Phase Locking of Auditory-Nerve Fibers Reveals Stereotyped Distortions and an Exponential Transfer Function with a Level-Dependent Slope

Adam J. Peterson¹ and Peter Heil^{1,2}

¹Leibniz Institute for Neurobiology, 39118 Magdeburg, Germany, and ²Center for Behavioral Brain Sciences, 39106 Magdeburg, Germany

Phase locking of auditory-nerve-fiber (ANF) responses to the fine structure of acoustic stimuli is a hallmark of the auditory system's temporal precision and is important for many aspects of hearing. Period histograms from phase-locked ANF responses to low-frequency tones exhibit spike-rate and temporal asymmetries, but otherwise retain an approximately sinusoidal shape as stimulus level increases, even beyond the level at which the mean spike rate saturates. This is intriguing because apical cochlear mechanical vibrations show little compression, and mechano-electrical transduction in the receptor cells is thought to obey a static sigmoidal nonlinearity, which might be expected to produce peak clipping at moderate and high stimulus levels. Here we analyze phase-locked responses of ANFs from cats of both sexes. We show that the lack of peak clipping is due neither to ANF refractoriness nor to spike-rate adaptation on time scales longer than the stimulus period. We demonstrate that the relationship between instantaneous pressure and instantaneous rate is well described by an exponential function whose slope decreases with increasing stimulus level. Relatively stereotyped harmonic distortions in the input to the exponential can account for the temporal asymmetry of the period histograms, including peak splitting. We show that the model accounts for published membrane-potential waveforms when assuming a power-of-three, but not a power-of-one, relationship to exocytosis. Finally, we demonstrate the relationship between the exponential transfer functions and the sigmoidal pseudotransducer functions obtained in the literature by plotting the maxima and minima of the voltage responses against the maxima and minima of the stimuli.

Key words: auditory nerve; exponential transfer function; gain control; harmonic distortions; phase locking; refractoriness

Significance Statement

Phase locking of auditory-nerve-fiber responses to the temporal fine structure of acoustic stimuli is important for many aspects of hearing, but the mechanisms underlying phase locking are not fully understood. Intriguingly, period histograms retain an approximately sinusoidal shape across sound levels, even when the mean rate has saturated. We find that neither refractoriness nor spike-rate adaptation is responsible for this behavior. Instead, the peripheral auditory system operates as though it contains an exponential transfer function whose slope changes with stimulus level. The underlying mechanism is distinct from the comparatively weak cochlear mechanical compression in the cochlear apex, and likely resides in the receptor cells.

Introduction

In the auditory system, phase locking refers to the fact that the probability of neuronal action potentials (spikes) varies as a func-

tion of the phase of low-frequency tones, low-frequency components of broadband sounds, or low-frequency fluctuations of the envelope (for review, see Joris et al., 2004; Heil and Peterson, 2015, 2017). In mammals, phase locking of type-I spiral ganglion neurons [auditory-nerve fibers (ANFs)] to stimulus fine structure is observed for frequencies up to several kilohertz and is considered a hallmark of the temporal precision of the auditory system. Temporal coding of fine structure probably underlies perceptual phenomena such as localization of low-frequency sounds (Grothe et al., 2010), detection of near-threshold low-

Received July 16, 2018; revised Feb. 28, 2019; accepted March 7, 2019.

Author contributions: P.H. designed research; P.H. performed research; A.J.P. and P.H. analyzed data; A.J.P. and P.H. wrote the paper.

This work was supported by grants of the Deutsche Forschungsgemeinschaft (Priority Program 1608 "Ultrafast and Temporally Precise Information Processing: Normal and Dysfunctional Hearing," He1721/11-1 and He1721/11-2 to P.H.). Data collection in the laboratory of Dexter R.F. Irvine at Monash University, Clayton, Victoria, Australia, was supported by other grants (He1721/5-1 and He1721/5-2 to P.H.). We are grateful to Dexter R.F. Irvine and Mel Brown for help with data collection.

The authors declare no competing financial interests.

Correspondence should be addressed to Peter Heil at peter.heil@lin-magdeburg.de.

<https://doi.org/10.1523/JNEUROSCI.1801-18.2019>
Copyright © 2019 the authors

Figure 1. Period histograms of ANFs are skewed or even exhibit peak splitting, but none show indications of peak clipping even at high SPLs for which the mean spike rate is saturated. **A**, Period histograms from a low-SR ANF in response to 1200 Hz CF stimuli at various SPLs (A7-U10-R1; SR, 0.2 spikes/s). Vector strengths, V , are also identified. **B**, Mean spike rate versus level function (black; left ordinate) and vector strength versus level function (gray; right ordinate) for the same ANF and stimulus frequency shown in **A**. Significant vector strengths computed from period histograms with ≥ 125 spikes are marked by gray dots. **C, D**, Corresponding plots from a high-SR ANF in response to 390 Hz CF stimuli (A3-U20-R4; SR, 55.1 spikes/s). **E**, Plot of the maximum significant vector strengths of the 78 ANFs as a function of stimulus frequency (gray dots). The mean maximum vector strengths, computed across approximately one-third-octave-wide bins, are also shown (white dots). They compare favorably with the means computed by Weiss and Rose (1988; black dots) from the data of Johnson (1980). **F, G**, Plots corresponding to those in **A–D**, from a high-SR ANF in response to 250 Hz stimuli one octave below CF (A2-U3-R10; SR, 104.2 spikes/s). The period histograms at high SPLs show peak splitting, which results in a reduction in vector strength. For display purposes, period histograms in **A, C**, and **F** were shifted by a constant phase for each ANF such that their peaks occur at $\sim \pi/2$.

frequency sounds (Huet et al., 2018), and perception of pitch and speech (Lorenzi et al., 2006; Moore, 2008) in quiet and in background noise.

ANF phase locking is often studied using tone stimuli because the temporal fine structure that drives the neuronal responses is known and periodic. Phase locking to tone stimuli is often documented using period histograms (Rose et al., 1967; Anderson et al., 1971; Johnson, 1980; Palmer and Russell, 1986; Palmer and Shackleton, 2009). Period histograms shown for three example ANFs in Figure 1A,C,F exhibit the major features of those published in the literature. The instantaneous spike rate can both increase and decrease relative to the spontaneous rate, with increases being more pronounced than decreases. This rate asymmetry is clearest at moderate and high stimulus levels (Littlefield, 1973; Johnson, 1980). Phase locking can be observed at stimulus levels ~ 15 – 20 dB below those at which the mean spike rate begins to increase notably (Kim and Molnar, 1979; Johnson, 1980;

Gleich and Narins, 1988; Hill et al., 1989; Littlefield, 1973; Salvi et al., 1992; Joris et al., 1994a,b; Köppl, 1997; Paolini et al., 2001; Dreyer and Delgutte, 2006; Fukui et al., 2006; Temchin and Ruggero, 2010; Fig. 1B,D,G). Period histograms can also be temporally asymmetric, such that the leading edge of the peak is steeper than the trailing edge. In some cases, usually at high stimulus levels, the peak splits into two or more separate peaks (Kiang and Moxon, 1972; Johnson, 1980; Ruggero and Rich, 1983, 1989; Ronken, 1986; Kiang, 1990; Joris et al., 1994a,b; Cai and Geisler, 1996; Ruggero et al., 1996; Fig. 1F), a phenomenon called “peak splitting.” It is intriguing that, aside from rate and temporal asymmetries, period histograms retain an approximately sinusoidal shape over several orders of magnitude of stimulus amplitude. In particular, they lack “peak clipping” (i.e., the instantaneous spike rate does not saturate during a cycle), even over the range of stimulus levels for which the mean spike rate has saturated (Fig. 1A–D). This is perhaps surprising given that apical cochlear mechanical re-

sponses show little, if any, compression (Robles and Ruggero, 2001), and mechano-electrical transduction in the peripheral auditory system is thought to obey a static sigmoidal nonlinearity (Meddis, 2006; Zilany et al., 2009; Altoè et al., 2018; Horst et al., 2018).

Here, we show that the lack of peak clipping is not due to ANF refractoriness or to mechanisms underlying spike-rate adaptation on time scales longer than the stimulus period. Instead, instantaneous rate and instantaneous pressure are related by an exponential function whose slope decreases as the stimulus level increases. This decrease saturates the mean rate while enabling the full range of instantaneous stimulus pressures to be coded by the instantaneous rate. We show that a sigmoidal function is formed when the maxima and minima reached in these exponential functions are connected across stimulus levels. We also show that relatively stereotyped harmonic distortions in the input to the exponential function could account for the temporal asymmetry of period histograms. Finally, we show that period histograms are compatible with published membrane-potential waveforms when assuming a power-of-three relationship to exocytosis.

Materials and Methods

The data for this study come from a set of recordings from 171 ANFs in five adult cats (three females, two males; weight, 3–3.5 kg) used previously in other studies (Heil et al., 2007, 2008, 2011; Peterson et al., 2014; Peterson and Heil, 2018). All procedures were approved by the Monash University Department of Psychology Animal Ethics Committee.

Surgery and recordings

The surgical and recording procedures are described in detail elsewhere (Heil and Irvine, 1997) and are summarized here. Each cat was anesthetized by an initial intraperitoneal injection of pentobarbitone sodium (40 mg/kg) and kept anesthetized throughout the experiment by intravenous injections of pentobarbitone sodium mixed with physiological saline containing 5% glucose. The electrocardiogram was continuously monitored, and rectal temperature was held at $\sim 38^\circ\text{C}$ by a thermostatically controlled blanket. A round-window electrode for monitoring the compound action potential (Rajan et al., 1991) and a length of fine-bore polyethylene tubing for equalization of static pressure within the middle ear were inserted through a small hole in the bulla on the recording side (left). The bulla was then resealed, and the external meatus cleared of tissue and trimmed to a short stub. The skull was trephined caudal to the tentorium, the dura was removed, the cerebellum over the cochlear nucleus was aspirated, and the auditory nerve was exposed near its exit from the internal auditory meatus.

All recordings were performed with the cat in an electrically shielded, sound-attenuated room. Micropipettes or glass-insulated tungsten microelectrodes with impedances of 4–7 M Ω at 1 kHz were used for the extracellular recordings. Signal-to-noise ratios with the two types of electrodes were approximately equivalent and generally excellent. No cochlear microphonic component was present in the recordings. Under visual control through an operating microscope, the electrode was advanced in a dorsoposterior-to-ventroanterior and slightly medial-to-lateral direction, similar to the approach of Liberman and Kiang (1978). Acoustic stimuli were digitally produced (Tucker-Davis Technologies) and presented to the left ear via a calibrated sealed sound-delivery system consisting of a Stax SRS-MK3 transducer in a coupler (Sokolich, 1981). Noise and tone bursts were used as search stimuli. When spikes having a large signal-to-noise ratio and consistent shape from a single ANF were encountered, the characteristic frequency (CF; the frequency to which the ANF is most sensitive) was determined by manually varying the frequency and amplitude of the search tones. Spike times were taken as the instants at which the amplified and filtered electrode signal crossed a Schmitt-trigger level. They were expressed relative to the onset of the AC voltage driving the speaker and stored digitally with a resolution of 1 μs , which is sufficiently high to precisely quantify phase locking over the frequency range where it occurs in ANFs (Johnson, 1978; Ashida and Carr, 2010).

Stimuli

For each ANF investigated, tone stimuli of a given frequency (typically first at CF) and duration (typically 100 ms, including 4.2-ms-long cosine-squared rise and fall times) were presented at a range of sound pressure levels (SPLs) increasing from low to high values (typically in 4 dB steps from below rate threshold to 80–100 dB SPL). All stimuli started with a positive zero crossing of the AC voltage driving the speaker. Each stimulus was repeated ≤ 200 times (typically 50 times) at a fixed rate (typically 4 Hz). The spontaneous activity of the fiber was subsequently recorded. If the recording conditions were still stable, another stimulus (differing from the previous one in frequency, duration, or rise and fall times) was selected and the protocol was repeated, and so on, until the fiber was lost.

Data analysis

All data analysis was performed using Matlab 2017 (The MathWorks). Spikes occurring during the first 10 ms of each stimulus repetition were excluded from all analyses unless stated otherwise. Although the stimulus rise time was typically only 4.2 ms, we excluded responses during the first 10 ms of each repetition (unless stated otherwise) to accommodate the delay in the response relative to the stimulus (Heil et al., 2008; Neubauer and Heil, 2008) and to ensure that the stimulus amplitude underlying the response was constant. In addition, we noticed in several poststimulus time histograms that the intervals between the first two or three peaks were shorter than the stimulus period (Fig. 2A); such onset effects are beyond the scope of the present study. Spike times were included in the analyses only if they were from complete cycles between 10 ms and the end of the stimulus.

Period histograms. Each spike time was expressed relative to the previous positive zero crossing of the AC voltage driving the speaker (relative spike time t , calculated as the remainder after dividing the spike time by the stimulus period). For analysis and modeling purposes, period histograms were constructed with a bin width of $\Delta t = 1 \mu\text{s}$, equal to the sampling rate of the recordings. For display purposes, wider bins were often chosen (Fig. 1). The instantaneous spike rate $r(t)$ (in spikes/s) was calculated as follows (Eq. 1):

$$r(t) = \frac{n(t)}{\Delta t \cdot m},$$

where $n(t)$ is the number of spikes at time t , and m is the number of stimulus periods included.

Vector strength and phase angle of the mean vector. The degree of phase locking is usually quantified by vector strength (Goldberg and Brown, 1969). Because period histograms of the recovered release events (see below) contain noninteger numbers in each bin, vector strength was calculated directly from the period histogram (Johnson, 1980; Ashida et al., 2010) using the following equation (Eq. 2):

$$V = \frac{1}{\sum_{i=1}^m q_i} \sqrt{\left(\sum_{i=1}^m q_i \cos \theta_i \right)^2 + \left(\sum_{i=1}^m q_i \sin \theta_i \right)^2},$$

where m is the total number of bins in the histogram, q_i is the recovered number (or rate) of release events in bin i , and θ_i is the phase angle of the given bin (in radians). Vector strength can attain values between 0 and 1. A value of 1 means that all events occur in the same bin, whereas a value of 0 typically means that events occur with equal probability across all bins.

