UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions

de Boer, E.; Nuttall, A.L.

DOI
10.1121/1.418319

Publication date
1997

Published in
The Journal of the Acoustical Society of America

Link to publication

Citation for published version (APA):
de Boer, E., & Nuttall, A. L. (1997). The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions. *The Journal of the Acoustical Society of America*, *101*(6), 3583-3592. https://doi.org/10.1121/1.418319

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Download date:12 Jun 2024

The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions[a)]

Egbert de Boer[b)]
Room D2-226, Academic Medical Center, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands

Alfred L. Nuttall[c)]
Oregon Hearing Research Center, NRC04, Oregon Health Sciences University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97201-3098 and Kresge Hearing Research Institute, University of Michigan, 1301 E. Ann Street, Ann Arbor, Michigan 48109-0506

(Received 14 March 1996; revised 9 December 1996; accepted 7 February 1997)

The purpose of this investigation is to present evidence from experimental as well as model results on temporal variations of the frequency of oscillation in the basilar membrane's impulse response. Stimuli were either clicks leading to a *direct* estimate of the impulse response, or bands of pseudo-random noise (one or two octaves wide) which lead to an *indirect* estimate of the impulse response via a cross-correlation procedure. The noise bands were centered at the best frequency of the BM location under observation. Responses were obtained from the basal turn of the guinea-pig cochlea, from a location with a best frequency (for the weakest stimuli) between 17.0 and 18.5 kHz. Data acquisition was done with a sample frequency of 208 kHz. Input–output cross-correlation functions were found to share with impulse responses the property that the initial oscillations have a noticeably lower frequency than the later ones. During the impulse response the frequency of oscillation increases gradually. This increase occurs and continues to beyond the time that the oscillations reach the largest amplitude. This frequency variation is called a ''glide.'' Using the ''analytic signal'' method the frequency of oscillation is found to increase continually throughout the duration of the main lobe of oscillation, even at the lowest tested stimulus intensities (about 20 dB SPL). At high stimulus intensity both the direct and indirect impulse response change their appearance drastically but the glide retains its basic form. In the case of the direct impulse response estimate the glide can be attributed to temporal variation of the degree of nonlinearity. For the indirect impulse response this is not true, because with a constant level noise stimulus there is no regular temporal variation of nonlinearity. In this case the glide should be interpreted as an intrinsic property of the cochlear system. From our and others' data the glide was found to exist over a topographic frequency range of best frequencies of at least from 1.76 to 18 kHz. Two examples of present-day models of the cochlea are discussed of which one is found to demonstrate the glide phenomenon in its response, and the other one does not. © *1997 Acoustical Society of America.*
[S0001-4966(97)04506-2]

PACS numbers: 43.64.Kc, 43.64.Bt [RDF]

GENERAL INTRODUCTION

As a biomechanical device, the cochlea shows an impressive degree of frequency selectivity. With the somewhat artificial types of stimulus (such as pure tones) that are usually employed in physiological experiments in the ear, the cochlea shows noticeable nonlinear behavior, and the nonlinearity is also found to depend strongly on frequency. In the actual habitat of animals most acoustical stimuli generally are not of the special types as used in physiological experiments and the same is true for the most important acoustical stimulus for man: speech. Typical ''natural'' sounds and speech signals have been used in physiological experiments on neural coding (starting with Aertsen and Johannesma, 1980; Aertsen *et al.*, 1980, 1981; Sachs and Young, 1979, 1980, and continuing until the present day) but the precise role played by cochlear distortion under such more general conditions has not precisely been analyzed. The same is true for understanding what processes are occurring, physically and physiologically, under such stimulation conditions. In the case of responses to impulse sounds, our understanding of cochlear functioning seems to be only marginally better than for sinusoidal stimuli, but this understanding is only qualitative, to say the least. Because stationary random (noise-like) stimuli form a bridge between sinusoidal and impulse stimuli on the one hand, and natural signals on the other, these signals may be of great value as tools for finding out the relative importance of nonlinear effects in the waveform of basilar-membrane movement.

In the present report data are presented on the mechani-

[a)]Part of this material has been presented at the 1995 Midwinter meeting of the ARO (abstract 747).
[b)]Electronic mail: e.deboer@amc.uva.nl
[c)]Electronic mail: nuttall@ohsu.edu

cal response of one location on the basilar membrane (BM) in the basal turn of the guinea-pig cochlea to various kinds of stimuli, including clicks and random-noise signals comprising a band of frequency components which is much wider than the passband of the location of measurement. The main goal of these experiments is to provide a common framework to describe the nature of linear and nonlinear aspects of the response to such stimuli. The emphasis in this paper is on temporal phenomena, such as impulse responses, but responses to stationary signals are also interpreted in the form of equivalent impulse responses. Thus, we describe properties of *direct* (from impulsive stimuli) and *indirect* (from noise stimuli) mechanical impulse responses of the cochlea.

In both direct and indirect estimates of the impulse response the initial oscillations are found to have a noticeably lower frequency than the later ones, and the frequency of oscillation gradually increases with time. This increase occurs and continues during and past the time that the oscillations reach the largest amplitude. We have termed this frequency variation a ''glide.'' We found this phenomenon at all levels of stimulation used. Robles *et al.* (1976) show evidence of it in (high-level) impulse responses measured with the Mössbauer method in the squirrel monkey. A similar phenomenon characterizes results described by Møller (1983). We quantified the glide by way of the ''analytic-signal'' technique. At high stimulus intensity the impulse response changes its appearance significantly; its oscillations appear to start earlier and to rise faster in amplitude. Since we can quantify the glide only in the period where the oscillations are maximal the glide appears to be shorter than at low stimulus levels. In the case of the direct impulse response estimate the glide can be attributed to temporal variation of the degree of nonlinearity. However, for the indirect impulse response this is not true, because with noise stimulation there is no regular temporal variation of nonlinearity. Therefore, the glide should be interpreted as an intrinsic property of the cochlear system. We collected evidence about the same phenomenon in data from experiments performed by other authors.

The scope of this paper does not permit us to discuss theoretical aspects of the origin of the glide too deeply. However, we briefly discuss two examples of present-day models of the cochlea of which one demonstrates the glide phenomenon, and the other one does not. This finding indicates that there is currently little or no theoretical basis for understanding the glide phenomenon.

I. DIRECT AND INDIRECT IMPULSE RESPONSES

In low-level stimulus situations the relative amount of cochlear nonlinearity is small. In those cases the mechanical transduction of the cochlea can be adequately described by an *impulse response function* (as if the system were linear). This function can be measured *directly*, i.e., by presenting an impulsive acoustical stimulus and recording the mechanical response, or *indirectly*, by presenting a wide-band noise stimulus and computing the cross-correlation function between stimulus and response signals (see Appendix A). Theoretically, when the system is linear or approximately linear, both methods lead to the same result. In the case of the cochlea, this is true when the peak acoustical stimulus level is below approximately 40 dB SPL. At higher levels the two estimates of the impulse response start to differ because nonlinear effects are more pronounced and express themselves differently in the impulse response than in the cross-correlation function.

