Physical basis of two-tone interference in hearing

Frank Jülicher*†, Daniel Andor*, and Thomas Duke*‡

*Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom; and †Institut Curie, PhysicoChimie Curie, Unité Mixte de Recherche/Centre National de la Recherche Scientifique/Institut Curie 168, 26 Rue d'Ulm, 75248 Paris Cedex 05, France

Communicated by T. Gold, Cornell University, Ithaca, NY, May 23, 2001 (received for review September 28, 2000)

The cochlea uses active amplification to capture faint sounds. It has been proposed that the amplifier comprises a set of self-tuned critical oscillators: each hair cell contains a force-generating dynamical system that is maintained at the threshold of an oscillatory instability, or Hopf bifurcation. While the active response to a pure tone provides frequency selectivity, exquisite sensitivity, and wide dynamic range, its intrinsic nonlinearity causes tones of different frequency to interfere with one another in the cochlea. Here we determine the response to two tones, which provides a framework for understanding how the ear processes the more complex sounds of speech and music. Our calculations of two-tone suppression and the spectrum of distortion products generated by a critical oscillator accord with experimental observations of basilar membrane motion and the nervous response. We discuss how the response of a set of self-tuned oscillators, covering a range of characteristic frequencies, represents the structure of a complex sound. The frequency components of the stimulus can be inferred from the timing of neural spikes elicited by the vibrating hair cells. Passive prefiltering by the basilar membrane improves pitch discrimination by reducing interference between tones. Our analysis provides a general framework for examining the relation between the physical nature of the peripheral detection apparatus and psychophysical phenomena such as the sensation of dissonance and auditory illusions.

The nonlinear nature of sound detection has been known for more than 250 years, ever since Tartini described the perception of combination tones that are not present in a complex sound stimulus (1). More recently, experimental techniques have made it possible directly to observe nonlinearities in the basilar membrane motion (2–8) and in the signals of the auditory nerve (9) and, in some instances, to trace them to the transduction process in the hair cells of the inner ear (10, 11). Nonlinearities imply that the response to two tones is not simply a superposition of single frequency responses—different frequencies interfere with and distort one another (10). Manifestations of this interference, which have been identified experimentally, include two-tone suppression (7–9) and the generation of distortion product (DP) frequencies (5, 6). Here we demonstrate that these physiological observations are a direct consequence of the active system of signal detection (12), which the ear uses to amplify weak signals. This insight permits us to draw conclusions about the way that the auditory system infers information about pitch and to shed light on a variety of psychophysical observations.

Active Amplification

That an active mechanism operates in hearing was predicted as long ago as 1948 by Gold (13), who argued that the cochlea might work analogously to a regenerative radio receiver and use a source of energy to counteract the limiting effects of friction and actively amplify the stimulus. General acceptance of this viewpoint has been hampered by the lack of a quantitative theory to confront with experimental data and also by doubts about how the hypothesized amplificatory feedback could be correctly regulated. The recently introduced notion of self-tuned criticality (14) addresses these issues. According to this concept, each hair cell contains a force-generating dynamical system that is poised on the verge of an oscillatory instability (a Hopf bifurcation) and is kept at that critical point by a self-adjustment mechanism. Such a critical oscillator is especially responsive to weak sinusoidal stimuli applied at its characteristic frequency, and calculations have demonstrated that a Hopf resonance provides frequency selectivity, extreme sensitivity, and a broad dynamic range as a result of nonlinear amplification (14, 15).

Here, we recall the response of a mechanical amplifier, which is self-tuned to a Hopf bifurcation, to a pure tone of frequency f, characterized by the amplitude F_f of a periodic stimulus force. The system responds with a deflection X, which is dominated by the Fourier amplitude X_f at the same frequency. Stimulus and response are related by an expansion of the form (14)

$$F_f = AX_f + B|X_f|^2X_f, \qquad [1]$$

where the principal nonlinearity is cubic. The complex coefficients $A(f)$ and $B(f)$ are frequency-dependent. The Hopf bifurcation is characterized by the fact that A vanishes for a characteristic frequency f_c:

$$A(f) \simeq \alpha(f - f_c), \qquad [2]$$

with a complex coefficient α. If the stimulus frequency is close to the characteristic frequency, the linear term in Eq. **1** is insignificant and the system displays a nonlinear amplified response, $X_f \sim F_f^{1/3}$. This occurs when $|f - f_c| < \Delta f_a$, where $\Delta f_a \equiv |7B^{1/3}F_f^{2/3}/(4\alpha)|$ denotes the bandwidth of active amplification, which depends on the stimulus amplitude. For frequencies outside this bandwidth, the response is linear, $X_f \sim F_f/|f - f_c|$. Within the sensitive bandwidth Δf_a the system amplifies with a gain $|X_f|/|F_f| \sim F_f^{-2/3}$, which becomes very high for weak stimuli. The power-law response of the system allows it to operate over a wide dynamic range; it compresses the 12 orders of magnitude in stimulus intensity that the ear can hear into deflections that vary by only a factor of 100.

Two-Tone Interference

Two-tone suppression and the generation of DPs can be explained by the two-tone interferences generated at a Hopf bifurcation. Because the response of a Hopf oscillator is generic near its critical point, we can describe the main features of these effects without detailed knowledge of the underlying molecular mechanisms.