Estimates of vector strength were deemed reliable when the period histogram contained ≥ 125 spikes. Vector strength was assessed using the Rayleigh test of uniformity and deemed significant when $V \geq \sqrt{z/n}$, where $z = 4.6052$ is the critical value for $p = 0.01$ and large n (Batschelet, 1965; Mardia and Jupp, 2000). For each period histogram with a reliable and significant vector strength, the phase of the mean vector, $\hat{\theta}$, was calculated as follows (Eq. 3):

$$\hat{\theta} = \tan^{-1} \left(\sum_{i=1}^n q_i \sin \theta_i / \sum_{i=1}^n q_i \cos \theta_i \right).$$

Estimating and removing the effects of refractoriness from period histograms. To determine the extent to which refractoriness affects period histograms, we used a method developed by Johnson and Swami (1983;

Figure 2. Recovery of the rate of release events from spike trains. **A**, Raster plot (top) and poststimulus time histogram (bottom) for an example spike train in response to 50 repetitions of 100 ms tones of 211 Hz and 70 dB SPL, presented at 4 Hz (A3-U8-R3-C10; CF, 353 Hz; SR, 89.8 spikes/s). Only the first 50 ms of the 150 ms interstimulus intervals are shown. Gray vertical lines in the poststimulus time histogram mark complete stimulus cycles between 10 and 100 ms. Note that, over this time range, the spacing of the peaks matches the stimulus period, but that the interval following the peak before 10 ms is shorter. **B**, The excitability function (i.e., the probability that the ANF is not refractory), shown for the first 30 ms of the example spike train (gray line; top), and collapsed into a single period for all complete cycles between 10 and 100 ms (gray lines; bottom). Spike times and corresponding excitabilities are marked by white dots. To aid visibility, the vertical portions of the gray lines following each spike are omitted from the bottom panel. The mean excitability was computed for each time point in the cycle (black line). **C**, Smoothed period histograms showing the instantaneous spike rate (black) and recovered instantaneous event rate (gray) for the example spike train. **D–G**, Simulations were performed to examine the validity of the recovery method. **D**, The simulated (true) instantaneous event rate (dotted line), the instantaneous spike rate resulting from application of refractoriness (black), the recovered instantaneous event rate (gray), and the absolute error of the recovery method (light gray line), for one example simulation. **E**, Relative error as functions of the maximum and mean event rates. The inset shows that the recovered event rates match the true event rates up to high values. **F**, Relative error as a function of the stimulus frequency. **G**, Relative error as functions of the absolute and mean relative refractory periods assumed in the recovery method (true values were 0.6 ms each).

Miller, 1985; Berry and Meister, 1998; Avissar et al., 2013) to remove the effects of refractoriness. The refractory period following each spike has been modeled as consisting of a fixed dead time having duration $t_D \geq 0$ and a random dead time whose duration is drawn from an exponential distribution with mean $t_R > 0$ (Li and Young, 1993; Heil et al., 2007; Peterson et al., 2014; Peterson and Heil, 2018). The probability $p(t')$ that the ANF is excitable (i.e., is not refractory) is expressed as follows (Eq. 4):

$$p(t') = 1 - e^{-(t'-t_D)/t_R} \text{ for } t' \geq t_D$$

$$p(t') = 0 \text{ for } t' < t_D,$$

where t' is the time since the previous spike. The first step of this method involves estimating the instantaneous probability that the ANF is excitable over the entire continuous recording of the spike train, rather than over the shorter segments commonly used to display responses to individual stimulus presentations (e.g., raster plots as in Fig. 2A). For a given spike train, the probability that the ANF is excitable was initially set to 1, and then followed the time course described by Equation 4 after each spike. Both t_D and t_R were set to 0.6 ms, which is consistent with the estimates obtained by Heil et al. (2007) and Peterson and Heil (2018) from analyses of the spontaneous activity of the same ANFs. Figure 2B (top) shows the initial 30 ms of the excitability function resulting from the response in Figure 2A. From each stimulus repetition, all complete cycles of the excitability function between 10 ms and the end of the stimulus were collapsed into a single period, $w(t)$ (Fig. 2B, bottom, gray lines), which is analogous to the construction of a period histogram. The mean of these functions, $\bar{w}(t)$, represents the mean excitability function (i.e., the probability that the ANF is not refractory during the course of a stimulus cycle; Fig. 2B, bottom, black line). The final step is to divide the

instantaneous spike rate from Equation 1 by the mean excitability function to yield the instantaneous rate $q(t)$, expressed as follows (Eq. 5):

$$q(t) = r(t) / \bar{w}(t).$$

This rate, the spike rate that would be expected if there were no refractoriness, has been referred to as the “free firing rate” (Berry and Meister, 1998; Avissar et al., 2013). We refer to it here as the “event rate” because it would also correspond to the rate of transmitter release events from the ribbon synapse in the inner hair cell (IHC) that drives the ANF, assuming that every release event triggers a spike if the ANF is not refractory. There is good evidence for such an assumption (Siegel, 1992; Rutherford et al., 2012; Zhang-Hooks et al., 2016). Each release event could represent the exocytosis of a single vesicle (Chapachnikov et al., 2014; Grabner and Moser, 2018; Huang and Moser, 2018; univesicular release) or a group of vesicles released simultaneously in a coordinated fashion (Glowatzki and Fuchs, 2002; Grant et al., 2010; Graydon et al., 2011; Kim et al., 2013; Rudolph et al., 2015; multivesicular release). Figure 2C shows the instantaneous spike rate $r(t)$ (black) and the corresponding instantaneous event rate $q(t)$ (gray) for the example in Figure 2A,B.

To examine how well this recovery method performs, we simulated phase-locked event trains, applied refractoriness to obtain corresponding spike trains, applied the method to recover the event rates from the spike trains, and then compared the recovered period histograms to those computed from the original event trains. Event trains were simulated as 2000-s-long inhomogeneous Poisson point processes in which the instantaneous rate was equal to the instantaneous value of a sinusoid passed through an exponential transfer function (as in Eq. 17; see Re-

sults). Unless otherwise specified, event trains were simulated for a stimulus of 400 Hz and mean rate of 200 events/s, with the slope of the exponential transfer function chosen such that the vector strength was 0.8. Spike trains were then obtained by applying the refractory function given by Equation 4 with both t_D and t_R equal to 0.6 ms. Also, unless otherwise specified, both t_D and t_R were assumed to be 0.6 ms when recovering the event rates. The results of this analysis are shown in Figure 2D–G. Figure 2D shows the simulated (true) instantaneous event rate (dotted line), the instantaneous spike rate resulting after having applied refractoriness (black), and the recovered instantaneous event rate (gray). The absolute error of the recovery method was computed as the absolute value of the difference between the true and recovered instantaneous event rates (light gray line). A relative error was obtained by integrating the absolute error and the true event rate over the stimulus period and then computing the ratio of these integrals. For the purpose of illustrating absolute and relative errors, the recovery for this example was performed with t_R set to an incorrect value of 0.3 ms, because the correct value of 0.6 ms yields errors too small to discern (Fig. 2G). Figure 2E shows the relative error as functions of the maximum and mean event rates. The error is very small (<0.01) for maximum rates $<\sim 2000$ events/s (or mean rates $<\sim 500$ events/s), beyond which it grows steeply. The reason is that when the event rate increases further, the spike rate saturates, making it impossible to recover the underlying event rate. The recovered maximum and mean event rates then fall short of the true rates (Fig. 2E, inset). Figure 2F shows the relative error as a function of the stimulus frequency. The error is very small (<0.01) over the frequency range covered by the data in this study (200–5000 Hz). Figure 2G shows the relative error as functions of the absolute and mean relative refractory periods assumed in the recovery method. In either case, the error function is V-shaped with a minimum of ~ 0.003 when the assumed refractory period matches the true value (0.6 ms). The relative error was independent of the vector strength (data not shown). The results of these simulations show that the recovery method works well and imply that it is suitable for use with our data.

Model fitting

Maximum likelihood estimation. We used maximum likelihood estimation to fit several models of phase locking (see Results) to the recovered instantaneous event rate. Spike counts are often modeled with the Poisson distribution, $P(n|\lambda) = \lambda^n e^{-\lambda}/n!$, where P represents the probability of observing n spikes in a particular bin, given an expected count λ . However, the factorial function in the denominator can accommodate only integer numbers, whereas the recovered instantaneous event rate typically yields noninteger numbers. We therefore used the continuous approximation f of the discrete Poisson distribution expressed as follows (Eq. 6):

$$f(n|\lambda) = \frac{\lambda^n e^{-\lambda}}{\Gamma(n+1)},$$

where Γ represents the gamma function, which is a continuous extension of the factorial function (note that $\Gamma(n+1) = n!$ for any integer value of n).

The goal is to find the set of parameter values having the maximum likelihood, given the observed numbers of events per bin in a period histogram. To accommodate computational limitations in representing very small numbers, we added an infinitesimal constant value to each probability density computed with Equation 6. We then computed the negative logarithm of the likelihood function (i.e., of the joint probability across all bins). For a period histogram with m bins having $n_1, n_2, \dots, n_m$ events, the negative log likelihood, $-\ln \mathcal{L}$, of a set of parameter values θ_λ that defines the expected numbers $\lambda_1, \lambda_2, \dots, \lambda_m$ is given by the following (Eq. 7):

$$-\ln \mathcal{L}(\theta_\lambda | n_1, n_2, \dots, n_m) = -\ln \prod_{k=1}^m f(n_k | \lambda_k) = -\sum_{k=1}^m \ln f(n_k | \lambda_k).$$

Fitting procedure. We used the Matlab function `fmincon` to estimate the set of model parameters θ_λ that yields the minimum negative log likelihood. For each period histogram, the expected count per bin, λ_k ,

was computed from the rate equation for the model being fitted (Eqs. 15, 17). On the basis of theoretical considerations and preliminary fits to period histograms, simple relationships were identified between each parameter and the stimulus amplitude, the vector strength, or the phase angle of the mean vector. These relationships were used to compute suitable initial values for the final fits. To increase the probability of finding the optimal solution, each histogram was also fitted with 10 additional sets of initial values obtained by adding random values between $+\pi/2$ and $-\pi/2$ to each phase parameter and scaling each remaining parameter by random values between 0.5 and 1.5. The fit having the highest likelihood was selected.

The exponential model

We will show in the Results that the period histograms can be well described by the exponential transfer function expressed as follows (Eq. 8):

$$R_{\text{model}}(t) = A \cdot e^{B \cdot P(t)},$$

where $R_{\text{model}}(t)$ is the rate as a function of time, A is a scaling parameter, B is a slope parameter, and $P(t)$ is the effective stimulus pressure as a function of time.

When $P(t)$ is a sinusoid, Equation 8 yields the von Mises distribution (Ashida et al., 2010, 2013; van Hemmen, 2013) scaled to represent an instantaneous rate. Such an equation has been used previously to model period histograms from ANF responses to low-frequency tones of low and moderate stimulus levels (Siebert, 1970; Colburn, 1973; Johnson, 1974; Heinz et al., 2001; Colburn et al., 2003; Horst et al., 2018). The concentration parameter of the von Mises distribution is equal to the dimensionless product BP_1 of the slope factor B and the amplitude of the sinusoid, P_1 . The vector strength is independent of the mean rate but is related to the concentration parameter by the following (Eq. 9):

$$V = \frac{I_1(BP_1)}{I_0(BP_1)},$$

where I_0 and I_1 are the 0th-order and first-order modified Bessel functions of the first kind (Ashida et al., 2013; van Hemmen, 2013). This result can be obtained by replacing the sums in Equation 2 with the corresponding integrals of $Ae^{BP_1 \sin(\theta)}$, $A \cos(\theta)e^{BP_1 \sin(\theta)}$, and $A \sin(\theta)e^{BP_1 \sin(\theta)}$ over the interval from 0 to 2π radians. The vector strength approaches 0 as BP_1 decreases, and it equals 0 when BP_1 equals 0 (which happens only in the absence of a stimulus when $P_1 = 0$). The mean rate R_{mean} is given by the following (Eq. 10):

$$R_{\text{mean}} = \frac{1}{2\pi} \cdot \int_0^{2\pi} A \cdot e^{BP_1 \sin(\theta)} d\theta = A \cdot I_0(BP_1),$$

such that the vector strength is also given by the following (Eq. 11):

$$V = I_1(BP_1) \cdot \frac{A}{R_{\text{mean}}}.$$

The ratio of the maximum to the minimum instantaneous rate is given by the following (Eq. 12):

$$\frac{R_{\text{max}}(t)}{R_{\text{min}}(t)} = \frac{A \cdot e^{BP_1}}{A \cdot e^{-BP_1}} = e^{2BP_1}.$$

This ratio is uniquely related to vector strength because both depend on only BP_1 . The rate asymmetry resulting from the exponential transfer function can be expressed as the ratio of the difference between the maximum rate reached and the rate at zero pressure to the difference between the rate at zero pressure and the minimum rate reached. This asymmetry measure is always >1 and is given by the following (Eq. 13):

$$Y = e^{BP_1}.$$

Results

We analyzed the responses of 83 ANFs to tone stimuli having frequencies ≤ 5 kHz, which is approximately the highest fre-

Figure 3. Refractoriness affects phase locking in a frequency-dependent manner. **A**, Ratio of the recovered mean event rate to the mean spike rate plotted as a function of mean spike rate. **B**, Ratio of the vector strength computed from the recovered instantaneous event rate to the vector strength computed from the instantaneous spike rate, plotted as a function of stimulus frequency. Ratios >1 represent a decrease of vector strength by refractoriness. Ratios <1 represent an increase of vector strength by refractoriness. Five ratios >1.1 are not shown. **C**, Difference between the phase angles of the mean vectors computed from the recovered instantaneous event rate and the instantaneous spike rate, expressed in units of time by dividing by 2π and the stimulus frequency f_1 . The difference is plotted as a function of stimulus frequency (horizontal axis) and the mean rate of release events (gray scale). This difference is always positive, meaning that refractoriness advances the phase angle of the mean vector. In **B** and **C**, the frequency was randomly jittered to aid visualization. Dashed vertical lines mark the inverse of the mean refractory period (1.2 ms) assumed in the recovery method. T denotes the stimulus period. The white dot marks the example shown in Figure 2A–C.

quency to which phase locking is observed in the responses of cat ANFs (Kiang et al., 1965; Rose et al., 1967; Johnson, 1980; Joris et al., 1994a). Stimuli were presented at various SPLs and frequencies (typically at or below CF) and repeated ≤ 200 times. Analysis was restricted to spike trains for which the computed period histograms contained ≥ 125 spikes and for which the vector strength was significant (2845 spike trains). We show results from the 78 ANFs for which estimates of the spontaneous rate were also available (2675 spike trains), with 38 of these ANFs yielding results for more than one stimulus frequency. The maximum vector strengths of these ANFs are shown as a function of stimulus frequency in Figure 1E (gray dots) and compare favorably with the data of Johnson (1974, 1978, 1980).

