary greatly, all nonlinear oscillators share many universal properties, providing certain degrees of freedom and also significant constraints, discussed next.

#### 4 Predicting Tonality

**Nonlinear resonance.** Nonlinear oscillators possess a filtering behavior, responding maximally to stimuli near their own eigenfrequency. This is sometimes referred to as frequency selective amplification, due to extreme sensitivity to low amplitude stimuli. The first simulation (Figure 3) modeled frequency transformation of a sinusoidal stimulus by a single layer network of critical nonlinear oscillators (Equation 8), to demonstrate some basic properties. For this simulation the parameter values  $\alpha = 0$ ;  $\omega = 2\pi$ ;  $\beta_1 = \beta_2 = \beta_n = -1$ ;  $\delta_1 = 1$ ;  $\varepsilon = 1.0$  were used, and  $\tau_i = 1/f_i$ , where  $f_i$  is the natural frequency of each oscillator in Hz. All other parameters were set to zero. The frequencies of the network were distributed along a logarithmic frequency gradient, with 120 oscillators per octave, spanning four octaves. The choice of  $\alpha = 0$ ;  $\beta_n < 0$  makes this a critical nonlinear oscillator, network similar to models that have been proposed for cochlear hair cell responses (Camalet, et al., 1999). No internal network connectivity was used in this simulation.

Figure 3 shows how a nonlinear oscillator bank responds as stimulus intensity varies. At low levels, high frequency selectivity is achieved. As stimulus amplitude increases, frequency selectivity deteriorates due to nonlinear excitation. As a nonlinear oscillator responds to a stimulus near its eigenfrequency, frequency entrains to that of the stimulating waveform, such that instantaneous frequency





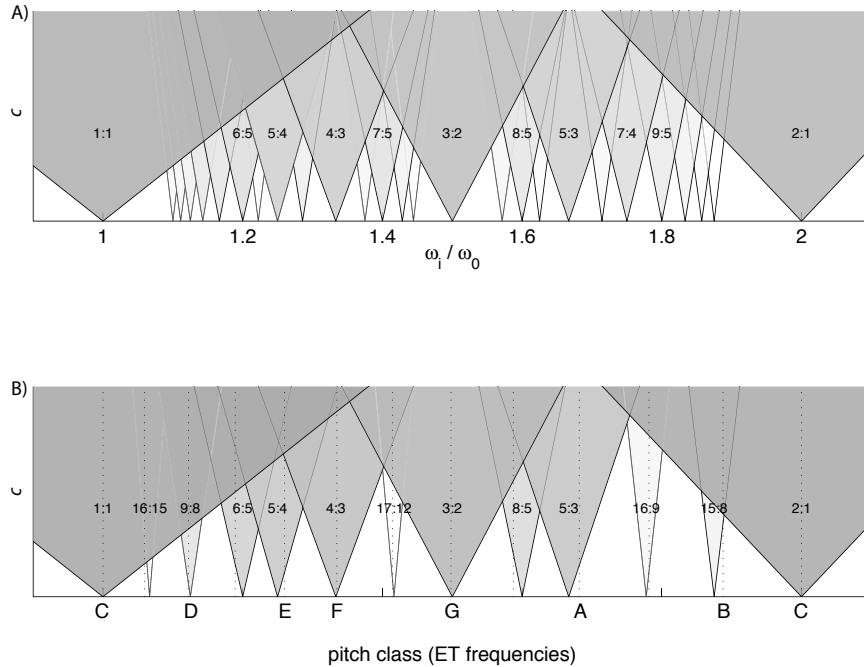
**Fig. 3** Response amplitudes,  $r$ , of a gradient frequency nonlinear oscillator array (frequencies  $250 \leq f \leq 4000$  Hz) to a sinusoid (frequency  $f_0 = 1000$  Hz) at three different stimulus amplitudes.

comes to match stimulus frequency. A nonlinear oscillator array also responds at frequencies that are not physically present in the acoustic stimulus. At low stimulus intensities, higher-order resonances are small; they increase with increasing stimulus intensity. The strongest response is observed at the stimulus frequency, and additional responses are observed at harmonics and subharmonics of the sinusoidal stimulus. The second sub- and super-harmonics (1:2 and 2:1) are the strongest resonances, predicting the universality of the octave. Additionally, the response frequency of the oscillator depends on the amplitude of the resonance, i.e., frequency changes as amplitude increases. Such *frequency detuning* can be seen in Figure 3 as a bend in the resonance curve as stimulus intensity, and therefore, response amplitude increases. Frequency detuning could predict systematic departures from ET (and JI), which are commonly observed in category identification experiments (Burns, 1999), including octave stretch, as discussed below.

