Traveling wave theory, and some shortcomings

This chapter summarises the existing traveling wave model of the cochlea, looks at some shortcomings, and suggests that many of the documented anomalies can be explained by assuming that outer hair cells are responsive to the fast pressure wave. It does not try to present an historical account of the development of the theory nor offer a comprehensive account of every conceivable refinement that has been attempted – a vast task[1]. Rather, it looks at the basic core of the theory and one modern account that is generally accepted as the standard picture. The modern version, due to Shera and Zweig (§I 3.1/d), adds two key elements – active properties and a reverse traveling wave – necessary to account for otoacoustic emissions. However, although generally successful, the modern version is still unable to account for the full range of cochlear phenomena, as we will see. Perhaps refinements can be made to overcome the shortcomings, but I want to suggest that the fault may lie in the basic reliance on differential pressure and that otoacoustic emissions could reflect a situation in which, at low sound pressure levels, the cochlea operates along pure local resonance principles and is responding to the fast pressure wave.

In some places the arguments I put forward rely on just sketching the outline of an alternative picture, as evidence is lacking to support what I admit is a non-conventional approach. Nevertheless, I have tried to make the alternative model as clear as I can, and I hope that others with more mathematical facility can place the model on a firmer footing if they see virtue in it. The intention is that by questioning the fundamentals of cochlear mechanics, progress in understanding may be made. I hope this sceptical approach will open up new avenues and therefore be more fruitful than simply accepting the textbook account on face value.

3.1 Formulation of the traveling wave equations

As described in §I 1.7, two different, but related, signals arise in the cochlea in response to sound stimulation. The first, p_+, is the common-mode pressure and the second, p_-, the differential pressure.

[1] For two accounts, see Zwislocki, J. J. (2002). *Auditory Sound Transmission: An Autobiographical Perspective*. (Erlbaum: Mahwah, NJ). // de Boer, E. (1996). Mechanics of the cochlea: modeling efforts. In: *The Cochlea*, edited by P. Dallos et al. (Springer: New York), 258-317.

To recapitulate, p_+ is the acoustic pressure wave that is created by the stapes vibrating backwards and forwards in the oval window. It spreads throughout the cochlear fluids at the speed of sound in water (1500 m/s), creating, nearly instantaneously, a quasi-static hydraulic pressure field that is an exact analog of stapes motion (and ear-canal pressure). This pressure wave depends on the mass and compliance of the cochlear fluids; after the wave has traversed the cochlea a number of times, the magnitude of the residual common-mode pressure depends crucially on the compliance of the round window.

The second signal, p_-, is the difference in pressure between the upper and lower galleries caused by the presence of the partition. Depending on its acoustic impedance, a pressure difference will occur across the basilar membrane, leading to a pressure p_v in the upper gallery (scala vestibuli) and a pressure p_t in the lower (scala tympani).

Thus, the common mode pressure p_+ is given by $p_+ = (p_v + p_t)\ /2$, whereas the differential pressure $p_- = (p_v - p_t)\ /2$.

The standard view is that differential pressure is the sole stimulus in the system, and so a traveling wave mechanism excites the hair cells and thence auditory nerve fibres. As foreshadowed in Chapter I1, I find this conclusion not fully justifiable, and here I want to put forward some reasons. I do not deny that a traveling wave mechanism may exist; but I think that the effects attributed to it have been exaggerated, and, at least at low sound pressure levels, are smaller than those due to excitation of the partition by outer hair cells in response to the fast pressure wave.

3.1/a The first transmission line model

Békésy provided no mathematical underpinning for his theory, leaving that to others. The first step towards a mathematical model was made by Wegel and Lane in 1924, who proposed that the cochlea operated like a tapered transmission line[2]. Their electrical network model looked like Fig. 3.1, and this representation is the essence

[2] Here I follow Allen, J. B. (2001). Nonlinear cochlear signal processing. In: *Physiology of the Ear (2nd ed.)*, edited by A. F. Jahn and J. Santos-Sacchi (Singular Thomson Learning: San Diego, CA), 393-442. Allen also points out (§1.1) that Fletcher deserves some credit too.

of the traveling wave formalism. Nearly the same arrangement is used today, albeit with additional serial and parallel elements; nowadays the mass (inductance) is usually taken to be more or less constant from base to apex[3].

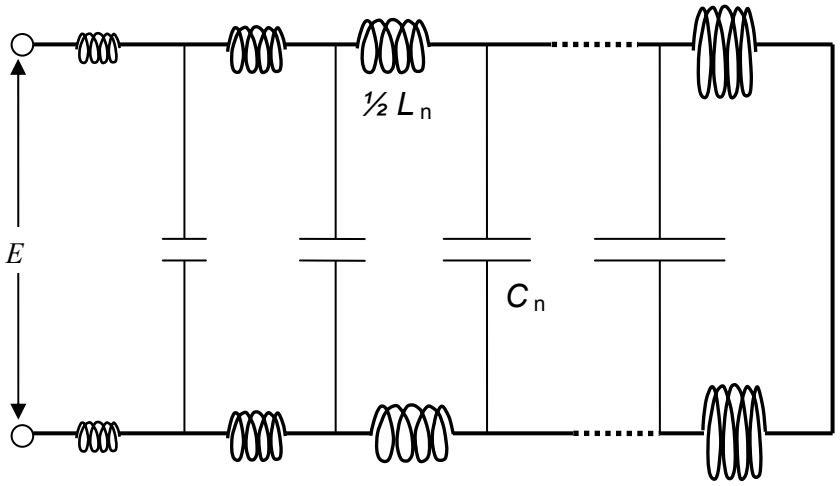

Fig. 3.1. An electrical network analog of the cochlea, the basis of all traveling wave models. The capacitances represent the compliances of the basilar membane (the partition taken as a whole). The inductances represent masses of fluid in the upper and lower galleries.

A modern-day treatment[4] of passive cochlear mechanics can be found in Fletcher (1992). A convenient analogue treatment giving a simple one-dimensional model is to take voltage to represent pressure and current to represent acoustic volume flow. Simplifying as much as possible, inductances represent the mechanical inertance, due to mass, of the fluid in the upper and lower galleries, which the stapes pressure encounters when the oval window pushes in and out; the capacitances represent the compliance of the basilar membrane, which tends to deflect in reaction to the pressure in the fluid moving along the galleries[5]. It is assumed that there is no mechanical coupling along the membrane itself, so that all coupling is due to the surrounding fluid. Dividing the cochlea into equal-length sections, the inductance, L_n, representing the mechanical impedance of each section is given approximately by

[3] Geisler, C. D. (1976). Mathematical models of the mechanics of the inner ear. In: *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer: Berlin), vol. 5.3, 391-415.
[4] Fletcher, N. H. (1992). *Acoustic Systems in Biology*. (Oxford University Press: New York). See Ch. 8 and Ch. 12.4.
[5] The total volume of incompressible fluid displaced by the stapes has to move the round window, and in so doing it either moves along the upper gallery to the lower through the helicotrema or takes a short cut by deflecting the basilar membrane. By 'basilar membrane' is meant the whole partition – the organ of Corti and all its supporting structures.

$$\tfrac{1}{2} L_n \approx \frac{\text{density of fluid} \times \text{section length}}{1/2 \text{ cross-section area of channel}} \qquad (3.1)$$

and the capacitance, C_n, is given by

$$C_n \approx \frac{\text{width of vibrating b.m.} \times \text{segment length}}{\text{stiffness of b.m.}} . \qquad (3.2)$$

As Fig. 3.1 illustrates schematically, both L_n and C_n increase as distance, x, from the base increases, in the first case because the cross-section of the channel decreases a little and in the second because the stiffness of the partition (essentially taken to be the basilar membrane) decreases and its width increases. The helicotrema (Fig. 3.7) is usually treated as a short circuit, although in practice there will be a small mechanical impedance associated with it.

The result is that the mechanical impedance, $Z(x,\omega)$, can be represented by an equation of the form

$$Z = i\omega m + K/(i\omega) + r \qquad (3.3)$$

where[6] m is the mass per unit length associated with each section (50 mg/cm^2 is typical), K is the stiffness (such that it decreases exponentially with distance like $K = 10^7 e^{-1.5x}$), and r is a damping term (in the manner of $r = 3000 e^{-1.5x}$).

At some angular frequency, ω, within the auditory range, the inertia and compliance of one section, taken to be the nth, will be in resonance so that $\omega = 1/(L_n C_n)^{1/2}$ and the section will have almost no impedance and look like a short circuit (a hole). The result is that all the flow passes through this section, causing large displacement of the partition, limited only by damping. On the apical side of this point, both L and C are large (large cross-section and low stiffness) and lie far from resonance so that the signal will, given the stiffness map, be attenuated about exponentially; very little will pass through the helicotrema. On the basal side, the two factors work together to produce a traveling wave which progresses along the partition, increasing gradually in amplitude to reach a broad peak, and dissipating

[6] Typical values as used in Lesser and Berkley (1972).

before[7] it reaches the resonant point. Each frequency will come to a peak at a particular point along the cochlea – its characteristic frequency: the lower the frequency, the further along the partition it will reach. At very low (subsonic) frequencies, ωL_n is very small and $1/\omega C_n$ very large, so fluid must then flow through the helicotrema.

Of course, this treatment is a simplification, and ignores active properties, but it gives a useful one-dimensional picture – the acoustic pressure is assumed to be a function of only the distance from the stapes – and provides an explanation of how tonotopic tuning can arise in the cochlea. It is the picture that naturally explains Békésy's stroboscopic observations on human cadavers at extreme sound levels and it remains the centre-piece of modern cochlear models.

More detailed treatments can be found in expositions by Lighthill[8], Zwislocki[9] and de Boer[10–13], and the accepted modern-day active model, due to Shera and Zweig, is outlined in §3.1/d below. Overall, none of these models deviate from the fundamental property that the stimulus travels through the network elements in series – a stimulus cannot reach its characteristic place on the partition without going through a cascade of circuit elements; thus for all audible sounds, there will be a significant time delay before a stimulus can reach a hair cell. The propagation speed of a traveling wave starts out at more than 100 m/s at the base and slows down to as low as 1 m/s at the apex. The time delay to the peak is typically 1 or 2 cycles, so that for a 1 kHz signal, the group delay will be 1 or 2 ms. A distinguishing feature of traveling waves is accumulating phase delay with frequency, until at the characteristic frequency many cycles of delay are apparent – the pivotal reason that resonance models, limited to $\pi/2$ delay, have been discarded[14].

---

[7] For discussion of this point, see Zwislocki (2002), Lighthill (1981, 1991), p. 9; Patuzzi (1996), p. 214; Withnell (2002), Fig. 3.
[8] Lighthill, J. (1981). Energy flow in the cochlea. *J. Fluid Mech.* 106: 149-213.
[9] Zwislocki, J. J. (1965). Analysis of some auditory characteristics. In: *Handbook of Mathematical Psychology*, edited by R. D. Luce et al. (Wiley: New York), 3, 1–97.
[10] de Boer, E. (1980). Auditory physics. Physical principles in hearing theory. I. *Physics Reports* 62: 87-174.
[11] de Boer, E. (1984). Auditory physics. Physical principles in hearing theory. II. *Physics Reports* 105: 141-226.
[12] de Boer, E. (1991). Auditory physics. Physical principles in hearing theory. III. *Physics Reports* 203: 125-231.
[13] de Boer (1995).
[14] Patuzzi (1996), p. 199.

Some physically important insights into traveling wave behaviour are given by Lighthill (1981). First, he points out (his Fig. 1) that the system differs from a standard electrical waveguide in that the cut-off is a high-frequency one (not low-frequency); hence a propagating wave will not be reflected as it will in the standard electrical analogue. Thus, he prefers to make the analogy (his section 4) with an atmospheric wave phenomenon called critical-layer resonance. Secondly, he highlights (p. 193) that traveling wave mechanics entails that stapes pressure cannot remain perfectly in phase with volume flow – the wave is somewhat decoupled from its driving force – and so this reduces the ability of the stapes to efficiently drive the basilar membrane. This means that a purely resonant interaction between the two is not possible, particularly at low frequencies, where the phase relationship approaches 90°. Finally, he underlines the importance of the fast wave, which carries off half of the stapes energy according to his reckoning (pp. 150, 176), and which is necessary to explain why high-frequency limits in the cochlea often plateau at phases with integer multiples of π, behaviour which is "inconceivable" in a traveling wave system (pp.153, 180).

3.1/b Differential pressure and common-mode pressure

In order to see how well the above description relates to the actual physics of the cochlea, we need to be sure that the equations we choose are comprehensive – as simple as possible, but no simpler, as Einstein expressed it. Fletcher (1992) provides a basic schema, but ignores common-mode pressure. In this thesis it is considered vital to set out a formalism that includes both differential and common-mode pressures. The first such approach was that due to Peterson and Bogert[15] (1950), who in fact introduced the notation p_+ and p_- for what they called 'longitudinal' and 'transverse' modes of pressure (and similarly u_+ and u_- for the associated particle velocities). They give an equivalent circuit (Fig. 3.2 below) that generates both common-mode pressure ($p_v + p_t$) and a differential pressure ($p_v - p_t$). Given certain boundary conditions, a set of equations were developed that mirror this circuit.