We are interested in the response to two tones (frequencies f_1 and f_2, frequency difference $\Delta f = |f_2 - f_1|$) with amplitudes F_{f_1} and F_{f_2} acting on a critical Hopf oscillator with $f_c = f_1$. The spectrum of the response contains the corresponding amplitudes X_{f_1} and X_{f_2}. Both amplitudes, however, are systematically smaller than each of them would be separately in the absence of the second tone, as is observed experimentally (7). This two-tone suppression is illustrated in Fig. 1, which displays the numerical solutions of a simple model for a Hopf bifurcation (see *Appendix A*). The diagram shows that the nonlinear amplification of a tone at the oscillator's characteristic frequency can be extinguished by

Abbreviation: DP, distortion product.

‡To whom reprint requests should be addressed. E-mail: td18@cam.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

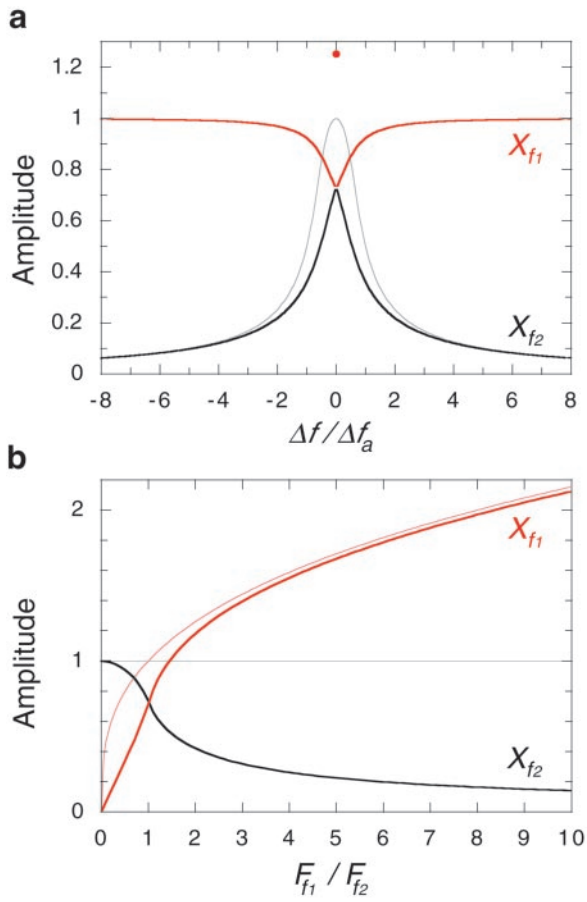

Fig. 1. Two-tone suppression at a Hopf bifurcation. A Hopf oscillator of characteristic frequency f_c is stimulated by an oscillating force that contains two frequency components: a fixed frequency $f_1 = f_c$ and a variable frequency f_2. The amplitudes of the stimulus force at frequencies f_1 and f_2 are denoted F_{f_1} and F_{f_2}, respectively. (*a*) Response of the system for the case of equal stimulus amplitudes $F_{f_1} = F_{f_2}$ as a function of the frequency difference $\Delta f = f_1 - f_2$. The bandwidth of nonlinear amplification of the Hopf oscillator is denoted Δf_a. Displayed are the Fourier amplitudes of the displacement at the two stimulus frequencies, X_{f_1} and X_{f_2}. As f_2 approaches f_1, the response X_{f_1} is suppressed. Similarly, the response X_{f_2} is suppressed as compared to the response of the system to a single tone at frequency f_2, shown as a faint line for comparison. For $\Delta f = 0$ (i.e. $f_1 = f_2$), only a single frequency is present; the Fourier amplitude of the response at this frequency is shown as an individual point at the top in the center. (*b*) Displacement amplitudes X_{f_1} and X_{f_2} of the response to a two-frequency stimulus with a small relative frequency difference ($\Delta f/\Delta f_a = 0.05$) as a function of the stimulus magnitude F_{f_1}, for fixed F_{f_2}. The response X_{f_2} is increasingly suppressed as F_{f_1} is raised. The response X_{f_1} is also diminished in comparison to the single-tone response (i.e. $F_{f_2} = 0$), which is displayed as a faint red line for comparison. The suppression of X_{f_1} is negligible when $F_{f_1} \gg F_{f_2}$; similarly, the suppression of X_{f_2} is negligible when $F_{f_1} \gg F_{f_2}$.

the presence of the second tone, especially when $F_{f_2} > F_{f_1}$ and $\Delta f < \Delta f_a$. Two-tone suppression near a Hopf bifurcation is generic and follows from nonlinearities in the expansion of the Fourier modes (see *Appendix B*). Analysis reveals that the presence of X_{f_2} in the response spectrum generates an effective linear term in the equation for X_{f_1}. This mode thus behaves as if the oscillator were not tuned precisely to the bifurcation point. The corresponding loss of amplification would lead to an increased detection threshold when a second tone (or noise) is introduced. This phenomenon is referred to as masking (16).

The nonlinearities of a Hopf bifurcation also generate new frequencies. The amplitudes X_{f_1} and X_{f_2} couple to the amplitudes with frequencies $2f_1 - f_2$ and $2f_2 - f_1$, which are therefore also present in the response (see *Appendix B*). They subsequently excite further DPs. This leads to a hierarchy of DPs with frequencies $f_k \equiv f_1 + (k-1)\Delta f$, where k is a positive or negative integer. For large $|k - 3/2| > \Delta f_a/\Delta f$ the amplitudes decrease exponentially

$$|X_{f_k}| \sim e^{-\lambda|k-3/2|}. \qquad [3]$$

This characteristic spectrum of DPs is apparent in Fig. 2, where the response of a simple model (see *Appendix A*) is displayed for three different values of Δf, together with the corresponding waveforms. The coefficient λ^{-1}, characterizing the number of strongly excited DPs, decreases as Δf increases. These findings are consistent with experimental data (6). For $\Delta f < \Delta f_a$, a large number of modes is excited and deviations from the exponential law appear. In the limit of vanishing Δf, a singular limit is attained for which the DP amplitudes decay as a power law $|X_{f_k}| \sim |k - 3/2|^{-\nu}$, with $\nu = 4/3$ (see *Appendix B*). This is confirmed by numerical solutions of the simple model described in *Appendix A,* for which we find $\nu \simeq 1.31 \pm 0.05$.