**Natural resonances.** For multi-frequency stimulation, the response of an oscillator network may include harmonics, subharmonics, integer ratios, and summation and difference tones, some of which are illustrated in Figure 3. To explore the natural resonances in a gradient frequency network a bifurcation analysis was used. Analysis of the higher-order resonances was based on the phase equations:

$$\begin{aligned}\dot{\phi}_1 &= \omega_1 + c_{12} \mathcal{E}^{(k+m-2)/2} \sin(k\phi_2 - m\phi_1) \\ \dot{\phi}_2 &= \omega_2 + c_{21} \mathcal{E}^{(k+m-2)/2} \sin(m\phi_1 - k\phi_2)\end{aligned}\quad (9)$$

where the frequency ratio is  $k:m$  and the effects of amplitude are neglected. Here  $c_{ij} \mathcal{E}^{(k+m-2)/2}$  is the strength, or relative stability, of the  $k:m$  resonance, where  $c_{ij}$



**Fig. 4** Resonance regions. A) Bifurcation diagram showing natural resonances in a gradient frequency nonlinear oscillator array as a function of connection strength,  $C$ , and frequency ratio  $\omega_i/\omega_0$ . An infinite number of resonances are possible on this interval; the analysis considered the unison (1:1), the octave (2:1) and the twenty-five most stable resonances in between. B) Bifurcation diagram for a nonlinear oscillator network with internal connectivity reflecting an equal tempered chromatic scale. Internal connectivity can be learned via a Hebbian rule given passive exposure to melodies. Resonance regions whose center frequencies match ET ratios closely enough are predicted to be learned.

is coupling strength, a parameter that would be learned, and  $\mathcal{E}$  is the degree of nonlinearity in the coupling. This bifurcation analysis (Figure 4) assumes  $\mathcal{E}=1$  (maximal nonlinearity) and plots resonance regions as a function of coupling strength on the vertical axis and relative frequency,  $\omega_i/\omega_0$ , on the horizontal axis. The phase equations (9) were used to derive the boundaries of the resonance regions, or Arnold tongues according to

$$\frac{k}{m} \pm c \frac{m+k}{mk}.$$

The analysis varied oscillator frequency and coupling strength, assuming equal stimulation to each oscillator at a fixed frequency (the tonic,  $\omega_0$ ). The result depicts the long-term stability of various resonances in the network, displayed as a bifurcation diagram called Arnold tongues (Figure 4A). It predicts how different pools of neural oscillators will respond by showing the boundaries of resonance

neighborhoods as a function of coupling strength and frequency ratio. The oscillators in each resonance region frequency-lock at a specific ratio with the stimulus.

An infinite number of resonances exist on the interval between 1:1 and 2:1 and are found at integer ratios; smaller integer ratios are more stable and therefore more likely to be observed in the limit. This analysis displayed the 25 largest resonance regions on this interval, to provide a picture of the “natural” resonances in such a network. From the point of view of a gradient-frequency oscillator network, the Arnold tongues can be thought of as displaying the resonance for each oscillator in the network interacting with an oscillator outside the network (for example, an oscillator providing afferent stimulation) whose frequency corresponds to the tonic (1:1). The analysis considers pairwise interactions only; it neglects interactions between the oscillators within the gradient frequency network. Thus, it provides a somewhat simplified picture of network behavior, but one that is highly informative.

Nonlinear resonance predicts a generalized preference for small integer ratios. This prediction does not correspond to any specific musical scale; rather, natural resonances predict constraints on which frequency relationships can be learned. When stimulus frequency does not form a precise integer ratio with the eigenfrequency of an oscillator, resonance is still possible, provided that coupling is strong enough. Resonances affect not only oscillators with precise integer ratios; they also establish patterns of resonant neighborhoods.

