Transactions of Mathematics and its Applications (2017) **1**, 1–37
doi: 10.1093/imatrm/tnx002

Mode conversion in the cochlea?

ROBERT S. MACKAY*
Mathematics Institute and Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, UK
*Corresponding author: R.S.MacKay@warwick.ac.uk

[Received on 23 August 2016, revised on 24 December 2016]

The cochlea is the part of the ear where sound is analysed into frequencies. It is partitioned into two channels by a flexible membrane, whose effective density varies little along the length but whose stiffness reduces by a factor 10^4 from base to apex. Via the stapes near the base of the cochlea, the incoming sound excites travelling waves on the cochlear partition. The standard theory is that they travel to a frequency-dependent place where the cochlear partition is resonant with the sound frequency and thus is excited to relatively large amplitude there. The response to high frequencies peaks a short distance from the base, that for lower frequencies peaks towards the apex of the cochlea. Thus the cochlea separates frequencies by location of the peak response along the cochlear partition, from where they are transduced into neural signals to the brain. Mathematically, such a scenario falls into the class of critical-layer resonance models, well known in plasma physics, fluid dynamics and atmospheric science.

It is suggested here, however, that the frequency selectivity of the ear may be based on mode conversion rather than critical-layer resonance. Mode conversion is a phenomenon discovered in plasma physics in which two modes of wave in an inhomogeneous medium completely interconvert. Waves entering in one mode travel to a frequency-dependent place where they turn round and travel back in the other mode. The wave amplitude grows to a peak at the turning point. Thus mode conversion is an equally good mechanism for frequency analysis.

By comparing the theory with some of the wealth of experimental data, several inconsistencies of critical-layer resonance with the cochlea are pointed out, and it is found that they can be removed by mode conversion. Furthermore, critical-layer models are not robust to addition of higher derivatives: they generically deform to mode conversion or a case with no singularity.

A proposal is made for a physiological mechanism for mode conversion to arise in the cochlea, based on forces produced by outer hair cells (which so far have primarily been invoked only to cancel damping). Other sources for mode conversion are possible too, however, notably two-membrane models. The analysis is carried out at the linearised level, though extension to incorporate nonlinear effects is an essential next step.

The key result of this article is to highlight mode conversion as a highly plausible mechanism for the frequency selectivity of the cochlea, much more likely than critical-layer resonance, and to stimulate further investigation of this option.

Keywords: cochlea; travelling wave; mode conversion.

1991 Mathematics Subject Classification: 43.64.Kc, 43.66.Ba

1. Introduction

The cochlea is the part of the ear where mechanical vibrations (forced by sound waves in the air) are ‘Fourier analysed’ into neural signals. There is a huge literature about it, both experimental and mathematical (for summaries, see Keener & Sneyd, 1998, Chapter 23; Robles & Ruggero, 2001; Frosch, 2010; Duifhuis, 2012, Chapter 3; or Reichenbach & Hudspeth, 2014). There are large differences between

the cochleae of different species (especially between mammals, birds and reptiles); in this article attention is restricted to mammalian and usually human cochleae.

Most current explanations for the function of the mammalian cochlea are based on 'critical-layer resonance', the build up of wave energy at a frequency-dependent place where the wavelength goes to zero, e.g. Lighthill (1981), de Boer (1996), Hubbard & Mountain (1996) and Nobili *et al.* (1998), even if the words and mathematical theory are not always used and it is disputed by some (e.g., the discussion in Section VII of Olson, 2001).

The goal of this article is to suggest that critical-layer resonance is not the right explanation: rather it might be 'mode conversion', the propagation of a wave to a frequency-dependent place where it turns round and comes back out in another wave mode, the energy density forming a peak at the turning point. Critical-layer resonance models fail to fit several experimental observations, whereas mode conversion produces consistent results.

A notable alternative to both critical-layer and mode-conversion models is given by de Boer & Nuttall (2000), to be commented on at various points.

Despite clear experimental evidence for nonlinear effects (e.g., Rhode & Robles, 1974; and summaries in Section 8 of de Boer, 1996; Robles & Ruggero, 2001), and the large amplitude that occurs near a mode-conversion point, this article is limited to treatment of the linear regime, as the first step in the analysis.

Critical-layer resonance modelling is reviewed in Section 2, with particular reference to the cochlea, mainly in the frequency domain but an appendix is added on the time domain. In Section 3, a range of inadequacies of critical-layer resonance models for the cochlea is listed. Section 4 surveys the phenomenon of mode conversion and Section 5 suggests that it could occur in the cochlea. Section 6 surveys some precedents. Section 7 is an interlude on two-mode models. Section 8 gives some ideas about how mode conversion could arise physiologically. Section 9 discusses stability constraints. A second appendix discusses the effects of longitudinal coupling in the cochlear partition.

2. Critical-layer resonance

First the phenomenon of critical-layer resonance is recalled, together with how it arises in a simple model of the cochlea. Quite a long treatment is provided, starting from elementary considerations, because it is important to appreciate how mode conversion differs from it, but readers who are already familiar with critical-layer resonance and its proposed occurrence in the cochlea can skip the section.

2.1. *Frequency domain*

The context for critical-layer resonance is waves in a medium whose properties vary smoothly in space and the local wavelength-squared goes through zero somewhere. If the medium can be treated as linear (i.e., any linear combination of two solutions is another solution) and time independent, one can analyse its behaviour by considering solutions with a single frequency ω, i.e., time-dependence like (the real part of) $e^{\mathrm{i}\omega t}$ (ω is the 'angular frequency'; some may prefer to reserve the word 'frequency' for $f = \omega/2\pi$). Then one can attempt to write solutions as superpositions of waves $e^{\mathrm{i}(\omega t - k.x)}$, or better, WKB[1] solutions $A(x,\omega)\, e^{\mathrm{i}(\omega t - \int^x k(x',\omega)\,\mathrm{d}x')}$ with A determined by the evolution of wave energy density along the rays, where the wave vectors $k(x,\omega)$ are solutions of the local dispersion relation for the medium. For introductions

[1] After Wentzell, Kramers and Brillouin; also attributed to Liouville and Green.

to linear wave theory in fluid mechanical contexts, see Whitham (1974) or Lighthill (1978). A *critical layer* is a place x (or line or surface, depending on the ambient dimension) where the local wave number $k(x, \omega)$ goes to infinity for some mode at the given frequency ω.

The phenomenon of critical-layer resonance is that waves of a given frequency propagating towards the critical layer slow down and take infinite time to reach it. Their energy density increases in inverse proportion to their group speed until damping effects take over. Virtually all the energy is absorbed in a zone near the critical layer and almost nothing is reflected. Waves can propagate on only one side of the critical layer, they evanesce on the other side.

Fluid mechanics attribute the discovery and analysis of the critical-layer phenomenon to Booker & Bretherton (1967), but it had already been done by Budden (1955) in plasma physics, and it had already been found numerically in a model of the cochlea by Peterson & Bogert (1950)!

The best way to illustrate the critical-layer phenomenon in the cochlear context is provided by the passive one-dimensional cochlear model of Peterson & Bogert (1950), so it is now recalled. Consider the cochlea to consist of a rigid tube separated into two by a flexible partition along its length, called the 'basilar membrane',[2] and the fluid motion to be incompressible,[3] and for the moment only longitudinal.[4] So the volume fluxes in the two tubes are equal and opposite, denoted by $\pm j(x, t)$ at longitudinal position x and time t, and the pressures are uniform over each part of the cross-section. Denote the area displaced by the membrane from its equilibrium position in the two parts of the cross-section by $\pm a(x, t)$ and the pressure difference across the membrane by $p(x, t)$ (Lighthill, 1981, denotes it $2p$). If for each x the membrane deforms in a mode shape $z = \zeta(y)$ with respect to lateral position y (the shape ζ can depend on x), then a is related to the displacement z_0 at a chosen y_0 (for example, on the centre line) by $a = \frac{z_0}{\zeta(y_0)} \int \zeta \, \mathrm{d}y$, so some authors convert a to z_0, to correspond more closely with what is measured (which is usually the velocity $\dot{z}_0$).

Then fluid mass conservation leads to

$$\frac{\partial a}{\partial t} = \frac{\partial j}{\partial x}, \tag{2.1}$$

and horizontal momentum conservation (ignoring viscous effects) to

$$\sigma \frac{\partial j}{\partial t} = -\frac{\partial p}{\partial x}, \tag{2.2}$$

where

$$\sigma(x) = \frac{\rho_1}{A_1} + \frac{\rho_2}{A_2}, \tag{2.3}$$

with ρ_i the fluid density in the two tubes and $A_i(x)$ their equilibrium cross-sectional areas. It would be reasonable to take $\rho_1 = \rho_2$, but the analysis is no different when $\rho_1 \neq \rho_2$.[5] Combining equations (2.1)

[2] Though it should include the organ of Corti and the tectorial membrane too, so would be better called the 'cochlear partition'.

[3] This is a good approximation as long as the time for acoustic waves in the fluid to equilibrate the pressure to that corresponding to the boundary conditions at the ends is small compared with the timescale on which the boundaries move, though there are authors who claim this fails above 7 kHz, referred to in Lighthill (1981).

[4] This is inaccurate, cf. Fig. 7 of Olson (2001), but the effects of two-dimensional and three-dimensional fluid flow are recalled shortly and make little qualitative difference.

[5] Actually the upper channel consists of two fluids of different ionic composition separated by a very flexible membrane, called Reissner's membrane.

and (2.2),

$$\frac{\partial^2 a}{\partial t^2} = -\frac{\partial}{\partial x}\frac{1}{\sigma}\frac{\partial p}{\partial x}. \tag{2.4}$$

The vertical momentum equation is

$$m\frac{\partial^2 a}{\partial t^2} = -p - \lambda a, \tag{2.5}$$

where $\lambda(x)$ is the stiffness of the membrane (meaning the pressure difference to produce unit area displacement in the plane at constant x) and $m(x)$ an effective 'density' of the membrane (though it is usually called mass). If the membrane deforms in mode ζ as above and has mass $\mu(x,y)$ per unit area in (x,y), then (Lighthill, 1981)

$$m(x) = \frac{\int \mu\zeta^2 \,\mathrm{d}y}{(\int \zeta \,\mathrm{d}y)^2}. \tag{2.6}$$

A more realistic evaluation of m requires incorporation of how the organ of Corti moves, which in the above treatment is regarded as moving rigidly with the basilar membrane.

In live cochlea, it is well established that there are additional forces generated by the outer hair cells (OHCs), e.g., as surveyed by Robles & Ruggero (2001). Effects of active feedback can be added to (2.5), as could any other contributions to the transverse force, but for simplicity of illustration of the phenomenon of critical layer resonance we continue without them.

Combining (2.4) and (2.5), and looking for solutions with time dependence $e^{\mathrm{i}\omega t}$, yields

$$\left(\frac{\lambda}{\omega^2} - m\right)\frac{\partial}{\partial x}\frac{1}{\sigma}\frac{\partial p}{\partial x} + p = 0. \tag{2.7}$$

Thus if the parameters σ, λ, m are treated as locally independent of x one obtains $p \propto e^{\pm \mathrm{i}kx}$ with dispersion relation

$$(\lambda - m\omega^2)k^2 - \sigma\omega^2 = 0. \tag{2.8}$$

The term $\sigma\omega^2$ proportional to ω^2 is sometimes described by saying the response of the longitudinal fluid flow to transverse membrane motion acts like an 'added mass' for the membrane, with density σ/k^2.

Lighthill (1981) surveyed ways to calculate two-dimensional and three-dimensional corrections to this added mass, which involve modifying (2.4) to give a as the convolution of p with a nonlocal version of the second derivative of a delta-function. So one no longer obtains an ordinary differential equation, but one can still perform WKB analysis, inserting

$$p = -\omega^2 I(k)a$$

into (2.5) with a function

$$I(k) \sim \begin{cases} \frac{\sigma}{k^2} \text{ for } kh \ll 1 \\ \frac{2\rho}{w|k|} \text{ for } kh \gg 1, \end{cases} \tag{2.9}$$

where w is an effective width of the membrane and h an effective height of the two channels. For example, in the case of a two-dimensional model with rectangular cross-sections of areas $A_1 = A_2 = wh$ then $I(k) = \frac{2\rho}{wk \tanh kh}$. The regime $kh \ll 1$ corresponds to using the one-dimensional fluid approximation.

Sticking to the one-dimensional model for ease of exposition, (2.8) yields

$$k^2 = \frac{\sigma\omega^2}{\lambda - m\omega^2}. \tag{2.10}$$

It is usually stated that σ and m do not vary much along the length of the cochlea. It seems difficult to estimate m; if μ were constant, the simplest guess from (2.6) would be that m is inversely proportional to the width w of the membrane, but μ is not constant. The cross-sectional area of the organ of Corti is often taken as a surrogate for the mass, and it varies in mature gerbils by only a factor of two (and non-monotonically) along the cochlear partition (Richter *et al.*, 2000). On the other hand, Fig. 1 of Mammano & Nobili (1993) shows a significant variation of A_1 and A_2, hence of σ. Nevertheless, the stiffness λ decreases by a factor of about 10^4 from base to apex (partly because the width of the membrane increases by a factor of about 4), so k^2 increases. Taking the direction of increasing x from the base to the apex, k^2 goes to $+\infty$ as x approaches a resonant location where

$$\lambda(x) = m(x)\omega^2, \tag{2.11}$$

and is thereafter negative (i.e., k is imaginary, so solutions grow or decay exponentially with respect to x locally). Thus there is a critical layer at the solution $x(\omega)$ of (2.11). It moves from base to apex as the frequency ω decreases from an upper limit to a lower limit.

It is useful to consider the *affine impedance model* ('affine' means constant plus linear) obtained from (2.7) by making the straight-line approximation

$$\frac{1}{\sigma(x)}\left(\frac{\lambda(x)}{\omega^2} - m(x)\right) \approx \beta(x(\omega) - x) \tag{2.12}$$

for x near $x(\omega)$, where $\beta = (m'(x) - \lambda'(x)/\omega^2)/\sigma(x)$ evaluated at $x(\omega)$, and supposing σ locally constant. In the scaled variable

$$X = \frac{x - x(\omega)}{\beta}, \tag{2.13}$$

this yields

$$Xp'' = p. \tag{2.14}$$

The solutions can be written in terms of Bessel functions of order 1 (de Boer & MacKay, 1980). Using (2.4) to convert p to a gives

$$a = \frac{\sigma p}{\omega^2\beta^2 X}, \tag{2.15}$$

and one obtains that the solutions on each side of the critical layer look like linear combinations of those in Fig. 1 or one which grows faster than exponentially as $X \to +\infty$ (not shown). In particular $a(X) \to \infty$ like $|X|^{-1}$ for most solutions as x approaches $x(\omega)$ from either side.

