UvA-DARE (Digital Academic Repository)

The mechanical waveform of the basilar membrane. II. From data to models-and back

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

DOI

[10.1121/1.428435](https://doi.org/10.1121/1.428435)

Publication date

2000

Published in

The Journal of the Acoustical Society of America

[Link to publication](#)

Citation for published version (APA):

de Boer, E., & Nuttall, A. L. (2000). The mechanical waveform of the basilar membrane. II. From data to models-and back. *The Journal of the Acoustical Society of America*, 107, 1487-1496. <https://doi.org/10.1121/1.428435>

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.

The mechanical waveform of the basilar membrane.

II. From data to models—and back

Egbert de Boer^{a)}

Room D2-226, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands

Alfred L. Nuttall^{b)}

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 10 May 1999; revised 1 October 1999; accepted 29 November 1999)

Mechanical responses in the basal turn of the guinea-pig cochlea are measured with low-level broad-band noise as the acoustical stimulus [for details see de Boer and Nuttall, *J. Acoust. Soc. Am.* **101**, 3583–3592 (1997)]. Results are interpreted within the framework of a classical three-dimensional model of the cochlea that belongs to a very wide class of nonlinear models. The use of linear-systems analysis for this class of nonlinear models has been justified earlier [de Boer, *Audit. Neurosci.* **3**, 377–388 (1997)]. The data are subjected to inverse analysis with the aim to recover the “effective basilar-membrane impedance.” This is a parameter function that, when inserted into the model, produces a model response, the “resynthesized” response, that is similar to the measured response. With present-day solution methods, resynthesis leads back to an almost perfect replica of the original response in the spatial domain. It is demonstrated in this paper that this also applies to the response in the frequency domain and in the time domain. This paper further reports details with regard to geometrical properties of the model employed. Two three-dimensional models are studied; one has its dimensions close to that of the real cochlea, the other is a stylized model which has homogeneous geometry over its length. In spite of the geometric differences the recovered impedance functions are very similar. An impedance function computed for one model can be used in resynthesis of the response in the other one, and this leads to global amplitude deviations between original and resynthesized response functions not exceeding 8 dB. Discrepancies are much larger (particularly in the phase) when a two-dimensional model is compared with a three-dimensional model. It is concluded that a stylized three-dimensional model with homogeneous geometric parameters will give sufficient information in further work on unraveling cochlear function via inverse analysis. In all cases of a sensitive cochlea stimulated by a signal with a stimulus level of 50 dB SPL per octave or less, the resulting basilar-membrane impedance is found to be locally active, that is, the impedance function shows a region where the basilar membrane is able to amplify acoustic power or to reduce dissipation of power by the organ of Corti. Finally, the influence of deliberate errors added to the data is discussed in order to judge the accuracy of the results. © 2000 Acoustical Society of America. [S0001-4966(00)02703-X]

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

INTRODUCTION

In the first paper of the present series (de Boer and Nuttall, 1997b-GLIDE) we presented experimental data on the response of the basilar membrane (BM) to click and noise stimuli. We now turn to the theoretical side. Quantitative interpretation of experimental data on the response of the cochlea is only possible within a well-defined *frame of reference*. In our case the proper frame of reference is *a model of the mechanics of the cochlea*, i.e., a conceptional construction describing how different mechanical parts of the cochlea interact and cooperate to achieve that organ's remarkable performance. For the work presented in this paper a model of the cochlea is utilized in which acoustic waves propagate

through fluid channels in close interaction with the mechanics of the BM and its associated structures. The model is briefly described in the Appendix. It is assumed that on its way the wave may be amplified in power by a specific frequency- and place-dependent mechanism [proposed by Kim *et al.* (1980), theoretically founded by de Boer, 1983, and coined as the “cochlear amplifier” by Davis (1983)]. The region where amplification occurs, the region of “activity,” is spatially limited and such a model is called “locally active” (de Boer, 1993, 1996).

Furthermore, it is assumed that the elements of this mechanism of the model are *nonlinear*, and that they are the *only* nonlinear elements of the model. The Appendix describes not only the model we used but also the class of nonlinear models to which it belongs (which encompasses almost all of the models published to date) and indicates under which conditions *linear-systems theory* can be applied

^{a)}Electronic mail: e.deboer@amc.uva.nl

^{b)}Electronic mail: nuttall@ohsu.edu

to a model of this class. The present paper will treat data obtained at low stimulus levels where nonlinear effects are small enough to be neglected. We will use three-dimensional models which support all types of waves, long, intermediate and short waves. In addition, we will briefly study a two-dimensional model. All models will be linear.

We performed physiological–mechanical experiments on the response of the BM using wide bands of flat-spectrum random noise as stimulus signals. The responses have been converted into input–output cross-correlation functions (ccfs). This technique is a prerequisite for interpretation of the data in terms of linear-systems theory (see the Appendix). A central place in our work is occupied by the inverse-solution method that has been described in three earlier papers, de Boer (1995a,b), de Boer and Nuttall (1999), henceforth to be called INV-1, INV-2, and INV-3, respectively. In the inverse procedure the measured response in the frequency domain for a fixed location x is first converted into a hypothetical response in the x domain for a fixed frequency. The inverse procedure yields the “effective BM impedance” $Z_{\text{BM}}(x, \omega)$ of the model, where the independent variable x is the longitudinal coordinate of the model and the parameter ω is the radian frequency. The variables x and ω are assumed to “scale,” i.e., one can be traded for the other using the assumed cochlear map. Furthermore, we computed (“resynthesized”) the response of a model which uses the BM impedance function $Z_{\text{BM}}(x, \omega)$, recovered by inverse analysis, as its BM impedance function.