**Learning.** Hebbian learning provides a theoretical basis for the acquisition of tonality relationships. Connections between oscillators can be learned via a Hebbian rule (Hoppensteadt & Izhikevich, 1996b), providing a mechanism for synaptic plasticity wherein the repeated and persistent co-activation of a presynaptic cell and a postsynaptic cell lead to an increase in synaptic efficacy between them. Between two neural oscillators four synapses are possible (Figure 2B, above), providing both a strength and a natural phase for the connection between neural oscillators (Hoppensteadt & Izhikevich, 1996a). Hebbian learning rules have been proposed for neural oscillators and the single-frequency case has been studied in some detail (Hoppensteadt & Izhikevich, 1996b). For the single-frequency case, the Hebbian learning rule can be written as follows:

$$\dot{c}_{ij} = -\delta c_{ij} + k_{ij} z_i \bar{z}_j \quad (10)$$

This model can learn both amplitude and phase information for two oscillators with a frequency ratio near 1:1. In the current study, learning of higher-order resonances is also of interest. The following generalization of the above learning rule enables the study of learning for higher order resonant relationships

$$\dot{c}_{ij} = -\delta c_{ij} + k_{ij} (z_i + \sqrt{\epsilon} z_i^2 + \epsilon z_i^3 + \dots) \cdot (\bar{z}_j + \sqrt{\epsilon} \bar{z}_j^2 + \epsilon \bar{z}_j^3 + \dots). \quad (11)$$

Coupling strength,  $c_{ij}$ , is the parameter that would be altered by learning. Due to computational complexity, extensive simulation of learning on melodies has not yet been carried out. However, analysis of multi-frequency Hebbian learning

(Eq. 11) shows that connections to near-resonant frequencies, such as integer ratios, can be learned in a gradient frequency network. Assuming stimulation with melodies using ET tone frequencies, connections would be learned between an oscillator at the frequency of the tonic (1:1) and the most stable resonances that approximate the ET tone frequencies closely. Thus, a second bifurcation analysis was performed, in which the  $k$  and  $m$  parameters were chosen as the largest resonance region (smallest integer ratio) that approximated the ET tone frequency to within 1%<sup>1</sup>. The result of this analysis is shown in Figure 4B. Learning would likely result in different coupling strengths for each resonance, therefore analysis shows each resonance region for a range of coupling strengths.

This analysis predicts that, as Western melodies are heard, the network would learn the most stable attractors whose center frequencies closely approximate the ET chromatic frequencies. Hebbian synaptic modification would effectively prune some resonances, while retaining others. The resulting resonances reflect the chromatic scale as shown in Figure 4B. In principle a similar learning analysis could be performed for any tuning system, such as gamelan, whose frequency ratios differ significantly from 12-tone ET.

**Perceptual Categorization.** In Figure 4B, the center frequencies of the resonances do not precisely match ET frequencies; however, as connection strength increases, larger regions of the network resonate, emanating from integer ratios, and encompassing ET ratios. Such regions predict perceptual categorization of musical intervals. Perceptual categorization and discrimination experiments reveal that musicians show categorical perception of melodic intervals (Burns & Campbell, 1994), and although such experiments are more difficult with non-musicians, Smith et al. (Smith, Nelson, Grohskopf, & Appleton, 1994) demonstrated that nonmusicians also perceive pitch categories. Dependence of frequency on amplitude further predicts that perceptual categories might not be precisely centered on integer ratios. In interval identification experiments, mean frequency deviates systematically from ET, although not always in the direction predicted by JI. Musicians prefer flatter small intervals and sharper large intervals (Burns, 1999). In fact, in many tuning systems (including the piano) octaves are stretched, i.e., tuned slightly larger than 2:1. In performance on instruments without fixed tuning (e.g., the violin, or the human voice), mean frequency also deviates systematically, similarly to perceptual categorizations (Loosen, 1993). More importantly, frequency variability is quite large in music performance, even during the “steady state” portions of tones, emphasizing the importance of pitch categorization in the perception of tonality.

