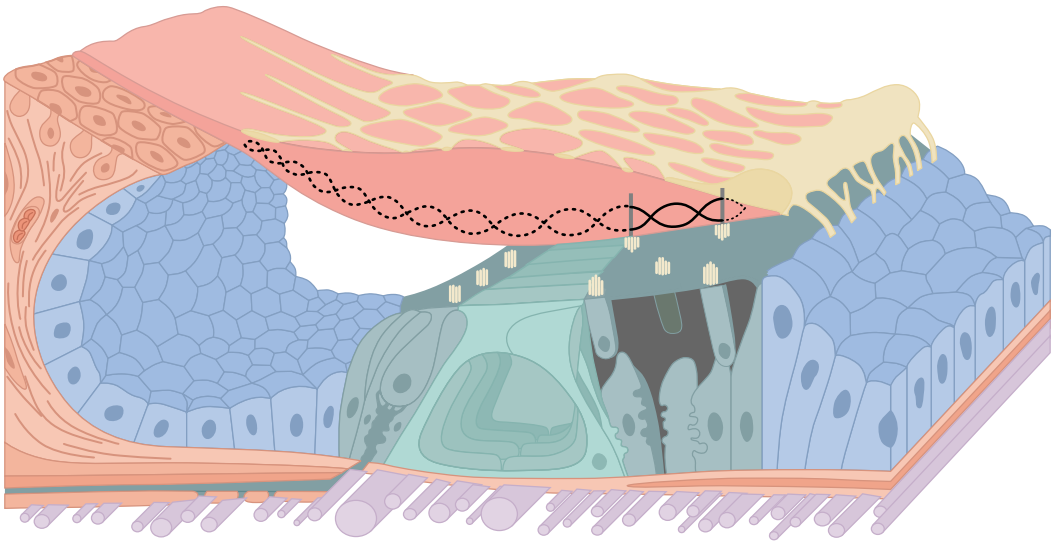
The Underwater Piano

Revival of the Resonance Theory of Hearing



Andrew Bell

The underwater piano: revival of the resonance theory of hearing

Andrew Bell

P.O. Box A348
Australian National University
Canberra, ACT 2601
Australia

e-mail: bellring@smartchat.net.au

CAPTION TO COVER FIGURE

A three-dimensional view of the cochlear partition showing a string of the 'underwater piano', represented as a resonant cavity in the tectorial membrane. The 'string' is a standing wave (shown as solid black lines between vertical marks) generated between the first and third row of outer hair cells (test-tube shaped olive-green cells at right bearing yellow tufts of stereocilia). The oscillating cavities are the ear's resonant elements, the 'piano strings' sought by Helmholtz last century, and they can be detected in the ear canal as a continuous faint ringing (spontaneous otoacoustic emission). Incoming sound is detected by inner hair cells (*middle*) as a disturbance to these resonant cavities much like a radio set detects electromagnetic waves using a regenerative receiver arrangement. An even closer analogy is the surface acoustic wave (SAW) generator, which generates electromagnetic ripples between two electrodes placed on the surface of a solid-state substrate.

Schematically, the waves are shown as transverse waves within this gelatinous structure, but physically they are more likely to propagate as capillary waves, or ripples, on its lower surface. The diagram shows the tectorial membrane sitting on top of the stereocilia, and ripples are initiated on its surface when the outer hair cells are stimulated by acoustic pressure variations (sound), in the same way as trembling willow branches overhanging a pool of water do. The pressurized cells can sense pressure variations across their cell walls and they express it as movement of their stereocilia. However, outer hair cells are reversible transducers, so that when a tuft senses a passing ripple, the movement causes an amplified cell response, or 'kick-back', which sends a ripple back in the opposite direction. The end result of this integral detector/motor system is that ripples, once generated by sound, end up reverberating between the rows of hair cells. We have a resonating cavity, like the plucked string of a piano, but in this case cellular energy is used to sustain the reverberation, enabling high Q's to exist in a watery environment (leading to the term 'underwater piano').

Like a laser cavity, oscillations can escape the resonant cavity, and in this case they travel through the tectorial membrane, past the flask-shaped inner hair cells (the ear's detectors, which signal the brain about changes in the strength of the laser beam), and reflect off the sharp edge of the inner spiral sulcus (top edge of blue area on left). When the returning wave reenters the cavity, it can give rise to an echo (evoked otoacoustic emission).

Drawing by Tara Goodsell, RSBS Graphic Design, after fig. 3 of Lim, D. J. 1980 J. Acoust. Soc. Am. 67, 1686. Used with permission of the author and the Acoustical Society of America.

May 2000

CONTEXT:

This paper outlines a radical new theory of how the ear works which reinstates the resonance model of hearing proposed by Helmholtz last century. The resonating elements, however, are not physical fibres, as Helmholtz thought, but reverberation between rows of outer hair cells, which both detect, and generate, ripples on the surface of the gelatinous tectorial membrane in response to incoming sound. Our eye can perceive sound by noting the pattern of ripples produced on the surface of a tray of water sitting on top of a loudspeaker; in a similar way, the ear can detect sound by sensing the ripples induced on the surface of a gelatinous 'pond' in the inner ear called the tectorial membrane.

Traditionally, hearing science has explained how the ear works in terms of incoming sound producing a 'traveling wave', which moves progressively from base to apex. Traveling wave theory says that hair cells detect the movement created by this wave. In contrast, the new resonance theory says that outer hair cells directly perceive sound (as pressure variations) entering the cochlea, although it is true that the summed response of these detectors does produce, at high intensities, a movement which can be described as a traveling wave (since the envelope of the response of a graded set of a resonators can be seen as a traveling wave). The difference is that the traditional explanation sees the movement first, then its detection, whereas the new theory has the detection first, then the movement. In other words, the traveling wave is an epiphenomenon and not causally efficacious – the reverse of existing theory which sees the traveling wave as providing movement which is detected by the hair cells.

Indeed, the new theory sees the movement of the basilar membrane on which the hair cells sit as a mechanism for *damping* excessive response of the detectors so that their sensing elements (the hair-like stereocilia) are not broken. Such a proposal has already been made by Martin Braun in 1996 (*Hearing Research* 97, 1–10). Vertical movement of the partition only begins at sound levels in excess of about 60 dB SPL.

Why introduce the new theory? Because the traveling wave theory is unable to satisfactorily explain why, when a microphone is placed in the ear canal, a faint, pure sound can be detected *coming out* of the ears of most subjects. Some convoluted explanations have been introduced for these so-called spontaneous otoacoustic emissions, but none are generally agreed upon or intuitive. In the new approach, the sound coming out is taken as the starting point: it is simply seen as the continual ringing of (somewhat over-active) resonant elements. And once the elements are identified as the reverberation between adjacent rows of hair cells, many previously puzzling phenomena have a natural explanation: the shape of the physical tuning curve of the cochlea (with its steep high-frequency slope and gently sloping tail); cochlear 'echoes' (when a sound is introduced to the cochlea, a tiny echo comes back a short time later); and even the occurrence of musical ratios in the spacing of hair cells.

In summary, the model can be seen as describing 'an underwater piano', a term used by Thomas Gold in 1948 when speculating how the ear, operating in fluid, could still provide long-lasting, sharply tuned resonance. He suggested that the ear operated like a 'regenerative receiver' found in radio circuits, by which some of the output is fed back to the detector circuit to enhance sensitivity and sharpen tuning. Indeed, the new theory presented here acts this way, separating the detector stage (the reverberating outer hair cells) from the output stage (the nearby, although separate, inner hair cells) which sends signals to the brain.

Another way of describing the resonant elements is to see a parallel with surface acoustic-wave resonators, solid-state devices usually used to generate frequencies in the megahertz range by sending relatively slow electromechanical pulses back and forth between two sensing and generating electrodes. An optical analogy is the Fabry-Perot etalon, in which light reflects back and forth between two lightly silvered mirrors.

Satisfyingly, the new proposal answers all the problems with current theory. Its drawbacks? The gel of the tectorial membrane must have special properties: the compliance, surface tension, or other properties must be such as to support a very low propagation speed of the ripples (or other wave propagation mode), for in this way the microscopic distance involved, some 30 μm, can be tuned to acoustic frequencies. Relevant properties of the tectorial membrane are presently unknown; nevertheless, this is question can be tested. The hypothesis is very much a live one, and, at the very least, this new theory should generate fruitful discussion and experiment. Your feedback on the proposal presented here is welcome.

Abstract: In 1857 Helmholtz proposed that the ear contained an array of sympathetic resonators, like piano strings, which served to give the ear its fine frequency discrimination. Since the discovery that most healthy human ears emit faint, pure tones (spontaneous otoacoustic emissions), it has been possible to view these narrowband signals as the continuous ringing of the resonant elements. But what are the elements? We note that motile outer hair cells lie in a precise crystal-like array with their sensitive stereocilia in contact with the gelatinous tectorial membrane. This paper therefore proposes that ripples on the surface of the tectorial membrane propagate to and fro between neighbouring cells. The resulting array of active resonators accounts for spontaneous emissions, the shape of the ear's tuning curve, cochlear echoes, and could relate strongly to music. By identifying the resonating elements that eluded Helmholtz, this hypothesis revives the resonance theory of hearing, displaced this century by the traveling wave picture, and locates the regenerative receiver invoked by Gold in 1948.

Introduction

To explain how the ear works, resonance theories of hearing — accommodating the ancient Greek idea that 'like is known by like' — have frequently been put forward. However, since Bauhin in 1605 formulated the first resonance idea on the basis of anatomy, the actual resonating elements have proved elusive (Wever 1949). First it was air-filled cavities; later, minute strings. But even Helmholtz (Helmholtz 1875), the major proponent of the resonance picture, found difficulty finding suitable candidates, at times favouring the arches of Corti, at others the transverse fibres of the basilar membrane. The problem is to find structures within the pea-sized cochlea that can resonate, like piano strings, over 3 decades of frequency.

The fibres of the basilar membrane have continued to remain the favoured tuning elements, even though it is difficult to make their combined stiffness and mass vary by the required 6 orders of magnitude (de Boer 1980; Hubbard & Mountain 1996). Moreover, these elements are closely coupled, so it is difficult to understand how the high Q that the ear displays (150 at 2.5 kHz; Gold & Pumphrey 1948) can arise. Nevertheless, it is this bank of graded resonators which auditory science has seen as the cause of the 'traveling wave', observed by von Békésy (von Békésy 1960), that underlies the stimulation of inner hair cells and the generation of neural impulses.

This paper suggests it is reverberation between neighbouring outer hair cells, communicated by capillary waves (ripples) on the tectorial membrane, that constitute the resonant elements. These elements resonate sympathetically with incoming sound energy and, because they are discrete, high Q's can be achieved.

It was just this problem of how the cochlea, immersed in fluid, could achieve the high Q's revealed by psychophysical experiments which led Gold and Pumphrey (1948) to declare that 'Previous theories of hearing are considered, and it is shown that only the resonance hypothesis of Helmholtz … is consistent with observation.'

Gold (1948) went on to posit that some sort of 'regenerative receiver' must be at work in the cochlea, and indeed searched for objective evidence for it by placing a microphone in ears which, with loud sounds, had been caused to 'ring', a phenomenon that Gold saw as a clear indication of a regenerative receiver operating with excessive positive feedback. The experiment did not meet with success, but it didn't prevent at least two other attempts to reinstate a resonance a theory of hearing: by Naftalin (Naftalin 1963; Naftalin 1981) and Huxley (Huxley 1969).

Some 30 years later, Gold's work received renewed attention when Kemp, with improved equipment, discovered that sound energy could be detected emerging from human ears when a sensitive microphone was placed in the ear canal (Kemp 1978). The sound can be observed either as an answering echo to a stimulus or, more revealingly, occur spontaneously as a continuous faint ringing now called spontaneous otoacoustic emission or SOAE (for a review see Probst et al. 1991).

Since that seminal discovery, the ear could be viewed as an active device, not a passive detector, and the motile properties of outer hair cells soon identified them as the locus of some sort of 'cochlear amplifier' (Davis 1983), although how these cells perform this function has not been clarified. This paper puts forward a physical model that unifies all these disparate features.

A physical model

The hypothesis calls on a particular feature of OHCs that has been overlooked: in all higher animals, including humans, OHCs lie in three or more rows in geometric alignment with their neighbours. Examination of published micrographs shows that the geometry is typically closely defined, much like that of a thin slice of crystal lattice, with regular alignments of hair cells in defined directions (**Fig. 1**).

This paper theorises that resonant cavities can form between lines of outer hair cells. Since OHCs are mechanical sensors/actuators of some sort, a wave disturbance in the gel of the tectorial membrane (in which the OHC stereocilia are embedded) could undergo successive amplification and reflection between the rows. This constitutes an acoustic surface wave resonator in which the stereocilia, connected to fast molecular motors in the hair-cell body, pump in acoustic energy.

A good analogy is the familiar solid-state surface acoustic wave (SAW) resonator which employs regularly placed electrodes on the surface of a crystalline material to generate (relatively slow) electromechanical ripples that resonate between the electrodes, giving stable frequencies in the megahertz range (Bell 1976).

Developing the idea of a resonant cavity between facing stereocilia, active resonators may form not only at right angles to the OHC rows but also at oblique angles where alignments of two or three hair cells occur. Herein is the genesis of the cochlea's typical tuning curve, of sets of spontaneous emissions, and perhaps, of musical ratios.

Perhaps Helmholtz was right: there are piano strings in the ear, but they are smaller and less conspicuous than he imagined. In fact, without appreciating the integral role of the tectorial membrane — called a 'peculiar' elastic membrane by Helmholtz (1875) — the resonating elements are nigh invisible. In considering Helmholtz's piano-string model, Gold (1987) asks 'how can a tiny structure of little strings immersed in liquid be so sharply tuned?' He draws an analogy to an underwater piano and points out that only by adding a positive feedback system to each string could such a device be made to work. This is what the current hypothesis does.

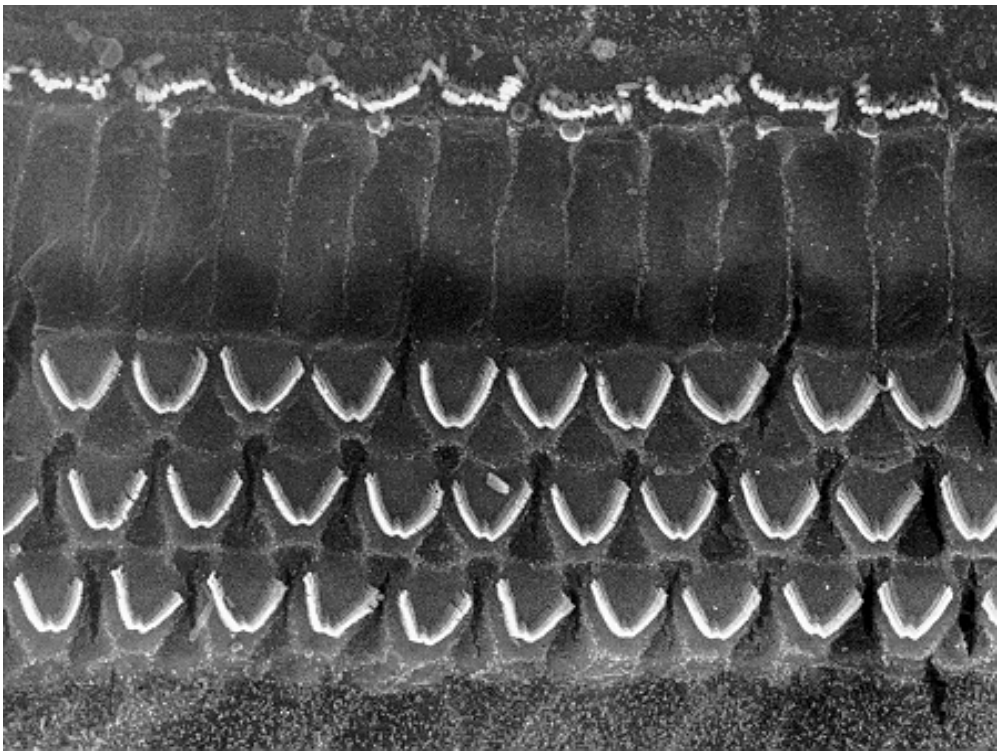


Figure 1. Geometrical arrangement of the hair cells of a rabbit, showing three rows of outer hair cells and one row of inner hair cells. Observe the regular face-centered orthorhombic arrangement of the OHCs.
[SEM courtesy of Allen Counter and the Karolinska Institutet and used with the permission of Elsevier Science Ireland Ltd. Reprinted from Counter, S. A., Borg, E. & Löfqvist, L. 1991 Acoustic trauma in extracranial magnetic brain stimulation. Electroencephalography Clin. Neurophys. 78, 173–184.]

(a) Geometry of the OHC lattice

When micrographs of the organ of Corti, as shown in Fig. 1, are examined, one is immediately struck by the regular parallel rows of OHCs (three or more) which run from the base, or high-frequency end, of the spiral cochlea to its apex, where low frequencies are detected. Not only is the inter-row spacing precisely defined (typically 15 μm in humans, but continuously graded from 10 μm at the base to 25 μm at the apex), but so too is the longitudinal spacing, usually 8–12 μm (Fig. 92 of Bredberg 1968). Cochlear geometry is fixed by birth, and stays constant throughout life (p.13 of Bredberg 1968), just like frequencies of SOAEs (Burns et al. 1994).

The typical OHC geometry of Fig. 1 is drawn schematically in **Fig. 2**, where we see that five adjoining cells can be grouped into a 'face-centered orthorhombic' unit cell with spacing a in the longitudinal direction and spacing b between OHC rows 1 and 3, an arrangement defining a diagonal at θ degrees to the transverse direction.

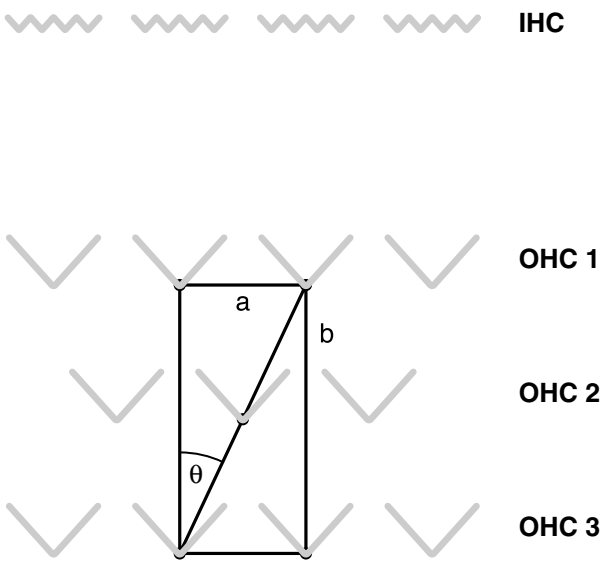


Figure 2. Schematic arrangement of hair cell geometry, showing longitudinal distance, *a*, between OHCs (along the length of the cochlea), and distance *b* between the first and third rows. The diagonal appears at an angle θ given by arctan *a*/*b*.

Measurement of a variety of published micrographs and maps of hair cell positions (cochleograms) shows that *a*/*b* centers around 0.35, so that θ, numerically arctan *a*/*b*, is usually about 20°. For humans, of 17 such examples, 12 returned a value of 20 ± 3°; for a wide variety of other vertebrate species (29 examples), more than half (16 cases) gave a value in this range (see Appendix §1). Narrow angles derive from apical regions and wide ones from basal locations; the median appears to represent the important mid-frequency region where, in humans, speech is detected and SOAEs are most prevalent.

The most common angle of about 20° means that the diagonal is (1/cos20°) times the length of the perpendicular, or 1.06. That number is a key one, for it is also the favoured ratio between neighbouring spontaneous emissions. The suggestion, detailed later, is that these two directions represent adjacent reverberating cavities.

(b) Role for the tectorial membrane

The tectorial membrane, a gelatinous acellular matrix permeated with fibres (Steel 1983), occupies a central place next to the hair cells of many animal ears, but its function in contemporary hearing theory has been secondary. This communication conjectures that its special role is as a medium supporting the propagation of slow surface waves, thereby allowing microscopic distances to be tuned to acoustic frequencies.