[15] Peterson, L. C. and B. P. Bogert (1950). A dynamical theory of the cochlea. *J. Acoust. Soc. Am.* 22: 369-381.

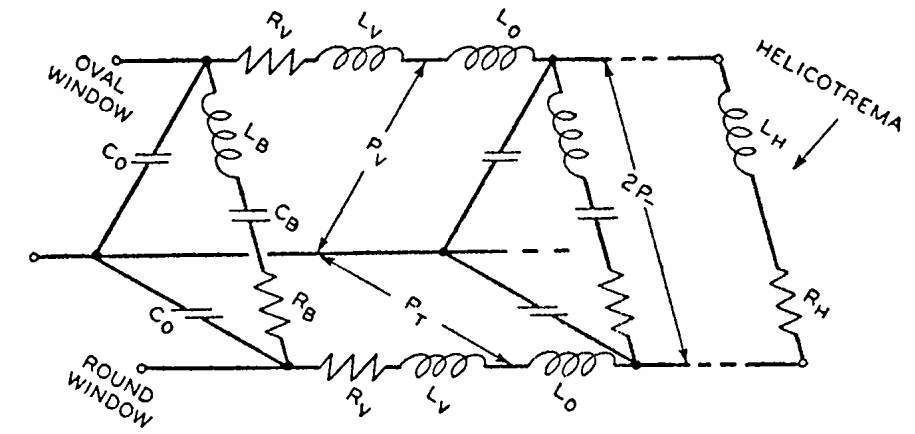

Fig. 3.2. The first equivalent circuit of the cochlea to include both common-mode and differential pressure. The three-terminal transmission line is from Fig. 22 of Peterson and Bogert (1950), and used with permission of the Acoustical Society of America.

The equations involving p_+, the instantaneous pressure, were

$$p_+ = P_+ e^{i\omega t} \tag{3.4}$$

(where P_+ is the pressure amplitude at the stapes and ω its frequency)

and $$\frac{1}{S(x)}\frac{\partial}{\partial x}\left(S(x)\frac{\partial p_+}{\partial x}\right) = \frac{1}{c^2}\frac{\partial^2 p_+}{\partial t^2} \tag{3.5}$$

where $S(x)$ is the cross-sectional area of each gallery at distance x from the base, and c is the velocity of sound in a free fluid. Given some simplifications, these equations, can be solved numerically. To do so, three boundary conditions are imposed: a fixed pressure of 2 dyne/cm^2 at the oval window; no pressure (but continuity of flow) across the helicotrema; and zero pressure at the round window. They therefore managed to derive a complex expression for $P_+(x,\omega)$ [their equation on bottom of p. 373)] which was independent of p_- and had a closed form solution involving Bessel functions. Another set of equations, independent of the first set, described the differential pressure, and these naturally lead to the standard traveling wave.

The numerical solutions provided a graph (their Fig. 4) of p_+ along the length of the cochlea. At low frequencies (some kilohertz), the average pressure is virtually constant along the partition at about 1 dyne/cm^2, but at higher frequencies a standing

wave begins to form[16] and so at 10 kHz the average pressure ranges from 1 dyne/cm^2 at the stapes to nearly 4 dyne/cm^2 at the apex. Similarly, they calculate the differential pressure, which, for all frequencies, ranges from 1 dyne/cm^2 at the base to zero (as specified) at the apex. For progressively higher frequencies, zero differential pressure occurs closer to the base, so that at 10 kHz (shown in Fig. 3.3 below), all differential pressure vanishes 10 mm beyond the stapes. Calculations of transit times of impulses through the system (their Table 1) appear to broadly match those seen by Békésy.

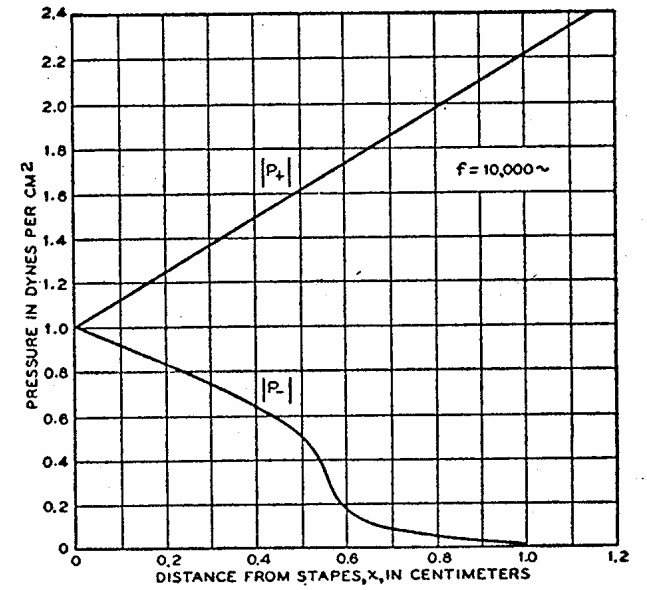

Fig. 3.3. Common-mode pressure p_+ and differential pressure p_- as calculated by Peterson and Bogert (1950) for a frequency of 10 kHz. Note that, given their boundary conditions, the magnitude of the former exceeds the latter. (Reproduced from their Fig. 9, and used with permission of the Acoustical Society of America)

Undoubtedly, the Peterson and Bogert paper is a major advance in understanding wave propagation in the cochlea. Given their consideration of common pressure, however, a peculiarity is that, in setting boundary conditions, they discard the round window membrane. "Since the round window membrane separates the fluid in the scala tympani from the air in the middle ear it is reasonable to assume that the acoustic impedance terminating the scala tympani is zero" they say (p. 373). But it is because of the round window's stiffness that common-mode pressure arises in the first place. It almost produces a physical contradiction, for unless the cochlear channels are especially long and narrow, and the partition unusually stiff, there is no

[16] Peterson and Bogert calculate (p. 373) that a quarter-wave resonance would appear at 12 kHz.

way that the pressure at the round window can remain zero when the stapes moves. Another consequence, of course, is that their formulation exaggerates the differential pressure, placing the full pressure generated at the stapes across the partition; it also has the effect of exaggerating the common mode pressure.

More than 20 years elapsed until Geisler and Hubbard (1972) appreciated the limitations of the Peterson and Bogert work and refined the analysis[17] to specifically include round window stiffness. They called p_+ the 'fast' wave and p_- the 'slow' wave. They pointed out that the round window has a compliance of between 10^{-9} and 10^{-10} cm^5/dyne (measured by Békésy [p. 435] and equivalent to 10^{-14} m^3/Pa). It has an area of about 2 mm^2, so that it has an acoustic stiffness of 2×10^7 dyne/cm^3 or (2×10^{-2} N/m^3). Geisler and Hubbard used the same equation (3.5) as their starting point, but effected a considerable simplification by replacing $S(x)$ with a constant S. In justification, they remark that the cross-sectional area of the human cochlea is almost constant along its length (and as a side-effect making L_n about constant in Eq. 3.1); it also means that S in the numerator and denominator of Eq. 3.5 cancel, and we are left with a standard wave equation and its solution is

$$p_+(x) = \cos[\omega\,(l-x)c]\mathrm{e}^{i\omega t}/\cos(\omega\,l/c) \qquad (3.6)$$

which does not differ appreciably from the more complex Peterson and Bogert result. Introducing the round window stiffness, but eliminating the variation in S, gives the solution

$$p_+(x) = A\,\exp[i\omega\,(t-x/c)] + B\,\exp[i\omega\,(t+x/c)], \qquad (3.7)$$

which is a familiar standing wave (two waves propagating in the $+x$ and $-x$ directions, with A and B complex constants). An interpretation is that the fast wave reflects multiple times in the cochlea and, since the cochlea is small and of irregular shape, forms a complex longitudinal pressure field.

Geisler and Hubbard show how the unknown constants can be found by applying boundary conditions. This results in a somewhat more complex expression for the fast wave, although still of the standing wave form:

[17] Geisler, C. D. and A. E. Hubbard (1972). New boundary conditions and results for the Peterson–Bogert model of the cochlea. *J. Acoust. Soc. Am.* 52: 1629-1634.

$$p_+(x) = \frac{P_0\{\rho\omega^2 c + Kc[P'_-(0)/P_-(0)]\}[e^{i\omega(t-x/c)} + e^{-2i\omega l/c}e^{i\omega(t+x/c)}]}{(1+e^{-2i\omega l/c})\{2\rho\omega^2 c + Kc[P'_-(0)/P_-(0)] + K\omega\tan(\omega l/c)\}}, \qquad (3.8)$$

where P_0 is the sinusoidal pressure applied to the stapes, K is the acoustic stiffness of the round window membrane, l is the length of the cochlea, $P'_-(x)$ is the spatial derivative of $P_-(x)$ at $x = 0$, and the other symbols have their normal meaning.

For completeness, the corresponding slow (traveling) wave equation can be solved numerically[18], but for the boundary conditions specified, the differential pressure at the stapes can be explicitly stated as

$$p_-(0) = \frac{P_0[\rho\omega^2 c + K\omega\tan(\omega l/c)]e^{iwt}}{2\rho\omega^2 c + Kc[P'_-(0)/P_-(0)] + K\omega\tan(\omega l/c)}. \qquad (3.9)$$

At any point x_0 in the cochlea, the pressure in the upper gallery will therefore be $p_+(x) + p_-(x)$, while the pressure in the lower will be $p_+(x) - p_-(x)$. In this case, the fast wave is no longer independent of the slow one, and the two waves are coupled. Notice that if $p_-(x)$ is small, the dominant signal in the cochlea will be $p_+(x)$, and vice versa. At the low frequency limit, $p_-(x)$ will be at its lowest and the pressure will be about constant throughout the whole cochlea. Thus, if we are looking for common-mode effects, they are more likely to be apparent at low frequencies. Geisler and Hubbard describe this situation as the cochlear fluids acting essentially as a tube of incompressible fluid, with the round window moving out when the stapes moves in, and vice versa.

Geisler and Hubbard conclude that at mid-frequencies the initial stapes stimulus is shared *about equally* between the two modes (not unlike Fig. 3.3), and just above 10 kHz a resonance occurs because at this frequency the length of the human cochlea is a quarter-wavelength of the pressure wave. The windows will usually act piston-like and 180° out of phase, but when the high frequency resonance is approached the relative phases of the windows will rapidly switch as the driving frequency passes through the resonance.

Geisler and Hubbard increased the stiffness by a factor of 5, and, apart from some frequency shifts, saw little change in the behaviour of their model. The input

[18] Geisler (1972), p. 1630; for a broader perspective see also Geisler (1976).

impedance of their model cochlea was comparable at low frequencies to that measured by Békésy (*EiH*, p. 436) in a cadaver with the partition removed (leaving only the fluid and round window membrane). They also point out the similarity of their model to the results[19] of Wever and Lawrence (1950) who measured the phase responses of the two windows in a cat and found resonance-like behaviour near 9 kHz. This important work will be discussed in more detail later (§D 8.1/b), since the observed antiphase motion of the windows, and the finding of a *minimum* in cochlear microphonic response when the windows are stimulated *in phase*[20], appears, prima facie, to contradict the idea that outer hair cells respond to common mode pressure.

In summary, the Geisler and Hubbard model gives a physically accurate insight into the mechanics of the actual cochlea. It describes both a fast wave and a slow wave, the first of which is associated with common mode pressure, and the second with differential pressure. A traveling wave emerges from the action of the differential pressure, and that slow wave has remained the focus of cochlear mechanics, generating more and more detailed models. *The surprise is the readiness with which the fast wave has been deemed irrelevant.*

3.1/c Discarding common mode pressure

Since consideration of common mode pressure is a major point of departure in this thesis, the literature's short treatment of the fast wave is worth documenting.

1. The first hint that the standard model may be inadequate came from reading the exposition[21] of cochlear mechanics by Zwislocki (1980). He speaks of the Peterson and Bogert paper and claims (p. 173) that the pressure difference across the basilar membrane must be very small *and* that the pressure amplitude of the compressional waves must be small (because of the low impedance of the round window). Having both of these quantities small seems an ineffectual and unlikely outcome, so perhaps his other conclusion is open to question too: "Because hair cells

[19] Wever, E. G. and M. Lawrence (1950). The acoustic pathways to the cochlea. *J. Acoust. Soc. Am.* 22: 460-467.
[20] In particular, Wever and Lawrence (1950) and subsequent work which is discussed in §D 8.1/b.
[21] Zwislocki, J. J. (1980). Theory of cochlear mechanics. *Hear. Res.* 2: 171-182.

are excited as a result of deflection of their stereocilia rather than by pressure, compressional waves cannot be expected to play any direct role in the hearing process." Note that direct pressure measurements may not answer the question satisfactorily because drilling a hole in the cochlea will disturb the pressure field.

2. Lighthill (1991) refers to the fast wave[22] and says (p. 4) it is "uninteresting in another way as producing no motion of the cochlear partition. Accordingly, the fast wave becomes quite unimportant and I shall omit any further mention of it".

3. Shera and Zweig[23] simply say (p. 1363) that "the inner ear responds only to the pressure difference $P_{ow} - P_{rw}$ between the oval and round windows and not the absolute pressure at either window."