**Attraction.** The theory also makes predictions about tonal attraction. In areas where resonance regions overlap (e.g., Figure 4B), more stable resonances overpower less stable ones, such that the instantaneous frequency of the population in the overlap region is attracted toward the frequency of the more stable resonance. To understand the implications for tonal attraction, a nonlinear

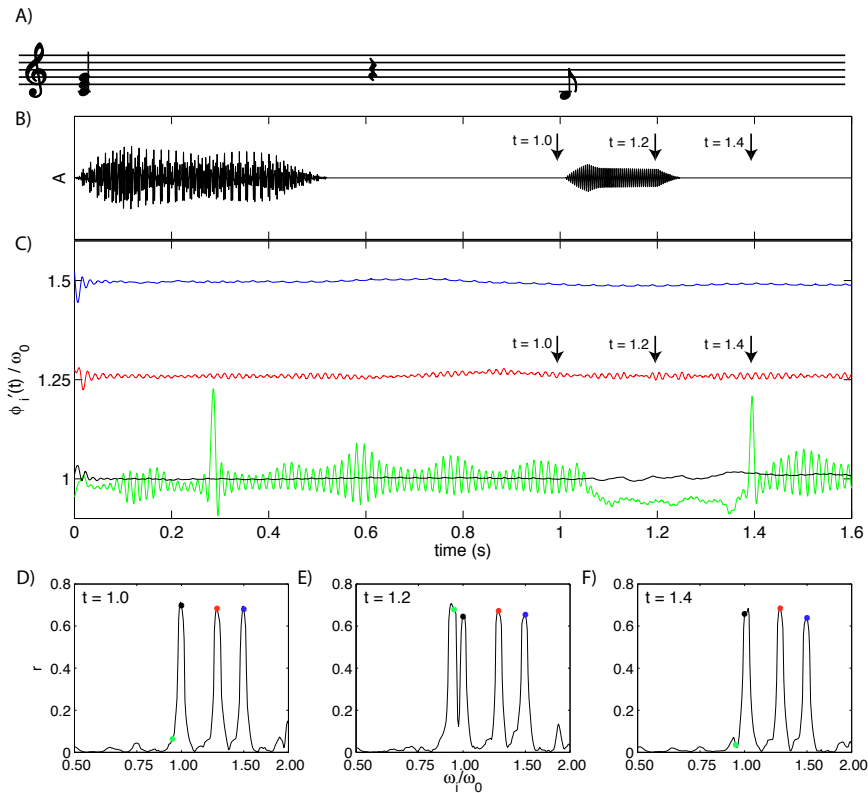
---

<sup>1</sup> Operationalization of “close” frequency as 1% is somewhat arbitrary, and different choices result in different resonance regions for the weaker resonances, changing the predictions slightly, but not altering the basic results.

oscillator network was simulated using internal connectivity among oscillators that reflected the structure of the major scale. The simulation was based on a two-layer network, modeling a cochlear transformation followed by a neural transformation. This minimal model provided a simple example in which tonal attraction can be observed. The first layer parameters were  $\alpha = -0.01$ ;  $\omega = 2\pi$ ;  $\beta_1 = \beta_2 = \beta_n = -1$ ;  $\delta_1 = \delta_2 = \delta_n = 0$ ;  $\varepsilon = 0.1$  and  $\tau_i = 1/f_i$ , where  $f_i$  is the frequency of each oscillator in  $Hz$ , similar to the first simulation, but without frequency detuning. The frequencies of the network were distributed along a logarithmic frequency gradient, with 120 oscillators per octave, spanning two octaves. A Gaussian kernel modeled local basilar membrane coupling (cf. Kern & Stoop, 2003).

The parameters of the second network were set to  $\tau_i = 1/f_i$  where  $f_i$  is the frequency of each oscillator in  $Hz$ , and  $\alpha = -0.4$ ;  $\beta_1 = 1.2$ ;  $\beta_2 = \beta_n = -1$ ;  $\delta_1 = -0.01$ ;  $\varepsilon = 0.75$ . All other parameters were set to zero. Again, the frequencies of the network were distributed along a logarithmic frequency gradient, with 120 oscillators per octave, spanning two octaves. The center frequency of both networks was chosen to match middle C. Afferent connectivity from the cochlear network was one-to-one, with oscillators of the cochlear network stimulating oscillators of the neural network according to frequency. In the second network internal connectivity was constructed to reflect learning of the ET scale, as described above. Each oscillator was connected to the others that are nearby in frequency, as well as to those whose eigenfrequencies approximated the frequency ratios of the scale. The main feature of this simulation is that the parameters of the second (neural) network are chosen to be near a degenerate Hopf bifurcation, also known as a Bautin bifurcation (Guckenheimer & Kuznetsov, 2007). For this reason, the nonlinear coupling allows amplitude peaks to self-stabilize at the frequencies of stimulation, such that after the stimulus is removed, the peaks remain. This behavior is seen in Panels D, E, and F of Figure 5. Individual peaks interact with one another as well. Interactions in the gradient frequency network are complex, and a complete analysis is beyond the scope of this chapter. However, self-stabilizing amplitude peaks have been observed and analyzed for single-frequency oscillator networks with nonlinear coupling near a Bautin bifurcation (Drover & Ermentrout, 2003).