With intense impulsive stimuli, the cochlea is partly saturated (overloaded) in the first part of the response, and thus has a lower best frequency (BF); later on, when much of the initial energy has been dissipated, the response gradually becomes that of the nonsaturated cochlea—which is tuned to a higher frequency. If this interpretation is correct, the directly measured impulse response reflects a degree of nonlinearity that varies with time. On the other hand, the indirect impulse response is a good descriptor of cochlear transduction (a kind of ''system function'') in cases where the ear is stimulated with stationary wide-band, noise-like, signals, and this is true for all levels of stimulation. In essence, this statement implies that with stimulation by random noise the system is linearized. What this means is formalized in the EQ-NL theorem (de Boer, 1997) which relates the indirect impulse response of a nonlinear cochlear model to the (indirect) impulse response of a linear model. We will come back to this topic later (Sec. IV A) and present a brief description. As a result of the application of this theorem, the indirect method of determining the impulse response leads to an interpretation in terms of an equivalent linear system even though the actual system (the cochlea) is nonlinear and the degree of nonlinearity is randomly fluctuating in time. In the present paper we describe properties of direct and indirect impulse response functions as they have been measured mechanically in the basal turn of the guinea-pig cochlea. We will concentrate on low-level phenomena and, through the use of the aforementioned EQ-NL theorem, interpret what happens at higher levels of stimulation where the cochlea is nonlinear.

II. DATA FROM CURRENT EXPERIMENTS

A. Experimental method

Recordings were made of the movements of the basilar membrane (BM) in the cochlea of anesthetized guinea pigs. This study was consistent with NIH guidelines for humane treatment of animals and was reviewed and approved by the University of Michigan Committee on Use and Care of Animals. Pigmented guinea pigs were surgically prepared to allow visual access to the scala tympani surface of the basilar membrane in the basal or first turn of the cochlea. A ventral and postauricular dissection exposed the cochlea leaving the middle ear ossicles and tympanic membrane intact. A small opening was made in the bony wall of the cochlea of the scala tympani in the first turn and gold-coated glass microbeads (diameter smaller than 20 μm) were deposited onto the scala tympani surface of the basilar membrane. The beam of a laser Doppler vibrometer (OFV 1102) was directed through a microscope and focused onto a glass bead, the movement of which registers the basilar membrane motion (Nuttall

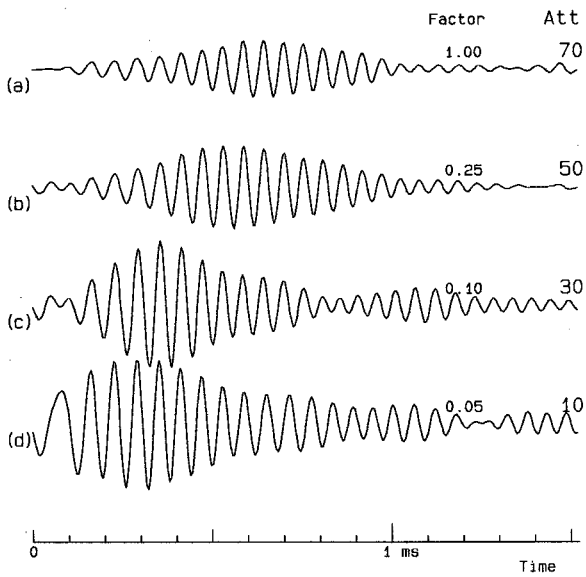

FIG. 1. Waveforms of impulse responses of the basilar membrane. Experiment: 0620 (guinea pig), file: #023. Impulse duration: 24 μs (5 sample periods). Intensity levels indicated by attenuation value (''Att'') with respect to 100 dB peak SPL. Response waveform shown over 1.52 ms. Signal values are plotted multiplied by the ''Factor'' indicated in the figure. Signals are corrected for the stapes response. (a) Att=70, a weak stimulus; (b) Att=50, some deviant behavior in the beginning; (c) Att=30, the initial part of the waveform shows a more clearly different behavior; (d) Att=10, initial oscillations are widely different from later ones.

et al., 1990, 1991). During recording a thin glass plate over the hole served to correct for optical distortion and to minimize vibration of the fluid surface.

Signal generation and data acquisition were done with our own custom-designed programming system, which operates with a clock frequency of 208 kHz. The system and our method of data acquisition are briefly described in Appendix B.

The total duration of an experiment with an ''intact'' (i.e., optimally functioning) cochlea varied between 4 and 13 h. After the experiment, the animal was sacrificed and a number of records was taken from the ''dead'' cochlea. Finally, the laser beam was directed at the stapes, and frequency response and impulse response were carefully measured again from a glass bead (diameter 100 μm) placed on the stapes footplate. In the present paper all our BM response data are corrected for the stapes response. This implies that our data are corrected for transducer response and middle ear transmission. Omitting this correction does not alter the nature of our results as described here because the stapes response is wide-band.

The programming language for experiments and analysis is Turbo Pascal® (version 7.0). Programming of data handling, experiments and analysis has mainly been done by the first author; the actual experiments have been carried out in the laboratory of the second author. Final analysis of the data was done in close collaboration.

B. Typical results—Direct impulse response

Figure 1 shows four direct impulse responses (BM velocity records), obtained at various stimulus intensities. These data are obtained from a cochlea that is (as far as we can judge) functioning optimally. For weak stimuli the best frequency (BF) of the cochlear location studied is 17.9 kHz. The stimulus signal was a pulse, 5 sample periods long (approximately 24 μs), which was repeated with a period of approximately 19.7 ms. It should be noted that in our figures stimulus intensity is always denoted by its complement, the attenuation (''Att'') with respect to the level of 100 dB. At 0-dB attenuation the sound pressure during the pulse is equal to the peak pressure of a sinusoidal signal of approximately 105 dB SPL. Some more details of data processing are given in Appendix B.

The figure shows the impulse response over most of its main lobe of oscillations. The response continues after the period shown, and has, in some of our animals tested, a few later lobes. We do not discuss these later lobes here. By choosing different multiplication factors (''Factor'' in the figure) we show the waveforms as having comparable amplitudes. One of the curves (panel a) has been recorded at such a low intensity that we may expect that nonlinear effects are minimal. In the other three (b, c, and d) the stimulus intensity increased in steps of 20 dB, and cochlear nonlinear effects are observable. With increasing intensity the impulse response appears to start earlier and earlier, and has a character quite different from the response at low stimulation levels. We also observe that the response appears to start abruptly at time zero. This phenomenon is due to the windowing that we applied in the frequency domain which is the same for all records (and rather severe, see Appendix B). At the higher stimulus intensities the bandwidth of the response is larger and the windowing has more effect. We consider as more important that *in the initial part of the response the frequency of oscillation is lower than in the remaining part.* A simple explanation would be one in terms of a temporally varying degree of nonlinearity (see Sec. I). To study this feature more accurately, we have applied a special method to isolate the ''instantaneous frequency of oscillation''; the technique used is based on the concept of the ''analytic signal'' and is briefly described in Appendix C. Figure 2 shows three of the four waveforms from Fig. 1, with plots of the instantaneous frequency (IF) added. The IF scale is linear but covers only a relatively small frequency region, namely, from 15 to 20 kHz. The IF is only drawn where the amplitude of the impulse response is less than 12 dB below the maximum amplitude, this is done because noise in the data prevents us from deriving a reliable IF in the periods with much smaller amplitudes. The same sensitivity to noise of the IF method shows up in the form of irregularities in the IF curves (in this as well as in later figures). The reader's attention is directed at the gradual course of the IF. It is seen that in the high-level record (c), the oscillations start with a low frequency and the IF increases throughout the main part of the response. What is particularly striking: the same is the case in the *low-level* recording (a). The IF increase appears to continue until after the amplitude of the waveform starts to decrease. In this, the low stimulus-level case, we have no reason to assume that the cochlea would react in a nonlinear way during the initial part of the response. We will refer to this phenomenon, the gradual increase of the IF during the