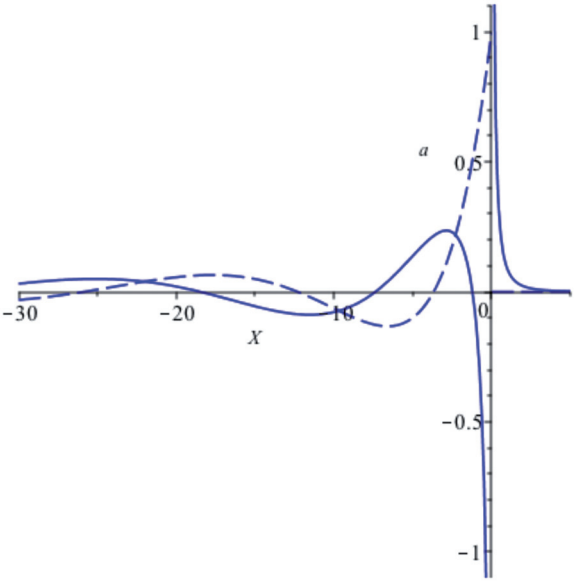

FIG. 1. Graph of $a(x)$ at two phases of oscillation for the affine-impedance approximation (2.14) to (2.7), plotted against $X = \frac{x-x(\omega)}{\beta}$. The solid curve is at phase $-\pi/2$. The dashed curve is at phase 0 and is zero on the right.

The differential equation (2.7) does not tell one how to connect the solutions on the two sides of the critical layer, because it is singular at $x = x(\omega)$ (the coefficient of the highest derivative goes through 0 there). Nevertheless, adding a little damping (e.g., a dissipative term $-\delta\frac{\partial a}{\partial t}$, $\delta > 0$) to (2.5) allows one to match the solutions on the two sides: for the affine approximation the answer is that the solution with no spatially growing mode as $X \to +\infty$ (a physically reasonable requirement) has only in-going waves on the other side (Budden, 1955; Stix, 1962; Booker & Bretherton, 1967): at one phase of the oscillation the solution for a is the solid curve of Fig. 1, and $\frac{\pi}{2}$ later it is the dashed curve (zero on the right). Up to scale, the exact solution of (2.14),[6] converted by (2.15), is

$$a(X) = \begin{cases} \frac{1}{\sqrt{-X}} H_1^{(1)}(2\sqrt{-X}) \text{ for } X < 0 \\ \frac{2\mathrm{i}}{\pi\sqrt{X}} K_1(2\sqrt{X}) \text{ for } X > 0, \end{cases} \tag{2.16}$$

where $H_1^{(1)} = J_1 + \mathrm{i}Y_1$ is a Hankel function.

The amplitude of oscillation $|a|$, denoted by $A(X)$, is drawn in Fig. 2(a). It goes to infinity at $X = 0$. Damping would make the maximum of A finite, albeit still large (inversely proportional to the damping strength) (de Boer & MacKay, 1980).

Similarly, one can plot the phase ϕ of the solution (Fig. 2(b)) (relative to a reference value chosen to be $-\pi$ for $X > 0$). Note the jump of $-\frac{\pi}{2}$ on crossing the critical layer, which is determined by considering the effect of small damping. One can compare the exact result with that of the WKB approximation.

[6] A factor of i was lost in de Boer & MacKay (1980).

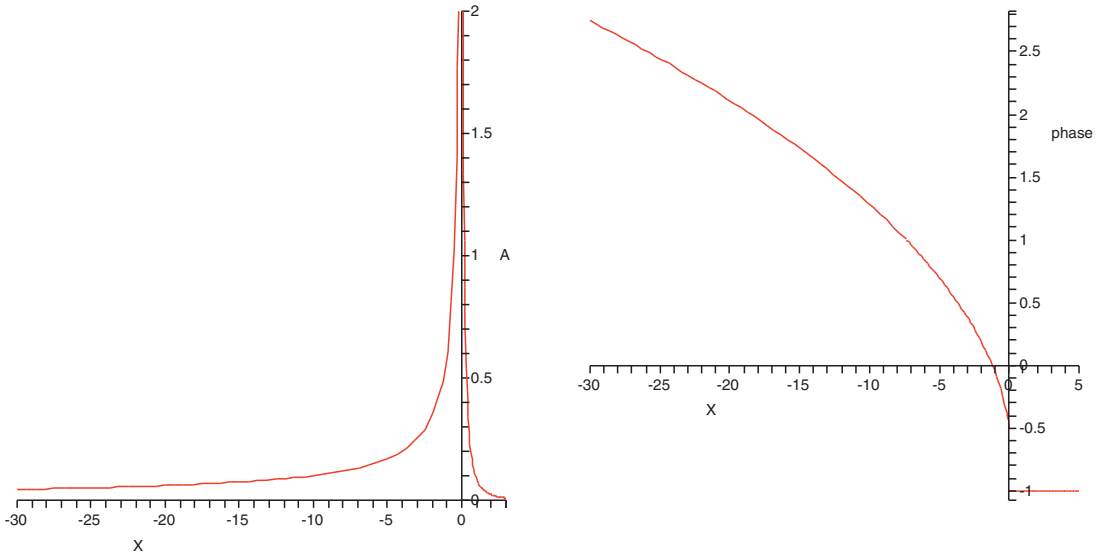

FIG. 2. (a) Amplitude $A(X)$ and (b) scaled phase $\phi(X)/\pi$ for the affine-impedance critical-layer model (2.14).

In particular, the phase of the exact solution differs from the WKB approximation $\phi = -\int k \, \mathrm{d}X = 2\sqrt{-X} - \frac{\pi}{4}$ (choosing the constant to fit the asymptotics of the Hankel function) by losing an extra $\pi/4$ in the non-WKB regime, followed by this jump of $-\pi/2$.

Somewhere to the left of the critical layer one has to match the solution of the affine-impedance model to the full model. This can be done with ease if the affine approximation is good over a range $x(\omega) - C\beta < x \leq x(\omega)$ for some C significantly larger than 1, because then it can be matched to a WKB solution coming in from $x < x(\omega) - C\beta$. The condition for validity of the WKB approximation is $k' \ll k^2$. It shall be assumed this holds up to a point where the affine impedance approximation (2.14) takes over (for the affine impedance model, $k'/k^2 = \frac{1}{2}|X|^{-1/2}$, so WKB is good for $|X| \gg \frac{1}{4}$).

When two-dimensional and three-dimensional effects for the motion of the fluid are taken into account, the local dispersion relation changes from (2.10) to

$$I(k) = \lambda/\omega^2 - m.$$

Then the criterion for the WKB approximation is just $\beta \ll \bar{A}/w$, where $\bar{A}$ is the harmonic mean of A_1 and A_2, which holds nearly everywhere in the cochlea (Lighthill, 1981) (if the membrane is assumed to fill the width of the cochlea then $\bar{A}/w \approx h$, so this shorthand will sometimes be used). It follows that the amplitude $|a|$ grows like $|X|^{-1}$ again near the critical layer, but the phase goes to $-\infty$ in proportion to $\log|X|$ as X increases through negative values to 0.

2.2. *Plausibility of critical-layer resonance in the cochlea*

Critical-layer resonance sounds a perfect explanation for the cochlea. Indeed, Lighthill (1981) argued that any model of the cochlea must be based on critical-layer resonance. Frequency ω is mapped to place $x(\omega)$, so a given inner hair cell (which transduces movement into neural signals to the brain) is stimulated mainly by Fourier components around only one frequency.

One can work out the frequency selectivity, assuming some reference for forcing amplitude. Sound pressure level in the ear canal is a common choice in experiments. To translate this into models one could assume conservation of wave energy through the ossicles and ear drum. The sound energy flux in the ear canal is $\frac{A}{\rho c}|p|^2$, where $|p|$ is the amplitude of pressure fluctuation, A the cross-sectional area of the ear canal, ρ the density of air and c the speed of sound in air. So constant sound pressure level corresponds to constant energy flux, independently of frequency. In the one-dimensional model, the energy flux Φ for a solution is the average $\langle \Re p\, \Re j \rangle$ over a cycle, where $\Re$ denotes real part. For time-dependence proportional to $e^{i\omega t}$, equation (2.2) implies $j = \frac{ip'}{\omega\sigma}$, so

$$\Phi = \frac{1}{\omega\sigma}\Im(p\bar{p}'), \tag{2.17}$$

where $\Im$ denotes imaginary part and over-bar denotes complex conjugate. For a right-going wave in the WKB regime like $e^{-i\int k\,dx}$ then $\bar{p}' = ik\bar{p}$, so $\Phi = \frac{k}{\sigma\omega}|p|^2$. Equation (2.4) implies $p = -\sigma(\frac{\omega}{k})^2 a$, so we can write Φ in terms of $|a|$ as

$$\Phi = \sigma\left(\frac{\omega}{k}\right)^3 |a|^2. \tag{2.18}$$

Thus the amplitude of oscillation as a function of incoming frequency ω for fixed incoming energy flux Φ in the WKB regime is (for the one-dimensional model)

$$|a| = \sqrt{\frac{\Phi}{\sigma}}\left(\frac{k}{\omega}\right)^{3/2} = \frac{\sqrt{\Phi}\sigma^{1/4}}{(\lambda - m\omega^2)^{3/4}}.$$

Matching this to the energy flux of the solution of (2.14) gives

$$|a| = \frac{\pi}{2}\sqrt{\frac{\Phi}{\sigma}}(\omega\beta)^{-3/2}A(X)$$

in the regime where (2.14) holds. Consequently, the frequency response at a given place looks like Fig. 3(a), computed for the *exponential model*:

$$\lambda(x) = C\,e^{-\alpha x},$$

σ and m constant (de Boer & MacKay, 1980).

The exponential model has a scaling symmetry so it suffices to study a single place for all frequencies or a single frequency for all places, and there is just one free parameter $\frac{\alpha^2 m}{\sigma}$, which is about $\frac{1}{40}$ in the cochlea (de Boer & MacKay, 1980); actually, the result in the WKB regime is independent of its value, so Fig. 3 is plotted for $\frac{\alpha^2 m}{\sigma} = \frac{1}{4}$, else the decay above the resonant frequency becomes too steep to see. Also the extension of the WKB curve to the left half of the affine region agreed to within plotting resolution with that for the affine model, so just the WKB curve and the right-hand half of the affine region are plotted.

A more common protocol in experiments is to measure the forcing amplitude required to produce a given amplitude of response at a given point, which also has the advantage for comparison with linear theory that the system is likely to remain in the linear regime. Also the measure of response amplitude

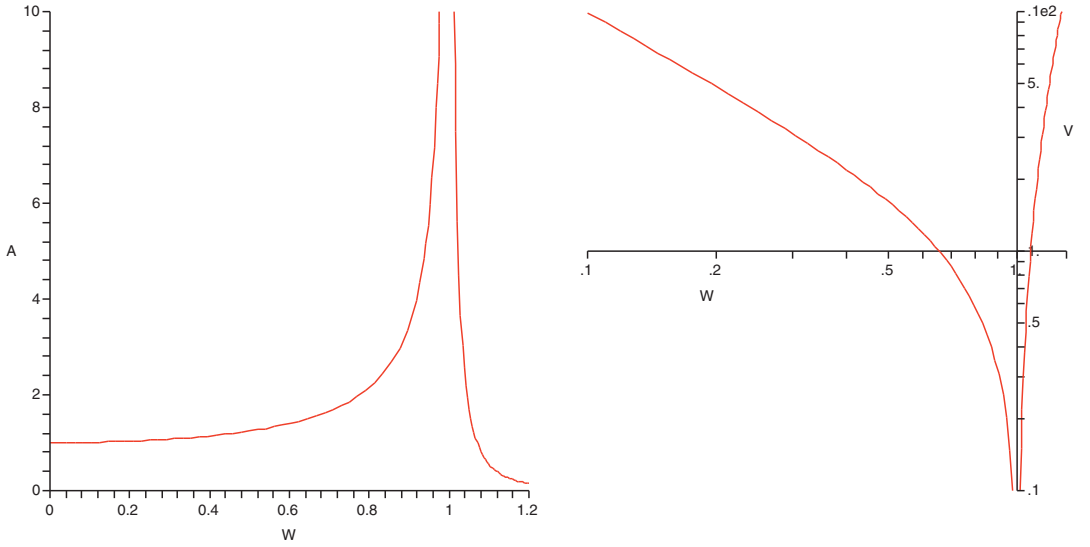

FIG. 3. Response of exponential model with $\frac{\alpha^2 m}{\sigma} = 1/4$: (a) Amplitude $A = |a|$ of response at a given place x to forcing of a given energy flux, as a function of scaled frequency $W = \frac{\omega}{\omega_r(x)}$, with $\omega_r(x)$ being the resonant frequency at x; (b) a tuning curve (log-log graph of basilar membrane velocity amplitude $V = |\dot{a}|$ against scaled frequency W).

is usually basilar-membrane velocity $\dot{z}$, which on linear theory is proportional to $\dot{a}$ at a given place. It is then plotted on a log–log scale and called a 'tuning curve' (though sometimes a linear scale is used for the frequency). A tuning curve for the exponential model is shown in Fig. 3(b). If the effects of damping are incorporated, which rounds off the spike, the resulting tuning curves look roughly like observed ones, e.g., Fig. 6 of Robles & Ruggero, 2001. One can also compute the phase at a given position relative to that at the stapes, as a function of input frequency, and incorporate modifications for the effects of two-dimensional and three-dimensional fluid flow.

Good fits to the observed frequency response of the cochlea have been obtained by suitable dependence of parameters on x, including damping (both positive and negative) (de Boer & Nuttall, 2000), though interestingly their best fits have no critical layer (it would be beyond the apex for all relevant frequencies). More comments on their approach will be given in Sections 3 and 8.

The discussion so far has been in the frequency domain, but it can also be useful to study the cochlea in the time domain, in particular for the response to impulses and to include nonlinear effects, so a summary of the time-domain approach to critical-layer resonance models of the cochlea is presented in Appendix A.

3. Inadequacies of critical-layer models

Critical-layer models of the cochlea, however, have several inadequacies.

Firstly, even after modification for two-dimensional and three-dimensional effects, adding damping processes like $-\delta \frac{\partial a}{\partial t}$ to (2.5) and $-\gamma j$ to (2.2), and many models for active processes (OHC feedback), de Boer (1996) claims that the peak in the response to periodic forcing does not come out the right shape if based on critical-layer resonance: 'either the amplitude of the peak remains too low or the phase variations in the peak region are too fast'. On the other hand, subsequent finite-element computer modelling produces

results that are claimed to be close (Kolston, 1999). Correct peak shape can be obtained by parameter fitting in a class of abstract models allowing a general basilar-membrane impedance-function (de Boer & Nuttall, 2000); but realistic electro-mechanical assumptions would place strong constraints on the joint frequency- and position-dependence and it is not clear that they are satisfied by the basilar-membrane impedance functions that are fit from data.

Returning to critical-layer models, let us discuss particularly the phase variation. On a critical-layer model, the wavelength should go to zero at the critical layer. Yet the wavelength (inferred as $2\pi/\phi_x$, where ϕ_x is the rate of change of phase with position) of the travelling wave along the basilar membrane is observed not to go to 0; indeed it seems to go to a minimum of about 0.5–1 mm at the resonance (Table 4 of Robles & Ruggero, 2001), with the shorter minimum for the higher frequencies. Now the one-dimensional models escape this criticism because even though the slope of the phase does go to infinity, it does so in such a way that the phase change up to the critical layer is finite (Fig. 2), so any smoothing effects (damping, nonlinearity, other forces, experimental error) will saturate the slope. Note also that if one considers the wavelength to be twice the distance between zeroes then the one-dimensional case escapes even without adding extra effects. This is because the WKB approximation is not valid for $|x(\omega) - x| \leq \beta$ so the approximation of slowly modulated waves is not correct there. It can be inferred from the Bessel functions behind Fig. 1 that there is a $c > 0$ (of order 1) such that there is at most one zero between $x(\omega) - c\beta$ and $x(\omega)$. An order of magnitude estimate for β, however, for a model with exponential stiffness-variation (de Boer & MacKay, 1980) is 0.075 mm (independent of frequency), which is much shorter than the observed minimum wavelength.