The inverse-solution method will be applied to various types of models, to be described in Sec. IB. The nature of the results remains the same, only details differ. The most severe test of the integrity of the entire chain of procedures, starting with the spectral response, going via inverse analysis to resynthesis of the spatial response and ending with the response in the frequency domain, is performed by computing the impulse response of the model, and comparing it with the impulse response in the data. The deviations are found to remain within reasonable boundaries, in particular, all the dominant features of the impulse response are preserved. In the first paper of the present series of papers (de Boer and Nuttall, 1997b-GLIDE)—which we will refer to as “Part I”—we reported on a typical frequency modulation in the impulse response of the BM, the “glide.” The resynthesized impulse response is found to include a good representation of the glide. That means that a cochlear model with the function $Z_{\text{BM}}(x, \omega)$ —as recovered by inverse analysis—is capable of simulating the glide. We reported earlier (Part I) that not all current models of the cochlea are able to do that.

I. METHOD AND RESULTS I—FROM THE DATA TO A MODEL

A. Data

We have collected data on movements of the basilar membrane (BM) in the basal turn of the guinea-pig cochlea with a laser velocimeter (cf. Nuttall *et al.*, 1990; Nuttall and Dolan, 1996). The measurement and data processing techniques used in the present work have been described in Part I. For the experiments described here¹ *bands of flat-spectrum*

pseudo-random noise were used as acoustical stimuli. “Stimulus level” is defined as the SPL of one octave of the stimulus signal around the best frequency (BF). We measured the velocity of the BM at a location tuned to a frequency between 16 and 18 kHz as a function of time and computed input–output cross-correlation functions (ccf’s). The accuracy of the spectrum of a ccf usually is sufficient only over a narrow range of frequencies around the BF. In order to obtain a useful ccf spectrum over a wider frequency range, we constructed *composite ccf spectra* by combining ccf spectra obtained with flat-spectrum stimuli of different bandwidths, center frequencies and intensities in the same animal (see INV-3 for details). The composite ccf spectrum should be interpreted as if it had been possible to measure the ccf spectrum in one experiment with constant stimulus power density over the entire frequency range from below 1 kHz to over 25 kHz. The composite ccf spectrum is measured and processed in such a way that it represents *the ratio of BM-to-stapes velocity*. For the description of the modest amount of smoothing applied to the composite spectrum we refer to INV-3, Appendix B. The corresponding ccf waveform is obtained from the composite ccf spectrum via (inverse) Fourier transformation. When the stimulus level is so low that the cochlea is operating like a linear system, we may consider the ccf waveform as an estimate of the impulse response.

B. Models

The underlying model is a member of the class of models described in the Appendix. In this paper we will assume that a linear model is sufficient. For higher stimulus levels a nonlinear model would be needed. However, for a nonlinear model, the composite ccf spectrum can still be interpreted as the input–output ccf spectrum of a *linear* model, namely, the “comparison model” as described in the Appendix. The corresponding ccf waveform then becomes the impulse response of the comparison model.

The model to be used is *three-dimensional* and supports long, intermediate as well as short waves. We will consider two versions of this model. In one, model “3dm,” the cross sections of the channels and the width of the BM all vary with x (the longitudinal coordinate), and we will refer to this model as the “realistic” model. Figure 1 shows the BM width $b_{\text{BM}}(x)$ and the channel cross-sectional areas as functions of the longitudinal coordinate x . We adopted the data for this figure from the paper by Mammano and Nobili (1993) who used data published by Fernández (1952). For our purpose (all relevant frequencies are above 5 kHz) it will be sufficient to include only the most basal part. Therefore, the model will have a length of 6 mm [$L = 6$ (mm)]. In Fig. 1 the corresponding segments of the curves are shown by solid lines. The remaining part of the length (curves drawn by dashed lines) is not included in our model. In the second version of the model, to be denoted by “3ds,” the cross sections of the two channels are the same and constant over the length L of the model. Likewise, the width of the BM (which is the movable part of the partition between the two channels) is constant, it is ε times the width b of the model. To conform with earlier work (INV-1 to INV-3) we made b

FIG. 1. Width of the basilar membrane (BM) and cross-section area of cochlear ducts for the guinea pig. Modified from Fig. 1 in Mammano and Nobili (1993). Dashed lines: original data. The segment shown by solid curves is used for the model in this paper. In this segment the data have been slightly smoothed to remove discontinuities.

equal to 1 (mm) and ε equal to 0.2. The height h of each channel is also 1 (mm). The dimensions differ considerably between the two models, particularly in the basal region, but we will see further on what the effect of this discrepancy is. We will call this model the “stylized” model. In addition to three-dimensional models we will consider a two-dimensional model, to be labeled “2d.” This model has the same outer dimensions as the “stylized” three-dimensional model but the BM is assumed to move *uniformly* over the entire width b of the model. A few words about units. We will use mm (millimeter) as the unit for length and g (gram) as the unit of mass. As a result the density of water is 10^{-3} (not 1 as in the cgs system).