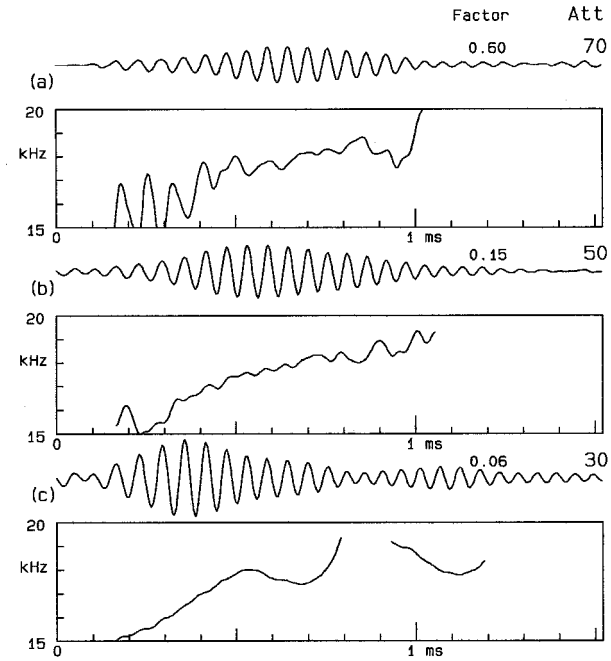

FIG. 2. Waveforms and instantaneous frequencies (IFs) of impulse responses of the basilar membrane. Experimental conditions the same as in Fig. 1. Below each waveform is shown the IF curve enclosed in a rectangle. The linear IF scale goes from 15 to 20 kHz. The time scale is indicated on the abscissa. The IF is only drawn where the amplitude is less than 12 dB below the maximum amplitude. Panels (a) to (c) as in Fig. 1; panel (d) is omitted.

main part of the impulse-response waveform, as the ''glide,'' and note that it appears to be present over a wide range of stimulus intensities. Figure 3 shows other examples of direct impulse responses and records of the IF for low-level

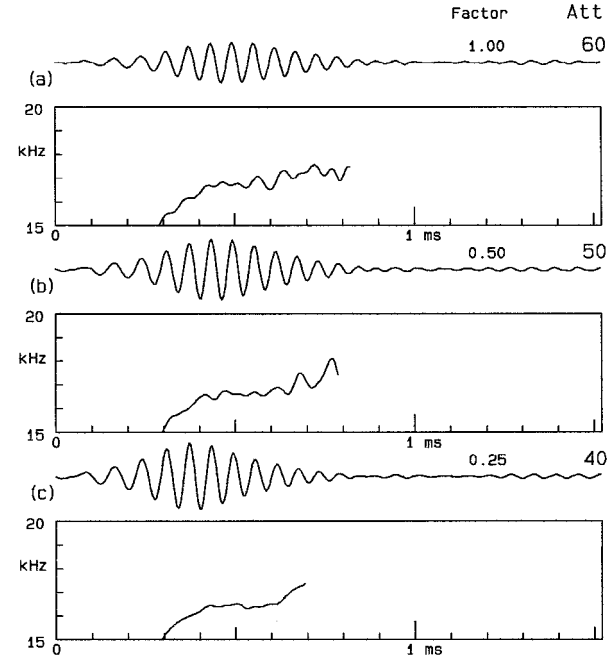

FIG. 3. Waveforms and instantaneous frequencies (IFs) of impulse responses of the basilar membrane; low-level case. Experiment: 0621 (guinea pig), file: #009. General layout as in Fig. 2.

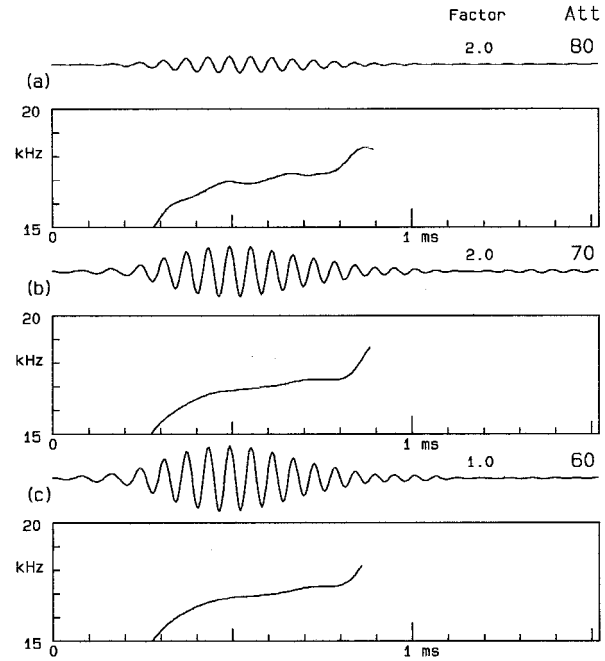

FIG. 4. Cross-correlation functions (ccfs) and instantaneous frequencies (IFs) of the basilar membrane. Experiment: 0621 (guinea pig), file: #001, optimally functioning cochlea, very weak stimuli. Stimulus: a band of noise two octaves wide, centered at the BF (18 kHz). ''Att'' values indicate attenuation in dB with respect to 95 dB over-all SPL of the noise stimulus. General layout as in Fig. 2.

stimuli, for another animal. In all three recordings we find a glide to be present, and the frequency of oscillations continues to increase throughout the main lobe of the response.

C. Typical results—*Indirect* impulse response (ccf)

We have complemented our findings with measurements of the indirect impulse response. That is, we applied a stimulus consisting of stationary wide-band noise, centered at the best frequency of the location studied, and computed the cross-correlation function (ccf) between stimulus and response signals. We believe that at low levels the cochlea operates as a linear system. Then, the ccf measured with wide-band noise should be indistinguishable from the system's impulse response (Appendix A), so it comes as no surprise that ccfs show the glide phenomenon too. Figure 4 shows three examples, from records measured at very low intensities in the same animal as used for Fig. 3. Figure 5 shows similar results, from a different animal, but (in one record) at a still lower level of stimulation. In all panels of this figure we observe the ''glide'' clearly.