It has previously been suggested that two-tone suppression might be explained by a passive nonlinearity sandwiched between two linear bandpass filters (17). Such a system also could produce a prominent DP at frequency $2f_1 - f_2$ on the assumption that the nonlinearity is cubic (17). This picture has no physical foundation, however, and also fails to account for the observed amplitude dependence of the DP spectrum. By contrast, a self-tuned Hopf bifurcation simultaneously describes two-tone suppression and DP generation, as well as explaining why the nonlinear response to a single tone at the characteristic frequency is cubic (14). Moreover, the model has a sound physical basis, because the presence of an active system of amplification is well-established (12).

Passive Prefilter

Although a Hopf bifurcation provides excellent amplification, especially for weak signals, its ability to filter frequencies is less impressive. An oscillator with $f_c = f_1$ can have a significant contribution X_{f_2} in its response spectrum, even when $\Delta f > \Delta f_a$. This interference, which would pose problems for the detection of complex sounds, can be reduced by prefiltering the stimulus. It is well-established that the mechanical properties of the basilar membrane provide such a filter (19), which, owing to the tonotopic organization of hair cells (20), is centered on the characteristic frequency of each oscillator. The bandwidth Δf_p of this passive prefilter sets a frequency interval above which two-tone interference is suppressed, whatever the level of the stimulus. This bandwidth is therefore roughly equal to the critical bandwidth Δf_{CB} (16) measured at moderate to high intensities, for which the active system does not significantly sharpen the tuning. At low intensities, $\Delta f_a \ll \Delta f_p$ and the active amplifier is itself very effective at suppressing interference; in this case we expect the critical bandwidth to diminish, as is indeed observed experimentally (18).

Pitch Extraction

Our analysis contributes to the longstanding debate about whether frequency is encoded by place or by timing. An example of the response of a full set of hair cells, from which frequency information must be derived, is shown in Fig. 3. Could frequency be represented by the spatial distribution of the neural response, which is maximal where the disturbance is greatest? This notion (which was originally promoted by Békésy's experiments on cochlear mechanics) suffers a drawback. The passive filter is too broad to account for the observed pitch discrimination (19) and, although the active amplifier certainly sharpens the tuning for weak stimuli, it has little effect at high intensities. Many have