The two outermost rows of OHCs are like the reflecting surfaces of a Fabry-Perot etalon, except in this acoustic analogue they are active and can supply energy upon reflection. It is proposed that a wave disturbance propagating in the tectorial membrane bends OHC stereocilia and, in response, a more powerful return stroke is executed, reflecting and amplifying the disturbance and initiating continuous oscillation between the rows. As with an etalon or Helmholtz resonator, multiple reflection naturally leads to high *Q*. The more reflections, the higher the *Q*.

The required ‘kick back’ effect has been identified in OHC stereocilia (Flock 1988), but another mechanism for pumping energy into the cochlear partition can be recognised. When OHC stereocilia are deflected, the body of the cell changes length, so that the entire cell expands and contracts (with very little time delay) in synchrony with the backwards and forwards deflection (Evans & Dallos 1993); sinusoidal movement of stereocilia by ±0.05° gives synchronised length changes of 10–30 nm. This ‘mechanomotility’ can be understood in

terms of stereocilia deflection changing the cell's membrane potential, which in turn drives a fast molecular motor in the cell wall.

The phase of the mechanomotility depends on cell polarization: hyperpolarized cells react 180° out of phase to depolarized ones (Fig. 1B of Evans & Dallos 1993). A particular role for the middle row of OHCs is therefore proposed: these cells, with distinctly different polarization, respond in antiphase to the flanking rows, an arrangement ideal for continuous oscillation of the cavity. Consider a wave propagating from one row to the next: this would create a half-period time delay and lead to a total phase shift of 360° and sustained oscillation. A surface acoustic wave propagating in the tectorial membrane would fulfill this requirement. In effect, the cavity would then resonate like a pipe open at both ends and sounding in its whole-wavelength mode. Indeed, two populations of OHCs, bearing opposite response polarities, have been observed. When isolated OHCs are electrically stimulated, some 80% elongate under positive potential gradients, while the remainder contract (Kachar et al. 1986).

In summary, waves in the tectorial membrane can be created by up-and-down movement of the OHCs, and if the propagating wavefronts are subsequently sensed by bending of their stereocilia, a simple self-sustaining (and self-limiting) oscillation is set up. Gain in the system depends on the size of the polarization offset from the neutral –70 mV resting level, a factor that could be regulated by efferent activity.

(c) Generation of ripples

Up-and-down movement requires that the tectorial membrane (TM) support some form of transverse wave. Various modes of wave propagation in a gel (a polymer swollen with fluid) are possible, depending on the gel's particular internal properties (Heinrich et al. 1988; Onuki 1993), and the physics of this as applied to the TM require further study. However, the simplest model is one involving familiar capillary waves, like ripples on the surface of water, in which surface tension provides the restoring force. We thus posit ripples propagating across the surface of the TM in response to sound-induced movement by the OHCs.

The speed of propagation, c, of a capillary wave is related to the surface tension, T, by

$$c = (2\pi T/\lambda\rho)^{1/2}$$

where λ is the wavelength and ρ is the density (Lighthill 1978). Ripples are dispersive, with the speed increasing as the wavelength decreases. On the surface of water, for example, the speed is 0.86 m sec^{-1} at 1 kHz and only 0.36 m sec^{-1} at 100 Hz. Note that these are very low values compared to the velocity of a compressional wave in water, some 1500 m sec^{-1}.

Capillary waves have just the right properties for the cochlear resonators: a very low propagation speed which, in order to tune the bank of resonating cavities, decreases steadily from base to apex. Calculating values, if a cavity 30 μm long is to oscillate at 1 kHz, a propagation speed of 30 μm per 0.5 ms would be required; that is, 0.06 m sec^{-1}. Although lower than the speed of ripples on the water–air interface, this value is reasonable for a water–gel interface and could occur if the surface tension between the TM and the cochlear fluids were about 4 μN m^{-1}. At higher frequencies nearer the base, say 10 kHz, the necessary speed would be faster, typically 20 μm per 50 μs, or 0.4 m sec^{-1}, and calling for a T of 130 μN m^{-1}; whereas near the apex, say 0.1 kHz, the requisite speed would fall to 50 μm per 5 ms, or 10 mm sec^{-1}, and T would need to be less than 1 μN m^{-1}.

No measurements of surface tension of the tectorial membrane are available. These calculations, however, indicate that small values of surface tension are involved, and this means low energies. However, cochlear sensitivity is remarkably high, and some calculations indicate that hair cells can detect sound energies as low as 1 eV and deflections of their stereocilia of less than 0.01° (Bialek & Schweitzer 1985). The advantage of dealing with short-

wavelength capillary waves is that their curvature is correspondingly high, vastly easing the stereocilia's task of detecting angular deflections.

Although this hypothesis calls for small values of surface tension, this is no doubt easier to achieve than uncommonly large ones. Indeed, it would be surprising if there were no surface tension between these surfaces, particularly since the environment is electrically charged. The requisite grading in surface tension between base and apex may be achieved through regulation of electrical potentials, and the cochlear tuning map may be adjusted by efferent activity in just this way.

The simplest form of the ripple hypothesis is one having isotropic wave propagation. In this connection, the surface of the TM is covered with a thin amorphous layer in which the OHC stereocilia are embedded (see Fig. 3 of Kimura 1966) and it is this isotropic medium in which the capillary waves propagate. One property of capillary waves is high attenuation at acoustic frequencies, but this should not be a problem when dealing with distances measured in micrometres. No attenuation could in fact lead to interference problems (see Appendix §2), and exponential attenuation is factored into a model of cochlear tuning described below.

(d) Cavities in several directions and the cochlear tuning curve

So far, a resonant cavity involving reverberation of wave energy between OHC1 and OHC3 has been postulated. However, the same process that creates reverberation at right angles to the rows could also work for hair cells that are obliquely disposed, as shown in **Fig. 3**, creating a series of resonators L_0, L_1, L_2, L_3, L_4, L_5, …

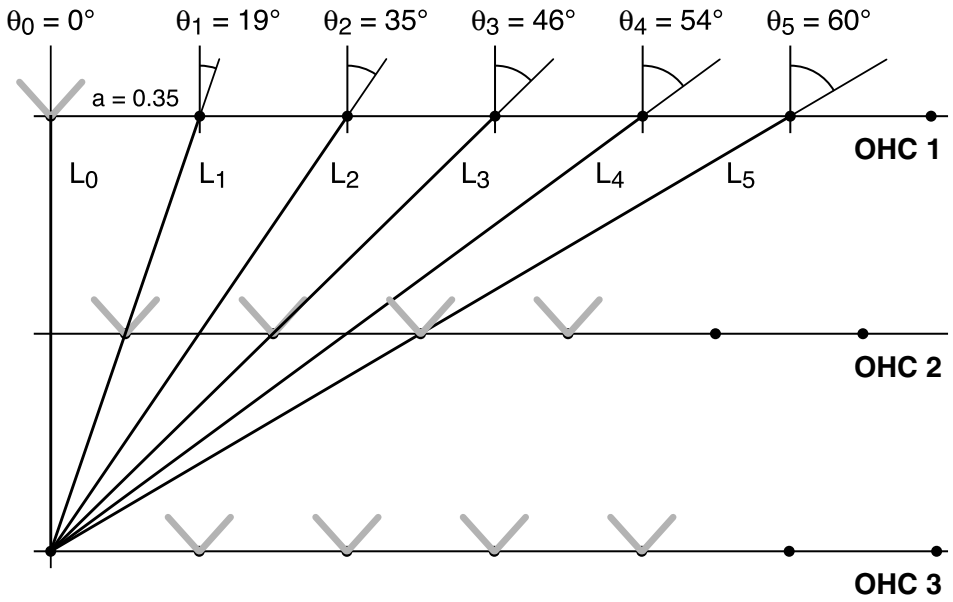


Figure 3. Geometry of outer hair cell array, elaborated from Fig. 1 with *a* = 0.35 and *b* = 1, showing multiple oblique alignments of hair cells. We obtain a set of alignments at angles θ_0, θ_1, θ_2, θ_3, θ_4, θ_5… with lengths L_0, L_1, L_2, L_3, L_4, L_5, …
The angles shown produce cavity lengths of 1.00, 1.06, 1.22, 1.44, 1.71, 2.00, …
which would have corresponding frequencies of 1.00, 0.94, 0.82, 0.69, 0.59, 0.50, …
It is noteworthy that L_1:L_0 is 1.06, close to a semitone and equal to the most common ratio between SOAEs, and that L_5:L_1 is 2:1 (an octave).

In Fig. 3, the angle between L_0 and L_1 is 19°, corresponding to the most commonly observed cochlear geometry. Accordingly, L_1 is 1.06 times longer than L_0, and so will have a resonance frequency 1.06 times lower. It is therefore hypothesised that the observed favoured ratio of 1.06 between SOAEs (Braun 1997) reflects the simultaneous excitation of these two cavities.

The L_0 mode, being the shortest, is generally expected to be the strongest, and can hence be associated with the characteristic frequency or tuning tip of the cochlear partition at that point. Neighbouring hair cells, acting like a phased array of transmitters/receivers, co-operatively generate a strong coherent wavefront. However, the strength of the first oblique mode, L_1, and the other odd-numbered alignments, is augmented because they have a middle row hair cell (in OHC 2) to help carry the wavefront from row 1 to 3 and back. OHC 2 may be considered a sort of traveling wave amplifier.

Examination of published micrographs, particularly the map of stereocilia positions for almost the entire cochleas of rhesus monkeys (Lonsbury-Martin et al. 1988), sometimes reveals a very well defined oblique. This example, and others like it, show the tendency for stereocilia arms to define certain oblique directions (that is, the arms sit at right angles to the axis of the cavity). Inspection of the rabbit cochlea in Fig. 1 indicates that the stereocilia arms are well placed to define the third and fourth oblique modes. That is, they are placed perpendicular to these cavities (which slant some 52° and 59° from the perpendicular).

If the first oblique mode (L_1 or L_{-1}) were stronger than the L_0 mode (because of the above factors), and an emission were associated with the former, then the tip of its suppression tuning curve might be expected to be ½–1 semitone higher, and be more sensitive, than at the emission frequency, and this has been observed (Bargones & Burns 1988; Abdala et al. 1996).

Significantly, multiple tips and notches are regularly seen in suppression tuning curves (Nuttall et al. 1997; Bargones & Burns 1988; Powers et al. 1995), appearing on the low-frequency or high-frequency slope (or both) depending, it is suggested, on whether the SOAE arises from L_0 or from an oblique resonator (see Appendix §3).

Of particular interest, if the response of all the resonators is summed, the result is the typical response curve of a point on the cochlear partition. That is, let us take a single high-Q resonator at L_0 with slopes of 100 dB/octave and add to it the response of the other associated cavities. We assume that the strength of a linear propagating wave front falls off, by attenuation, as a simple exponential and is further weakened, because of circular expansion, by a $1/r^2$ factor. Longer resonators will therefore make successively weaker contributions at frequencies the inverse of their length. In this model, the effect of OHC2 has been ignored for the sake of simplicity. The summation, shown in **Fig. 4** (and Appendix §4), exhibits a sharply tuned tip flanked by a very steep high-frequency slope and a more gently sloping, although somewhat notched, low-frequency tail. This curve resembles the psychophysical (de Boer 1980) and mechanical (Nutall et al. 1997) tuning curve of the cochlea. In particular, it explains the notches that are commonly seen.

Perhaps one of the clearest instances of a tuning curve in which the contributions of the individual resonators can be seen is a recent laser-beam investigation of the guinea pig cochlea (Nutall et al. 1997). In this study, tiny glass beads were placed on the basilar membrane and their movement detected with a laser doppler velocimeter. The core of that work, shown in **Fig. 5** here, shows the response of a bead to broad-band noise, and it is clear that the typical shape of the cochlea's mechanical response is generated. Note the distinct, reproducable peaks. The position of these peaks is consistent with oblique alignments of hair cells based on an orthorhombic alignment with a/b of 0.338 ($\theta_1 = 18.7°$). The predicted positions are shown on the figure, and they match the actual peaks to within 6% in frequency. These positions represent alignments in which response to imposed sound is enhanced. Observe that the odd-numbered peaks are larger than the even-numbered ones: the larger peaks correspond to alignments where a middle-row OHC occurs, facilitating ripple transmission.

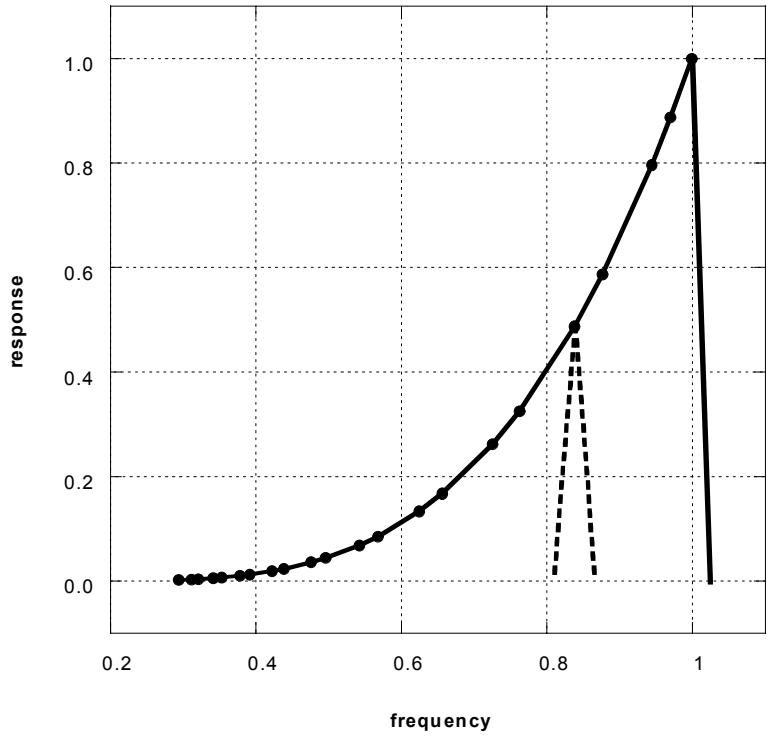


Figure 4. Summing the response of each of the cochlear resonators produces, using simple assumptions, a curve that resembles the mechanical response of the cochlear partition (and, inverted, the typical cochlear neural threshold curve). Here, each resonator L_0, L_1, L_2, L_3, …L_{11} of a set similar to that in Fig. 2 is arbitrarily assigned a *Q* of 50 derived from multiple reflections between hair cells. [The actual set is that found in Appendix 4.] Each member of the set (•) produces a peaked response (like that shown dotted for one representative member – similar peaks can be found whenever multiple reflections give rise to standing waves, such as in organ pipes or plucked strings). The L_0 resonator is assigned a response of 1 at a relative frequency of 1; other longer resonators act at progressively lower frequencies corresponding to the inverse of the cavity length (that is, the X-axis is simply the inverse of the cavity length). The Y-axis response is based on the simple attenuation of capillary waves with distance as they travel between one outer hair cell and its partner, and is therefore a simple function of resonator length; given that the amplifying ability of the hair cell is, for simplicity, taken as constant (that is, independent of the particular resonator considered, meaning that the orientation of the stereocilia arm with respect to the resonator axis is ignored), the strength of the response is assumed to diminish as a negative exponential of the cavity length, and is further weakened because of circular expansion of the wavefront by a $1/r^2$ factor. (See Appendix §4 for numerical details.) The final response envelope is shown as the full curve, although in actuality there will be notches between the points – as is frequently observed in the cochlea, and in particular, Fig. 5. Given the simplifying assumptions used, there is excellent agreement with Fig. 5, particularly the general shape of the curve and the range of frequencies which contribute to it.

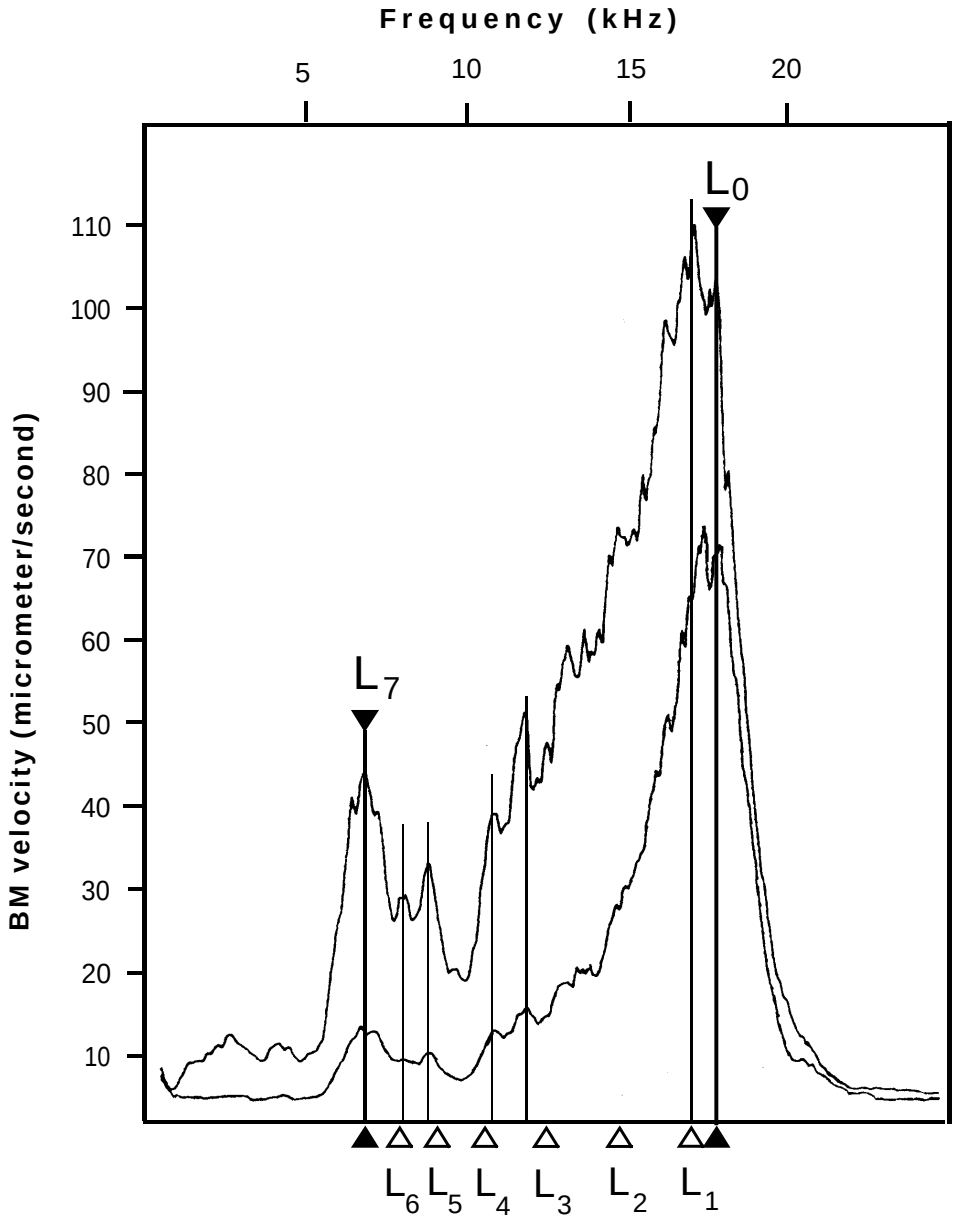


Figure 5. Peaks in the mechanical response of a guinea pig cochlea match the expected response from an outer hair cell array with $a/b = 0.338$ (or $\theta_1 = 18.7°$), a value consistent with measurements of micrographs. The response of glass beads to wide-band noise was measured with a laser doppler velocimeter (Nutall et al. 1997). Note the general reproducibility of peaks, marked with vertical lines, between stimulation at 90 dB SPL (top curve) and 80 dB (bottom). The peak at 17.8 kHz is here assumed to be the resonance associated with the L_0 cavity, and that at 6.9 kHz to be the L_7 cavity. Then L_1 to L_6 fall at the positions marked with open arrow heads: there is less than 6% deviation from the marked peaks, and as expected the odd-numbered peaks show more strongly than the even. The higher than normal response of peaks attributed to L_5, L_6, and L_7 could well come from a favourable orientation of stereocilia arms for these particular resonators (that is, the arms are approximately at right angles to these cavities).

(Adapted from Fig. 5 of Nuttall et al. 1997 and used with permission of Elsevier Science Ireland Ltd.)

(e) Implications for cochlear distortion

Summarising so far, an array of active resonant cavities stretching from one end of the cochlea to the other has been assembled, with the tuning governed by the graded propagation speed of a transverse wave in the tectorial membrane. The resonance of the shortest cavity, L_0, can be associated with the 'characteristic frequency' or tuning tip of the cochlear partition at a certain point.

