

## EXTENDING THE HUMAN ACOUSTIC MODEL

We have seen that we can create microphones and loudspeakers that capture and reproduce sound sources placed in space. These reproduced sources could then be treated as real – measured and recorded, or listened to. We have avoided the human perception aspect. We now look at this.

Lord Rayleigh published his paper on the directional nature of sound perception in 1907 [1]. This work explained binaural hearing behaviour using a model considering the ears as simple pressure sensing microphone elements. Two variables were used – Inter-aural Time Delay (ITD) for the additional distance the sound wave had to travel around the head and Inter-aural Intensity (level) Difference (IID or ILD) for the sound source proximity to the nearest ear.

This model essentially recognised that two pressure sensing elements separated by a small distance along a line can selectively determine the direction of a sound source.

### A primitive model of hearing

We will use an electro-acoustic analogy of human acoustic perception to help determine exactly what the essential “system” requirements are. This will provide further insights into how acoustic perception could work.

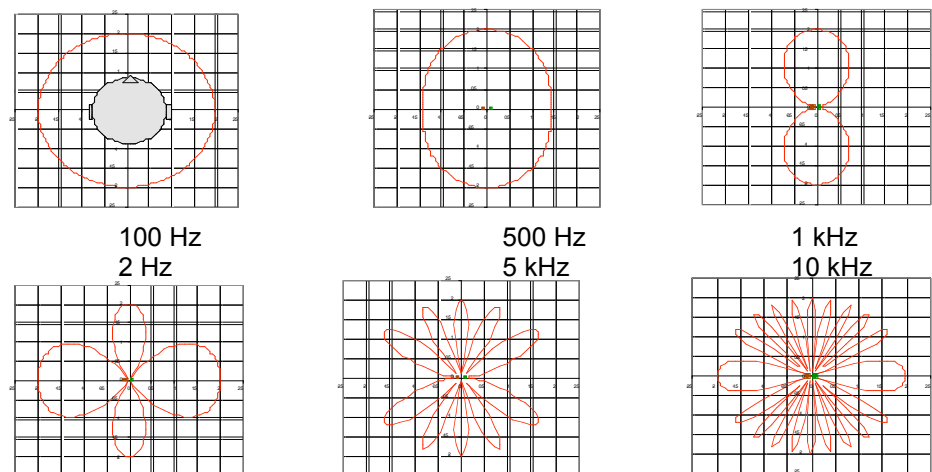


FIGURE 1 SIMPLE TWO MICROPHONE DIRECTIONALITY

We start with the basic configuration described above. Two microphone elements are spaced 170 mm apart in free space to model a primitive human head. The electrical outputs from the microphone elements are summed.

Figure 1 shows the directional behaviour of such a model. The graphs represent the measured output from a point sound source moved around the “head”, measured at various frequencies. To clarify the nature of the graphs with respect to path difference the sound level received at each element has been adjusted in the model to be the same. In practice the levels are different because of proximity and shadowing. This means that the nulls are not as deep or as regular in appearance. The full three-dimensional graphs extend the lobes to toroidal or doughnut shaped responses.

The maximum path difference between the elements occurs when the sound source is to the extreme left or the right. At approximately 1 kHz, the path difference from the source to each



element has become half the wavelength and so the summed signal has a cancellation null or dip appear that occurs symmetrically to the left and the right of the elements (the head).

As the frequency is further increased the half wave cancellation path difference is decreases and so the null or dip location “moves around the head” with front-back symmetry. At the same time the next highest frequency of cancellation corresponding to three half wavelengths occurs to the left and the right and so multiple nulls now “appear”, and move to the front and back. As the frequency is increased, more and more nulls or dips appear and move to the front and the back. At any nulls, the particular frequency is theoretically missing completely in that direction – there is an acoustic blind spot. No amount of electrical equalisation of the overall or summed output signals will recover signals from such nulls.

Human hearing does not exhibit these periodic complete nulls. The primitive model will need augmenting by some means to more accurately represent human hearing, but the result of path differences can be clearly seen.