Before the inverse solution can be applied, it is necessary to transform the response, measured at one location as a function of *frequency* (f), to the response at one frequency (e.g., the best frequency, BF) as a function of *location* (x). We assume that frequency and place “scale” in the sense described by (among others) Zweig (1976) and Zweig *et al.* (1976). In particular, the (radian) frequency $\omega(x)$ ($2\pi f$) corresponding to location x is

$$\omega(x) = \omega_{\max} \exp(-\frac{1}{2}\alpha x), \quad (1)$$

where $\omega_{\max}$ is the radian frequency corresponding to the location of the stapes ($x=0$) and α is a constant. Scaling implies that the BM impedance function $Z_{\text{BM}}(x, \omega)$ has the corresponding property in its dependence on x and ω . It is found as the impedance for one frequency, as a function of x ,

but it can be transformed to a function of ω for a given value of x . We use the following parameters:

$$\alpha = 0.5 \text{ (mm}^{-1}\text{)}, \quad \omega_{\max} = 2\pi 45 \text{ (kHz)}, \quad L = 6 \text{ (mm)}, \quad N = 700. \quad (2)$$

The value of α is “borrowed” from Greenwood (1961), and that of $\omega_{\max}$ is extrapolated from Cooper and Rhode’s (1992) findings (their Fig. 10, corrected for a shift of 1.7 mm).² The mapping is assumed to be the same in all models that we will consider. The parameter N is the number of sections in which the length L is divided.

The model equation has been solved with the full-matrix technique described in INV-3, which is based on papers by Allen (1977) and by Mammano and Nobili (1993). This technique was chosen because an alternative, faster and more economical technique (de Boer, 1998) could only be used for model “3ds.” In the inverse solution the boundary condition at the stapes side is formulated in terms of the “virtual stapes velocity” which is defined as the stapes velocity for which the volume of fluid in one channel of the model is conserved (INV-3). This definition can be used for model “3ds.” For model “3dm” we have to reinterpret the virtual stapes velocity because the BM width $b_{\text{BM}}(x)$ varies with x . In our formulations the BM velocity $v_{\text{BM}}(x, \omega)$ is defined as the *point velocity* of the center of the BM. For model “3dm” the virtual stapes velocity must then be computed as (minus) the fluid *volume* velocity density—which is $b_{\text{BM}}(x)v_{\text{BM}}(x, \omega)$ —integrated over x divided by the area of the stapes (which is taken as the area of the upper channel in Fig. 1 at $x=0$).

Executing a forward solution for the chosen model (resynthesis) is a way of “going from the model back to the data.” When the *same* model is used for inverse analysis and resynthesis, the resulting response is identical (within plotting accuracy) to the response used as input to the inverse procedure. As a result of smoothing the BM impedance function small deviations occur (cf. INV-3), and we have seen to it that these are less than 2 dB in amplitude. We will explore more subtle relations between experimental data and resynthesized response in Secs. II and III.

C. Results (I)

The *upper* panel (a) of Fig. 2 shows, for one experiment, the measured composite ccf spectrum transformed from the frequency to the place domain. The abscissa is the longitudinal coordinate x (labeled “location”). We have divided the length L of the model (6 mm) into N (700) sections. The solid line shows the amplitude, and the dashed line the phase. The amplitude is normalized to 0 dB at the peak (see the legend for the maximal BM-to-stapes velocity ratio³ and the virtual stapes velocity). The response shown is the response of the model under consideration, for stimulation by a tone with its frequency equal to the BF.

In the *lower* panel (b) of Fig. 2 the computed BM impedance $Z_{\text{BM}}(x, \omega)$ resulting from the inverse solution is presented, the solid line showing the real part and the dashed line the imaginary part. The impedance function is computed for model “3dm” (the “realistic” model). As in earlier papers (INV-1 to INV-3), the ordinate scale for the impedance

FIG. 2. Upper panel: Measured response (amplitude and phase) transformed to the x domain. Experiment 7619, stimulus level 50 dB. Maximal response amplitude level in this figure is normalized to 0 dB on the ordinate. Lower panel: BM impedance (real and imaginary parts) resulting from inverse solution for model “3dm” (the “realistic” model). The impedance scale is nonlinear: values from -1 to $+1$ are shown linearly, larger values are logarithmically compressed. The norm “1” corresponds to $2 \text{ g mm}^{-2} \text{ s}^{-1}$, or $2000 \text{ kg m}^{-2} \text{ s}^{-1}$. Best frequency (BF): 16.6 kHz. Maximal BM-to-stapes velocity ratio: 112.4. Virtual stapes velocity: $2.26-0.5i$. The filled area shows the region where the BM impedance is “active” (i.e., where the real part is negative). The right-hand end point of this region coincides with the response peak (cf. de Boer, 1983). Length (L) of model: 6 mm. Number of sections (N): 700.

is nonlinear: small values are plotted linearly and large values are compressed logarithmically. See the legend to the figure.

In earlier work (e.g., INV-2 and INV-3) it was found that a *locally active model* is needed to simulate data from a good preparation at low stimulus levels. Figure 2 and later figures confirm this: there invariably is a region of x where the real part of the BM impedance is negative. This region lies basalward from—in our figure to the left of—the location of the largest response, and extends from the point where the response amplitude starts its final rise toward the peak to the location of the peak itself. In Fig. 2 the “active” region is indicated by shading.