But as well as this primary resonance, each cavity carries with it a set of oblique resonators, some pointing towards the base and some towards the apex. When the L_0 cavity is energised, it cannot help but excite associated oblique resonators (and vice versa) because they have hair cells in common. This arrangement renders the cochlea naturally liable to high levels of intermodulation (that is, distortion). The 'essential nonlinearity' of the cochlea, in which distortion can be detected even at the lowest stimulus levels (Goldstein 1967), may be seen as distortion remaining at the intrinsic idling levels of the active resonators.

This intrinsic linking of resonators also suggests that when an SOAE arises in one cavity, it is likely to generate other (weaker) SOAEs in neighbouring cavities. This process would explain the occurrence of linked bistable emissions (Burns et al. 1984), many of which appear at a ratio of about 1.06 (see Appendix §5).

Indeed, this linking process could underlie the observation of extensive sets of SOAEs containing several emissions. Some linking appears to arise when an extensive data set of SOAEs (Russell 1992) is examined. When the highest-frequency (perhaps shortest-cavity) SOAE is taken as a starting point, there appears to be a statistical preference for SOAEs to arise not only at the expected lower ratio of 0.95, but also at 0.77 and 0.31 (see Appendix §6; but why these particular ratios are favoured is not clear).

It is of particular import that interactions take place via resonators that are always longer (lower in frequency) than the characteristic frequency. Audiological texts describe how combination tones (involving non-linear interaction of two primary tones in the cochlea) are audible as difference frequencies (such as 2f1–f2) but sum tones (f1+f2, for example) are never heard; the paradox is that a non-linearity should generate both types (de Boer 1984). An explanation lies in seeing that interaction between the two primaries at one point on the partition can only occur via longer (lower frequency) cavities, which allows the difference tones to physically excite a resonator, but there are no such resonators higher in frequency to carry the sum tones.

(f) Deviations from regularity

There are 4–5000 sets of OHC 'triplets' ranged along the length of the organ of Corti, each possessing a characteristic frequency and carrying a dozen or more associated frequencies with it. If these banks of oscillators were perfectly placed along the partition, their summed response would be nearly complete cancelation. However, if some irregularity in the frequency–place mapping were to occur, cancelation would be imperfect and certain frequencies would come to dominate (Sutton & Wilson 1983; Wit et al. 1994). It is therefore no coincidence that humans have both the highest prevalence of spontaneous emissions compared to other animals (Probst et al. 1991) and the most irregular arrangement of outer hair cells (Bredberg 1968; Lonsbury-Martin et al. 1988), often possessing extra or missing cells. A link between these two facts has already been suggested (Manley 1983).

In terms of this hypothesis, it means that a centre of energetic activity, which gives rise to a discrete set of SOAEs, can arise as much from a gap as from a supernumary cell. It is presumed that some integrated activity over certain irregular cochlear regions is at work in generating detectable SOAEs.

(g) The gel as a delay line: evoked emissions

The resonant cavity construction also accounts for that other enigmatic phenomenon arising from an active cochlea, evoked otoacoustic emissions.

When a sound burst is conveyed to the ear, a delayed form of it, a 'Kemp echo', can be recorded some time later (Kemp 1978). A key property of this echo is that the delay is surprisingly large, typically 7 ms or 10–15 cycles (Wilson 1980). It is difficult to accommodate this delay as that incurred by the delay of a traveling wave in the forward and reverse directions (O'Mahoney & Kemp 1995). Significantly, the echo sometimes recirculates, with a fixed cycle time of some 6 ms (7.4 periods in one clear instance; Wit & Ritsma 1980). Although the envelope delay changes with stimulus intensity, the wave delay remains constant (Wilson 1980).

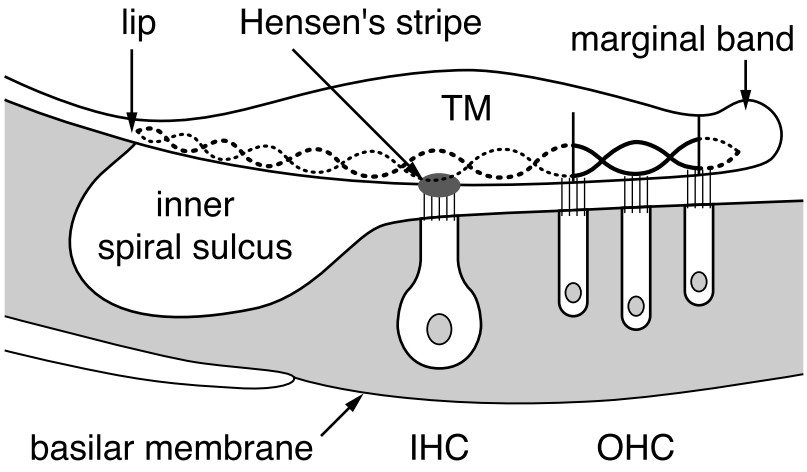


Figure 6. Radial cross-section of the tectorial membrane showing excitation of the resonant cavity between the rows of outer hair cells. Most of the energy — propagating as a slow capillary wave — is absorbed at Hensen's stripe, stimulating the inner hair cells. Some continues on towards the lip of the inner sulcus where it is reflected and re-enters the cavity, creating evoked emissions ('Kemp echoes') with about a 10-wave delay. The cycle can repeat, causing reverberation.

Looking at **Fig. 6**, the incoming compressional wave in the cochlear fluids can be pictured as immediately stimulating the resonant cavity at the edge of the tectorial membrane. However, energy emerging from the cavity would first pass the inner hair cells and then continue on towards the inner edge of the tectorial membrane. There it would encounter a sharp edge where the TM overlies the space of the inner spiral sulcus, and a wave encountering this discontinuity would be reflected back to its source. In this way, re-excitation of the cavity could occur, leading to repeated echoes. Note that the distance from OHC 1 to the limbal edge is about 5 cavity lengths, giving the right round-trip delay (about 10 cycles). Again the slow wave speed in the TM has been called upon, this time to produce a delay line.

Discussion

This paper fulfils Helmholtz's quest for resonant elements in the ear. That these elements resemble a self-sustaining tuning fork, an electromagnetically driven version of which he built and described (Fig. 33 of Helmholtz 1875), would no doubt have appealed to him. The hypothesis also satisfies Gold's demand for some type of regenerative receiver in the cochlea. Gold knew in 1947 that one would not wish to put a 'detector' — that is, a nerve fibre — right at the front end of a receiver (Gold 1989), and this hypothesis clearly separates the regenerative stage (the outer hair cells) from the detector stage (the inner hair cells).

It is therefore presumed, as a corollary, that it is the outer hair cells themselves which capture the incoming sound energy, amplify it, and pass it to the inner hair cells. The idea is expanded on in the Appendix [§§ 7&8], but it is sufficient here to recognise that OHCs are compressible elements (Zenner 1992) immersed in virtually incompressible fluid, and so the first step in transduction involves acoustic energy being converted into oscillatory energy of the OHC's cytoskeletal spring (Holley & Ashmore 1988). Individually, such an oscillation would lack adequate Q: Nature's answer has been to link two (or three) OHCs via ripples on the tectorial membrane to give time delays large enough to tune the system and sustain high-Q mechanical resonance in a fluid-filled environment — the 'underwater piano' described by Gold (1987).

In the context of pianos, one other intriguing aspect of this hypothesis demands attention, and that is the presence of musical ratios. As well as containing the semitone, the cochlear geometry in Fig. 3 contains the octave as the length of L_5:L_0. (Note also that in Fig. 5 L_5:L_0 is 1.96.) A preliminary examination of the maps of stereocilia position in a monkey (Lonsbury-Martin et al. 1988) shows that the lengths of obliques commonly involve ratios of 2:1 or 3:2 (see Appendix §9). Note also in Fig. 1 alignments of 0°, 23°, 40°, 52°, and 59° to the radial, giving corresponding lengths of 1.00, 1.09, 1.31, 1.62, and 1.97. Significantly, 1.62/1.09 is close to 3:2, and 1.97/1.00 is nearly 2:1. Similarly, the length ratios found in the Fig. 5 geometry include, as well as the octave, some other small-integer ratios (e.g., L_4:L_0 = 1.68, close to 5:3; L_2:L_0 = 1.21, close to 5:4; L_6:L_0 = 2.26, close to 9:8).

By adjusting the a/b ratio, and by tilting the unit lattice a few degrees, it is possible to create many small-integer ratios of musical significance (Appendix §§9 & 10). The question, of course, is whether the ear in fact uses such a scheme — detecting simultaneous excitation in the two arms of an outer hair cell — to detect harmonic ratios. Only extended measurements on hair cell geometries can decide the issue, but if confirmed, it would open a startling new window on music. No doubt Helmholtz would have been delighted to find musical ratios lying hidden within the cochlear geometry.

The proposal for reverberant cavities in the cochlea outlined here solves a number of puzzles in auditory theory and produces many testable predictions, and for these reasons deserves further investigation. But if validated, the idea calls for a major shift in our understanding of cochlear mechanics, replacing the dominant traveling wave picture with one involving true resonance. This is not the place to begin a critique of traveling wave theory, which presents a number of theoretical problems, but it is perhaps worth noting that a graded delay in a bank of resonators can be seen as a traveling wave, so the difference is more one of interpretation and underlying mechanism. Thus, it is still possible to hold to aspects of the older concept, but the new proposal gives us a clearer insight of what might be going on at the micromechanical level.

Acknowledgements

I thank my wife and children for support; CSIRO and many of its staff for generous cooperation, particularly Ken Hews-Taylor; and Neville Fletcher for discussion and comment. Martin Braun pointed me to Russell's data set. May no animal suffer or be sacrificed for knowledge's sake.

References:

Abdala, C.,Y. Sininger, S., Ekelid, M. & Zeng, F. G. 1996 Distortion product otoacoustic emission suppression tuning curves in human adults and neonates. Hearing Res. 98, 38–53.

Bargones, J. Y. & Burns, E. M. 1988 Suppression tuning curves for spontaneous otoacoustic emissions in infants and adults. J. Acoust. Soc. Am. 83, 1809–1816.

Bell, D. T. 1976 Surface-acoustic-wave resonators. Proc. IEEE 64, 711–721.

Bialek, W. & Schweitzer, A. 1985 Quantum noise and the threshold of hearing. Phys. Rev. Lett. 54, 725–728.

Braun, M. 1997 Frequency spacing of multiple spontaneous otoacoustic emissions shows relation to critical bands: a large-scale cumulative study. Hearing Res. 114, 197–203.

Bredberg, G. 1968 Cellular pattern and nerve supply of the human organ of Corti. Acta Otolaryngol. Suppl. 236.

Burns, E. M., Campbell, S. L. & Arehart, K. H. 1994 Longitudinal measurements of spontaneous otoacoustic emissions in infants. J. Acoust. Soc. Am. 95, 385–394.

Burns, E. M., Strickland, E. A., Tubis, A. & Jones, K. 1984 Interactions among spontaneous otoacoustic emissions. I. Distortion products and linked emissions. Hearing Res. 16, 271–278.

Davis, H. 1983 An active process in cochlear mechanics. Hearing Res. 9, 79–90.

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

de Boer, E. 1984 Auditory physics. Physical principles in hearing theory. II. Phys. Rep. 105, 141–226.

Evans, B. N. & Dallos, P. 1993 Stereocilia displacement induced somatic motility of cochlear outer hair cells. Proc. Natl. Acad. Sci. U.S.A. 90, 8347–8351.

Flock, Å. 1988 Do sensory cells in the ear have a motile function? Prog. Brain Res. 74, 297–304.

Gold, T. & Pumphrey, R. J. 1948 Hearing. I. The cochlea as a frequency analyzer. Proc. Roy. Soc. B, 135, 462–491.

Gold, T. 1948 Hearing. II. The physical basis of the action of the cochlea. Proc. Roy. Soc. B, 135, 492–498.

Gold, T. 1987 The theory of hearing. In Highlights in Science (ed. H. Messel), pp. 149–157. Sydney: Pergamon.

Gold, T. 1989 Historical background to the proposal, 40 years ago, of an active model for cochlear frequency analysis. In Cochlear Mechanisms: Structure, Function and Models (ed. J. P. Wilson & D. T. Kemp), pp. 299–305. New York: Plenum.

Goldstein, J. L. 1967 Auditory nonlinearity. J. Acoust. Soc. Am. 41, 676–699.

Heinrich, G., Straube, E. & Helmis, G. 1988 Rubber elasticity of polymer networks: theories. Adv. Polym. Sci. 85, 33–87.

Helmholtz, H. L. F. 1885 On the Sensations of Tone as a Physiological Basis for the Theory of Music (trans. A. J. Ellis), pp. 174–226. London: Longmans, Green, and Co.

Holley, M. C. & Ashmore, J. F. 1988 A cytoskeletal spring in cochlear outer hair cells. Nature 335, 635–637.

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

Huxley, A. F. 1969 Is resonance possible in the cochlea after all? Nature 221, 935–940.

Kachar, B., Brownell, W. E., Altschuler, R. & Fex, J. 1986 Electrokinetic shape changes of cochlear outer hair cells. Nature 322, 365–371.

Kemp, D. T. 1978 Stimulated acoustic emissions from within the human auditory system. J. Acoust. Soc. Am. 64, 1386–1391.

Kimura, R. S. 1966 Hairs of the cochlear sensory cells and their attachment to the tectorial membrane. Acta Oto-laryng. 61, 55–72.

Lighthill, J. 1978 Waves in Fluids, pp. 221–229. Cambridge University Press.

Lonsbury-Martin, B. L., Martin, G. K., Probst, R. & Coats, A. C. 1988 Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy. Hearing Res. 33, 69–93.

Manley, G. A. 1983 Frequency spacing of acoustic emissions: a possible explanation.In Mechanisms of Hearing (ed. W. R. Webster & L. M. Aitken), pp. 36–39. Clayton: Monash University Press.

Naftalin, L. 1963 The transmission of acoustic energy from air to the receptor organ in the cochlea. Life Sciences 2, 101–106.

Naftalin, L. 1981 Energy transduction in the cochlea. Hearing Res. 5, 307–315.

Nuttall, A. L., Guo, M., Ten, T. & Dolan, D. F. 1997 Basilar membrane velocity noise. Hearing Res. 114, 35–42.

O'Mahoney, C. F. & Kemp, D. T. 1995 Distortion product otoacoustic emission delay measurements in human ears. J. Acoust. Soc. Am. 97, 3721–3735.

Onuki, A. 1993 Theory of phase transitions in polymer gels. Adv. Polym. Sci. 109, 63–121.

Powers, N. L., Salvi, R. J., Wang, J., Spongr, V. & Qui, C. X. 1995 Elevation of auditory thresholds by spontaneous cochlear oscillations. Nature 375, 585–587.

Probst, R. Lonsbury-Martin, B. L., & Martin, G. K. 1991 A review of otoacoustic emissions. J. Acoust. Soc. Am. 89, 2027–2067.

Russell, A. F. 1992 Heritability of Spontaneous Otoacoustic Emissions (PhD thesis, U. of Illinois). Ann Arbor: UMI.

Steel, K. P. 1983 The tectorial membrane of mammals. Hearing Res. 9, 327–359.

Sutton, G. J. & Wilson, J. P. 1983 Modelling cochlear echoes: the influence of irregularities in frequency mapping on summed cochlear activity. In Mechanics of Hearing (ed. E. de Boer & M. A. Viergever), pp. 83–90. Delft University Press.

von Békésy, G. 1960 Experiments in Hearing (trans. E. G. Wever), pp. 485–510. New York: McGraw-Hill.

Wever, E. G. 1949 Theory of Hearing 9–75 New York: Dover.

Wilson, J. P. 1980 Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus. Hearing Res. 2, 233–252.

Wit, H. P. & Ritsma, R. J. 1980 Evoked acoustical responses from the human ear: some experimental results. Hearing Res. 2, 253–261.

Wit, H. P., van Dijk, P. & Avan, P. 1994 On the shape of (evoked) otoacoustic emission spectra. Hearing Res. 81, 208–214.

Zenner, H. P., Gitter, A. H., Rudert, M. & Ernst, A. 1992 Stiffness, compliance, elasticity and force generation of outer hair cells. Acta Otolaryngol. 112, 248–253.

Appendix

Bell, A. The underwater piano: revival of the resonance theory of hearing.

Appendix 1a Measured cochlear geometry in humans

Measurement of angles of a variety of published micrographs, tracings, and maps of hair cell positions (cochleograms) shows that a/b centers around 0.35, so that θ, numerically arctan a/b, is usually about 20°. For humans, of 17 such examples, 12 returned a value of 20 ± 3°, as the following table reveals.

Table A-1a.
First oblique angle in OHC cell geometry of the adult human organ of Corti

author	figure number	location	a (μm)	b (μm)	a/b	arctan a/b (degrees)
Tracings and maps of hair cell positions (cochleograms)						
Retzius (1884)	Fig. 8				0.36 ±0.02	19.7 ±0.7
Bredberg *et al.* 1965	Fig. 19	base–middle			0.36, 0.37, 0.36, 0.37, 0.37	19.8, 20.1, 19.8, 20.1, 20.1
Bredberg (1968)	Fig. 40	base			0.37 ±0.02	20.1 ±0.7
Micrographs						
Kimura *et al.* (1964)	Figs 3A, 3B	base	7.7	19	0.41, 0.32	22.5, 18.0
Johnsson &	Fig. 10a	base	15	31	0.49	26.3
Hawkins (1967)	Fig. 10b	middle	12	49	0.25	14.2
	Fig. 10c	apex	11	44	0.25	14.2
Bredberg (1968)	Figs 39, 41A, 42			20 (base) –50 (apex)	0.32, 0.45, 0.39	17.9, 24.6, 21.2
Wright (1984)	Figs 1, 2	middle	9.2	35	0.27, 0.37	14.9, 20.1

Table A-1a caption: Measurements of the first oblique angle in published micrographs and in tracings of cell positions in the adult human organ of Corti. The angle was derived by measuring the average longitudinal spacing, *a*, and radial spacing, *b*, of hair cells in OHC rows 1 and 3 and taking the arctan of that ratio. Distance *a* was taken as the average over the number of equi-spaced (no missing or supernumerary cells) cells visible; *b* was derived by drawing lines by eye through the rows of hair cells and measuring their separation. Reference point on all hair cells was the junction of the two stereocilia arms. Although a range of angles occurs, there is a cluster of values near 20°, which also represents the median. Narrow angles originate from the apex and broad ones from the base. This paper suggests it is significant that when the fifth oblique is at 60° (L_5 = 2.0, an octave), the first oblique is at 19.1° (a semitone); therefore, it is possible to explain angles near 14° from recognising that when the *seventh* oblique is at 60°, the first oblique is at 13.9°. [For a pair of SOAEs f_0 and f_1 with approximate semitone spacing and relative frequency 1.00 and r_1 (=f_1/f_0), simple trigonometry gives $\theta_1 = \arccos r_1$. Subsequent frequencies in the set will then occur at $r_n = \cos \theta_n = \cos(\arctan(n \tan(\arccos\theta_1)))$, where n is 2, 3, 4, 5, …]

References: G. Retzius (1884) in G. Bredberg, *Acta Otolaryngol.* Suppl. **236**, (1968), p. 41; G. Bredberg, H. Engstrom, H. W. Ades, *Arch. Otolaryng.* **82,** 462 (1965); G. Bredberg, *ibid.;* R. S. Kimura, H. F. Schuknecht, I. Sando, *Acta Otolaryngol.* **58,** 390 (1964); L.-G. Johnsson and J. E. Hawkins, *Arch. Otolaryng.* **85,** 43 (1967); A. Wright, *Hearing Res.* **13,** 89 (1984).

Appendix 1b. Measured cochlear geometry in some non-human vertebrates

Measurement of angles of a variety of published micrographs, tracings, and maps of hair cell positions (cochleograms) shows that a/b centers around 0.35, so that θ, numerically arctan a/b, is usually about 20°. For a wide variety of non-human vertebrate species (29 examples), more than half (16 cases) gave a value in the range 20 ± 3°, as the following table shows. Values close to 30° (from near the base) may reflect the relationship that $\theta_1 = 30°$ if $L_3 = 2.0$.