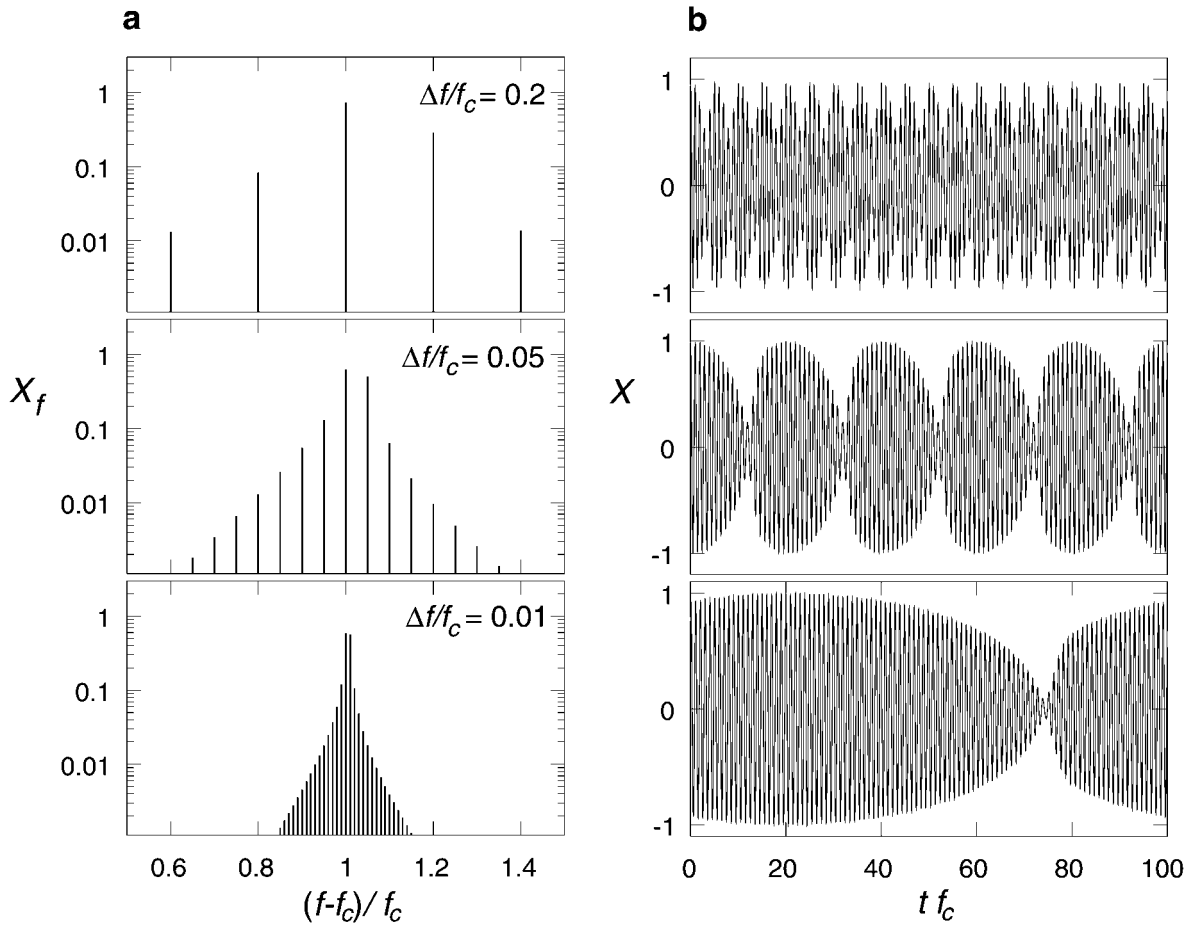

Fig. 2. Two-tone distortion products at a Hopf bifurcation. A Hopf oscillator of characteristic frequency f_c is stimulated by an oscillating force that contains two frequency components, $f_1 = f_c$ and f_2, of equal magnitude. (*a*) Spectral representation of the amplitude X_f of the response as a function of frequency f for several values of the frequency difference $\Delta f = f_1 - f_2$ ($\Delta f/f_c = 0.2$, 0.05 and 0.01). The response exhibits distinct spikes corresponding to distortion products of frequencies $f_k = f_1 + (k - 1)\Delta f$, where k is an integer. The examples shown correspond to a high-level stimulus (see *Appendix A*). (*b*) Complex waveforms of the response X as a function of time t that correspond to the spectra displayed in *a*.

argued that the majority of information about frequency is derived from the detailed time course of the response of the hair cells, via the timing of neural spikes (21, 22). When stimulated by two tones, each excited oscillator vibrates in a pattern that contains components at f_1 and f_2 and also at the DP frequencies. Both the absolute and the relative sizes of each component vary along the cochlea, as the characteristic frequency of the hair cells changes (Fig. 3). The nervous system is provided with only partial information about each of these complex waveforms, in the form of a time series of nervous spikes. We argue below that clear identification of a frequency component is possible only if it is dominant in the motion. This suggests a role for basilar membrane resonance in contraposition to place coding. By prefiltering the stimulus and limiting interference, it permits accurate inference of frequency from nervous timing.

We base our model of pitch discrimination on the generic nervous response of hair cells: (*i*) spikes are elicited whenever the hair bundle deflection traverses a threshold (for frequencies below 5 kHz, at least); and (*ii*) hair cells with a range of different thresholds are present at each characteristic frequency (23). How might the nervous system extract information about frequency and intensity from the resulting spike trains? A simple, but effective, algorithm is to compute a histogram of inter-spike intervals T, summing over all the hair cells in the cochlea (and integrating over a fraction of a second). Perceived tones correspond to peaks in the histogram and are assigned pitch $1/T$, while the perceived loudness of a tone is related to the height of the peak. Examples of this procedure are shown in Fig. 4*a*. Three regimes are apparent, depending on the ratio of the stimulus frequencies. (*i*) When the two frequencies are very close, $\Delta f \ll \Delta f_p$, a single tone is perceived with pitch $(f_1 + f_2)/2$. In addition, there are strong loudness fluctuations at the beat frequency $f_2 - f_1$. The quality of the beats depends on Δf; the ratio of silence to loudness diminishes as the frequencies approach one another (reflecting the waveform in Fig. 2*b*). (*ii*) When the two frequencies differ by a few percent, $\Delta f < \Delta f_p$, the perceived pitch is more ambiguous. The histogram becomes much broader, making pitch assignment less accurate. Furthermore, in the case where two pitches can be discriminated, both undergo rapid fluctuations in loudness with different phases. The resulting variability of the perceived pitch would account for the roughness of sensation that is experienced in this situation (24). (*iii*) At larger frequency differences, $\Delta f > \Delta f_p$, the two pitches f_1 and f_2 are accurately and clearly distinguished. Although we do not rule out the possibility that the nervous system uses a more sophisticated algorithm to infer pitch, the results summarized in Fig. 4*b* concur with a wide variety of psychophysical observations (16).

They also add to our understanding of the enigmatic relation between harmony and the ratio of small integers, on which musical scales are based. Helmholtz (1) overturned the Pythagorean doctrine by arguing that consonant intervals are not perfect harmonies, but simply less jarring dissonances. He

