

# Illusions and *Ghost* Resonances: How We Could See What Isn't There.

Dante R. Chialvo

*Department of Physiology, University of California,  
Los Angeles, California 90024, USA  
Email: dchialvo@ucla.edu. Web: <http://www.bol.ucla.edu/~dchialvo>*

**Abstract.** To judge with certainty which information our senses are paying attention to is a hard problem, because often what we see, hear, or feel is not out there. In this paper, objective explanations are discussed for these tantalizing illusions demonstrating how relatively simple nonlinear stochastic transformations make the solution possible.

## INTRODUCTION

Despite important advances, is still not well understood what a sensory neuron encodes. How to be certain that a given spike-train from a sensory neuron is encoding a particular aspect of the physical world? In this contribution we choose to discuss this aspect by way of an example in which the brain reads something that objectively isn't there ... The unsolved problem is how the brain extracts the pitch of complex sounds, and the solution proposed is a stochastic nonlinear mechanism which is biologically plausible. Work in this problem can be traced back to Pythagoras [1], who believed that music was more than entertainment. For him music was the expression of a divine principle *-harmonia-* that is able to bring order to chaos. Pitch is the subjective attribute of a sound sensation by which it may be ordered on a scale from high to low. In the simplest case, when the sound is composed of a single tone, like one made by a tuning fork, the pitch equals the frequency of the tone. However, natural sounds more often are complex, i.e., are comprised of the superposition of many tones. Pythagoras was interested in understanding the production of sounds with different pitch, and looked for the physical mechanism involved in the generation of complex tones. It is said that he was passing by a blacksmith's shop and heard the sound of hammers striking on a piece of iron. He noticed that, although the sounds made by the different hammers were all different, they were all in harmony with each other, except one. After toying around with other possibilities, he discovered that the pleasant harmony was related with certain ratios of the weight of the hammers, which were twelve, nine, eight and six pounds. He noted that striking together those with weights ratios of  $1/2$  - the six and twelve pound ones-, produced the intervals of an octave, using those with ratios of  $2/3$  (the eight and twelve pound) produced the sound of a fifth, and using a ratio of  $3/4$  -the nine and twelve pound hammers- produced the sound of a fourth. These were pleasant tones, derived all from the integers 1, 2, 3, and 4, which were

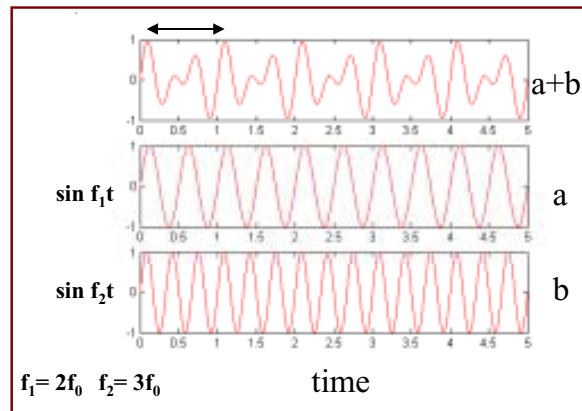
later replicated by Pythagoras using strings stretched by similar weights. At that moment in history, by his understanding of how to produce consistent musical consonant intervals, the world of sensations stopped being magic and became physics.

As often happens, one happy discovery reveals another puzzling mystery. Pythagoras' discovery opened the yet unresolved mystery of the location of the "consonance detector", i.e., which brain area finds the harmony by computing these pleasant magic ratios? Furthermore, how is it that even untrained individuals, from disparate cultures, all share the same consonance appreciation? Finding universality in this problem, belongs to the yet unwritten pages in the theory of music, awaiting to be formally solved as any other complicated problem.

The scope of these pages is to discuss how the brain could extract the pitch of say just one of Pythagoras' hammers. Of course, it will also bear relevance into the larger problem of harmony, but we will leave that discussion out for the moment.