Table A-1b. First oblique angle in published micrographs of the organ of Corti from non-human vertebrates

author	figure number	animal	*a* (μm)	*b* (μm)	arctan *a*/*b* (degrees)
Engström *et al.* (1966*a,b*)	Figs 4, 24, 111	monkey			21.4, 21.1, 25.0
Lonsbury-Martin *et al.* (1988)	Fig. 3	monkey			
		(80% from apex)	8.7	14	31.8
		(46% from apex)	9.5	25	20.8
		(20% from apex)	8.5	30	15.8
Altschuler & Fex (1986)	Fig. 4a	guinea pig	4.2	18	16.3, 13.2, 21.4
Bredberg (1968)	Fig. 7	g.p.	7.0	17	22.7
Veldman *et al.* (1990)	Fig. 11a	g.p.			18.3
Zhou & Pickles (1996)	Figs 2, 3	g.p.			20.1 (av.) at apex – 36.9 at base
Engström *et al.* (1996*b*)	Figs. 13, 14, 26, 133,	g.p.			17.1, 18.0, 21.9, 30.4
Harada (1983)	Fig. 107	rabbit	7.8	23	18.6
Engström *et al.* (1966*b*)	Figs 19, 20, 21, 27, 33, 34	rabbit			20.4, 21.2, 21.2, 21.9, 28.8, 33.8
Counter (1993)	Fig. 7	rabbit			23.0
Altschuler *et al.* (1995)	Fig. 1D	rat			18.6
Richardson *et al.* (1989)	Fig. 3b	mouse	7.7	22	19.0
Harrison & Hunter-Duvar (1988)	Fig. 6	chinchilla	5.0	14.8	18.5
Bredberg (1968)	Fig. 95D	dog			12.3
Keidel *et al.* (1983) [supplied by Bredberg]	Fig. 27	cat			21.5
Engstrom et al. (1966b)	Fig. 31	cat			14.9

Table A-1b caption. As for Table A-1a, but for a range of non-human vertebrates.

References: H. Engstrom, H. H. Lindeman, H. W. Ades, in *Second Symposium on the Role of the Vestibular Organs in Space Exploration*, (NASA, Washington, 1966*a*), pp. 33–46; H. Engström, H. W. Ades, A. Andersson, *Structural Pattern of the Organ of Corti* (Almqvist & Wiksell, Stockholm, 1966*b*); B. L. Lonsbury-Martin, G. K. Martin, R. Probst, A. C. Coats, *Hearing Res.* **33,** 69 (1988); R. A. Altschuler, and J. Fex, in *Neurobiology of Hearing: The Cochlea*, R. A. Altschuler, D. W. Hoffman, R. P. Bobbin, Eds, (Raven Press, New York, 1986), pp. 383–396; G. Bredberg, *Acta Otolaryngol.* Suppl. 236 (1968); J. E. Veldman, F. M. J. Albers, P. R. W. J. Ruding, J. C. M. J. de Groot, E. H. Huizing, *Adv. Otorhinolaryngol.* **45,** 154 (1990); S. Zhou, and J. O. Pickles, *Hearing Res.* **100,** 33 (1996); Engström et al. (1966b) ibid.; Y. Harada, *Atlas of the Ear by Scanning Electron Microscopy* (MTP Press, Lancaster, 1983); Engström et al. (1996*b*) *ibid.*; S. A. Counter, *Scandinavian Audiology* **22** Suppl. 37 (1993); R. A. Altschuler, Y. Raphael, H. H. Lim, J. Dupont, K. Sato, J. M. Miller, in *Active Hearing,* Å. Flock,.D. Ottoson, M. Ulfendahl, Eds, (Pergamon, Oxford, 1995) pp. 239–255; G. P. Richardson, I. J. Russell, R. Wasserkort, M. Hans, in *Cochlear Mechanisms*, J. P. Wilson and D. T. Kemp, Eds, (Plenum, New York, 1989) pp. 57–65; R. V. Harrison and I. M. Hunter-Duvar, in *Physiology of the Ear,* A. F. Jahn, and J. Santos-Sacchi, Eds, (Raven Press, New York, 1988) pp. 159–171; G. Bredberg, (1968) *ibid.*; W. D. Kiedel, S. Kallert, M. Korth, *The Physiological Basis of Hearing* (Thieme-Stratton, New York, 1983) p. 33.

Appendix 2. Diplacusis echotica: justification for a very low propagation speed

The simplest form of the ripple hypothesis is one having isotropic wave propagation. In this connection, the surface of the TM is covered with a thin amorphous layer in which the OHC stereocilia are embedded (see Fig. 3 of Kimura, 1966) and it is this isotropic medium in which the capillary waves propagate. One property of capillary waves is high attenuation at acoustic frequencies, but this should not be a problem when dealing with distances measured in micrometres. No attenuation could in fact lead to interference problems, as one curious phenomenon — diplacusis echotica — seems to demonstrate.

Diplacusis echotica is a phenomenon in which subjects complain of double hearing: they hear both a sound and its echo. It occurs in less than 1 in 500 audiology patients (Götze, 1963) and has been recorded in audiology textbooks for more than a century (Urbantschitsch, 1910), although given little attention. The echo occurs as long as ½ to 1 second after the sound (Götze, 1963; Shambaugh, 1940). Such a large delay appears impossible acoustically unless we take the very low propagation speed proposed, little attenuation, and longitudinal transmission of wave energy along the TM. The length of the human TM is about 30 mm, so that with a wave speed of 40 mm/s or less, delays of the order of 1 second are indeed possible acoustically.

Close audiological examination of individuals with double hearing would provide good evidence for the theory presented here.

Götze, Á. Clinical observations concerning the question on diplacusis and echoacusis. *Int. Audiology* **2,** 214–216 (1963). "The time between the first sound and echo is much shorter than 1 second." (p. 215)

Kimura, R.S. Hairs of the cochlear sensory cells and their attachment to the tectorial membrane. *Acta Oto-laryngol.* **61,** 55–72 (1966).

Shambaugh, G. E. Diplacusis: a localizing symptom of disease of the organ of Corti. *Arch. Otolaryngol.* **31,** 160–184 (1940). "…sounds being heard a second later in the affected ear" (p. 166)

Urbantschitsch, V. *Lehrbuch der Ohrenheilkunde* (5th edn), 66–67 (Urban and Schwarzenberg, Berlin, 1910).

Appendix 3. Multiple tips in tuning curves

It is significant that multiple tips and notches are regularly seen in suppression tuning curves, appearing on the low-frequency or high-frequency slope (or both) depending, it is suggested, on whether the SOAE arises from L_0 or from an oblique resonator. These observations suggest that the suppressing tone is interacting with both the primary (near-perpendicular) resonator and its allied oblique resonators. The dynamics are such that the tone's entrainment of any one of these resonators is done at the expense (suppression) of the others. Such intimate linking suggests close physical coupling: one energetic resonator and its allied oblique counterparts.

The most detailed data on multiple resonators is given in the paper as Fig. 5. However, suppression data are also compatible with this picture.

Martin *et al.* (1988)* show suppression curves for three SOAEs in a monkey (their Fig. 7) and notches appear at frequencies 1.30 ±0.05 and 1.60 ±0.05 times lower than the highest notch, corresponding to the ratios of the L_2, L_3, and L_0 cavity lengths emanating from a geometry with an a/b ratio of 0.42. Measurement of this monkey's hair cell geometry (mapped in Fig. 7 of Lonsbury-Martin *et al.*, 1988†) for the region 54–59% from the apex — the 2–4 kHz region, wherein the highest notch appears — gives an a/b ratio of 0.38–0.41.

In the case of the 1413-Hz suppression curve (Fig. 7A), the two higher frequency notches correspond with observed SOAE frequencies.

One good illustration of multiple tips appears in Powers *et al.* (1995)‡ , in which the major dip of a chinchilla's suppression curve can be interpreted using the geometry of Fig. A-4a. (Although

* 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.

† Lonsbury-Martin, B.L., Martin, G.K., Probst, R., & Coats, A.C. 1988 Spontaneous otoacoustic emissions in a nonhuman primate. II. Cochlear anatomy. *Hearing Res.* **33,** 69–93.

‡ Powers, N.L., Salvi, R.J., Wang, J., Spongr, V. & Qui, C.X. 1995 Elevation of auditory thresholds by spontaneous cochlear oscillations. *Nature* **375,** 585–587.

containing some flexible parameters, this geometry is typical, and indicates an approach — adjusting the parameters of a/b and tilt — that can be taken in matching peaks in relative frequency.)

Thus, using this tilted geometry and taking the SOAE observed by Powers *et al.* to correspond to the L_{-1} resonator (relative frequency of 1.00), then each of the associated dips in the suppression curve (at relative frequencies of 1.03, 1/1.26, 1/1.40, 1/1.66, and 1/1.93) are good matches to the L_1, L_3, L_4, L_5, and L_6 resonator lengths (1/1.03, 1.24, 1.44, 1.66, 1.90) given in the last column of Table A-4.

Appendix 4. Elaboration of the cochlear geometry to include skew of the crystal lattice

Most micrographs examined appear to show an orthorhombic crystal lattice, with a hair cell in OHC 1 lining up with another in OHC 3, so that the line between them is at right angles to the rows. However, some examples show a slight tilt in the lattice to produce a rhombic arrangement. For example, one micrograph (Fig. 39 of Bredberg, 1968*) taken 17 mm from the base of a human cochlea (2 kHz region) shows an a/b ratio of 0.32 and a skew of 3–4°. This is small, but it has a measurable effect on the length of the resulting resonators. It is this set of lengths which are used in calculating the cochlear response in Fig. 4 of the main text.

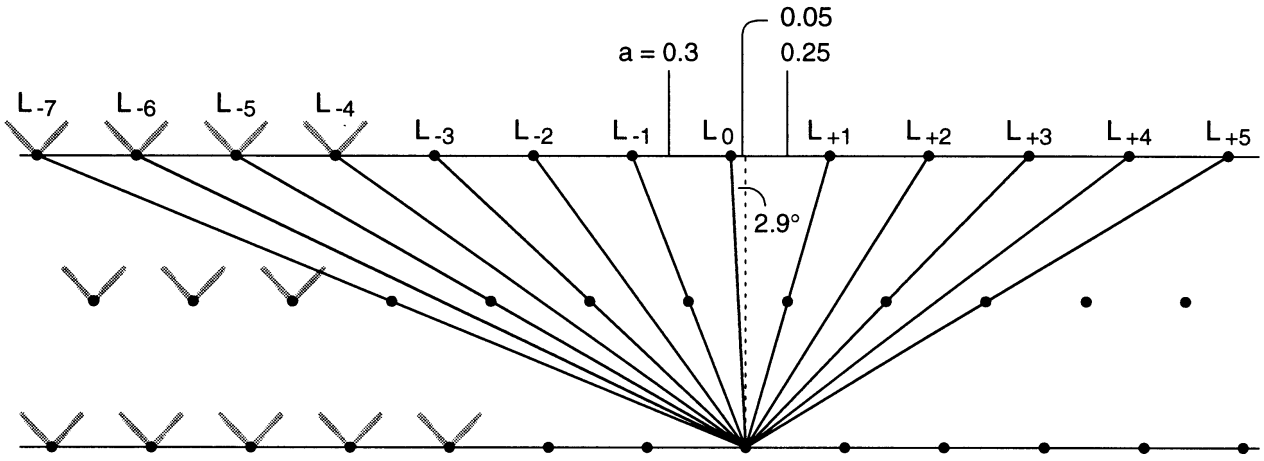


Fig. A-4a. Cochlear geometry where a/b is 0.3 and OHC row 1 is skewed 3° relative to OHC 3. This geometry generates a set of lengths L_1, L_2, L_3, …in one direction and L_0, L_{-1}, L_{-2}, L_{-3}, …in the other, and these are tabulated in Table A-4.

* Bredberg, G. 1968 Cellular pattern and nerve supply of the human organ of Corti. *Acta Otolaryngol.* Suppl. 236.

Table A-4. Theoretical resonant cavity lengths

cavity	length	length re L_0 = 1.001 [reciprocal length]	length re L_{-1} = 1.0595 [reciprocal length]
L_0	1.001	1.000	
L_{+1}	1.031	1.030 [0.97]	
L_{-1}	1.059	1.058 [**0.95**]	1.000
L_{+2}	1.141	1.140 [0.88]	1.077 [0.93]
L_{-2}	1.193	1.192 [0.84]	1.126 [0.89]
L_{+3}	1.312	1.311 [0.76]	1.239 [0.81]
L_{-3}	1.379	1.378 [0.73]	1.302 [**0.77**]
L_{+4}	1.524	1.522 [0.66]	1.439 [0.69]
L_{-4}	1.601	1.599 [0.63]	1.512 [0.66]
L_{+5}	1.761	1.759 [0.57]	1.663 [0.60]
L_{-5}	1.845	1.843 [0.54]	1.742 [0.57]
L_{+6}	2.016	2.014 [0.50]	1.904 [0.53]
L_{-6}	2.103	2.101 [0.48]	1.986 [0.50]
L_{+7}	2.281	2.279 [0.44]	2.154 [0.46]
L_{-7}	2.371	2.369 [0.42]	2.239 [0.45]
L_{+8}	2.554	2.551 [0.39]	2.412 [0.41]
L_{-8}	2.646	2.643 [0.38]	2.499 [0.40]
L_{+9}	2.832	2.830 [0.35]	2.672 [0.37]
L_{-9}	2.926	2.923 [0.34]	2.763 [0.36]
L_{+10}	3.115	3.112 [0.32]	2.939 [0.34]
L_{-10}	3.210	3.207 [**0.31**]	3.031 [0.33]
L_{+11}	3.400	3.397 [0.29]	3.208 [**0.31**]

CAPTION: Cavity lengths generated by the skewed geometry of Fig. A-4a (a = 0.3, b = 1, and tilt of 2.9°) are tabulated here. The cavities are ranked from shortest to longest and are labelled L_n as per the figure. Columns 3 and 4 give the lengths relative to cavity L_0 and L_{-1}, and the reciprocal of these lengths (related to frequency) is shown in square brackets. Figures in **bold** indicate matches to peaks in relative frequency shown in Fig. A-6a.

Appendix 5. Linked bistable emissions

The intrinsic linking of various perpendicular and oblique resonators has been put forward in this work as a possible explanation of linked bistable emissions, a phenomenon described in Burns *et al.* (1984)* in which two SOAEs are found to jump backwards and forwards between two fixed frequencies at an irregular rate. Most of the human data for this phenomenon involve switching at a ratio of about 1.06, as Table A-5a shows. This number is interpreted as the ratio of the lengths of the L_0 and L_1 cavities, and the widespread occurrence of a switching ratio of 1.06 follows naturally from the assumption that these two cavities are the ones most frequently energised.

The assumption is reasonable, given that the most common ratio between simultaneously occurring SOAEs is also about 1.06 (Table A-5b).

* Burns, E.M., Strickland, E.A., Tubis, A. & Jones, K. 1984 Interactions among spontaneous otoacoustic emissions. I. Distortion products and linked emissions. *Hearing Res.* **16,** 271–278.

Table A-5a. Linked bistable emissions

author	f_1 (Hz)	f_2 (Hz)	f_2/f_1	semits
Keefe et al. (1990)	1595.6	1701.8	1.0666	1.12
	1408.1	1524.1	1.0824	1.37
	1330.6	1410	1.0597	1.00
Wit (1990)	1612	1700	1.0546	0.92
Wilson et al. (1988)	3002 ±5	3233 ±5	1.077	1.28
Bell (unpublished)	2165.5 ±0.1	2295.6 ±0.1	1.0601	1.01
van Dijk et al. (1996) [barn owl]	8544	9018	1.055	0.93
Zurek and Clark (1981) [chinchilla]	4730	5680	1.20	3.2
Ohyama et al. (1991) [guinea pig]	1438	1489	1.0355	0.60

References: Keefe, D. H., Burns, E. M., Ling, R. & Laden, B. in *Mechanics and Biophysics of Hearing* (eds Dallos, P., Geisler, C. D., Matthews, J. W., Ruggero, M. A. & Steele, C. R.) 194–201 (Springer-Verlag, Berlin, 1990); Wit, H. P., op. cit. 259–268; Bell, unpublished; Wilson, J. P., Baker, R. J. & Whitehead, M. L. in *Basic Issues in Hearing* (eds Duifhuis, H., Horst, J. W. & Wit, H. P.) 80–87 (Academic Press, London); van Dijk, P., *J. Acoust. Soc. Am.* **100,** 2220–2227; Zurek, P. M. & Clark, W. W. *J. Acoust. Soc. Am.* **70,** 446–450; Ohyama, K., Wada, H., Kobayashi, T. & Takasaka, T. *Hearing Res.* **56,** 111–121 (1991).

Table A-5b

Most common interval in spacing of spontaneous otoacoustic emissions

author	peak spacing (equivalent)
Dallmayr (1985)	0.35–0.40 Bark (1.05–1.07)
Dallmayr (1987) [SFOAE]	0.35 Bark (1.05)
Zwicker & Peisl (1990) (SOAE, SEOAE, DEOAE)	0.3–0.5 Bark) (1.05–1.08))
Lind & Randa (1990)	0.3–0.4 Bark (1.05–1.07)
He & Schmiedt (1993) [DPOAE]	3/32 octave (1.067)
Engdahl & Kemp (1996) [DPOAE]	160 Hz @ 2.2 kHz – 290 Hz @ 4.6 kHz (1.06–1.07)
Talmadge et al. (1993)	0.4 mm (Bark) (1.06)
Braun (1997)	100 cent (1.059)
this work (from Russell (1992))	0.99 semitone (1.059)

Table A-5b caption: All authors agree that the peak spacing occurs very near 1.06, a value very close to an equal-tempered semitone. This ratio often appears in a different guise in the literature where it is expressed as 0.4 Bark. A Bark is a psychophysical unit (named after Barkhausen) designed to roughly correspond with a critical band on the partition (a distance of some 0.5–0.9 mm). According to Zwicker & Terhardt, Bark = 8.7 + 14.2 $\log_{10}(f_{kHz})$; therefore an interval of 0.4 Bark = 1.067.

References: C. Dallmayr, *Acustica* **59**, 67 (1985); C. Dallmayr, *Acustica* **63**, 243 (1987); E. Zwicker and W. Peisl, *Hearing Res.* **44**, 209 (1990); O. Lind and J. S. Randa, *J. Otolaryngol.* **19**, 252 (1990); N. He and R. A. Schmiedt, *J. Acoust. Soc. Am*. **94**, 2659 (1993); B. Engdahl and D. T. Kemp, *J. Acoust. Soc. Am*. **99**, 1573 (1996); C. L. Talmadge, G. R. Long, W. J. Murphy, A. Tubis, *Hearing Res*. **71**, 170 (1993); M. Braun, *Hearing Res.* **114,** 197 (1997); A. F. Russell, *Heritability of Spontaneous Otoacoustic Emissions,* PhD thesis, U. of Illinois, 1992.

The last line of the table derives from an analysis of the SOAE data of Russell (1992). When the ratios between neighbouring SOAEs are examined, the following distribution is found. The mean gap is 0.99 semitone.

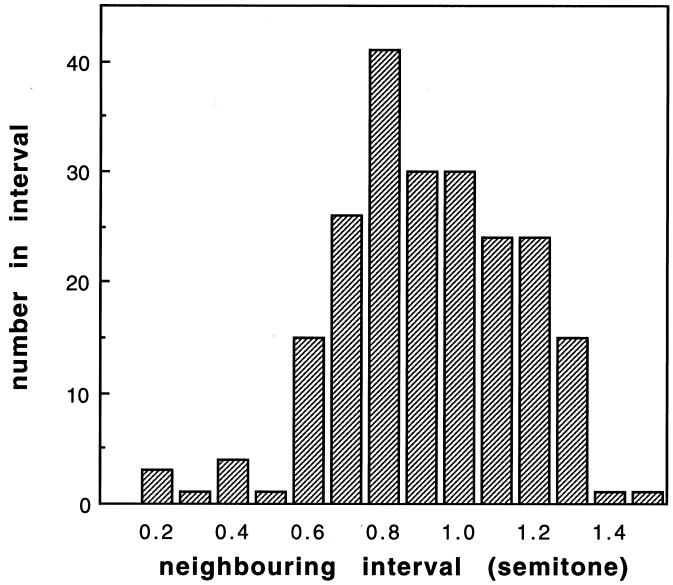


Fig. A-5a. Distribution of nearest neighbour ratios between simultaneous SOAEs taken from the data of Russell (1992).