Figure 3 shows response and BM impedance functions for four other experiments. In each of the four panels the upper and lower panels of Fig. 2 are merged. The ordinate scales for amplitude and phase are shown on the sides of the figure. The BM impedance is plotted on the same nonlinear ordinate axis as in Fig. 2; the impedance scale is placed in the middle. Of the abscissa only the region from 1 to 5 mm is shown to present more details in the region of the response peak. Response amplitude is plotted by solid and response phase by dashed lines. The real part of the BM impedance is plotted by solid and the imaginary part by dashed lines

whereby the impedance curves are drawn with thickened lines. Curves are labeled in the upper left panel, the general shape and the line thickness serve to identify the curves in the other panels. See the legend for experiment codes, BFs, maximum BM-to-stapes ratios and values of the virtual stapes velocity for the four experiments. Figure 3 has been prepared for the same “realistic” model (“3dm”) as Fig. 2.

Figures 2 and 3 serve to illustrate the large variability in the BM response and, especially, in the recovered impedance function. We observe that in all records the real part of the BM impedance is negative (the model is “locally active”) in the region where the response amplitude rises to its peak. Over this region the BM is enhancing the power of the cochlear wave, and it has been demonstrated earlier that actual amplification does occur in the model, but also that it is less than anticipated (INV-3). In panels (b) and (c) we observe secondary lobes where the real part of the BM impedance $Z_{\text{BM}}(x, \omega)$ is negative. Such lobes always correspond to regions where the response amplitude is rising rapidly. We should not expect appreciable amplification in these regions, however, because the imaginary part of the impedance is too large.⁴

In some experiments we found that the “active” region extends all the way to the left, and includes the location of the stapes ($x=0$). We have come to the conclusion that this type of finding is most probably due to errors in the data. The arguments for this conclusion come from the study described in Sec. IV. Here we recall that the influence of data errors generally increases from the right (the region just beyond the peak) to the left (the region of the stapes) (an effect amply illustrated in INV-2 and INV-3). That the model has to be “locally active” in the response peak region is a much more “robust” finding than a similar property in the basal region.

II. RESULTS II—BACK TO THE DATA: “CROSS-FERTILIZATION”

Our next task is to explore differences between various models. In order to find out what is essential and what is not, we derive the BM impedance function for one model and execute resynthesis with another one. We will call this procedure “cross-fertilization.” We found it more rewarding than mere inspection of BM impedance functions, as will become apparent further on. It is recalled that, when the *same* model is used for inverse analysis and resynthesis, the resynthesized response is virtually identical to the response that serves as the input to the inverse procedure (INV-3).

Figure 4 shows response and BM impedance function for one of our experiments, the same as in Fig. 2. The BM impedance is computed for the “realistic” model (“3dm”), as before. Next, resynthesis is done, but this time with the “stylized” model (“3ds”), in which model the geometry is independent of x . The result is labeled with the code “3dm-3ds,” and it is seen to be similar in shape to the input response. It has a somewhat smaller slope of the amplitude curve and also a steeper phase curve. These two features are obviously due to the taper in dimensions that the “3dm” model exhibits and that is absent in the “3ds” model. Deviations in the amplitude are limited to 8 dB.

FIG. 3. Response and BM impedance in four experiments. The two panels of Fig. 2 are merged into one. Of the abscissa only the section from 1 to 5 mm is shown. Inverse solution for model “3dm” (as in Fig. 2). Panels (a)–(d). Experiment: 7620, 1020, 8925 and 6616. Best frequency (BF): (a) 16.3, (b) 16.0, (c) 16.7, and (d) 17.6 (kHz). Maximal BM-to-stapes velocity ratio: (a) 289, (b) 377, (c) 1003, and (d) 335. Virtual stapes velocity: (a) $1.26 - 0.474i$, (b) $0.338 - 0.038i$, (c) $1.059 - 1.27i$, and (d) $0.77 + 0.33i$.

The converse procedure is illustrated by Fig. 5. The BM impedance is computed for the “stylized” model (“3ds”). The first thing that meets the eye is that the BM impedance function in Fig. 5 is not much different from the one shown in Fig. 4. Resynthesis is done with the “realistic” model (“3dm”), and, as before, the differences between the models turn out to be moderate. This time the resynthesized response (labeled “3ds-3dm”) is about 5 dB larger in amplitude at the response peak, and the phase slope is seen to be slightly smaller than for model “3ds.” In both Figs. 4 and 5 the *variations* of the amplitude in the region of the response peak are reasonably well reconstructed, and the same is true for the *variations* of the phase. It should now be clear that “cross fertilization” is better suited to illustrate differences between models than mere observation of the BM impedance function. We conclude that the effect of the taper in model dimensions on the response is relatively small. Sizable effects in the stapes region (near $x=0$)—where model “3dm” is highly inhomogeneous—can only be expected for much lower frequencies (cf. the “breaking of symmetry” treated by Shera and Zweig, 1991).

We now turn to the *two-dimensional model* (see Sec. IB). Figure 6 shows, in the lower panel, the BM impedance function for the three-dimensional “stylized” model (“3ds”), the same function as in Fig. 5. Resynthesis is carried out for the two-dimensional model (“2d”), again with the full-matrix method (see Allen, 1977 and Sondhi, 1978). The resynthesized response is labeled “3ds-2d” (upper panel). Now we see larger differences, especially in the *initial slope* of the phase curve. In point of fact, this comparison is somewhat unfair. In the basal region of this two-dimensional model, the width of the BM is 1 mm, which is five (or more) times that in the other two models (see Sec. IB). This implies that in the long-wave approximation the phase variations will be more than two times what they are in

the three-dimensional model [de Boer (1996), Eq. (4.2.8) with H replaced by the quotient of channel cross-section area and BM width], and this is indeed what we observe.