Refractoriness affects phase locking in a frequency-dependent manner

Refractoriness limits the spike rate of ANFs and also slightly alters the shapes of the period histograms relative to those that would be observed in the absence of refractoriness (Avissar et al., 2013). Refractoriness has also been proposed to be responsible for the temporal asymmetry of period histograms (Horst et al., 2018). We removed the effects of refractoriness from each spike train to yield period histograms showing the recovered instantaneous rate of presumed synaptic release events rather than the rate of spikes (Fig. 2C). Figure 3A shows that the recovered mean event rate q_{mean} exceeds the mean spike rate r_{mean} by a factor that increases with increasing mean spike rate from just above 1 to ~ 2.3 . Figure 3B shows, as a function of stimulus frequency, the ratio of the vector strength computed from the recovered instantaneous event rate to the vector strength computed from the instantaneous spike rate. Ratios >1 represent a decrease of vector strength by refractoriness, and ratios <1 represent an increase of vector strength by refractoriness. Refractoriness typically decreased vector strength for frequencies $< \sim 400$ Hz and increased it for frequencies between ~ 400 and ~ 800 Hz. In either case, however, the effect is small (typically $< 5\%$). For frequencies > 800 Hz, refractoriness has virtually no effect on vector strength. These findings are consistent with observations from chicken ANFs (Avissar et al., 2013, their Fig. 7b) when expressed with respect to the duration of the refractory period: vector strength is decreased by refractoriness when frequencies are low enough that the stimulus period is considerably longer than the mean refractory period; vector strength is increased by refractoriness when

the stimulus period is only somewhat longer than the mean refractory period; and vector strength is unaffected by refractoriness when the stimulus period is shorter than the mean refractory period. Figure 3C shows, as a function of stimulus frequency, the difference between the phase angles of the mean vectors computed from the recovered instantaneous event rate and the instantaneous spike rate, expressed in units of time. This difference is always positive, meaning that refractoriness advances the phase angle of the mean vector (Fig. 2C). The advancement is typically small (< 0.1 ms) and decreases with increasing stimulus frequency. On this double logarithmic plot, a slope of -1 would correspond to a constant difference in the phase angles when expressed in radians. Because the slope of the decrease is steeper than -1 , the difference in the phase angles decreases with increasing frequency. The difference also decreases as the rate of release events decreases (gray scale).

Together, these analyses show that the dominant effect of ANF refractoriness is to scale down the rate (Fig. 3A). The small changes in vector strength (Fig. 3B) imply only small changes in the shapes of the histograms, regardless of any phase shifts (Fig. 3C). Refractoriness therefore is not the mechanism responsible for skewing the period histograms, in contrast to what has been proposed (Horst et al., 2018), or for preventing peak clipping at high SPLs.

Phenomenological model of phase locking

We seek to relate the period histogram to the instantaneous pressure during a cycle of the tone stimulus having peak amplitude P_1 and frequency f_1 (Fig. 4A). It is not our intention to present a biophysical model of ANF phase locking that incorporates detailed cochlear mechanics, IHC physiology, synaptic transmission, and spike-generating mechanisms. Instead, we attempt to use the simplest and fewest assumptions possible. The first necessary assumption is that the response driven by the stimulus is delayed relative to the AC voltage driving the speaker (Fig. 4B, top). Such a delay is accounted for by a phase-shift parameter (φ_1). The second necessary assumption is that the stimulus driving the response acquires harmonic distortion components, each with its own frequency ($f_2 = 2f_1, f_3 = 3f_1$, etc.), amplitude (P_2, P_3 , etc.), and phase (φ_2, φ_3 , etc.; Fig. 4B, middle and bottom). These distortion components are not present in the acoustic stimulus (Johnson, 1980), but are instead generated by nonlinearities in the auditory periphery (Kim and Molnar, 1979; Sachs and Young,

Figure 4. Phenomenological model of phase locking, illustrated using results obtained from a single example spike train (A6-U14-R1-C20; CF, 1300 Hz; SR, 48.3 spikes/s; f_1 , 1300 Hz; 76 dB SPL). **A**, Acoustic stimulus. **B**, Phase-shifted stimulus of frequency f_1 (top) and harmonic distortion components of frequencies $f_2 = 2 \cdot f_1$ and $f_3 = 3 \cdot f_1$ (middle and bottom). **C**, Mechanical drive obtained by summing the individual frequency components in **B**. **D**, IHC transfer function relating the mechanical drive to the rate of release events. Although the model includes a first-order Boltzmann function (gray, shown here with $\alpha = 1$), the system only uses a portion in the lower tail, such that it can be well described by an exponential function (black). The vertical dotted line marks zero. **E**, Model event rate (black) resulting from the exponential transfer function in **D**. The gray bars show the period histogram of events recovered by removing the effects of refractoriness (**F**) from the empirical period histogram of spikes (**G**).

1980; Ruggero and Rich, 1983; Zwislocki, 1986; Cody and Mountain, 1989; Dallos and Cheatham, 1989; Kiang, 1990; Rhode and Cooper, 1996; Cheatham and Dallos, 1998; Cooper, 1998; Khanna and Hao, 1999; Mountain and Cody, 1999; Guinan, 2012). Consistent with the reported peripheral nonlinearities, we assume that distortion components are present already in the forces deflecting the IHC stereocilia. For convenience, we describe these forces using the units of the acoustic stimulus [pascals (Pa)] and refer to them as the “mechanical drive.” The waveform of the mechanical drive, $P(t)$ (Fig. 4C), is therefore as follows (Eq. 14):

$$P(t) = P_1 \cdot \sin(2\pi f_1 t + \varphi_1) + P_2 \cdot \sin(2\pi f_2 t + \varphi_2) + P_3 \cdot \sin(2\pi f_3 t + \varphi_3) + \dots$$

Kiang (1990) proposed the existence of a second fundamental component to account for a notch in the rate-level function and an $\sim 180^\circ$ phase shift of the response to the fundamental at very high stimulus levels (Kiang, 1984; Liberman and Kiang, 1984). We do not include a second fundamental component because the sum of two sine waves having the same frequency yields, regardless of their amplitudes and phases, a single sine wave with that frequency. Therefore, for the purpose of fitting the model to period histograms, a single fundamental component suffices. In addition, we find no clear notches in the rate-level functions in our data, perhaps because stimulus levels were not high enough. The mechanical drive deflects the IHC stereocilia to give rise to a mechano-electrical transducer (MET) current. We assume that the relationship between the mechanical drive and the MET current is described by a first-order Boltzmann function, as expected for a two-state channel that transitions between a closed and an open state (Fettiplace and Kim, 2014). Several alternative functions were also examined (see below Initial model fits reveal that saturation is not reached and that the model’s implementation can be simplified). The asymmetry of the MET function around its operating point creates a DC component. The MET current activates a variety of voltage-gated K^+ channels and changes the

membrane potential, which is also characterized by AC and DC components (Russell and Sellick, 1978, 1983; Dallos, 1985; Dallos and Cheatham, 1989; Johnson, 2015; Altoè et al., 2018). For simplicity, we assume that changes in the membrane potential are proportional to changes in the MET current. Low-pass filtering by the IHC membrane (not modeled here) will attenuate the AC component in a frequency-dependent way and thereby affect the AC/DC ratio (Palmer and Russell, 1986). The change in the membrane potential gives rise to a change in the Ca^{2+} signal that initiates transmitter release at the ribbon synapse. For simplicity, we assume that changes in the Ca^{2+} signal are proportional to changes in the membrane potential (e.g., consistent with the finding of Li et al., 2014; see their Fig. 5). Given the assumed relationships, the mechanical drive and the synaptic Ca^{2+} signal are also related by a first-order Boltzmann function. The rate of synaptic release events is assumed to be proportional to some power of the Ca^{2+} signal (Beutner et al., 2001; Brandt et al., 2005; Johnson et al., 2005, 2008, 2009, 2010; Goutman and Glowatzki, 2007; Beurg et al., 2008; Heil and Neubauer, 2010), such that the instantaneous event rate of the model, $R_{\text{model}}(t)$, is given by a first-order Boltzmann function raised to the power α (Fig. 4D, gray function, shown for $\alpha = 1$), expressed as follows (Eq. 15):

$$R_{\text{model}}(t) = \left(\frac{S_{\text{max}}}{1 + e^{b \cdot (P_{0.5} - P(t))}} \right)^\alpha.$$

Here, S_{max} is the maximum possible Ca^{2+} signal (in events s^{-1}), $P_{0.5}$ is the mechanical drive (in Pa) at which the Ca^{2+} signal is half the maximum value (i.e., the inflection point of the Boltzmann function), and b is the slope factor (in Pa^{-1}). The maximum possible instantaneous event rate is $R_{\text{max}} = S_{\text{max}}^\alpha$.

Figure 4E shows the resulting instantaneous event rate (black line) as a function of time in the stimulus cycle. It also shows the period histogram of the events (gray bars) recovered by removing the effects of refractoriness (Fig. 4F) from the period histogram of the spikes (Fig. 4G). Because we have an analytical function for the instantaneous event rate (Eq. 15), but not for the instant-

neous spike rate, the model was fitted to the period histogram of the recovered instantaneous event rate. Nonetheless, we also fitted the model equation to the instantaneous spike rates, ignoring the effects of refractoriness. All results were qualitatively very similar to those obtained from fits to the recovered instantaneous event rates (reported below), which is consistent with the finding that the major effect of refractoriness is to scale down the rate (Fig. 3).

Initial model fits reveal that saturation is not reached and that the model's implementation can be simplified

Initial fitting of the model to the recovered instantaneous event rates revealed that it was impossible to reliably estimate $S_{\max}$, α , or $R_{\max}$ (which is derived from $S_{\max}$ and α ; see above Phenomenological model of phase locking). To investigate why this was the case, we systematically fitted the model with α fixed to 1 and $R_{\max}$ (equal to $S_{\max}$ when $\alpha = 1$) fixed to several different values. Two harmonic distortion components were included in the model, resulting in seven free parameters (P_2 , P_3 , φ_1 , φ_2 , φ_3 , b , and $P_{0.5}$). Figure 5A shows the resulting transfer functions for a single example histogram (gray lines; assuming two upper harmonic components). The vertical dotted lines mark the portion of each transfer function covered by the acoustic stimulus. Notably, aside from the four or five functions that saturate at the lowest values of $R_{\max}$, the shapes of the transfer functions are virtually identical over this portion. Furthermore, this portion is well described by an exponential function (black line). A similar exercise was repeated using seven log-spaced values of $R_{\max}$ from 250 to 16,000 events/s for all 2675 period histograms that satisfied our inclusion criteria. For each fit, the negative log likelihood was computed and then normalized by dividing it by the minimum obtained across all $R_{\max}$. Figure 5B shows these normalized values as functions of $R_{\max}$ (gray lines). If there were an optimal value for $R_{\max}$, these functions would be U-shaped. But this is not the case. Instead, they tend to decrease rapidly and monotonically toward 1, and the negative log likelihood hardly changes once $R_{\max}$ is large enough, which explains why $R_{\max}$ could not be estimated reliably. This shows that the data are well described by an exponential transfer function for all SPLs examined (≤ 100 dB SPL). Replacing the Boltzmann function in Equation 15 with the exponential function yields the following (Eq. 16):

$$R_{\text{model}}(t) = (C \cdot e^{D \cdot P(t)})^\alpha = C^\alpha \cdot e^{\alpha D \cdot P(t)},$$

where C is a scale factor and D is a slope factor. The normalized negative log likelihoods for this exponential transfer function (Fig. 5B, exp) are virtually identical to those for the Boltzmann transfer function and large $R_{\max}$. The log likelihoods obtained with Equation 16 are identical for all values of α , because C^α represents a single scaling factor and αD represents a single slope factor, such that C and D can perfectly compensate α . This explains why α (and $S_{\max}$) could not be estimated reliably. The model therefore reduces to a single exponential transfer step (Fig. 4D, black function) between the instantaneous mechanical drive $P(t)$ and the instantaneous event rate $R_{\text{model}}(t)$, expressed as follows (Eq. 17):

$$R_{\text{model}}(t) = A \cdot e^{B \cdot P(t)},$$

which has two fewer parameters than the full model (Eq. 15). Here, A is the scale factor that specifies the rate (in events/s) at $P(t) = 0$ Pa (i.e., the operating point) and accounts for any DC component of the response, and B is the slope factor (in Pa^{-1}) such that R_{model} changes by a factor of e^B when the pressure

changes by 1 Pa. Note that a DC shift in the input to the exponential transfer function would also be included in the scale factor A .