When the same data set is plotted as a ranking of ratios, it is apparent that distinct gaps appear (Fig. A-5b). The principal gap, at about 1.057 ± 0.001, seems to be statistically significant ($p<0.05$). This ratio is remarkable in that it is close to the conjectured ratio of L_0 to L_{-1}. The gap could be interpreted to mean that it is not possible to have two connected cavities active at the same time and at the same spot on the cochlear partition — there must be a small gap along the partition from where the L_0 cavity is active to where the L_{-1} cavity is at work. Because of the slightly different tuning conditions at the two respective spots (according to the tuning map from base to apex), their frequency ratio will never be exactly the ratio of their lengths.

This interpretation meshes with the observation of linked bistable emissions: if there are two excited cavities at the same position on the partition, they will alternate rather than coexist.

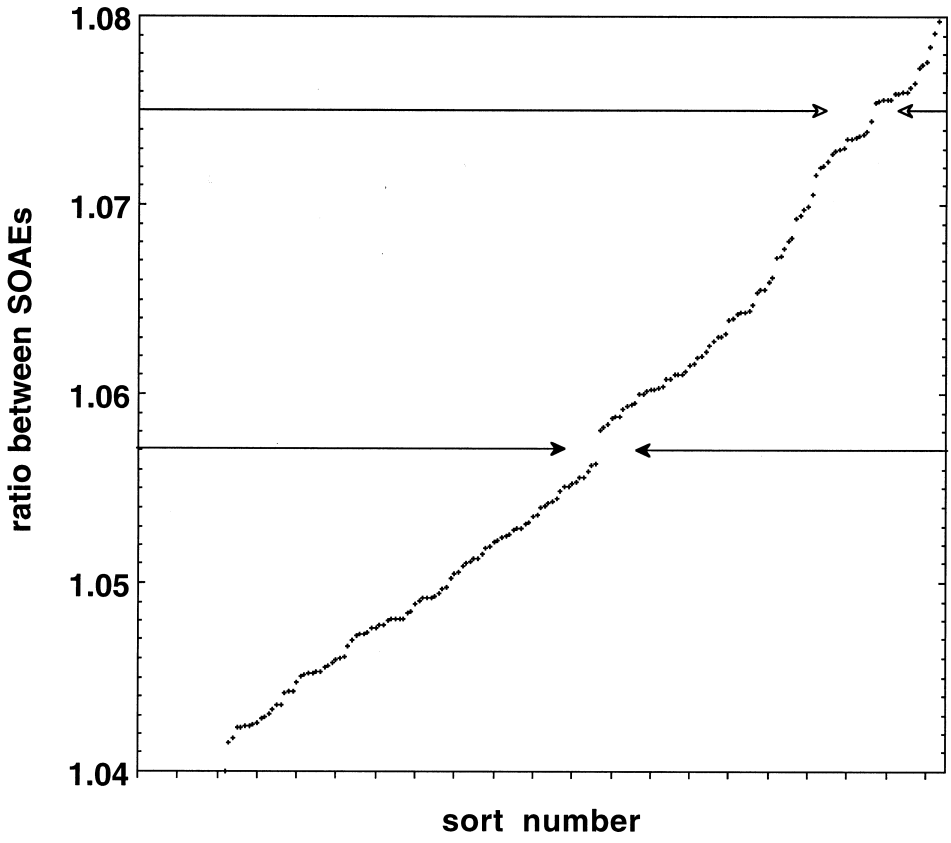


FIGURE A-5b. Frequency ratios between all neighbouring SOAEs given in Russell (35) are here plotted in ascending order. The distribution is mostly uniform, but a distinct gap occurs at a ratio of 1.057 ±0.001 (solid-headed arrows). Another possible gap occurs at 1.075 ±0.001 (open-headed arrows). These ratios match the ratio of cochlear cavity lengths based on the geometry of Fig. 5. The first represents the ratio of the L_0 and L_{-1} cavities; the second the ratio of L_{-1} and L_{-2}. The inference is that it is not possible to have two simultaneous cavities active at the same cochlear location, as discussed in the text. (Note that the accuracy of the SOAE ratios here is better than 1 in 1000).

Appendix 6. Other preferred ratios between SOAEs

This section investigates whether favoured spacings between SOAEs in a large data set can be matched to hair cell geometry. We have already seen in the previous section that the most common spacing of SOAE frequencies (0.94) correlates well with the lengths of two adjacent resonant cavities (1.06), and this process is extended to other, long-range, favoured SOAE ratios. A suggestive, although, given the available data, not definitive, correlation is obtained.

The data set used is that of Russell (1992)*. This extensive set lists 791 emission frequencies from 60 sets of twins, and includes 121 ears with 2 or more emissions. When these multiple emissions are expressed as ratios to the highest detected emission, they fall into a pattern (Fig. A-6a) showing a number of preferred ratios.

The highest detected frequency is presumed to be a short, near-perpendicular resonator, and other lower frequency emissions are assumed to be longer, oblique resonators.

Analysis of Russell's data confirms that the most favoured nearest-neighbour ratio is between 1.05 and 1.06, as shown in the figure above. The reciprocal of 1.055 (0.95) appears as a major peak.

However, other substantial peaks are seen. In particular, a strong peak is seen at a ratio of 0.31. Such a long-range peak has not been previously observed. Braun (1997)† looked for preferred spacings, but confined his analysis to less than 800 cents. He found the prominent semitone peak, and a number of secondary ones, but none of these smaller peaks were common to males and females, suggesting they could have occurred randomly. His method of analysis, taking all ratios, up and down in frequency, between all SOAEs, also tends to cloud the picture. The peak found here at 0.31 is statistically significant (2.7 standard deviations above trend

* Russell, A. F. 1992 *Heritability of Spontaneous Otoacoustic Emissions* (PhD thesis, U. of Illinois). Ann Arbor: UMI.

† Braun, M. 1997 Frequency spacing of multiple spontaneous otoacoustic emissions shows relation to critical bands: a large-scale cumulative study. *Hearing Res.* **114,** 197–203.

line). Another clear peak occurs at 0.77 (2.9 standard deviations), and may well be the same peak found by Braun (1997) at a ratio of 9:7.

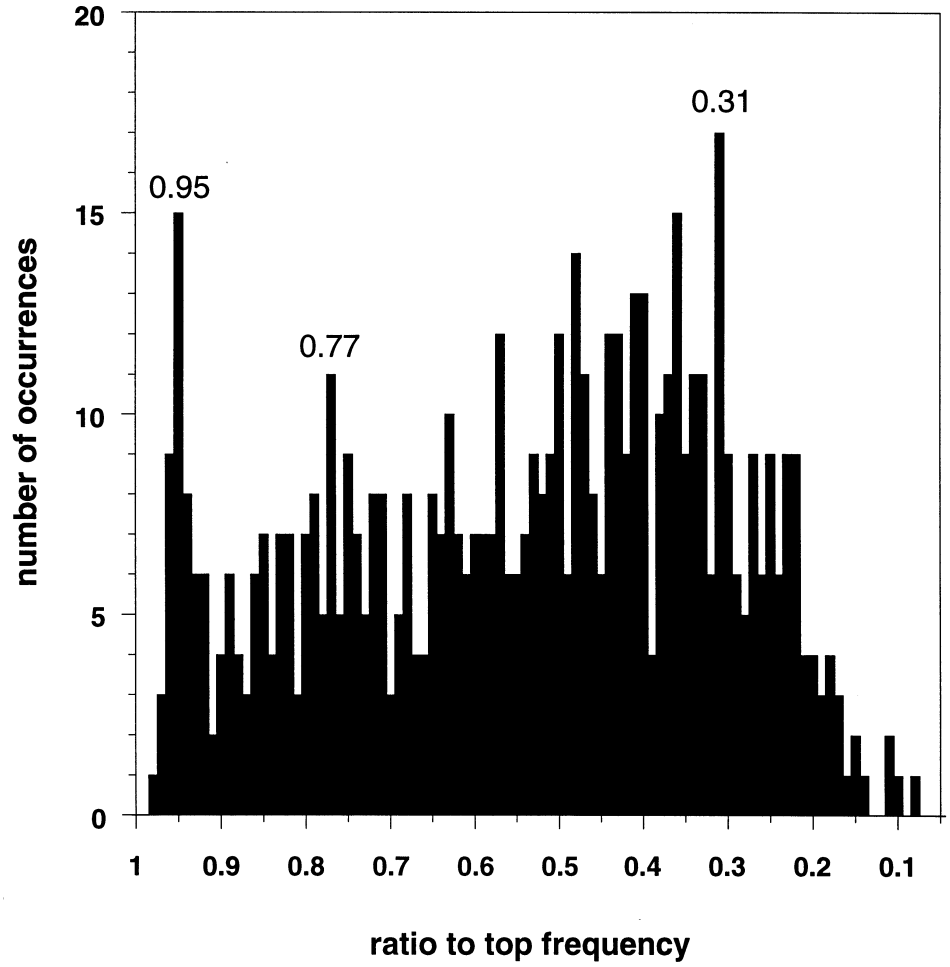


FIGURE A-6a. Distribution of frequencies of SOAEs in 121 ears that have at least 2 emissions expressed as a ratio to the highest frequency detected in each ear. A major peak occurs at 0.95 corresponding to the commonly found interval of 1.05–1.06. Two other prominent peaks occur at ratios of 0.77 and 0.31. Other peaks occur at 0.88, 0.63, 0.57, 0.50, 0.48, 0.44, 0.40, and 0.36. Not all these are statistically significant; nevertheless, the peaks at 0.77 and 0.31, previously unrecognised, appear to be real (see below). All these 11 peaks can be made to correspond to the inverse of the lengths of the cochlear cavities displayed in Fig. A-4a and tabulated in Table A-4. Data from Russell (1992).

Confirmation that the 0.77 and 0.31 peaks are real comes from looking at the phenomenon of 'mirroring' in which SOAEs in both left and right ears have the same frequency (Braun, 1998)*. Braun looked only at occurrences of mirroring at single frequencies, but examination of the Russell data set shows that it can occur at multiple frequencies simultaneously. A good example is a subject (Russell's DZF7A) where at least 5 SOAEs occur in each ear at frequencies that differ by less than 1% (see following table).

Table A-6

Mirroring in subject DZF7A of Russell (1992)

SOAEs in left ear (Hz)	relative frequency to left	presumed cavity	relative frequency to right	SOAEs in right ear (Hz)
3959				
3383	**1.0000**	L_{-1}	**1.0000**	3374
			0.9795	3305
3175	0.9385	L_{+2}	0.9395	3170
2703	0.7990	L_{+3}	0.8139	2746
2603	**0.7694**	L_{-3}	**0.7712**	2602
2434	0.7195			
2263	0.6689			
1613	0.4768			
		L_{-7}	0.4446	1500
1415	0.4183			
		L_{-8}	0.4034	1361
1287	0.3804	L_{+9} (?)	0.3800	1282
1216	0.3594	L_{-9}		
1145	0.3385	L_{+10}		
1072	**0.3169**	L_{+11}	**0.3154**	1064
etc.	… …			

Given that there are less than six SOAEs per octave (average interval of about 200 cents) in this subject, it is clear that the 5 instances of mirroring (accurate to better than 13 cents) did not occur by chance ($p<0.05$).

* Braun, M. Accurate binaural mirroring of spontaneous otoacoustic emissions suggests influence of time-locking in medial efferents. *Hearing Res.* **118,** 129–138 (1998).

Expressed as ratios to the top mirroring frequency, the mirroring in this subject occurs at 1.000, 0.939, 0.769, 0.380, 0.317, and 0.172 in the left ear and 1.000, 0.940, 0.771, 0.380, 0.315, and 0.173 in the right. Moreover, when all SOAEs in which frequencies in left and right ears match to better than 1.5% are taken together and expressed as ratios to the top matching frequency, the following figure results.

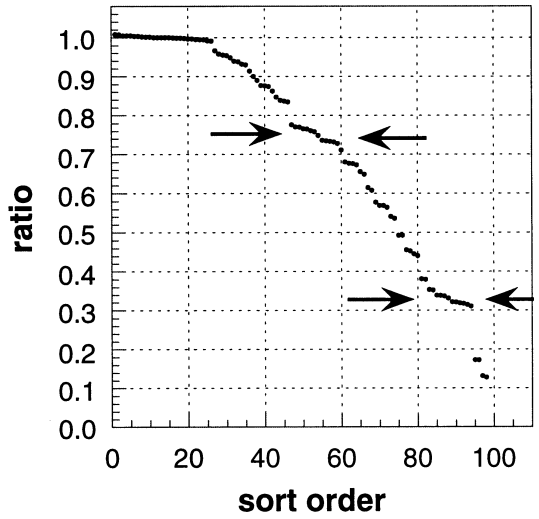


Fig. A-6b. All examples of mirroring in Russell's data set (involving at least 2 pairs of matching frequencies) are calculated as ratios to the top frequency and ranked in size. A tendency for ratios near 0.75 ± 0.02 and 0.32 ± 0.01 (arrows) is evident.

The tendency for ratios to appear near 0.75 and 0.32 suggests that the peaks found at 0.77 and 0.31 in Fig. A-6a have a physical basis.

A further demonstration of a preferred ratio at 0.77 comes from looking at how frequently it recurs in some ears. Thus, the ratio 0.77±0.01 appears 7 times among the 23 SOAEs of the left ear of subject DZF7A; 6 times among the 17 SOAEs in the right ear of MZF13A; and 6 times among the 17 SOAEs in the left ear of MZF13B. These rates are about 4 times higher than expected by chance.

It is possible to see these favoured ratios as further instances of resonant cavities, but at longer inter-hair cell distances than the common 1.06 ratio. Thus, the cochlea-like geometry shown in Fig. A-4a produces cavity lengths (Table A-4) which are the inverse of the frequencies of SOAE peaks at 0.95, 0.77, 0.31, and others. This geometry is based on a 17° orthorhombic lattice (a/b

= 0.30) tilted 3° as shown. The resulting ratio between the shortest cavity and its neighbour is 1.058, close to the required 1.06, and the tilt provides a match between the inverse length of the other cavities and the main peaks in frequency ratios shown in Fig. A-6a. (Indeed, this geometry also accounts for the other 8 less statistically significant peaks.) The presumed geometry closely accords with a micrograph (Fig. 39 of Bredberg 1968*) taken 17 mm from the base of a human cochlea (2-kHz region) in which the measured a/b is 0.32 and the skew is 3–4°. The data of Russell used here predominantly have top frequencies in the 2–5kHz band.

This correspondence strengthens the suggestion that the cavities are oscillating like plucked strings. The interpretation is that generation of SOAEs at one location on the partition, by the oscillation process described, often creates a number of other reverberating cavities at or near the same spot. This generating mechanism can occur at a number of unrelated active sites, and this is why the baseline in Fig. A-6a tends upwards at lower frequencies, reflecting the occurrence at places nearer the cochlear apex of further short-cavity emissions unrelated to the original top frequency.

Table A-4 indicates that many cavities are possible, but Fig. A-6a suggests that only certain ones are favoured — perhaps those that are picked out by the particular orientation of the stereocilia arms. Table A-4 shows how a number of preferred ratios in SOAE frequency could arise as ratios of the specified cavity lengths to that of the shortest cavity (L_0). There is presently insufficient data to show clear preferred ratios other than at 0.95, 0.77, and 0.31, but it suggests that the L_{-1} and two or more longer cavities are prime candidates in generating SOAEs. The hypothesis therefore predicts that stereocilia arms are frequently found at angles defining these cavities (i.e., arms in rows 1 and 3 should be parallel to each other and perpendicular to these cavities) which would occur when cavity angles lie at about 19°, 44°, and 72° from the transverse direction.

* Bredberg, G. 1968 Cellular pattern and nerve supply of the human organ of Corti. *Acta Otolaryngol.* Suppl. 236.

Appendix 7 Sketch for a revised cochlear mechanics

The essential element in this new formulation of the resonance theory of hearing is to understand that OHCs are compressible elements (Zenner *et al.*, 1992)* immersed in virtually incompressible fluid, and so the first step in transduction involves acoustic energy being converted into oscillatory energy of the OHC's cytoskeletal spring (Holley & Ashmore, 1988)†.

In substantiating this picture, the following points are noteworthy.

- Small static pressures (2–5 Pa) applied to an isolated OHC produced 10–20% shortening of the cell (Ashmore 1987)‡.

- Using a fluid jet directed at the body of the OHC, Brundin & Russell (1993)§ detected mechanically induced length changes to pressures of about 1 mPa. They conclude that OHCs in the cochlea may change length in direct response to natural sound stimulation.

- Isolated OHCs have been observed to respond directly to sound in a tuned manner (Brundin *et al.*, 1989)**.

- In other studies of isolated OHCs (Santos-Sacchi & Dilger 1988)††, the mechanical response of cells lacking stereocilia was indistinguishable from those with sterocilia when both were electrically stimulated, suggesting that stereocilia deflection may not be the primary stimulus in hearing.

---

* Zenner, H.P., Gitter, A.H., Rudert, M. & Ernst, A. 1992 Stiffness, compliance, elasticity and force generation of outer hair cells. *Acta Otolaryngol.* **112,** 248–253.

† Holley, M.C. & Ashmore, J.F. A cytoskeletal spring in cochlear outer hair cells. *Nature* **335,** 635–637 (1988).

‡ Ashmore 1987 A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J. Physiol.* **388,** 323–347.

§ Brundin, L. & Russell, I. 1993 Sound-induced movements and frequency tuning in outer hair cells isolated from the guinea pig cochlea. In *Biophysics of Hair Cell Sensory Systems*, ed. H. Duifhuis et al. Singapore: World Scientific.

** Brundin, L., Flock, Å., Canlon, B. 1989 Sound-induced motility of isolated cochlear outer hair cells is frequency-specific. *Nature* **342,** 814–816.

†† Santos-Sacchi, J. & Dilger, J. P. 1988 Whole cell currents and mechanical responses of isolated outer hair cells. *Hearing Res.* **35,** 143–150.

- A model of cochlear mechanics has already been proposed (Kolton et al. 1989)* in which "excitation for the OHC comes from movement of its body rather than from bending of its stereocilia", with motile force generation occuring "through the stereocilia, since this would keep the input and the output more effectively isolated."

- Both stereocilia length and OHC length are graded from one end of the cochlea to the other (Pujol *et al.,* 1992)†, allowing them to be directly (but broadly tuned) to incoming sound energy.

- Finally, OHC are surrounded by fluid: the spaces of Nuel. Why else would these cells, unlike virtually every other cell in the body, be in contact with cochlear fluids (not neighbouring cells) except that it enables them to react directly (by compression) to hydraulic pressure oscillations imposed by the stapes?

The primary stimulus is therefore taken to be a compressional wave which acts directly on the OHC, causing them to shorten and lengthen in synchrony with the pressure. However, because of mechanomotility, this causes their stereocilia to deflect backwards and forwards, generating ripples on the tectorial membrane which are amplified by the feedback process operating between neighbouring cells. This process is the same as that by which a regenerative receiver works.

The end result of all the OHCs responding in this way is whole-scale movement of the cochlear partition — the observed traveling wave. The traveling wave is in fact the graded delay in a bank of resonators. However, the important difference between this picture and traveling wave theory as usually conceived is that it is the resonance of the tuned OHC elements that is the primary response: only later does this movement initiate up and down movement of the partition, and it is to be expected that, for a given place on the partition, the OHC tuning tip will differ from

* Kolston, P. J., Viergever, M. A., de Boer, E. & Diependaal, R. J. 1989 Realistic mechanical tuning in a micromechanical cochlear model. *J. Acoust. Soc. Am.* **86,** 133–140.
† Pujol, R., Lenoir, M., Ladrech, S., Tribillac, F. & Rebillard, G. 1992 Correlation between the length of outer hair cells and the frequency coding of the cochlea. *Adv. Biosci.* **83,** 45–52.

tuning derived from the mass and compliance of the basilar membrane fibres.

Indeed, this difference could well explain the mysterious 'half-octave shift' in which a loud sound creates a temporary hearing loss, not at its own frequency but half an octave higher. This could be explained if the (broad) basilar membrane tuning of a place on the partition were approximately half an octave higher than the tuning of the OHC elements at this place. Thus, it is excessive movement of the partition that desensitises the OHC tuning elements at that location, not the high sound energy impinging directly on the OHC cells (as we have seen, the OHC are part of a regenerative receiver that can turn down its gain when high sensitivity is not required).