4. de Boer (1984) makes a one-sentence statement[24]: "The mechanical impedance of the round window is assumed to be zero". In his 1996 exposition, he devotes a paragraph to the "compressional wave"[25], but notes that the instantaneous pressure associated with it will be the same everywhere; thus, this component is considered "totally uninteresting" and not considered further.

5. Lindgren and Li (2003) began work with a double-sided transmission line model of the cochlea[26] that followed Peterson and Bogert's original Fig. 22 and so specifically included the compliance of the round window (see Fig. 3.4a). However, they are soon led to say that the stiffness of the round window is small compared to other stiffnesses and so they considered the pressure at the round window to be zero (p. 6). Thus, the round window disappears (see Fig. 3.4b).

---

[22] Lighthill, J. (1991). Biomechanics of hearing sensitivity. *Journal of Vibration and Acoustics* 113: 1-13.
[23] Shera, C. A. and G. Zweig (1992). Middle-ear phenomenology: the view from the three windows. *J. Acoust. Soc. Am.* 92: 1356-1370.
[24] de Boer (1984), p. 162.
[25] de Boer (1996), p. 263.
[26] Lindgren, A. G. and W. Li (2003). Analysis and simulation of a classic model of cochlea mechanics via a state-space realization. *unpublished manuscript*: http://www.ele.uri.edu/SASGroup/cochlea.html.

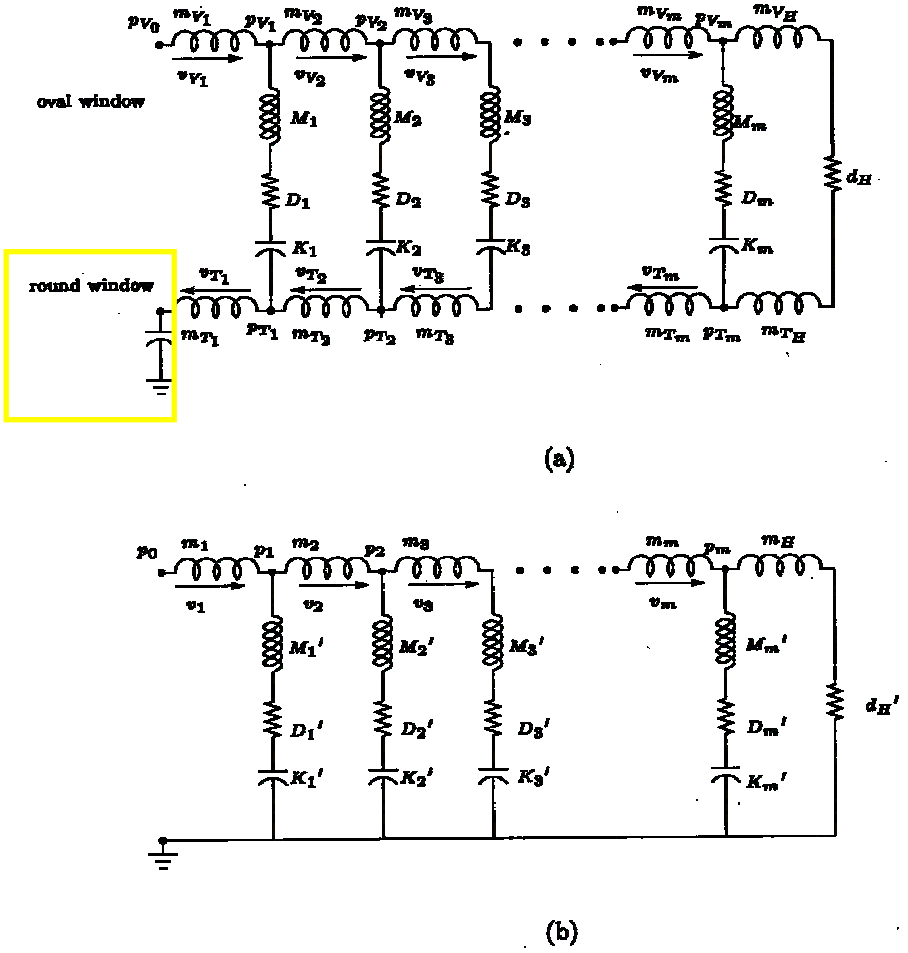

Fig. 3.4. A reasonably accurate model of the cochlea, the two-sided transmission line (top). Note the presence of a compliance (highlighted) representing the round window membrane which appears in series with the cochlear input. (Subscript *V* refers to scala vestibuli, *T* to scala tympani; the line is terminated by helicotrema mass m_H and damping d_H.) The authors also present an "equivalent" one-sided model (*b*) which, in omitting the round window, differs physically from (*a*). The diagrams are from Lindgren and Li (2003), and used with permission.

6. Baker (2000) sets out (his §3.3) to present[27] a mathematical development of the compressive wave. His aim is to develop piezoelectric amplification models of the cochlea. He notes that the compressive pressure field is symmetric about the partition, whereas the traveling wave of basilar membrane displacement is antisymmetric – thus "if one is interested in modelling basilar membrane motion, then one need not consider the compressive pressure wave. However, if one is interested in modelling fluid pressure measurement with the cochlear duct, then one must consider the compressive wave's contributions as well" (p. 63). He proceeds to

[27] Baker, G. J. (2000). *Pressure-feedforward and piezoelectric amplification models for the cochlea.* PhD thesis, Department of Mechanical Engineering, Stanford University.

develop governing equations but sets a boundary condition that "At the round window, the total pressure should be zero or very nearly zero" (p. 65). Evidence for this is that "the round window membrane is relatively large and compliant, and that the cochlear fluids do not flow out when the round window membrane is carefully removed". The first reason provides a useful simplifying assumption, but it tends to militate against the setting up of pressure fields. The round window's stiffness is important in allowing common mode pressure to exist at all (and in some creatures the round window is remarkably small[28] or stiff[29]). On the other hand, the mass of the cochlear fluids (and their small compliance) allows for pressure fields to establish at all frequencies above zero. The second reason has the limitation of applying only to static pressures and ignores surface tension effects. Overall, once it is acknowledged that outer hair cells may contribute significant amount of compressibility to the system, there are many possibilities for setting up a complex pressure field within the cochlea.

3.1/d The modern standard model

As said earlier, there is no intention of giving here a complete historical development of traveling wave theories. Allen (2001) provides a good perspective on the evolution of the field, and he discusses the way in which two- and three-dimensional models can improve the match between theory and experiment. However, to ward off complacency, he underlines (his §2.1) that "even a 3D model, no matter how much more frequency selective it was compared to the 1D model, would not be adequate to describe either the newly measured selectivity, or the neural tuning."

de Boer also gives a wide-ranging summary[30] of cochlear modelling, prefaced with the warning "*How can we be sure that we are extracting the "true" information or drawing the "right" conclusions?* [p. 259, emphasis in original]. He discusses the

[28] Gulick, W. L., et al. (1989). *Hearing: Physiological Acoustics, Neural Coding, and Psychophysics*. (Oxford University Press: Oxford). [p. 115]
[29] In whales and bats it is funnel-shaped, like a loudspeaker cone [Reysenbach de Haan (1956), pp. 83, 89-90]
[30] de Boer (1996).

intricacies of long- and short-wave models of the cochlea[31], as well as two- and three-dimensional models, second filters, active contributions from outer hair cells, and nonlinearity. Even so, longitudinal coupling, a real complication, needed to be ignored, and the chapter ends with a list of unsolved problems. The question posed again is (p. 307), "Haven't we left out something essential?" The following section of this thesis (§I 3.2) takes this question seriously.

Nevertheless, despite acknowledged limitations, traveling wave models have captured major features of cochlear behaviour. If there is one accepted standard modern model it is probably the 'coherent reflectance filtering' (CRF) model due to Shera and Zweig[32–35]. This model incorporates active elements and reverse traveling waves, for without both these features otoacoustic emissions could not arise within a traveling wave picture. The CRF model assumes that activity on the partition – mediated by outer hair cells – can cause a traveling wave to propagate in reverse towards the windows, where it is reflected at the stapes, and returns, via a traveling wave, to where it came. By multiple internal reflection, energy can in this way recirculate inside a longitudinally resonant cochlear cavity – and the end result is otoacoustic emissions.

Because of the appreciable length of the cochlear channels – some tens of millimeters – the theory establishes itself as a 'global oscillator' model, in contrast to the 'local oscillator' models of Gold and the like (included in which would be this thesis) where an oscillation emerging from the cochlea is traced back to a small group of outer hair cells on the partition. Because the traveling wave is broad, the CRF model cannot identify any single reflection point. It assumes that there is some 'spatial corrugation' or 'distributed roughness' inside the cochlea, so that scattering of a traveling wave occurs with a certain spatial regularity. The scattered wavefronts end up adding coherently in the opposite direction, and the result of this coherent reflection is acoustic emissions. The frequencies are not harmonically related, but there are an integer number of wavelengths in the round trip.

---

[31] The former still stands on a pedestal (p. 270).
[32] Shera, C. A. (2003). Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves. *J. Acoust. Soc. Am.* 114: 244-262.
[33] Zweig, G. and C. A. Shera (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J. Acoust. Soc. Am.* 98: 2018-2047.
[34] Shera, C. A. and J. J. Guinan (2003). Stimulus-frequency-emission group delay: a test of coherent reflection filtering and a window on cochlear tuning. *J. Acoust. Soc. Am.* 113: 2762-2772.
[35] Shera, C. A. and J. J. Guinan (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *J. Acoust. Soc. Am.* 105: 782-798.

The theory therefore places great emphasis on the relative phase of cochlear activity. A microphone in the ear canal measures regular peaks and valleys in pressure as frequency is swept, and CRF views these as an interference pattern produced by interaction of the forward and backward waves. Reflectance of waves at the stapes, R, will therefore have the form (Shera and Guinan, 1999, p. 795)

$$R \approx R_0 \, \mathrm{e}^{-2\pi i f \tau} \tag{3.10}$$

where f is frequency and τ is a time constant. Experimentally, from investigation of stimulus frequency otoacoustic emissions (SFOAEs), τ appears to be about 10 ms at 1500 Hz. Since it has the form of a delay, "it is natural to associate that delay with wave travel to and from the site of generation of the re-emitted wave" (ibid., p. 785). The phase of the reflectance therefore rotates rapidly, going through one full period over the frequency interval $1/\tau$, which corresponds to the spacing between neighbouring otoacoustic emissions. That is, near 1500 Hz, the interval will be about 100 Hz, so that neighbouring emissions will occur in the frequency ratio 1600/1500 ≈ 1.07.

Another way of expressing τ is in terms of the number of periods of the traveling wave in the recirculating loop, so that $\tau(f) = N/f$, and experiment shows (Fig. 3 of Shera and Guinan, 2003) that in humans N ranges from about 5 (at 500 Hz) to near 30 (at 10 kHz). That is, the cochlea stores between 5 and 30 cycles of acoustic signal. The phase can also be expressed in the following way (Shera and Guinan, p. 785)

$$\angle R = \Delta\theta_{\text{forward-travel}} + \Delta\theta_{\text{re-emission}} + \Delta\theta_{\text{reverse-travel}} \tag{3.11}$$

in which $\angle R$, the phase unwrapped from 3.10, is taken to be the sum of three phase delays, the forward travel time of the traveling wave, a phase lag due to the signal passing through the cochlear filter, and a phase delay for the reverse traveling wave. Zweig and Shera (1995) have emphasised that the cochlea possesses scaling symmetry, so that the number of waves in any traveling wave is about constant: a low frequency wave will travel further along the cochlea than a high frequency one and will require a longer time to reach its peak, but in terms of total phase shift it is

about the same in the two cases. That means that the first and last terms on the right-hand side of 3.11 are about constant, and means that nearly all of the observed phase variation seen from the ear canal must derive from the second term. It is my contention that in fact the first and third terms are practically zero and that *nearly all* of the observed phase derives from the high Q of the cochlear resonators.

The same point can be approached from a different direction. Konrad-Martin and Keefe (2005) consider the Q of the cochlear filters[36] in terms of the 'round-trip latency'[37] of Shera et al. (2002). Applied to SFOAEs, the latency amounts to Tf cycles of signal, where T is the measured latency and f is the frequency, and according to the Shera model, half of that latency ($Tf/2$) derives from the forward trip, and the other half (also $Tf/2$) from the reverse trip. Now the Q of the cochlear filters can be expressed as

$$Q = kTf/2 \qquad (3.12)$$

where k is a dimensionless measure of the filter shape. Experimentally, k is found to be about 2 when the basilar membrane delay is assumed to be *half* the SFOAE delay, making $Q \approx Tf$, which is just what we expect from a simple resonating filter, since the Q is equivalent to the number of cycles of build up and decay[38]. But the same result applies if we were to take k as 1 and the basilar membrane delay as simply *identical to* the filter delay. That is, the same results obtain whether k is set to be 1 (local oscillator model) or 2 (forward and reverse wave model).

Irrespective of what model one uses to interpret the results, the paper by Shera et al. (2002) is of interest in demonstrating that basilar membrane tuning in humans is appreciably sharper than previously thought. They used psychophysical studies conducted *near threshold* to show that the Q of the human cochlea is in the region of 15–20, values that are not as large as those calculated by Gold and Pumphrey, but indicative of high tuning nonetheless.