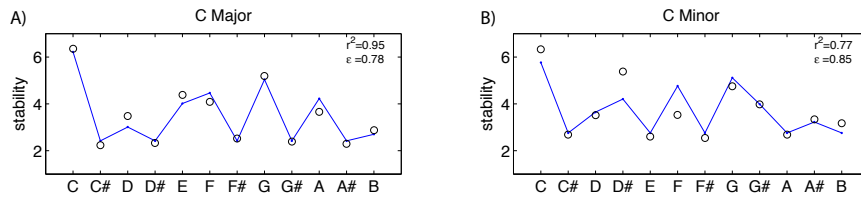
The network was stimulated with a C-major triad (the notes C, E, and G), which was followed after a delay by a leading tone (the note B; Figure 5A & B), and the instantaneous frequencies (Panel C) and amplitudes (Panels D, E, F) of the oscillators in the network were measured. The stimulus was prepared using Finale Notepad Plus 2005a, saved as a MIDI file, and rendered to a digital audio file as pure tones. After stimulation with the triad, a dynamic field self-stabilized to embody a set of resonant frequencies that was consistent with the prior stimulation, embodying a memory of the stimulus. Immediately before introduction of the leading tone ( $t = 1.0$ ), stable amplitude peaks corresponding to the populations of oscillators surrounding C, E and G are observed (Panel D), and the instantaneous frequencies of the three oscillators at C, E and G also appear stable (Panel C,  $t = 1.0$ ). After the leading tone is introduced, a corresponding



**Fig. 5** Attraction. A network with ET scale connectivity is stimulated with a C major triad (C, E, G; scale degrees 1, 3, 5), followed after a delay with a leading tone (B, scale degree 7). A) Musical notation. B) Stimulus waveform. C) Instantaneous frequency of four oscillators (out of 241 in the network) closest in eigenfrequency to scale degrees 7 (green), 1 (black), 3 (red), 5 (blue). After the triad is silenced, the dynamic field remains stable (D) with peaks corresponding to scale degrees 1, 3, and 5. When stimulation at scale degree 7 begins, a corresponding peak arises (E) and its frequency stabilizes (green). After stimulation ceases, the oscillator with eigenfrequency near scale degree 7 loses stability, the peak dies away (F), and its instantaneous frequency is attracted toward the tonic.

amplitude peak is observed (Panel E) and the instantaneous frequency is stabilized by the external stimulus (Panel C;  $t = 1.2$ , green). The important observation is that when the external stimulus is removed, this oscillation loses stability ( $t = 1.4$ ) and its frequency is attracted toward the note C, the tonic frequency. This network behavior predicts a physical correlate for the perceived attraction of the leading tone to the tonic, and in general for expectation of what should follow in a tonal context.

**Stability.** The next analysis asked whether theoretical stability of higher order resonances could predict perceived stability of tones (Krumhansl, 1990). For the



**Fig. 6** Comparison of theoretical stability predictions and human judgments of perceived stability for two Western modes. A) C major, B) C minor.

stability analysis, frequency ratios were fixed according to the previous analysis of learning (Figure 4B). It was further assumed that all the non-zero  $c$  were equal, effectively eliminating one free parameter (although in principle, the coupling strength,  $c$ , could be different for each resonance as a result of learning). Relative stability of each resonance was predicted by  $\varepsilon^{(k+m-1)/2}$ , where  $k$  and  $m$  are the numerator and denominator of the frequency ratio, respectively, and  $0 \leq \varepsilon \leq 1$  is a parameter that controls nonlinear gain (Hoppensteadt & Izhikevich, 1997). The analysis assumed that each tone listeners heard as part of the context sequence was stabilized in the network, and those that were not heard were not stabilized. This assumption reflects the behavior of the network simulated in the previous analysis. Thus, each context tone received a stability value of  $\varepsilon^{(k+m-1)/2}$ , and those that did not occur in the context sequence received a stability value of 0. For major and minor Western modes, the parameter  $\varepsilon$  was chosen to maximize the correlation ( $r^2$ , different from oscillator amplitude,  $r$ , used previously) with the stability ratings of listeners. This provides a single parameter fit to the perceptual data on stability, shown in Figure 6.