Allen and Fahey (1993)* suggested, on the basis of all distortion products consistently appearing strongest half an octave below CF, that there was a second frequency–place map in the cochlea. They attributed it to a mass–compliance resonance involving the outer hair cell stereocilia and fluid below the TM and drew a correlation between the tip and tail of the neural tuning curve. These authors fitted a power-law to their data, but for the limited data shown for humans, a straight line gives an equally good fit. Significantly, its slope is 0.69, which is close to half an octave. We may therefore be seeing in the half-octave shift a manifestation of the dual cochlear tunings: the primary OHC resonator tuning and the secondary cochlear partition tuning.

There are thus two conceivable ways by which the partition at one location can begin to move. One is in response to activation of the OHC tuning elements in response to sound (a 'jiggling' of the surface that gets transmitted to the body of the partition — not an efficient process, particularly because of the half-octave disparity). The second is that the partition can respond weakly to sound in the manner in which Bekesy describes. The partition movement corresponds to the broad tail of the tuning curve; this is a passive process that can be observed in the cochleas of dead animals. By

* Allen, J.B. & Fahey, P.F. 1993 A second cochlear-frequency map that correlates distortion product and neural tuning measurements. *J. Acoust. Soc. Am.* **94,** 809–816.

contrast, the active process derives from the regenerative receiver of the OHC elements, and produces the tuning tip (Davis, 1983)*.

The cochlear partition responds because the tectorial membrane presents a barrier to acoustical energy. The tectorial membrane is a gel, substances that generally possess very low Young's moduli. We therefore expect the velocity of sound propagation through the TM (by a standard compressional wave) to be appreciably lower than in the surrounding watery fluids of the cochlea. The corollary is that the TM presents an acoustical barrier to the sound which is propagating at about 1 km/sec from the oval window, through the cochlear ducts, to the round window. Its acoustical resistance is high, unlike the watery tissues of the basilar membrane, through which sound passes with virtually no interaction at all. In other words, the TM has the intrinsic property of impeding sound transmission across the partition, resulting in the creation of appreciable acoustical forces. The result is that acoustical energy, at high levels, can initiate some movement of the partition.

However, it is important to see why this is a weaker response than direct compression of the OHCs: because of short-circuiting by the helocotrema. This diminishes the pressure response according to the helicotrema's conductance, making the volume velocity of the scala fluids the controling factor. As Naftalin (1964)† points out, relying on translational movement of fluid in scala vestibuli to create stimulation of the hair cells must be an inefficient process when stapes movements of the order of atomic dimensions are involved.

> The theory of travelling waves in the basilar membrane is based on this assumption [that the fluid in the scala vestibuli undergoes a translational movement], since to create the travelling wave the perilymph has to be moved. Von Bekesy himself did not consider that his observations decided the path of energy flow in the cochlea as Davis (1957) points out in his review of the subject, but with the theory of resonance out of favour, the travelling wave theory has been made to carry the transfer from acoustic wave energy to mechanical translation of the structures on the basilar membrane and hence the shearing action of the tectorial membrane on the hair

* Davis, H. 1983 An active process in cochlear mechanics. *Hearing Res.* **9,** 79–90.
† Naftalin, L. 1963 The transmission of acoustic energy from air to the receptor organ in the cochlea. *Life Sciences* **2,** 101–106.

> processes. It may be reiterated, however, that where the travelling wave theory may very well describe the distribution of resonating elements, it can no more account for the quantitative transference of acoustic energy than does the classical resonance theory. (Naftalin, ibid., p. 105)

The traveling wave — a secondary phenomenon

At high sound pressure levels, acoustic pressure effects and viscous coupling between adjacent TM oscillators will move the partition and create the familiar traveling wave which Bekesy first saw. Although he thought the basilar membrane was responsible for the wave, he actually made his observations on Reissner's membrane, so any part of the organ of Corti could have been the causal agent. However, it is important to realise that the traveling wave is secondary: it is either the movement of the partition in response ot high sound pressure, or conceivably the response to the combined action of many individual high-*Q* oscillators. On the resonance picture, the traveling wave is not the primary causal mechanism. The TW is, from one viewpoint, merely a description of how the amplitude envelope of the bank of oscillators appears — an epiphenomenon, with no causal efficacy, but describing the phase relations between the high-frequency oscillators that quickly respond to the stimulus and the low-frequency ones that require more time to build up amplitude.

The *Q* of an oscillator can be expressed as the number of cycles required to set it into full-amplitude motion or, alternatively, for it to die away. Indeed, it was Pumphrey and Gold's incisive experiments in 1948 on the 'phase memory' of the ear that first gave conclusive evidence that the ear possessed some type of high-*Q* (>100) oscillator. Spontaneous emissions often show bandwidths of less than 1 Hz, giving an equivalent *Q* of 1000 or more. Helmholtz, too, recognised the fundamental incompatibility between a device that has very good frequency-resolving power (and hence high *Q*) and at the same time good temporal resolution. He investigated how fast notes could be played on the piano without appearing blurred, and with a 'shake' of about 10 per second calculated that the width of the resonators must be no more than a semitone (6%, or *Q* of at least 17).

The ear overcomes the paradox by having it both ways: a set of IHC that are heavily damped working in concert with a set of highly tuned OHCs that can ring like a bell. Since OHCs are supplied with an abundance of efferent nerve endings, gain control is probably very tight, allowing the information-retrieval capabilities of the system to be optimised.

Of course, at high SPL, above about 80 dB, the oscillators couple their energy and whole-scale movement of the partition begins. Here we do need to consider viscous coupling between adjacent resonators, and recognise direct excitation of the IHCs. But, to reiterate, where there is no coupling there is no energy in the travelling wave.

The important point, often forgotten, is the primacy of the acoustical pressure across the partition as the causal agent. Bekesy himself acknowledged this, pointing out that the TW was simply a descriptive shorthand of what was happening[*], but it needs restating. And so we notice that SOAEs respond virtually instantaneously to a suppressing tone. Zero response times have also been seen by Wilson (1980)[†] and Brown & Kemp (1985)[‡]. One cannot have an SOAE at zero SPL linked to a traveling wave mechanism, although attempts have been made to do so by calling for a standing wave between the generation site and the stapes end of the partition.

Cogent and long-standing arguments against the BM being the frequency-resolving component in the ear still stand. In brief, they are that it is difficult for the BM fibres to possess the required range in mass and compliance; a number of animals, including frogs and lizards, possess a TM (and display spontaneous emissions) but have no BM; and even in some humans, the organ of Corti at certain places rests upon bone, not the BM. Some birds have been observed to normally have a perforation in the basilar membrane near the stapes, yet their hearing is apparently not

[*] Wever, E.G., Lawrence, M. & von Bekesy, G. 1954 A note on recent developments in auditory theory. *Proc. Nat. Acad. Sciences* **40,** 508–512.
[†] Wilson, J.P. 1980 Model for cochlear echoes and tinnitus based on an observed electrical correlate. *Hearing Res.* **2,** 527–532.
[‡] Fig. 2 of Brown, A.M. & Kemp, D.T. 1985 Intermodulation distortion in the cochlea: could basal vibration be the major cause of round window CM distortion.

affected. Braun (1996)* puts forward the idea that the BM is there as an energy-absorbing structure and called into play at high sound pressure levels. Dancer and Franke (1989)† have also made limited attempts to reintroduce a resonance theory based on the tuned response of individual outer hair cells.

* Braun, M. 1996 Impediment of basilar membrane motion reduces overload protection but not threshold sensitivity: evidence from clinical and experimental hydrops. *Hearing Res.* **97,** 1–10.

† Dancer, A. ***** Franke, R. 1989 Mechanics in a "passive" cochlea: travelling wave or resonance? *Il Valsalva* **54,** suppl. 1, 1–5.

Appendix 8. Points of clarification

1. Although the cavities have been schematised as originating from a single cell in OHC 3, this is for illustrative convenience, and it is likely that the excited cavities are disjoint, and are scattered over a circumscribed length of the partition. Equally, the focus of the cavities may just as easily be in OHC1 as much as in OHC3.

2. In situations where additional gain might well be required, such as at low frequencies (at the apex), four or more rows are often found (see Fig. A-9 here, also Bredberg, 1968). Constant inter-row spacing is retained, and the fourth-row elements appear at positions to augment the dominant left- and right-facing cavities. The resonant cavities are therefore three half-wavelengths long instead of two, creating the same resonant frequency but with additional gain. Some animals, such as the echidna, have 5 or more rows (Pickles, 1992). Multiple rows are found in birds, frogs, and lizards, and a similar resonating mechanism is proposed involving rows of alternating polarity.

3. Surface acoustic wave devices can also be configured as delay lines, convolvers, filters, and frequency analysers (Maines & Paige, 1976).

4. As well as OHC having stereocilia arms purposely arranged at right angles to the required oblique laser cavities, the TM is permeated by fibres which have been observed to run at an angle of about 30° from the radial (towards the apex) (Steel 1983; Morisaki *et al.,* 1991). This would facilitate the strong L_2 mode at this angle, channeling energy in this direction rather than allowing it to spread omnidirectionally.

5. Afferent nerves connect to OHCs in a generally oblique manner, beginning at the outermost row and taking an inward course, slanting towards the apex, until they reach the modiolus some millimetres away (Spoendlin, 1986)**.** Such an arrangement is ideally suited to pick up correlated activity in oblique resonators (as well as presumed orthogonal connections).

6. When maps of hair cell positions are examined, hair cells occupying the fourth row almost invariably augment oblique alignments (even though the fourth row itself may appear irregular in the longitudinal direction) (Lonsbury-Martin *et al.,* 1988). Moreover, when the positions of clusters of missing (or sparsely remaining) hair cells are mapped, they preferentially define oblique alignments (e.g., Fig. 40 of Bredberg, 1968).

7. Ripples impressed permanently into the tectorial membrane by acoustic overstimulation have been observed (Fig. 6 of Morisaki *et al.* 1991). After they delivered a 137-dB report from a starting pistol to a guinea pig, its TM showed permanent wave-like impressions. The ripples appear to originate from OHCs and travel in the direction of the TM fibres. However, their wavelength is much less than a cavity length, indicating that, for overstimulation, unusual behaviour has resulted. For normal-intensity stimuli, however, involvement of the fundamental mode is expected and no residual deformation of the TM would take place.

8. The TM can be seen as a key set of 'crystal oscillators' in the system. In the same way as a crystal-controlled oscillator relies largely on the physical properties of the crystal to set the oscillation frequency, so the OHC/tectorial membrane oscillator relies in large measure on the physical properties of the TM, not on the physiological properties of the OHCs. Indeed, SOAE frequencies are remarkably stable, remaining virtually constant even after large doses of aspirin sufficient to abolish emissions (McFadden & Plattsmier, 1984); similarly, stimulus-frequency emissions do not change frequency despite aspirin administration (Brown *et al.,* 1993). Factors that can affect frequency include intra-cochlear pressure (Bell, 1992), temperature (O'Brien, 1994), and the presence of nearby tones (Rabinowitz & Widin, 1984), which causes small amounts of frequency 'pushing' and 'pulling'. Pressure and temperature might be expected to slightly change the stiffness of the TM and affect wave transmission speed, in the same way as SAW resonators respond to these variables (Reeder & Cullen, 1976).

9. In the new picture, a traveling wave is no longer needed to account for the observed delay in evoked emissions. After incoming sound energy excites the OHC resonators, the oscillation travels across the TM and is reflected off the sharp edge of the inner spiral sulcus before reentering the resonant cavity, where it creates an echo. The echo can then be considered to appear almost instantly at the ear canal by virtue of a fast pressure wave traveling through the cochlear fluids. This idea relies on a small change in volume of outer hair cells upon excitation (Wilson, 1980) and it simplifies the mechanics of the situation considerably. We noted earlier that anomalous 'zero delays' have been observed in a number of evoked emission experiments.

The delay-line picture allows us to view stimulus-frequency emissions as a case of entrainment to an external tone by matching cochlear resonators. The signal observed in the ear canal will be the sum of all the contributing resonators, which will include not only the orthogonal at its characteristic place but also oblique resonators at other locations (more basally) that match the frequency of the incoming tone. Clearly, the phases between all these contributors will vary, and so the summed response detectable in the ear canal will be complex (although stable and repeatable), as observed.

10. In a similar fashion, distortion product emissions may be considered as the interaction of two stimulus-frequency emissions at f1 and f2 where the detection frequency is set at, typically, 2f1–f2. For reasons not perfectly clear, distortion reaches a maximum for f2/f1 ratios of 1.22–1.25 (Harris et al., 1989; Gaskill & Brown, 1990). A significant clue to what is going on, however, comes from noting that the trace of distortion against frequency possesses stable fine structure (Gaskill & Brown, 1990), and that particularly sharp notches occur at certain favoured ratios (Harris *et al.*, 1989). These ratios depend on the subject, but it is significant that values close to those found in Table 3 recur in the results of Harris *et al.*, 1989: 1.03, 1.08, 1.14, and 1.19.

In other words, distortion is at a minimum when two tones can excite two coupled sympathetic resonators in the cochlea — one the strong L_0 or L_1 resonator, the other one of its allied resonators.

In the same way, distortion between an SOAE and an external tone shows prominent peaks and dips, with a distinct minimum in one particular instance at a frequency 1/1.39 times (c.f. the L_{-3} length of 1.38) the SOAE (Norrix & Glattke 1996). Conversely, where such matching resonators do not exist at a single location on the partition, it is supposed that the two tones must force slightly unmatched resonators into oscillation, creating distortion.

Note that the greatest opportunity for mismatch comes where there is the biggest frequency gap between adjacent resonators — that is, between the L_{-3}/L_4 resonators, where there is a maximum gap of 1.105 (Table 3). In fact, this region (1.38–1.52 times L_0) has been variously identified as the locus of 'a second filter' about half an octave (1/1.41) below the higher frequency (Brown & Williams, 1993) or the place of a second frequency–place map (Allen & Fahey, 1993). Not only is it the site of maximum 2f1–f2 distortion, but of all other (less prominent) higher-order distortion products as well (Allen & Fahey, 1993). The human data given in Fig. 3 of Allen and Fahey (1993), showing the frequency of maximum distortion plotted against the upper frequency generating the distortion, can be fitted by a straight line of slope 0.69, equivalent to a resonator length of 1.45, close to the geometric mean of the L_{-3} and L_4 lengths.

General conclusions

This work began by uncovering a pattern in spontaneous otoacoustic emissions and relating this to the regular geometry of outer hair cells. Such a matching requires slow wave propagation in the tectorial membrane and an active process in which stereocilia are constantly in motion, creating and responding to an unceasing criss-crossing of rippling wavefronts. Like a laser cavity in an omnidirectional field, this dynamic process picks out the most periodic wavefront and amplifies it.

This conjecture calls for a major shift in our thinking about cochlear mechanics, transferring our focus away from the basilar membrane and its traveling waves to the tectorial membrane and its sympathetic resonators. The acoustic laser means the traveling wave and its complex dynamics (Patuzzi, 1996) is no longer

needed as the primary causal mechanism in cochlear mechanics; on this alternative view, a traveling wave can be considered an epiphenomenon, a reflection of graded phase delays in a bank of resonators responding directly to intracochlear pressure fluctuations (Wever et al., 1954; p. 188 of Patuzzi, 1996; this author, in prep.). The basilar membrane, then, may be just an energy absorber, not a frequency analyser (Braun, 1996).

While simplifying matters, this new picture also provides greater explanatory power. It is now possible to explain many auditory phenomena that were obscure on the traveling wave interpretation. Some have been discussed here, although space precludes a detailed exposition. But just on a basic level, all the following can be consistently explained: spontaneous, evoked, and distortion-product emissions; diplacusis echotica, favoured ratios between SOAEs; the shape of the ear's auditory filter; its essential non-linearity; the basis of music, and the geometry, disposition, and function of the major organ of Corti elements.

It is interesting to note that not long after Helmholtz formulated his theory, Hasse in 1867 (see Wever, 1949) proposed the TM as the resonator, followed by Shambaugh in 1907, and Naftalin in 1964 (see Naftalin, 1977). Shambaugh (1907) provides a compelling argument in favour of the tectorial membrane as the ear's resonating element. His detailed drawings and description invoke a picture of "an immense number of delicate lamellae", akin to a soft feather, which together form the basis for a series of resonators. Hardesty's fine illustrations showing organised fibre directions (Hardesty, 1908) provide tantalising clues to conceivable propagation modes, and his description of strong surface tension adds to the possibilities. Why the tectorial membrane ever lost its central role to the basilar membrane (which Hardesty [1908] likened to a board in comparison to the silk-like tectorial membrane) makes an interesting question.

Over the last decade or so, a number of researchers (Zwislocki & Kletsky, 1979; Brown *et al.*, 1992; de Boer, 1993; Markin & Hudspeth, 1995) have called on resonance between TM and OHC elements to generate secondary resonance and sharper tuning. The idea has recently been put forward that, in lizards, hair cells

continuously drive side-to-side movement of discrete blobs of tectorial membrane and cause spontaneous emissions (Manley *et al.,* 1996).

Now we return once again to the neglected tectorial membrane and find on its surface delicate ripples, the strings of a remarkable underwater piano.

ADDITIONAL REFERENCES

Allen, J. B. and P. F. Fahey, *J. Acoust. Soc. Am.* **94**, 809 (1993).

Bell, A. *Hearing Res.* **58,** 91 (1992).

Bell, D.T. and R. C. M. Li, *Proc. IEEE* **64,** 711 (1976).

Braun, M. *Hearing Res.* **97**, 1 (1996).

Bredberg, G. *Acta Otolaryngol.* Suppl. **236** (1968).

Brown, A. M., D. M. Williams, S. A. Gaskill, *J. Acoust. Soc. Am.* **93,** 3298 (1993).

Brown, A. M., S. A. Gaskill, D. M. Williams, *Proc. R. Soc. Lond. B* **250**, 29 (1992).

Brown, A. M. and D. M. Williams, in *Biophysics of Hair Cell Sensory Systems*, H. Duifhuis, J. W. Horst, P. van Dijk, S. M. van Netten, Eds, (World Scientific, Singapore, 1993), pp.72–77.

Brundin, L., Å. Flock, B. Canlon, *Nature* **342,** 814 (1989).

Burns, E. M., Campbell, S. L. & Arehart, K. H. *J. Acoust. Soc. Am.* **95,** 385–394 (1994).

Dallos, P. and B. N. Evans, *Science* **267,** 2006 (1995).

de Boer, E. *J. Acoust. Soc. Am.* **93**, 2845 (1993).

de Boer, E. *Phys. Rep.* **105**, 141 (1984).

Fletcher, N.H. *Acoustic Systems in Biology* (Oxford University Press, New York, 1992), pp. 70–109.

Gaskill, S. A. and A. M. Brown, *J. Acoust. Soc. Am.* **88,** 821 (1990).

Hardesty, I. *Am. J. Anat.* **8,** 109 (1908).

Harris, F. P., B. L. Lonsbury-Martin, B. B. Stagner, A. C. Coats, G. K. Martin, *J. Acoust. Soc. Am.* **85,** 220 (1989).

Hartmann, W. M. *J. Acoust. Soc. Am.* **93,** 3400 (1993).

Kemp, D. T. in *Tinnitus*, D. Evered and G. Lawrenson, Eds, (Pitman, London, 1981), pp. 54–81.

Kössl, M. & G. Frank, in *Advances in Hearing Research*, G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, H. Oeckinghaus, Eds, (World Scientific, Singapore, 1995), pp. 125–135.

Maines, J. D. and E. G. S. Paige, *Proc. IEEE* **64,** 639 (1976).

Manley, G. A., L. Gallo, C. Köppl *J. Acoust. Soc. Am*. **99**, 1588 (1996).

Markin, V. S. and A. J. Hudspeth, *Biophys. J.* **69**, 138 (1995).

Martin, G. K., B. L. Lonsbury-Martin, R. Probst, A. C. Coats, *Hearing Res.* **33,** 49 (1988).

McFadden, D. and H. S. Plattsmier, *J. Acoust. Soc. Am.* **76,** 443 (1984).

Morisaki, N., Y. Nakai, H. Cho, S. Shibata, *Acta Otolaryngol.* Suppl. **486,** 19 (1991).

Naftalin, L. *Physiol. Chem. Physics* **9**, 337 (1977).