From the figures we see that, when the cochlea is strongly stimulated, the impulse response appears to start earlier than when the stimulus is weak. It may be asked whether the glide that we observe with weak stimuli can be considered as an *extrapolation* of the glide in the response with strong stimuli (or, even of the dead cochlea). To study this subject deeper, we present Fig. 6 showing results over a very wide range (80 dB!) of intensity levels. Note that the IF frequency scale has been made wider, from 10 to 20 kHz. In all panels the IF function for the 70-dB attenuation response is shown by a dashed line, and this serves as a guide. In

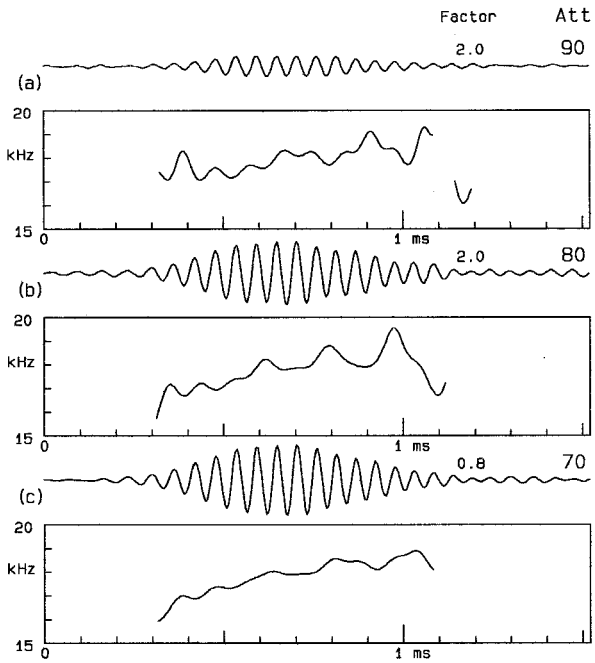

FIG. 5. Cross-correlation functions (ccfs) and instantaneous frequencies (IFs) of the basilar membrane. Experiment: 0620 (guinea pig), file: #019, optimally functioning cochlea, very weak stimuli. Layout as Fig. 4.

panel (c), the high-level stimulus case, the IF is seen to rise from 10 kHz, almost immediately after the oscillations start. The IF curve is soon interrupted because the amplitude of the waveform starts to decrease early in this case. In the other panels, referring to weaker stimuli, the IF curve starts later because the waveform amplitudes rise only slowly to their

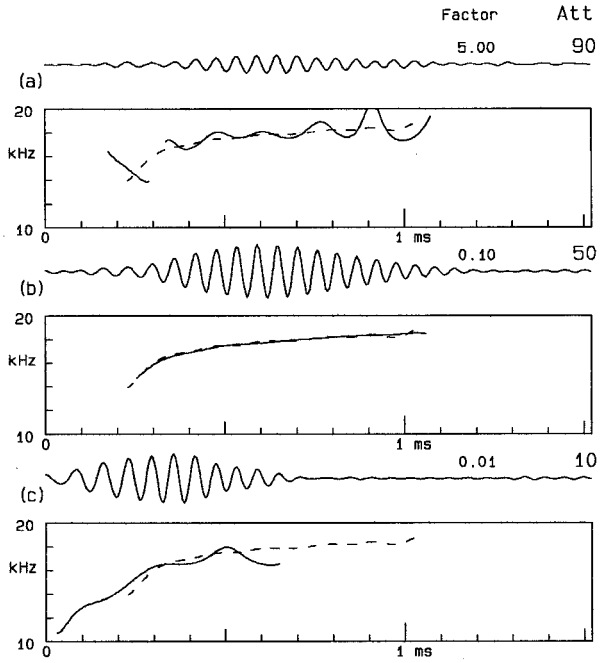

FIG. 6. Cross-correlation functions (ccfs) and instantaneous frequencies (IFs) of the basilar membrane. Experiment: 0628 (guinea pig), file: #035. This figure illustrates the effect of stimulus intensity over a very wide range. Layout as Fig. 4. The dashed line shows the IF function for the 70-dB attenuation condition.

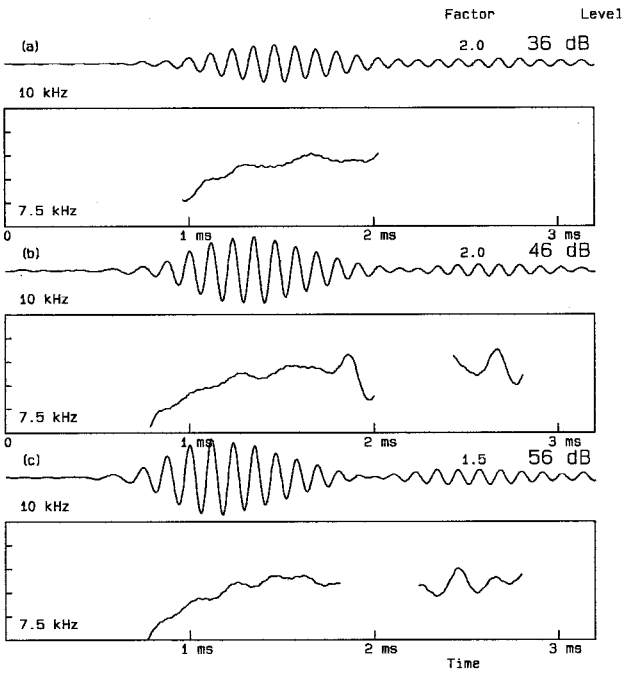

FIG. 7. Impulse response and IF curve for a location in the basal turn of the chinchilla cochlea (Courtesy Mario Ruggero). Layout as in Fig. 2, note that IF scale encompasses lower frequencies. The dB values are peak-SPL values, see the reference. At the lowest level the best frequency is 9.0 kHz.

maximum amplitude. The rise of the IF is continued to beyond the time of the maximum amplitude. In fact, the rise of the IF continues over nearly the entire time that the response is observable, its slope decreasing slowly. Closely related to the similarity of the IF functions is the property that the phase of the oscillations does not depend much on stimulus level. In point of fact, the individual lobes of the ccf waveform, considered for different stimulus levels, line up rather neatly. This is a most interesting and intriguing subject by itself which will be a topic of study in later work.

III. DATA FROM OTHER EXPERIMENTS

A. Mechanical measurements

The authors have been provided with data from experiments performed by colleagues, and these data have been processed with the aim of trying to discover similar trends in equivalent impulse responses. In one set of data, direct impulse responses in the basal turn of the chinchilla cochlea (Ruggero *et al.*, 1992, courtesy Mario Ruggero), this was particularly easy. We applied high-pass filtering with a cut-off frequency of 3 kHz, and low-pass filtering with a cut-off frequency of 10 kHz to a subset of his data. Figure 7 shows a typical example of low-level impulse responses. It is clear that a similar type of ''glide'' is present in the response but note that the location of measurement has a best frequency (BF) of 9.0 kHz, *about one octave lower* than in our earlier figures (for the guinea pig).