A large part of the argument for assuming that the traveling wave delay is not zero rests on showing that

---

[36] Konrad-Martin, D. and D. H. Keefe (2005). Transient-evoked stimulus-frequency and distortion-product otoacoustic emissions in normal and impaired ears. *J. Acoust. Soc. Am.* 117: 3799-3815.
[37] Shera, C. A., et al. (2002). Revised estimates of human cochlear tuning from otoacoustic and behavioural measurements. *Proc. Nat. Acad. Sci.* 99: 3318-3323.
[38] Fletcher, *Acoustic Systems in Biology*, p. 26.

$$\tau(f) = 2 \times \tau_{BM}(f) \qquad (3.13)$$

where $\tau_{BM}(f)$ is the group delay of the basilar membrane. The factor of 2 is what one expects if the traveling wave carries the signal around the loop. The theory is open to the criticism that, experimentally, the appropriate factor is somewhat less than 2, with the weight of evidence pointing to an actual factor of 1.7±0.2 (in the cat[39]), 1.6±0.3 (guinea pig[40]), and 1.86±0.22 (chinchilla and guinea pig[41]). However, a recent paper[42] claims that the discrepancy can be accounted for by use of a more realistic two-dimensional model.

Finally, the CRF theory introduces one distinctive mechanical feature of the cochlea which is worthy of note. Phase measurements reveal that while SFOAEs show the expected rapid rotation with frequency, the behaviour of distortion product otoacoustic emissions (DPOAEs) is radically different[43]. DPOAEs appear to be due to the interaction of the rapid rotation (slow time constant) with a very slow one (fast time constant). On this basis, Shera and Guinan identify two fundamentally different mechanisms: OAEs that arise by linear reflection and those that derive from nonlinear distortion. They set out a 'taxonomy' for acoustic emissions as set out in the table below.

[39] Shera and Guinan (2003), p. 2765.
[40] Shera and Guinan (2003).
[41] Cooper, N. P. and C. A. Shera (2004). Backward-traveling waves in the cochlea? *Association for Research in Otolaryngology, Midwinter Meeting, Abstract 342.* This reference concludes that its results rule out the pressure wave hypothesis, but in this it only treats the hypothesis in its one-way guise: the DPOAEs travel from basilar membrane to ear canal via a pressure wave, but the traveling wave is still considered to take the signal back the other way. This is the original picture of Wilson (1980), but the model I want to promote is that the pressure wave acts in both directions, and that the "basilar membrane delay" is in fact all filter delay (see §I 3.2/k).
[42] Shera, C. A., et al. (2005). Coherent reflection in a two-dimensional cochlea: short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions. *J. Acoust. Soc. Am.* 118: 287-313. See also §D 10.1/b.
[43] Shera and Guinan (2003), p. 2764.

A taxonomy for mammalian acoustic emissions (*Shera and Guinan, 1999*)	
Reflection source	**Distortion source**
Linear	Nonlinear
Rapid phase rotation	Slow wave rotation
SFOAEs, SOAEs, and TEOAEs	DPOAEs
Same frequency as stimulus (derive from near CF)	Frequency not in stimulus (require overlap of different TW peaks)
"place fixed"	"wave fixed"
High amplitude in humans, low in rodents	Maximum amplitude when $f_1/f_2 \approx 1.2$

Physically, the interpretation of the reflection source emissions (left column) is the one given above, in which there is one reverberating loop. By way of contrast, distortion sources (right column) arise, in the CRF view, from overlapping of the f_1 traveling wave peak and the f_2 peak, generating components at $2f_1 - f_2$ which travel to their own traveling wave maximum. The interactions become complicated, but the end result is a "wave fixed" emission that doesn't depend on a single place on the partition in the way that "place fixed" emissions do. Importantly, the DPOAE emissions can be separated into a quickly rotating component (slow wave) and a slowly rotating one (fast wave).

The rapidity of the fast wave is highlighted in §I 3.2/k, and a 'local' model for generation of DPOAEs is put forward in Chapter R7. It seems much more straightforward to see practically all the phase delay as deriving from the filter delay of a local resonator.

At this point we bring discussion of traveling wave theories to an end. We have enough detail to convey a picture of the traveling wave running forth (and back) along the basilar membrane, generating responses in hair cells above. This background has been preparation for listing situations where traveling wave theories cannot give a comprehensive account of cochlear mechanics.

3.2 Anomalies in traveling wave theory

The traveling wave model has been the mainstay in interpreting the results of cochlear experiments. The model seems to fit, in the main, and there have been some notable achievements in matching theory and experiment. And yet, there are recurring disparities that suggest that our understanding is not quite right. By outlining these major points of departure, the hope is that the underlying root of the problem may come to the fore. As someone once remarked, "paradox is truth standing on its head in order to draw attention to itself."[44] With that in mind, let us delve into the literature.

3.2/a The peak is so sharp

For a long time, the broad peak of the traveling wave was considered a virtue, for its associated low Q meant that hearing of transient sounds could begin and end quickly, without lag or overhang. But as improved experimental techniques showed increasingly sharp tuning of the basilar membrane[45], the problem became one of explaining how the traveling wave can give such a narrowly defined peak.

Some modern defining results include the following.

- Ren (2002) observed a traveling wave in a gerbil cochlea in response to 16 kHz tones and reported[46] that it occurred over a very restricted range (0.4–0.5 mm), even when the intensity varied from 10–90 dB SPL. Following death of the animal, response of the membrane was nearly undetectable and its tuning was lost.
- Nilsen and Russell (2000) saw sharp peaks in the tuning of a guinea pig basilar membrane[47] and evidence of radial phase differences.

[44] This saying is due, I believe, to Alan Watts (1916–1973).
[45] For quite some time mechanical tuning has been seen to be as sharp as neural tuning. Khanna, S. M. and D. G. B. Leonard (1982). Basilar membrane tuning in the cat cochlea. *Science* 215: 305-306.
[46] Ren, T. (2002). Longitudinal pattern of basilar membrane vibration in the sensitive cochlea. *Proc. Nat. Acad. Sci.* 99: 17101-17106.
[47] Nilsen, K. E. and I. J. Russell (2000). The spatial and temporal representation of a tone on the guinea pig basilar membrane. *Proc. Nat. Acad. Sci.* 97: 11751-11758.

When the animal died, responses dropped by up to 65 dB and phase gradients disappeared.

- Russell and Nilsen (1997) observed[48] in a similar investigation that the response to 15 kHz tones narrowed as intensity was reduced so that at 15 dB SPL, the peak was only 0.15 mm wide (the width of 14 inner hair cells). At 60 dB, the peak was more than a millimetre wide.
- Lonsbury-Martin et al. (1987) found histologically (Fig. 3.5) that the damage to a monkey's organ of Corti after exposure to loud pure tones was restricted to localised regions only 60–70 μm wide[49].

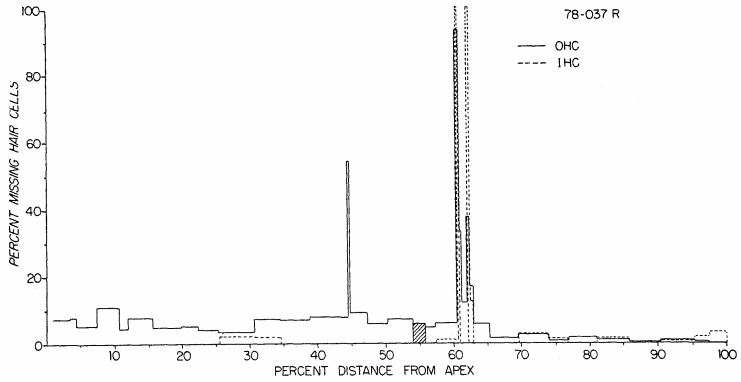

Fig. 3.5. Loss of inner and outer hair cells in the right ear of a monkey exposed long term to a wide range of pure tones at 100 dB SPL. Note the three sharp regions of high loss. The unexposed left ear showed no such peaks. [From Lonsbury-Martin et al. (1987) and used with the permission of the Acoustical Society of America]

- Lindgren and Li (2003) noted the discrepancy in the extent of excitation between their traveling wave model and the results of Ren (2002), but left it as inexplicable.
- Cody (1992) was puzzled[50] that neurally sharp tuning could remain in close proximity to regions damaged by overly loud sound. In one guinea pig, normal tuning and sensitivity were found within 0.5 mm of where 97% of outer hair cells were either missing or showed severe stereociliar damage.

[48] Russell, I. J. and K. E. Nilsen (1997). The location of the cochlear amplifier: spatial representation of single tone on the guinea pig basilar membrane. *Proc. Nat. Acad. Sci.* 94: 2660-2664.
[49] Lonsbury-Martin, B. L. and G. K. Martin (1987). Repeated TTS exposures in monkeys: alterations in hearing, cochlear structure, and single-unit thresholds. *J. Acoust. Soc. Am.* 81: 1507-1518.
[50] Cody, A. R. (1992). Acoustic lesions in the mammalian cochlea: implications for the spatial distribution of the 'active process'. *Hear. Res.* 62: 166-172.

In the opinion of Allen (2001), "The discrepancy in frequency selectivity between basilar membrane and neural responses has always been, and still is, the most serious problem for the cochlear modeling community. *In my view, this discrepancy is one of the most basic unsolved problems of cochlear modeling*."[51] While 2-D and 3-D models have improved matters, they have not narrowed tuning down to neural bandwidths. Active cochlear properties have opened the door to a gamut of signal processing strategies, but in Allen's view, a theory and computational model are still desperately needed to tie it all together. He lists a number of anomalies between basilar membrane and neural responses (his §2.2.6) which we do not have the space to consider in detail. However, to mention an issue that relates to the resonance mechanism examined in this thesis, he calculates that, despite the best 3-D models, the deficiency in "excess gain" – the additional basilar membrane gain at the characteristic frequency (compared to its surrounding frequencies) – is out by a factor of between 10 and 100 (20 to 40 dB) when compared to nerve fibre data[52].

de Boer (1996) also noted the poor match between models and experiment, even with short-wave 2-D and 3-D models. In no case does the response peak rise more than 10–15 dB above its surroundings. We might manipulate the parameters of the model, he observes (p. 281), but the dilemma is that either the amplitude of the peak remains too low or the phase variations in the peak region become too fast. He blames fluid damping, and makes a passing reference to the poor sound of an underwater piano (or carillon).

It is not often appreciated (or made clear by modelers) that there is flexibility in adjusting parameters to fit experimental data. Lesser and Berkley (1972) clearly spelt out that the process of matching experimental data to models is tricky[53]. They pointed out (p. 509) that the resistance term in Eq. 3.3 is not readily amenable to independent measurement and so, following Zwislocki's initial work, it is adjusted so as to yield agreement with the data. The mass term usually ends up larger than is physically plausible[54], even though some fluid will move with the partition.

[51] §2.1, italics in original.
[52] Last sentence of §2.2.6.
[53] Lesser, M. B. and D. A. Berkley (1972). Fluid mechanics of the cochlea. Part 1. *Journal of Fluid Mechanics* 51: 497-512.
[54] de Boer (1980), p. 166.

Similarly, Allen and Sondhi (1979) adjusted the damping term[55] until the best fit near CF was attained. Zwislocki (2002) adjusted parameters in an attempt to match his model to Békésy's narrow cochlear filter bandwidths; however, because the attempt failed (p. 156), Zwislocki was more inclined to suspect that the data was awry rather than consider his model wrong[56].

Active models provide even more adjustable parameters. Essentially, the active models allow for amplification stages between one transmission line stage and the next. This can work well in tuning frequency responses but it detracts from the physical realism of the model – in that the actual cochlea must make good use of all the signal energy available[57]. It cannot afford, like the 120-section electronic analogue[58] of Lyon (1988), to employ a cascaded amplifier gain of 1800 just to prevent the traveling wave from dying out. Again, Hubbard and Mountain describe[59] an active model by Neely and Kim (1986) in which a power gain[60] of 30 000 is called for. Zweig and Shera[61] have commented on the enormous gains typically required in active models to match theory with experiment. Gold, of course, would be quick to point out the danger of boosting a signal by 90 dB in the presence of unavoidable noise (§I 1.4).

3.2/b Doubts about the adequacy of the stiffness map

Even when the focus is kept on the stiffness of the embedded fibre, there are doubts that it can vary sufficiently between base and apex to tune the cochlea over 3

[55] Allen, J. B. and M. M. Sondhi (1979). Cochlear mechanics: time-domain solutions. *J. Acoust. Soc. Am.* 66: 123–132. [p. 128]

[56] Even though he gives the caveat (p. ix-x) that in auditory science "mathematical theory is often ignored or at least distrusted. In part, this is justified because the history of auditory research is full of examples of unrealistic mathematical and conceptual models that ignore existing experimental evidence" and contradict fundamental physical laws. Sentimentally, perhaps, he says (p. ix) that the simple classical picture of cochlear mechanics now has to be "reluctanctly" abandoned in the light of Kemp's findings.

[57] Although in passive (linear) transmission line models, the primary wave suffers a power loss of about 15 dB before it reaches its best frequency [de Boer (1980), p. 160. See also de Boer, p. 267 of *Mechanics and Biophysics of Hearing*, edited by P. Dallos et al. (Springer: New York, 1990)].