### **Improving the situation with electrical filtering**

We now consider whether electrically treating the signals from each element separately before summing would affect the directionality. The answer is that it can.

The simplest example to show how this could work only changes the response for sounds in one particular direction.

The signal from the nearest element is delayed by the propagation time to the furthestmost element. The time delay would effectively mean that there was no path difference between the elements and so the output from the combined elements was independent of frequency. This would only apply for a point sound source in one direction. As the source rotated away from this direction, nulls would reappear. This is not a very useful apparatus unless you can continually more the array to align with the source (turn the head), have some way of deriving delayed signals for every direction and then determining the relevant delay to use for each source direction or have the luxury of signals from many “ears” around the head to treat so that any nulls are outside the frequency range of consideration, but it does demonstrate three important aspects:

1. Microphone design strategies can either minimise or compensate for design dimensions.
2. The directional response can be altered by electrical signal processing of the individual element signals in isolation. Nulls and dips can be shifted but not necessarily eliminated because, for this geometry of elements the effect of the path difference cannot be removed in all directions simultaneously by this form of signal processing alone.
3. The electrical outputs of the microphone elements can be *simultaneously* electrically processed to obtain more than one output, each with a different directional behaviour. In the above case both the untreated response with nulls and the delayed response with no nulls to the left (or the right) would be available.

A different approach will be required if acoustic blind spots (nulls) and dips are to be eliminated, but electrical processing is always available to assist with processing should the need arise.



When this model is translated to the human hearing case the necessary electrical signal processing becomes the responsibility of the brain. In particular, where frequency dependent processing is required, such as in the case of time delay and phase filtering of multiple frequencies, the brain will need to receive signals analysed into frequency bands by some means (cochlea). Significant signal processing “brainpower” would be required to undertake the processing, and some forms of both long term and dynamic temporal storage of the processing artifacts would need to be included in the brain along the way.

### Improving the situation with directional acoustic filters

It is also possible to treat acoustic waves arriving at each microphone element with external, purely acoustic filters.

External acoustic filters have a significant advantage over pure electrical processing in that they can be physically constructed to vary phase and amplitude with direction. Electrical filter processing is limited to working with the signals from the microphones and therefore has limited processing capability for direction. Arrays of elements can be used to synthesise directional response by electrical means alone, but the highest frequency of operation is limited by the element spacing and so multiple elements are required, particularly if either high directivity, full coverage over all space or a combination of both is required. External acoustic filters will of necessity have to take into account the acoustic behaviour of the overall microphone structure. This is also the case with multi-element arrays using solely electrical processing.

Many external acoustic filter treatments are possible. For example, with a spherical geometry head having just two elements, it is possible to remove periodic frequency dependent nulls and dips completely using an acoustic filter that exactly compensates for the path length difference between the elements around the sphere. A suitable filter for each element is based on a quarter wave transmission tube. Whilst this filter structure is not in itself directional, it does create a directional response on the sphere. On the median plane there is no path difference, and on the axial line the path length difference is exactly compensated. There will still be amplitude and phase errors with direction (and distance) between these locations, but these are now less severe. Figure 2 shows a graph of the summed element