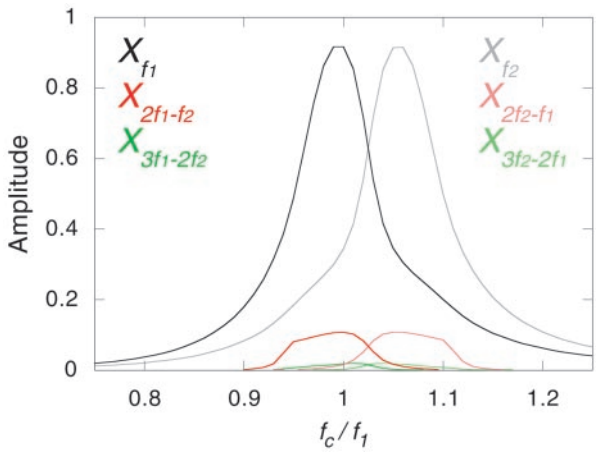

Fig. 3. Two-tone response of a set of Hopf oscillators. The oscillators cover a range of critical frequencies f_c. They are forced by an oscillating stimulus that contains two frequency components f_1 and f_2 ($f_1/f_2 = 1.05$). Both components have an equal amplitude of moderate level (see *Appendix A*). Shown are the Fourier amplitudes X_f of the displacement at the two stimulus frequencies and at the frequencies of the principal distortion products ($2f_1 - f_2$, $2f_2 - f_1$, $3f_1 - 2f_2$ and $3f_2 - 2f_1$), as a function of the critical frequency f_c of the oscillators. In the cochlea, the critical frequency varies with position. The diagram can thus be interpreted as the predicted cochlear response along the basilar membrane. Over a wide range of f_c, the excited oscillators respond with three significant, but unequal, frequency components and therefore display complex, nonlinear waveforms.

ascribed dissonance to close, but inexact, matches in frequency between some of the harmonics that are generated when notes are played on a musical instrument. Subsequent experiments using pure tones lent weight to his argument (24). No preference for integer frequency ratios was expressed; rather, the roughness of two tones was found to be most intrusive when the difference amounted to a few percent and to diminish smoothly as Δf increased. Helmholtz attributed the roughness to beats, but his explanation is unsatisfactory because it fails to explain why dissonance persists above absolute frequencies of 1 kHz, when the beats are too fast to be distinguished. It is little surprise that his linear argument proves inadequate for a nonlinear system. We argue that dissonance arises from the difficulty of inferring frequency components from partial information about a complex waveform, which results in an indeterminancy of pitch. Both dissonance and pitch discrimination depend on the degree of interference between the two tones in the cochlea, so we would expect the interval of dissonance to have the same dependence on frequency as the interval of pitch discrimination, as is indeed observed (24). We also predict that these intervals should diminish at low intensities, when the active amplifier is most effective at sharpening the filter.

In addition, our analysis accounts for two different types of auditory illusion and indicates that they have distinct origins. Many investigators have confirmed that the DP frequencies can be heard (25). Although the DPs are generated by the Hopf resonance, they are not produced as the strongest component. Thus the active amplifier cannot, by itself, account for their audibility. Nevertheless, any nonlinearity in the prefilter would generate small DP components (1), which subsequently would be magnified by oscillators of the corresponding characteristic frequency. The Tartini effect, then, can be explained by the combination of a nonlinear prefilter and active amplifier. The second type of illusion is the residue pitch (26, 27). When f_1 and f_2 are neighboring harmonics, this pitch is identified as the missing fundamental Δf. But when the frequencies are less

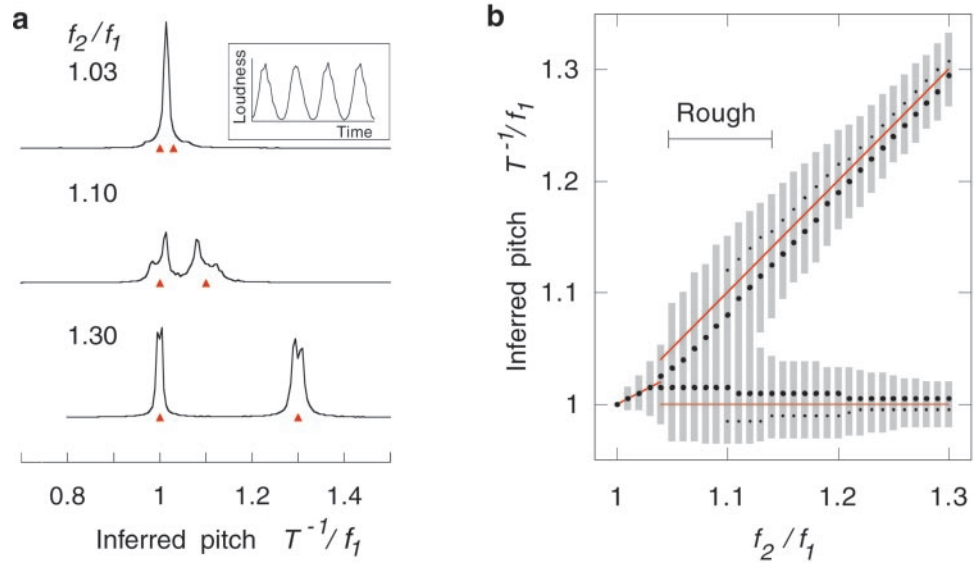

Fig. 4. Pitch determination in the presence of a two-tone stimulus. The response of a Hopf oscillator can be used to generate a model nervous spike train by assuming that a spike is elicited whenever the displacement amplitude X of the oscillator traverses a threshold value. The time intervals T between consecutive spikes provide information about the frequency components in the stimulus. (*a*) Pitch extraction. Histogram of the reciprocal, T^{-1}, of the inter-spike intervals for three different values of f_2/f_1, where f_1 and f_2 are the frequency components in the stimulus (marked by red arrows). The histogram is summed over a number of oscillators with different characteristic frequencies and a range of thresholds (see *Appendix A*). For $f_1/f_2 = 1.3$, two pitches corresponding to the stimulus frequencies can be clearly identified. For $f_1/f_2 = 1.1$, two pitches can be distinguished, but the corresponding peaks are broadened and distorted, suggesting greater ambiguity in pitch assignment. For sufficiently close frequencies $f_1/f_2 = 1.03$, only one pitch, corresponding to the mean stimulus frequency, is extracted. (*Inset*) The time-dependent height of this histogram peak, indicating the occurrence of beats in the loudness. (*b*) Pitch discrimination. The locations of maxima in the histograms of inverse inter-spike intervals are marked by points (for larger values of f_1/f_2, when two main peaks are clearly resolved, each peak is split; the subsidiary maximum is marked by a smaller point). Gray bars provide an indication of the width of the peaks: they extend over the range for which the histogram height exceeds one-tenth of the peak height. The diagram shows that for $f_1/f_2 > 1.15$, the two pitches can be clearly discriminated. For small $f_1/f_2 < 1.05$, a single pitch is inferred and beats occur. In the intermediate range, two pitches are distinguished but they are more ambiguous, suggesting a sensation of roughness.