III. RESULTS III—BACK TO THE DATA: IMPULSE RESPONSE

When inverse solution and resynthesis are done with the same mathematical formalism, the resynthesized response is virtually identical to the response that has served as the input to the inverse solution. This input response was, as recalled, the response imposed on the model in the *place* domain (the x domain). Likewise, the resynthesized response is a model response in the x domain for a fixed frequency. However, we originally started from a response measured in the *frequency* (f) domain, for a fixed location, and we may well inquire into the properties of the resynthesized response in the f domain.

For a sensitive test we turn to the *time* (t) domain. Let the input signal to the analysis be the impulse response corresponding to the composite ccf spectrum and assume that we have obtained the resynthesized response in the x domain. To obtain the resynthesized *impulse* response it is first necessary to retransform the resynthesized response from the x to the f domain. This is done by inverting the frequency-to-place transformation that the original response underwent. That procedure includes undoing the amplitude compensation described in INV-2 and the phase compensation described in INV-3. The resynthesized frequency response is now known from the frequency f_0 corresponding to $x=0$ upward but *it is unknown for lower frequencies*, f_0 being of the order of 5 kHz. Without further ado, the transformed function is linearly extrapolated down to zero frequency with constant phase and amplitude proportional to frequency. A Fourier transformation then produces the resynthesized impulse response. The so-obtained impulse response, (b), is

FIG. 4. "Cross fertilization I." Experiment 7619. Upper panel: original response, plotted with respect to stapes velocity but with the amplitude shifted down by 10 dB. Lower panel: BM impedance computed for the "realistic" model ("3dm"). Resynthesis is done with the "stylized" model ("3ds"), the amplitude is labeled "3dm-3ds" in the upper panel. The resynthesized phase is shown by the lower dashed curve.

FIG. 6. "Cross fertilization III." Experiment 7619. Upper panel: original response as in Fig. 4. Lower panel: BM impedance computed for the "stylized" model ("3ds"). Resynthesis is done with the two-dimensional model ("2d"), amplitude and phase are labeled "3ds-2d" in the upper panel. Note the large deviations in the phase slope (see text).

FIG. 5. "Cross fertilization II." Experiment 7619. Upper panel: original response as in Fig. 4. Lower panel: BM impedance computed for the "stylized" model ("3ds"). Resynthesis is done with the "realistic" model ("3dm"), the amplitude is labeled "3ds-3dm" in the upper panel. The resynthesized phase is shown by the upper dashed curve.

shown in Fig. 7, together with the original one, (a), on vertically displaced coordinate axes. For this figure the "stylized" model is used. Amplitude scaling is the same for the two responses. We observe a very close correspondence be-

FIG. 7. Resynthesis in the time domain. Experiment "1016." Maximal BM-to-stapes velocity ratio: 291. Best frequency (BF): 16.6 kHz. Virtual stapes velocity: $1.23-0.49i$. Curve (a): original impulse response (i.e., the ccf that corresponds to the composite ccf spectrum). Curve (b): resynthesized impulse response for the "3ds" model as described in the text. Curves (a) and (b) have the same normalization factor. Curve (c): "cross fertilization," from model "3ds" to model "3dm." Curve (d): "cross fertilization," from model "3ds" to model "2d." Normalization factors of curves (c) and (d) are selected for clarity (see text).

FIG. 8. Effect of deliberate errors added to the ccf data. Response and impedance panels merged as in Fig. 3. Panels (a) and (b): two experiments, 1016 and 1019. In each panel five response and impedance curves have been superimposed, four of them corrupted by noise (see text). Note how the computed BM impedance is the most stable against these errors in the region of the response peak. Deviations become larger when going to the left, a well-known effect from earlier work (see references in text).

tween the curves, in particular with respect to the timing of the zero crossings. Clearly, the rather crude way we handled frequencies below 5 kHz has little influence on the ultimate impulse response. We have obtained equivalent results with the “3dm” model.

It is interesting to consider effects of “cross fertilization” on the impulse response. Curve (c) of Fig. 7 shows the impulse response arising when we start with the “3ds” model and do resynthesis with the “3dm” model (as we did in Fig. 5). We show this impulse response with a 6 dB smaller amplitude (compare Fig. 5). The result is of the same general type as curves (a) and (b) but shows a few characteristic differences. The small deviation in phase slope that we saw in the peak region in Fig. 5 translates itself into a small shift of the group delay and a change in the shape of the waveform envelope. The timing of the zero-crossings of the original impulse response is not preserved in this response. Curve (d), finally, shows the impulse response corresponding to Fig. 6, resynthesized for the two-dimensional model. To compensate for the large amplitude difference seen in Fig. 6, we show the impulse response with a 12 dB larger amplitude. Apart from the initial oscillations and the shift in group delay, this curve resembles the original one reasonably well. The same applies to curve (c). This correspondence results from good agreement in the peak region which in its turn is due to the fact that two- and three-dimensional models support *short waves* equally well (note that in the peak region short waves dominate, see, e.g., INV-3, Sec. II).

On close inspection it is seen that during the course of the impulse response the frequency of the oscillations increases gradually. This effect is called the “glide” and has been extensively reported in Part I. This feature is retained in the resynthesized impulse response [curve (b)] (as well as in the other resynthesized impulse responses). In fact, this means that we have here a model of the cochlea that faithfully

simulates all major and many subtle aspects of the cochlear response. The resulting model is characterized by its BM impedance function, and not by being constructed from elements that functionally correspond to mechanical structures of the cochlea. In this sense the model is abstract.