Models with two-dimensional or three-dimensional fluid flow do not escape this criticism, because they predict that the phase goes to infinity on approaching the critical layer, which disagrees strongly with observations. One could argue that the observations are masked by the effects of damping, which would reduce the amplitude to almost zero near the resonant location because the wave takes infinite time to reach it, but experimentalists can still measure the phase beyond the resonant location so the amplitude is not zero. Or one could argue that the observations are masked by an additional component after the resonant location ('cochlear fast wave'), due to the differences between the oval and round windows, so that the infinite phase-change is truncated to some large value. Indeed, the phase beyond the resonant location is always observed to adopt a specific relation modulo 2π with that of the stapes, so perhaps this is a valid escape. The observed sharpness of tuning at low amplitudes, however, suggests that damping is almost cancelled and then the phase should go to infinity.

Similarly, one can compare with experiments observing the response at fixed place to varying frequency. The accumulated phase (usually measured as the phase difference between response at a fixed place and forcing at the stapes, on increasing the frequency ω from 0), does not go to infinity as ω approaches the critical frequency for the given place; instead it plateaus at around 3–4 cycles, e.g., Fig. 7 of Robles & Ruggero (2001) and Fig. 5 of Shera & Guinan (1999). This is consistent with a one-dimensional critical-layer model but not a two-dimensional or three-dimensional one.

Secondly, a common objection to such critical-layer models (e.g., Naidu & Mountain, 1998) is that the observed variation in stiffness (by a factor of 10^4, e.g., Olson & Mountain, 1991) from one end of the membrane to the other is inadequate to account for the observed auditory range of frequencies (a ratio of about 10^3 from highest to lowest). The resonant frequency at a given place is $\sqrt{\frac{\lambda}{m}}$, which would change by a factor of only 10^2 from one end to the other (as already mentioned, the effective membrane density m is usually assumed not to vary much). Some authors try to get round this by proposing alternative resonances below the minimum on the basilar membrane, to extend the range, for example using the helicotrema (Lighthill, 1981), but the effects look weak. On the other hand, Emadi *et al.*, 2004 claim

that stiffness variation could just about account for the auditory range. Yet Reichenbach & Hudspeth (2014) write 'discrimination at low frequencies [less than 1 or 2 kHz] must ... function through a distinct mechanism that remains elusive'.

Thirdly, nearly all the models ignore longitudinal coupling in the membrane, which, although apparently very small (Voldrich, 1978; Emadi *et al.*, 2004), cannot logically be neglected in models which predict k^2 to go through infinity at some point and the basilar membrane displacement to suffer an infinite discontinuity there! The effects of addition of longitudinal coupling are analysed in Appendix B. It would add terms like $f_1k^4+f_2k^6$ to the dispersion relation (2.8) with $f_1, f_2 \geq 0$, so instead of going through infinity as x increases, $k^2(x)$ would rise steeply and then plateau. One would obtain an enhanced amplitude as the wave approaches the resonant location but the amplitude would remain high thereafter, unless damping was stronger for such short wavelength modes.

Critical-layer models in any domain of science suffer from a generalisation of this problem: the phenomenon of critical-layer resonance is not robust to addition of many other physical effects, however small (Stix, 1992). Generically, a critical layer is replaced by transition to a different wave mode, which is either (i) on the other side of the resonant location, or (ii) on the same side but with reversed group velocity. Indeed, as just remarked, the addition of longitudinal coupling yields case (i). Case (ii) is called 'mode conversion' and is the main subject of this article. To some extent the fragility of critical layer models is mitigated by damping, because addition of further physical effects weaker than damping does not change the results much. In the (live and oxygenated) cochlea, however, damping can be considered to be very small because of the active feedback provided by outer hair cells, which among other things is likely to nearly cancel (or exceed) damping. The removal of the critical layer by addition of other effects to cochlear models can be seen in work of de Boer (1990, 1993), even if not mentioned there.

Fourthly, the impulse response is often observed to have a double-lobed structure, e.g., Fig. 9 of Robles & Ruggero (2001), which is hard to explain with critical-layer models. On a critical-layer model, as discussed in Appendix A, an impulse at the stapes should produce a response at a given location starting with low frequencies and increasing to the resonant frequency for that location and then ringing at that frequency for a number of cycles depending on the damping. The double-lobed response suggests that a second packet of waves arrives later and interferes with the first. From where does it come?

Fifthly, critical-layer models have trouble explaining otoacoustic emissions (OAE). These are sounds which arise in the ear canal when vibrations are transmitted backwards from the cochlea (Kemp, 2002). There are various sorts (Probst *et al.*, 1991):

Spontaneous: occurs in about one third of ears, at one or more well defined frequencies, e.g., figures in McFadden & Plattsmier (1984), Martin *et al.* (1988) and Fig. 5.6 of de Boer (1991). A case with 23 has been reported (Probst *et al.*, 1991). Note that although sometimes correlated with tinnitus, a subjectively audible high frequency whistle, opinion seems to be that tinnitus is a neurological rather than mechanical phenomenon (Probst *et al.*, 1991).

Transiently evoked: response to a click or short tone-burst; it usually has a fairly well defined frequency and delay time (around 10 ms) before starting. The reflected energy can reach 100% of the incident energy at low stimulus amplitudes, e.g., Figs 5 and 8 of Wilson (1983). This was the first form of OAE to be discovered (Kemp, 1978).

Tone evoked: response at the same frequency; interference with the incident waves causes modulation in the frequency response of the ear canal with a period of around 100 Hz, which approaches nearly complete cancellation at low amplitudes, e.g., Fig. 1 of Shera & Guinan (1999).

Evoked distortion products: response at a combination frequency to input at two frequencies (as a clearly nonlinear effect, this will not be considered here, but see Ren, 2004; He *et al.*, 2008, for some challenges to cochlear models raised by experiments on these).

Spontaneous OAEs (SOAEs) and the nearly full energy-reflection of transiently and tone-evoked OAEs were the first evidence for an active component to the mechanics of the cochlea. Active processes are not a problem for critical-layer models, but the first three classes of OAEs are incompatible with critical-layer models in several ways. Critical-layer models can produce one eigenvalue (associated with the flow through the helicotrema, see Appendix A), but SOAEs in some subjects occur at more than one frequency. On linear theory, all those places with negative damping would produce SOAEs corresponding to their resonant frequency, and nonlinear effects could lock them to a discrete set of frequencies (or produce chaotic output), but the set of frequencies would be highly sensitive to physiological conditions, in contrast to the observed stability (Martin *et al.*, 1988). It is usual to try to explain transiently and tone-evoked OAEs in terms of errors in the WKB approximation, arising from non-smooth spatial variations of the medium, or nonlinear effects producing some equivalent. Shera & Zweig (1993) (also Zweig & Shera, 1995), and then Talmadge *et al.* (1998), carried out particularly thorough attempts to make this approach work. Such explanations encounter a problem, however: the delay time for transiently evoked OAEs is around 10–14 periods (de Boer, 1991, p. 172), which was considered there to be 'rather large' and 'hard to explain'.

Sixthly, if the active processes over-compensate for damping at the resonant location (as is proposed by de Boer & Nuttall (2000), in a zone just before the resonant location), then the connection formula for critical-layer resonance would change, so that an incoming wave produces a response growing rapidly towards the apex and the solution decaying towards the apex matches to an outgoing wave from the critical layer. One might say this is SOAEs, but then the cochlea should be unable to match in-going waves at unstable frequencies to bounded displacement at the apex. Although nonlinearity of the outer hair cell response can be expected to saturate any instabilities, using a model with such extreme sensitivity to the balance between damping and anti-damping feels risky. It is a good general principle that the behaviour of models should be robust to small modifications unless there are strong constraints on their form.

In contrast, mode conversion is robust to modelling errors, and it will be argued that it can work with the observed range of stiffness, can explain multiple SOAEs and tone-evoked OAEs, and has the potential to give the right tuning curves.

4. Mode conversion

The context for mode conversion is a wave-bearing medium which varies smoothly in space in such a way that the square k^2 of the local wavenumber is a multi-valued function of position x, with a fold point at a positive value of k^2 where two real values of k^2 merge. In Fig. 4, this is contrasted with the case of a critical layer. They share the mode with smaller k for large negative x, but instead of going to infinity at a certain x as for a critical layer ($x = 0$ in the figure), k turns around into a different mode at an earlier x.

It is crucial to distinguish mode conversion from the well-known case where k^2 decreases with non-zero slope through 0, which is usually called a 'cutoff' (because waves cannot propagate beyond it) or 'turning point' (because it gives a fold in the graph of $k(x)$ but at the value $k = 0$, or because it gives a reflection of the waves) and plays a fundamental role in semiclassical mechanics, radio propagation and many other domains of science. To avoid confusion, from now on the place where the graph of $k^2(x)$ folds will be referred to as a 'fold point', although both 'turning point' and 'cutoff' would a priori have been equally good descriptions.

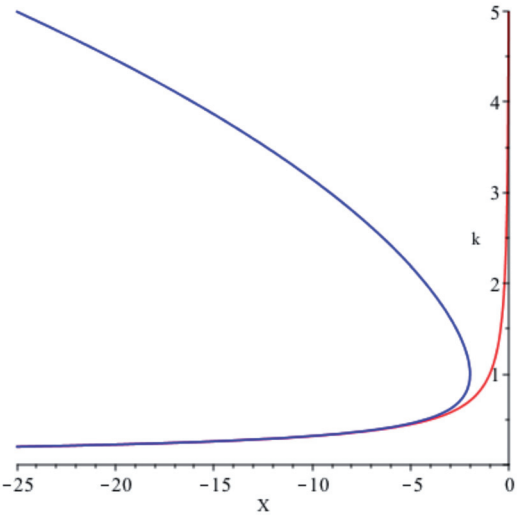

FIG. 4. Wavenumber $k > 0$ as a (two-valued) function of position x for given frequency for a system exhibiting mode conversion (blue), contrasted with critical-layer resonance (red) where k goes to infinity at $x = 0$ with k^{-2} going linearly through 0. For this illustration, $\omega^2 = 1 - x - k^{-2} - \eta k^2$ with $\omega = 1$, and $\eta = 1$ for mode conversion, 0 for critical layer.

The phenomenon of mode conversion is that a wave entering in one mode (one branch of the blue curve) slows down to zero group-velocity c_g at the fold point, because $\frac{\partial k}{\partial x}$ goes through infinity there and

$$c_g = \left.\frac{\partial \omega}{\partial k}\right|_x = -\left.\frac{\partial \omega}{\partial x}\right|_k \Big/ \left.\frac{\partial k}{\partial x}\right|_\omega , \tag{4.1}$$

and then turns into the other mode and propagates back out the way it came. Equation (4.1) shows that the group velocities on the two branches have opposite signs, despite having the same sign of phase velocity $c_p = \omega/k$. The possibility of group velocity opposite to phase velocity may be unfamiliar to some readers, but was known to Lamb (1904), plasma physicists in the 1960s, and is currently all the rage in active optical media (e.g., Nemirovsky *et al.*, 2012; Ye *et al.*, 2013). It invalidates the common statement in the cochlear literature (e.g., Ren, 2004; He & Ren, 2013) that the direction of travel of a wave can be inferred from the slope of the curve of phase versus longitudinal location.

If damping can be neglected, the intensity builds up in inverse proportion to the group speed (by conservation of energy), so produces a maximum at the fold. In contrast to a critical layer, however, the time taken to reach the fold is finite and the amplitude at the fold is finite.

This phenomenon of mode conversion was found and analysed by Stix (1965) in a plasma waves context.[7] In magnetised plasmas (ionised gases), it occurs near lower- and upper-hybrid resonances, the perpendicular ion-cyclotron resonance and the Buchsbaum two-ion resonance (Stix, 1992). It should also occur in many other contexts, e.g., ultrasound in elastic plates with thickness gradients, cf. Fig. 1 of Prada *et al.* (2005).

[7] Warning: Stix called the fold a 'critical layer', which again is reasonable terminology but was subsequently used by Booker & Bretherton (1967) for the case of $k^{-2}(x)$ passing through 0 and the latter usage has dominated.

Stix considered equations of the form (superscript-(4) denotes the fourth derivative):

$$\eta p^{(4)} - Xp'' + p = 0 \tag{4.2}$$

representing the leading terms for the response $p(X)$ of an inhomogeneous lossless medium at a given frequency near a point ($X = 0$) where the coefficient of the second derivative changes sign.[8] The spatial coordinate is denoted X to show how (4.2) is related to (2.14). The case of small positive η was studied earlier by Wasow (1950) in connection with Orr–Sommerfeld theory for linear stability of parallel shear flows (though application to that problem requires taking η imaginary).

Take $\eta > 0$ (the case $\eta < 0$ has no fold and corresponds to longitudinal coupling, see Appendix B). The local dispersion relation is

$$\eta k^4 + Xk^2 + 1 = 0, \tag{4.3}$$

so it has local dispersion curve

$$k^2 = \frac{-X \pm \sqrt{X^2 - 4\eta}}{2\eta},$$

equivalently

$$X = -(\eta k^2 + k^{-2}), \tag{4.4}$$

with a fold at $X = -2\eta^{1/2}$, $k = \eta^{-1/4}$. Stix was motivated by waves in a warm magnetised plasma, but (4.2) can be viewed as a (singular) perturbation of the affine impedance cochlear model (2.14).

The general solution of (4.2) can be written as a linear combination of solutions of the form

$$p(X) = \int_\Gamma \mathrm{d}z \exp\left(\frac{\eta}{3z^3} - \frac{X}{z} - z\right), \tag{4.5}$$

for various choices of contour Γ in the complex z-plane (some are shown in Stix (1992) in the plane of $u = \mu z$). Asymptotic analysis of these solutions showed that the solutions going to 0 on the far right do so like

$$\mathrm{i}\sqrt{\frac{2X}{\pi}} K_1(2\sqrt{X}),$$

and connect to linear combinations of solutions asymptotic to

$$-\sqrt{\frac{-\pi X}{2}} H_1^{(1)}(2\sqrt{-X}) + \frac{\eta^{3/4}}{\sqrt{2}(-X)^{5/4}} e^{\mathrm{i}\left(\frac{2}{3}\eta^{-1/2}(-X)^{3/2} - \frac{\pi}{4}\right)}$$

[8] Actually, he used a parameter μ and spatial coordinate u, related to ours by $\eta = \mu^3$ and $X = -\mu u$, but the above form is more convenient for present purposes.

and its complex conjugate on the far left (a subdominant term on the right determines the linear combination). As before, K_1 and $H_1^{(1)}$ are Bessel functions: K_1 decays to 0 as its argument goes to infinity, $H_1^{(1)}$ oscillates with clockwise-rotating phase.