### Sometimes what is heard is not what is being listen to

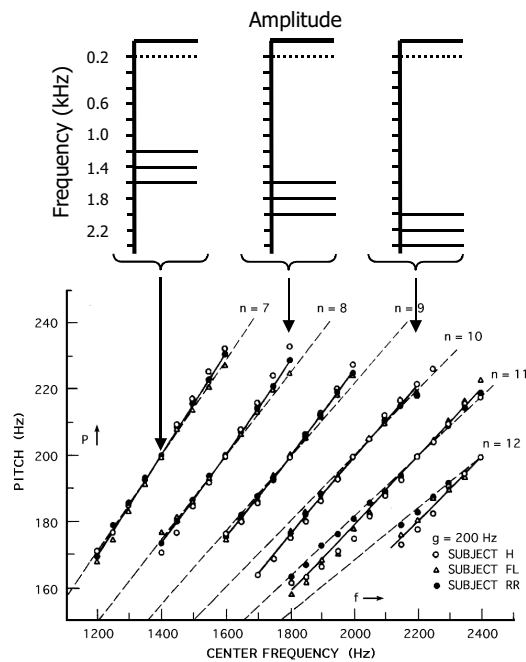
Natural sounds, as birdsongs, speech or those produced by a hammer striking on a piece of iron are complex, in the sense that they are formed by several discrete spectral lines. As it was said above, if pitch is a unitary attribute of a sensory experience, then how is it that a complex sound can be perceived as a unique entity? In other words, what is the neural process by which the brain fuses all the sound's frequency components into the perceived single pitch?



**Figure 1:** A complex tone (top trace) is constructed by adding two sinusoids (a and b) which are the first and second harmonics of  $f_0$ , respectively. When played, at relatively low volume, the resulting time series produces an  $f_0$  pitch (the double-headed arrow indicates the peaks spaced by  $1/f_0$ ).

Some inspiration came from simple inspection of complex tones. Consider the example in Fig. 1 where two sinusoids (with frequencies multiples of a third one) are added. It is well known, that when the resulting complex tone is played, the third frequency, though absent in the sound, is heard. The heard frequency corresponds to the intervals between the peaks of the time series, indicated by the double-headed

arrow in the figure. Because the missing frequency is also, from the Fourier viewpoint, the fundamental, this psychoacoustic phenomenon is often called the “*missing fundamental*” illusion [2-4]. Only if the tones are harmonic, like in the example of Fig. 1, (i.e., multiples of  $f_0$ ), the fundamental and the difference between the tones are the same. In any case the peaks are produced by constructive interference between the tones, a process first demonstrated by Thomas Young in 1800 when explained the origin of Newton’ rings [5] and set the basis of the wave nature of light: two periodic processes will add up around their common subharmonic frequency.



**Figure 2:** Results from Schouten’s pitch shift experiments [4] using three-tones sounds with center frequency from 1.2 up to 2.4 kHz. The top diagrams depict the line amplitude spectra for three examples of the complex sounds listened by the subjects. Symbols (circles and triangles) in the bottom graph indicate the pitch heard by each of the three subjects for each complex sound. (Redrawn and modified from ref. [4]) The dotted lines (in the original publication) were meant to convey visually the idea that a  $1/n$  function consistently underestimates the linear relationship between frequency and pitch shift. ( $g=200$  Hz is “ $f_0$ ” and “ $n$ ” is “ $k$ ” in our notation)

It was assumed, for a long time, that what it is heard is the *frequency difference* between the tones and that the brain could compute that simply by passing the complex tones through a simple nonlinearity. But Schouten et al [4], four decades ago, has clearly demonstrated that this is not so, by frequency shifting an initially harmonic complex stimulus. Specifically, the complex tones they used were composed by three frequencies:

$$f_1 = kf_0 + \Delta f, \quad f_2 = (k+1)f_0 + \Delta f, \quad f_3 = (k+2)f_0 + \Delta f \quad (1)$$

If the brain computes differences then, with this stimulus, one should hear a constant pitch since this experimental paradigm preserved the equal spacing between tones. Instead, as shown in Figure 2, perception was found to shift linearly with the shift  $\Delta f$ , in the region explored (between 1.2 and 2.4 kHz). Another important observation these authors made was to note that the slope of the linear relation between  $\Delta f$  and perceived shift decreases for increasing  $k$ . In addition, they noted that a  $1/n$  function was underestimating the experimental data (see Fig. 2). Finally, a qualitative, but not less important, observation was made: there was a peculiar ambiguity in the pitch judgment. As illustrated in the plot, there were always between two to four perceived pitches for any given sound.