We have also processed data from mechanical frequency-response measurements in the basal turn of the guinea pig from another laboratory (Sellick *et al.*, 1983). These data were interpolated and converted to reasonably

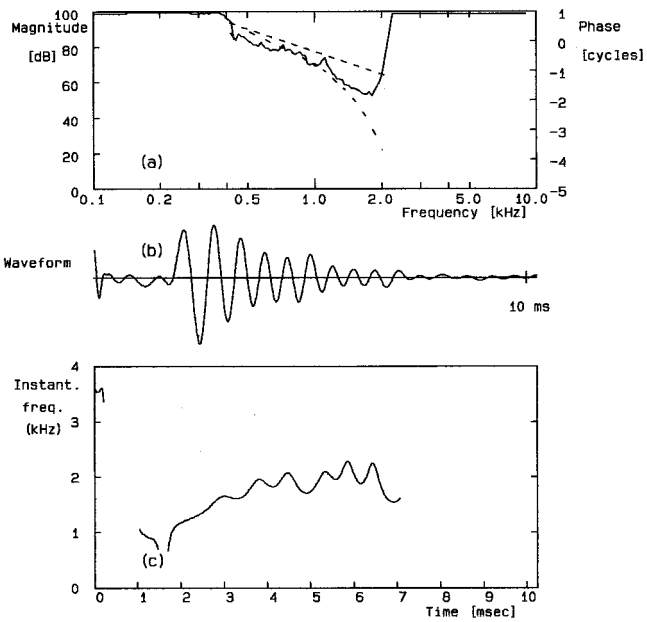

FIG. 8. Frequency threshold curve for neural response (cat data, courtesy Jont Allen/Paul Fahey) converted into impulse response. Unit: 9101, CF: 1.76 kHz, clock period used: 20 μs, corrected for a delay of 2.0 ms. Panel (a): intensity and phase versus frequency. Solid line: frequency-threshold curve. The straight dashed straight line indicates the trajectory along which the phase data are taken. The curved dashed line shows the phase (i.e., the average phase of firing). Panel (b): reconstructed impulse response. Panel (c): instantaneous frequency versus time. Note: ordinate scale is linear and is centered near the CF of the unit.

smooth functions of frequency by a procedure described by Diependaal *et al.* (1987a). A simple Fourier transform was then applied to obtain a time response. The response functions were put at our disposition (courtesy Rob Diependaal) and processed. The results for nearly all of the data were entirely comparable to the ones we presented in Figs. 1 to 3, and thus we refrain from showing results.

B. Neural data

Data for mechanical BM responses in regions of lower frequencies would be most useful to study. However, suitable data are not available in mechanical form, therefore, we have to resort to other sources. In particular, we can draw information from recordings of neural activity. The first example comes from neural ''revcor functions'' (reverse-correlation functions). The data were collected in the first author's laboratory by A. J. Breed (other, related, results of this work have been reported in Breed *et al.*, 1992, and de Boer and Breed, 1993). The experimental animal was the Mongolian gerbil, and recordings were made of single fibers of the auditory nerve. The findings corresponded fully to the results reported by Møller (1983), and we may conclude that the glide phenomenon persists at least down to the frequency range of 3 kHz.

Other neural data come from experiments (Allen, 1983) on frequency threshold curves (tuning curves) and response phase. The data were put at the authors' disposal (courtesy Jont Allen and Paul Fahey), in the form of amplitude and phase records, and converted into equivalent impulse responses. Characteristic frequencies (CFs) for these data varied from below 1 kHz to over 4 kHz (at higher frequencies no reliable phase data could be obtained). Appendix D gives some details about data processing.

In the interest of brevity we present only one figure. Figure 8 shows results for a primary fiber with a rather low CF (1.76 kHz). Panel (a) shows threshold level versus frequency, in the form of a solid line, and the phase as a *curved* dashed line. Phase was measured after the frequency threshold curve was determined; the *straight* dashed line shows the course of level versus frequency used in measuring the phase. Note that the phase data are taken at levels near and just above the threshold of the neural response. Panel (b) of the figure shows the reconstructed impulse response, and panel (c) the IF. The IF scale is again linear; it is centered at the CF but it encompasses the full range from zero. Again, we observe a clear glide.

IV. INTERPRETATION AND EVALUATION

A. Interpretation, the EQ-NL theorem

In the case of the direct impulse response a possible interpretation of the glide at high levels involves the notion that the degree of cochlear nonlinearity varies with time (see Sec. I). The same interpretation of a temporally varying degree of nonlinearity might be used in the case of the indirect impulse response (measured as a ccf) but the variations are random. However, this does not lead to more insight because the variations are difficult to predict quantitatively. There is another, and possibly better, interpretation. The word ''interpretation'' is used here in the sense that we consider a model of the cochlea and that we explain the observed phenomena in terms of phenomena occurring in that model. Let us consider a cochlear model that is linear as far as the fluid and the basilar-membrane proper are concerned, but that contains outer hair cells (OHCs) which act as nonlinear and compressive transducers. Without the contribution of the OHCs the response amplitude at a certain location x has a peak at a certain frequency f_0. The notion ''frequency selectivity'' expresses the fact that the response (at the same place and for a constant stimulus level) is smaller at frequencies away from f_0. At low stimulation levels the OHCs function to enhance frequency selectivity, i.e., to make the response variations around the peak frequency f_0 larger. OHCs are modeled to do this by producing a force (or pressure) that is an instantaneous function of the displacement of their stereocilia. This type of model is usually called ''active'' but a better term is ''locally active'' because at each frequency the response-enhancing activity of the OHCs is only needed over a limited region of the BM, basal to the location of the response peak (see de Boer, 1991). With strong stimulation the OHCs tend to become saturated and thus will contribute less to frequency selectivity. Call this nonlinear model M_1.

For this type of nonlinear model of the cochlea a theorem has been derived, called the EQ-NL theorem (de Boer, 1997), which describes a property of the input–output ccf for wide-band signals. Explicitly, the theorem states that *the ccf for the nonlinear model* M_1 *is equal to the ccf for a comparison model* M_2 *that has exactly the same structure and the same parameters with the exception that it contains linear*

OHC transducers. The second part of the theorem states that *in the comparison model* M_2 *the efficiency of all OHCs is reduced by a constant factor, which is the same for all OHCs.*

Very briefly, the idea behind the theorem is the following. Consider each outer hair cell (OHC) as a memoryless nonlinear device embedded in a linear system (all frequency dependence of hair-cell transduction being transported to the ''outside world,'' i.e., the linear parts of the model). With a noise signal as input, the average compression of signal components by the nonlinear device is the same for all frequencies. In this sense the OHC transducer can thus be represented by a linear transducer with a frequency-independent attenuation.

When the nonlinear model is stimulated with strong sounds, the transduction of the OHCs will saturate (partly), and the extra force or pressure produced by them is reduced in a relative sense. As a consequence, frequency selectivity deteriorates. The same reduced selectivity will be demonstrated by the linear comparison model M_2 and this will also show up in its impulse response. In terms of the EQ-NL theorem, the ccf of the nonlinear system M_1 is equal to the impulse response of the linear comparison system M_2. The glide can thus always be interpreted as a property of a *linear* system, and this is true at all levels of stimulation used.