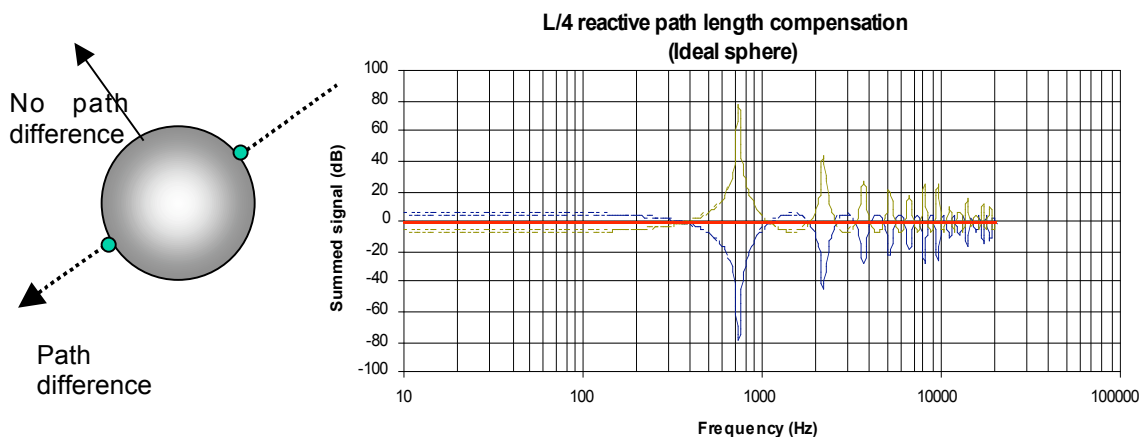


FIGURE 2 TWO ELEMENT ARRAY MICROPHONE  
PATH LENGTH RELATED ANOMALIES AND  $\lambda/4$  WAVE CORRECTION

output response for a two element spherical microphone showing the response anomalies related to path length differences on-axis, an ideal  $\lambda/4$  wave compensation filter response and the resultant overall response (red).



The remaining frequency response errors that occur with direction can now be treated by using additional acoustic filters that provide phase shift with frequency that varies with direction. A separate filter will be required for each element. These filters will be symmetrical both with regard to each other and independently as they are each correcting for a symmetrical geometry spherical head.

When a restricted number of sensing elements (two) are used, passive acoustic structures are necessary as these are the only method of providing the necessary direction dependent phase shifts. Electrical processing of the outputs from the two elements cannot achieve this.

Once the response correction has been achieved in all directions, no null or dip producing phase shifts will occur in the summed output of the two elements as sound sources move around the microphone and as the sources move closer and further away. It is still possible to process the difference signal and obtain divergence (sound source distance) information at the same time, though the use of two elements placed on a diameter would limit the usefulness of the approach for capturing source distance information as no divergence information would be able to be captured anywhere on the median plane.

To overcome this limitation, the microphone elements can be moved away from their diametrically opposed locations. This relocation will require adjustment of the acoustic equalisation filters as now the path length differences have changed to become dependent on direction. Many combinations of  $\lambda$  wave equalisers and directional phase filters are possible. The  $\lambda$  wave stub equalisers could be adjusted to compensate the dominant path difference that is now the shortest distance on the sphere surface. The directional phase filters on each element can then equalise the remaining directionally dependent phase shifts, noting that now the filters will be mirror images of each other but independently asymmetric.

By this means it is possible to approach the uniform response capability of the symmetric case, but it may not be as effective in removing all peaks and dips with direction because of the variable shadowing now present. The advantage, however, is the now the difference signals will provide divergence (distance related) output in all directions.

The acoustic filters correct the amplitude/phase differences *between* the two microphone elements and so would need to consider both the basic physical configuration including sphere diameter and microphone offset from the diameter. An iterative design strategy is thus most likely required. Where the basic shape varied from a pure sphere, the principles outlined above could be applied to develop the appropriate equalisation that preserves both distance and direction information for sound sources placed anywhere.

Sets of filters could be designed such that there was no direction over 360 degrees in all planes for either sensor where unacceptable loss of frequency band amplitude and phase information occurred, and that allowed determination of directional and distance related information. These filters would exhibit physical mirror symmetry when used to compensate spherical shapes.

This is fully analogous to the nature of human hearing, the ear canals and the ears on the human head. In profile it can also be seen that the human head indeed does have offset ear locations.



## The electronic head

For humans, certain directions may need preferable coverage at the expense of worse performance in other directions, and certain frequency bands may be of greater interest than others. This would enable simplification and customisation of the structures and the subsequent signal filters.

Electrical signal processing of the outputs from the individual elements by the digital signal processing “brain” would also be available for assistance in the design process [2].

