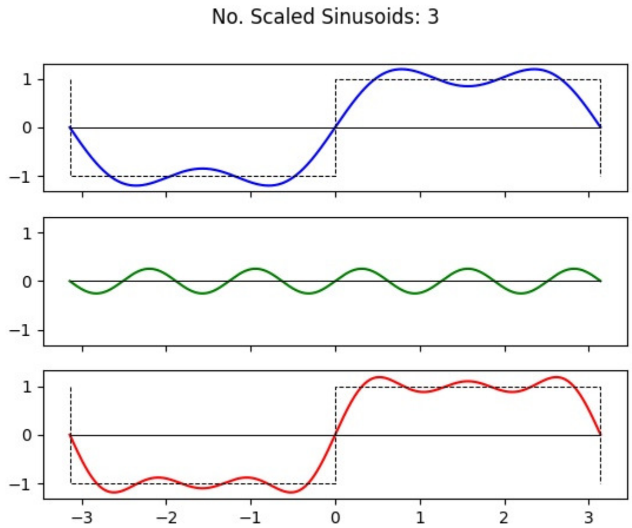

SIGNAL PROCESSING INSIGHTS INTO HEARING LOSS

- and lived experience of its mitigation

ABSTRACT

An arbitrary sound can be represented by summing its constituent pure-tone frequencies of different volumes. The hearing process uses this phenomenon, receiving information of volumes of frequency-specific signals from the cochlea known as a sound spectrum (singular) or spectra (plural). The brain stores memories of sound spectra, which it uses to recognise incoming sound. Elements of signal processing are presented to offer a context for explanations of how sound spectra are the key to understanding the effects of hearing loss, hearing aids and cochlear implants. The paper concludes with lived experience of using hearing aids and cochlear implants, with emphasis on rehabilitation techniques following an implant. Knowledge of the biology of the cochlea is limited, with some important details omitted in the interests of presenting a basic understanding of the signal processing of hearing.

Author and Notes

Chris Satchwell PhD

The author is a hearing aid and cochlear implant user, and engineer. Relevant background includes a working knowledge of some signal processing techniques and artificial neural networks.

Introduction

Many people with hearing loss seek a greater understanding of their condition. In the case of biological aspects, there are a number of good web sites which usually end at the point where the cochlea sends frequency-specific signals of their volumes to the auditory nerve. From that point onwards some knowledge of signal processing becomes useful. Usual descriptions pre-suppose a knowledge of advanced mathematics, which creates a barrier to understanding for most people. The present document attempts to offer a purely visual explanation, which aims to increase the depth of understanding of both hearing and how hearing-corrective devices work.

Basic Concepts of Sound

A good explanation of sound can be found at: https://en.wikipedia.org/wiki/Sound. To summarise: in air, sounds are longitudinal pressure fluctuations defined by their frequencies and amplitudes. The easiest sounds to define are pure tones which have a single frequency and volume.