Norrix, L. W. and T. J. Glattke, *J. Acoust. Soc. Am.* **100**, 945 (1996).

O'Brien, A. J. *Br. J. Audiol.* **28,** 281 (1994).

O'Mahoney, C. F. and D. T. Kemp, *J. Acoust. Soc. Am.* **97,** 3721 (1995).

Patuzzi, R. in *The Cochlea*, P. Dallos, A. N. Popper, R. R. Fay, Eds, (Springer, New York, 1996), pp. 186–257.

Pickles, J.O. *Adv. Biosci.* **83,** 101 (1992).

Pujol, R., M. Lenoir, S. Ladrech, F. Tribillac, G. Rebillard, *Adv. Biosci.* **83,** 45 (1992).

Rabinowitz, W. M. and G. P. Widin, *J. Acoust. Soc. Am.* **76,** 1713 (1984).

Reeder T. M. and D. E. Cullen, *Proc. IEEE* **64,** 754 (1976).

Shambaugh, G. E. *Am. J. Anat.* **7,** 245 (1907).

Smith, W. R. in *Physical Acoustics: Principles and Methods, vol. 15*, W. P. Mason and R. N. Thurston, Eds, (Academic Press, New York, 1981), pp. 99–189.

Spoendlin, H. *Scand. Audiol.* Suppl. **25,** 27 (1986).

Steel, K. P. 1983 *Hearing Res.* **9,** 327.

Wever, E. G., M. Lawrence, G. von Békésy, *Proc. Natl. Acad. Sci. U.S.A.* **40**, 508 (1954).

Wilson, J. P. *Hearing Res.* **2,** 233 (1980).

Wilson, J. P. *Hearing Res.* **2,** 527 (1980).

Wit, H. P. and R. J. Ritsma, *Hearing Res.* **2,** 253 (1980).

Zhou, S. and J. O. Pickles, *Hearing Res.* **100,** 33 (1996).

Zwislocki, J. J., and E. J. Kletsky, *Science* **204**, 639 (1979).

Appendix 9. Simple musical ratios in ratios of cavity lengths

Table A-4 (column 4) presents ratios close to the major second, major third, fifth, major sixth, major seventh, and octave. Russell's data shows that L_0/L_{-1} is almost indistinguishable from an equal-tempered semitone, and so it has been set at 1.0595 (100 cents). Then L_{-2}/L_{-1} is only 1 cent larger than a just major second (9:8), L_3/L_{-1} is 15 cents smaller than a just major third (5:4), L_{-4}/L_{-1} is 14 cents larger than a just fifth (3:2), L_5/L_{-1} is 8 cents below a natural seventh (7:4), and L_{-6}/L_{-1} is 12 cents below an octave (2:1).

In general, the probability of small-integer ratios (m : n) appearing by chance in the range 200–1000 cents is considerable ($p \approx 0.5$ for $m, n \leq 10$). Nevertheless, if measurement of a number of different hair cell patterns repeatedly produced ratios near 2:1 and 3:2, it would be powerful evidence for the hypothesis presented here.

(facing page)

FIGURE A-9. Tracing of stereocilia array in a monkey (from Fig. 7 of Lonsbury-Martin *et al.* 1988*) showing a geometry similar to that in Fig. A-4a, except with a supplementary fourth row. The ratio between spacing of the cells within a row and across two rows is 0.34, and the arctan of this ratio is 19°. Cells separated by two rows are not quite aligned perpendicularly (in the radial direction): instead they are sheared some 4° away from this direction. Note the way the positioning of cells in the fourth row emphasises the first oblique (L_{-1}) mode (grey lines) in the –23° direction, and how the right arms of the stereocilia are generally at right angles to this cavity. In a similar way, the left arms point most strongly towards the L_5 cavity at +53° (thin black lines), although some L_7 (at 64°) and L_9 (72°) cavities (not marked) can also be discerned. It is significant that the average length of the L_5 cavities is 1.51 ±0.03 times that of the L_{-1} cavities, a ratio close to 3:2 (and that the ratio of L_7/L_{-1} is close to 2 and the ratio of L_{-1}/L_9 is about 3), an arrangement ideally suited to simple and rapid detection of harmonics.
(Figure adapted from Lonsbury-Martin *et al., Hearing Research* **33**, 81 (1988), with kind permission of Elsevier Science Ireland Ltd.)

* Lonsbury-Martin, B. L., Martin, G. K., Probst, R. & Coats, A. C. 1988 Spontaneous otoacoustic emissions in a non-human primate. II. Cochlear anatomy. *Hearing Res.* **33,** 69–93.

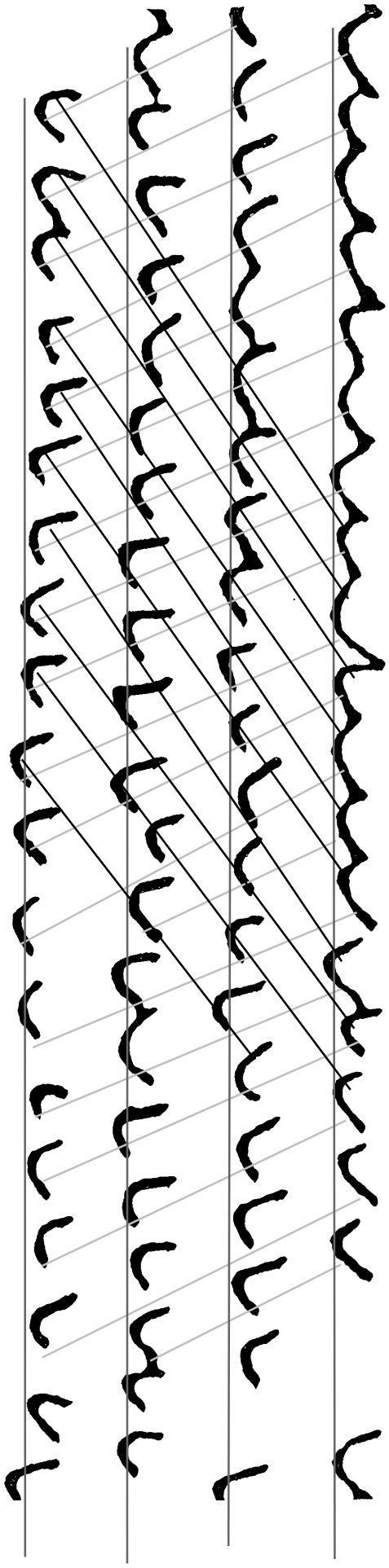


FIG. A-9

Appendix-10

Is cochlear geometry the basis of music? Outline of a theory

Pythagoras discovered that dividing a stretched string into simple integer ratios produced musical intervals, but why music should relate to such special numbers is obscure. Current-day inclination is to locate our wonderful musical sense in a neural correlation network somewhere in the brain. In an earlier paper, it was hypothesised that spontaneous otoacoustic emissions could be accounted for by cooperative functioning of outer hair cells in the cochlea. This paper shows that the typical geometric pattern in which the OHC lie generates important musical ratios such as the semitone and octave — and can produce an entire chromatic scale in certain cases — and the idea is put forward that musical intervals are sensed by the two arms of the OHC stereocilia acting as ratio detectors.

Spontaneous otoacoustic emissions (SOAEs) are faint, pure tones emitted by the ear which can be detected with a microphone placed in the ear canal[1]. Since their discovery by Kemp in 1979 (ref. 2), they have been taken as good evidence that the cochlea is very finely tuned, but it has been difficult to say what the tuned elements are. Recently, this author[3] suggested that the elements were reverberating cavities in the tectorial membrane between rows of active outer hair cells, an arrangement that worked as a surface acoustic wave resonator.

The main paper pointed out that the most common ratio between neighbouring SOAEs was 1.06 ± 0.01, a number that is very close to an equal-tempered semitone ($2^{1/12}$ = 1.05946..). In fact, as Figure A-5a shows, the average of 216 neighbouring intervals in a particularly rich data set[4] was 0.99 semitone, with a spread from 0.6–1.3 semit. This result is explained as arising from an approximate 19° angle between the primary transverse resonator and the first oblique one, an angle defined by the face-centered orthorhombic geometry of the outer hair cell array.

The appearance in cochlear tuning elements of a ratio that is very close to the basic musical interval of the semitone immediately raises the idea that this alignment may have musical relevance. This notion is strengthened when we discover that, for a 19.1° first oblique angle, another vital musical interval, the octave, appears in the same OHC pattern as the length of the fifth oblique resonator. The aim of this paper is to place the set of lengths generated by the OHC array in a musical context, and show that this schema can give rise to the entire 12-tone chromatic scale.

How can the ear detect such intervals? The stereocilia of OHCs appear as two prominent arms, and it is proposed that this unique construction is designed to detect ratios in the rate of oscillation of each arm.

The musical ear

When a stretched string is divided into two lengths of ratio 2:1, a pure octave is formed. Similarly, a ratio of 3:2 creates a musical fifth. As Pythagoras showed, these two intervals are sufficient to create, by a repeated cycle of fifths, the 5-note pentatonic scale (in

'Pythagorean' tuning)[5,6]; extension of this cycle produces the 7-note diatonic (major) scale, and (allowing for the small 'comma of Pythagoras'[5]) the complete 12-note chromatic scale with which we are familiar.

Many cultures have come by a similar arrangement, and it has often been asked why. A naturalistic answer has been keenly sought, but no physical structure responsible for our pre-disposition towards this set of simple ratios has been found[7,8]. The matter has usually been displaced to the realm of higher order signal processing, variously assigned to the brain, mind, or soul[7]. Some have even denied that music resides in our make up, merely that it reflects cultural conditioning[9,10].

This paper provides an explanation for our musical sense which involves the special geometry of the outer hair cell pattern, an arrangement that can deliver important musical ratios.

There is a major difficulty in accepting a naturalistic basis for the 12-tone scale: whence come the simple integer ratios? While some intervals clearly have simple integer ratios, not all do. Thus, the octave (2:1), fifth (3:2), and fourth (4:3) bear simple ratios in the Pythagorean scale, but an insistence that the ratios be powers of 2 or 3 gives uncomfortably complex ratios like 81:64 for the major third and 243:128 for the major seventh. The 'just' intonation scheme (which we owe to Ptolemy and Ramos) allows for powers of 5 (refs 11, 5, 6) and diminishes the magnitude of the integers involved, but a ratio of 45:32 for the tritone is not really simple, and there is a reluctance to admit ratios involving the larger primes (7, 11, 13, and so on) because if all the consonances are made just, the resulting scale would be unstable under changes of key[12]. There seems to be arbitrariness in deciding which ratio is sufficiently simple from a theoretical viewpoint while in practice still sounding accurate and consonant. The contradiction always remains that no simple ratio can exist for the semitone which, when raised to the 12th power, returns us to the octave. In practice, the ear tolerates irrational approximations sufficiently well[13] that the equal-tempered scale (semitone of $2^{1/12} = 1.05946..$) has become nearly universal despite its sharp major third and other infelicities[6,14].

The overwhelming advantage of equal temperament is that it allows instant modulation from one key to another without retuning the entire instrument. And while studies have shown that musical intonation on unaccompanied voice or instruments such as the violin tends towards Pythagorean tuning[11], it doesn't demand exact small-integer ratios either[9].

This paper provides a possible answer to this difficulty in which the set of musical intervals is derived from a geometric arrangement of the ear's regular crystalline lattice. For certain lattice parameters this leads to specific cell alignments which represent musical intervals. The interval is usually a small-integer ratio, for it is designed to detect harmonics which naturally occur in the aural environment. However, each interval does not build upon its predecessor by using a constant multiplier; rather, it involves a triangular geometric construction, and alignments of hair cells can occur for irrational intervals.

Put this way, a scale is a set of favoured integral frequency ratios which fall into a coherent perspective, somewhat like viewing the shadows of a picket fence. Many members of the set correspond with, or come close to, ratios found in 'just' intonation, but not all of them; noteworthy are ratios involving 7, which have conventionally been banished from the musical stage[15].

Musical trigonometry

A new description of cochlear mechanics involving a special quasi-crystalline OHC geometry was described earlier[3]. That paper called for outer hair cells to create a perpendicular resonator, which defined the characteristic frequency of the cochlear partition. However, because of the regular geometrical arrangement of the cells, a set of oblique resonances could also appear at particular angles to the perpendicular, generating five major oblique cavities of length L_1, L_2, L_3, L_4, and L_5**.** The frequency of a cavity is, of course, the inverse of its length.

For a face-centered arrangement of outer hair cells, as observed, the relative lengths of the oblique resonators depend on the ratio between the lengthwise spacing of OHCs along the partition and the distance between rows 1 and 3. This ratio can be specified as

the angle, θ_1, between the orthogonal and the first oblique resonator. That is,

$$\theta_1 = \arccos (L_0/L_1) ,$$

where L_0 ($\equiv$1) and L_1 are the lengths of the orthogonal and first oblique resonators. The relation between lengths of the other oblique resonators is then given as:

$$L_n = 1/\cos \theta_n = 1/\cos (\arctan(n \tan \theta_1)) ,$$

where L_n is the length of the cavity n hair cells away from the perpendicular and θ_n is the angle enclosed between the resonator and the perpendicular. A precise octave (θ_5 = 60° and L_5 = 2.000) therefore arises from a first oblique angle of 19.1° (1.058, or 0.98 semit). From the same schema, it is calculated that the 0.99 semit average, seen in the data of ref. 4, corresponds to an angle of 19.2° and gives an octave of 2.007, only 6 cents too large (1 octave = 1200 cents).

Investigating further, it can be seen that, based on this geometry, a small disparity exists between the size of this L_1 semitone and its L_5 'octave'. An exact equal-tempered semitone ($2^{1/12}$) translates to a first oblique angle of 19.29° (that is, $L_1 = 1.059 \times L_0$), but then $L_5 = 2.015 \times L_0$, meaning a slightly enlarged octave of 1213 cents. In this connection, experiments show that even the best musicians judge 'an octave' to be slightly larger than the theoretical 2:1 ratio[11,16]. The data indicate they prefer a stretched octave of close to 2.01 (1210 cents), a value very close to the calculated disparity.

For θ_1 = 19.1°, the corresponding lengths are 1.06, 1.22, 1.44, 1.71, and 2.00. Musically, only the first and last in this series carry musical significance, the others represent unmusical intervals of 339, 634, and 928 cents.

However, by taking L_5 to be a semitone lower, a particularly auspicious arrangement of hair cells emerges. This arrangement, illustrated in Figure A-10a, is based on L_5 = 2.125, meaning that θ_1 = 20.56° and L_1 = 1.068 (114 cents, a Pythagorean semitone and very close to a just semitone of 16:15 or 1.067). Using the equation for L_n given earlier, L_2 = 1.250, exactly a just major third; L_3 =

1.505, only 6 cents above a just fifth; and L_4 = 1.803, only 2 cents above a minor seventh. (Note that, for ease of comparison, we talk in terms of lengths, even though the frequency is the inverse of this. As an historical point of interest, Pythagoras explained his ideas in terms of string lengths, not frequency, and considered the scale to progress by increasing string length, as the scale here does.)

If it is supposed that the ear has some mechanism for comparing the length (frequency) of the L_1 cavity with the others, a rendition of the entire musical scale can be created, as shown in Fig. A-10a and Table A-10. A remarkable aspect of this scheme is that 8 of the 12 intervals turn out to be within 9 cents of the just scale values, 10 of them are within 12 cents of simple integer ratios, and all are within 34 cents of their equal-tempered equivalents.

As foreshadowed, this scheme has the virtue that we are not constrained to build up the scale by reiteration of one or two intervals; instead the triangular geometry spells out the five basic musical ratios directly. This occurrence of a complete musical scale from directly observable cochlea elements seems to be beyond coincidence, and it is postulated that the physical basis of our musical appreciation may have been uncovered. This inner ear instrument possesses strings not unlike the piano or harp strings Helmholtz was seeking, and it seems that at least some of them are musically tuned.

A string model of this 'ear harp' has been constructed based on a θ_1 of 20.6° (Fig. A-10b), and when tuned according to Table A-10 it is conveys musicality, despite its deviations from the conventionally accepted scale.

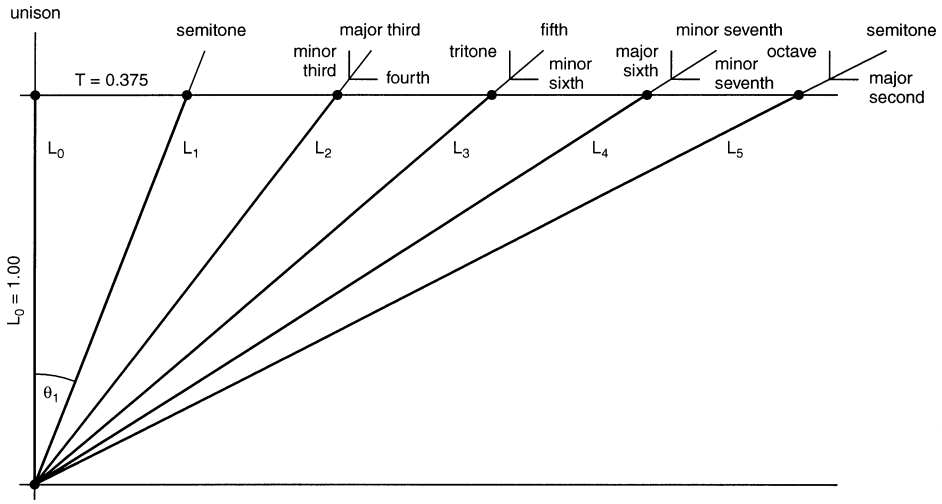


FIGURE A-10a (above) and TABLE A-10 (below). Geometry of the outer hair cell array can, in one particular case, produce every musical interval. The alignments are based on the general scheme of the face-centered orthorhombic lattice outlined earlier, but in this case the ratio of horizontal to vertical spacing of hair cells was chosen so as to give a particular first oblique angle, θ_1, of 20.56° (that is, θ_1 = arctan 0.375). For this angle, the oblique lengths L_1 to L_5 produce five musical intervals accurate to within 6 cents of the just scale intonation. In addition, by comparing the ratios of these lengths with L_1 (in a manner described in the text), the remaining musical intervals are created with no more than 34 cents deviation from their equal-tempered equivalents. A model of this configuration has been constructed with guitar strings, and it plays music. Details of the lengths are given below.

TABLE A-10

L_n	length	size (cents)	close ratio	musical interval	deviation from ratio (¢)	deviation from ET (¢)
L_0	1.000	0	1:1	unison	0	0
L_1	1.068	114	16:15	semitone	+2	+14
	1.17	272	7:6	minor third	+5	–28
L_2	1.250	386	5:4	major third	0	–14
	1.34	500	4:3	fourth	+2	0
	1.41	594	7:5	tritone	+11	–6
L_3	1.505	708	3:2	fifth	+6	+8
	1.61	822	8:5	minor sixth	+8	+22
	1.69	907	5:3	major sixth	+23	+7
L_4	1.803	1020	9:5	minor seventh	+2	+20
	1.93	1134	15:8	major seventh	+46	+34
			17:9		+33	
	1.99	1191	2:1	octave	–9	–9
L_5	2.125	1305	17:8	semitone	0	+5
	2.27	1419	16:7	major second	–12	+19

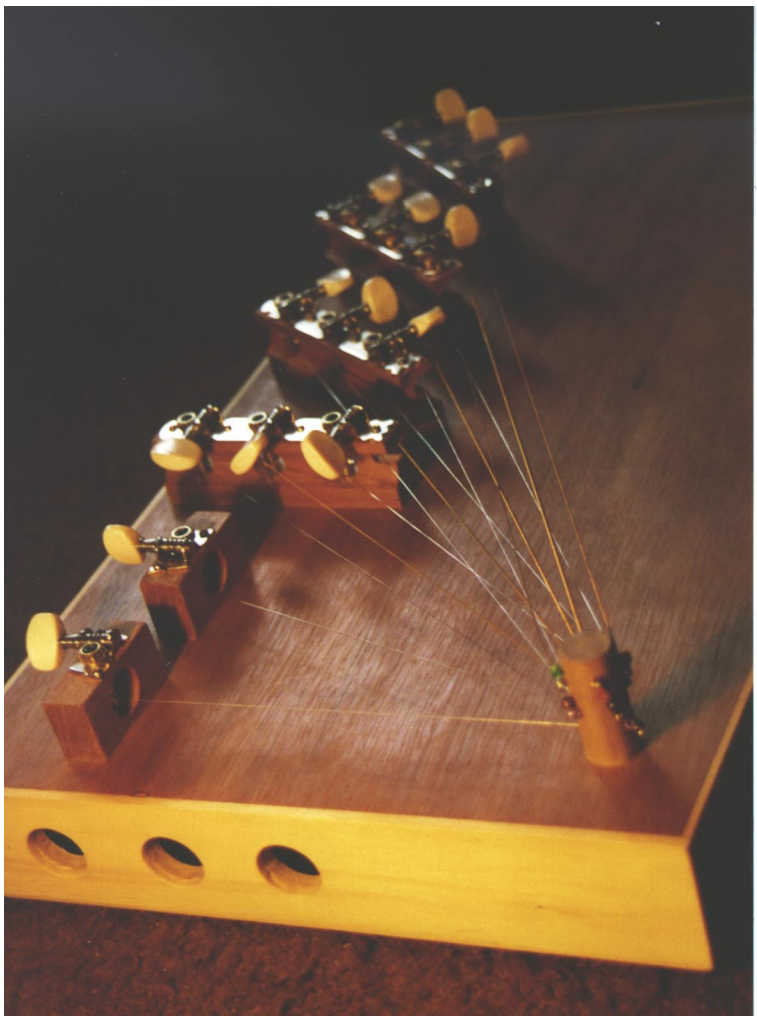


Fig A-10b. A guitar-string model based on the cochlear geometry of Fig. A-10a. This 'ear harp' renders music.