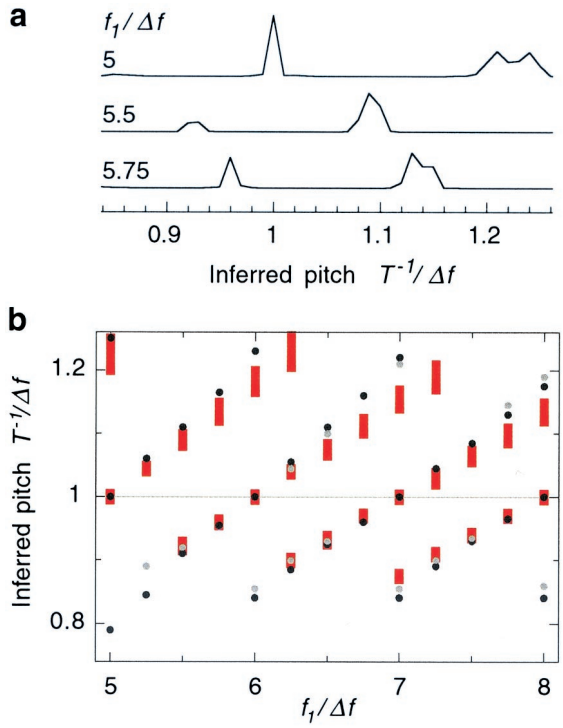

Fig. 5. Residue pitch. In the presence of two stimulus frequencies f_1 and $f_2 = f_1 + \Delta f$, one or more pitches in the vicinity of the frequency Δf can be extracted. (*a*) Displayed are subsidiary peaks in the histogram of the reciprocal, T^{-1}, of the inter-spike intervals, in the vicinity of the pitch that corresponds to Δf. The responses to three different combinations of stimuli frequency are shown (f_2:f_1 = 120:100, 130:110 and 135:115), indicating that the inferred residue pitch varies as a function of $f_1/\Delta f$. (*b*) Comparison between the location of histogram peaks (red bars) and the experimentally determined perceived pitch of residues (black and gray circles; data of ref. 27 for two different subjects).

simply related, one or more pitches close to, but not equal to, Δf are heard. These pitches are derived from the complex waveform of hair-bundle motion and appear as peaks in the histogram of spike intervals (Fig. 5). This adds substance to previous suggestions (26–28) that they are artefacts that arise from the coding of hair-bundle vibration as a time series of nervous spikes.

Summary

In summary, we have provided a unifying physical description of two-tone interference effects and shown that many aspects of our perception of sound may be traced to the physiology of the inner ear, where they originate in one (or a combination) of the three stages of sound detection: prefiltering by the basilar membrane, active amplification by hair cells, and neural coding of hair-bundle motion.

Appendix A: Numerical Model

Active Amplifier. Numerical results are solutions of the complex differential equation

$$\tau\frac{dZ}{dt} = (2\pi i\tau f_c - \varepsilon)Z - |Z|^2Z + \tilde{F}_{f_1}e^{2\pi i f_1 t} + \tilde{F}_{f_2}e^{2\pi i f_2 t},$$

where τ is a time scale and the Hopf bifurcation occurs for $\varepsilon = 0$. We use frequencies for which $f_1/\Delta f$ is integer and define the response $X = x_0\mathrm{Re}(Z)$. The amplitudes $X_f = 2\Delta f\int_0^{1/\Delta f} dtX(t)e^{-2\pi i f t}$, obtained by a Fourier transform of the limit cycle, satisfy Eqs. **1**, **2**, **A4**, and **A5** with $\alpha = (2\pi i\tau/x_0)(F_f/\tilde{F}_f)$ and $B = \bar{B}/2 = C = \alpha/(2\pi i\tau x_0^2)$. The active bandwidth Δf_a is defined as the value of $|f - f_c|$ for which the single-frequency response falls to half the peak amplitude. It varies with force, $\Delta f_a \sim F_f^{2/3}$: comparison with the experimental basilar membrane response (4) suggests that $\Delta f_a/f_c = 0.01$, 0.1, and 1 correspond, respectively, to sound pressure levels 10, 40, and 70 db. Data in Figs. 1 and 2 were obtained with $\tau f_c = 1$, $x_0 = 1$, slightly on the oscillating side of the bifurcation, $\varepsilon = -10^{-3}$. In Fig. 2, $\tilde{F}_{f_1} = \tilde{F}_{f_2} = 0.5$, giving $\Delta f_a/f_c \simeq 1.22$ (a high-level stimulus). In Figs. 3 and 4, a moderate stimulus was used, such that $\Delta f_a/f_c \simeq 0.05$. Data in Fig. 5 were obtained by using stimuli of slightly higher amplitude, corresponding to $\Delta f_a/f_c = 0.1$.