We also examined alternative functions used in other models, specifically the second-order Boltzmann function expected for a three-state channel that transitions between two closed states and an open state (Sumner et al., 2002; for review, see Fettiplace and Kim, 2014; Altoè et al., 2018) and a logarithm-based function (Zhang et al., 2001; Zilany et al., 2009; Bruce et al., 2018). Despite introducing additional free parameters, these functions generally did not perform better than the exponential function (data not shown). The tangent hyperbolic function used by Carney (1993) can be scaled and shifted to be mathematically equivalent to a first-order Boltzmann function and would therefore yield identical fits and parameter values.

Two harmonic distortion components are sufficient to describe most responses

We next systematically examined how many harmonic distortion components would be needed to satisfactorily account for the data using the exponential transfer function. We fitted Equation 17 to each period histogram of recovered events, assuming zero, one, two, or three distortion components in the mechanical drive (Eq. 14). Figure 5C–E shows results for a representative set of data recorded from one ANF in response to CF stimuli of 1300 Hz with levels from 4 to 76 dB SPL. Each function shows the residuals for one stimulus level. Residuals were computed by subtracting the number of events obtained from the model from the number of events obtained from the data. Without distortion components included (Fig. 5C), the residuals are pronounced, particularly for high SPLs, and oscillate around zero with a frequency of two times the stimulus frequency. When one distortion component is included (Fig. 5D), the residuals are much smaller and oscillate around zero with a frequency of three times the stimulus frequency. When two (Fig. 5E) or three (data not shown) distortion components are included, the residuals are smaller still and without a clear oscillatory pattern. These observations emphasize the need for including distortion components in the model. The transfer functions obtained from the model fits for this example depend little on the number of distortion components included (Fig. 5F, shown for 76 dB SPL). Figure 5G shows, for all period histograms, the normalized negative log likelihoods as functions of the number of distortion components (gray lines). These functions tend to decrease markedly with the inclusion of the first distortion component, and then less so with each subsequent component. Although a third component meaningfully improved the fits for a few histograms, we included only two distortion components to avoid overfitting. The relative improvement of the fits when adding distortion components tends to be greatest for responses to low stimulus frequencies (data not shown). Note that although the first and second distortion components would contribute to the Fourier spectrum of the period histogram, the spectrum will also reflect additional components introduced by the exponential transfer function, including additional first and second distortion components differing in phase from those in Equation 14. This means that the distortion components estimated from the model fits are not identical to those seen in Fourier spectra computed from period histograms (Sachs and Young, 1980). Detailed examination of the distortion components estimated from the model might therefore provide new insights into the mechanical drive to the IHC stereocilia.

Figure 5. Justification for the use of an exponential transfer function and two distortion components. **A**, First-order Boltzmann transfer functions (gray lines) estimated for a single example histogram (A6–U14–R1–C20; CF, 1300 Hz; SR, 48.3 spikes/s; f_1 , 1300 Hz; 76 dB SPL), with α fixed to 1 and $R_{\max}$ fixed to several different values. Aside from the four or five functions having the lowest values of $R_{\max}$, the shapes of the functions are virtually identical over the portion used in the fits (between vertical dotted lines) and are well described by an exponential function (black line). **B**, Normalized negative log likelihoods for all period histograms (gray lines) and their geometric mean (circles and black line), plotted as a function of $R_{\max}$. These functions tend to decrease rapidly and converge to the result for an exponential transfer function (exp). **C–E**, Residuals (data minus model) for spike trains recorded from the same ANF as in **A**, in response to CF stimuli presented at 4–76 dB SPL in 4 dB steps. Residuals are shown with no (**C**), one (**D**), or two (**E**) distortion components included. The mean number of events per cycle ranged from 0.042 at low stimulus levels to 0.12 at high stimulus levels. **F**, Exponential transfer functions obtained from the model fits (shown for 76 dB SPL) depend little on the number of distortion components included. **G**, Normalized negative log likelihoods for all period histograms (gray lines) and their geometric mean (circles and black line), plotted as a function of the number of distortion components.

Harmonic distortions in the mechanical drive are stereotypical, both in the absence and presence of peak splitting

All period histograms were fitted by the exponential model (Eq. 17) with the fundamental (f_1) and two distortion components

(f_2 and f_3) in the mechanical drive (Eq. 14). This model has seven free parameters (P_2 , P_3 , φ_1 , φ_2 , φ_3 , A , and B). Figure 6 shows the resulting parameter estimates of the f_2 and f_3 components (P_2 , P_3 , φ_2 , φ_3) for all examples having a significant vector strength for the respective component. Figure 6A,B shows the estimated ampli-

Figure 6. Parameter estimates of the harmonic distortion components. Results are shown for period histograms having significant vector strengths for both the fundamental (f_1) and the respective (f_2 or f_3) distortion component. **A, B**, Estimated amplitudes, P_2 and P_3 , of the distortion components relative to the amplitude P_1 of the fundamental (the stimulus amplitude). The ratios are expressed in dB and plotted as functions of the stimulus amplitude in dB SPL. Color indicates the vector strength for the respective distortion component (see scale in **A**). Parameter values circled in black indicate period histograms characterized by peak splitting. **C, D**, Relative amplitudes of the f_2 and f_3 components plotted as functions of the stimulus frequency f_1 (**C**) and the distance in octaves of f_1 from CF (**D**). In **A–D**, six values of P_2/P_1 and nine values of P_3/P_1 are < -40 dB and are not shown. **E, F**, Estimated phases of the f_2 and f_3 components (horizontal axis), relative to the phase of the fundamental and plotted as functions of tone frequency (vertical axis). The horizontal axis spans one period of f_1 and therefore spans two periods of f_2 and three periods of f_3 . Color indicates the vector strength for the fundamental. Parameter values circled in black in one of the bands indicate period histograms characterized by peak splitting. **G**, Polar histogram of the relative phases from **E** (blue) and **F** (pink). **H**, The typical waveform (black) that results from summing the fundamental component (gray) and the typical f_2 (blue) and f_3 (pink) components.

tudes of these components, P_2 and P_3 , relative to the stimulus amplitude P_1 [in dB, computed as $20 \cdot \log(P_i/P_1)$] and plotted as functions of P_1 (in dB SPL). The color indicates the vector strength for f_2 and f_3 , not for f_1 (Fig. 6A, scale). Period histograms characterized by clear peak splitting (Fig. 1F) yield parameter values circled in black. Figure 6A shows that P_2 tends to be ~ 15 dB below P_1 at low stimulus levels and increases to ~ 10 dB below P_1 at high stimulus levels. Figure 6B shows that P_3 tends to be ~ 20 dB below P_1 at all stimulus levels. The vector strengths for f_2 and f_3 also increase with increasing stimulus level and with increasing relative amplitude of these components. Data points from period histograms characterized by peak splitting do not form separate clusters, but instead tend to lie near the upper edges of the bands of data points. In these cases, P_2 approaches or even exceeds P_1 and the vector strengths for f_2 and f_3 are rather high. Figure 6C shows P_2/P_1 and P_3/P_1 plotted as functions of the stimulus frequency. These relative amplitudes decrease only minimally with increasing frequency. Although significant phase locking is sometimes obtained for distortion components having relatively high frequencies (Fig. 6C, rightmost pink dots, for which f_1 is ~ 2 – 4 kHz and f_3 is therefore ~ 6 – 12 kHz), a visual inspection of the corresponding period histograms leads us to suspect these results are spurious. Figure 6D shows P_2/P_1 and P_3/P_1 plotted as functions of the distance in octaves of the stimulus frequency from CF. Relative amplitudes decrease only minimally over the range of distances examined. Figure 6E, F shows the estimated phases of the f_2 and f_3 components, φ_2 and φ_3 (horizontal axis), relative to the phase of the fundamental, φ_1 , and plotted as functions of tone frequency (vertical axis). The relative phases are given respectively by $\varphi_2/2 - \varphi_1$ and $\varphi_3/3 - \varphi_1$. They are fairly similar

across ANFs and conditions, such that the data points form a narrow band in each plot. The horizontal axis spans one period of f_1 and therefore spans two periods of f_2 and three periods of f_3 , such that the same bands are shown two and three times, respectively. Color indicates the vector strength for f_1 , and not for f_2 and f_3 as in Figure 6A, B. For both the f_2 and f_3 components, there is a subtle but clear tendency for the relative phase to increase with increasing frequency. These effects might be due to frequency-dependent phase shifts introduced by low-pass filtering (e.g., due to the IHC membrane capacitance). Most of the low vector strengths obtained for frequencies < 1 kHz are from period histograms with peak splitting, marked by black circles in one band of data points in Figure 6E, F. For the f_2 component, the relative phases from these histograms align with the left edge of the band. This is because such a phase relationship produces two prominent peaks in the mechanical drive at values of P_2/P_1 lower than those required with other phase relationships. Figure 6G shows a polar histogram of the relative phases from Figure 6E, F. The relative phase of the f_2 component (blue) leads the fundamental (black) by $\sim \pi/12$ radians, and the relative phase of the f_3 component (pink) leads the fundamental by $\sim \pi/7$ radians. Figure 6H shows the typical waveform (black) that results from summing the fundamental (gray) and the typical f_2 (blue) and f_3 (pink) components (the typical values obtained from the model fits were as follows: $P_2/P_1 = -13.1$ dB and $P_3/P_1 = -20.9$ dB, $\varphi_2/2 - \varphi_1 = 2.88$ radians and $\varphi_3/3 - \varphi_1 = 1.64$ radians). The instantaneous amplitude of the typical waveform has a steeper leading edge and a shallower trailing

Figure 7. *A–F*, Instantaneous event rate as a function of instantaneous mechanical drive estimated from model fits without (rows *A, C, E*) and with (rows *B, D, F*) two harmonic distortion components. Examples are the same three as in Figure 1. Model fits are shown as solid lines, whereas data are shown as symbols connected by dashed lines in the form of Lissajous figures. Data are shown using 30 bins per cycle. Without distortion components included in the model, the data can deviate from the model in a systematic loop, most obviously for the ANF showing peak splitting (row *E*; Fig. 1*F*). Systematic deviations are reduced or eliminated when distortion components are included. Also note that the steepness of the functions (and the scaling of the horizontal axes) varies with SPL.

edge than a sinusoid of the same frequency, which is similar to most period histograms (Fig. 1*A,C*).

Harmonic distortions in the mechanical drive eliminate the need for assuming hysteresis in the transfer function

Figure 7 shows the instantaneous event rate plotted as a function of the instantaneous mechanical drive estimated from the model

fits without (rows *A, C, E*) and with (rows *B, D, F*) harmonic distortion components included. All examples are from the same three ANFs as in Figure 1. Data are shown for the majority of SPLs for which there were ≥ 125 spikes and for which the vector strength was significant. Most panels include results for two adjacent SPLs. Model fits are shown as solid lines, whereas data are shown as symbols connected by dashed lines to form Lissajous

Figure 8. *A–C, E, F, H, I.* Dependence of the slope factor B on stimulus level, stimulus frequency, and spontaneous rate. Gray lines connect adjacent stimulus levels. The color indicates the vector strength for the stimulus frequency (f_1). One example function in each panel is marked by squares. The dotted negative diagonal in each panel represents $BP_1 = 1$. The arrows in *B* show how sensitivity and vector strength increase within this coordinate system in the absence of harmonic distortions in the mechanical drive. *D.* Combinations of stimulus frequency and ANF spontaneous rate for which parameter estimates were obtained, with corresponding B -versus-level functions shown in the panel indicated in parentheses. *G.* Vector strength as a function of BP_1 . Color indicates the relative amplitude of the f_2 component. The white line represents the theoretical relationship between vector strength and BP_1 in the absence of distortions.

figures. For display purposes, data are plotted using 30 bins per cycle. At each SPL, the data are well described by the exponential transfer function. Note that the steepness of the function (and the scaling of the horizontal axis) varies with SPL, an issue we return to in the next section. With only the fundamental included (three free parameters: φ_1 , A , and B), the data can form loops that systematically deviate from the model function. This effect is most obvious for the ANF showing peak splitting (Figs. 1*F*, 7*E*), but is also visible in the other examples at high SPLs. Such looping behavior in the data could be interpreted as hysteresis in the transfer function (Patuzzi and Moleirinho, 1998; Brown et al., 2009; Meenderink and van der Heijden, 2011; Horst et al., 2018). However, the loops are reduced or eliminated with f_2 and f_3 components included in the model (Fig. 7, compare rows *B*, *D*, *F*, rows *A*, *C*, *E*).

The exponential transfer function changes with stimulus level

Figure 7 shows that the model with two distortion components describes the data well and that the exponential transfer function changes with stimulus level (except at low SPLs in Fig. 7*D*).