Predicted stability matched the perceptual judgments well (C major:  $r^2 = .95$ ,  $p < .0001$ ,  $\varepsilon = 0.78$ ; C minor:  $r^2 = .77$ ,  $p < .001$ ,  $\varepsilon = 0.85$ ), as shown in Figure 6. In other words, the theoretical stability of higher-order resonances of nonlinear oscillators predicts empirically measured tonal stability for major and minor tonal contexts. This result is significant because it does not depend on the statistics of tone sequences, but instead it predicts the statistics of tone sequences, which are known to be highly correlated with stability judgments (Krumhansl, 1990).

## 5 Discussion

While the properties of nonlinear resonance predict the main perceptual features of tonality well, this theory makes two additional significant predictions: 1) that nonlinear resonance should be found in the human auditory system and 2) that animals with auditory systems similar to humans may be sensitive to tonal relationships. Recently, evidence has been found in support of both predictions.

Helmholtz (1863) described the cochlea as a time-frequency analysis mechanism that decomposes sounds into orthogonal frequency bands for further analysis by the central auditory nervous system. Von Bekesey (1960) observed

that human cadaver cochlear responses behave linearly over the range of physiologically relevant sound intensities. However, the weakest audible sounds impart energy per cycle no greater than that of thermal noise (Bailek, 1987), and the system operates over a range of intensities that spans at least 14 orders of magnitude. Moreover, laser-interferometric velocimetry performed on living, intact cochleae has revealed exquisitely sharp mechanical frequency tuning (Ruggero, 1992). Recent evidence, including the discovery of spontaneous otoacoustic emissions (Kemp, 1979), suggest that the sharp mechanical frequency tuning, exquisite sensitivity and operating range of the cochlea can be explained by critical nonlinear oscillations of hair cells (Choe, Magnasco, & Hudspeth, 1998). Thus, the cochlea performs an active, nonlinear transformation, using a network of locally coupled outer-hair cell oscillators, each tuned to a distinct eigenfrequency.

There is a growing body of evidence consistent with nonlinear oscillation in the central auditory system as well. In mammals, action potentials phase-lock to both fine time structure and temporal envelope modulations at many different levels in the central auditory system, including cochlear nucleus, superior olive, inferior colliculus (IC), thalamus and A1 (Langner, 1992), and recent evidence points to a key role for synaptic inhibition in maintaining central temporal representations. Hyperpolarizing inhibition is phase-locked to the auditory stimulus and has been shown to adjust the temporal sensitivity of coincidence detector neurons (Grothe, 2003), while stable pitch representation in the IC may be the result of a synchronized inhibition originating from the ventral nucleus of the lateral lemniscus (Langner, 2007). Moreover, neurons in the IC of the gerbil have been observed to respond at harmonic ratios (3:2, 2:1, 5:2, etc.) with the temporal envelope of the stimulating waveform (Langner, 2007). *Pollimyrus*, an fish that lacks a sophisticated peripheral structure for mechanical frequency analysis, has modulation-rate selective cells in the auditory midbrain that receive both excitatory and inhibitory input, and are well described as nonlinear oscillators (Large & Crawford, 2002). Finally, residue pitch shift – a central auditory phenomenon – is consistent with 3-frequency resonances of nonlinear oscillators, making nonlinear resonance viable as a candidate for the neural mechanism of pitch perception in humans (Cartwright, Gonzalez, & Piro, 1999).

If key aspects of tonality depend directly on auditory physiology, one would predict that non-human animals might be sensitive to certain tonal relationships. Wright et al. tested two rhesus monkeys for octave generalization in eight experiments by transposing 6- and 7-note musical passages by an octave and requiring *same* or *different* judgments (Wright, Rivera, Hulse, Shyan, & Neiworth, 2000). The monkeys showed complete octave generalization to childhood songs (e.g., "Happy Birthday") and tonal melodies (from a tonality algorithm). They showed no octave generalization to random-synthetic melodies, atonal melodies, or individual notes. Takeuchi's Maximum Key Profile Correlation measure of tonality, based on human tonality judgments (Takeuchi, 1994) accounted for 94 percent of the variance in the monkey data. These results provide evidence that tonal melodies retain their identity when transposed with whole octaves, as they do for humans. Adult listeners can recognize transpositions of tonal but not atonal



melodies (Dowling & Fujitani, 1971). Preverbal infants can recognize transposed tonal melodies (Trehub, Morrongiello, & Thorpoe, 1985), and melody identification is nearly perfect for octave (2:1 ratio) transpositions, even for novel melodies (Massaro, Kallman, & Kelly, 1980), as is also the case for macaques.