In an extreme case, where the OHCs are entirely put out of action, we would observe the impulse response of a model simulating the ''passive'' (i.e., nonactive) cochlea. It is known that the impulse response of such a passive model has an intrinsic frequency modulation, a glide. See for the case of a passive long-wave model de Boer (1980), Figure 6.9. Let us now return from the passive case to the case of the normally functioning, active, cochlea, and do this in a gradual manner. With increasing degree of activity, the OHCs come into play to improve frequency selectivity more and more. Apparently, *this process occurs in such a way that the glide character is maintained.* In other terms, the impulse response is not swamped by the newly occurring, sharply frequency-selective, action of the OHCs. Nor is the new OHC action so purely that of a simple resonator that it dominates the response in all respects (a single resonator does not show a glide). What is even more surprising is that the course of the IF is nearly independent of stimulus level. In terms of the comparision model M_2, the IF function is nearly independent of the efficiency of the OHCs.

B. Two locally-active cochlear models

We will next illustrate to which degree the glide is demonstrated by the response of two current locally active cochlear models. Both models are linear, and operate with the long-wave approximation. If the models were made nonlinear by assuming that OHCs are nonlinear transducers, both models would belong to the class of models to which the EQ-NL theorem (de Boer, 1997) applies.

We first choose a member of the class of *classical* models, a class characterized by the assumption that mechanical longitudinal coupling in the BM is absent and that longitudinal coupling occurs only via the cochlear fluid (cf. Viergever, 1986). In the Neely–Kim model (Neely and Kim, 1986) OHCs are assumed to contribute to the sound pressure in the cochlear channels. In order to achieve realistic local activity each OHC is assumed to be stimulated by movements that undergo extra filtering. The filtering is due to a resonance assumed to be exhibited by the tectorial membrane (TM), and the local TM resonance frequency is a (more or less fixed) fraction of the local resonance frequency of the BM. The operation of the OHCs is controlled by a coefficient γ that can be chosen between 0 and 1; $\gamma=0$ yields a passive model (no OHC action) and $\gamma=1$ produces a model of which the response is well comparable with experimental findings at low stimulation levels in the intact cochlea. The Neely–Kim model (note that it is formulated as a long-wave model) was set up with the parameters as published by Neely and Kim (1986), except that the height of the model was chosen as 0.4 (cm)—this was done in an effort to simulate that in a three-dimensional model the BM occupies only a fraction of the channel width. The impulse response was calculated with the robust solution method described by Diependaal *et al.* (1987b). Instead of an impulse, a doublet pulse (two consecutive time samples were given the values $+1$ and -1) was used as the stimulus signal; this serves to depress persisting low-frequency components in the response. Hence, it is the time derivative of the impulse response that has been computed. The OHC effectivity coefficient γ was made equal to 0.98 (higher values tend to cause computational instability). Figure 9 shows the result, for the location that has its best frequency (BF) equal to 18.1 kHz. The figure shows the impulse response and the course of the instantaneous frequency (IF) computed with the same method as before. It is seen that a very fast rise of the IF occurs in the initial period where the oscillations are too small to be visible (in this figure we can show the IF curve over a much wider range of amplitude values because the noise in the impulse response is very small). In the period where we observe the oscillations to rise in amplitude toward their maximum, the IF is seen to reach a maximum and to decrease—in other words, there is an *inverted* glide. In the final phase of the impulse response, the IF function tends to

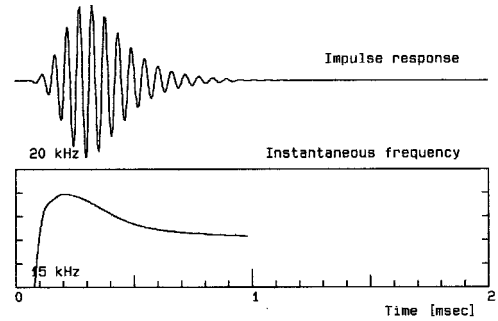

FIG. 9. Impulse response (upper panel) and IF curve (lower panel) of the Neely-Kim model. Model parameters as in Neely and Kim (1986); the effective channel height (channel area divided by BM width) was made equal to 0.4 (cm). Length of model divided into 500 sections. Time step: 2 μs. OHC effectivity coefficient γ: 0.98. The IF curve is shown over the period where the signal is less than 40 dB below its maximal value. Average best frequency (BF) of impulse response at location selected: 18.1 kHz, BM resonance frequency at this location: 39.6 kHz.

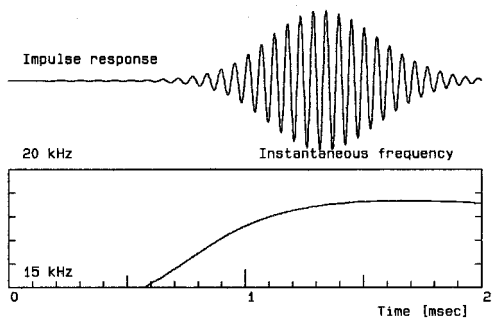

FIG. 10. Impulse response (upper panel) and IF curve (lower panel) of Geisler and Sang's feed-forward model (Geisler and Sang, 1995), TM resonance included. Model parameters as in the original paper, except for the effective channel height which was taken as 0.4 (cm) and the length of the model which was shortened to 1.5 (cm). Length of model divided into 250 sections. OHC effectivity coefficient γ: 0.45. The IF curve is shown over the period where the signal is more than 0.03 times its maximal value. Average best frequency (BF) of impulse response: 17.5 kHz.

settle at a constant value. We found similar behavior when the efficiency coefficient γ was given a lower value (for instance, 0.75, which corresponds to a loss of sensitivity in the response-peak region of approximately 30 dB). Note that, compared to the data, the group delay of this model is (much) too small, this can partly be attributed to our choice of the effective height. We conclude that the parameters of the Neely–Kim model are such that (at least for $\gamma > 0.75$) the impulse response does not show a glide as our (and others') data do. We do not know whether the model would perform better in this respect with different parameters.

The second model is the *nonclassical* model described by Geisler and Sang (1995) and conceived as a ''feed-forward'' model. In this model the OHCs receive their input from a location that is slightly more basal than the point where they are located. With this expedient the model shows many realistic properties. We used the form of the model with a secondary resonance in the TM (similar to that in the Neely–Kim model) to produce Fig. 10. Again, we chose the parameters as used by the authors but made the height equal to 0.4 (cm). We selected the efficiency factor so as to simulate the response of the intact cochlea at very low stimulation levels. However, we used a different (in fact, simpler) method for computing the impulse response because it is not known whether the Diependaal *et al.* (1987b) method remains robust when modified for nonclassical models. We computed the model response for 128 frequencies from 0 to 30 kHz, and stored the velocity response at a selected location of the BM in an array. Next we performed a Fourier transform on that array. The result is shown in Fig. 10, for the location that has 17.5 kHz as its best frequency (BF). Clearly, a very realistic glide is present, and we found it to change little in appearance when the efficiency of the OHCs in the model was reduced in order to simulate conditions where the cochlea is stimulated by stronger stimuli. Note that, in view of our data, the group delay of this model is too large (see comments in Geisler and Sang, 1995).