The slope factor

Figure 8 shows the dependence of the slope factor B on stimulus level, stimulus frequency, spontaneous rate, vector strength, and relative amplitude of the f_2 component. Figure 8*D* shows the combinations of stimulus frequency and ANF spontaneous rate for which parameter estimates were obtained. Horizontal dashed lines demarcate spontaneous-rate (SR) groups of <1 spike/s, between 1 and 10 spikes/s, and >10 spikes/s, whereas vertical dashed lines demarcate stimulus-frequency groups <500 Hz, between 500 and 2000 Hz, and >2000 Hz. For each point, the corresponding B -versus-level function is shown in the panel indicated in parentheses. Gray lines connect adjacent stimulus levels. Data points for one B -versus-level function are marked by squares. The example marked in Figure 8*A* is that shown in Figures 1*C*, *D* and 7*C*, *D*. The example marked in Figure 8*H* is that shown in Figures 1*A*, *B* and 7*A*, *B*. The color (except in Fig. 8*G*) indicates the vector strength for the stimulus frequency and shows that vector strength tends to increase with increasing stimulus level, decreasing stimulus frequency, and decreasing spon-

Figure 9. *A*, Example BP_1 -versus-level functions corresponding to the B -versus-level functions in Figure 8*B*, with color also indicating vector strength. *B*, Maximum BP_1 obtained from each BP_1 -versus-level function, plotted as a function of stimulus frequency.

taneous rate, which is consistent with findings in other studies (for review, see Heil and Peterson, 2017).

For all ANFs and stimulus frequencies, B decreases with increasing stimulus level. For ANFs having high SRs and stimulated with tones of low or intermediate frequencies (Fig. 8*A,B*), B changes little at low stimulus levels and then decreases with a slope of approximately -1 in these log-log plots (i.e., B is approximately proportional to the inverse of the stimulus amplitude; we refer to this as a “linear decrease”). For ANFs having high SRs and stimulated with tones of relatively high frequencies (Fig. 8*C*) and for ANFs having intermediate (Fig. 8*E,F*) and low (Fig. 8*H,I*) SRs, B decreases linearly with stimulus amplitude over the entire range of stimulus levels for which parameter estimates were obtained. Nonetheless, these B -versus-level functions probably also have flat portions, which fail to show up here due to our criteria of significant vector strength and ≥ 125 spikes, a conclusion supported by the data of Horst et al. (2018). When stimulus frequencies are high (Fig. 8*C,F,I*), low-pass filtering (e.g., by the IHC membrane capacitance) would attenuate the AC component and prevent vector strengths from reaching significance when stimulus levels are low and spikes are too few. In other cases, vector strength is already near its maximum at stimulus levels at which the criterion of ≥ 125 spikes is reached (Fig. 1*B*).

The flat portion of a B -versus-level function corresponds to the range of stimulus levels over which the slope of the transfer function changes little or not at all, but the dimensionless product BP_1 of the slope factor and the stimulus amplitude (i.e., the concentration parameter) increases with increasing stimulus level (Fig. 9*A*). Over this range of stimulus levels, the vector strength increases with level and there is an overlap of the functions relating the recovered instantaneous event rate to the instantaneous mechanical drive for different stimulus levels (as in Fig. 7*C,D* at low stimulus levels; Horst et al., 2018). The linearly decreasing portion of each B -versus-level function corresponds to the range of stimulus levels over which BP_1 is approximately constant (Fig. 9*A*). This range probably extends to higher stimulus levels than were used here, because none of the B -versus-level functions shows any sign of asymptoting at high stimulus levels. These results are consistent with the proposals of Colburn (1973) and Heinz et al. (2001), who assert that the concentration parameter of the von Mises distribution increases with stimulus level before

saturation. In the absence of distortion components, a constant BP_1 means that the shape of the period histogram and the vector strength are constant across stimulus levels due to the relationship between vector strength and BP_1 (Eq. 9). Figure 8*G* shows this relationship (white line) and also shows the vector strengths obtained from the period histograms of recovered events (symbols, where the color indicates P_2/P_1). Weak distortions in the mechanical drive yield vector strengths near the function described by Equation 9, whereas strong distortions yield predominantly lower vector strengths.

The positions of the B -versus-level functions in Figure 8 along the negative diagonal (e.g., the positions of the knee points) might relate to the sensitivities of the ANFs to the stimulus frequency (Fig. 8*B*, left arrow). Such sensitivity is a function of CF and a function of the distance of the stimulus frequency to the CF. The linearly decreasing portions of the B -versus-level functions also differ in their positions along the positive diagonal. The higher the function along this diagonal, the larger BP_1 and therefore the higher the maximum vector strength (Figs. 8*B*, right arrow; 9*A*, vertical arrow). The positions of the functions along this diagonal, and therefore the maximum vector strength, depend little on spontaneous rate, which is consistent with the small effect of spontaneous rate on maximum vector strength reported in other studies (Johnson, 1980; Joris et al., 1994b; Louage et al., 2004; Dreyer and Delgutte, 2006; Temchin and Ruggero, 2010; for review, see Heil and Peterson, 2017). The position of the functions along the positive diagonal, and therefore the maximum vector strength, is primarily a function of stimulus frequency (Fig. 1*E*; see Fig. 11, compare columns; Palmer and Russell, 1986; Javel and Mott, 1988; Weiss and Rose, 1988; Joris et al., 1994a; Köppl, 1997; Temchin and Ruggero, 2010; Versteegh et al., 2011; Heil and Peterson, 2017) because low-pass filtering reduces B (and therefore BP_1) as frequency increases.

Figure 9*A* shows example BP_1 -versus-level functions corresponding to the B -versus-level functions in Figure 8*B*. Figure 9*B* shows the maximum BP_1 obtained from each BP_1 -versus-level function, plotted as a function of stimulus frequency. This plot is analogous to Figure 1*E*, which shows the maximum vector strength as a function of stimulus frequency. The maximum BP_1 decreases with increasing frequency over the entire frequency range and more clearly than does the vector strength. This is a consequence of the fact that the vector strength saturates but BP_1

Figure 10. *A–C, E, F, H, I*, Dependence of the scale factor A on stimulus level, stimulus frequency, and spontaneous rate. Gray lines connect adjacent stimulus levels. The color indicates the vector strength for the stimulus frequency (f_s). One example function in each panel is marked by squares. *D*, Combinations of stimulus frequency and ANF spontaneous rate for which parameter estimates were obtained, with corresponding A -versus-level functions shown in the panel indicated in parentheses. *G*, Vector strength as a function of A/R_{mean} . Color indicates the relative amplitude of the f_2 component. The white line represents the theoretical relationship between vector strength and A/R_{mean} in the absence of distortions.

does not. Using BP_1 as a measure of phase locking avoids the need to rescale vector strength near its upper limit when higher resolution is desired (Johnson, 1978, 1980; Joris et al., 1994a,b).

The event rate at the operating point

Figure 10 shows the dependence of the scale factor A on stimulus level, stimulus frequency, spontaneous rate, vector strength, and P_2/P_1 , in the same format as Figure 8 and with the same examples marked. For ANFs having high SRs and stimulated with tones of low frequencies, A tends to decrease slightly with increasing stimulus level (Fig. 10A). Several of the functions are rather erratic, particularly at high stimulus levels. For ANFs having high SRs and stimulated with intermediate (Fig. 10B) and high (Fig. 10C) frequencies, A respectively changes little or increases slightly with increasing stimulus level. For ANFs having intermediate SRs, A increases more clearly with increasing stimulus level before saturating (Fig. 10E,F). For ANFs having low SRs, this increase is even more pronounced (Fig. 10H,I). For all ANFs and stimulus frequencies, A would have to approximately equal the spontaneous rate at stimulus levels near the rate threshold. This is seen in the data of Horst et al. (2018), but is not shown in Figure 10 because period histograms for such

low levels did not satisfy our criteria of significant vector strength and ≥ 125 spikes. At high stimulus levels, A saturates at values considerably higher than the spontaneous rate. The ratio of the value of A at saturation to the spontaneous rate is largest for low-SR ANFs and smallest for high-SR ANFs. Vector strength is a function of A/R_{mean} (Eq. 11), which is shown in Figure 10G (white line). Weak distortions in the mechanical drive yield vector strengths near the function described by Equation 11, whereas strong distortions yield predominantly lower vector strengths. When both A and BP_1 are constant (as is often the case at moderate and high stimulus levels; compare Figs. 8, 10), and provided the relative amplitudes and phases of the distortion components remain constant, the mean rate, the shape of the period histogram, and the vector strength are constant across stimulus levels. This is equivalent to complete compression of the response (i.e., a growth of 0 dB/dB).

Exemplary changes in the transfer function with stimulus level

Figure 11 shows, for the examples marked in the middle and right columns of Figures 8 and 10, how the model exponential transfer function changes with stimulus level. The instantaneous event rate is plotted on a logarithmic axis, such that the transfer func-

Figure 11. A–F, Changes of the model transfer function with stimulus level, shown for the examples marked in the middle and right columns of Figures 8 and 10. Note that the instantaneous event rate is plotted on a logarithmic axis, such that the exponential transfer functions form straight lines. Color indicates stimulus level, with the lowest in blue and the highest in red. The extent of each line shows the range of the transfer function used. Dashed horizontal lines mark the estimated spontaneous rate of release events.

tions form straight lines. Color indicates stimulus level. The extent of each line shows the used range of the transfer function. This range is not necessarily symmetrical around $P(t) = 0$ Pa, because the minimum and maximum of the mechanical drive depend on the distortion components present (Fig. 6H). In the absence of distortions, however, the drive would be symmetrical. Furthermore, when plotted on a logarithmic axis, the instantaneous event rate in the period histogram would vary symmetri-

cally around the value at the operating point. In each example in Figure 11, the transfer functions are steepest (i.e., the sensitivity is highest) at low stimulus levels, such that small changes in the pressure lead to large changes in the response. With increasing stimulus level, the transfer functions become shallower (i.e., the sensitivity decreases). This gain control is observed well below CF (Fig. 11A), at CF (Fig. 11B–E), and above CF (Fig. 11F). The ratio of the maximum to minimum instantaneous rates becomes at-

Figure 12. Vector strength and mean number of spikes per cycle as functions of the cycle number. **A–C**, Data from three spike trains for which the vector strength remains relatively constant across cycles while the spike rate adapts. **D**, Data from a spike train for which the vector strength decreases over time. **E**, Data from a spike train for which the vector strength increases over time. In **A–E**, vector strength is plotted only when significant. **F**, Values of Spearman's rank correlation coefficient ρ computed for each function relating significant vector strength to cycle number (as in **A–E**), plotted as a function of stimulus frequency.

tenuated when the stimulus frequency increases (e.g., across columns in Fig. 11), likely due to low-pass filtering. Low-SR ANFs show pronounced growth of the DC component with increasing stimulus level, and therefore their ratio of the mean rate to the spontaneous rate (Fig. 11, dashed horizontal lines) becomes higher.

Rate adaptation does not account for the change in the slope of the transfer function

Here, we examine whether the change in the slope factor B can be accounted for by spike-rate adaptation occurring on time scales longer than the stimulus period. For each cycle of the tone stimulus, starting from the onset, we computed the vector strength and the mean number of spikes per cycle across all repetitions (proportional to the rate). This analysis was restricted to spike trains in which at least half of the cycles contained ≥ 10 spikes and had significant vector strength (622 spike trains). Figure 12A–E shows the results from five example ANFs, with only significant vector strengths shown. There is pronounced spike-rate adaptation in each example, which is consistent with previous reports (for review, see Heil and Peterson, 2015). Figure 12A–C shows data from ANFs for which the vector strength remains relatively constant across cycles. Vector strength was similarly constant for the other stimulus levels not shown here. Delgutte (1980, his Fig. 13) and Joris et al. (1994a,b, their Fig. 7 in each) also show examples of approximately constant vector strength across cycles despite spike-rate adaptation. Assuming distortions are constant (or absent) over time, the finding that vector strength is constant across cycles means that the slope factor B is also constant. This is because vector strength is a function of BP_1 (Eq. 9) and P_1 is constant (except during the brief stimulus onset and offset). Figure 12D shows data from an ANF for which the vector strength decreases over time, which could correspond to a gradual decrease in B or a gradual change in distortions. Figure 12E shows data from an ANF for which the vector strength increases over time, which could correspond to a gradual increase in B or a

gradual change in distortions. More repetitions than were collected would be needed to investigate these possibilities. In any case, the changes in B over time that would potentially account for the changes in vector strength are small compared with the orders-of-magnitude changes in B occurring with changes in stimulus level (Fig. 8).

To examine whether monotonic relationships between vector strength and cycle number are typical at the population level, we computed Spearman's rank-order correlation coefficient for each spike train. Figure 12F shows that most (532 of 622) coefficients scatter around zero and are not significant ($p > 0.01$), indicating that vector strength might typically be constant across cycles. Thirty of the 34 significant negative coefficients were obtained from a single ANF (A2-U3, same ANF as in Fig. 1F). Forty-nine of the 56 significant positive coefficients were observed for frequencies > 1 kHz and were obtained from seven ANFs. These data show that the slope factor B is largely set by mechanisms independent of those responsible for spike-rate adaptation occurring on time scales longer than the stimulus period. The mechanism must also be very fast, because in most cases B is constant starting from the beginning of the response.

The power-law relationship between the synaptic Ca^{2+} signal and the rate of release events

Because period histograms are well described by an exponential transfer function, it is impossible to use fits of the model to the period histograms of recovered release events or spikes to estimate the power α in the model's relationship between the synaptic Ca^{2+} signal and the rate of release events (Eq. 16). To estimate the power, it would be necessary to know the Ca^{2+} signals at individual synapses in response to the acoustic stimuli. Such data are not available, but the assumed linearity between the membrane potential and the Ca^{2+} signal, and the experimental support for this assumption (Li et al., 2014), means that the membrane potential is a reasonable proxy for the Ca^{2+} signal.