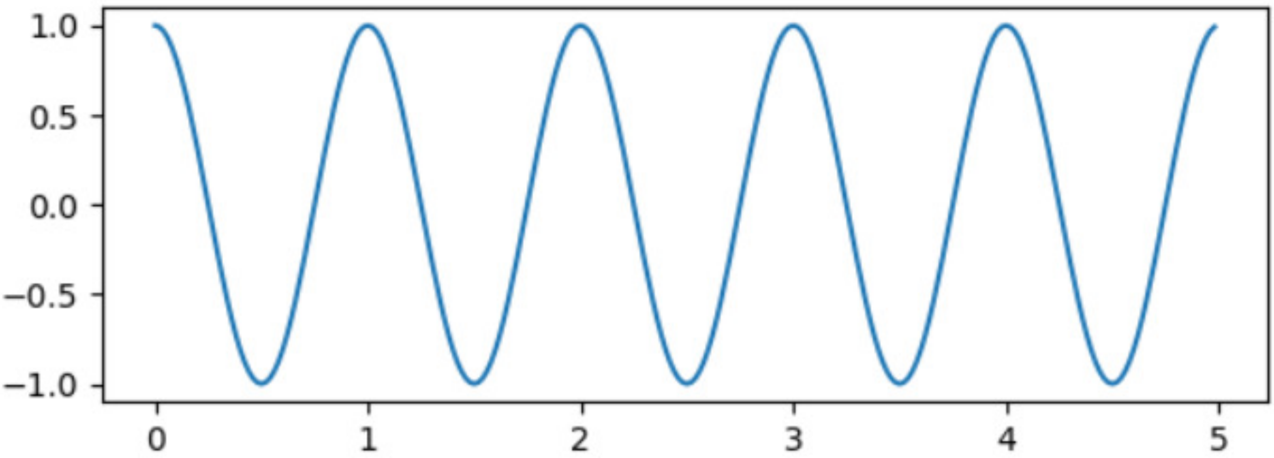

Fig 1 Plot of change-of-pressure v. time for a pure-tone sound

The name of the waveform of a pure tone is a "sinusoid" and the adjective relating to it "sinusoidal".

Later we will see that scaled pure tones are an essential part of signal processing because they can be appropriately scaled and then summed to form arbitrary complex sounds. Scaling factors are known as Fourier Coefficients, named after Joseph Fourier who first published the relevant mathematics around 1820. They can be thought of as a proxy for audio volume. Mathematically, any pure wave-form is a trigonometric sine function.

Understanding How the Brain "Hears"

To understand how the brain "hears" we first need to appreciate that signals sent from the cochlea can be related to scaling factors for multiple, pure-tone sinusoids, each of which is a component of an arbitrary sound waveform. We shall demonstrate this in Figures 2 to 7, leading to a bar chart of maximum pressure changes for each frequency in Figure 8 and its equivalent in decibels in Figure 9. This is a key result for understanding both how the brain "hears" and how hearing aids and cochlear implants work.

Firstly, we need to offer some visual insight into a contentious claim that a sum of scaled sinusoids can represent a non-sinusoidal arbitrary waveform; which in this example is a "square" wave. In part, this is chosen because details can be independently checked in Kreszig's "Advanced Engineering Mathematics". Details are beyond the scope of this document. The immediate objective is to offer a credible visual demonstration of how scaled, time-domain sinusoids can be summed to form a close approximation to an arbitrary non-sinusoidal waveform: to deliver a sound spectrum

The demonstration starts with the top graph of Figure 2, where the black dotted line shows a "square" wave as a "target" for a sum of scaled sinusoids to represent. In the series of Figures that follow, the blue line shows what we already have (in the case of Figure 2 nothing), the middle graph a green scaled-sinusoid to be added to the line above, with the result of their sum shown in the bottom graph as a red line. In subsequent figures, the top graph shows the result of summing all scaled sinusoids prior to the current frequency.