Schouten et al. [4] findings became the litmus test for theories of pitch perception and despite an extensive literature on the subject no satisfactory quantitative theory exists to explain the origins of these scaling relationship between complex tones and perceived pitch.

The specific nature of the neural activity involved in pitch processing has long been the subject of discussions (for reviews see [6-8]). One view proposes that pitch of complex sounds is extracted by the auditory system by first deriving a spectral profile from a (tonotopic) frequency-specific auditory input, followed by some pattern-matching mechanisms [9-10]. Others have suggested mechanisms based on the timing of the auditory nerve fiber activity, irrespective of any frequency organization [11-14]. A serious drawback in deciding the merits of these theories is the absence of quantitative predictions and/or the ambiguity as to where in the brain a proposed computation is being done.

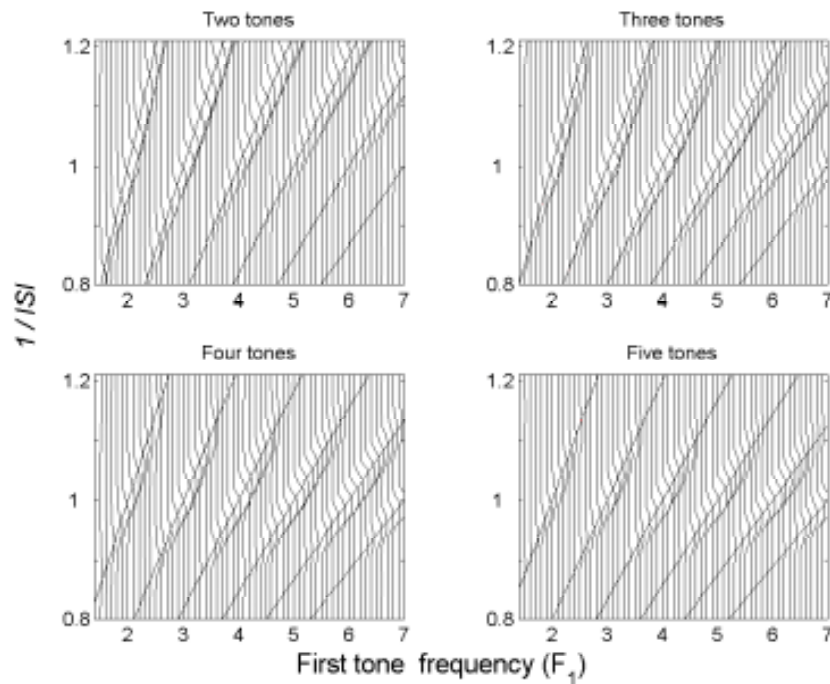
Departing from previous approaches, in the next section, a conjecture is discussed that interprets complex tones perception as the byproduct of a *linear* process of interference (or coincidence) of waves and a *nonlinear* detection of the interference peaks by a noisy threshold.

## Robust ghost stochastic resonance

The main idea is that the illusion we hear is uniquely determined by the peaks of the complex tone (as the example in the top time series of Figure 1). Although it is a trivial fact that neurons can easily detect those peaks, it was not anticipated how well the quantitative aspects of the psycho acoustic phenomenon can be reproduced by this simple assumption, as was discussed in ref [15-19]. In that work, a simple toy model of a sensory neuron driven by complex tones was considered. Briefly, the system was a non-dynamical threshold device [16] that reduces the most elementary neuron dynamics to a set of rules comparing the input signal  $x(t)$  with a threshold  $U_{th}$  like:  $x(t) > U_{th}$  or  $< U_{th}$ . Anytime  $x(t)$  crosses the fixed threshold  $U_{th}=1$ , a “spike” is emitted. The only quantities of relevance for the problem here are the inter-spike intervals (ISI). The model is driven with combinations of sinusoidal of various frequencies, i.e., simulating complex sounds used in experiments [4,13,14]:

$$x(t) = A(\sin(f_1)2\pi t + \sin(f_2)2\pi t + \dots \sin(f_n)2\pi t) \frac{1}{n} + \varepsilon(t) \quad (2)$$

where  $f_1 = kf_0$ ,  $f_2 = (k+1)f_0$ , ...,  $f_n = (k+n-1)f_0$  and  $k > 1$ . The term  $\varepsilon(t)$  is white noise with zero mean and Gaussian distribution with variance  $\sigma$ . The region of parameters of concern here is  $A$  small (a rule of thumb is  $A=0.9U_{th}$ ) such that the deterministic forcing alone is not enough to cross the threshold and fire a spike. The range of  $\sigma$  values need to be explored on the region where the timing of the spikes is expected that will become more coherent with the input signal. Notice that the deterministic terms in Eq. 2 represent the complex tones where the fundamental  $f_0$  is absent. Since, as was commented above, sounds of this form are perceived with a pitch equal to  $f_0$ , i.e., the “missing” fundamental, numerical experiments were aimed to investigate the response of the model to these signals.



**FIGURE 3:** Probability density function of the inter-spike intervals of the model in response to the two, three, four and five frequencies tones. The figure uses the same format as in Fig. 2, i.e., the vertical axis is the instantaneous frequency of the model spikes ( $1/\text{inter-spike intervals}$ , which should correlate with pitch) and the horizontal axis the driving frequency (using the 1<sup>st</sup> of the two, three, four or five components) for which the computation was done. Superimposed straight lines are the theoretical expected resonance (i.e., from Eq. 5), showing a remarkable agreement.

Choosing initially signals composed by  $n$  subthreshold periodic terms and increasing noise it was found: 1) there is a wide range of noise for which the neuron spikes are spaced preferably  $\sim 1/f_0$  [15-19]. Because there is an optimum noise intensity for which the system emits spikes at a frequency that is not in the input we also call this phenomenon “ghost” stochastic resonance [17-19]. Thus, analogous to the commented psycho-acoustic experiments the neuron’s strongest resonance occurs

for the frequency missing in the input. II) This resonance to the “missing”  $f_0$  is not analogous to the difference  $f_2 - f_1$ . This was found by studying the response to inharmonic signals constructed by shifting all components of the harmonic signal by the same amount  $\Delta f$  :

$$f_1 = kf_0 + \Delta f, \quad f_2 = (k+1)f_0 + \Delta f, \quad \dots, f_n = (k+n-1)f_0 + \Delta f \quad (3)$$

In this case, despite that the difference  $f_n - f_{n-1}$  remains constant, the location of the strongest resonance was found to shift linearly as  $f_p = f_0 + \Delta f / (k+1/2)$  for  $n=2$  and  $f_p = f_0 + \Delta f / (k+1)$  for  $n=3$ . Furthermore, it is possible to generalize predicting the response of the neuron to stimuli composed of  $N$  sinusoidal signals of frequencies:

$$kf_0 + \Delta f, \quad (k+1)f_0 + \Delta f, \quad \dots, (k+N-1)f_0 + \Delta f \quad (4)$$

and the resonance will be expected to occur at frequencies given by:

$$f_r = f_0 + \frac{\Delta f}{k + (N-1)/2} \quad (5)$$

Figure 3 shows four representative examples of numerical results using complex tones composed of 2, 3, 4 and 5 sinusoids (using noise variance  $\sigma = 0.1U_{th}$ ). The graphs represent the probability of observing spikes with a given instantaneous firing frequency  $f_r$  ( $1/ISI$  in the ordinate) as a function of the frequency ( $f_i$ , abscissa) of the lowest of two (three, four or five) components of the driving signal. The lines are the theoretical prediction of Eq. 5 that over-imposes exactly with the simulation results. Note, from Eq. 5, that the neural responses to any combinations of sinusoids will be represented by only one of two diagrams shifted horizontally by an amount equal to  $f_0$ , i.e., tones composed of *even* numbers of sinusoids produce spikes spaced at intervals with statistics as the ones in the left panels; tones with *odd* numbers produce responses like those in the right panels. Thus, the responses will always fall on lines with slopes  $2/3, 2/5, 2/7, 2/9, 2/11$ , etc for even number of tones, and  $1/2, 1/3, 1/4, 1/5, 1/6$ , etc for odd numbers.