The interpretation is that an incoming wave from large negative X in a mode where $\eta p^{(4)}$ is negligible (the term involving $H_1^{(1)}$; this is the usual mode approaching a critical layer (2.16)) produces an evanescent wave for $X > 0$ (the term involving K_1; just as for a critical layer) and an outgoing wave for $X < 0$ in a mode where the term p in (4.2) is negligible, resulting from a balance between $\eta p^{(4)}$ and Xp'' (this would be a new mode in the cochlear context).

Note that the problem depends on the parameter η which cannot be removed by scaling, so unlike critical-layer solutions, there is not a universal form for mode-conversion solutions.

On the $k' \ll k^2$ criterion, the WKB approximation is good for $|2\eta k^3 - \frac{2}{k}| \gg 1$. This corresponds to avoiding sufficiently a neighbourhood $[k_1, k_2]$ of the wavenumber at the fold, with $k_{1,2} \sim \eta^{-1/4} \mp \frac{1}{8}\eta^{-1/2}$ for $\eta \gg \frac{1}{16}$ (the value of η at which $k^2/k' = 2(k^{-1} - \eta k^3)$ transitions from being approximately linear in k on the scale of 1 to substantially the sum of two powers), $k_1 \sim 2, k_2 \sim (2\eta)^{-1/3}$ for $\eta \ll \frac{1}{16}$, equivalently, to X being to the left of X_1 on the lower branch, X_2 on the upper branch, with $X_{1,2} \sim -2\sqrt{\eta} - \frac{1}{16}$ for $\eta \gg \frac{1}{16}, X_1 \sim -\frac{1}{4}, X_2 \sim -(\frac{\eta}{4})^{1/3}$ for $\eta \ll \frac{1}{16}$. Since the crossover occurs at a small value of η, it is valid to use the large η formulae even for η of order 1.

The theory of mode conversion was extended to contexts in which the two propagating modes reappear after a gap in x, the archetypal equation for this being called the tunnelling equation (e.g., Swanson, 1998). In this case, the conversion to the reverse mode is only partial: some energy is transmitted to the forward mode on the other side of the gap, but the amount is exponentially small as the gap grows. The tunnelling equation also has a parameter regime in which there is no mode conversion in the sense of this paper: instead two modes which exist for all x have an avoided crossing of their wave numbers. Some energy may transfer there between the modes. Confusingly, this phenomenon has also been called mode conversion by some authors, but is not what is proposed here. The scenarios for the tunnelling equation will be revisited in Section 7.

5. Mode conversion in the cochlea

To see how mode conversion could apply to the cochlea, assume that there is some physical effect which produces a force on the membrane of the form

$$\frac{\partial}{\partial x}(\eta p^{(3)}) \tag{5.1}$$

with $\eta > 0$ (putting η inside the first derivative allows for η to depend on x while respecting Newton's third law). Assuming one-dimensional fluid flow for simplicity and adding this to (2.7) produces the equation

$$\frac{\partial}{\partial x}\left(\eta(x)\frac{\partial^3 p}{\partial x^3}\right) - g(x)\frac{\partial^2 p}{\partial x^2} + p = 0 \tag{5.2}$$

with $g(x) = (m(x) - \lambda(x)/\omega^2)/\sigma$, which is increasing. This section predicts to what it would lead for the cochlea.

For illustration, take η constant and use the affine approximation (2.12) for g, the scaled coordinate X (2.13) and denote η/β^4 by η. Thus this is the Wasow equation (4.2). Use (2.4) to convert p to a for the

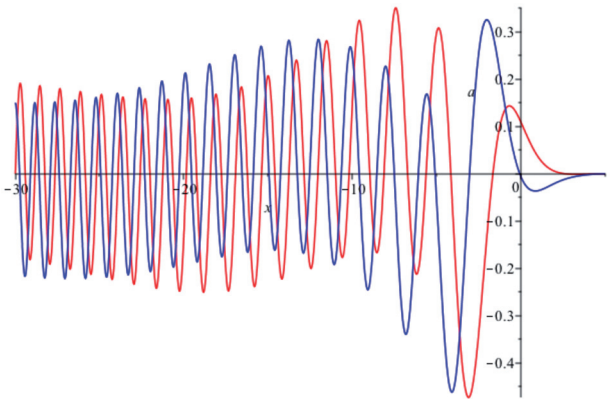

FIG. 5. Area displacement $a(X)(\propto p''(X))$ for two solutions of (4.2) for $\eta = 1$. The red curve is about $\pi/2$ ahead of the blue one in phase.

pictures (two-dimensional and three-dimensional effects are addressed later, which also modify (5.2)). To select solutions of (5.2) which decay to the right, one has to shoot from small final conditions on the right, or better treat it as a two-point boundary value problem with small or zero values of p and p' at the right. Figure 5 shows two solutions determined the latter way in Maple, differing by a phase of about $\pi/2$: one can see they are superpositions of two wave modes on the left – the underlying long one corresponds to that for the critical-layer model (cf. Fig. 1) and the superimposed larger amplitude short one is the new mode. Note that the phase velocities of both modes are positive.

One might object that the large-amplitude short-wave mode is not observed in the cochlea, but perhaps one would need greater than usual spatial resolution to see it; probably most current measurement procedures would average away the short wave leaving only the long one. More pertinent is that for short waves the fluid flow is not one-dimensional, so equation (5.2) and the conversion (2.4) need modifying. This will be done later in this section. Another resolution is that the basilar membrane is not a continuum but is made up of cochlear segments of width about 0.01 mm (Mammano & Nobili, 1993), which would prevent wavelengths from going smaller than twice this. In a similar vein, it may be more appropriate to make a model with more components than just basilar membrane motion. The reticular lamina is observed to move more than the basilar membrane (Ren *et al.*, 2016b), by a factor of up to 10, so the shortwave mode might be carried more by the reticular lamina than the basilar membrane. It should be understood that a fourth-order differential equation can result from two coupled second-order ones by expressing all in terms of one variable. Also, by comparison with negative group-velocity waves in other contexts like those near absorption bands in optical media, it is possible that the short mode is damped relatively fast so decays to the left for a solution with input power coming in the long mode from the left. Finally, the solutions depend on the parameter η (see Section 8 for a formula, but it requires knowledge of OHC responses), though numerics did not reveal a great dependence on η other than setting the wavelengths of the two modes. These issues are set to one side for the moment.

To start some analysis of the shape of the solutions, note that (5.2) has a conserved quantity, 'power flow' or 'wave energy flux' for complex solutions:

$$\Phi = \frac{1}{\sigma\omega}\Im(p\bar{p}' - \eta p''\bar{p}^{(3)}) \tag{5.3}$$

(the factor $\frac{1}{\sigma\omega}$ is included to correspond with the physically derived (2.17) in the case $\eta = 0$). This can be checked by differentiation.

Alternatively, it can be derived as a consequence of the following Hamiltonian formulation for (5.2). Equation (5.2) defines a non-autonomous linear dynamical system with respect to position x on the space of $(p, p', p'', p_3) \in \mathbb{R}^4$, where $p_3 = \eta p^{(3)}$ (the coefficients g and η are allowed to depend on x). For a solution p let

$$H = \frac{1}{2}(p'^2 - 2pp'' + gp''^2 - p_3^2/\eta),$$

and for two solutions p, q let

$$\Omega(p, q) = (pq' - qp') - (p''q_3 - q''p_3).$$

Then Ω is a symplectic form (closed non-degenerate antisymmetric function of pairs of displacements) on $\mathbb{R}^4$, and the Hamiltonian vector field of H with respect to Ω (i.e., solution V of $\Omega(V, \xi) = dH(\xi)\ \forall \xi \in \mathbb{R}^4$) can be checked to be equivalent to (5.2). Conservation of Ω on pairs of solutions is automatic; this is the Hamiltonian way to view conservation of Wronskians in lossless one-dimensional systems. In particular, for any complex solution p, then $P = \Omega(\Im p, \Re p)$ is conserved, which can be written as

$$P = \Im(p\bar{p}' - p''\bar{p}_3).$$

Up to multiplication by a physical factor, this P is the power flow Φ. It is not necessary for H to be conserved (nor even possible, because $\frac{dH(p)}{dx} = \frac{\partial H}{\partial x} = -g'p''^2 \neq 0$ in general). Such a Hamiltonian formulation is the mathematical expression for the physical concept of a 'lossless medium' at the linear time-invariant level.

Now study what the WKB approximation tells about the solutions, using conservation of power flow. For example, for the affine impedance approximation, as soon as $\eta \gg \frac{1}{16}$ the domain of validity of the WKB approximation is all but an interval of length about $\frac{1}{16}$ in X to the left of the fold point (and even for smaller η it is all but an interval of about $\frac{1}{4}$ in X).

For solutions bounded on the right, hence going to zero, then $\Phi = 0$, so it follows that the incoming and mode-converted waves carry equal and opposite power flows, which abusing notation will also be written as Φ. Then in the WKB regime for each mode, inserting $p \propto \exp(-\mathrm{i}\int k\, dx)$ into (5.3) leads to

$$|p|^2 = \frac{\sigma\omega\Phi}{|1 - \eta k^4|k}. \tag{5.4}$$

Using (2.4) to convert p to a in one-dimensional,

$$|a|^2 = \frac{k^3\Phi}{\sigma\omega^3|1 - \eta k^4|} \tag{5.5}$$

(one could redo this to allow for two-dimensional and three-dimensional fluid flow effects, by taking $\eta k^2 = -g(x) - \frac{2\rho}{w|k|}$ for $kh \gg 1$).

Take the affine impedance approximation $g(x) = \beta x$ (including the factor β this time but shifting the origin to $x(\omega)$) and η constant. Then (5.2) becomes

$$\eta p^{(4)} - \beta x p'' + p = 0.$$

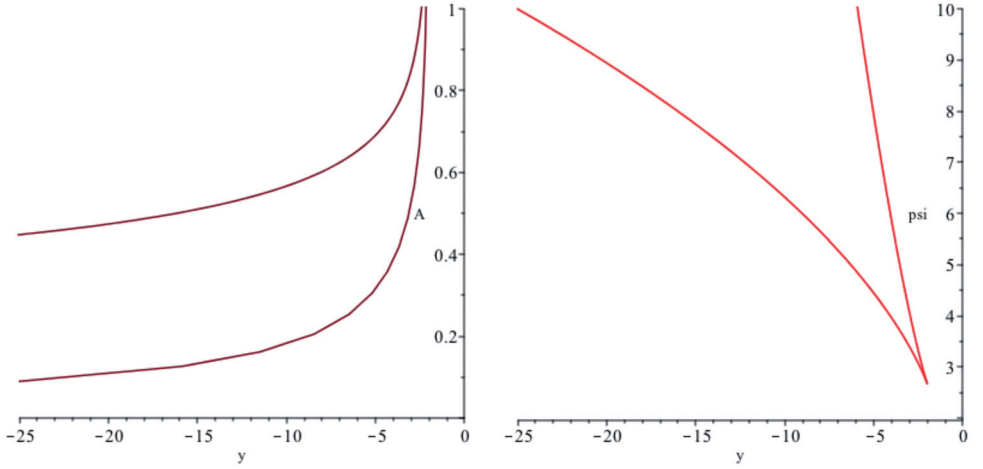

FIG. 6. Amplitude $A = |a|$ and scaled phase $\psi = \phi/\eta^{1/4}$ of oscillation as functions of $y = \beta x/\eta^{1/2}$ for the WKB approximation of (5.2). The long-wave mode has the lower amplitude and shallower phase curve.

The WKB approximation leads to local dispersion relation

$$\eta k^4 + \beta x k^2 + 1 = 0.$$

The constants η and β can be scaled out by writing $x = \eta^{1/2} y/\beta$ and $k = \eta^{-1/4}\kappa$. Then

$$\kappa^4 + y\kappa + 1 = 0,$$

or

$$y = -\kappa^2 - \kappa^{-2}. \tag{5.6}$$

The amplitude is given by

$$|a|^2 = \frac{\eta^{-3/4}\kappa^3 \Phi}{\sigma\omega^3|1-\kappa^4|}.$$

The left panel of Fig. 6 shows the WKB amplitude as a function of y for the two waves (connected through the fold even though the WKB approximation is not formally valid there). Note that on the lower branch, as $y \to -\infty$, $\kappa \sim (-y)^{-1/2}$ and up to scale, $|a| \sim \kappa^{3/2} \sim (-y)^{-3/4}$. On the upper branch, $\kappa \sim (-y)^{1/2}$ and $|a| \sim \kappa^{-1/2} \sim (-y)^{-1/4}$. Correction to WKB makes the amplitude bounded.

The phase of the waves from WKB theory is $\phi = -\int k \, \mathrm{d}x$. Using the scalings, equation (5.6) and integration by parts, one obtains $\phi = \eta^{1/4}\psi/\beta$ with

$$\psi = 2\left(\frac{\kappa^3}{3} + \frac{1}{\kappa}\right),$$

relative to the value $8/3$ at the fold point ($\kappa = 1, y = -2$). This is plotted in the right panel of Fig. 6. For the phase of the true problem, there is a non-WKB correction and a non-trivial approach to a constant on the right.

Since the maximum amplitude in mode conversion occurs near the fold rather than at the resonance, it is feasible for the range of mode-converting frequencies to be broader than that for critical-layer resonances. In particular, the frequency range extends considerably downwards for mode conversion because for low frequencies one can have a fold on the basilar membrane at a frequency whose resonant location would be off the apical end.

With addition of suitable anti-damping to the incoming wave, damping to the outgoing one and three-dimensional effects, it is likely that a fit with observed tuning curves could be obtained.

Mode conversion could explain transiently and tone-evoked OAEs, because it gives complete conversion of the incoming wave into an outgoing one, and the properties of the second mode can be widely different from the first, e.g., much slower group velocity, so it could fit the long time-delay for the echo and the accumulated-phase curves.

As a step in fitting mode conversion to the cochlea, consider a specific form for the additional term (5.1), namely that from adding a force

$$\frac{\partial^2}{\partial t^2}\frac{\partial}{\partial x}\nu\frac{\partial a}{\partial x}$$

with $\nu > 0$ to the right hand side of (2.5), to be justified physiologically in Section 8. Let us compute the phase shift and group delays for the case with exponential stiffness-variation but other properties constant. Three-dimensional fluid-flow effects are allowed, but success or failure of this particular model should not be taken as definitive for the whole class of mode-conversion models. Thus take $\lambda(x) = C\,e^{-\alpha x}$ and m, ν and the function I of (2.9) independent of x. So the dispersion relation is

$$\nu k^2 - \left(\frac{C}{\omega^2}\,e^{-\alpha x} - m\right) + I(k) = 0.$$

This inherits the simplifying feature from the exponential critical-layer model that a multiplicative change in ω is equivalent to a shift in x. It follows that

$$e^{-\alpha x} = \frac{\omega^2}{C} f(k) \tag{5.7}$$

where $f(k) = \nu k^2 + m + I(k)$.