[58] Lyon, R. F. and C. Mead (1988). An analog electronic cochlea. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36: 1119-1134.

[59] Hubbard, A. E. and D. C. Mountain (1996). Analysis and synthesis of cochlear mechanical function using models. In: *Auditory Computation*, edited by H. L. Hawkins et al. (Springer: New York), 62–120. [p. 97]

[60] The ratio of power entering the system a given frequency to the power dissipated at the characteristic frequency.

[61] Zweig and Shera (1995), p. 2039.

orders of magnitude. The issue was first raised in connection with the tuning range of resonating fibres[62], and is summarised in Fig. 6.3 of de Boer (1980) and its associated discussion[63].

In general, a broad trend linking upper and lower hearing limits and cochlear width and thickness can be discerned across species, but the correlation is poor and is contradicted by certain specialised animals like horse-shoe bats and elephants[64]. In a developmental study of gerbil cochleas, it was found that a region that codes for the same frequency can have basilar membranes of very different dimensions, depending on age[65]. Treating the basilar membrane as having simple mass–spring resonance leads to difficulties. To vary the frequency by 10^3 means that the combined mass and stiffness needs to vary by a factor of 10^6. Since the mass is generally accepted as more or less constant[66], this requires stiffness (measured in terms of resistance to displacement by a probe, the 'point stiffness', divided by the width of the membrane) to vary a million-fold.

Measurements show that the stiffness of the basilar membrane varies by less than this. Békésy, for example, measured a stiffness variation (using a fluid pressure of 1 cm water, which generated about 10 µm deflection) of only a hundred-fold[67]. One possible avenue is to go beyond the simple two-dimensional picture and call on three-dimensional fluid–membrane interactions[68], although such a solution is by no means universally accepted.

The summary figure of de Boer (1980) shows stiffness variations (and characteristic frequency) plotted against distance from the stapes. Although the 100-fold variation of Békésy is depicted, his three data points obtained by pressing a hair on the membrane are also shown, and these are preferred because they show a 2.5 order of magnitude variation in stiffness over a similar variation in frequency – even

---

[62] The discussion in Chapter 2 following Equation 2.1.
[63] de Boer (1980). Auditory physics. Physical principles in hearing theory. I. Physics Reports 62, 87-174.de Boer Auditory physics. Physical principles in hearing theory. I.
[64] Echteler, S. M., et al. (1994). Structure of the mammalian cochlea. In: *Comparative Hearing: Mammals*, edited by R. R. Fay and A. N. Popper (Springer: New York), 134–171.
[65] Schweitzer, L., et al. (1996). Anatomical correlates of the passive properties underlying the developmental shift in the frequency map of the mammalian cochlea. *Hear. Res.* 97: 84-94. With age, the position representing a given frequency (11.2 kHz) shifted along the cochlea, being 90% from the base (near birth) and shifting to 65% (adult). To preserve place coding in accordance with traveling wave theory, the authors suggest that the *stiffness* of the partition must have changed.
[66] de Boer (1980), p. 166; Naidu and Mountain (1998), p. 130; Allen (2001), §1.3.1.
[67] Békésy (1960), p. 476.
[68] Steele, C. R. (1999). Toward three-dimensional analysis of cochlear structure. *ORL – Journal for Oto-Rhino-Laryngology and Its Related Specialities* 61: 238-251.

though this data requires an assumption that the point stiffness, which varies by 1.7 orders, can be realistically converted into an area modulus. Extrapolating this limited data appears to give a mapping with a suitably steep slope.

Work after Békésy was largely confined to measurements at or near the base until a provocative paper[69] by Naidu and Mountain (1998) confirmed Békésy's original findings: in experiments on isolated gerbil cochleas, they could only measure a variation of 56 in the pectinate zone of the basilar membrane (below the outer hair cells) and a factor of 20 in its arcuate zone. Making allowance for variations in the width of the basilar membrane, they found a final volume compliance ratio of about 100 between base and apex. They conclude (p. 130) that "conventional theories that explain cochlear frequency analysis based on an enormous stiffness gradient and simplistic motion of the OC require substantial modification."

One attempt at explaining cochlear tuning is due to Wada et al. (1998) who measured thickness and length along the whole of the guinea pig cochlea[70]. Based on a computerised reconstruction and beam model, they found that the natural frequency at the basal turn was only 3.1 times that at the apical turn, assuming that the Youngs modulus and diameter of the constituent fibres was constant. Given that the variation was inadequate to produce wide-range tuning, the authors conclude that the assumption must be wrong, and that the modulus must vary. Unfortunately, direct evidence (which they cite on p. 5) shows that the Youngs modulus of human basilar membrane only varies by 50% between base and apex, so the question remains.

Inadequate variation in tuning also emerged from another finite-element model of the cochlea[71]. In this case, geometry alone gave a 2-fold change, and allowing for stiffness variations a 20-fold difference between base and apex resulted. A way around the limitation is, the authors suggest, to suppose – ad hoc – that hair cells in the apex respond to a first vibrational mode while hair cells in the apex respond to a second.

---

[69] Naidu, R. C. and D. C. Mountain (1998). Measurements of the stiffness map challenge a basic tenet of cochlear theories. *Hear. Res.* 124: 124–131.

[70] Wada, H., et al. (1998). Measurement of guinea pig basilar membrane using computer-aided three-dimensional reconstruction system. *Hear. Res.* 120: 1–6.

[71] Zhang, L., et al. (1996). Shape and stiffness changes of the organ of Corti from base to apex cannot predict characteristic frequency changes: are multiple modes the answer? In: *Diversity in Auditory Mechanics*, edited by E. R. Lewis et al. (World Scientific: Singapore), 472-478.

An effort to meet Naidu and Mountain's challenge was made by Emadi et al. (2004), who used a vibrating stiffness probe on the basilar membrane of gerbils[72] at various radial and longitudinal locations. In their unidirectional measurements, they focus on the minimum of the parabolic stiffness, values of which they took to reflect the basilar membrane fibres[73]. Of the four radial positions at which they took readings, three of them gave a longitudinal gradient comparable to those of Naidu and Mountain. However, the fourth, measured at the mid-pectinate location, gave a steeper longitudinal gradient (–5.7 dB/mm) than Naidu and Mountain (–3.0 dB/mm), and the authors argue that this set of data is the most relevant[74]. Putting this value into a simple resonance model and into a 3D fluid model, they calculate an excellent match between stiffness variation and frequency ratio between base and apex.

As a critique, I would argue that, since fluid pressure over the whole membrane is the physiological stimulus, an average of all positions would be more representative. Moreover, the statistics of the analysis are marginal, in that the gradient of the line through the three error-barred points in their Fig. 5D carries large uncertainties. The primary author says[75] that the 95% confidence limits on the gradient are –6.2 and –3.0 dB/mm, the last figure corresponding to the gradient they wish to dispute. Moreover, the figures derive from averaging, after 5-point (5-μm) smoothing, all data from 1 μm deflection to 17 μm, and this processing may not yield the physiologically relevant value, particularly when most of the curves shown in Fig. 5B have non-linear slope (either less or more than 1 dB/dB, as shown in Fig. 5C). The non-linearity is a good reason to suspect that the statistical model applied to the data is not valid.

Nevertheless, it is true that fluid models do provide a way of expanding the tuning range for a given stiffness range. The model[76] used by Emadi et al., and its later form of development[77,78], do give wide-range tuning; the difficulty is accepting

---

[72] The experiments were done both in vivo (base only) and in vitro (on a hemicochlea). As mentioned in §I 3.2/b, these authors used *unidirectional* probing of the basilar membrane.

[73] Although they acknowledge (p. 483) that the physiologically relevant stiffness may occur at smaller tissue deflections and be buried in the noise.

[74] The authors should not have expressed their findings in decibels, which applies to power, but their meaning, in terms of a ratio change of stiffness per millimetre, is clear enough.

[75] Personal communication to T. Maddess 2005/02/05.

[76] Steele, C. R. and J. G. Zais (1983). Basilar membrane properties and cochlear response. In: *Mechanics of Hearing*, edited by E. de Boer and M. A. Viergever (Delft University Press: Boston, MA), 29-36.

[77] Steele Toward three-dimensional analysis of cochlear structure.

the underlying finite-element model, which has some peculiar features. For example, the fluid pressure in the spiral sulcus is assumed central in stimulating inner hair cells, so that Steele argues (p. 241) that the whole purpose of the organ of Corti is to develop that pressure. He also uses (his Table 1) a Youngs modulus of 1 GPa for *all* parts of the organ of Corti, including the tectorial membrane but excepting Hensen cells, which seems overly simplistic. For example, in Chapter 5 measurements of the stiffness of the tectorial membrane are examined and values in the region of some kilopascals seem most appropriate.

In conclusion, therefore, real doubts remain about being able to achieve a satisfactory range of tuning and, as Allen (2001) remarks, 3D models do not, without some radical assumptions, provide adequate sharpness.

3.2/c The spiral lamina is flexible

The basilar membrane is supported on its outer side by the spiral ligament and on its inner side by the (osseous) spiral lamina. While the width of the partition is about constant along its length, the basilar membrane is relatively wide at the apex and tapers to its narrowest at the base. This arrangement suggested to Helmholtz, and to many since, that the basilar membrane is tonotopically tuned via its width. The problem, as pointed out by Kohllöffel[79] (1983), is that the spiral lamina is in many animals as flexible as the basilar membrane. This author says (p. 215) that in unfixed human preparations the spiral lamina deflected as much as the basilar membrane (over the region 3–14 mm from the base when vibrated at frequencies up to 1 kHz). Using a hair probe, the human spiral lamina deflected nearly as much as the round window membrane.

Interestingly, the flexibility of the spiral lamina was noted as early as 1680 and formed the basis of DuVerney's cochlear frequency analysis idea in 1684. Its

[78] Steele, C. R. and K.-M. Lim (1999). Cochlear model with three-dimensional fluid, inner sulcus and feed-forward mechanism. *Audiol. Neurootol.* 4: 197-203.
[79] Kohllöffel, L. U. E. (1983). Problems in aural sound conduction. In: *Mechanisms of Hearing*, edited by E. de Boer and M. A. Viergever (Delft University Press: Delft), 211-217.

flexibility is confirmed by a recent study[80] in which the amplitude and tuning of the lamina in human cadavers was examined with a laser vibrometer. When exposed to air-conduction stimuli, the motion of the lamina (at 12 mm from the round window) was comparable to – and at some frequencies exceeded – the motion of the adjacent basilar membrane[81].

If the whole partition is flexible (and nearly constant in width), it removes one more factor by which tonotopic tuning can be produced.

3.2/d The basilar membrane rests on bone

If the basilar membrane were essential for hearing, as the traveling wave theory supposes, then we would invariably find it present in a functioning cochlea. That is not always the case.

In some cases we find a well formed organ of Corti, but it rests on solid bone, not the basilar membrane. Fig. 3.6 shows a microscopic section made by Shambaugh (1907) of the organ of Corti of a pig[82] sitting upon solid bone, one of several observations of the basilar membrane that made Shambaugh think that its "thick, inflexible character" makes it an unsuitable candidate as a vibrating structure. He thought the tectorial membrane, which was always associated with the organ, a much better candidate.

One may be tempted to argue that the pig was deaf. However, work in the 1930s by Crowe, Guild, and Polvogt (cited by Tonndorf[83] 1959) indicates otherwise. In a post mortem study of human temporal bones, Polvogt and colleagues compared the results with audiograms taken before death and found that the person with a similar bony projection could hear, at least for frequencies lower than those corresponding to the site of the abnormality.