Figure 13. Recordings of the IHC receptor potential are consistent with a power of three, but not with a power of one, in the model relationship between the synaptic Ca^{2+} signal and the rate of release events. **A**, Vector strength as a function of the asymmetry of the output of each model transfer function. Color indicates the relative amplitude of the f_2 component. The white line represents the theoretical relationship in the absence of distortions. **B**, IHC receptor potential recordings from Russell and Sellick (1983, their Fig. 1) in response to 300 Hz tone stimuli presented at 90 (top) and 80 (bottom) dB SPL. **C**, Top, Two cycles of the waveform corresponding to the membrane potential of the model if the power were $\alpha = 1$. Gray lines show the waveforms obtained from the histograms within the region marked by the rectangle in **A**, each normalized to its peak value. The red line shows the mean waveform, and the dotted red line shows its value at 0 Pa. Bottom, Corresponding waveforms if the power were $\alpha = 3$. **D**, Transfer functions (gray, red and blue lines) and their operating points (white circles), which give rise to the waveforms in **C**. For $\alpha = 3$, the transfer functions are shallower and the waveforms more sinusoidal than for $\alpha = 1$. **E**, Six cycles of the mean model waveforms for $\alpha = 1$ (red line) and $\alpha = 3$ (blue line) overlaying the center six cycles of the membrane-potential recordings (**B**, boxed region) of Russell and Sellick. The model waveforms have been shifted and scaled to match the amplitude of the recordings. The red and blue dotted lines mark the outputs of the respective model transfer functions at 0 Pa. The white dots mark rough estimates of the output at 0 Pa from the data of Russell and Sellick. The function and operating point for $\alpha = 3$ closely matches the waveform and operating point of the recordings of Russell and Sellick, whereas the function for $\alpha = 1$ has a much sharper peak and a much broader trough, and its operating point is too low.

Only few *in vivo* recordings of the membrane potential of IHCs in response to low-frequency tone stimuli are available. These come from the pioneering work of Russell and Sellick (1983, their Figs. 1, 2) and of Dallos (1985, his Fig. 5; Dallos, 1986, his Fig. 6), and Dallos and Cheatham (1989, their Figs. 1, 10), all in guinea pig. No recordings have been published for cat. Sharp electrodes such as those used for these recordings likely alter the total conductance of the IHC. Such alterations in turn alter the magnitude of the changes in the membrane potential (Zeddes and Siegel, 2004; Altoè et al., 2018), but are unlikely to change the shape of the waveform. The waveforms shown in the literature are similar to one another, regardless of whether recorded from basal or apical IHCs. We therefore chose to compare our model results to the receptor potential recordings shown in Figure 1 of Russell and Sellick (1983), because this figure offers the highest resolution. Figure 13B shows these basal IHC recordings in response to 300 Hz stimuli presented at 90 (top) and 80 (bottom) dB SPL. Figure 13A shows, for our data, vector strength as a function of the rate asymmetry, which is the ratio of the difference between the maximum rate reached and the rate at zero pressure to the difference between the rate at zero pressure and the minimum rate reached. The color of the symbols indicates P_2/P_1 . The white line shows an analogous function constructed using Equations 9 and 13. Because the receptor potential contains negative values, it is impossible to calculate vector strengths for the recordings from Russell and Sellick (1983) and to determine their corresponding positions along this function. Nonetheless, we can compare the

shapes of these recorded waveforms to the shapes of the waveforms derived from our model fits. For this comparison, we chose to use period histograms having high vector strengths and relatively low distortions, located in the region marked by the small rectangle in Figure 13A. Figure 13C (top) shows, for each histogram in this region, two cycles of the model waveform corresponding to the membrane potential if the power were $\alpha = 1$ (gray lines). An approximately linear relationship between the Ca^{2+} signal and the rate of release events has been proposed and interpreted to indicate a nanodomain control of release (Brandt et al., 2005; Kim et al., 2013). Each waveform is normalized to its peak value. The mean waveform (red line) and its value at 0 Pa (red dotted line) are also shown. Figure 13D (top) shows the normalized transfer functions (gray lines) that give rise to the waveforms in Figure 13C. The mean transfer function (red line) and its operating point (white circle) are also shown. Figure 13C,D (bottom) show the corresponding waveforms and transfer functions if the power were $\alpha = 3$ (obtained by taking the cube root of the waveforms and transfer functions for $\alpha = 1$). A power of ~ 3 , derived previously from the dependence of ANF spike rate on stimulus level (Heil et al., 2011) and of ANF first-spike latency on stimulus level and onset duration (Heil et al., 2008; Neubauer and Heil, 2008; Heil and Neubauer, 2010), has been interpreted to indicate a microdomain control of release. Note that for $\alpha = 3$, the transfer function is shallower and the waveform more sinusoidal than for $\alpha = 1$. Figure 13E shows six cycles of the mean waveforms for $\alpha = 1$ (red line) and $\alpha = 3$ (blue line) overlaying

Figure 14. Relationship between the exponential transfer functions and pseudotransducer functions. **A, C**, Model IHC receptor potentials estimated from fits to the responses of two ANFs. For clarity, the estimated distortions and phase shifts are omitted. **B, D**, Corresponding exponential transfer functions relating the instantaneous model receptor potential to the instantaneous pressure. White dots connected by black lines represent pseudotransducer functions analogous to those in the literature.

the center six cycles of the membrane-potential recordings of Russell and Sellick (gray lines, replotted from the boxed region in Fig. 13B). The model waveforms have been scaled to approximately match the peaks and troughs of the recordings. Note that the function for $\alpha = 3$ matches the waveform of the recordings well, whereas the function for $\alpha = 1$ has a much sharper peak and a much broader trough. The red and blue dotted lines mark the outputs of the respective transfer functions at 0 Pa. We also obtained rough estimates of the output at 0 Pa from the data of Russell and Sellick (1983). Precise estimates are impossible to derive due to the presence of distortions that affect the locations of the peaks and troughs (Fig. 6H). The rough estimates were obtained by identifying the voltages (white dots) at the time points halfway between the maxima and minima. The estimates for the leading and trailing edges differ due to the distortions, but the true value is likely within the range identified. These estimates are also better matched by the model when $\alpha = 3$ than when $\alpha = 1$. These comparisons show that, given our model, the IHC membrane-potential recordings and our data would be inconsistent if there were a linear relationship between the Ca^{2+} signal and the rate of release events, but would be consistent if there were a power-of-three relationship.

Relationship between the exponential transfer functions and pseudotransducer functions

Figure 14A,C shows model IHC receptor potentials estimated from model fits to the responses of two ANFs to tones having a range of stimulus levels. For clarity, the estimated distortions and phase shifts are omitted. Receptor potentials were calculated assuming a power-of-three relationship between the Ca^{2+} signal and the rate of release events. The model receptor potentials are similar to the receptor potentials measured *in vivo* in response to acoustic stimulation (Russell and Sellick, 1983; Dallos, 1985, 1986; Russell et al., 1986a; Cody and Russell, 1987; Dallos and Cheatham, 1989, 1990) or *in vitro* in response to sinusoidal deflections of the stereocilia (Russell et al., 1986a,b; He et al., 2004). The AC component increases with stimulus level before saturating. Despite this saturation, no peak clipping occurs. Figure 14B,D shows the corresponding exponential transfer functions relating the receptor potential to the instantaneous pressure. Their slopes decrease with increasing stimulus level. Several investigators have plotted the maxima and minima of the voltage responses of IHCs and outer hair cells (OHCs) against the maxima and minima of the stimulus amplitudes (Russell and Sellick, 1983; Russell et al., 1986a,b; Dallos and Cheatham, 1989; Cheatham and Dallos, 2000). The resulting functions have sometimes

been referred to as “pseudotransducer” functions (Dallos and Cheatham, 1989, 1990; Cheatham and Dallos, 2000) and, aside from effects such as low-pass filtering by the hair-cell membrane, are assumed to approximate the true transducer functions (Russell, 2008, his Figs. 4, 18). In Figure 14*B,D*, the white dots connected by black lines represent pseudotransducer functions analogous and similar to those in the literature. These pseudotransducer functions fail to describe the transfer functions obtained for individual stimulus levels (colored lines), except those for low levels that operate far away from the saturating regions of the pseudotransducer functions. As in Figure 14*D*, many model pseudotransducer functions are nonmonotonic, with a sharp dip at low negative instantaneous pressures. Such dips occur because the transfer function must operate near the resting value when stimulus levels are low, whereas a rapid growth of the DC component can result in transfer functions that operate entirely above the resting value when stimulus levels are higher (Fig. 11*C–F*). Although the sharp dips are not apparent in published pseudotransducer functions, they should also occur in such cases.

Discussion

Period histograms of phase-locked responses of cat ANFs to low-frequency tones are well described by stereotyped harmonic distortions at the input to an exponential transfer function whose slope and operating point change with stimulus level. For an exponential transfer function, a given change in the instantaneous mechanical drive leads to the same relative change in the instantaneous event or spike rate throughout the cycle, such that the shape of the input waveform is faithfully represented by the shape of the logarithm of the output waveform. The finding that the series of transfer stages operating in the auditory periphery behave exponentially overall constrains the sorts of transfers possible at individual stages, but such stages cannot be distinguished using our modeling approach.

The slope of the exponential transfer function

At low stimulus levels, the slope of the exponential transfer function is nearly constant, such that functions relating instantaneous rate and instantaneous mechanical drive for different levels overlap (Fig. 7*C,D*), confirming the finding of Horst et al. (2018). Consistent with Littlefield (1973), such functions can often be approximated by straight lines. At moderate and high levels, the slope of the exponential function decreases with increasing level, such that the functions for different levels do not overlap, consistent with the few observations made at such levels by Horst et al. (2018). The mechanisms underlying the change in slope are independent of ANF refractoriness because such changes persist even after removing refractory effects (Figs. 7, 8, 11), contrary to the proposal of Horst et al. (2018). They are also largely independent of synaptic mechanisms determining spontaneous rate (Frank et al., 2009; Heil and Neubauer, 2010; Wu et al., 2016) because slope-versus-level functions are largely independent of spontaneous rate (Fig. 8). They must further be independent of mechanisms responsible for spike-rate adaptation occurring on time scales longer than the stimulus period (Fig. 12). Such slow adaptation mechanisms might include mechanisms within the ANF (the effects of which are observed during electrical stimulation; Woo et al., 2009; Boulet and Bruce, 2017); synaptic mechanisms, such as vesicle-pool depletion (Westerman and Smith, 1988; Sumner et al., 2003; Goutman and Glowatzki, 2007; Frank et al., 2010; Bruce et al., 2018) and postsynaptic receptor desensitization (Goutman and Glowatzki, 2007; Goutman, 2017); and

mechanisms that cause adaptation of the IHC membrane potential, such as basolateral K^+ conductances (Zeddes and Siegel, 2004; Lopez-Poveda and Eustaquio-Martín, 2006; Johnson, 2015; Altoè et al., 2018). The mechanisms underlying the change in slope must be very fast, because the slope is typically constant from the beginning of the response to a given stimulus (Fig. 12). The change in the slope with stimulus level is similar for frequencies near CF, above CF, and well below CF (Fig. 11). In the cochlear apex, compressive nonlinearities in the vibrations of the basilar (Robles and Ruggero, 2001) or tectorial (Rhode and Cooper, 1996) membranes are weak or absent. The change in slope at moderate and high stimulus levels, equivalent to complete compression when the scale factor is constant, can therefore be at most partially accounted for by such nonlinearities. The primary mechanisms might reside in the IHC transduction machinery (Hudspeth, 2014; for review, see Fettiplace and Kim, 2014; Corey et al., 2017). For example, if the microscopic stiffness of the hypothetical gating spring decreased with increasing stimulus level, then a larger displacement of the stereocilia would be required to yield a given open probability of the MET channels. Rapid gain control in transduction has been proposed for basal OHCs to account for changes in the instantaneous input–output functions derived from cochlear microphonic recordings (Patuzzi and Moirinho, 1998; Meenderink and van der Heijden, 2011). Alternatively, the change in slope could reflect a static saturating nonlinearity that produces increasing clipping of MET currents with increasing stimulus level, but only if later filtering converts such currents into sinusoid-like membrane potentials. The *in vitro* data of Russell et al. (1986a, their Fig. 5) are consistent with either possibility because they reveal that increasing the amplitude of a sinusoidal displacement increases the instantaneous displacement needed to produce a given change in the instantaneous membrane potential. More biophysically oriented modeling is required to address these issues.

Regardless of the underlying mechanism, the effective change in the slope of the transfer function nonetheless prevents peak clipping and enables the full range of the instantaneous pressure to be coded by instantaneous spike rate at all stimulus levels. For responses to high-level sinusoidally amplitude-modulated tones, such gain control would cause peak clipping in period histograms computed for the modulation frequency, but not in those computed for the carrier frequency, just as observed by Joris and Yin (1992).

The operating point of the exponential transfer function

The event rate at the operating point varies with stimulus level in a way that depends on stimulus frequency and ANF spontaneous rate (Figs. 10, 11). The event rate at the operating point is near the spontaneous rate at low stimulus levels (Horst et al., 2018), but then increases with increasing level, particularly for low-SR and medium-SR ANFs. Because ANFs having different spontaneous rates can innervate the same IHC (Liberman, 1982; Wu et al., 2016), spontaneous-rate-dependent differences in the change of the operating point with stimulus level likely arise at the synapse and are therefore unlikely to be accounted for by changes in the IHC transduction machinery with stimulus level or by the low-pass filtering by the IHC membrane capacitance. In the low-CF, high-SR ANFs, the event rate at the operating point initially decreases with increasing level (Fig. 10*A*). This decrease might relate to the finding of Johnson (2015) that the membrane potential of apical IHCs changes less in the depolarizing than in the hyperpolarizing direction in response to tone-like sinusoidal current stimuli (Johnson, 2015, his Fig. 7*B,D*). In this special case, the

transfer function between the stimulus and the membrane potential is expansive below, and compressive above, the operating point, opposite to the transfer function assumed in the model. This would force the estimate of the event rate at the operating point to a value below the true rate (and might underlie the erratic estimates in Fig. 10A).