Firstly, compute the phase shift from the base to the fold in the long-wave mode (using the WKB approximation):

$$\Delta\phi^f = -\int_0^{x_f(\omega)} k\ \mathrm{d}x,$$

where the position $x_f(\omega)$ of the fold is given by $e^{-\alpha x_f(\omega)} = \frac{\omega^2 M}{C}$ with $M = \min_k f(k)$. Define $k_-(\omega) < k_+(\omega)$ to be the positive roots of $f(k) = \frac{C}{\omega^2}$ for $\omega < \sqrt{C/M}$ (the maximum frequency for which the fold falls in the domain $x \geq 0$) and k_f to be the minimiser of f. Note that this is independent of ω, whereas experiment shows the wavenumber at the characteristic place to increase weakly with frequency

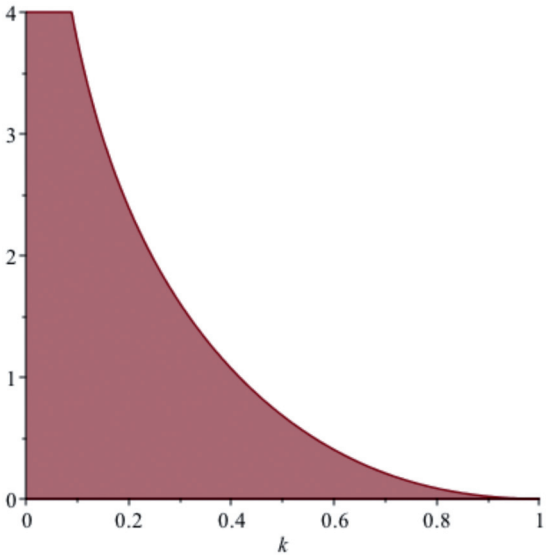

FIG. 7. The area entering the calculation of the phase delay from stapes to fold for the exponential mode-conversion model. The plotted function $g(k) = \log \frac{f(k)}{M}$, the horizontal cutoff is at height $\log \frac{C}{M\omega^2}$, and the integral extends to the minimum of g (which by construction has $g = 0$). The plot was made using $\nu = 1$ and $I(k) = \frac{1}{k \tanh k}$.

(Robles & Ruggero, 2001), but variation of k_f could be achieved by allowing ν, m or I to vary with x, at the cost of more complicated calculations. Using integration by parts, one obtains

$$\Delta\phi^f = -k_f x_f(\omega) + \int_{k_-}^{k_f} x \, \mathrm{d}k = -\frac{1}{\alpha}\left(k_- \log \frac{C}{M\omega^2} + \int_{k_-}^{k_f} \log \frac{f(k)}{M} \, \mathrm{d}k\right).$$

This gives a phase delay of magnitude $1/\alpha$ times the area shown in Fig. 7. As $\omega \to 0$ the area converges to a finite non-zero limit $L = \int_0^{k_f} \log \frac{f(k)}{M} \, \mathrm{d}k$. Thus for ω sufficiently below $\sqrt{C/M}$ the phase shift is close to $-L/\alpha$. This corresponds with observations of a phase delay from stapes to characteristic place (of about 23 radians) independent of frequency (Shera & Guinan, 1999; Robles & Ruggero, 2001). The present model predicts a reduction in the magnitude of the phase delay to the characteristic place as frequency increases, at a rate

$$\tau_p^f = -\frac{\partial \Delta\phi^f}{\partial\omega} = \frac{2k_-(\omega)}{\alpha\omega} \approx \frac{2}{\alpha}\sqrt{\frac{\sigma}{C}}$$

for low frequencies, with the phase delay going to 0 as ω approaches the maximum $\sqrt{C/M}$. It would be interesting to know whether such a change is observed. There are articles, e.g., Rhode (2007) and Ren *et al.* (2016b), reporting the phase difference ϕ from malleus to a given place as a function of frequency. This does not, however, allow one to see how the phase change to the characteristic place changes with frequency, because the characteristic place itself moves with frequency. Figure 3b of He & Ren (2013) shows phase delays to two locations with characteristic frequencies 12.4 kHz and 15 kHz as functions

of frequency for gerbils, but the phase changes at the characteristic frequencies in the two cases are very close (about -24 radians), so it is not easy to see a difference.

Secondly, compute the group delay τ_g^f from stapes to fold in the long-wave mode. Disturbances with frequency ω propagate at the group velocity $c_g = \frac{\partial \omega}{\partial k}$. Thus the time delay for a disturbance to propagate from stapes to fold is

$$\tau_g^f = \int_0^{x_f(\omega)} \frac{\mathrm{d}x}{c_g} = \int_{k_-}^{k_f} \frac{\partial x}{\partial k} \Big/ \frac{\partial \omega}{\partial k} \, \mathrm{d}k,$$

the derivative in the numerator being performed at constant ω and that in the denominator at constant x. This evaluates to

$$\tau_g^f = \frac{2}{\alpha\omega}(k_f - k_-(\omega)) \sim \frac{2k_f}{\alpha\omega}$$

for ω sufficiently below $\sqrt{C/M}$. Thus the model predicts a group delay from stapes to characteristic place inversely proportional to frequency (because k_f is independent of ω), except for frequencies near the maximum audible, where the group delay goes to zero. The group delay to a given place can be inferred from data on the phase delay as a function of frequency, because the derivative of the phase delay with respect to frequency is precisely the group delay τ_g for that frequency. This relation has been well known in cochlear mechanics for many years, but it is worth recalling the simple proof. The phase change ϕ from x_1 to x_2 is $\phi = -\int_{x_1}^{x_2} k \, \mathrm{d}x$. So its derivative with respect to frequency is

$$\frac{\partial \phi}{\partial \omega} = -\int_{x_1}^{x_2} \frac{\partial k}{\partial \omega} \, \mathrm{d}x = -\int_{x_1}^{x_2} \frac{\mathrm{d}x}{c_g} = -\tau_g.$$

Estimating the slope by eye from Fig. 3b of He & Ren (2013) for gerbils (and converting to angular frequency) shows a group delay of about 0.8 ms to the characteristic place for 12.4 kHz and about 0.55 ms to that for 15 kHz. Thus the group delay to the characteristic place is indeed less for higher frequencies but the data is too limited to check the prediction of inverse proportion. Note that the group delay (around 10 cycles) is much larger than the phase shift (3–4 cycles); mixing the two has been a source of confusion.

Thirdly, compute the phase shift for a wave entering in the long-wave mode and coming out in the short one (or equivalently, up to sign change, entering in the short-wave mode and coming back in the long-wave mode). The WKB phase shift is, integrating by parts,

$$\Delta\phi^0 = -\int k \, \mathrm{d}x = \int_{k_-}^{k_+} x \, \mathrm{d}k = -\frac{1}{\alpha}\int_{k_-}^{k_+} \log \frac{\omega^2 f(k)}{C} \, \mathrm{d}k$$

(perhaps one should add something for the non-WKB region near the fold). This is a phase advance, because the argument of the logarithm is less than 1 in the given range. Up to an additive constant and ambiguity in sign, one can infer it from the observed approximately 100 Hz modulation in the frequency response of the ear canal (Fig. 4 of Shera & Zweig, 1993). The interpretation of the oscillations in the response as a function of frequency is that a wave is coming back out of the cochlea with a phase shift which goes through multiples of 2π producing constructive interference, and odd multiples of π producing destructive interference. Thus counting the oscillations in the response between two frequencies provides the difference in their round trip phase shift. Perhaps the tidiest way to use the data is to estimate the

derivative of the phase change with respect to frequency, by taking the reciprocal of the difference in frequency between successive maxima (or minima). By the same argument as above, this derivative is the roundtrip group delay, so the 100Hz modulation observed in the frequency range 1.4–2 kHz in humans gives a roundtrip group delay of 10 ms. At lower frequencies, the modulation is slightly faster, around 75 Hz for 1.2 kHz, giving a longer group delay of about 13 ms. For the model, this derivative is

$$\tau_g^0 = -\frac{\partial \Delta\phi^0}{\partial \omega} = \frac{2}{\alpha\omega}(k_+(\omega) - k_-(\omega)).$$

For ω sufficiently below $\sqrt{C/M}$ we have $k_+(\omega) \sim \sqrt{\frac{C}{\nu}}\frac{1}{\omega}$ and k_- much smaller, so $\tau_g^0 \sim \frac{2}{\alpha}\sqrt{\frac{C}{\nu}}\frac{1}{\omega^2}$, and as ω approaches the maximum $\omega_m = \sqrt{C/M}$, then $\tau_g^0 \propto \sqrt{\omega_m - \omega}$. This decreases in magnitude as frequency increases but too fast compared with observations. As mentioned, the Shera & Zweig (1993) data show a decrease of τ_g^0 as frequency increases but suggest τ_g^0 is roughly constant over the range 1.4–2 kHz. Wilson (1980) gives $\omega\tau_g^0$ approximately constant at about $15 \times 2\pi$, but that is still one power of ω short of the prediction. So perhaps some modifications of the model are necessary, e.g. ν could vary exponentially with x too (the effects of the OHCs are thought to be larger at the base than the apex).

The results of He *et al.* (2008) and He & Ren (2013) on distortion product OAEs could also be reinterpreted in the light of mode conversion, because nonlinear effects could create outgoing waves along both branches.

In principle, one could work out a formulation in the time domain for a mode-conversion system, analogous to that of Appendix A, but further qualitative features can already be deduced without this. For example, consider generation of SOAEs. In contrast to critical-layer models, mode-conversion models have a discrete set of resonant frequencies, much as in a wind instrument but using reflection between the two modes at the fold and a combination of reflections at the base. Of these, it could be that some are unstable or at least so close to unstable that they amplify noise and hence give rise to an array of SOAEs, cf. Fig. 1 of Martin *et al.* (1988). Figure 9 of Wilson (1983) suggests a situation close to a subcritical Hopf bifurcation for one of the modes.

Finally, consider the impulse response for a mode-conversion model. An impulse at the stapes will generically produce waves in both modes, of which the short-wave mode is slower. Thus at a given location one should expect to see first a fast wave arrive, with frequency increasing to the value for the fold at that location and then decreasing as we get to see waves turning at more apical folds and coming back in the other mode, followed by a slow wave, with frequency also increasing to the value for the fold at that location and then decreasing. This could explain the two-lobed waveform of Recio *et al.* (1998). If one waits longer one might also see the effects of waves reflecting at the stapes and coming round again.

To obtain realistic predictions of the shape of the tuning curves, impulse response and OAEs one would have to make realistic models, including realistic variation of the parameters along the length, active feedback and all the relevant transfer functions, and probably nonlinear effects too.

6. Precedents

The idea of mode conversion has already been proposed in cochlear mechanics, by Huxley (1969), albeit without the terminology or results. He pointed out that it was inconsistent to leave out the effects of longitudinal coupling along the membrane and suggested that to obtain the observed frequency-to-place response a path for waves from the incoming mode to an exponentially decaying one should be included. He suggested this could be achieved by longitudinal compression of the membrane (indeed it would make

a fold in $k^2(x)$ of the type for mode conversion, though whether longitudinal compression is physical is another matter) or an effect of the spiral geometry of the cochlea (the effects of spiral geometry have been pursued by Manoussaki & Chadwick (2000), Cai *et al.* (2005), Manoussaki *et al.* (2008) and references cited therein, but mode conversion is not among the proposed consequences). Then addition of longitudinal coupling would produce a second fold in the other direction, probably before the stapes and so physically irrelevant, and give rise to an irrelevant third mode with much shorter wavelength.

The idea of studying models with more than one mode has been proposed several other times (see Section 7.1 of de Boer (1996)). In addition to the references given there, Kolston (1988) proposed an outer hair cell arcuate-pectinate (OHCAP) model, which has two modes of deformation of the basilar membrane, de Boer (1990) proposed a two-membrane model (basilar membrane and reticular lamina), de Boer (1993) proposed a model with pressure variations both across the basilar membrane and in the sulcus, and Hubbard (1993) proposed a two-mode model. Also Wilson & Bruns (1983) observed two modes of deformation of the basilar membrane in a bat. These and others are surveyed in Section 4.3 of Hubbard & Mountain (1996).

None of these authors appears to have commented on the resulting possibility of mode conversion, however. Interestingly, de Boer's 1990 model can exhibit mode conversion, namely where $hZ_{BM} = 4h_3(Z_{OC} - Z_{HC})$, though this might not occur in the physiological range (and de Boer includes damping effects in h_3 so it is complex, whereas mode conversion is clearest for light damping). Also de Boer's 1993 model can exhibit mode conversion if the coefficients have appropriate signs. He obtained

$$\begin{aligned} Gp_1'' &= Cp_1 + Dp_2 \\ Fp_2'' &= Ap_1 + Bp_2 \end{aligned} \tag{6.1}$$

with p_1, p_2 being two pressure differences and the coefficients A to G being functions of position and frequency. On the WKB approximation, the possible wavenumbers k are given by an eigenvalue problem for $-k^2$ with the result that

$$k^2 = -\frac{1}{2}\left(\frac{C}{G} + \frac{B}{F}\right) \pm \sqrt{\frac{1}{4}\left(\frac{C}{G} - \frac{B}{F}\right)^2 + \frac{AD}{FG}}.$$

Thinking of real coefficients, there is mode conversion wherever the argument of the square root passes linearly through 0, provided $\frac{C}{G} + \frac{B}{F} < 0$ there. Note in particular that this requires $\frac{AD}{FG} < 0$.

The 'second mode' of Watts (1993) is not a candidate for mode conversion in the sense here (despite his use of the term), because his is just the evanescent wave on the apical side of a critical layer. Nor is the 'fast wave' of Lighthill (1981) a candidate, because it has a much longer wavelength, this being an acoustic wave in the fluid supported by its slight compressibility.

More recently, however, Lamb & Chadwick (2011a,b) have reanalysed de Boer's 1990 two-membrane model, though with tectorial membrane rather than reticular lamina, and raised the possibility of mode conversion,[9] but their mode conversion is different from that described here. Their model fits in the class of de Boer's 1993 model but has $\frac{AD}{FG} > 0$ (damping gives $C/G - B/F$ an imaginary part, but their two

[9] Note that Lamb's comment in the discussion notes of Lamb & Chadwick (2011a): 'the MacKay paper has a single propagating mode convert to a sort of reverse evanescent mode at the cutoff. Neither addresses the problem that we raise of two propagating modes transferring energy' is totally wrong. MacKay (2006), as here, has two propagating modes which transfer their energy at the mode conversion point.

modes each have positive group velocity). So they have the case of an avoided crossing, described at the end of Section 4. It might be a good model of the cochlea, but is not mode conversion in the sense used here. Lamb & Chadwick (2014) extended their model to one with basilar membrane, reticular lamina and tectorial membrane, which appears to be still based around an avoided crossing. Reichenbach & Hudspeth (2014) studied a model for interaction between waves on the basilar membrane and Reissner's membrane, which may occur in the apical region of the cochlea for frequencies less than 1 or 2 kHz, and raise the question of how the interaction may influence cochlear mechanics. A natural answer might be by mode conversion, hence resolving the mystery of how the cochlea works for low frequencies. Chadwick *et al.* (2014) produced a cochlea-inspired toy model to demonstrate their form of mode conversion. van der Heijden (2014) made a similar one, which also has avoided crossing, referring to the phenomenon as 'mode shape swapping'.

Thus, apart from Huxley's prescient article, the author is not aware of anyone else having proposed mode conversion for the cochlea in the sense used here.