---

[80] Stenfelt, S., et al. (2003). Basilar membrane and osseous spiral lamina motion in human cadavers with air and bone conduction stimuli. *Hear. Res.* 181: 131-143.
[81] Stenfelt (2003), Fig. 5a.
[82] Shambaugh, G. E. (1907). A restudy of the minute anatomy of structures in the cochlea with conclusions bearing on the solution of the problem of tone perception. *Am. J. Anat.* 7: 245–257 (+ plates).
[83] Tonndorf, J. (1959). The transfer of energy across the cochlea. *Acta Otolaryngol.* 50: 171–184.
[p. 182]

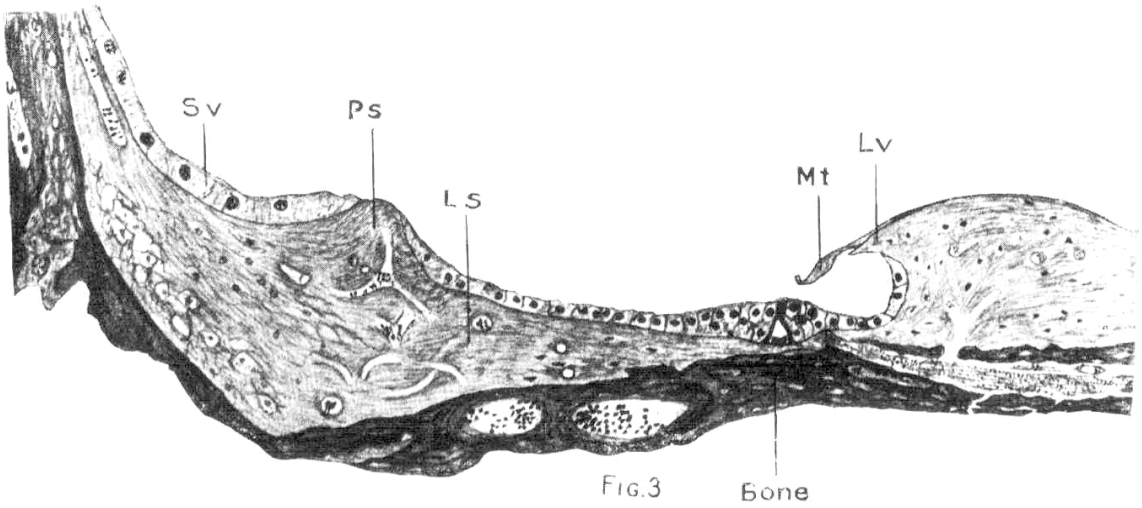

Fig. 3.6. A pig's organ of Corti, perfectly formed, sitting upon a solid bony plate. [From Fig. 3 of Shambaugh (1907)]

3.2/e Holes in the basilar membrane

1. That same temporal bone study[84] found other malformations in which there was either a hole in the basilar membrane or the bone separating one cochlear turn from another was lacking. In the first case there was open communication between the upper gallery of the first turn and the lower one of the second; in the second, two cochlear ducts stretched across one common channel. Again, the hearing thresholds of the affected ears were indistinguishable from those in the opposite, normally constructed, ears. One might predict that such holes would short-circuit a traveling wave, destroying sensitivity to all frequencies apical to the hole, but this did not happen. Another experiment reported by Tonndorf (loc. cit.) leads in a similar direction: Tasaki, Davis, and Legouix (1952) induced open communication between the adjacent turns of a guinea pig cochlea and found that cochlear microphonics apical to the injury site were unaffected.

2. In many species of birds there is a naturally occurring shunt through the basilar membrane called the ductis brevis[85,86]. In contrast to the helicotrema, it connects the galleries *at the basal end*. According to Kohllöffel, it is variable in size

[84] Polvogt, L. M. and S. J. Crowe (1937). Anomalies of the cochlea in patients with normal hearing. *Ann. Otol. Rhinol. Laryngol.* 46: 579-591.
[85] Kohllöffel (1983)
[86] Kohllöffel, L. U. E. (1984). Notes on the comparative mechanics of hearing. II. On cochlear shunts in birds. *Hear. Res.* 13: 77-81.

and occurrence, being absent in owls and extremely narrow in turkey, pheasant, and quail; in contrast, it is present in pigeon, woodpecker, duck, and songbirds, and is especially wide in goose, reaching a diameter of 0.6 mm. The anatomy of birds forced Helmholtz to reconsider his theory, and it appears these creatures once more prompt us to re-examine our models.

3. Finally, let us look more closely at normal human anatomy. We tend to accept the presence of the helicotrema as a convenient way for pressure in the two galleries to be equalized. The hole, about 0.4 mm^2 in area, connects two ducts of about 1.2 mm^2 (*EiH*, p. 435).

But the form of the hole, as shown in Fig. 3.7, invites comment. There is no differential pressure at the helicotrema – it behaves hydraulically as a short circuit – and yet the organ of Corti retains the same form here as it does elsewhere in the cochlea: positioned near the apex of the triangular cochlear duct, but *without a basilar membrane underneath*. The question needs to be asked, are the hair cells at the helicotrema functional, because, if they are, they do not appear to be stimulated by motion of a basilar membrane.

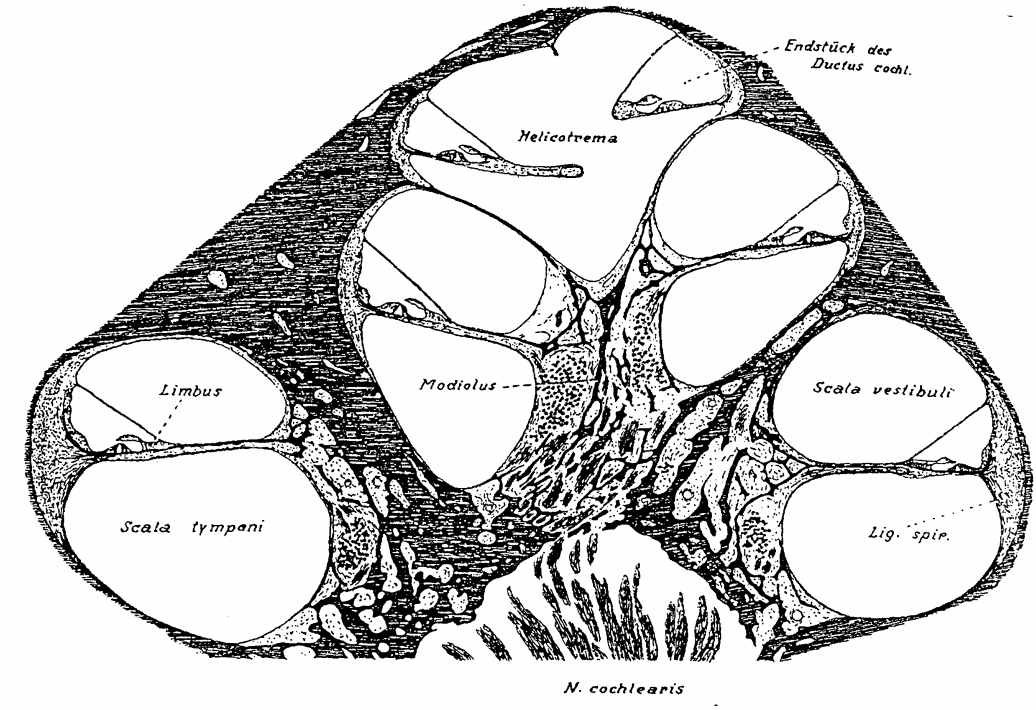

Fig. 3.7. Human cochlea, showing the form of the cochlear duct at the helicotrema[87]. The organ of Corti retains its standard form, even though the differential pressure is zero. [From Fig. 9 of Neubert (1950) and reproduced with permission of Springer-Verlag]

[87] Neubert, K. (1950). Die Basilarmembran des Menschen und ihr Verankerungssystem: ein morphologischer Beitrag zur Theorie des Hörens. *Zeitschrift für Anatomie und Entwicklungsgeschichte* 114: 539-588. A similar picture is depicted in de Boer (1984), Fig. 3.1a.

3.2/f Zero crossings

As pointed out by Shera (2001), the cochlea possesses a remarkable symmetry[88]. As the intensity of stimulation increases, the zero crossings of the basilar membrane response (and acoustic nerve firings) stay fixed. Although the waveform's centre of gravity moves to shorter times, the zero points stay put, as Fig. 3.8 makes plain. The phenomenon rules over nearly the entire dynamic range of the cochlea.

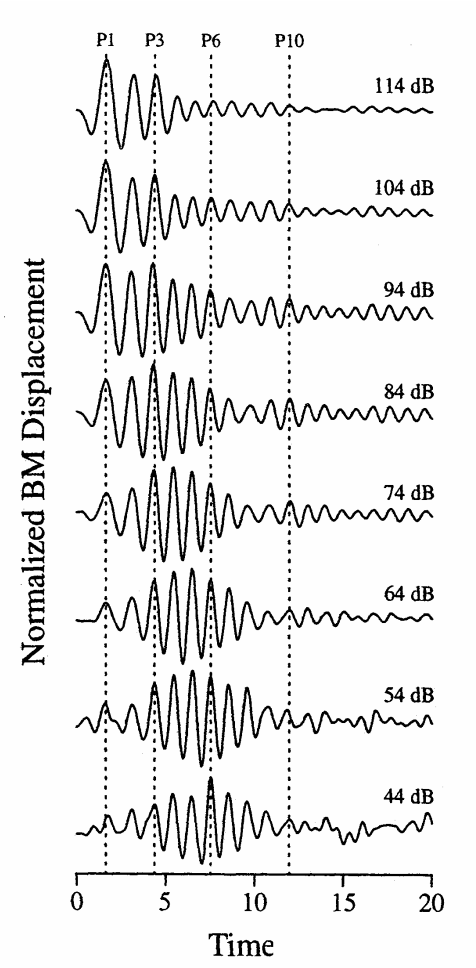

Fig. 3.8. Fixed zero crossings. As the intensity of a 1-kHz tone was raised from 44 to 114 dB, the basilar membrane motion of a chinchilla was monitored by a laser vibrometer. The time is in periods of 14.5 kHz (CF). The structure of the wave form in the time domain stays virtually constant. [From Recio and Rhode (2000) via Shera (2001), and used with permission of the Acoustical Society of America]

[88] Shera, C. A. (2001). Intensity-invariance of fine-structure in basilar-membrane click responses: implications for cochlear mechanics. *J. Acoust. Soc. Am.* 110: 332-348.

Physically, the effect only makes sense, says Shera, if the local resonant frequencies of the partition are nearly independent of intensity. This places strong constraints on the way that outer hair cells work, calling for the cochlear amplifier not to affect the natural resonant frequency of its surroundings as it works to supply feedback forces. In fact, *it contradicts many, if not most, cochlear models* (pp. 332, 345). In particular, it rules out all those models that require the outer hair cells to alter the stiffness (and impedance) of the partition.

Shera presents a detailed mathematical analysis of how a harmonic oscillator interacts with a traveling wave, and how the dispersion of the latter introduces time and frequency effects. He sets out certain conditions under which the oscillator's poles may stay fixed, but in general a traveling wave model will fail this requirement[89]. On the other had, it seems clear that a pure resonance model – such as the SAW model – will cope much better in meeting this condition: the independent oscillators will just gain strength as stimulus intensity is raised and the frequency (and time) structure will be preserved.

de Boer and Nuttall (2003) recognise the peculiarity of the zero crossings[90], but cannot suggest an answer. In fact, their active 'feed-forward' model doesn't help because it is 'non-causal', meaning that motion of the basilar membrane at one point would instantaneously affect points further away[91]. Chadwick (1997) saw the drawback of such non-causality[92], and remarked that it would mean a non-unique, non-realizable, and less useful model. I agree that this way of refining traveling wave models strains understanding, although it may be useful to see that from a traveling wave perspective a fast pressure wave is in fact non-causal.

---

[89] Cooper (2004) explains how low-frequency components will travel further and slightly faster than the high-frequency components, and low-intensity sounds will travel slightly further and slightly more slowly than higher intensity ones. [Cooper, N. P. (2004). Compression in the peripheral auditory system. In: *Compression: From Cochlea to Cochlear Implants*, edited by S. P. Bacon et al. (Springer: New York), 18-61.] It makes one ask how the auditory system, on this basis, can disentangle the components.

[90] de Boer, E. and A. L. Nuttall (2003). Properties of amplifying elements in the cochlea. In: *Biophysics of the Cochlea: From Molecules to Models*, edited by A. W. Gummer (World Scientific: Singapore), 331-342.

[91] A system is causal if it doesn't depend on future values of the input to determine its output. A non-causal system senses an input coming and gives an output before it does (Antoulas and Slavinksy, http://cnx.rice.edu/content/m2102/latest/)

[92] Chadwick, R. S. (1997). What should be the goals of cochlear modeling? *J. Acoust. Soc. Am.* 102: 3054. Subsequently, de Boer defended his model [de Boer, E. (1999). Abstract exercises in cochlear modeling: reply. *J. Acoust. Soc. Am.* 105: 2984.]

3.2/g Hear with no middle ear

Before the days of antibiotics, it would be common for a middle ear infection to escalate to the point where there was total loss of the middle ear, including ear drum. The result was that the person was left only with oval and round windows, which opened directly to the ear canal. Surprisingly, such people do not suffer total hearing loss; they lose some 20–60 dB in sensitivity, but they can still hear, more so at low frequencies than high. In terms of the traveling wave theory, that is a major anomaly, because there should be no pressure difference across the partition to generate a stimulus, and any phase difference between the windows should virtually disappear at low frequency.

Békésy recognised the contradiction, and sought to explain it (*EiH*, p. 105–108). He suggested that the cochlea was not incompressible, so that even when sound impinged on the two windows in phase, the pressure could cause some movement of the windows. He imagined that some of the cochlear fluids could surge in and out of the cochlea through blood vessels or the "third windows" of the vestibular and cochlear aqueducts. If fluid flow is easier on the stapes side (it short-circuits stapes pressure), the round window pressure will force fluid to deflect the basilar membrane in a direction opposite to the usual – and hence the phase perception will be 180° different. Sound localisation experiments indeed show that, remarkably, people with only one middle ear hear sound 180° out of phase in that ear (*EiH*, Fig. 5-12), which tends to confirm Békésy's conjecture[93].

Through introducing this mechanism, traveling wave theory can avoid an inherent contradiction. However, it is mentioned here as a signal that an alternative explanation is possible: that the outer hair cells can be stimulated directly by pressure.

No matter what model one chooses, the "middleless" ear configuration provides major constraints on the compressibility of the cochlea, as Shera (1992) calculated[94]. He uses a network model and Békésy's data to show that the degree of

[93] However, this fact is interpreted differently in Ch. 9 where it is used as evidence that the ear uses two detection systems: a pressure-detection one involving the outer hair cells, and a deflection mechanism involving the inner hair cells.