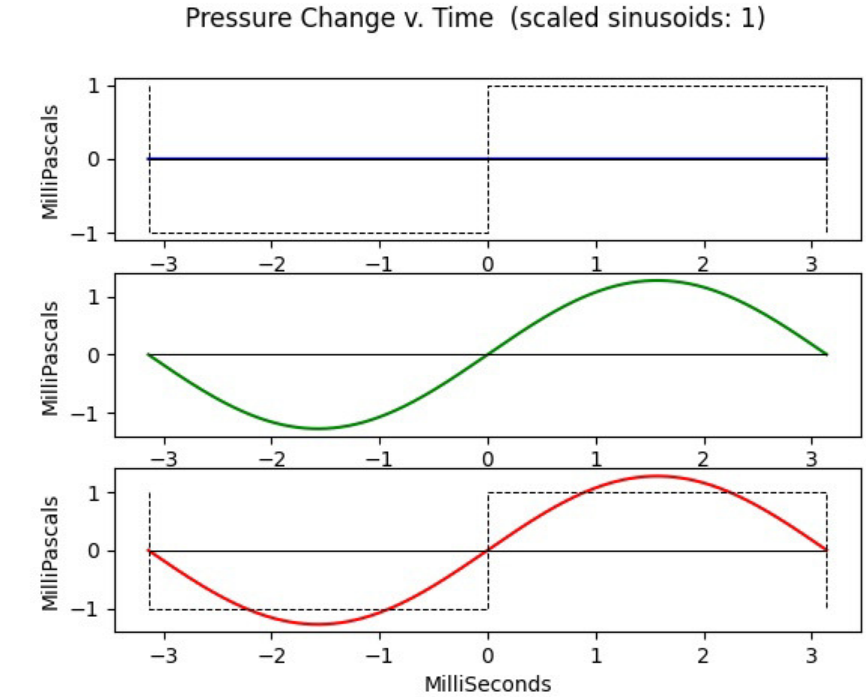

Figure 2 Initial sinusoid (green) is added to blue line (above) to produce the red line (lower graph).

In Figure 2, the single (red) sinusoid and square wave (dotted black line) are a long way apart.

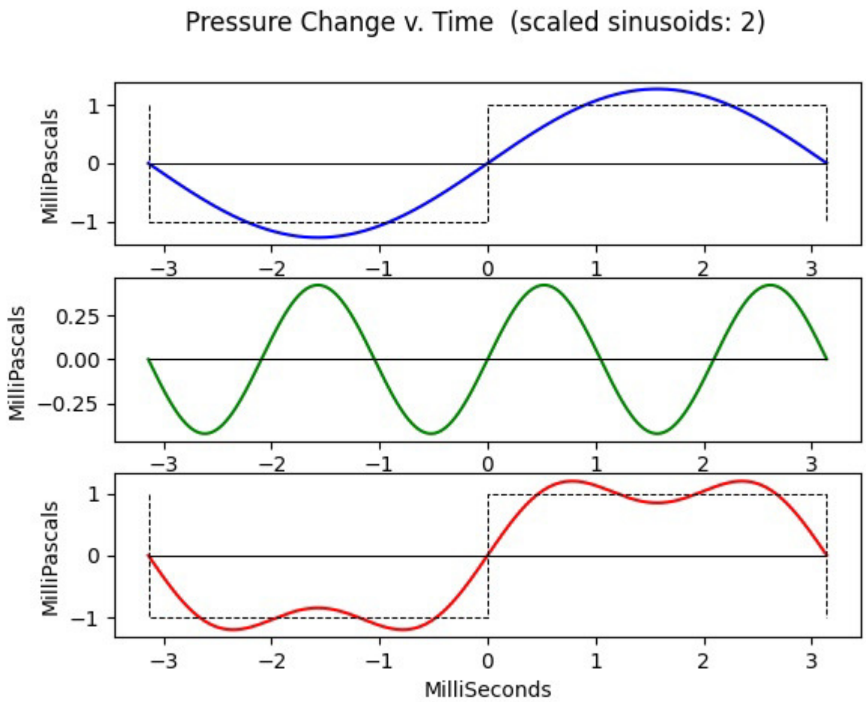

Figure 3: The next green line now has two complete cycles. When added to the result of Figure 2 (shown by the blue line above) the resulting red line produces a slightly better fit to the target square wave (dotted black line).