The glide phenomenon has been shown to be a fundamental property of data obtained from the cochlea. As such, we propose that it becomes a standard to which models of the cochlea should be compared. It is interesting that the linear, passive long-wave model (de Boer, 1980, Figure 6.9) exhibits the glide while a later, locally active long-wave model, like the Neely and Kim (1986) model, fails the glide validation test (this paper, Fig. 9). Clearly, Geisler and Sang's (1995) model passes the proposed test (Fig. 10), although one cannot overlook the unreasonably long delay that is characteristic of this model. Both locally active models contain features that are not realistic, for instance, the assumed mass of the BM is too small. In view of our experimental and modeling results it appears that further research in the area of the glide could substantially enhance our understanding of the way the cochlea works.

ACKNOWLEDGMENTS

The authors wish to thank (in alphabetical order) Jont Allen, Arjan Breed, Rob Diependaal, Paul Fahey, Dan Geisler, Karl Grosh, Luc Johan Kanis, Dick Lyon, Bob Masta, Darren Miller and Mario Ruggero for their contributions to the present report. In addition, two anonymous reviewers have contributed much to the final presentation. This work has been supported by the Netherlands Foundation for Scientific Research (NWO), No. SLW 01.011, and by NIH, No. NIDCD-DC-00141.

APPENDIX A: AUTO- AND CROSS-CORRELATION FUNCTIONS

Consider a linear and time-invariant system H with impulse response $h(t)$ and apply an input signal $x(t)$ in the form of a white-noise signal with a power of unity. The *cross-correlation function* (abbreviated: ccf) $\varphi(t)$ between input signal $x(t)$ and output signal $y(t)$ is defined as

$$\varphi(\tau) = \mathscr{E}\{x(t+\tau)y(t)\}, \tag{A1}$$

where the symbol $\mathscr{E}\{.\}$ stands for ''ensemble average'' (or ''expectation value''). This means an average over the ensemble of $x(t)$ signals having the same (given) probability distribution and the associated response signals $y(t)$. Instead of this procedure, averaging is usually carried out over time. Denoting the averaging time by T, Eq. (A1) then reads:

$$\varphi(\tau) = \int_0^T x(t+\tau)y(t)(dt/T) \qquad (\text{for } T \to \infty). \tag{A2}$$

When the signals $x(t)$ and $y(t)$ refer to the same signal, the ccf is known as the *auto-correlation function* (acf).

For the system H under consideration the following relation holds between the ccf and the impulse response:

$$\varphi(\tau) = h(-\tau). \tag{A3}$$

In this paper Eq. (A3) is used to relate direct and indirect impulse responses to one another. This relation can easily be derived by using the convolution relation between $x(t)$ and $y(t)$:

$$y(t) = \int_0^\infty x(t-\tau)h(\tau)d\tau. \tag{A4}$$

Relation (A3) is strictly true only for stimulation with white noise having unity power. When the input signal is not white noise, Eq. (A3) receives a correction including the acf of $x(t)$. However, when the bandwidth of the noise is considerably larger than that associated with $h(t)$, the correction is unimportant. This is the case for the low-level records used in the present paper. It is not entirely true for the high-level records and this results in visible distortion of the ccfs.

APPENDIX B: PROCESSING MECHANICAL RESPONSES

Our custom-designed programming system utilizes A/D and D/A converters located on PC boards (Computer Boards, type CIO-DAS 16/330i and cSBX-DDA4, respectively). The computer used was a Gateway 2000 (Intel 486, 50 MHz). Both D/A and D/A converters have a resolution of 12 bits, and operate with a clock frequency of 208 kHz. The stimulus signal consists of either pulses or a pseudo-random noise signal; one full period of the waveform is stored in a 4096-word array in memory (the ''input-signal array'') and is sequentially read out to the D/A converter. The array is used as a *circular buffer*: When the end of the array is reached, the next sample is taken from the beginning, and so on. Consequently, the periods of the stimulus are presented consecutively, without any gaps or irregularities. Synchronously with the D/A conversions the signal from the laser velocimeter's output amplifier is sampled by the A/D converter. The resulting numbers are stored in a temporary array. At regular intervals (synchronously with the periods of the stimulus) the contents of this array is added to the contents of another array, called the ''output-signal array,'' and a count (N) is maintained of the number of additions. In this way the output-signal array will, in due course, contain N times the average recorded BM velocity. Of course, this array is also interpreted as a circular buffer. In this paper we speak of ''the response'' as the average BM velocity recorded in this way, averaged after many (more than 1000) repetitions.

The D/A and A/D conversions occur concurrently and synchronously, both at a rate of 208 k samples per second. Since the arrays are 4096 samples long the duration of one period is 19.7 msec and the frequency components of the signal generated will be spaced by approx. 50.8 Hz. In the experiments the number N of repetitions is equal to or larger than 1000, hence each record requires a measuring time of at least 20 seconds.

The signal generated by the D/A converter is passed via a computer-controlled attenuator (Wilsonics PATT) and an amplifier to a condenser microphone (B&K type 4134, 0.5″ diameter) serving as a loudspeaker. A small acoustic coupler ensures optimal coupling to the ear drum membrane of the animal under test. No attempt has been made to compensate for the inherent quadratic distortion of a condenser microphone used as a loudspeaker. We checked that the quadratic distortion of the loudspeaker was smaller than -40 dB with respect to the primary level at the highest stimulus levels used, and we consider this small enough so as not to influence the results. Cubic distortion was smaller.

Experimental data have been acquired in arrays of 4096 long-integer words but are converted into (single-precision) real variables for analysis purposes. Cross-correlation functions (ccfs) are computed via the digital Fourier transform of Eq. (A1) with the sign of τ inverted:

$$\Phi(\omega)=\mathscr{E}\{X(\omega)^*Y(\omega)\}, \tag{B1}$$

where the asterisk denotes the complex conjugate. This procedure is justified by the fact that the sampling frequency is well above the Nyquist frequency. Before further processing, low- and high-frequency components are reduced in amplitude. Except where indicated otherwise, the high-pass cut-off frequency is 6 kHz, and the low-pass cut-off frequency 20 kHz; the asymptotic slopes are 48 dB per octave for the direct impulse response and 4 times higher for the indirect impulse response. For the figures in this paper based on our own data, a correction for the stapes response has been applied. This we did via division of the respective spectra. For the direct impulse response (Figs. 1 to 3) we used the impulse response of the stapes. For the indirect impulse response we compensated via the ccf of the stapes response for a band-limited stimulus signal. It is because in this case the stapes ccf is a band-limited signal, and we divided by its spectrum, that we had to use higher cut-off slopes.