Stereotyped harmonic distortion components in the mechanical drive

Harmonic distortion components in the input to the transfer function are present even after removing the effects of refractoriness and tend to have relatively consistent amplitude and phase relationships to the fundamental component (Fig. 6). The temporal asymmetry of period histograms, including peak splitting, is due to these distortions and not to the transfer function or ANF refractoriness (Horst et al., 2018). Lissajous figures show hysteresis when the instantaneous cochlear microphonic potential (Patuzzi and Moleirinho, 1998; Brown et al., 2009; Meenderink and van der Heijden, 2011) or the instantaneous ANF spike rate (Horst et al., 2018) is plotted against the instantaneous stimulus pressure. In our data, hysteresis is reduced or eliminated by assuming distortion components in the mechanical drive (Fig. 7).

The power relating calcium signal and release rate

The relationship between the whole-cell Ca^{2+} signal and the whole-cell rate of exocytosis at IHC–ANF (and other) synapses is commonly quantified by a power law with powers of ~ 0.5 – 4 observed in different preparations (Beutner et al., 2001; Brandt et al., 2005; Johnson et al., 2005, 2008, 2009, 2010; Schnee et al., 2005; Goutman and Glowatzki, 2007; Beurg et al., 2008, 2010; Dulon et al., 2009; Vincent et al., 2014). Powers near 3 – 4 have been interpreted as reflecting the biochemical cooperativity of the Ca^{2+} sensor for exocytosis (Roux et al., 2006; Michalski et al., 2017; otoferlin in mammalian IHCs) in microdomain coupling between Ca^{2+} channels and the sensor. Powers near 1 have sometimes been interpreted as reflecting nanodomain coupling (Brandt et al., 2005; Kim et al., 2013), although low powers might also arise from summing saturating higher-power relationships from individual synapses having different sensitivities (Heil and Neubauer, 2010). In our study, the power cannot be derived from the period histograms. However, because the membrane potential and Ca^{2+} signal are approximately proportional (Li et al., 2014), *in vivo* recordings of IHC membrane potentials can be compared with the model Ca^{2+} signal for given powers (Fig. 13). We found that a power of three, but not one, matches the *in vivo* recordings. A power of three was obtained from the changes in ANF first-spike latency and spike rate with stimulus level, independent of CF and spontaneous rate (Heil et al., 2008, 2011; Neubauer and Heil, 2008). It is near that found by Johnson et al. (2008) in mature apical IHCs in gerbil, but is inconsistent with proposed nanodomain coupling in the cochlear apex inferred from the weak effects of the slow Ca^{2+} buffer EGTA (Johnson et al., 2017). However, estimating coupling distances using EGTA might be problematic (Nakamura et al., 2018).

References

- Altòe A, Pulkki V, Verhulst S (2018) The effects of the activation of the inner-hair-cell basolateral K^+ channels on auditory nerve responses. *Hear Res* 364:68–80.
- Anderson DJ, Rose JE, Hind JE, Brugge JF (1971) Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: frequency and intensity effects. *J Acoust Soc Am* 49:Suppl 2:1131+.
- Ashida G, Carr CE (2010) Effect of sampling frequency on the measurement of phase-locked action potentials. *Front Neurosci* 4:172.
- Ashida G, Wagner H, Carr CE (2010) Processing of phase-locked spikes and periodic signals. In: *Analysis of parallel spike trains* (Grün S, Rotter S, eds), pp 59–74. New York: Springer.
- Ashida G, Funabiki K, Carr CE (2013) Theoretical foundations of the sound analog membrane potential that underlies coincidence detection in the barn owl. *Front Comput Neurosci* 7:151.
- Avissar M, Wittig JH Jr, Saunders JC, Parsons TD (2013) Refractoriness enhances temporal coding by auditory nerve fibers. *J Neurosci* 33:7681–7690.
- Batschelet E (1965) *Statistical methods for the analysis of problems in animal orientation and certain biological rhythms*. Washington, DC: The American Institute of Biological Science.
- Berry MJ 2nd, Meister M (1998) Refractoriness and neural precision. *J Neurosci* 18:2200–2211.
- Beurg M, Safieddine S, Roux I, Bouleau Y, Petit C, Dulon D (2008) Calcium- and otoferlin dependent exocytosis by immature outer hair cells. *J Neurosci* 28:1798–1803.
- Beurg M, Michalski N, Safieddine S, Bouleau Y, Schneggenburger R, Chapman ER, Petit C, Dulon D (2010) Control of exocytosis by synaptotagmins and otoferlin in auditory hair cells. *J Neurosci* 30:13281–13290.
- Beutner D, Voets T, Neher E, Moser T (2001) Calcium dependence of exocytosis and endocytosis at the cochlea inner hair cell afferent synapse. *Neuron* 29:681–690.
- Boulet J, Bruce IC (2017) Predictions of the contribution of HCN half-maximal activation potential heterogeneity to variability in intrinsic adaptation of spiral ganglion neurons. *J Assoc Res Otolaryngol* 18:301–322.
- Brandt A, Khimich D, Moser T (2005) Few $\text{Ca}_v1.3$ channels regulate the exocytosis of a synaptic vesicle at the hair cell ribbon synapse. *J Neurosci* 25:11577–11585.
- Brown DJ, Hartsock JJ, Gill RM, Fitzgerald HE, Salt AN (2009) Estimating the operating point of the cochlear transducer using low-frequency biased distortion products. *J Acoust Soc Am* 125:2129–2145.
- Bruce IC, Erfani Y, Zilany MSA (2018) A phenomenological model of the synapse between the inner hair cell and auditory nerve: implications of limited neurotransmitter release sites. *Hear Res* 360:40–54.
- Cai Y, Geisler CD (1996) Temporal patterns of the responses of auditory-nerve fibers to low-frequency tones. *Hear Res* 96:83–93.
- Carney LH (1993) A model for the responses of low-frequency auditory-nerve fibers in cat. *J Acoust Soc Am* 93:401–417.
- Chapochnikov NM, Takago H, Huang CH, Pangršič T, Khimich D, Neef J, Auge E, Göttfert F, Hell SW, Wichmann C, Wolf F, Moser T (2014) Uniquantal release through a dynamic fusion pore is a candidate mechanism of hair cell exocytosis. *Neuron* 83:1389–1403.
- Cheatham MA, Dallos P (1998) The level dependence of response phase: observations from cochlear hair cells. *J Acoust Soc Am* 104:356–369.
- Cheatham MA, Dallos P (2000) The dynamic range of inner hair cell and organ of corti responses. *J Acoust Soc Am* 107:1508–1520.
- Cody AR, Mountain DC (1989) Low-frequency responses of inner hair cells: evidence for a mechanical origin of peak splitting. *Hear Res* 41:89–99.
- Cody AR, Russell IJ (1987) The responses of hair cells in the basal turn of the guinea-pig cochlea to tones. *J Physiol* 383:551–569.
- Colburn HS (1973) Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. *J Acoust Soc Am* 54:1458–1470.
- Colburn HS, Carney LH, Heinz MG (2003) Quantifying the information in auditory-nerve responses for level discrimination. *J Assoc Res Otolaryngol* 4:294–311.
- Cooper NP (1998) Harmonic distortion on the basilar membrane in the basal turn of the guinea-pig cochlea. *J Physiol* 509:277–288.
- Corey DP, Maoileidigh DÓ, Ashmore JF (2017) Mechanical transduction processes in the hair cell. In: *Understanding the cochlea* (Manley GA, Gummer AW, Popper AN, Fay RR, eds), pp 75–111. New York: Springer.
- Dallos P (1985) Response characteristics of mammalian cochlear hair cells. *J Neurosci* 5:1591–1608.
- Dallos P (1986) Neurobiology of cochlear inner and outer hair cells: intracellular recordings. *Hear Res* 22:185–198.
- Dallos P, Cheatham MA (1989) Nonlinearities in cochlear receptor potentials and their origins. *J Acoust Soc Am* 86:1790–1796.
- Dallos P, Cheatham MA (1990) Effects of electrical polarization on inner hair cell receptor potentials. *J Acoust Soc Am* 87:1636–1647.

- Delgutte B (1980) Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers. *J Acoust Soc Am* 68:843–857.
- Dreyer A, Delgutte B (2006) Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: implications for sound localization. *J Neurophysiol* 96:2327–2341.
- Dulon D, Safieddine S, Jones SM, Petit C (2009) Otoferlin is critical for a highly sensitive and linear calcium-dependent exocytosis at vestibular hair cell ribbon synapses. *J Neurosci* 29:10474–10487.
- Fettiplace R, Kim KX (2014) The physiology of mechano-electrical transduction channels in hearing. *Physiol Rev* 94:951–986.
- Frank T, Khimich D, Neef A, Moser T (2009) Mechanisms contributing to synaptic Ca^{2+} signals and their heterogeneity in hair cells. *Proc Natl Acad Sci U S A* 106:4483–4488.
- Frank T, Rutherford MA, Strenzke N, Neef A, Pangršič T, Khimich D, Fetjova A, Gundelfinger ED, Liberman MC, Harke B, Bryan KE, Lee A, Egner A, Riedel D, Moser T (2010) Bassoon and the synaptic ribbon organize Ca^{2+} channels and vesicles to add release sites and promote refilling. *Neuron* 68:724–738.
- Fukui I, Sato T, Ohmori H (2006) Improvement of phase information at low sound frequency in nucleus magnocellularis of the chicken. *J Neurophysiol* 96:633–641.
- Gleich O, Narins PM (1988) The phase response of primary auditory afferents in a songbird (*Sturnus vulgaris* L.). *Hear Res* 32:81–91.
- Glowatzki E, Fuchs PA (2002) Transmitter release at the hair cell ribbon synapse. *Nat Neurosci* 5:147–154.
- Goldberg JM, Brown PB (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 32:613–636.
- Goutman JD (2017) Mechanisms of synaptic depression at the hair cell ribbon synapse that support auditory nerve function. *Proc Natl Acad Sci U S A* 114:9719–9724.
- Goutman JD, Glowatzki E (2007) Time course and calcium dependence of transmitter release at a single ribbon synapse. *Proc Natl Acad Sci U S A* 104:16341–16346.
- Grabner CP, Moser T (2018) Individual synaptic vesicles mediate stimulated exocytosis from cochlear inner hair cells. *Proc Natl Acad Sci U S A* 115:12811–12816.
- Grant L, Yi E, Glowatzki E (2010) Two modes of release shape the postsynaptic response at the inner hair cell ribbon synapse. *J Neurosci* 30:4210–4220.
- Graydon CW, Cho S, Li GL, Kachar B, von Gersdorff H (2011) Sharp Ca^{2+} nanodomains beneath the ribbon promote highly synchronous multivesicular release at hair cell synapses. *J Neurosci* 31:16637–16650.
- Grothe B, Pecka M, McAlpine D (2010) Mechanisms of sound localization in mammals. *Physiol Rev* 90:983–1012.
- Guinan JJ Jr (2012) How are inner hair cells stimulated? Evidence for multiple mechanical drives. *Hear Res* 292:35–50.
- He DZ, Jia S, Dallos P (2004) Mechano-electrical transduction of adult outer hair cells studied in the gerbil hemicochlea. *Nature* 429:766–770.
- Heil P, Irvine DR (1997) First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* 78:2438–2454.
- Heil P, Neubauer H (2010) Summing across different active zones can explain the quasi-linear Ca^{2+} -dependencies of exocytosis by receptor cells. *Front Synaptic Neurosci* 2:148.
- Heil P, Peterson AJ (2015) Basic response properties of auditory nerve fibers: a review. *Cell Tissue Res* 361:129–158.
- Heil P, Peterson AJ (2017) Spike timing in auditory-nerve fibers during spontaneous activity and phase locking. *Synapse* 71:5–36.
- Heil P, Neubauer H, Irvine DR, Brown M (2007) Spontaneous activity of auditory-nerve fibers: insights into stochastic processes at ribbon synapses. *J Neurosci* 27:8457–8474.
- Heil P, Neubauer H, Brown M, Irvine DR (2008) Towards a unifying basis of auditory thresholds: distributions of the first-spike latencies of auditory-nerve fibers. *Hear Res* 238:25–38.
- Heil P, Neubauer H, Irvine DR (2011) An improved model for the rate-level functions of auditory-nerve fibers. *J Neurosci* 31:15424–15437.
- Heinz MG, Colburn HS, Carney LH (2001) Rate and timing cues associated with the cochlear amplifier: level discrimination based on monaural cross-frequency coincidence detection. *J Acoust Soc Am* 110:2065–2084.
- Hill KG, Stange G, Mo J (1989) Temporal synchronization in the primary auditory response in the pigeon. *Hear Res* 39:63–73.
- Horst JW, McGee J, Walsh EJ (2018) Input–output curves of low and high spontaneous rate auditory nerve fibers are exponential near threshold. *Hear Res* 367:195–206.
- Huang CH, Moser T (2018) Ca^{2+} regulates the kinetics of synaptic vesicle fusion at the afferent inner hair cell synapse. *Front Cell Neurosci* 12:364.
- Hudspeth AJ (2014) Integrating the active process of hair cells with cochlear function. *Nat Rev Neurosci* 15:600–614.
- Huet A, Desmadryl G, Justal T, Nouvian R, Puel JL, Bourien J (2018) The interplay between spike-time and spike-rate modes in the auditory nerve encodes tone-in-noise threshold. *J Neurosci* 38:5727–5738.
- Javel E, Mott JB (1988) Physiological and psychophysical correlates of temporal processes in hearing. *Hear Res* 34:275–294.
- Johnson DH (1974) The response of single auditory-nerve fibers in the cat to single tones: synchrony and average discharge rate. PhD thesis, Massachusetts Institute of Technology.
- Johnson DH (1978) The relationship of post-stimulus time and interval histograms to the timing characteristics of spike trains. *Biophys J* 22:413–430.
- Johnson DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am* 68:1115–1122.
- Johnson DH, Swami A (1983) The transmission of signals by auditory-nerve fiber discharge patterns. *J Acoust Soc Am* 74:493–501.
- Johnson SL (2015) Membrane properties specialize mammalian inner hair cells for frequency or intensity encoding. *eLife* 4:e08177.
- Johnson SL, Marcotti W, Kros CJ (2005) Increase in efficiency and reduction in Ca^{2+} dependence of exocytosis during development of mouse inner hair cells. *J Physiol* 563:177–191.
- Johnson SL, Forge A, Knipper M, Münkner S, Marcotti W (2008) Tonic variation in the calcium dependence of neurotransmitter release and vesicle pool replenishment at mammalian auditory ribbon synapses. *J Neurosci* 28:7670–7678.
- Johnson SL, Franz C, Knipper M, Marcotti W (2009) Functional maturation of the exocytotic machinery at gerbil hair cell ribbon synapses. *J Physiol* 587:1715–1726.
- Johnson SL, Franz C, Kuhn S, Furness DN, Rüttiger L, Münkner S, Rivolta MN, Seward EP, Herschman HR, Engel J, Knipper M, Marcotti W (2010) Synaptotagmin IV determines the linear Ca^{2+} -dependence of vesicle fusion at auditory ribbon synapses. *Nat Neurosci* 13:45–52.
- Johnson SL, Olt J, Cho S, von Gersdorff H, Marcotti W (2017) The coupling between Ca^{2+} channels and the exocytotic Ca^{2+} sensor at hair cell ribbon synapses varies tonotopically along the mature cochlea. *J Neurosci* 37:2471–2484.
- Joris PX, Yin TC (1992) Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* 91:215–232.
- Joris PX, Carney LH, Smith PH, Yin TC (1994a) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol* 71:1022–1036.
- Joris PX, Smith PH, Yin TC (1994b) Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *J Neurophysiol* 71:1037–1051.
- Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* 84:541–577.
- Khanna SM, Hao LF (1999) Nonlinearity in the apical turn of living guinea pig cochlea. *Hear Res* 135:89–104.
- Kiang NYS (1984) Peripheral neural processing of auditory information. In: *Handbook of physiology*, Vol III, Pt 2, Section 1: The nervous system. (Brookhart JM, Mountcastle VB, eds), pp 639–674. Bethesda, MD: American Physiological Society.
- Kiang NY (1990) Curious oddments of auditory-nerve studies. *Hear Res* 49:1–16.
- Kiang NY, Moxon EC (1972) Physiological considerations in artificial stimulation of the inner ear. *Ann Otol Rhinol Laryngol* 81:714–730.
- Kiang NYS, Watanabe T, Thomas EC, Clark LF (1965) Discharge patterns of single fibers in the cat's auditory nerve. Cambridge, MA: MIT.
- Kim DO, Molnar CE (1979) A population study of cochlear nerve fibers: comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones. *J Neurophysiol* 42:16–30.
- Kim MH, Li GL, von Gersdorff H (2013) Single Ca^{2+} channels and exocytosis at sensory synapses. *J Physiol* 591:3167–3178.
- Köppl C (1997) Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, *Tyto alba*. *J Neurosci* 17:3312–3321.