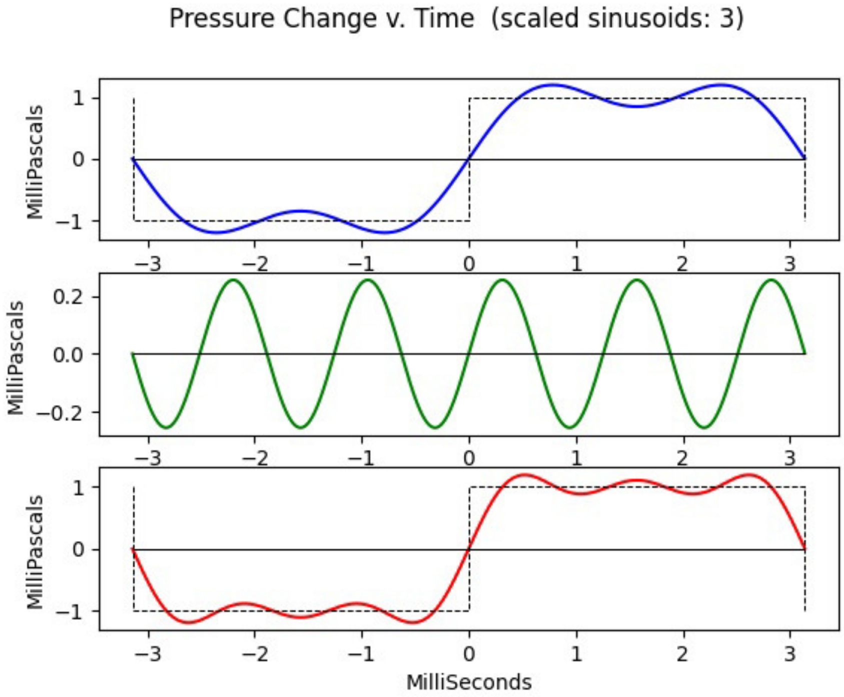

Figure 4: If a 3-cycle sinusoid is added to the result of fig 3, the (red line) fit to the square wave (black dotted line) improves.

The process of adding higher-frequency scaled sinusoids continues, with progressively closer fits to the target.

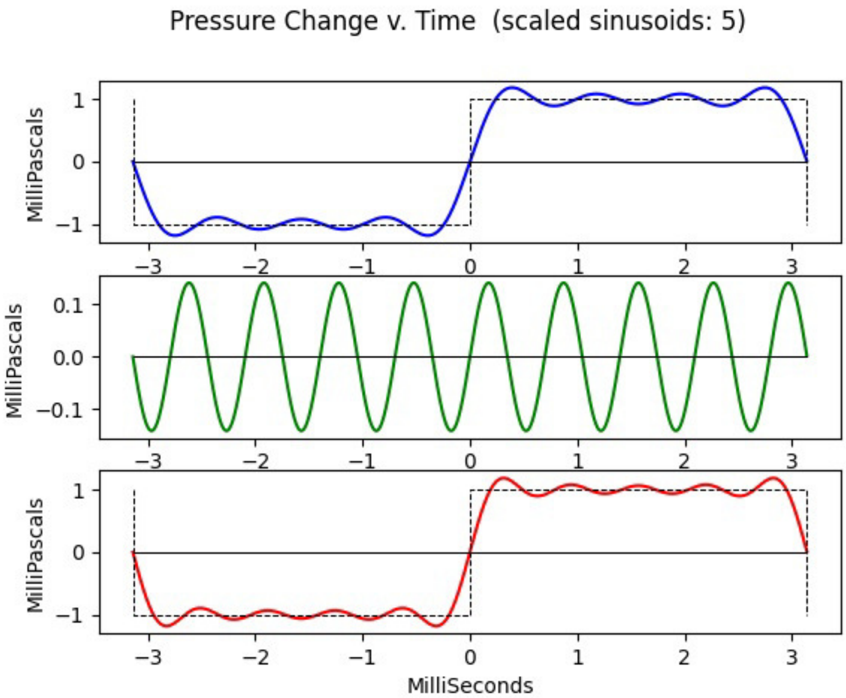

Figure 5: Result of adding five scaled sinusoids.