In conclusion, Schouten et al. observations were in the right direction: the function best explaining their data wasn't of the form  $1/n$ ; now, from Eq. 5, we know that it is  $1/(n+1)$  [15,19].

### What about phase, about binaural, and why noise?

The main argument sustains that all the illusions are related to a linear interference process that is nonlinearly detected by a noisy threshold. It could be argued that the shape of the time series is sensitive to phase, and unless the detection mechanism accounts for that, the spike timings can be affected. While it is true that the shape of the time series changes with phase, this doesn't affect the timing of the peaks, which remains separated by the same intervals. Note that for inharmonic signals phases are continuously changing, relative to each other. The only case in which the phase relationships are fixed is in perfectly harmonic signals, which hardly exist outside of the computer. Coincidentally it should be noted that in reality the phenomenon of the missing fundamental *is* phase insensitive.

A similar (although weaker) missing fundamental illusion exists as well when each of the two tones enters through a different ear. The neural substrate of this binaural phenomenon must be at some central structure. As discussed in [19] the same

mechanism and arguments hold to explain this illusion. In brief, in analogy with the scenario described for sinusoids, the two spike trains coming from each cochlear nerve *add and interfere* to be subsequently *nonlinearly detected* by the noisy threshold of some neuron. The only difference is that spikes, instead of sinusoids, are added but the rest of the dynamics are equivalent, with a “coincidence detection” flavor.

Finally, why noise? After all, one could design a deterministic peak detector that works as well as the noisy threshold discussed here. The answer is that a deterministic mechanism requires additional parameters to control at which value a spike is generated. Empirically, we found that it can be an important problem. Instead, a noisy mechanism warrants that the largest peaks of the interference have always the largest probability to induce spikes. This is important, such that even in the cases where the amplitude of the signal fluctuates, or when the shape of the signal is distorted by the phase effects commented above, a “hierarchy” is safely maintained: the biggest peak produced by the interference process will always be the most likely to produce a spike. Paradoxically, by adding noise, and without extra neuronal mechanisms, detecting in this way the missing fundamental is a robust process.

## CONCLUSIONS

The views discussed in this paper emphasize simple nonlinear transformations by which sensory neurons can fabricate illusions. In line with previous ideas, is the assumption that the pitch information is in the temporal envelope of the signal, but in contrast, previous proposals have suggested mechanisms using relatively sophisticated structures. These include delay lines [11]; oscillators in combination with integration circuits [20], neural networks [12], and timing nets [7]; but the search for such neural structures has been unsuccessful. Others, [21,22] have instead proposed more abstract alternatives for this problem. Insofar as the mechanism we propose is concerned, there is nothing on it that is peculiar of the auditory system. Then, there is the possibility of identical phenomenology in other sensory modalities such as visual, touch, etc. [23,24], as well as outside biology, where ghost resonance should be an ubiquitous dynamics of driven nonlinear systems, as in lasers [18] and electronic circuits as well [17].

Finally, these illusions stress the difficulty in judging which information our senses are paying attention to when what we see, hear, or feel is not out there. This can be of some importance for the design of optimum sensory implants. Consider, for instance that the calculation of the most usual Shannon information transmission measures, between neuronal input (complex tones) and its output (spike trains) will reflect, for the examples discussed in this paper, the wrong conclusion that no information is being encoded (is just an illusion after all!). The solution reveals itself, as soon as we abandon the wrong assumption that through our senses we asymptotically access a *linear* picture of the physical world.

## ACKNOWLEDGMENTS

This work was supported in part by grants from the National Institute of Health and by the Veterans Administration Greater Los Angeles Healthcare System.