Passive Prefilter. Results in Figs. 3 and 4 were obtained by using a passive prefilter $\chi(f)$, which multiplies F_f before excitation of the Hopf resonance. This is equivalent to the coefficients α and B in Eqs. **1** and **2** having the functional form $\chi^{-1}(f)$. The form $\chi(f) = [(f - f_c)^2/(\Delta f_p)^2 + 1]^{-1}$ with bandwidth $\Delta f_p/f_c = 0.15$ is a fair approximation of the postmortem amplitude-response of the basilar membrane (3). The prefilter suppresses two-tone interferences for $\Delta f > \Delta f_p$. It does not affect the relative amplitudes of DPs.

Neural Response. Data in Figs. 4 and 5 were obtained by using a set of 100 oscillators with characteristic frequencies in the range $f_c/f_1 = 0.5 - 1.5$. A neural spike was elicited by an oscillator every time its response X traversed a given threshold. Histograms of inter-spike intervals were constructed, averaging over all positive thresholds for each oscillator and then summing over all oscillators.

Appendix B: Generic Two-Tone Distortions Near a Hopf Bifurcation

In the presence of a stimulus

$$F(t) = F_{f_1}e^{2\pi i f_1 t} + F_{f_2}e^{2\pi i f_2 t} \qquad \textbf{[A1]}$$

containing two different frequencies f_1 and f_2, the response of a dynamical system close to a Hopf bifurcation contains all Fourier amplitudes with frequencies $f = nf_1 + mf_2$, where n and m are positive or negative integers. For simplicity, we choose stimuli with commensurate frequencies, for which $f_1/\Delta f$ is integer and the response is given by

$$X(t) = \sum_k X_{f_k}e^{2\pi i f_k t} \qquad \textbf{[A2]}$$

with $f_k = f_1 + (k - 1)\Delta f$.

Close to the Hopf bifurcation and for small amplitudes X_f, we can write a general expansion of the form

$$F_{f_k} = \mathcal{A}(f_k)X_{f_k} + \sum_{mn}\mathcal{B}(f_k, f_m, f_n)X_{f_k - f_m - f_n}X_{f_m}X_{f_n} + \ldots \qquad \textbf{[A3]}$$

Throughout this paper, we ignore quadratic terms that occur if the symmetry $X \to -X$ of the active system is broken; they renormalize the coefficients B and $\bar{B}$ of the cubic terms in Eq. **A4** below and are only involved in the generation of higher harmonics and difference tones. For small stimulus and large Δf, the dominant terms in the expansion (**A3**) for the response at frequencies f_1 and f_2 is given by

$$F_{f_1} = A(f_1)X_{f_1} + B(f_1)|X_{f_1}|^2X_{f_1} + \bar{B}(f_1, f_2)|X_{f_2}|^2X_{f_1}$$

$$F_{f_2} = A(f_2)X_{f_2} + B(f_2)|X_{f_2}|^2X_{f_2} + \bar{B}(f_2, f_1)|X_{f_1}|^2X_{f_2}, \qquad \textbf{[A4]}$$

where $A(f) = \mathcal{A}(f)$, $B(f) = \mathcal{B}(f, f, f)$ and $\bar{B}(f_1, f_2) = 2\mathcal{B}(f_1, f_1, f_2)$. We choose one of the stimuli to be at the characteristic frequency, $f_1 = f_c$, therefore $A(f_1) = 0$. For $F_{f_1} = F$ and $F_{f_2} =$

0, it follows that $X_{f_2} = 0$ and the nonlinear single-tone response $|X_{f_1}| \simeq |F^{1/3}B^{-1/3}|$ is recovered. The stimulus F_{f_2} generates X_{f_2}, which in the linear regime obeys $X_{f_2} \simeq F_{f_2}/\alpha\Delta f$. The mode X_{f_2} creates an effective linear term for X_{f_1} with $A_{\text{eff}} \simeq \bar{B}|X_{f_2}|^2$. The mode X_{f_1} therefore behaves as if the system were mistuned and displays a suppressed response that is linear for small F_{f_1}. Similarly, X_{f_1} creates an effective linear term that renormalizes $A(f_2)$ and suppresses the response X_{f_2}.

Distortion products result from nonlinear terms $X_{f_1}^2 X_{f_2}^*$ and $X_{f_2}^2 X_{f_1}^*$, which couple to the Fourier modes with frequencies $2f_1 - f_2$ and $2f_2 - f_1$. For example, in the absence of a stimulus of frequency $2f_1 - f_2$, i.e. $F_{2f_1-f_2} = 0$, the corresponding amplitude obeys

$$0 = A\, X_{2f_1-f_2} + B|X_{2f_1-f_2}|^2 X_{2f_1-f_2} + C\, X_{f_1}^2 X_{f_2}^*, \qquad \textbf{[A5]}$$

where $C = \mathcal{B}(2f_1 - f_2, f_1, -f_2)$. This leads to a DP amplitude $|X_{2f_1-f_2}| \simeq |C/\alpha\Delta f||X_{f_1}^2 X_{f_2}|$. This mode together with X_{f_1} generates, via the same coupling, the DP at $3f_1 - 2f_2$. Recursively, we therefore obtain a hierarchy of DPs with frequencies $f_k = f_1 + (k-1)\Delta f$, whose amplitudes decay exponentially according to Eq. **3** with $\lambda \simeq \ln(7\Delta f/12\Delta f_a)$ for $\Delta f \gg \Delta f_a$.