- Li GL, Cho S, von Gersdorff H (2014) Phase-locking precision is enhanced by multiquantal release at an auditory hair cell ribbon synapse. *Neuron* 83:1404–1417.
- Li J, Young ED (1993) Discharge-rate dependence of refractory behavior of cat auditory-nerve fibers. *Hear Res* 69:151–162.
- Lieberman MC (1982) Single-neuron labeling in the cat auditory nerve. *Science* 216:1239–1241.
- Lieberman MC, Kiang NY (1978) Acoustic trauma in cats. cochlear pathology and auditory-nerve activity. *Acta Otolaryngol Suppl* 358:1–63.
- Lieberman MC, Kiang NY (1984) Single-neuron labeling and chronic cochlear pathology. IV. Stereocilia damage and alterations in rate and phase-level functions. *Hear Res* 16:75–90.
- Littlefield WM (1973) Investigation of the linear range of the peripheral auditory system. PhD thesis, Washington University.
- Lopez-Poveda EA, Eustaquio-Martín A (2006) A biophysical model of the inner hair cell: the contribution of potassium currents to peripheral auditory compression. *J Assoc Res Otolaryngol* 7:218–235.
- Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BC (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U S A* 103:18866–18869.
- Louage DHG, van der Heijden M, Joris PX (2004) Temporal properties of responses to broadband noise in the auditory nerve. *J Neurophysiol* 91:2051–2065.
- Mardia KV, Jupp PE (2000) Directional statistics. New York: Wiley.
- Meddis R (2006) Auditory-nerve first-spike latency and auditory absolute threshold: a computer model. *J Acoust Soc Am* 119:406–417.
- Meenderink SWF, van der Heijden M (2011) Dynamic aspects of cochlear microphonic potentials. *AIP Conf Proc* 1403:53–58.
- Michalski N, Goutman JD, Auclair SM, Boutet de Monvel J, Tertrais M, Emptoz A, Parrin A, Nouaille S, Guillon M, Sachse M, Ciric D, Bahloul A, Hardelin JP, Sutton RB, Avan P, Krishnakumar SS, Rothman JE, Dulon D, Safieddine S, Petit C (2017) Otoferlin acts as a Ca^{2+} sensor for vesicle fusion and vesicle pool replenishment at auditory hair cell ribbon synapses. *eLife* 6:e31013.
- Miller MI (1985) Algorithms for removing recovery-related distortion from auditory-nerve discharge patterns. *J Acoust Soc Am* 77:1452–1464.
- Moore BC (2008) The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol* 9:399–406.
- Mountain DC, Cody AR (1999) Multiple modes of inner hair cell stimulation. *Hear Res* 132:1–14.
- Nakamura Y, Reva M, DiGregorio DA (2018) Variations in Ca^{2+} influx can alter chelator-based estimates of Ca^{2+} channel–synaptic vesicle coupling distance. *J Neurosci* 38:3971–3987.
- Neubauer H, Heil P (2008) A physiological model for the stimulus-dependence of first-spike latency of auditory-nerve fibers. *Brain Res* 1220:208–223.
- Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res* 24:1–15.
- Palmer AR, Shackleton TM (2009) Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency. *J Assoc Res Otolaryngol* 10:233–250.
- Paolini AG, FitzGerald JV, Burkitt AN, Clark GM (2001) Temporal processing from the auditory nerve to the medial nucleus of the trapezoid body in the rat. *Hear Res* 159:101–116.
- Patuzzi R, Moleirinho A (1998) Automatic monitoring of mechano-electrical transduction in the guinea pig cochlea. *Hear Res* 125:1–16.
- Peterson AJ, Heil P (2018) A simple model of the inner-hair-cell ribbon synapse accounts for mammalian auditory-nerve-fiber spontaneous spike times. *Hear Res* 363:1–27.
- Peterson AJ, Irvine DR, Heil P (2014) A model of synaptic vesicle-pool depletion and replenishment can account for the interspike interval distributions and nonrenewal properties of spontaneous spike trains of auditory-nerve fibers. *J Neurosci* 34:15097–15109.
- Rajan R, Irvine DR, Cassell JF (1991) Normative N1 audiogram data for the barbiturate-anesthetized domestic cat. *Hear Res* 53:153–158.
- Rhode WS, Cooper NP (1996) Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo. *Aud Neurosci* 2:101–121.
- Robles L, Ruggero MA (2001) Mechanics of the mammalian cochlea. *Physiol Rev* 81:1305–1352.
- Ronken DA (1986) Anomalous phase relations in threshold-level responses from gerbil auditory nerve fibers. *J Acoust Soc Am* 79:417–425.
- Rose JE, Brugge JF, Anderson DJ, Hind JE (1967) Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* 30:769–793.
- Roux I, Safieddine S, Nouvian R, Grati M, Simmler MC, Bahloul A, Perfettini I, Le Gall M, Rostaing P, Hamard G, Triller A, Avan P, Moser T, Petit C (2006) Otoferlin, defective in a human deafness form, is essential for exocytosis at the auditory ribbon synapse. *Cell* 127:277–289.
- Rudolph S, Tsai MC, von Gersdorff H, Wadiche JI (2015) The ubiquitous nature of multivesicular release. *Trends Neurosci* 38:428–438.
- Ruggero MA, Rich NC (1983) Chinchilla auditory-nerve responses to low-frequency tones. *J Acoust Soc Am* 73:2096–2108.
- Ruggero MA, Rich NC (1989) “Peak splitting”: intensity effects in cochlear afferent responses to low-frequency tones. In: *Cochlear mechanisms—structure, function and models* (Wilson JP, Kemp DT, eds), pp 259–266. London: Plenum.
- Ruggero MA, Rich NC, Shivapuja BG, Temchin AN (1996) Auditory-nerve responses to low-frequency tones: intensity dependence. *Aud Neurosci* 2:159–185.
- Russell IJ (2008) Cochlear receptor potentials. In: *The senses: A comprehensive reference* (Basbaum AI, Kaneko A, Shepherd GM, Westheimer G, Albright TD, Masland RH, Dallos P, Oertel D, Firestein S, Beauchamp GK, Bushnell MC, Kaas JH, Gardner E, eds), pp 319–358. San Diego: Elsevier.
- Russell IJ, Sellick PM (1978) Intracellular studies of hair cells in the mammalian cochlea. *J Physiol* 284:261–290.
- Russell IJ, Sellick PM (1983) Low-frequency characteristics of intracellularly recorded receptor potentials in guinea-pig cochlear hair cells. *J Physiol* 338:179–206.
- Russell IJ, Cody AR, Richardson GP (1986a) The responses of inner and outer hair cells in the basal turn of the guinea-pig cochlea and in the mouse cochlea grown in vitro. *Hear Res* 22:199–216.
- Russell IJ, Richardson GP, Cody AR (1986b) Mechanosensitivity of mammalian auditory hair cells in vitro. *Nature* 321:517–519.
- Rutherford MA, Chapochnikov NM, Moser T (2012) Spike encoding of neurotransmitter release timing by spiral ganglion neurons of the cochlea. *J Neurosci* 32:4773–4789.
- Sachs MB, Young ED (1980) Effects of nonlinearities on speech encoding in the auditory nerve. *J Acoust Soc Am* 68:858–875.
- Salvi RJ, Saunders SS, Powers NL, Boettcher FA (1992) Discharge patterns of cochlear ganglion neurons in the chicken. *J Comp Physiol A* 170:227–241.
- Schnee ME, Lawton DM, Furness DN, Benke TA, Ricci AJ (2005) Auditory hair cell-afferent fiber synapses are specialized to operate at their best frequencies. *Neuron* 47:243–254.
- Siebert WM (1970) Frequency discrimination in the auditory system: place or periodicity mechanisms? *Proc IEEE* 58:723–730.
- Siegel JH (1992) Spontaneous synaptic potentials from afferent terminals in the guinea pig cochlea. *Hear Res* 59:85–92.
- Sokolich WG (1981) Closed sound delivery systems. U.S. Patent 4251686.
- Sumner CJ, Lopez-Poveda EA, O’Mard LP, Meddis R (2002) A revised model of the inner-hair cell and auditory-nerve complex. *J Acoust Soc Am* 111:2178–2188.
- Sumner CJ, Lopez-Poveda EA, O’Mard LP, Meddis R (2003) Adaptation in a revised inner-hair cell model. *J Acoust Soc Am* 113:893–901.
- Temchin AN, Ruggero MA (2010) Phase-locked responses to tones of chinchilla auditory nerve fibers: implications for apical cochlear mechanics. *J Assoc Res Otolaryngol* 11:297–318.
- van Hemmen JL (2013) Vector strength after Goldberg, Brown, and von Mises: biological and mathematical perspectives. *Biol Cybern* 107:385–396.
- Versteegh CPC, Meenderink SWF, van der Heijden M (2011) Response characteristics in the apex of the gerbil cochlea studied through auditory nerve recordings. *J Assoc Res Otolaryngol* 12:301–316.
- Vincent PF, Bouleau Y, Safieddine S, Petit C, Dulon D (2014) Exocytotic machineries of vestibular type I and cochlear ribbon synapses display similar intrinsic otoferlin-dependent Ca^{2+} sensitivity but a different coupling to Ca^{2+} channels. *J Neurosci* 34:10853–10869.
- Weiss TF, Rose C (1988) A comparison of synchronization filters in different auditory receptor organs. *Hear Res* 33:175–179.
- Westerman LA, Smith RL (1988) A diffusion model of the transient re-

- sponse of the cochlear inner hair cell synapse. *J Acoust Soc Am* 83:2266–2276.
- Woo J, Miller CA, Abbas PJ (2009) Simulation of the electrically stimulated cochlear neuron: modeling adaptation to trains of electric pulses. *IEEE Trans Biomed Eng* 56:1348–1359.
- Wu JS, Young ED, Glowatzki E (2016) Maturation of spontaneous firing properties after hearing onset in rat auditory nerve fibers: spontaneous rates, refractoriness, and interfiber correlations. *J Neurosci* 36:10584–10597.
- Zeddies DG, Siegel JH (2004) A biophysical model of an inner hair cell. *J Acoust Soc Am* 116:426–441.
- Zhang X, Heinz MG, Bruce IC, Carney LH (2001) A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J Acoust Soc Am* 109:648–670.
- Zhang-Hooks Y, Agarwal A, Mishina M, Bergles DE (2016) NMDA receptors enhance spontaneous activity and promote neuronal survival in the developing cochlea. *Neuron* 89:337–350.
- Zilany MS, Bruce IC, Nelson PC, Carney LH (2009) A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *J Acoust Soc Am* 126:2390–2412.
- Zwislocki JJ (1986) Are nonlinearities observed in firing rates of auditory-nerve afferents reflections of a nonlinear coupling between the tectorial membrane and the organ of corti? *Hear Res* 22:217–221.