APPENDIX C: INSTANTANEOUS FREQUENCY AND THE ''ANALYTIC SIGNAL''

Consider a real signal $x(t)$ which is zero for $t<0$, and of which the Fourier transform $X(\omega)$ exists. The real part of this Fourier transform is an even, and the imaginary part an odd function of ω. The signal $x(t)$ may be an impulse response or a ccf. Because $x(t)$ is one-sided, real and imaginary parts of $X(\omega)$ are Hilbert transforms of one another. Now construct a new, one-sided, spectrum $Z(\omega)$ from $X(\omega)$ obeying the following conditions:

$$\begin{aligned} Z(\omega)&=2X(\omega), && \text{for} \quad \omega>0,\\ Z(\omega)&=0, && \text{for} \quad \omega<0,\\ Z(\omega)&=X(\omega), && \text{for} \quad \omega=0. \end{aligned} \tag{C1}$$

Consider the inverse Fourier transform $z(t)$ of $Z(\omega)$. This function will have real and imaginary parts that are Hilbert transforms of one another because $Z(\omega)$ is one-sided. The real part of $z(t)$ is equal to $x(t)$ since for positive ω its spectrum is the average of $Z(\omega)$ and $Z(-\omega)$. Hence, the imaginary part of $z(t)$, let us call it $x_{\mathrm{Hil}}(t)$, is the Hilbert transform of $x(t)$. The complex signal $z(t)$ is known as the ''analytic signal.'' It can be written as

$$z(t)=x(t)+ix_{\mathrm{Hil}}(t), \tag{C2}$$

or, alternatively, as

$$z(t)=a(t)\cos[\varphi(t)], \tag{C3}$$

where $a(t)$ is the magnitude of $z(t)$ and $\varphi(t)$ the phase. The derivative of $\varphi(t)$ with respect to time t is the instantaneous frequency (IF). The just-outlined procedure is quite general but useful applications exist only for ''narrow-band'' signals, where variations of the magnitude $[a(t)]$ and the IF $[d\varphi(t)/dt]$ with time are slow with respect to the average IF.

This definition provides the path to compute the IF from a given real function of time $x(t)$. First, determine the Fou-

rier transform, apply the transformation given by Eq. (C1), and construct the analytic signal $z(t)$. Then, determine the phase $\varphi(t)$ and compute its time derivative (which means one has to correct for $\pm 2\pi$ phase jumps). This recipe has been followed to arrive at the figures showing the IF as a function of time; the application of the technique to sampled data presents no special problem. Prior to the computation of the IF the signal $x(t)$ (impulse response or ccf) undergoes extra smoothing, see Appendix B.

APPENDIX D: PROCESSING OF TUNING-CURVE DATA

Tuning-curve data come in the form of records of the measurement of the frequency threshold curve and, separately, records of the phase. Both are functions of frequency but the frequencies used are not the same, nor do the frequency values always form a regular series. Therefore, it is necessary to interpolate the data. The raw amplitude data give the series of turning points in the threshold measurement procedure, level and frequency. These data are interpolated in a linear way on a decibel scale. The series of frequencies used in measuring the phase is more regular, and phase-data interpolation is simple (provided appropriate corrections in steps of 2π are applied). Phase data are corrected for a delay as given in the figure legend, that delay includes the synaptic delay and (part of the) traveling-wave delay.

From the amplitude and phase data an array (of length 1024) is filled with the corresponding complex numbers. Finally, the so-obtained frequency response is windowed: components with frequencies lower than 2 octaves below the characteristic frequency (CF) are diminished in amplitude, and the same is done with components more than one octave above the CF. A Fourier transform is all that is needed to obtain the equivalent impulse response.

Aertsen, A. M. H. J., and Johannesma, P. I. M. (**1980**). ''Spectro-temporal receptive fields of auditory neurons in the grassfrog. I. Characterization of tonal and natural stimuli,'' Biol. Cybern. **38,** 223–234.

Aertsen, A. M. H. J., Johannesma, P. I. M., and Hermes, D. J. (**1980**). ''Spectro-temporal receptive fields of auditory neurons in the grassfrog. II. Analysis of the stimulus-event relation for tonal stimuli,'' Biol. Cybern. **38,** 235–248.

Aersten, A. M. H. J., Olders, J. H. J., and Johannesma, P. I. M. (**1981**). ''Spectro-temporal receptive fields of auditory neurons in the grassfrog. III. Analysis of the stimulus-event relation for natural stimuli,'' Biol. Cybern. **39,** 195–209.

Allen, J. B. (**1983**). ''Magnitude and phase-frequency response to single tones in the auditory nerve,'' J. Acoust. Soc. Am. **73,** 2071–2092.

Boer, E. de (**1980**). ''Auditory physics. Physical principles in hearing theory. I,'' Phys. Rep. **62,** 87–174.

Boer, E. de (**1991**). ''Auditory physics. Physical principles in hearing theory. III,'' Phys. Rep. **203**, 125–231.

Boer, E. de (**1997**). ''Connecting frequency selectivity and nonlinearity for models of the cochlea,'' Aud. Neurosci. (in press).

Boer, E. de, and Breed, A. J. (**1993**). ''Encoding of 'nothing' in the peripheral auditory pathway: Nervous activity associated with a spectral gap,'' in *Brain Theory—Spatio-Temporal Aspects of Brain Function*, edited by A. Aertsen (Elsevier, Amsterdam), pp. 67–89.

Breed, A. J., Kanis, L. J., and Boer, E. de (**1992**). ''Cochlear nonlinearity for complex stimuli,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford), pp. 189–195.

Diependaal, R. J., de Boer, E., and Viergever, M. A. (**1987a**). ''Cochlear power flux as an indicator of mechanical activity,'' J. Acoust. Soc. Am. **81,** 184–186.

Diependaal, R. J., Duifhuis, H., Hoogstraten, H. W., and Viergever, M. A. (**1987b**). ''Numerical methods for solving one-dimensional cochlear models in the time domain,'' J. Acoust. Soc. Am. **82,** 1655–1666.

Geisler, C. D., and Sang, C. (**1995**). ''A cochlear model using feed-forward outer-hair-cell forces,'' Hearing Res. **86,** 132–146.

Møller, A. R. (**1983**). *Auditory Physiology* (Academic, New York), pp. 220ff and Fig. 3.13.

Neely, S. T., and Kim, D. O. (**1986**). ''A model for active elements in cochlear biomechanics,'' J. Acoust. Soc. Am. **79,** 1472–1480.

Nuttall, A. L., Dolan, D. F., and Avinash, G. (**1990**). ''Measurements of basilar membrane tuning and distortion with laser doppler velocimetry,'' in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer-Verlag, Berlin), pp. 288–295.

Nuttall, A. L., Dolan, D. F., and Avinash, G. (**1991**). ''Laser Doppler velocimetry of basilar membrane vibration,'' Hearing Res. **51,** 203–214.

Robles, L., Rhode, W. S., and Geisler, C. D. (**1976**). ''Transient response of the basilar membrane measured in squirrel monkeys using the Mössbauer effect,'' J. Acoust. Soc. Am. **59,** 926–939.

Ruggero, M. A., Rich, N. C., and Recio, A. (**1992**). ''Basilar membrane responses to clicks,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, London), pp. 85–91.

Sachs, M. B., and Young, E. D. (**1979**). ''Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate,'' J. Acoust. Soc. Am. **66,** 470–479.

Sachs, M. B., and Young, E. D. (**1980**). ''Effects of nonlinearities on speech encoding in the auditory nerve,'' J. Acoust. Soc. Am. **68,** 858–875.

Sellick, P. M., Yates, G. K., and Patuzzi, R. (**1983**). ''The influence of Mössbauer source size and position on phase and amplitude measurements of the guinea pig basilar membrane,'' Hearing Res. **10,** 101–108.

Viergever, M. A. (**1986**). ''Cochlear macromechanics—a review,'' in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. Hubbard, S. T. Neely, and A. Tubis (Springer-Verlag, Berlin), pp. 63–72.