For smaller Δf, the response involves a large number of terms of the expansion (**A3**) and deviations from a pure exponential appear. An interesting limit occurs for $\Delta f \to 0$ where the linear coefficients $A(f_k) \simeq \alpha\Delta f(k-1)$ and the frequency dependence of $\mathcal{B}(f, f', f'')$ can be neglected for a large number of modes around f_1. With this approximation Eq. **A3** becomes

$$F_{f_k} \simeq B(f_1) \sum_{mn} X_{f_k - f_m - f_n} X_{f_m} X_{f_n}, \qquad \textbf{[A6]}$$

which corresponds to $X(t) \simeq B^{-1/3}F^{1/3}(t)$. The spectrum of $X(t)$ exhibits, for large $|k|$, a power law decay of DPs, $X_{f_k} \sim |k - 3/2|^{-\nu}$, centered around the critical frequency. This can be demonstrated explicitly in the case $F_{f_1} = F_{f_2} = F$, for which the response to a two-tone stimulus for small Δf is $X(t) \simeq (2F/B)^{1/3}\cos^{1/3}(2\pi(f_1 + \Delta f/2)t)\cos^{1/3}(\pi\Delta f t)$. The spectrum X_{f_k} in this limit is therefore simply given by the convolution of the spectrum C_n with itself, where C_n are the Fourier components of $C(t) \equiv \cos^{1/3}(2\pi f t) = \Sigma_n C_n e^{2\pi i n f t}$. These Fourier components decay for large n as $C_n \sim n^{-\nu}$ with $\nu = 4/3$. This power law reflects singularities of dC/dt. Indeed, at the zeros t_0 for which $C(t_0) = 0$, $dC/dt \sim (t - t_0)^{-3/2}$ diverges with a power law that determines the exponent ν.

We thank S. Camalet, E. F. Evans, A. J. Hudspeth, P. Martin, and J. Prost for fruitful discussions. F.J. is grateful for the hospitality of the Cavendish Laboratory and acknowledges support from the Engineering and Physical Sciences Research Council. T.D. is a Royal Society University Research Fellow.

1. Helmholtz, H. L. F. (1954) *On the Sensations of Tone* (Dover, New York).
2. Rhode, W. S. (1971) *J. Acoust. Soc. Am.* **49,** 1218–1231.
3. Sellick, P. M., Patuzzi, R. & Johnstone, B. M. (1982) *J. Acoust. Soc. Am.* **72,** 131–141.
4. Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S. & Robles, L. (1997) *J. Acoust. Soc. Am.* **101,** 2151–2163.
5. Robles, L., Ruggero, M. A. & Rich, N. C. (1991) *Nature (London)* **349,** 413–414.
6. Robles, L., Ruggero, M. A. & Rich, N. C. (1997) *J. Neurophysiol.* **77,** 2385–2399.
7. Ruggero, M. A., Robles, L. & Rich, N. C. (1992) *J. Neurophysiol.* **68,** 1087–1099.
8. Rhode, W. S. & Cooper, N. P. (1993) *Hearing Res.* **66,** 31–45.
9. Sachs, M. B. & Kiang, N. Y. S. (1967) *J. Acoust. Soc. Am.* **43,** 1120–1129.
10. Jaramillo, F., Markin, V. S. & Hudspeth, A. J. (1993) *Nature (London)* **364,** 527–529.
11. Hudpeth, A. J., Choe, Y., Mehta, A. D. & Martin, P. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 11765–11772.
12. Dallos, P. (1992) *J. Neurosci.* **12,** 4575–4585.
13. Gold, T. (1948) *Proc. R. Soc. London Ser. B* **135,** 492–498.
14. Camalet, S., Duke T., Jülicher, F. & Prost, J. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3183–3188.
15. Eguiluz, V. M., Ospeck, M., Choe, Y., Hudspeth, A. J. & Magnasco, M. O. (2000) *Phys. Rev. Lett.* **84,** 5232–5235.
16. Zwicker, E. & Fastl, H. (1999) *Psychoacoustics* (Springer, Berlin).
17. Pfeiffer, R. R. (1970) *J. Acoust. Soc. Am.* **48,** 1373–1378.
18. Bourbon, W. T., Evans, T. R. & Deatherage, B. H. (1968) *J. Acoust. Soc. Am.* **43,** 56–59.
19. von Békésy, G. (1960) *Experiments in Hearing* (McGraw–Hill, New York).
20. Russel, I. J. & Nilsen, K. E. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 2660–2664.
21. Rose, J. E., Brugge, J. F., Anderson, D. J. & Hind, J. E. (1967) *J. Neurophysiol.* **30,** 769–793.
22. Evans, E. F. (1992) *Philos. Trans. R. Soc. London B* **336,** 295–306.
23. Winter, I. M., Robertson, D. & Yates, G. K. (1990) *Hearing Res.* **45,** 191–202.
24. Plomp, R. & Levelt, W. J. M. (1965) *J. Acoust. Soc. Am.* **38,** 548–560.
25. Plomp, R. (1965) *J. Acoust. Soc. Am.* **37,** 1110–1123.
26. Schouten, J. F., Ritsma, R. J. & Lopes Cardozo, B. (1962) *J. Acoust. Soc. Am.* **34,** 1418–1424.
27. Smoorenburg, G. F. (1970) *J. Acoust. Soc. Am.* **48,** 924–942.
28. Rose, J. E., Brugge, J. F., Anderson, D. J. & Hind, J. E. (1968) *J. Neurophysiol.* **32,** 402–423.