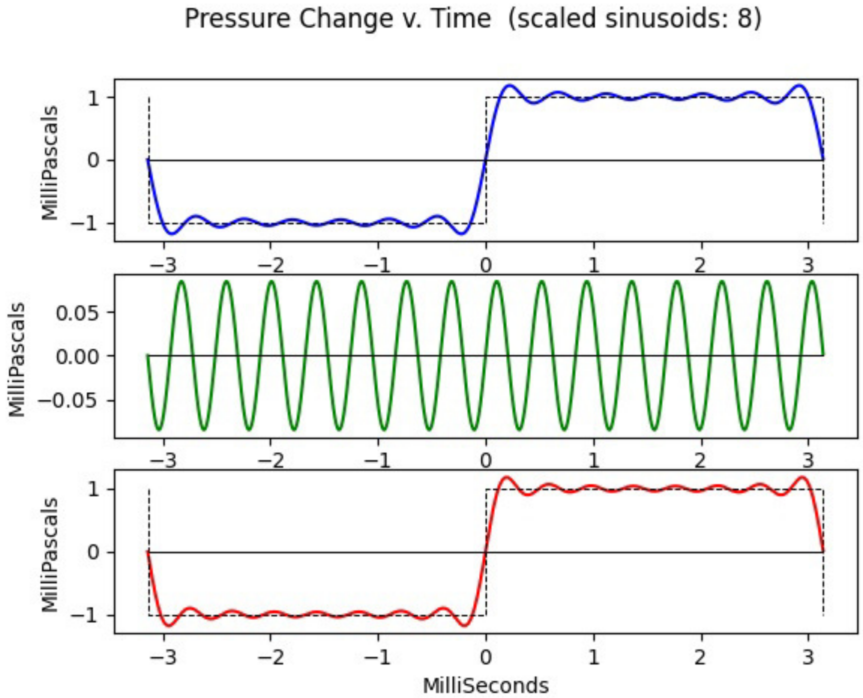

Figure 6: Result of adding eight scaled sinusoids.

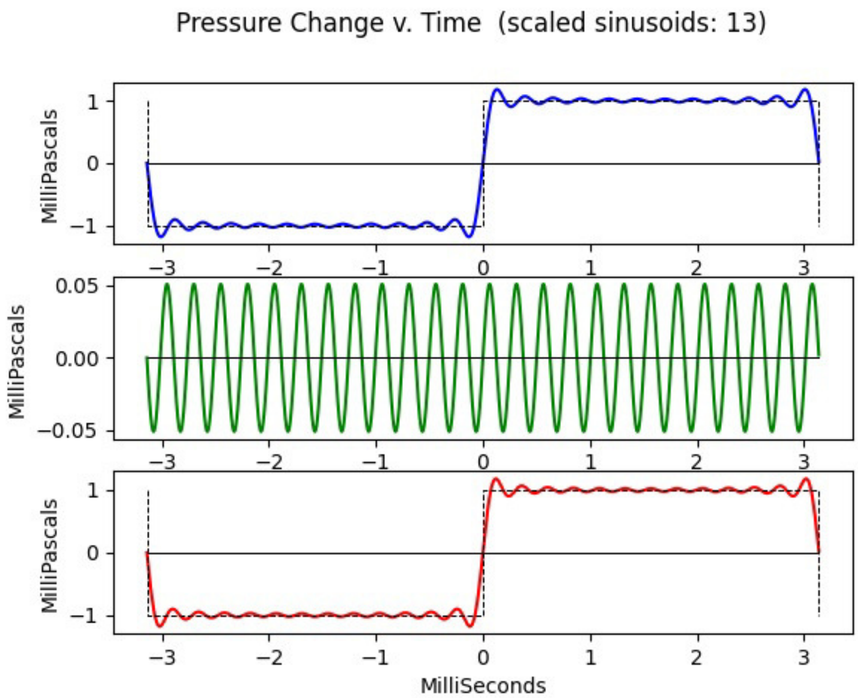

Figure 7: Result of adding thirteen scaled sinusoids. More scaled sinusoids produce an even closer fit.

To recap: the aim of the demonstration is to show that an arbitrary waveform (in this case a square wave) can be approximated by adding suitably-scaled sinusoids; i.e. if the scaling factors are known then the waveform can be reconstructed as it has been in this demonstration. This leads on to the next topic of the frequency domain and sound spectra.

Frequency Domain and Sound Spectra

So far, in figures 2 to 7, graphs of coefficient-weighted sinusoids have been plotted against a time axis, both individually and progressively summed. There is another way of looking at this data, which is important because it is the way the cochlea communicates with the brain. A time-dependent wave-form approximation is uniquely defined by frequency/scaling factor pairs; so instead of time, why not plot frequency on the 'x' axis and scale factors on the 'y'? This is called the frequency domain.