Zuckermandl (1956) argued that the dynamic quality of musical tones "...makes melodies out of successions of tones and music of acoustical phenomena." The current approach predicts that the perceived dynamics of tonal organization arise from the physics of nonlinear resonance. Thus, nonlinear resonance may provide the neural substrate for a substantive musical universal. In some ways, this idea is similar to the concept of universal grammar in linguistics (Prince & Smolensky, 1997). However, in the case of music, these perceptual universals are predicted by universal properties of nonlinear resonance, offering a direct link to neurophysiology. Learning would alter connectivity to establish different resonances, and different tonal relationships. Higher-order resonances may give rise to dynamic tonal fields in the central nervous system, with localized patterns of activation self-stabilizing to embody the musical system of the listener's culture. Musical melodies would be perceived in relation to the tonal field, creating a dynamic context within which perception of tone sequences takes place.

**Acknowledgments.** This research was supported by NSF CAREER Award BCS-0094229, and NSF grant BCS-1027761. The author would like to thank Justin London, Carol Krumhansl, John Iversen, Frank Hoppensteadt, and Eugene Izhikevich for insightful comments on an earlier draft. Thanks also to Felix Almonte, Marc Velasco and Nicole Flaig for many interesting discussions, and for their invaluable help with this manuscript.

## References

- Bailek, W.: Physical limits to sensation and perception. *Annual Review of Biophysics and Biophysical Chemistry* 16, 455–478 (1987)
- Bharucha, J.J.: Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology* 16, 485–518 (1984)
- Bharucha, J.J., Stoeckig, K.: Reaction-Time and Musical Expectancy - Priming of Chords. *Journal of Experimental Psychology-Human Perception and Performance* 12, 403–410 (1986)
- Burns, E.M.: Intervals, scales, and tuning. In: Deutsch, D. (ed.) *The Psychology of Music*, pp. 215–264. Academic Press, San Diego (1999)
- Burns, E.M., Campbell, S.L.: Frequency and frequency ratio resolution by possessors of relative and absolute pitch: Examples of categorical perception? *Journal of the Acoustical Society of America* 96, 2704–2719 (1994)
- Camalet, S., Duke, T., Julicher, F., Prost, J.: Auditory sensitivity provided by self tuned critical oscillations of hair cells. *Proceedings of the National Academy of Sciences* 97, 3183–3188 (1999)
- Cartwright, J.H.E., Gonzalez, D.L., Piro, O.: Nonlinear Dynamics of the Perceived Pitch of Complex Sounds. *Physical Review Letters* 82, 5389–5392 (1999)
- Choe, Y., Magnasco, M.O., Hudspeth, A.J.: A model for amplification of hair-bundle motion by cyclical binding of Ca<sup>2+</sup> to mechano-electrical-transduction channels. *Proceedings of the National Academy of Sciences* 95, 15321–15336 (1998)