One approach would be to design acoustic filters that minimise the amount of subsequent processing power required to extract the required source location information. Acoustic filters are just topological structures and can be replicated, whereas “brainpower” is usually in demand for other tasks.

Each directional filter microphone (ear) should now be able to discern direction information on its own, taking into account the acoustic influence of the head and torso shape. The mechanism would use the processing analysed into frequency bands and memory requirements identified previously. This is the first significant extension to the Rayleigh model.

The originally simple two-microphone element model can now be split into the arriving acoustic wavefront part (that is already known to carry information about both the distance and the direction of all sources), directional acoustic filters, the microphone elements with spectral analysis capability and the processing required by the central processor “brain”.

This logically leads to considering what information can be extracted and what subsequent “brainpower” processing is needed.

We have previously seen that a divergent acoustic wavefront contains both distance and direction information, so we now consider a perception model based on these parameters - a distance and direction model. We already know that both distance and direction will need to be considered together in order to fully locate sound sources. Their separate consideration will now lead to a significant potential extension of the present hearing models.

## Direction sensing issues

If a point source with sufficiently broad spectral content is measured by a single microphone element on a sphere with external directional acoustic filters as described previously, sufficient information is available from the wavefront to determine the source direction. This requires the filter/element combination to have been previously calibrated for directionality at all frequencies, and these calibrations stored in a suitable form to be available at all times for comparison. In humans this is described as a learned experience

### *Verifying human acoustic perception capability*

We can verify human distance perception capability in three easy steps.

1. Each ear should be directional with frequency. We should be able to test each ear independently to verify this. We also know that the spread of frequencies in the test signal should cover the range of directionality of the ear. A broadband point source such as a sharp finger click could be used.
2. The combination of two ears should be able to discern distance. Again the spread of frequencies in the test signal should cover the range of directionality of each ear.
3. Disabling one ear should remove the ability to perceive distance, yet retain direction sensing.

When these three tests are done it is indeed found that human hearing can discern the distance and direction of sound sources, but this tells little of how this is done by the brain inside the head.



The frequency band outputs will then only be consistent for one physical direction, and so the source direction can be uniquely determined.

The source has been assumed to be well behaved and at a point. The nature of real sources can disrupt the direction determination under certain conditions:

- The source does not contain sufficient spectral information
- The source is not a point
- The source is corrupted by specular reflections
- The source varies its phase response with frequency

Reflections or coherent phase anomalies in the propagation path can also disrupt the determination.

Direction determination will require significant processing in the “brain” and reference to a learned memory. This processing will also need to split the incoming signals into frequency bands during processing in order to determine the direction.

Thus there are two significant and distinct processing activities required, and each activity will need access to stored information – the compensation for the acoustic “ear” filter response, and the split frequency band processing needed to locate each broadband sound source direction. Each processing step will require some reference to learned or stored information.

The significance of this will be seen in the next section.

### **Distance sensing by series processing**

We have seen that each of the filter/element combinations can independently determine direction. With two “ears”, two directional determinations are thus available. It should be possible to determine source distance from these two determinations essentially by trigonometry with suitable additional “brainpower” processing and with reference to a learned distances/directions memory table.

This would be a significant third processing activity and memory requirement over the previous filter calibration and direction determination requirements.

#### ***A simple test for existence of a distance processing stream***

You will need an assistant.

Human directional acoustic sensing can be overloaded and thus effectively at least partially disabled by rapidly and repetitively turning the head from side to side.

Whilst this is being done you can still quite clearly determine distance of a finger click approaching and receding, or placed near or far, but you need to use the full divergence sensing array – both ears.

This shows that separate processing streams are possible, but more elaborate tests are required to prove this.

Most listeners will innately want to stop head turning and get a fix on the sound when it approaches. Simply stop clicking and berate the listener if this occurs!