7. Two-mode models

Many models can be put into the following slight variant of de Boer's 1993 form:

$$\begin{aligned}(Gp_1')' &= Cp_1 + Dp_2\\ (Fp_2')' &= Ap_1 + Bp_2.\end{aligned} \tag{7.1}$$

This includes de Boer (1990, 1993), Lamb & Chadwick (2011a,b), Chadwick *et al.* (2014) and the version of (5.2) to be discussed in the next section:

$$\begin{aligned}\left(\frac{1}{\sigma}p'\right)' &= \omega^2 a\\ (\omega^2 \nu a')' &= -p - (m\omega^2 - \lambda)a.\end{aligned} \tag{7.2}$$

It is useful to consider the subclass of (7.1) given by Hamiltonian dynamics in space with Hamiltonian

$$H(\pi_1,\pi_2,p_1,p_2) = \frac{1}{2}\left(\frac{\pi_1^2}{m_1} + \frac{\pi_2^2}{m_2} + ap_1^2 + 2bp_1p_2 + cp_2^2\right),$$

where m_1, m_2, a, b, c may depend on frequency and position, and the standard symplectic form $\Omega = dp_1 \wedge d\pi_1 + dp_2 \wedge d\pi_2$. The spatial dynamics of H has $\pi_j = m_j p_j'$ for $j = 1, 2$, and

$$\begin{aligned}(m_1p_1')' &= -ap_1 - bp_2\\ (m_2p_2')' &= -bp_1 - cp_2.\end{aligned} \tag{7.3}$$

This is (7.1) with the symmetry $A = D$. Multiplication by constants can hide the symmetry. Thus for example, (7.2) can be put in this form with $p_1 = p$, $p_2 = \omega^2 a$, $m_1 = 1/\sigma$, $m_2 = -\nu$, $a = 0$, $b = -1$, $c = m - \lambda/\omega^2$. The Hamiltonian formulation is equivalent to that in section 5, except that σ was assumed constant there.

One advantage of recognising the Hamiltonian form is that conservation of symplectic form on pairs of solutions implies in particular for the imaginary and real parts of a complex solution that the 'power flow'

$$P = \frac{1}{\omega}\Im(m_1 p_1 \bar{p}_1' + m_2 p_2 \bar{p}_2')$$

is conserved, in the same way as for (5.2), which allows one to determine the amplitudes of WKB solutions.

Another advantage is that the question of whether mode conversion can occur has a straightforward answer: mode conversion requires the Hamiltonian to have signature[10] $++--$. The signature of the part depending on π_j is easily read off as the signs of m_1, m_2. The signature for the part depending on p_j is $+-$ if $b^2 > ac$; if $b^2 < ac$ then it is $--$ if $a, c < 0$, $++$ if $a, c > 0$. Thus signature $++--$ arises from opposite signs of m_j combined with $b^2 > ac$ or from same signs of $m_j, -a, -c$ with $b^2 < ac$. The signature can change only when a wavenumber goes through 0. Each propagating mode contributes $++$ or $--$ to the signature according as the group velocity is in the same or opposite direction to the phase velocity (under the convention that the frequency $\omega > 0$). Each decaying or growing mode contributes $+-$. Two $++$ propagating modes cannot mode convert (in the sense of this paper), nor can two $--$ modes. The Hamiltonian structure also helps one understand the generic behaviour when two modes have wave numbers that approach a common non-zero value (as position or frequency or other parameters are changed). If they have the same signature they generically make an avoided crossing. If they have opposite signature they generically make a bubble of instability bounded on each side by mode conversion.

These results are a translation of MacKay (1986) from a temporal to a spatial context. In particular, it follows from there that the signature of the Hamiltonian for a mode w with eigenvalue $-\mathrm{i}k$ can be written as the signature of $-\mathrm{i}k\Omega(w, \bar{w})$, but this is $k\omega P$. So the signature for the mode is the product of the sign of the phase velocity and the direction of power flow, which is that of the group velocity for positive energy waves as here (to justify the latter requires multisymplectic formulation of the spatiotemporal problem, which would takes paper too far afield).

Note that Hamiltonian structure for spatial evolution and the interpretation of the symplectic form as energy flow are general features of lossless wave systems (Bridges, 1992), following the specific example of uniformly travelling water waves by Baesens & MacKay (1992), where the Hamiltonian is the flow force invariant of Benjamin (1984), which has its roots in Benjamin & Lighthill (1954). In general, the spatial Hamiltonian has the interpretation of momentum flux.

8. Physiological origin of mode conversion?

What the mode-conversion explanation requires before anything else is a plausible physical reason for a fold in the wavenumber as a function of position. Two-mode (including two-membrane) models can produce mode conversion. One needs to be sure it occurs in the physiological range, but the recent experiments of Ren *et al.* (2016b) are consistent with mode conversion involving motion of the basilar membrane and the reticular membrane.

[10] The signature of a quadratic form $Q(x)$, $x \in \mathbb{R}^n$, is the list of signs $\sigma_i \in \{-1, 0, +1\}$ in its Lagrange normal form $Q(x) = \sum_{i=1}^n \sigma_i x_i^2$, in arbitrary order.

Here is a conjectured mechanism to create a fold (MacKay, 2006). The OHCs exert a force (Brownell *et al.*, 1985) on the basilar membrane. The functional form of the forces the OHCs produce is a matter of continuing debate (see Section 7.2 of de Boer (1996) for a survey up to that date, Fettiplace & Hackney (2006) for a more recent review, and Ren *et al.* (2016b) for recent experimental results showing that the reticular lamina moves more than the basilar membrane as a result of OHC contraction), but since they are contained within the organ of Corti they cannot produce a net force nor a net longitudinal torque on the basilar membrane, unless they pull on the tectorial membrane via the hairs. This problem was understood by Allen (1989) and de Boer (1993), and attempts to address it have been reviewed and extended by Fukazawa (2002). It follows that in a model where all is referred to the basilar membrane displacement, the force must be a second derivative with respect to x (or a change in transverse mode shape, but ignore this option). This could be a natural result of evolution of the hair cells sharpening the response, since such a force acts as amplification of transverse displacement with inhibition. There are three rows of OHCs and they are slanted differently along the membrane so could easily produce a second derivative response (indeed, a first derivative response was already proposed by Kolston *et al.* (1989), Steele *et al.* (1993)). In any case, a second derivative in x of the given sign creates a fold in $k^2(x)$, analogous to Huxley's suggestion of longitudinal compression.

Let us suppose the OHC force has value

$$F = -\omega^2 \frac{\partial}{\partial x} \nu \frac{\partial a}{\partial x}, \tag{8.1}$$

for some positive function $\nu(x)$ of position, allowing for the observed variation in properties of the OHCs along the membrane (especially their length, Pujol, 2013; Oghalai, 2004). The factor ω^2 is required to keep the membrane stable (stability is analysed in the next section). It is also reasonable because Mammano & Nobili (1993) claim that despite the OHCs being low-pass filters, the 'inertial reaction of the tectorial membrane makes the triggering mechanism of outer hair cells increase as the square of frequency over a wide range'. This picture might also fit with those who see the OHCs as having a natural frequency, as for frogs and turtles (Duke & Jülicher, 2003): balance between $\omega^2 \frac{\partial}{\partial x} \nu \frac{\partial a}{\partial x}$ and $(\lambda - m\omega^2)a$ could give oscillations at frequency $\omega_0 = \sqrt{\frac{\lambda}{\nu k_0^2 + m}}$ with k_0 corresponding to the minimum possible wavelength of two segments of the basilar membrane (20 μm, Mammano & Nobili, 1993), to achieve zero group speed.

Note that the proposed OHC force (8.1) is reactive rather than resistive. The possibility of a reactive component was anticipated in Gold (1948). Of course there could be, and almost certainly is, an (anti-) resistive component too to cancel damping but it might be that the principal response of the OHCs is reactive (Kolston & Smoorenburg, 1990) (this would also be sensible to reduce the power requirement of the OHCs, which is a prohibitive concern for many models). The work of de Boer & Nuttall (2000) inferring the basilar membrane impedance by fitting data to a three-dimensional critical-layer model does not allow one to settle this question, because the separation of the imaginary part into active and passive parts is not possible. Also they obtain anti-damping only in the region just before the peak response at the tested frequency: surely the anti-damping should be ready for any frequency of input and therefore distributed along the whole membrane? Nevertheless, see Sections 6.1 and 8.3 of de Boer (1996) for a distinction between 'undamping' and 'local activity': the idea is that undamping is a force proportional to a velocity whereas local activity can depend on other time and space derivatives.

Assuming the OHC force law (8.1), one obtains dispersion relation

$$\nu\omega^2 k^4 - (\lambda - m\omega^2)k^2 + \sigma\omega^2 = 0$$

(neglecting the possible contribution of longitudinal coupling which would just reduce slightly the OHC effect, and using one-dimensional fluid flow for simplicity). If m, σ and ν are taken to be roughly constant and λ to decrease with x, then one obtains a fold as in Fig. 4 at the position x where $\lambda(x) = (m+2\sqrt{\nu\sigma})\omega^2$, and the wave number at the fold is $k_f = (\frac{\sigma}{\nu})^{1/4}$. Thus k_f is roughly independent of ω, as observed, though there should be corrections from three-dimensional fluid flow and an effect from variation of ν with position.

Taking the affine approximation (2.14), the parameter η of (4.2) is given by $\eta = \frac{\nu}{\sigma\beta^4}$, evaluated near the fold.

One feature of this model is that existence of the return mode depends on OHC activity (Brownell *et al.*, 1985), so if it were inhibited (e.g., by oxygen deprivation or aspirin) then the fold would be replaced by the critical-layer curve of Fig. 4 and there would be no transiently or tone-evoked OAEs, consistent with observations (Wier *et al.*, 1988; Martin *et al.*, 1988); but of course, oxygen deprivation would also reduce the cancellation of damping and it might just be that the OAEs become more damped so not noticeable.

One refinement that can be made is to replace the derivatives by finite differences, to reflect the fact that the basilar membrane is made up of segments of width about 10 μm (Mammano & Nobili, 1993). This would saturate the wavelength of the new mode at two segment widths. Another consideration is that in addition to active OHC length changes there is also active hair bundle motion (e.g., Kennedy et al., 2006), which might contribute to the function.

There are many other directions for refinements. In particular, it may be better to view the action of the OHCs more as changing the shape of the organ of Corti than exerting force on the basilar membrane. This would produce fluid flow in the gap between the reticular lamina and the tectorial membrane which could serve to amplify the input to the inner hair cells. de Boer, 1990b already addressed the question of shape changes in the organ of Corti, and the recent paper Ren *et al.* (2016b) sheds more light on this. It is possible that the result would still be a mode-conversion model, but with the short-wave mode carried more by the reticular lamina than the basilar membrane.

The role of the tectorial membrane is not clear and probably should be included (as did Lamb & Chadwick, 2014). Two articles addressing it are Legan *et al.* (2005) and Russell *et al.* (2007).

Lastly, one should mention that there is still no agreement on the mode in which vibrations are transmitted back out of the cochlea. In addition to the articles of Ren (2004), He *et al.* (2008) and He & Ren (2013) on which I have commented, a sample of further literature is Dong & Olson (2008), de Boer *et al.* (2008), Sisto *et al.* (2011) and Ren *et al.* (2016a).

9. Local stability

A fundamental requirement of a cochlear model is that the undisturbed state normally be stable, though spontaneous OAEs show it to be close to the threshold of instability. To obtain a first idea one can study 'local stability' meaning that all properties like λ, σ, m are treated as locally constant along the membrane.

Longitudinal compression K (one of Huxley's suggestions) would lead to instability, because it would modify (2.5) to

$$m\frac{\partial^2 a}{\partial t^2} = -K\frac{\partial^2 a}{\partial x^2} - \lambda a - p,$$

which with (2.4) gives frequencies ω for real wave number k according to

$$\omega^2 = \frac{\lambda k^2 - Kk^4}{\sigma + mk^2}.$$

This is negative for $k^2 > \lambda/K$ so the membrane is unstable (to short wave buckling) (though perhaps the spiral geometry could stabilise it?).

In contrast, the proposed OHC force proportional to the second derivative in both time and space would give a stable result, because

$$m\frac{\partial^2 a}{\partial t^2} = \frac{\partial^2}{\partial t^2}\frac{\partial}{\partial x}\nu\frac{\partial a}{\partial x} - \lambda a - p \tag{9.1}$$

with (2.4) gives

$$\omega^2 = \frac{\lambda k^2}{\sigma + mk^2 + \nu k^4},$$

which is positive for all real k.

To treat the case of λ, σ, m, ν depending on position one ought to examine the spectrum of the generalised eigenvalue problem (eigenvalue ω^2)

$$\left(\frac{\lambda}{\omega^2} - m\right)\frac{\partial}{\partial x}\frac{1}{\sigma}\frac{\partial p}{\partial x} - \frac{\partial}{\partial x}\nu\frac{\partial^2}{\partial x^2}\frac{1}{\sigma}\frac{\partial p}{\partial x} + p = 0$$

with appropriate boundary conditions, for which the eigenfunctions cannot be assumed to be of the form e^{ikx}, but the above treatment should suffice to gain a first impression. In fact, consideration of the correct boundary conditions leads to the requirement to include the volume displacement at the oval window (and its opposite at the round window) as an additional dynamical variable, as was done for critical-layer models in Appendix A.

10. Conclusion

The inadequacy of critical-layer resonance models of the cochlea has been detailed and it is proposed that they be replaced by mode-conversion models.

Mode-conversion models can explain the double-lobed impulse response, the resonances in stimulated otoacoustic emissions, the minimum observed wavelength roughly independent of frequency, and the operation of the cochlea at low frequencies.

The article's main prediction for experimentalists is that a short-wave mode should occur in the cochlea in addition to the 'usual' one, depending on a balance between membrane stiffness and outer hair cell forces. It would have group velocity opposite to phase velocity. This mode might be observable in transiently evoked oto-acoustic emission experiments, but the wavelength will be shorter than those observed so far, so it would require high spatial resolution to detect.

An important extension for modellers is to analyse the effects of nonlinearity on mode-conversion models.

Acknowledgements

This article is dedicated to the people who have inspired me about waves in inhomogeneous media or cochlea mechanics: Ken Budden, Michael McIntyre, James Lighthill, Egbert de Boer, Tom Stix, Ted Evans, Andrew Huxley and Pat Wilson. I have been terribly slow about writing it (I could have written most of it in 1994, see letter to Lighthill in Fig. 8) and revising it into a shape acceptable to reviewers, and I'm afraid Budden, Lighthill, Stix and Huxley have not lived to see this version. I'm grateful to Kevin Painter for getting Huxley's paper from the library and computing some solutions of (4.2) for his Applied Mathematics undergraduate project with me in 1994, to Jonathan Ashmore for discussions in 2004 and suggesting in 2005 it could still be useful to write it up, to David Kemp for important comments in 2006, to IHES (France) for hospitality during the writing of a first version (MacKay, 2006), to anonymous reviewers for comments on versions submitted for publication in 2006 and 2008 and this one, to Robert Szalai for comments in 2011, and to the University of Warwick for study leave in 2015/6 which enabled me to make the time to revise it. The figures were produced in Maple.

References

Allen, J. B. (1989) Is basilar membrane tuning the same as neural tuning—where do we stand? *Cochlear Mechanisms: Structure, Function and Models* (J. P. Wilson & D. T. Kemp eds). New York: Plenum, pp. 453–463.

Baesens, C. N. & MacKay, R. S. (1992) Uniformly travelling water waves from a dynamical systems viewpoint: some insights into bifurcations from the Stokes family. *J. Fluid Mech.*, **241**, 333–347.