Ratio detectors

Providing our ear with a multiple set of discrete tunings — preferentially musical — at each point along the partition permits direct frequency comparisons of the components of complex tones to be carried out, notably those at musical intervals. But how could this be done?

The explanation calls on the unique construction of OHCs which allows them to act as ratio detectors. Conspicuously, each cell possesses two arms, equipping it with the ability to detect the ratio with which they beat. For example, when one stereocilia arm responds to a θ_1 mode at 19°, and the other to the θ_3 mode at 46° (Fig. A-9c), the OHC would be able to signal this occurrence.

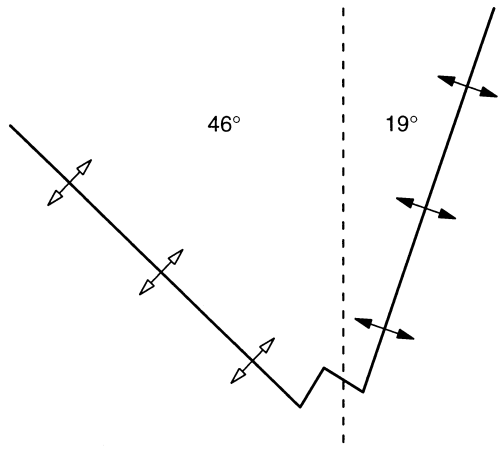


FIGURE A-10c. An outer hair cell has two stereocilia 'arms' which appear to be articulated at the join with a concertina coupling that allows independent to and fro movement of the arms. Although basically a 'V' shape, the hinge creates the often noted 'W' arrangement. The angle of the V changes from base to apex, and a distribution of angles always seem to be present at any point on the partition. The two arms are often found angled in such a direction that a resonant cavity will be formed between them.

Of particular import, the arms of an OHC appear to be articulated at the join with a concertina coupling that allows independent to-and-fro movement of the arms. Although basically a 'V' shape, the hinge creates the often-noted 'W' arrangement[17]. Detection could be carried out somewhere near the intersection of the two arms; the distinct basal body at this point (where, in developmental terms, the kinocilium used to be) is a likely spot[18].

On a knife edge

Evidence for this dual-frequency behaviour of OHCs comes from observations of bi-stable SOAEs, which are found to jump backwards and forwards between two fixed frequencies at an irregular rate (every second or so, more or less). Table A-10b lists observations of this phenomenon, and we see that the human data involve switching between semitone intervals, a distinctive behaviour that can be interpreted as switching between an orthogonal and an oblique resonator. Moreover, the table indicates that at least some of these instances involve switching between frequencies that are integer-related.

Table A-10b
Bistable emissions

author	f_1 (Hz)	f_2 (Hz)	f_2/f_1	semits	Δf (f_2–f_1)	$f_1/\Delta f$	$f_2/\Delta f$	f_1:f_2
Keefe et al. (1990)	1595.6	1701.8	1.0666	1.12	106.2	15.024	16.024	15:16
	1408.1	1524.1	1.0824	1.37	116.0	12.139	13.139	85:92
	1330.6	1410	1.0597	1.00	79.4	16.758	17.758	67:71
Wit (1990)	1612	1700	1.0546	0.92	88	18.32	19.32	55:58
Wilson et al. (1988)	3002 ±5	3233 ±5	1.077	1.28	231 ±10	13.0	14.0	13:14
Bell (unpublished)	2165.5 ±0.1	2295.6 ±0.1	1.0601	1.01	130.1 ±0.2	16.62–16.67	17.72–17.67	50:53
van Dijk et al. (1996) [barn owl]	8544	9018	1.055	0.93	474	18.03	19.03	18:19
Zurek and Clark (1981) [chinchilla]	4730	5680	1.20	3.2	950	4.98	5.98	5:6
Ohyama et al. (1991) [guinea pig]	1438	1489	1.0355	0.60	51	28.2	29.2	28:29

Caption: Bistable emissions appear to show simple integer ratios. In some cases, because of limited measurement accuracy, the ratios in the final column are in doubt, but taken together there is a clear tendency towards simple ratios. As discussed in the text, this behaviour can be taken as evidence of locking between two cavity modes (oblique resonators) of fluctuating strength, presumably the L_0 and L_1 cavities.

References: Keefe, D. H., Burns, E. M., Ling, R. & Laden, B. in *Mechanics and Biophysics of Hearing* (eds Dallos, P., Geisler, C. D., Matthews, J. W., Ruggero, M. A. & Steele, C. R.) 194–201 (Springer-Verlag, Berlin, 1990) (ref. 18); Wit, H. P., op. cit. 259–268; Bell, unpublished; Wilson, J. P., Baker, R. J. & Whitehead, M. L. in *Basic Issues in Hearing* (eds Duifhuis, H., Horst, J. W. & Wit, H. P.) 80–87 (Academic Press, London); van Dijk, P., *J. Acoust. Soc. Am.* **100,** 2220–2227; Zurek, P. M. & Clark, W. W. *J. Acoust. Soc. Am.* **70,** 446–450; Ohyama, K., Wada, H., Kobayashi, T. & Takasaka, T. *Hearing Res.* **56,** 111–121 (1991).

This 'mode hopping' behaviour is just what we expect from a self-excited, multiple-frequency system, and locking to integer ratios is another tell-tale characteristic. Mode locking occurs when a single active system generates two incommensurate frequencies, and is commonly observed in nonlinearly excited musical oscillators[19]. At a frequency where harmonics of the two driving frequencies nearly match, the system couples (due to a degree of feedback to the frequency-generating process). It is suggested that this is what happens when the two stereocilia arms oscillate at two incommensurate frequencies, say at the frequencies of the L_1 cavity in the left arm and the L_3 cavity in the right. Although the two arms can undergo largely independent oscillation, the cell is a unitary structure, and some degree of coupling must exist. Because of this coupling, the arms will tend to lock. The simpler the ratio, the stronger the locking.

The integers displayed in the table range from 15 to nearly 100; however, because of uncertainty in some of the measured frequencies, there is doubt whether some of the values should be multiples or sub-multiples of what is listed. Clearly, high accuracy measurements would be required to firmly establish that switching always involves discrete integers. Nevertheless, Table A-10b does show several unambiguous cases of integer ratios. These integers are not always small, although their magnitude — up to 92 — corresponds with the harmonic partials uncovered in two sets of SOAEs[20] and in the pattern of saxophone multiphonics[21].

Mode locking is possible in the OHC system because the frequency of an acoustic laser can be slightly altered, as demonstrated by the well-known entrainment phenomenon in which an SOAE locks to an externally presented tone over several hertz[22]. This behaviour is just what is needed for the beating frequency of each arm to be 'pushed' or 'pulled' into partial synchrony with the adjacent arm and lock to the nearest integer ratio. (Incidentally, it also demonstrates that the frequency of the acoustic laser is not determined exclusively by its length; there must be a small phase (time) delay built in to the kick-back response of the stereocilia when responding to a bend-inducing

wave-front, so that a small 'end correction' of the cavity length may sometimes be called for.)

When sets of SOAEs are analysed using an 'error-function method'[20] which detects common sub-harmonics in the frequencies of SOAEs, the appearance of sharp dips (corresponding to apparent 'fundamental' frequencies) indicates mode locking. A simpler, less rigorous method is to divide all SOAE frequencies by the smallest difference neighbouring interval and look for results close to integers or small fractions. Both methods have been applied to a number of SOAE sets and recurring integer relationships have been seen.

Examination of the ratio distribution shown in Fig. A-5b reveals a number of plateaus, and these preferred ratios can be interpreted as evidence of mode locking.

Variations on a geometrical theme

We have portrayed a precise crystalline lattice with constant geometry, but an examination of micrographs shows that this precision can be disturbed to greater or lesser degrees. In practice, therefore, there is a range in a/b ratios, and hence, in first oblique length and in the length of related longer cavities. Thus, Fig. A-5a shows that L_1 varies from 0.6 semit (corresponding to a θ_1 of 15° and an L_5 of 1.67) to 1.35 semit (θ_1 = 22° and L_5 of 2.25). Put another way, L_5 could range from a major sixth to a major tenth. Probably, the ear does some global average of perceived ratios in stereocilia arm beatings, and works on a semitone very close to that of the equal tempered scale (the average interval in Fig. A-5a is 0.99 semit). Without this averaging process, musical performance and appreciation would be a very precarious affair, requiring exact intonation.

The distribution of spacings allows many ratios to appear, and it is supposed that, during development, afferent nerve fibres preferentially connect to the simple-ratio (musical) cavities. With a fixed a/b (to date we have used a nominal spacing based on L_5 = 2.0) it would be impossible to arrange tuning of all the cavities to give exact, simple ratios, but a small variation in the θ_1 angle

allows generation of musically preferred intervals. For example, a θ_1 of 18.43° gives a semitone of 91 cents, an L_2 of 1.202 (minor third), L_3 of 1.414 (tritone), and L_4 of 1.667 (major sixth).

Examination of the extensive data set on stereocilia positions in the cochlea of a monkey[23] reveals this variation directly. Thus, in Fig. 7 of this reference, when we take for example the segment of the cochlea representing 54–59% of the distance from the apex, the 104 clearly defined L_1 cavities (leaning right from OHC3) show a mean length of 18.2 arbitrary units with a standard error of 1.0 unit. At the same time, the 71 comparable L_5 cavities (leaning in the opposite direction) reveal a length of 35.9 ±1.7 units, giving an average ratio of 1.97 (close to an octave). Similarly, the average L_3 (to the left) measures 26.7 ±1.4 units and L_1 to the right 17.6 ±0.9, giving a ratio of 1.52 (musical fifth). In general, small-integer ratios can be discerned on most of the other sections mapped, although the stereocilia pattern is never completely regular for more than 10–20 cells, and more basal sections show a much more periodic pattern than apical ones.

If it is supposed that the wiring of the nervous system is self-adaptive, and works to pick out the ratios of importance, then the regularity of the geometry which generates the ratios is not so critical. The 'irregularity' can be viewed as a means of picking out ratios which are not ordinarily defined by the regular crystalline lattice. Moreover, the lattice parameters themselves can change from one point to another.

At the middle of the cochlea, first oblique angles generally centre around 19–20°, meaning that the octave arises from the L_5 cavity. However, at the apex and base of the cochlea first oblique angles as narrow as 14° and as wide as 26° are found, and this non-standard geometry is conceivably why normal musical perception rapidly dissolves at low and high frequencies[24]. In other animals, the angles range from 13° to 37°, with both extremes coming from guinea pigs (Table A-1b). Although detailed measurements of these narrower and wider examples are called for, it seems possible that these cases represent arrangements where the basic octave length derives from L_7 (13.9°) or L_6 (16.1°) for narrower cases, or L_4 (23.4°) or L_3 (30°). Indeed, measurement of Fig. 1 in

ref. 3 shows that L_4 relates to a length of 1.98 ±0.12. Observed degrees of shearing of the OHC crystal cell[23] alter the situation slightly, and make the left and right oblique cavities of somewhat different length (as in the monkey hair cell distances given above). However, the factor that controls the available ratios that can be sensed remains the lengths of the dominant left- and right-pointing cavities within which a hair cell finds itself. Because of the assistance of the centre-row hair cell, the dominant cavities are usually the L_1, L_3, or L_5, but the direction in which the stereocilia arms point must be considered too, and this angle changes from cell to cell[23]. Three left-pointing and three right-pointing cavities allow for a total of 9 ratios to be sensed.

How does the cochlea achieve the precise geometry on which musical perception depends? Outer hair cells are supported in the cupped recesses of Dieters cells, and these supporting cells also extend flower-like phalangeal processes that, by reaching across and surrounding OHCs three or more cells distant, define the placement of OHCs in the reticular lamina[25]. It would seem that musical ratios arise from simple horizontal-to-vertical ratios of hair cell spacing: thus, the standard spacing arises from a particular spacing of 7:4 (θ_5 = arctan 1.75), the favourable arrangement of Fig. 2 derives from a ratio of 3:8 (θ_5 = arctan 0.375), and the agreeable 18.4° example above from 1:3 (θ_1 = arctan 0.333).

The exact interplay between efferent and afferent, IHC and OHC, is still unclear. However, is seems highly significant that in all published cochleograms and micrographs where long sections of the cochlea can be observed, the ratio between numbers of IHC and OHC is integral (commonly 6:5 or 5:4) and the major OHC alignments of L_3 and L_5 intersect IHC positions 4 and 7 cells away from the perpendicular (see, for example, Fig. 19 of ref. 26).

That is, the organism could also recognise an octave by detecting the simultaneous firing of IHCs 7 cells apart –similarly, 4 cells apart could designate a fifth, and so on. Interval detection could therefore be achieved by a close-neighbour analysis of IHC firings as well as by a few OHC afferents.

Coda: **integers rule**

Arguments have been exchanged, and volumes written, about which tuning temperament — Pythagorean, just, equal, or some other — is 'best'[6]. It looks from Table A-5a that a kind of just scale, involving simple ratios, comes closest to what the ear is doing. However, unlike the traditional just scale, the ear does not eschew the number 7, as some music theorists have wished[27], and close approximations to 7:6, 7:5, and 16:7 can be seen in Table A-5a. It was Archytas, Plato's friend, who first introduced 7 into music, and a fascinating account of how Plato adopted this scheme into a tuning system by which integral musical ratios arise from a right-angled triangle with all-integral sides is given in ref. 28. Helmholtz studied the seventh partial, but concluded, by considerations of harmony with the rest of the scale, that it could not be used in modern music[29].

Musical appreciation, this paper proposes, is a function of the OHC crystal lattice, an arrangement which, acting like a diffraction grating, allows musical intervals in complex tones to be sensed. It has been known for some time that pitch has dual aspects: pitch height, which provides a sense of high and low; and pitch chroma, that allows 'fifthness' and other musical intervals to be appreciated[30,31]. The former is presumably a function of distance along the cochlear partition; the latter, it has been suggested, involves the OHCs providing instantaneous frequency comparisons and the IHCs detecting the result using close neighbour analysis. To this we owe the beauty and power of music.

Some 2500 years after Pythagoras, a possible physical basis for our musical sense has been discovered and, naturally enough, it resides in our ears. We can see an elegant crystalline structure that is configured to detect the semitone, octave, and a range of simple ratios in between. The surprise is that in terms of cochlear geometry the semitone seems to be as prominent an interval as the octave, if not more so, although it never sounds that way to a listener, no matter what the culture.

If the cochlea provides us with the basic analyser for harmony and discord, it follows that novel contrived scales, like those based on 31 or 53 tones, are never likely to gain much favour.

Pythagoras said music was mathematics that could be heard. Looking closely at the cochlea, we see that it is indeed a matter of (Pythagorean) trigonometry.

ACKNOWLEDGEMENTS. To Libby, for faith; Neville Fletcher and Ken Hews-Taylor, for comments; Michael Delaney for construction of the ear harp. For the animals, who, like us, have ears and consciousness.

1. Probst, R., Lonsbury-Martin, B. L. & Martin, G. K. 1991 *J. Acoust. Soc. Am.* **89,** 2027–2067.
2. Kemp, D. T. 1979 *Arch. Otorhinolaryngol.* **224,** 37–45.
3. Bell, A. 1998 'The underwater piano: model for a regenerative receiver in the cochlea.' MS submitted to *Proc. Roy. Soc.* B.
4. Russell, A. F. 1992 *Heritability of Spontaneous Otoacoustic Emissions* (PhD thesis, U. of Illinois, publ. by UMI, Ann Arbor).
5. Jeans, J. H. 1947 *Science and Music* 160–190 (Cambridge University Press, Cambridge).
6. Hall, D. E. 1980 *Musical Acoustics* 436–461 (Wadsworth, Belmont).
7. Parncutt, R. 1989 *Harmony: A Psychoacoustical Approach* 5–11 (Springer-Verlag, New York).
8. Plomp, R. & Levelt, W. J. M. 1965 *J. Acoust. Soc. Am.* **38,** 548–560.
9. Burns, E. M. & Ward, W. D. 1982 in *The Psychology of Music* (ed. D. Deutsch) 241–269 (Academic Press, New York).
10. Terhardt, E. 1974 *J. Acoust. Soc. Am.* **55,** 1061–1069.
11. Ward, W. D. 1970 in *Foundations of Modern Auditory Theory* (ed. Tobias, J. V.) 407–447 (Academic Press, New York).
12. Walker, D.P. 1978 *Studies in Musical Science in the Late Renaissance* 1–13 (The Warburg Institute, London, 1978).
13. Vos, J. P. 1982 *Percept. Psychophys.* **32,** 297–313.
14. Redfield, J. 1928 *Music: A Science and an Art* 84–85 (Tudor, New York).
15. Levarie, S. & Levy, E. 1968 *Tone: A Study in Musical Acoustics* 31–40, 46-54 (Kent State University Press, Kent, Ohio).
16. Hartmann, W. M. 1993 *J. Acoust. Soc. Am.* **93,** 3400–3409.
17. Engström, H., Ades, H. W. & Hawkins, J. E. 1962 *J. Acoust. Soc. Am.* **34,** 1356–1363.
18. Hackney, C. M. & Furness, D. N. 1995 *Am. J. Physiol.* **268,** C1–C13.
19. Fletcher, N. H. 1978 *J. Acoust. Soc. Am.* **64,** 1566–1569.
20. Keefe, D. H., Burns, E. M., Ling, R. & Laden, B. 1990 in *Mechanics and Biophysics of Hearing* (eds Dallos, P., Geisler, C. D., Matthews, J. W., Ruggero, M. A. & Steele, C. R.) 194–201 (Springer-Verlag, Berlin).
21. Keefe, D. H. & Laden, B. 1991 *J. Acoust. Soc. Am.* **90,** 1754–1765.
22. Wilson, J. P. & Sutton, G. J. 1981 in *Tinnitus* (ed. Evered, D. & Lawrenson, G.) 82–107 (Pitman, London).
23. Lonsbury-Martin, B. L. Martin, G. K., Probst, R. & Coats, A. C. 1988 *Hearing Res.* **33,** 69–94.
24. Biasutti, M. 1997 *Hearing Res.* **105,** 77–84.
25. Bloom, W. & Fawcett, D. W. 1968 *Textbook of Histology* (9th edn) 826–828 (W. B. Saunders Co., Philadephia).
26. Bredberg, G., Engström, H. & Ades, H. W. 1965 *Arch. Otolaryng.* **82,** 462–469.
27. Shirlaw, M. 1917 *The Theory of Harmony* 78–79 (Da Capo Press, New York, 1969, reprint of the 1917 edn).
28. McClain, E. G. 1978 in *The Pythagorean Plato: Prelude to the Song Itself* 117–126 (Nicolas Hays, Stony Brook).
29. Helmholtz, H. L. F. 1875 *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Ellis, A. J., trans.) 228 (Longmans, Green, and Co., London).
30. Bachem, A. 1937 *J. Acoust. Soc. Am.* **9,** 146–151.
31. Licklider, J. C. R. 1951 *Experientia* **7,** 128–134.