IV. RESULTS IV—BACK TO THE DATA: VARIABILITY

The inverse-solution method is often described as an “ill-posed problem.” In INV-3 it has been argued why this epithet does not always automatically apply to cochlear mechanics. Here we will illustrate the same topic from another side. We will take the data in the form of a composite ccf spectrum (obtained, processed and smoothed according to the standard procedure described in this and previous papers) and *deliberately corrupt them by adding a random noise signal*. The ccf spectrum has components spaced by approximately 50 Hz (see, for instance, Part I), and each of these components is multiplied by a complex number $(1 + \beta c)$, where β is a small coefficient and c is a random complex number with unity amplitude and random phase. The values of c for the different components are independent of each other. The resulting spectrum is not smoothed in any additional way, and is used as the input signal of the inverse procedure.

Typical results are shown by Fig. 8, panels (a) and (b), for two experiments. In each panel five sets of response and impedance curves have been superimposed; one is the original set of curves, the other four result from corruption by noise. The impedance function corresponding to the original response is shown by two thickened lines (real and imaginary parts). The impedance functions due to corruption of the response are shown by normal lines. For this figure β was chosen as 0.2, implying average amplitude errors of the order of 1.5 dB. In both panels the impedance functions are observed to overlap well in the region of the response peak

and slightly to the left of it. This strengthens the conclusion that “local activity” is a robust feature of our procedure. More to the left of the peak the impedance due to the corrupted response starts to deviate markedly from the original function, and this is true over the entire range down to the stapes region. Near the stapes the real part tends to become negative, and the imaginary part to become less negative. A definite trend in accumulation of errors from right to left is evident. This behavior is typical for results of the inverse procedure (see for the explanation INV-1 and INV-3).

In going from the raw data to the composite ccf spectrum, a moderate amount of smoothing has been applied (see INV-3, Appendix B). As a result the errors in the original data (the ccf spectrum) are not independent from frequency to frequency. Therefore, the original response curves are smooth and the average amplitude error appears as less than 1 dB. The artificial errors that we have introduced in this section are much more “severe” in that they are *statistically independent*. The findings in Fig. 8 illustrate how critical is the method used in smoothing data.

We now come back to the point described in Secs. I C and II: in some of our data sets the BM impedance appears to be “active” all the way to the stapes location. By adding random “errors” of the same magnitude to such data it proved almost always possible to reverse this and to achieve a situation where there is no “activity” in the basal region. In this way we became convinced that “activity” in the stapes region is a variable and unstable property. It is only in the region of the response peak where the property of “local activity” is robust.

Possible influences of measurement errors can generally be estimated from repeating the same experiment a few times. In our experiments there also is a systematic error: we used *pseudo-random noise*, which means that one segment of a noise signal (in our case 20 ms long) is repeated over and over, without gaps (see Part I). We performed a few experiments in which four or five *different* noise signals were used as stimulus; these signals were based on different noise period waveforms. The BM velocity records were obtained within minutes of each other, and each record underwent the same inverse analysis. One of the aims of this test was to try to find evidence of corrugations in the BM mechanics. The term corrugation is used here in the sense discussed by Zweig and Shera (1995) as a component of BM impedance that is a function of location x , differs from animal to animal, and *does not scale*. The results were disappointing in that the processed ccf spectra showed random deviations not exceeding 1 dB, and the BM impedance functions showed corresponding deviations. No particular trend that could possibly be ascribed to irregularities of BM mechanics could be discerned.

V. REVIEW, CONCLUSIONS AND OUTLOOK

In combining experimental findings and theory we should keep one thing in mind: All our conclusions will only be valid under the assumption *that the real cochlea operates as the type of model that we are considering*. In the course of the analysis we convert the response, measured as a function of frequency at one location, to a model response distributed

over the length of the model for a particular frequency (the BF). Implicit is here that the cochlea “scales,” i.e., that it converts frequency to place in a regular manner. We have found that inverse analysis is relatively insensitive to details of the conversion (i.e., the form of the cochlear map), yet scaling is a prerequisite for the entire procedure. Of course, this implies that we must interpret the resulting BM impedance, which is a function of x and ω , in the same “scaling” sense. The recovered BM impedance is found as a function of x for fixed ω . Conversely, for a fixed value of x , the BM impedance should show the corresponding behavior as a function of $(\log) \omega$. This means that for a higher frequency the region of activity lies more basalward, etc.⁵ The same concept of inverse scaling is involved in the computation of the resynthesized impulse response (Sec. III).

In this paper we first went “from data to models.” We applied inverse analysis to data on the movement of the basilar membrane (BM) in response to noise stimuli. Two three-dimensional models were studied in this paper; one had its dimensions close to that of the real cochlea, the other one had homogeneous geometry over its length. Additionally, a two-dimensional model was treated. With all three models the BM impedance has to be “locally active” in order to match the measured response (at the levels employed in this study). It was found that in three-dimensional models the taper of BM width and cross-sectional area does not create more than a global effect. The two-dimensional model (in the form defined in Sec. I B) does not support the long-wave part of the cochlear wave correctly, but gives a fair representation of the short-wave part. Our conclusion at this point is that, to study global properties of cochlear models, it is sufficient to use the “stylized” model, “3ds.” This has the additional advantage that a very efficient and universal solution method is available (de Boer, 1998).