## REFERENCES

1. Strohmeier J. and P. Westbrook, *Divine Harmony. The life and Teachings of Phytagoras.* (1999), Berkeley, CA. Berkeley Hills Books. 159.
2. von Helmholtz H. (1895) *On the Sensations of Tone as a Physiological Basis for the Theory of Music.* trans. Ellis AJ. Longmans, Green and Co., New York.
3. de Boer E., *On the "residue" and auditory pitch perception.* In *Handbook of Sensory Physiology* (1976), eds. Keidel W.D., Neff W.D. Springer-Verlag, Berlin.
4. Schouten J.F., R.J. Ritsma and BL Cardozo, *Pitch of the residue.* J. Acoust. Soc. Am. **34**, 1418-1424 (1962).
5. Gamow G. (1961). *The great physicists from Galileo to Einstein*, Mineola, NY Dover Publications, Inc.
6. Greenberg S., J.T. Marsh, W.S. Brown and J.C. Smith, *Neural temporal coding of low pitch. I. Human frequency following responses to complex tones.* Hearing Res. **25**, 91–114 (1987).
7. Cariani P.A., *Temporal coding of periodicity pitch in the auditory system: an overview.* Neural Plast., **6**(4):147-172 (1999).
8. Tramo M.J., P.A. Cariani, B. Delgutte and L.D. Braidia, *Neurobiological foundations for the theory of harmony in western tonal music.* Ann N Y Acad Sci **930**:92-116, (2001).
9. Goldstein J., *An optimum processor theory for the central formation of pitch of complex tones.* J. Acoust. Soc. Am. **54**, 1496–1516 (1973).
10. Cohen M., S. Grossberg and L. Wyse, *A spectral network model of pitch perception.* J. Acoust. Soc. Am. **98**, 862–879 (1995).
11. Licklider, J.C.R. *A duplex theory of pitch perception.* Experientia **7**, 128–134 (1951).
12. Meddis, R. and J. Hewitt, *Virtual pitch and phase sensitivity of a computer model of the auditory periphery I: pitch identification.* J. Acoust. Soc. Am. **89**, 2866–2882 (1991).
13. Cariani P.A., and B. Delgutte, *Neural correlates of the pitch of complex tones. I. Pitch and pitch salience.* J. Neurophysiol. **76**(3): 1698-1716, (1996).
14. Cariani P.A. and B. Delgutte, *Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch.* J. Neurophysiol. **76**(3): 1717-1734, (1996).
15. Chialvo D.R., O. Calvo, D.L. Gonzalez, O. Piro and G.V. Savino, *Subharmonic resonance and synchronization in neuronal systems.* Phys. Rev. E, **65**, 050902 (2002).
16. Gingl Z., L.B. Kiss and F. Moss, *Non-dynamical stochastic resonance - theory and experiments with white and arbitrarily colored noise.* Europhys. Lett., **29**, 3, 191-196 (1995).
17. Calvo O. and D.R. Chialvo, *Ghost stochastic resonance on an electronic circuit.* Submitted to Electronics Letters (2002).
18. Buldu J.M, D.R. Chialvo, C.R. Mirasso, M.C. Torrent and J. Garcia-Ojalvo, *Ghost Resonance in a semiconductor laser with optical feedback.* Submitted to Phys. Rev. Letters, (2002).
19. Chialvo D.R., *A neural mechanism for the missing fundamental illusion and the perception of pitch.* Submitted to Nature, (2002).
20. Langner G., *Neural Processing and Representation of Periodicity Pitch.* Acta Otolaryngol (Stockh); Suppl **532**: 68-76 (1997).
21. Cartwright J.H., D.L. Gonzalez and O. Piro, *Pitch perception: a dynamical-systems perspective.* Proc. Natl. Acad. Sci. U.S.A., **98**, 9, 4855-4859 (2001).
22. Julicher F., D. Andor and T. Duke, *Physical basis of two-tone interference in hearing.* Proc. Natl. Acad. Sci. U. S.A. **98**,16, 9080-9085. (2001).
23. Kingdom F.A. and D.R. Simmons, *The missing-fundamental illusion at isoluminance.* Perception, **27**, 12, 1451-1460 (1998).
24. Fujii K., S. Kita, T. Matsushima, and Y. Ando, *The missing fundamental phenomenon in temporal vision.* Psychol Res. **64**(2), 149-154 (2000).