- Cuddy, L.L., Lunney, C.A.: Expectancies generated by melodic intervals: Perceptual judgements of melodic continuity. *Perception & Psychophysics* 57, 451–462 (1995)
- Dowling, W.J.: Scale and contour: Two components of a theory of memory for melodies. *Psychological Review* 85, 341–354 (1978)
- Dowling, W.J., Fujitani, D.S.: Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America* 49, 524–531 (1971)
- Drover, J.D., Ermentrout, B.: Nonlinear coupling near a degenerate Hopf (Bautin) Bifurcation. *SIAM Journal On Applied Mathematics* 63, 1627–1647 (2003)
- Eerola, T., Toivainen, P.: Finnish Folksong Database (2004), [https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/collection\\_download](https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/collection_download) (Retrieved October 13, 2010)
- Grothe, B.: New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience* 4, 540–550 (2003)
- Grothe, B., Klump, G.M.: Temporal processing in sensory systems. *Current Opinion in Neurobiology* 10, 467–473 (2000)
- Guckenheimer, J., Kuznetsov, Y.A.: Bautin bifurcation. *Scholarpedia* 2, 1853 (2007)
- Helmholtz, H.L.F.: On the sensations of tone as a physiological basis for the theory of music. Dover Publications, New York (1863)
- Hoppensteadt, F.C., Izhikevich, E.M.: Synaptic organizations and dynamical properties of weakly connected neural oscillators I: Analysis of a canonical model. *Biological Cybernetics* 75, 117–127 (1996a)
- Hoppensteadt, F.C., Izhikevich, E.M.: Synaptic organizations and dynamical properties of weakly connected neural oscillators II: Learning phase information. *Biological Cybernetics* 75, 126–135 (1996b)
- Hoppensteadt, F.C., Izhikevich, E.M.: *Weakly Connected Neural Networks*. Springer, Heidelberg (1997)
- Kameoka, A., Kuriyagawa, M.: Consonance theory part II: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America* 45, 1460–1471 (1969)
- Kemp, D.T.: Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. *European Archives of Oto-Rhino-Laryngology* 224, 370 (1979)
- Kern, A., Stoop, R.: Essential role of couplings between hearing nonlinearities. *Physical Review Letters* 91, 128101–128104 (2003)
- Knopoff, L., Hutchinson, W.: Entropy as a measure of style: The influence off sample length. *Journal of Music Theory* 27, 75–97 (1983)
- Krumhansl, C.L.: *Cognitive foundations of musical pitch*. Oxford University Press, NY (1990)
- Krumhansl, C.L., Kessler, E.J.: Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review* 89(4), 334–368 (1982)
- Langner, G.: Periodicity coding in the auditory system. *Hearing Research* 60, 115–142 (1992)
- Langner, G.: Temporal processing of periodic signals in the auditory system: Neuronal representation of pitch, timbre, and harmonicity. *Z Audiol.* 46, 21–80 (2007)
- Large, E.W., Almonte, F., Velasco, M.: A canonical model for gradient frequency neural networks. *Physica D: Nonlinear Phenomena* 239(12), 905–911 (2010)
- Large, E.W., Crawford, J.D.: Auditory temporal computation: Interval selectivity based on post-inhibitory rebound. *Journal of Computational Neuroscience* 13, 125–142 (2002)

- Larson, S.: Musical Forces and Melodic Expectations: Comparing Computer Models and Experimental Results. *Music Perception* 21, 457–498 (2004a)
- Lerdahl, F.: *Tonal Pitch Space*. Oxford University Press, New York (2001)
- Loosen, F.: Intonation of solo violin performance with reference to equally tempered, Pythagorean and just intonations. *Journal of the Acoustical Society of America* 93, 25–539 (1993)
- Massaro, D.W., Kallman, H.J., Kelly, J.L.: The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology: Human Learning and Memory* 6, 77–90 (1980)
- Plomp, R., Levelt, W.J.M.: Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America* 38, 548–560 (1965)
- Prince, A., Smolensky, P.: Optimality: From Neural Networks to Universal Grammar. *Science* 275, 1604–1610 (1997)
- Ruggero, M.A.: Responses to sound of the basilar membrane of the mammalian cochlea. *Current Opinion in Neurobiology* 2, 449–456 (1992)
- Smith, J.D., Nelson, D.G., Grohskopf, L.A., Appleton, T.: What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition* 52, 23–54 (1994)
- Tillmann, B., Bharucha, J.J., Bigand, E.: Implicit learning of tonality: A self-organizing approach. *Psychological Review* 107, 885–913 (2000)
- Trehub, S.E., Morrongiello, B.A., Thorpoe, L.A.: Children's perception of familiar melodies: The role of interval contour and key. *Psychomusicology* 5, 39–48 (1985)
- von Békésy, G.: *Experiments in Hearing*. McGraw-Hill Book Co., New York (1960)
- Wiggins, S.: *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer, New York (1990)
- Wilson, H.R., Cowan, J.D.: A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80 (1973)
- Wright, A.A., Rivera, J.J., Hulse, S.H., Shyan, M., Neiwirth, J.J.: Music perception and octave generalization in Rhesus monkeys. *Journal of Experimental Psychology: General* 129, 291–307 (2000)
- Zuckermandl, V.: *Sound and Symbol: Music and the External World* (WR Trask, Trans.). Princeton University Press, Princeton (1956)