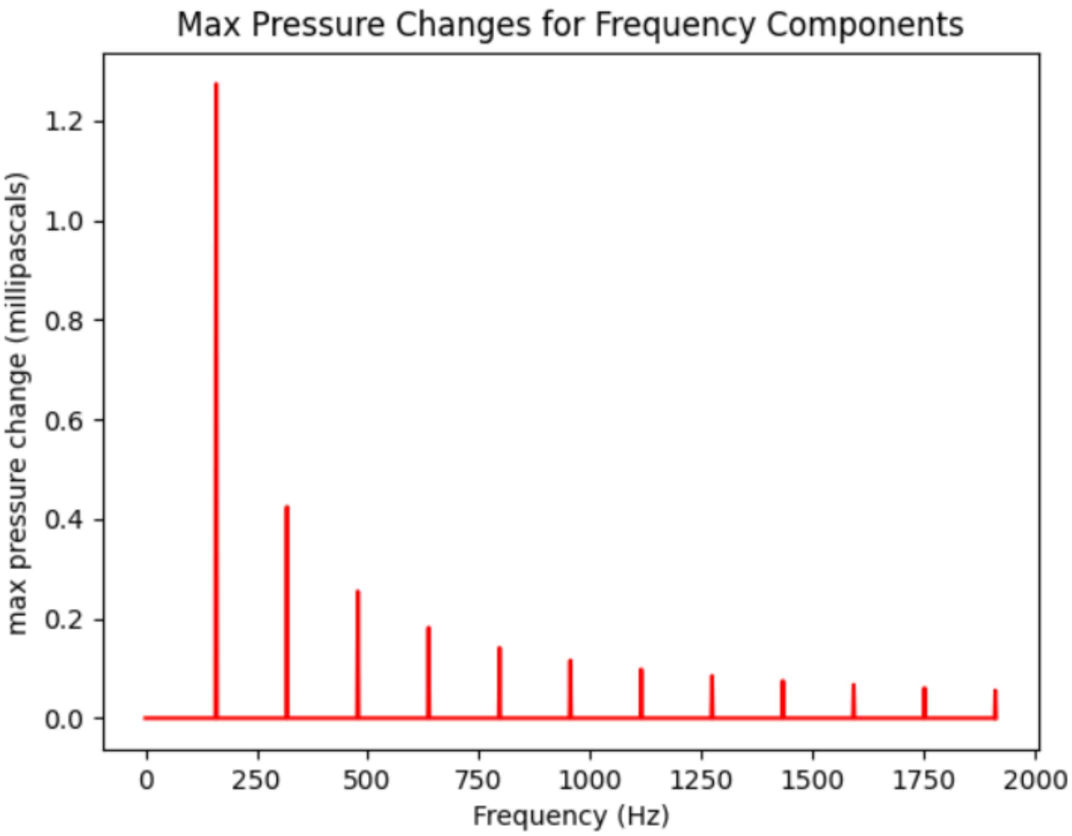

Figure 8: *A bar chart of scaling factors (maximum pressure changes) for different frequencies is an alternative way of describing a graph of pressure change against time (such as the dotted lines in figures 2 to 7).*

Note that the x axis is now in Hertz (Hz). From Fig 2, the first sine wave completed a full cycle in 2π milliseconds, making the scale factor to Hertz: 1000/(2π).

For audio signals information such as that contained in Fig 8 is called a sound spectrum, the plural of which is sound spectra.

Up until the point that sound reaches the ear, for the type of analysis just described, results need to be kept in units of pressure such as millipascals. In the brain, the sound is perceived as a logarithm of pressure, so a new unit, the decibel, is used to measure it. This is a tenth of the logarithm to the base 10 of the quotient of squares of root mean square of sound pressure (p) and a reference pressure (p_{ref}), which is commonly 20 microPascals. More can be found at https://en.wikipedia.org/wiki/Sound where a formula for decibels (dB) is given as $dB = 20 \log_{10}(p/ p_{ref})$. For sinusoids, the root mean square pressure (p) is 0.7071 of the maximum.