Apart from “cross fertilization” (Sec. II), we went “back to the data” in another way: we considered the resynthesized *impulse response*. We reestablished the consistency of the entire procedure, going from a measured response in the f domain, via the f -to- x transformation and the inverse solution to the BM impedance function in the x domain, via the forward solution for the model and, finally, via the x -to- f transformation back to the f domain. Resynthesis of impulse responses proved to be faithful to the original [Fig. 7, curves (a) and (b)].

All resynthesized impulse responses show the “glide.” Compared to the narrow-band ccf spectra analyzed in Part I, the composite ccf spectrum as it is used here has a much more extended low-frequency segment. We have ascertained that this feature has an enhancing effect on the glide (not shown). Thus the glide is not an artifact of the data processing method. On the contrary: when we have data covering a wider frequency range, the glide is slightly more pronounced.

From the results in Fig. 8(a) and (b) we conclude that the “wiggles” and random deviations that we normally find in the recovered BM impedance function are mainly due to measurement errors. No recognizable component of these fluctuations can be attributed to spatial irregularities in the

mechanical properties of the BM. The influence of errors is particularly large in the basal region.

In all experiments described here the stimulus level was low enough to neglect the contribution of distortion products to the response to noise stimuli. For stimuli of higher levels the cochlea will certainly show more pronounced nonlinearity. In this case linear analysis is justified, too—but only for certain purposes (the EQ-NL theorem, see the Appendix). Activity has been found to decrease with increasing stimulus level (de Boer and Nuttall, 1997a-DAM). A most remarkable nonlinear effect of the cochlea is the property that, as the stimulus level varies, the timing of the individual cycles of the impulse-response waveform is almost invariant [de Boer and Nuttall (1997b-GLIDE) and references cited therein; see also Recio *et al.*, 1998]. In this connection, we stress that zero-crossings of the impulse response are preserved in re-synthesis [curves (a) and (b) of Fig. 7]. Results on nonlinear effects associated with changes in stimulus level will be published elsewhere.

As a final note, it should be stressed that our model contains an abstract function, the BM impedance, $Z_{\text{BM}}(x, \omega)$, as its main parameter function. Our model remains abstract since it is not yet composed of elements of which each would replicate a specific mechanical component of the cochlea. The work presented here has laid solid foundations for further explorations in this direction. We are now more confident in relating certain components of the BM impedance to the dynamics of mechanical elements in the cochlea.

ACKNOWLEDGMENTS

The authors are indebted to (in alphabetical order) Gary Dootz, Anders Fridberger, Meng He Guo, Bob Masta, Darren Miller, Irina Omelchenko, Edward Porsov and Jiefu Zheng, for their help in setting up and performing the experiments. Thanks are due to Jont Allen, Renato Nobili, Fabio Mammano and Christopher Shera who shared their experience in MATLAB programming with us. Three JASA reviewers presented us with useful feedback designed to make our work more accessible to readers who are not thoroughly familiar with cochlear modeling. This work was supported by NIH (No. NIDCD-DC-00141), and VA Rehabilitation Research & Development Center (Grant RCTR-597-0160, Portland, VAMC), and is a continuation of project No. SLW 01.011 (NWO, The Netherlands).

APPENDIX: CLASS OF COCHLEAR MODELS, AN OVERVIEW

In the type of model used in this paper the basilar membrane (BM) forms a part of the cochlear partition that is located between two narrow fluid-filled channels stretched out in the x direction. It is assumed that outer hair cells (OHCs) of the organ of Corti are able to cause a local sound pressure $p_{\text{OHC}}(x, t)$ (an oscillating pressure, on a cycle-to-cycle basis) which, in its turn, gives rise to an additional component $p_{\text{add}}(x, t)$ of the sound pressure $p(x, t)$ near the BM in the fluid. By way of this pressure the OHCs are thought to enhance and sharpen the frequency response of

the system to a degree compatible with results from recent mechanical measurements of BM motion at low stimulation levels—including our own.

The basic form of the model is linear. In the *nonlinear* version of the model the “active” process in the outer hair cells (OHCs) is nonlinear. It is assumed that there exists a *nonlinear instantaneous* relation between the (radial) stereociliary deflection $d_{\text{cilia}}(x, t)$ (considered as a function of time t) and the pressure $p_{\text{OHC}}(x, t)$ that is locally produced by the OHCs. The nonlinear relation should be compressive. All frequency dependence around OHC-bound processes is included in two *linear frequency- and location-dependent* transformations:

- between BM displacement $d_{\text{BM}}(x, t)$ and ciliary excitation $d_{\text{cilia}}(x, t)$, and
- between OHC output $p_{\text{OHC}}(x, t)$ and the corresponding component $p_{\text{add}}(x, t)$ of the channel pressure $p(x, t)$ near the BM.

It is for nonlinear models of this class that the EQ-NL theorem holds (de Boer, 1997). This theorem is formulated in terms of the input–output cross-correlation function (ccf) for a wideband noise signal with uniform spectral density as input. The theorem states that the ccf of the nonlinear model is equal to the ccf of a *linear* “comparison model” and it defines exactly how that comparison model must be constructed. If the actual cochlea functions as the model considered here, we will use the EQ-NL theorem to interpret the data. That interpretation will always be in the language of *linear-systems theory*. Then, it is legitimate to use all concepts of linear-systems theory including “impedance” and “impulse response,” and, in particular, it is permitted to use the inverse-solution method to find the BM impedance $Z_{\text{BM}}(x, \omega)$. Formally, this impedance is the BM impedance of the “companion model” mentioned earlier. On a final note, we will tacitly assume that the model is zero-point stable, which means that it does not go into spontaneous oscillation. On the experimental side, this implies that we assume that spontaneous otoacoustic emissions of the cochlea are either absent or suppressed by the stimulus we used in our experiments.