A surprising result is found if the proximity test is repeated approaching behind the listener’s head. The sensitivity to proximal sounds is commonly greater behind the listener. This would probably be consistent with the audible survival alerts of our ancestors!

It is a good idea for the listener to sit down whilst the test is done to avoid falling over.



There is another way.

### Distance sensing by simplified parallel processing

We know that the divergence or swelling of the acoustic wavefront contains all the necessary information to determine source distance.

The previous, simple acoustic filtered model could use another processing method to more directly determine distance information using the divergence of the wavefront.

The divergence or gradient information essentially considers the change in the field and thus would logically start with the difference signal from the two acoustic filtered elements. The directivity nulls will still need to be removed with acoustic filters, as was the case for the directional determination above, but then the difference signal could be used for determination of distance.

This will use some processing brainpower, but the directional filters are already available from the direction determination processing and these could be used, minimising the additional brainpower required. The assumptions would be that either the acoustic filter correction was able to be done before the sum and difference processing, that the filter corrections were duplicated or that some form of synchronised parallel streams were used for the processing.

The key point is that after acoustic correction to remove nulls and dips, the difference signal should contain the divergence information for distance in all directions and this should not need a great deal of further processing. In particular it does not need spectral analysis.

A significant processing load for decoding and comparison could now be left out by the "brain", compared with processing the two directional parts of the summed signals [3].

From the model, it then makes sense that the primary human acoustic cue is distance as this requires less "computer power" than direction based processing. Sound source distance should show as a faster perceptual awareness, and could thus be a trigger for cognitive control of direction processing.

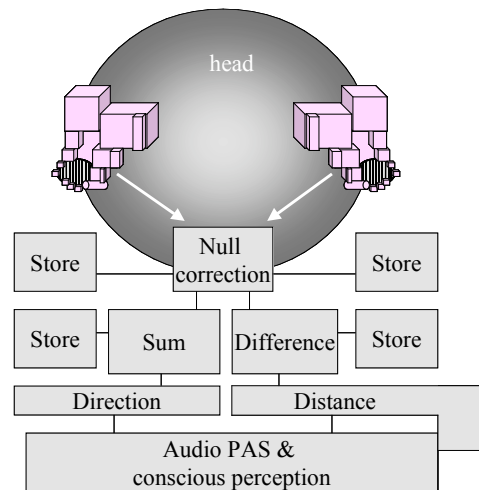


FIGURE 3 HEARING SCHEMATIC DIAGRAM

Some evidence of this has already been found for the Pre-associative Acoustic Store in humans [4]. The simple proximal test described in the box above also supports this.

An overall "wiring diagram" of human perception is suggested in figure 3, starting from the "pink" acoustic correction filters on either side of a head and progressing through to conscious perception.



The Rayleigh model can now be extended to include symmetric directional acoustic filter structures, and we can be on the lookout for parallel concurrent processing and both distance and direction perception streams in the brain.

### Graeme Huon HuonLabs 2008

---

The information contained herein is copyright to HuonLabs. No material can be reproduced in its totality or in part or without the express permission of HuonLabs Pty Ltd. Any reference to this material must quote the HuonLabs source. Trade marks and Patent applications apply to most aspects of the work disclosed here. Contact HuonLabs for further details, product information or licensing enquiries.

---

1 J. W. Strutt (3<sup>rd</sup> Baron Lord Rayleigh), “On our perception of sound direction,” *Philos. Mag.*, Vol. 13, pp. 214–232, 1907.

2 When the true geometry of human hearing is considered, the propagation behaviour of the head as a “bluff and shadowing body” would also need to be considered as part of the acoustic filters. This is not being introduced at this time to ensure that the concepts of hearing directionality are clearly understood.

3 Brain-centric evolution would have it that brains are either smart and lazy or just too popular and so evolve bodies as support systems that minimise their own processing load!

4 Crowder, R G and Morton, J. Precategorical Acoustic Storage (PAS). *Percept. Psychophys.* 5: 365-373, 1969. (Dept. Psychology, Yale University, New Haven CT)