Benjamin, T. B. (1984) Impulse, flow force and variational principles. *IMA J. Appl. Math.*, **32**, 3–68.

Benjamin, T. B. & Lighthill, M. J. (1954) On cnoidal waves and bores. *Proc. Roy. Soc. Lond. A*, **224**, 448–460.

Booker, J. R. & Bretherton, F. P. (1967) The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.*, **27**, 513–539.

Bridges, T. (1992) Spatial Hamiltonian structure, energy flux and the water-wave problem. *Proc. Roy. Soc. Lond. A*, **439**, 297–315.

Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. (1985) Evoked mechanical responses of isolated cochlear outer hair cells. *Science*, **227**, 194–196.

Budden, K. G. (1955) The non-existence of a "fourth reflection condition" for radio waves in the ionosphere. *Physics of the Ionosphere.* Physical Society of London, p. 320.

Cai, H., Manoussaki, D. & Chadwick, R. (2005) Effects of coiling on the micro mechanics of the mammalian cochlea. *J. Roy. Soc. Interface*, **2**, 341–348.

Chadwick, R. S., Fourney, M. E. & Neiswander, P. (1980) Modes and waves in a cochlear model. *Hear. Res.*, **2**, 475–483.

Chadwick, R. S., Lamb, J. S. & Manoussaki, D. (2014) Stimulated acoustic emissions from coupled strings. *J. Eng. Math.*, **84**, 147–153.

de Boer, E. (1990) Wave-propagation modes and boundary conditions for the Ulfendahl-Flock-Khanna preparation. *The mechanics and biophysics of hearing*, (P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero & C. R. Steele eds), Lect. Notes Biomath., vol. 87. New York: Springer, pp. 333–339.

de Boer, E. (1990b) Can shape deformations of the organ of Corti influence the travelling wave in the cochlea?. *Hear. Res.*, **44**, 83–92.

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

de Boer, E. (1993) The sulcus connection, on a mode of participation of outer hair cells in cochlear mechanics. *J. Acoust. Soc. Am.*, **93**, 2845–2859.

de Boer, E. (1996) Mechanics of the cochlea: modeling efforts. *The Cochlea* (P. Dallos, A. N. Popper & R. R. Fay eds). New York: Springer, pp. 258–317.

de Boer, E. & MacKay, R. (1980) Reflections on reflections. *J. Acoust. Soc. Am.*, **67**, 882–890.

DE BOER, E. & NUTTALL, A. L. (2000) The mechanical waveform of the basilar membrane II: from data to models and back. *J. Acoust. Soc. Am.*, **107**, 1487–1496.

DE BOER, E., ZHENG, J., PORSOV, E. & NUTTALL, A.L. (2008) Inverted direction of wave propagation (IDWP) in the cochlea. *J. Acoust. Soc. Am.*, **123**, 1513–1521.

DONG, W. & OLSON, E. S. (2008) Supporting evidence for reverse cochlear travelling waves. *J. Acoust. Soc. Am.*, **123**, 222–240.

DUIFHUIS, H. (2012) *Cochlear Mechanics: Introduction to a Time Domain Analysis of the Nonlinear Cochlea*. New York: Springer.

DUKE, T. & JÜLICHER, F. (2003) Active traveling wave in the cochlea. *Phys. Rev. Lett.*, **90**, 158101.

EMADI, G., RICHTER, C.-P. & DALLOS, P. (2004) Stiffness of the gerbil basilar membrane: radial and longitudinal variations. *J. Neurophysiol*, **91**, 474–488.

FETTIPLACE, R. & HACKNEY, C. M. (2006) The sensory and motor roles of auditory hair cells. *Nat. Rev. Neurosci.*, **7**, 19–29.

FROSCH, R. (2010). *Introduction to Cochlear Waves.* Zurich: VDF.

FUKAZAWA, T. (2002) How can the cochlear amplifier be realised by the outer hair cells which have nothing to push against?. *Hear. Res.*, **172**, 53–61.

GOLD, T. (1948) Hearing II: the physical basis of the action of the cochlea. *Proc. Roy. Soc. Lond. B*, **135**, 492–498.

HE, W., FRIDBERGER, A., PORSOV, E., GROSH, K. & REN, T. (2008) Reverse wave propagation in the cochlea. *Proc. Natl. Acad. Sci. U.S.A*, **105**, 2729–2733.

HE, W. & REN, T. (2013) Basilar membrane vibration is not involved in the reverse propagation of otoacoustic emissions. SCI. REP. **3**, 1874 (7 pages).

HUBBARD, A. (1993) A traveling-wave amplifier model of the cochlea. *Science*, **259**, 68–71.

HUBBARD, A. E. & MOUNTAIN, D. C. (1996) Analysis and synthesis of cochlear mechanical function using models. *Auditory Computation* (H. L. Hawkins, T. A. McMullen, A. N. Popper & R. R. Fay eds). New York: Springer, pp. 62–120.

HUXLEY, A. F. (1969) Is resonance possible in the cochlea after all?. *Nature*, **221**, 935–940.

KEENER, J. & SNEYD, J. (1998) *Mathematical Physiology*. New York: Springer.

KEMP, D. T. (1978) Stimulated acoustic emission from within the human auditory system. *J. Acoust. Soc. Am.*, **64**, 1386–1391.

KEMP, D. T. (2002) Otoacoustic emissions, their origin in cochlear function, and use. *Br. Med. Bull.*, **63**, 223–241.

KENNEDY, H. J., EVANS, M. G., CRAWFORD, A. C. & FETTIPLACE, R. (2006) Depolarization of cochlear outer hair cells evokes active hair bundle motion by two mechanisms. *J. Neurosci.*, **26**, 2757–2766.

KOLSTON, P. J. (1988) Sharp mechanical tuning in a cochlear model without negative damping. *J. Acoust. Soc. Am.*, **83**, 1481–1487.

KOLSTON, P. J. (1999) Comparing *in vitro*, *in situ* and *in vivo* experimental data in a three-dimensional model of mammalian cochlear mechanics. *Proc. Natl. Acad. Sci. U.S.A*, **96**, 3676–3681.

KOLSTON, P. J. & SMOORENBURG, G. F. (1990) Does the cochlear amplifier produce reactive or resistive forces? *The mechanics and biophysics of hearing*, (P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero & C. R. Steele eds), Lect. Notes Biomath., vol. 87. New York: Springer, pp. 96–105.

KOLSTON, P. J., VIERGEVER, M. A., BOER, E. DE, & DIEPENDAAL, R. J. (1989) Realistic mechanical tuning in a micromechanical cochlear model. *J. Acoust. Soc. Am.*, **86**, 133–140.

LAMB, H. (1904) On group velocity. *Proc. Lond. Math. Soc.*, **1**, 473–479.

LAMB, J. S. & CHADWICK, R. S. (2011a) Unraveling traveling waves using WKB modelling. *What Fire is in Mine Ears. Proceedings of 11th Internat Mechanics of Hearing Workshop* (C. A. Shera & E. S. Olson eds). American Institute of Physics, pp. 97–103.

LAMB, J. S. & CHADWICK, R.S. (2011b) Dual travelling waves in an inner ear model with two degrees of freedom. *Phys. Rev. Lett.*, **107**, 088101.

LAMB, J. S. & CHADWICK, R. S. (2014) Phase of shear vibrations within cochlear partition leads to activation of the cochlear amplifier. *PLoS ONE*, **9**, e85969.

LEGAN, P. K., LUKASHKINA, V. A., GOODYEAR, R. J., LUKASHKIN, A. N., VERHOEVEN, K., CAMP, G. VAN, RUSSELL, I. J. & RICHARDSON, G. P. (2005) A deafness mutation isolates a second role for the tectorial membrane in hearing. *Nat. Neurosci*., **8**, 1035–1042.

LIGHTHILL, J. (1978) *Waves in Fluids*. Cambridge: Cambridge University Press.

LIGHTHILL, J. (1981) Energy flow in the cochlea. *J. Fluid Mech.*, **106**, 149–213.

MACKAY, R. S. (1986) Stability of equilibria of Hamiltonian systems. *Nonlinear Phenomena and Chaos* (S. Sarkar ed). Bristol: Adam Hilger, pp. 254–270.

MACKAY, R. S. (2006) Mode conversion in the cochlea? linear analysis. IHES preprint P/06/31, preprints.ihes.fr/2006/P/P-06-31.pdf (last accessed 29 July 2017); revised in 2008. University of Warwick, UK: Warwick Mathematics Institute.

MAMMANO, F. & NOBILI, R. (1993) Biophysics of the cochlea: linear approximation. *J. Acoust. Soc. Am*., **93**, 3320–3332.

MANOUSSAKI, D. & CHADWICK, R. S. (2000) Effects of geometry on fluid loading in a coiled cochlea. *SIAM J. Appl. Math*., **61**, 369–386.

MANOUSSAKI, D., CHADWICK, R. S., KETTEN, D. R., ARRIDA, J., DIMITRIADIS, E. K. & O'MALLEY, J. T. (2008) The influence of cochlear shape on low-frequency hearing. *Proc. Natl. Acad. Sci. U.S.A*, **105**, 6162–6166.

MARTIN, G. K., LONSBURY-MARTIN, B. L., PROBST, R. & COATS, A. C. (1988) Spontaneous otoacoustic emissions in a nonhuman primate I. Basic features and relations to other emissions. *Hear. Res*., **33**, 49–68.

MCFADDEN, D. & PLATTSMIER, H. S. (1984) Aspirin abolishes spontaneous oto-acoustic emissions. *J. Acoust. Soc. Am.*, **76**, 443–448.

NAIDU, R. C. & MOUNTAIN, D. C. (1998) Measurements of the stiffness map challenge a basic tenet of cochlear theories. *Hear. Res*., **124**, 124–131.

NEMIROVSKY, J., RECHTSMAN, M. C. & SEGEV, M. (2012) Negative radiation pressure and negative effective refractive index via dielectric birefringence. *Optic. Express*, **20**, 8907–8914.

NOBILI, R., MAMMANO, F. & ASHMORE, J. (1998) How well do we understand the cochlea?. *Trends Neurosci*., **21**, 159–167.

OGHALAI, J. S. (2004) The cochlear amplifier: augmentation of the traveling wave within the inner ear. *Curr. Opin. Otolaryngol. Head Neck Surg.*, **12**, 431–438.

OLSON, E. S. (2001) Intracochlear pressure measurements related to cochlear tunings. *J. Acoust. Soc. Am.*, **110**, 349–367.

OLSON, E.S. & MOUNTAIN, D. C. (1991) In vivo measurements of basilar membrane stiffness. *J. Acoust. Soc. Am.*, **89**, 1262–1275.

PETERSON, L. C. & BOGERT, B. P. (1950) A dynamical theory of the cochlea. *J. Acoust. Soc. Am.*, **22**, 369–381.

PRADA, C., BALOGUN, O. & MURRAY, T. W. (2005) Laser-based ultrasonic generation and detection of zero-group velocity Lamb waves in thin plates. *Appl. Phys. Lett*., **87**, 194109.

PROBST, R., LONSBURY-MARTIN, B. L. & MARTIN, G. K. (1991) A review of otoacoustic emissions. *J. Acoust. Soc. Am.*, **89**, 2027–2067.

PUJOL, R. (2013). Available at www.cochlea.eu/en/hair-cells/outer-hair-cells-ohcs.

RECIO, A., RICH, N. C., NARAYAN, S. S. & RUGGERO, M. A. (1998) Basilar-membrane responses to clicks at the base of the chinchilla cochlea. *J. Acoust. Soc. Am.*, **103**, 1972–1989.

REICHENBACH, T. & HUDSPETH, A. J. 2014, The physics of hearing: fluid mechanics and the active processes of the inner ear. *Rep. Prog. Phys*., **77**, 076601.

REN, T. (2004) Reverse propagation of sound in the gerbil cochlea. *Nat. Neurosci*., **7**, 333–334.

REN, T., HE, W. & BARR-GILLESPIE, P. G. (2016a) Reverse transduction measured in the living cochlea by low-coherence heterodyne interferometry. *Nat. Commun.*, **7**, 10282.

REN, T., HE, W. & KEMP, D. T. (2016b) Reticular lamina and basilar membrane vibrations in living mouse cochleae. *Proc. Natl. Acad. Sci. U.S.A*, **113**, 9910–9915.

RHODE, W. S. (2007) Basilar membrane mechanics in the 6-9 kHz region of sensitive chinchilla cochleae. *J. Acoust. Soc. Am.*, **121**, 2792–2804.

RHODE, W. S. & ROBLES, L. (1974) Evidence from Mössbauer experiments for nonlinear vibration in the cochlea. *J. Acoust. Soc. Am.*, **55**, 588–596.

RICHTER, C.-P., EDGE, R., HE, D. Z. Z. & DALLOS, P. (2000) Development of the gerbil inner ear observed in the hemicochlea. *J. Assoc. Res. Otolaryngol.*, **1**, 195–210.

ROBLES, L., RHODE, W. S. & GEISLER, C. D. (1976) Transient response of the basilar membrane measured in squirrel monkeys using the Mössbauer effect. *J. Acoust. Soc. Am.*, **59**, 926–939.

ROBLES, L. & RUGGERO, M. A. (2001) Mechanics of the mammalian cochlea. *Physiol. Rev.*, **81**, 1305–1352.

RUSSELL, I. J., LEGAN, P. K., LUKASHKINA, V. A., LUKASHKIN, A. N., GOODYEAR, R. J. & RICHARDSON, G. P. (2007) Sharpened cochlear tuning in a mouse with a genetically modified tectorial membrane. *Nat. Neurosci.*, **10**, 215–223.

SHERA, C. A. & GUINAN, J. J. (1999) Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *J. Acoust. Soc. Am.*, **105**, 782–798.

SHERA, C. A. & ZWEIG, G. (1993) Noninvasive measurement of the cochlear traveling-wave ratio. *J. Acoust. Soc. Am.*, **93**, 3333–3352.

SISTO, R., SHERA, C. A., MOLETI, A. & BOTTI, T. (2011) Forward–and reverse–traveling waves in DP phenomenology: does inverted direction of wave propagation occur in classical models? *What Fire is in Mine Ears*. Am Inst Phys Conf Proc 1403, pp. 584–589.

STEELE, C. R., BAKER, G., TOLOMEO, J. & ZETES, D. (1993) Electro-mechanical models of the outer hair cell. *Biophysics of Hair-Cell Sensory Systems* (H. Duifhuis, J. W. Horst, P. van Dijk & S. M. van Netten eds). Singapore: World Sci Publ Co, pp. 207–214.

STIX, T. H. (1962) *Theory of Plasma Waves*. New York: McGraw-Hill.

STIX, T. H. (1965) Radiation and absorption via mode-conversion in an inhomogeneous collision-free plasma. *Phys. Rev. Lett.*, **15**, 878–882.

STIX, T. H. (1992) *Waves in Plasmas*. New York: American Institute of Physics.

SWANSON, D. G. (1998) *Theory of Mode Conversion and Tunneling in Inhomogeneous Plasmas*. New York: Wiley.

TALMADGE, C. L., TUBIS, A., LONG, G. R. & PISORSKI, P. (1998) Modeling otoacoustic emission and hearing threshold fine structures. *J. Acoust. Soc. Am.*, **104**, 1517–1543.