¹This 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 and the Oregon Health Sciences Committee on the Use and Care of Animals.

²In earlier work (INV-3) a higher value was selected for ω_{max} , namely $2\pi 60$ (kHz). This was done to extend the region over which the wave in the model is of the long-wave type.

³In INV-3 data from the same animal were used, at the stimulus level of 40 dB and unfiltered. Here the data are taken at 50 dB, and the response is filtered; as a result, the maximal BM-to-stapes ratio is smaller than in INV-3—where it was 177.

⁴Assume the imaginary part to dominate the BM impedance. In the short-wave region the amplification in dB per mm then is inversely proportional to the square of the imaginary part.

⁵Scaling of the impedance involves an extra factor ω , see Eqs. (2.2.a, b and c) in de Boer (1991).

- Cooper, N. P., and Rhode, W. S. (1992). "Basilar membrane mechanics in the hook region of cat and guinea-pig cochleae: Sharp tuning and nonlinearity in the absence of baseline position shifts," *J. Acoust. Soc. Am.* **63**, 163–190.
- Davis, H. (1983). "An active process in cochlear mechanics," *Hear. Res.* **9**, 79–90.
- de Boer, E. (1983). "No sharpening? A challenge for cochlear mechanics," *J. Acoust. Soc. Am.* **73**, 567–573.
- de Boer, E. (1991). "Auditory physics. Physical principles in hearing theory. III," *Phys. Rep.* **203**, 125–231.
- de Boer, E. (1993). "Some like it active," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 3–21.
- de Boer, E. (1995a). "The 'inverse' problem solved for a three-dimensional model of the cochlea. I. Analysis," *J. Acoust. Soc. Am.* **98**, 896–903 (INV-1).
- de Boer, E. (1995b). "The 'inverse problem' solved for a three-dimensional model of the cochlea. II. Application to experimental data sets," *J. Acoust. Soc. Am.* **98**, 904–910 (INV-2).
- de Boer, E. (1996). "Mechanics of the cochlea: modeling efforts," in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 258–317.
- de Boer, E. (1997). "Connecting frequency selectivity and nonlinearity for models of the cochlea," *Aud. Neurosci.* **3**, 377–388.
- de Boer, E. (1998). "A method for forward and inverse solutions of a three-dimensional model of the cochlea," *J. Acoust. Soc. Am.* **103**, 3725–3728.
- de Boer, E., and Nuttall, A. L. (1997a-DAM). "On cochlear cross-correlation functions: connecting nonlinearity and 'activity'," in *Diversity in Auditory Mechanisms*, edited by E. R. Lewis, G. R. Long, R. F. Lyon, P. M. Narins, C. R. Steele, and E. Hecht-Poinar (World Scientific, Singapore), pp. 291–297.
- de Boer, E., and Nuttall, A. L. (1997b-GLIDE). "The mechanical waveform of the basilar membrane. I. Frequency modulations ('glides') in impulse responses and cross-correlation functions," *J. Acoust. Soc. Am.* **101**, 3583–3592.
- de Boer, E., and Nuttall, A. L. (1999-INV-3). "The 'inverse problem' solved for a three-dimensional model of the cochlea. III. Brushing up the solution method," *J. Acoust. Soc. Am.* **105**, 3410–3420.
- Fernández, C. (1952). "Dimensions of the cochlea (guinea-pig)," *J. Acoust. Soc. Am.* **24**, 519–523.
- Greenwood, D. D. (1961). "Critical bandwidth and the frequency coordinates of the basilar membrane," *J. Acoust. Soc. Am.* **33**, 1344–1356.
- Kim, D. O., Neely, S. T., Molnar, C. E., and Matthews, J. W. (1980). "An active cochlear model with negative damping in the partition: Comparison with Rhode's ante- and post-mortem observations," in *Psychophysical, Physiological and Behavioural Studies in Hearing*, edited by G. v. d. Brink and F. A. Bilsen. (Delft U.P., Delft), pp. 7–14.
- Mammano, F., and Nobili, R. (1993). "Biophysics of the cochlea: Linear approximation," *J. Acoust. Soc. Am.* **93**, 3320–3332.
- Nuttall, A. L., and Dolan, D. F. (1996). "Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig," *J. Acoust. Soc. Am.* **99**, 1556–1565.
- 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.
- Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (1998). "Basilar-membrane responses to clicks at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **103**, 1872–1889.
- Shera, C. A., and Zweig, G. (1991). "A symmetry suppresses the cochlear catastrophe," *J. Acoust. Soc. Am.* **89**, 1276–1289.
- Sondhi, M. M. (1978). "Method for computing motion in a two-dimensional model," *J. Acoust. Soc. Am.* **63**, 1468–1477.
- Zweig, G. (1976). "Basilar membrane motion," *Cold Spring Harbor Symp. Quant. Biol.* **40**, 619–633.
- Zweig, G., and Shera, C. A. (1995). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," *J. Acoust. Soc. Am.* **98**, 2018–2047.
- Zweig, G., Lipes, R., and Pierce, J. R. (1976). "The cochlear compromise," *J. Acoust. Soc. Am.* **59**, 975–982.