[94] Shera, C. A. and G. Zweig (1992). An empirical bound on the compressibility of the cochlea. *J. Acoust. Soc. Am.* 92: 1382-1388.

compressibility (ε, the ratio between the stiffness of the organ of Corti to the compressional stiffness of scala media) must be less than a few percent, but greater than zero (p. 1385). Similarly, Ravicz and colleagues[95] performed experiments on cochleas of human cadavers and expressed the compressibility as an upper bound on a parameter α, where α was the ratio of the motion of the stapes with the round window blocked to the motion normally. They found α to range from 0.015 (at 500 Hz) to 0.5 (at 30 Hz).

The result obviously depends on the model, but it is worth keeping in mind as we consider how compressible living outer hair cells might be (Ch. D8) and evidence advanced against any compressibility whatever (§I 3.3/c below).

3.2/h Hear with blocked round window

The traveling wave depends on a pressure difference between the oval window and the round window. Blocking the round window may then be expected to be a recipe for total loss of hearing. Again, that isn't the case.

The most startling and clear-cut instances are congenital, in which a person is born without a round window, an uncommon malformation called round window atresia[96]. Instead, the round window niche is filled with bone and the person suffers a hearing loss of 30–40 dB.

Martin et al. (2002) confirmed the diagnosis in a bilateral case with high-resolution CT scans and found associated hearing losses of about 40 dB[97]. They held back from surgery because they regard fenestration as likely to cause hearing loss rather than improvement. They mention (p. 801) that when round window absence is found in combination with stapes fixation, surgery is likely to lead to *hearing loss* (that is, they advise not interfering because the person is still able to hear). Linder et al. (2003) also did CT scans to confirm the condition in two cases and expressed puzzlement (p. 262) at the limited hearing deficit (40 dB or so), as they expected

---

[95] Ravicz, M. E., et al. (1996). An upper bound on the compressibility of the human cochlea. Midwinter Meeting, Association for Research in Otolaryngology (http://www.mit.edu/~ajmiller/epl/RaviczARO1996.pdf)

[96] Linder, T. E., et al. (2003). Round window atresia and its effect on sound transmission. *Otology and Neurotology* 24: 259-263.

[97] Martin, C., et al. (2002). Isolated congenital round window absence. *Ann. Otol. Rhinol. Laryngol.* 111: 799-801.

complete conductive hearing loss. They conclude that, instead of a traveling wave, the cochlea is sensitive to an alternative way of stimulation[98].

On the other hand, it is recognised that the pressure detection idea is not immune from criticism on this count either. If it were valid, then blocking the round window might be expected to increase hearing sensitivity, and this clearly does not happen. This problem is addressed in §D 8.2.

3.2/i Hear with no tectorial membrane

According to the standard model, motion of the basilar membrane causes shear between the stereocilia of the hair cells and the tectorial membrane. In a situation where the tectorial membrane is absent, one might reasonably conclude that all hearing would be absent, but again the real situation confounds this outlook.

It is possible to produce a knock-out gene in mice that leads to the loss of α-tectorin, an essential component of the tectorial membrane matrix[99]. The result is that the affected mice have a completely detached tectorial membrane, and suffer a hearing loss of 35–40 dB. In terms of the standard model, however, we might expect to see total hearing loss, yet the authors report that outer hair cells still respond to sound with a reduced cochlear microphonic. How is that possible? One commentator[100] suggested that the stereocilia must be moved by fluid drag, which could happen if the hair cells were moving side to side as well as up and down.

However, in terms of the SAW model, a 40 dB loss is understandable because an essential feedback path is absent. However, it is still possible for the outer hair cells to provide an electrical response because they are reacting to the pressure wave, not a traveling wave. However, since the cells can no longer interact across rows,

[98] In the older literature, a mass of conflicting results on the role of the round window can be found, and not much ground can be gained by discussing it here. Many experimenters attempted to block the existing round windows of animals, but the problem is the extreme difficulty of achieving a true block, as even a minute air cavity would, given the vanishingly small displacements involved in hearing, provide sufficient compliance in the system. Nevertheless, they raise a real doubt about the necessity of a round window. One major doubt-raising paper [Hallpike, C. S. and P. Scott (1940). Observations on the function of the round window. *J. Physiol.* 99: 76-82.] favoured Pohlman's idea of pressure-sensitive cochlear receptors (p. 81).

[99] Legan, P. K., et al. (2000). A targeted deletion in α-tectorin reveals that the tectorial membrane is required for the gain and timing of cochlear feedback. *Neuron* 28: 273-285.

[100] Corey, D. P. (2000). Sound amplification in the inner ear: it takes TM to tango. *Neuron* 28: 7-9.

DPOAEs are impossible (see Chapter 7), and this is just what Legan et al. found (p. 276).

3.2/j The casing of the cochlea is exceptionally hard

The otic capsule in which the cochlea and its fluid contents are encased is noteworthy in its own right. The capsule is made of bone and, in humans, sits within the base of the skull. Remarkably, this bone is immensely hard, and this ivory-like bone is the hardest in the human body. Despite its small size, the otic capsule derives developmentally from 14 distinct ossification centres and the initial fetal architecture is maintained throughout adult life[101].

In whales, the inner ear is separate from the skull, so that it forms a spherical mass, the os perioticum, that has been described as a very compact, stony-hard 'glasslike' bone[102], the densest and hardest bone known in the animal kingdom. It is so solid that opening a specimen is a difficult and tedious job, often resulting in fracturing. When an animal dies, its soft spongy skeleton soon decays, but the os perioticum remains, littering the sea floor for millennia[103]. The question therefore arises, why is the otic capsule so hard?

In terms of the standard model, this design effort is superfluous: all it is has to do is contain the fluid contents and be adequately stiffer than the basilar membrane. Given the high compliance of the round window, the dense bone seems unnecessarily hard.

On the other hand, if the cochlea is designed to detect acoustic pressure, then the outlook changes considerably. In this case, the difference in acoustic impedance between the cochlear fluids and that of the capsule is crucial: acoustic energy will leak out of the cochlea unless there is a large ratio in acoustic impedance between the two. In terms of the proportion of energy reflected (R) at a boundary, Fletcher (1992) shows that (p. 98)

---

[101] Jahn, A. F. (1988). Bone physiology of the temporal bone, otic capsule, and ossicles. In: *Physiology of the Ear*, edited by A. F. Jahn and J. Santos-Sacchi (Raven: New York), 143-158.

[102] Reysenbach de Haan, F. W. (1956). Hearing in whales. *Acta Oto-Laryngologica Suppl.* 134: 1-114. [p. 44]

[103] Fossilised inner ear bones of whales, millions of years old, are therefore common [as a Google search confirms] and can be bought over the internet for a few dollars.

$$R = (z_2 - z_1)^2 / (z_1 + z_2)^2 \qquad (3.14)$$

where z_1 and z_2 are the acoustic impedances of the two media. To prevent loss of acoustic energy from the cochlear fluids (z = 1.5 × 10^6 rayl) into the skull, it therefore becomes important to make the acoustic impedance (ρc) of the surrounding bone as high as possible. The speed of sound in typical bone is about twice that in water, and its density greater by a similar factor, so that, using Eq. 3.14, we find that the energy reflected at the interface is only about a third. By making the density of cochlear bone double that of ordinary bone, the reflectance figure can be made to exceed 50%, which is then beginning to become a useful figure for containing the pressure wave. This is even more vital in the case of whales, because the loss will be not just to the skull, but to the surrounding water[104].

Interestingly, in guinea pigs and other rodents, the cochlea is encased only in thin bone and projects from the skull of the animal, a configuration greatly different to that in humans[105] and leading to the suspicion that the hearing process in these animals may differ in important respects from ours.

The issue of optimal design of the cochlea is a major one, and will be dealt with in the concluding chapter. Nevertheless, it is worth noting here that if the cochlea is configured to detect the fast pressure wave, then the presence of a round window nearby to an oval window appears counterproductive: if the windows move out of phase to each other, then the pressure waves generated by each would tend to cancel, leaving an evanescent wave which should decay rapidly with increasing distance from the windows. To avoid cancellation, a better design for pressure detection would be to have the round window at the far (apical) end of the cochlea, but the possibility is nonetheless worth exploring.

3.2/k Fast responses

Traveling waves progress relatively slowly, starting at the base at about 100 m/s and slowing down until at the apex their speed is more like 1 m/s. The slow

[104] Whale hearing is discussed in more detail in §3.3, where the special arrangement of their round window is outlined.
[105] Wysocki, J. (2005). Topographical anatomy of the guinea pig temporal bone. *Hear. Res.* 199: 103-110.

speed is understood as the basis of the time delay in Kemp echoes, where the travel time to a tuned place on the partition and its return to the stapes underlies the long delays observed[106]. Here we note cases where the cochlear responses are too fast to be mediated by a traveling wave, and instead a fast pressure wave must be carrying them.

1. *Wilson hair-cell swelling model.* The credit for seeing an association between outer hair cells and fast pressure waves must, in modern times, go to Wilson. Although he kept to the idea of traveling waves exciting the cells, he proposed that the activity of outer hair cells produced a change in volume, so that in this way a pressure wave could be *generated* and return almost instantly to the ear canal, thereby explaining Kemp echoes. This *hair-cell swelling model* was presented in a 1980 paper[107] as a way of explaining why the cochlear microphonic recorded at the round window in response to a low-level 800-Hz tone burst appeared to occur simultaneously with the ear canal pressure recorded with a microphone.

Since the electrical signal was instantaneous, so too must have been the acoustic signal conveyed from the outer hair cell, and a pressure wave is the only signal carrier fast enough. The data in the paper are not convincing by themselves, but the idea is an engaging one. However, in going against the mainstream it has not caught on[108].

If there is a negligible reverse travel time, where do long response times come from? Wilson still retained the idea that acoustic stimulation reached the outer hair cells via the traveling wave, but, even allowing for that forward propagation path, the long total echo delay meant an extra source of delay must be sought. He attributed this to the delay inherent in building up oscillation in a narrowly tuned filter – the 'second filter' delay. Noting that such an extended delay means the corresponding tuning would be uncommonly narrow – narrower than observed inner hair cell tuning – he suggested that the outer hair cells are actually more sharply tuned than inner hair cells. We are back into Gold territory.

---

[106] Even though making the time delay equal to twice the travel time is not as straightforward as it seems. For example, Shera and Guinan (2003) find that the delay is about 1.3–1.9 times the one-way travel time.

[107] Wilson, J. P. (1980). Model for cochlear echoes and tinnitus based on an observed electrical correlate. *Hear. Res.* 2: 527-532.

[108] It was considered briefly on pp. 522-523 of de Boer, E. (1980). Nonlinear interactions and the 'Kemp echo'. *Hear. Res.* 2: 519-526. Wilson's name was not attached to it, but the idea was considered unlikely and not considered further.

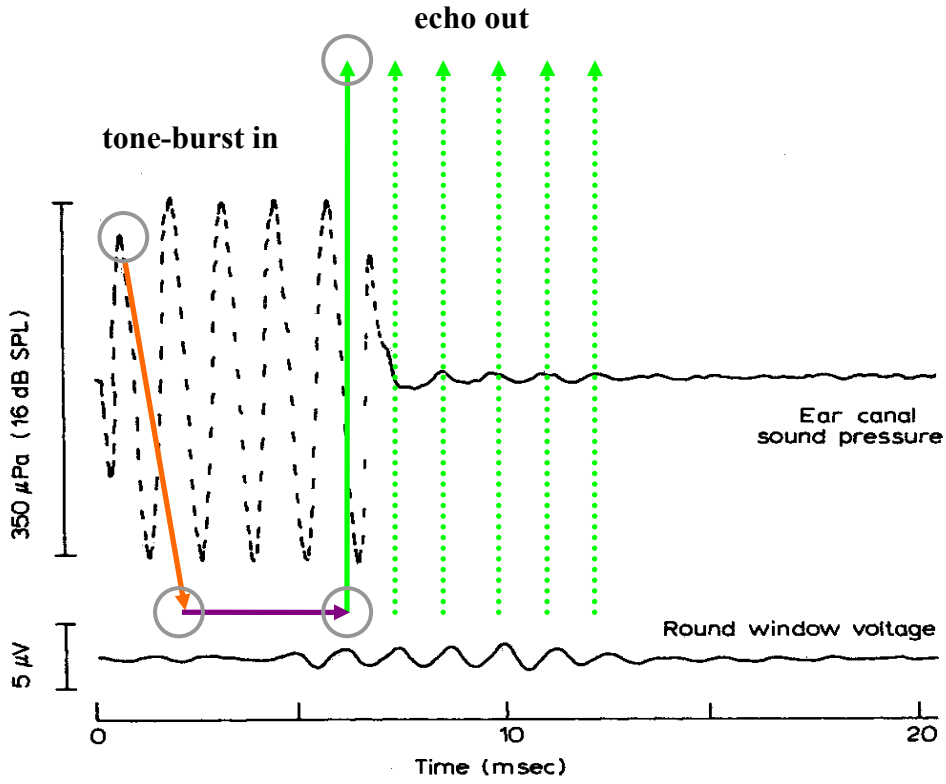

Fig. 3.9. In response to a tone-burst stimulus, the ear canal pressure of the echo (recorded with a microphone) occurred simultaneously with the voltage due to the cochlear microphonic (recorded with an electrode). This suggested to Wilson that a fast pressure pulse (nearly as fast as the electrical correlate) was conveying the echo from the 800-Hz position on the basilar membrane to the ear canal. Stimulus enters cochlea as traveling wave (orange), builds up oscillation in second filter (maroon), and exits cochlea as pressure wave (green). [Adapted from Fig. 2 of Wilson (1980), with permission of Elsevier Science]

In earlier chapters the suggestion was made that Wilson's model be modified so as to introduce a simple symmetry: not only do outer hair cells produce a fast pressure wave but in addition the fast pressure wave directly stimulates the outer hair cells, in this way bypassing the traveling wave (at least at low sound pressure levels). Consequently, all the delays we see are now due to filter delay (the orange line in Fig. 3.9 becomes vertical, and the maroon line doubles in length), but given what we know about the sharpness of cochlear tuning, the additional factor of 2 should be able to be accommodated with a suitably high Q filter. Wilson calculates (the basis of which is not given) that the volume change necessary to produce an ear canal

pressure of 20 dB SPL would require less than a 0.01% volume change in 300 outer hair cells (p. 530) and my own calculations suggest that this may be reasonable[109].