VAN DER HEIJDEN M. (2014) Frequency selectivity without resonance in a fluid wave-guide. *Proc. Natl. Acad. Sci. U.S.A*, **111**, 14548–14552.

VOLDRICH, L. (1978) Mechanical properties of basilar membrane. *Acta Otolaryngol.*, **86**, 331–335.

WASOW, W. (1950) A study of the solutions of the differential equation $y^{(4)} + \lambda^2(xy'' + y) = 0$ for large values of λ. *Ann. Math.*, **52**, 350–361.

WATTS, L. (1993) Cochlear mechanics: analysis and analog VLSI, PhD thesis, CalTech; also Watts L, The mode-coupling Liouville-Green approximation for a two-dimensional cochlear model. *J. Acoust. Soc. Am.*, **108**, 2266–2271.

WHITHAM, G. B. (1974) *Linear and Nonlinear Waves*. New York: Wiley.

WIER, C. C., PASANEN, E. G. & MCFADDEN, D. (1988) Partial dissociation of spontaneous otoacoustic emissions and distortion products during aspirin use in humans. *J. Acoust. Soc. Am.*, **84**, 230–237.

WILSON, J. P. (1980) Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus. *Hear. Res.*, **2**, 233–252.

WILSON, J. P. (1983) Active processes in cochlear mechanics. *Physiologie et Physiopathologie des recepteurs auditifs*. Collège de France, pp. 103–125.

WILSON, J. P. & BRUNS, V. (1983) Basilar membrane tuning properties in the specialised cochlea of the CF-bat, Rhinolophus ferrumequinum. *Hear. Res.*, **10**, 15–35.

YE, D., ZHENG, D., WANG, J., WANG, Z., QIAO, S., HUANGFU, J. & RAN, L. (2013) Negative group velocity in the absence of absorption resonance. *Sci. Rep.*, **3**, 1628.

ZWEIG, G. & SHERA, C. A. (1995) The origin of periodicity in the spectrum of evoked otoacoustic emissions. *J. Acoust. Soc. Am.*, **98**, 2018–2047.

Appendix A. Time domain

On the view of the ear as a frequency analyser it is natural to work in the frequency domain, as this article has mostly done. Some issues are better discussed in the time-domain, however, such as response to an impulse ('click') (Robles *et al.*, 1976; Recio *et al.*, 1998), even though for a time-invariant linear system this is in principle deducible from the frequency response. Almost any treatment of nonlinear effects requires formulation in the time-domain (as emphasised by Duifhuis (2012)). Even the stability problem is best treated in the time-domain. So here the system of Section 2, with or without two-dimensional or three-dimensional effects and OHC forces, is formulated as a time-evolution. This point of view was taken by Mammano & Nobili (1993).

Given the area-displacement a of the membrane as a function of x along the membrane and boundary conditions at the base and apex, the pressure field is determined by the assumptions of incompressible and irrotational flow, in particular the pressure difference p across the membrane is determined as a function of x. The simplest case is the one-dimensional fluid flow approximation and no OHC force. Then the equation determining p is

$$m\frac{\partial}{\partial x}\frac{1}{\sigma}\frac{\partial p}{\partial x} - p = \lambda a. \tag{A.1}$$

This does not involve time. Physically what happens is that the pressure equilibrates rapidly to a solution of (A.1) via acoustic waves in the fluid. The situation is analogous to water waves (Whitham, 1974; Lighthill, 1978) where the velocity potential under the surface is determined instantaneously by a boundary condition at the surface, and it is another surface boundary condition that determines the evolution.

Equation (A.1) needs supplementing by boundary conditions at the base and apex, however, denoted $x = 0$ and x_h, respectively. An appropriate boundary condition at the apex, where there is a hole in the membrane called the helicotrema which permits fluid to flow from one side to the other, is that $p = \frac{\rho}{d}\frac{\partial j}{\partial t}$ where d is the diameter of the hole (it has area about 0.06 mm^2) (Lighthill, 1981). Under the one-dimensional flow approximation, (2.2) makes the boundary condition at $x = x_h$

$$p + \frac{\rho}{\sigma_h d}p' = 0, \tag{A.2}$$

where $\sigma_h = \sigma(x_h)$.

At the base, the channels end in flexible membranes called the oval and round windows. Actually these are on the sides of the channels near the base, but let us model as if at their ends. Both can flex in and out of the middle ear, and the oval window moves the stapes which moves the other ossicles and hence the ear drum (this is usually described in reverse). Although much is written about the impedance of the stapes, I am not aware of a good treatment of this boundary condition, so here is a proposed treatment.

Denote by v the volume of fluid displaced by the oval window (positive if displaced into the scala vestibula). By the assumed compressibility of the fluid and rigidity of the bone around the cochlea, an equal and opposite volume is displaced by the round window, and $j(0) = \dot{v}$. Using (2.2) again, this gives

$$p'(0) = -\sigma_0 \ddot{v}, \tag{A.3}$$

where $\sigma_0 = \sigma(0)$. Now suppose change of v and associated motion of the windows, ossicles and eardrum has an effective 'mass' M (in g cm^{-2} because with respect to a volume-coordinate) and 'elasticity' μ, so

Professor R.S.MacKay
phone: (0203) 523737; secretary: 523584
FAX: (0203) 524182
telex: 31406 COVLIB G
e-mail: robert@maths.warwick.ac.uk

Nonlinear Systems Laboratory
Mathematics Institute
University of Warwick
Coventry CV4 7AL
ENGLAND

Prof Sir James Lighthill
Provost
University College
Gower St
London WC1E 6BT

28 April 1994

Dear Professor Lighthill,

Cochlear dynamics

Trevor Stuart told me that you have done some more work on cochlear dynamics since your JFM 1981 paper. I would be most interested to learn about it. If you have further publications could you send copies to me?

I have also been looking at cochlear dynamics again, together with an undergraduate project student here. We wished to address two important points:
1. Longitudinal stiffness can not be ignored; in particular, experiments show that the wavelength never gets less than 0.7mm
2. Experiments show that the frequency selectivity depends sensitively on the oxygen supply to the cochlea.

We propose that realistic models should have:
1. a mode conversion point instead of a critical layer; this is plasma physics terminology for a fold in k^2 at a non-zero value, as a (multi-valued) function of x, for fixed ω.
2. a combination of damping and positive feedback which results in amplification of waves on the lower branch and decay of waves on the upper branch.

Maybe you have already come to the same conclusions, but I would be most interested in your comments in any case.

Yours sincerely,

Robert MacKay

Robert MacKay

Dear Robert,
Thanks very much for your interesting comments. My own recent work on cochlear biomechanics is all included in the enclosed 2 papers, and I'm delighted to send you ...
Yours 29/4/94

FIG. 8. Letter to the founding IMA president from a future one, with the founder's handwritten reply.

that

$$M\ddot{v} = -\mu v - Ap + F$$

where A is an effective mean area of the two windows and F is any external force due to interaction of the eardrum with sound waves in the outer or middle ear. Using (A.3), this gives boundary condition at $x = 0$

$$Ap - \frac{M}{\sigma_0}p' = -\mu v + F. \tag{A.4}$$

The ratio $\frac{M}{A\sigma_0}$ is an effective length of the stapes (if replaced by a cylinder of material of the density of the fluid with the same volume).

Equation (A.1), with boundary conditions (A.2) and (A.4), determines the function p in terms of the function a and the numbers v and F. Then a and v evolve in time according to

$$\begin{aligned} \frac{\partial^2 a}{\partial t^2} &= -\frac{\partial}{\partial x}\frac{1}{\sigma}\frac{\partial p}{\partial x} \\ \frac{\partial^2 v}{\partial t^2} &= -\frac{1}{\sigma_0}p'(0) = -\frac{1}{M}(\mu v + Ap(0) - F), \end{aligned} \tag{A.5}$$

the first being the result of volume and horizontal momentum conservation in the one-dimensional approximation (as in (2.4)), and the second being (A.3). The analysis of Mammano & Nobili (1993) leads to a similar formulation.

In reality, the evolution equations are more complicated. One has to allow three-dimensional fluid flow (as Mammano & Nobili (1993) point out, that is in principle straightforward, using a boundary integral formulation for solutions of Laplace's equation, though in practice requires numerical computation if one wants to include the real geometry of the cochlea), add damping effects and OHC forces, and one has to determine the radiation of sound by the eardrum and corresponding reaction force to include in F, but let us stick with system (A.1, A.2, A.4, A.5) for this section.

A first issue about the system is its stability. It is a question of the spectrum of the operator taking (a, v) to the right-hand side of (A.5) being entirely in the left half-plane. Thus, stability is decided not purely by the basilar membrane dynamics: interaction with the stapes plays a role. Since there is a critical layer for every frequency of interest, we should expect the spectrum to consist mainly of a continuum (in contrast to acoustic waves in brass and woodwind instruments, for example, where the non-zero wave speed and finite length of the instrument makes the spectrum discrete), but there should also be at least one eigenvalue, corresponding to the discrete variable v. It would be interesting to estimate the corresponding frequency.

A second issue is its impulse response. This separates into a wave train with locally defined (in space and time) frequency and wave number, and energy flow given by the group velocity. In WKB approximation, a group of waves with given frequency ω moves with the group velocity $c_g = \frac{\partial\omega}{\partial k}$. Thus for the one-dimensional critical-layer model one obtains $c_g = \frac{(\lambda - m\omega^2)^{3/2}}{\lambda\sigma^{1/2}}$, which slows to zero at the resonant location. It slows even more when two-dimensional or three-dimensional effects are included and $kh \gg 1$, to $c_g = \frac{(\lambda - m\omega^2)^2 w}{4\rho\lambda\omega}$. This makes the time for a wave-group from the stapes to reach a given x near the critical layer be $t = \frac{2\lambda}{\omega^3\sigma\beta^2\sqrt{-X}} + cst$ for the one-dimensional model provided $X < -1$, and

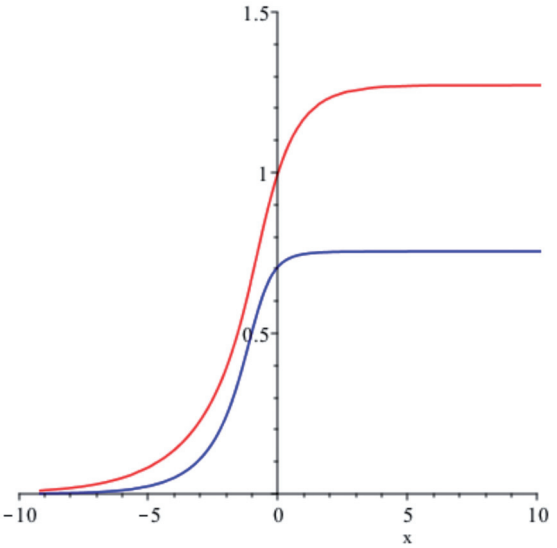

FIG. A1. The wavenumber k (in red) and the amplitude $|a|$ (in blue) for the exponential stiffness model with longitudinal coupling under the WKB approximation.

$\frac{4\lambda\rho}{\omega^3\sigma\beta^3 w|X|} + cst$ with two-dimensional or three-dimensional corrections, provided $-(\frac{h}{\beta})^2 < X < 0$, with X being the scaled position (2.13).

Appendix B. Effects of longitudinal coupling

Early models of the cochlea (Huxley, 1969; Chadwick *et al.*, 1980) took into account longitudinal stiffness of the cochlear partition. Strong anisotropy is expected because of the transverse collagen fibres, so the longitudinal stiffness should be much less than transverse. Nevertheless there must be some form of longitudinal coupling in the cochlear partition to resist extreme longitudinal gradient of displacement. The literature is unclear about what form of longitudinal coupling is believed in the cochlea partition: tension or elasticity or other? Emadi *et al.* (2004) measure a longitudinal 'space constant' of $21\mu m$ in gerbil cochlea but what is its mechanical interpretation?

Consider a general linear model for longitudinal coupling, namely for pressure $p(x) = p_0 e^{-ikx}$ (on a WKB ansatz), the area displacement is

$$a(x) = -\frac{p_0\, e^{-ikx}}{f(k,x)}$$

for some 'compliance' function f, with $f(0,x) = \lambda(x)$ as in (2.5) to correspond to the limit of uniform pressure. Then the vertical momentum equation (2.5) becomes

$$m\frac{\partial^2 a}{\partial t^2} = -p - f(k,x)a, \tag{B.1}$$

under the WKB approximation.

To proceed further, it is sensible to expand f in powers of k^2, so that to fourth order

$$m\frac{\partial^2 a}{\partial t^2} = -p - (\lambda + f_1 k^2 + f_2 k^4)a. \tag{B.2}$$

Longitudinal tension would produce f_1 equal to the integrated tension over a cross-section and $f_2 = 0$. A longitudinal component of elasticity would produce both f_1 and f_2 positive. The result depends on the boundary conditions for the edges of the basilar membrane. They are often described as cantilevered on the inside edge and clamped at the outside edge. For expositional purposes, make the assumption that f_1 dominates, so f_2 can be neglected anyway.

Stepping back into the space domain, that corresponds to

$$m\frac{\partial^2 a}{\partial t^2} = -p - \lambda a + \frac{\partial}{\partial x}\left(f_1 \frac{\partial a}{\partial x}\right), \tag{B.3}$$

where f_1 has been put inside the first derivative to satisfy Newton's third law.

Supposing $e^{i\omega t}$ time-dependence in (2.4) yields

$$a = \frac{1}{\omega^2}\frac{\partial}{\partial x}\frac{1}{\sigma}\frac{\partial p}{\partial x},$$

so combining with (B.3) one obtains

$$\left(\frac{\lambda}{\omega^2} - m\right)\frac{\partial}{\partial x}\left(\frac{1}{\sigma}\frac{\partial p}{\partial x}\right) + p + \frac{\partial}{\partial x}\left(f_1 \frac{\partial^2}{\partial x^2}\frac{1}{\sigma}\frac{\partial p}{\partial x}\right) = 0. \tag{B.4}$$

If σ is assumed constant, this equation is of the same type as that for mode conversion except for the sign of $\eta = -\frac{f_1}{\sigma\omega^2} < 0$.

The dispersion relation is

$$1 + \left(m - \frac{\lambda}{\omega^2}\right)\frac{k^2}{\sigma} - |\eta|k^4 = 0,$$

which has a single positive solution for k^2, and (with decrease of λ as x increases) leads to the wave number k increasing towards the resonant location and then plateauing afterwards (at $k^2 = \frac{m}{\sigma|\eta|} = \frac{m\omega^2}{f_1}$), though perhaps f_1 should be allowed to depend on x. The equation has the same conserved energy flux $\Phi \propto \Im(p\bar{p}' + |\eta|p''\bar{p}^{(3)})$ as derived in (5.3), so under the WKB approximation

$$|a|^2 = \frac{k^3\Phi}{\sigma\omega^3(1 + |\eta|k^4)}. \tag{B.5}$$

The results are illustrated in Fig. A1 for the case of exponential stiffness variation. Thus, if uncompensated by other effects, longitudinal coupling produces a large amplitude beyond the resonant location. This can be turned into a peak of response around the resonant location if short wavelengths are damped. But one would still have to explain the 100 Hz modulation in the frequency response of the ear canal and other results.