All the remaining material in this section is consistent with a two-way fast pressure wave and a slow build up in a highly tuned resonator. The filter could equally be considered a first filter as a second one. A reinterpretation of Shera and Zweig along these lines was given in §I 3.1/d.

2. *Fast distortion products.* It has been known for a long time that cochlear distortion products can occur remarkably quickly. In 1985, Brown and Kemp measured distortion products in gerbils, both acoustically in the ear canal and electrically via the cochlear microphonic[110]. They found several instances of delays at a small fraction of a millisecond (their Fig. 2), but in the main the delays were between 0.5 and 1.3 ms, slightly shorter than a two-way traveling wave. They were more puzzled that the delays in upper sideband distortion products ($2f_2 - f_1$) were shorter than for upper ($2f_1 - f_2$), and suggested that the acoustic distortion product may reach the base as 'fluid borne sound waves' (p. 197), in accordance with Wilson's proposal. Even so, the extremely short delays they measured were not highlighted.

A reason too much store has not been placed on distortion product measurements is that the mechanism is not well understood[111]. It is generally supposed that the interaction of the two primary frequencies takes place where the two traveling wave envelopes overlap on the partition. The observed result is an extended series of peaks and troughs, including much fine structure. The rapid phase and amplitude variations make data collection difficult and it has been easier to ignore discordant data.

---

[109] An OHC 50 µm long and 10 µm in diameter occupies a volume of 4×10^{-15} m^3. When 300 of them change volume by 0.01%, the volume change is 10^{-16} m^3. The cochlear fluid is incompressible, so this volume will displace the stapes and round window (about equally) and cause 15× this volume change at the ear drum. If the ear canal occupies a microphone-sealed volume of 1 cm^3, the ear canal volume will change by about 1 part in 10^8. Changing the ear canal pressure of 10^5 Pa by 1 in 10^8 gives a pressure of 10^{-3} Pa, and this is an SPL of 34 dB.
A change of 0.01% in volume could be effected by tilting the hinged cuticular plate by 0.1°. The cuticular plate can tilt by up to 15° [Zenner et al. (1988), p. 234].

[110] Brown, A. M. and D. T. Kemp (1985). Intermodulation distortion in the cochlea: could basal vibration be the major cause of round window CM distortion? *Hear. Res.* 19: 191-198.

[111] A proposal is made in Chapter D7.

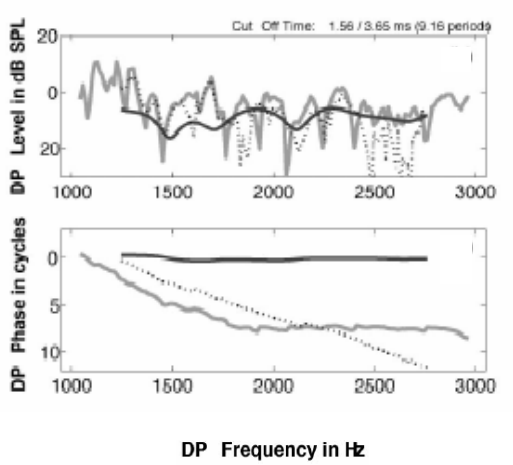

Fig. 3.10. Grey = black + dotted. How an actual DPOAE (grey line) can be separated into two components, a low-latency contribution (black line) and a high-latency one (dotted line). [Data is for a $2f_1 - f_2$ DPOAE with $f_2/f_1 = 1.2$ and $f_1 = 51$ dB and $f_2 = 30$ dB SPL. From Fig. 1 of Mauermann and Kollmeier (2004) with permission of the Acoustical Society of America]

The whole picture has recently been brought into sharp focus by a trio of papers[112,113,114] that demonstrate that a given DPOAE signal can be separated into two discrete components. Goodman et al. (2003) used an inverse FFT method on guinea pig data, Shera used a suppressor tone near the $2f_1–f_2$ frequency in a human, and Mauermann and Kollmeier (2004) employed a time-windowing procedure, again with human data. In such ways, the experimenters were able to separate DPOAEs into (using the most recent paper's terminology) a long-latency 'reflection' component (RCOAE, almost equivalent to an SFOAE) and a short-latency distortion component (DCOAE). The analyses reveal that a DPOAE is actually due to the interference of two components arising from two separate mechanisms, a conclusion that is not entirely new but which had not before been clearly demonstrated. As Fig. 3.10 illustrates, the separation provides a simpler picture of what may be going on.

[112] Goodman, S. S., et al. (2003). The origin of SFOAE microstructure in the guinea pig. *Hear. Res.* 183: 7-17.
[113] Shera, C. A. (2004). Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions. *Ear Hear.* 25: 86-97. See also Fig. 9 of Shera and Guinan (1999).
[114] Mauermann, M. and B. Kollmeier (2004). Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source. *J. Acoust. Soc. Am.* 116: 2199-2212.

In terms of amplitude variations, we see that instead of wide-ranging peaks and nulls the amplitude excursions are generally less. More dramatically, we see that the DPOAE phase curve is decomposed into two almost linear components: an RCOAE plot with a slope of about –1 cycle per 138 Hz, and a DCOAE plot, *almost horizontal*, with a slope of no more than about –1 cycle per 10 000 Hz. A major conclusion for the discussion here is that the horizontal component corresponds to a delay of no more than 0.1 ms and possibly much less, whereas the slow component corresponds to a delay of about 7.2 ms. The Shera paper shows a phase lag of less than 30° over the frequency range 2.6–7 kHz, corresponding to a time delay of about 34 μs, whereas the Goodman et al. paper show an average phase slope of –55° over 2–10 kHz, equivalent to a lag of some 19 μs. These are extremely small delays, suggestive of a fast mechanism – perhaps, I would venture, a compression wave. Shera has argued[115] against any naïve equivalence between DPOAE phase gradients and wave travel times, but there is reason to think[116] that the low phase gradient in fact represents the action of the pressure wave. The complete inference (drawn in more detail in Chapter R7) is that DPOAEs are due to the interaction of two tones at a single place on the partition where there are two highly tuned resonators; the tones enter and exit the cochlea via fast compression waves, and in between they interact through the slow build-up of highly tuned resonators.

Such a picture may be disputed, but it does at least suggest that further examination of the phase slope of DCOAEs is warranted, as it certainly looks like a fast wave underlies them[117]. A recent paper[118] confirms such an interpretation: it observed SFOAE latencies and found that "many of the latencies were too short to be considered valid" and were "apparently inconsistent with the reflection source hypothesis" (p. 3811). One of the explanations considered for the recurring "invalid" data is the reverse transmission path through the fluid, although the authors underline the need to reconcile this with long-latency observations, which are consistent with the conventional round-trip concept. The model in the preceding paragraph does this.

---

[115] Footnote 27 of Shera and Guinan (1999).
[116] Fuller, R. B. (1975). *Synergetics: Explorations in the Geometry of Thinking*. (Macmillan: New York). [p. xix]
[117] In all the papers showing a near-horizontal phase plot, none has sought to provide a physical interpretation.
[118] Konrad-Martin and Keefe Transient-evoked stimulus-frequency and distortion-product otoacoustic emissions in normal and impaired ears.

3. *Fast suppression*. The amplitude and frequency of a spontaneous emission can be disturbed by projecting a suppressing tone into the cochlea, and recording the time course of the process gives useful information about the oscillator dynamics. In one such study, the experimenters found that just prior to release from suppression, a prominent short-lived dip appeared[119]. The time-constant of the dip was found (p. 3718) to be 0.03 ms, a lag which strains an explanation based on interactions of traveling waves but matches the figures derived from the DPOAE phase measurements.

3.2/l A bootstrap problem

An interesting perspective on the traveling wave theory is given[120] by Fukazawa (2002). If the fundamental stimulus to the outer hair cells is differential pressure – that is, pressure across the partition – how can the outer hair cells, embedded *in* the membrane, cause any change in that differential pressure? That is, the outer hair cells have nothing to push against: any force generated by them will pull down on the plateau of Corti at the same time as they push against the basilar membrane and the forces will cancel. It is the classic bootstrap problem in which internal forces can never change the momentum of a system. Fukazawa concludes that the cochlear amplifier can never get off the ground.

3.2/m No backward traveling wave

Kemp introduced the idea of a backward traveling wave in order to explain cochlear echoes, and the concept is now a standard part of the modern traveling wave model (see §I 3.1/d), even though a backward traveling wave has never been directly observed. Wilson queried the concept[121] in 1988, but it has persisted because without it active cochlear mechanics is left high and dry. A provocative discussion on the

[119] Fig. 2c (p. 3714) of Murphy, W. J., et al. (1995). Relaxation dynamics of spontaneous otoacoustic emissions perturbed by external tones. II. Suppression of interacting emissions. *J. Acoust. Soc. Am.* 97: 3711-3720.
[120] Fukazawa, T. (2002). How can the cochlear amplifier be realized by the outer hair cells which have nothing to push against? *Hear. Res.* 172: 53–61.
[121] p. 114 of *Basic Issues in Hearing*, ed. Duifhuis, Horst, Wit (Academic, London, 1988).

issue can be found on pp. 583–586 of *Biophysics of the Cochlea*[122] in which Dallos concurs that "there is absolutely no experimental evidence that shows there is a backward travelling wave."[123]

It is impossible to prove a negative, since lack of evidence is not evidence of lack. Perhaps the traveling wave is at such a low level it is beyond observation. Nevertheless, the longer it remains unobserved, the more doubts grow. The latest observations to draw a blank[124] were done on the gerbil cochlea by Ren (2004) using a scanning laser interferometer to detect vibration of the stapes and of the partition. He projected two tones into the cochlea and detected the DPOAE at $2f_1–f_2$. Significantly, the interaction generated a forward traveling wave – he could see the stapes vibrate at the $2f_1–f_2$ frequency 50 μs before the basilar membrane at its best frequency did – but he could not see any motion on the basilar membrane before that stapes vibration. So what caused the stapes to vibrate? There didn't seem to be any backward traveling wave, so Ren proposes a compression wave, even though he acknowledges it contradicts current theory. In supplementary material on the *Nature* web site, Ren discusses similar work by Narayan et al. (1998) and concludes that "This unambiguous finding in different species of experimental animals by two independent laboratories clearly demonstrates that the stapes vibration at the emission frequency and the consequent resulting otoacoustic emission in the ear canal are not mediated by the hypothetical backward travelling wave."

The problem of anomalous round trip travel times was raised in §I 3.1/d, and this also strengthens doubts about the existence of backward traveling waves.

In brief, there are certain difficulties underlying traveling wave theory, and these are magnified when trying to sustain the case for a backward-traveling version.

---

[122] Ed. A. W. Gummer (World Scientific, Singapore, 2003).

[123] Gummer (2003) p. 584. Long candidly admits that "I think it goes back to most published experiments – they only show the beautiful and simple results. Every time I have done transient-evoked emissions and even in some of our distortion-product emissions, although the major returning wave has a two-times travel-time, there are indications of a one-time travel-time, and [even though] Pat Wilson talked about it [and] other people have talked about it, it is not talked about much, we can't explain it, so we don't tend to stress it, but I am sure that it is there." [ibid., p. 585]

[124] Ren, T. (2004). Reverse propagation of sound in the gerbil cochlea. *Nat. Neurosci.* 7: 333-334.

3.3 Summary

This wide-ranging discussion has highlighted the major limitations of the traveling wave theory and has shown how these may have to do with neglecting the round window membrane and its generation of common-mode pressure. A later part of this thesis (Chapter D8) sets out evidence that outer hair cells do in fact appear to act as pressure detectors, and a mechanism is described whereby this could happen. However, at this point the hypothesis that outer hair cells detect acoustic pressure is taken as a viable theory, and a model of how this process – likened to a surface acoustic wave (SAW) resonator – could operate in the cochlea is presented in the following chapter. We will return to considerations of pressure detection in the cochlea in an assessment of supporting evidence for the model in Chapter D9.