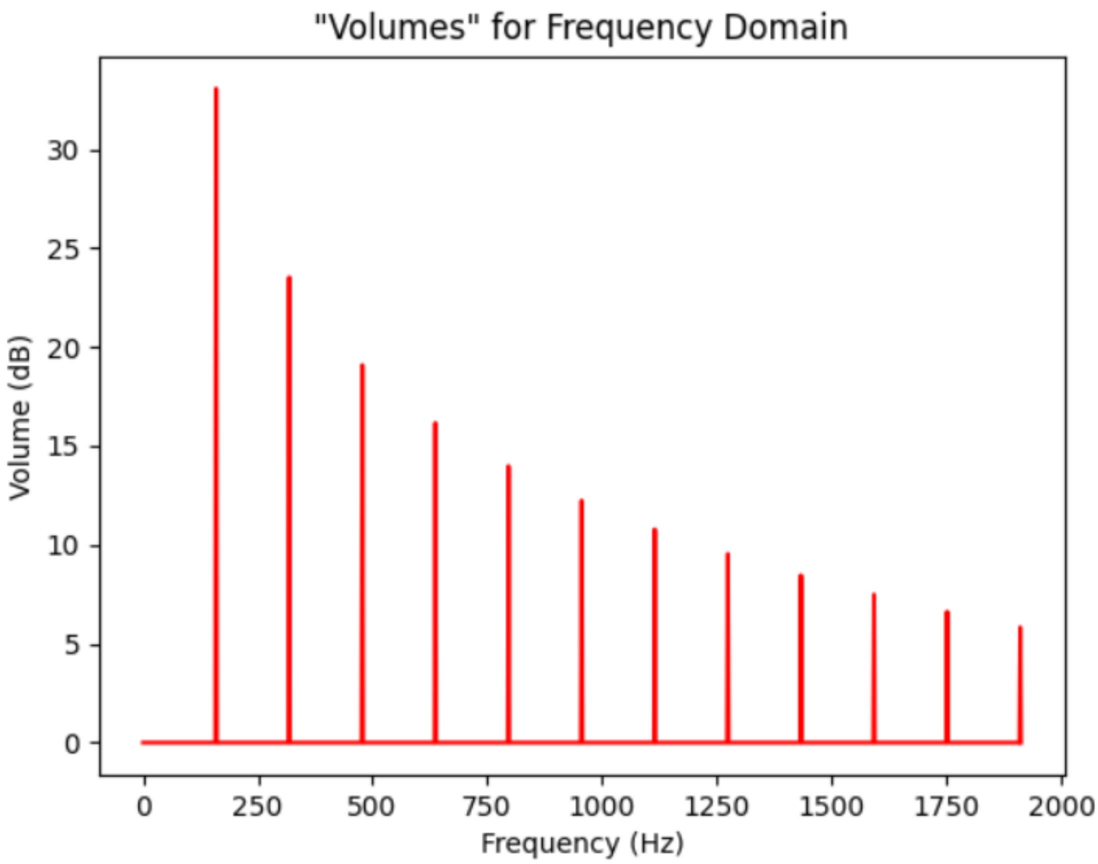

Fig 9: This frequency-domain bar graph (based on Fig 8) is in decibels (dB), which is a more useful measure in reflecting the brain's perception of loudness.

Its significance for hearing is that it reflects the way the brain stores memories of sounds (e.g. phonemes of speech) against which the inputs of sounds heard are compared for recognition. It is also a key to understanding how hearing aids and cochlear implants work.

The Cochlea

The cochlea is a complex organ whose details are described in https://en.wikipedia.org/wiki/Sound. What follows is a partial and a very basic description of its functionality: the "what" more than the "how". More accurate details can be found in literature on biology or audiology.

To recap from the earlier section on How the Brain Hears: "we first need to appreciate that signals sent from the cochlea can be related to scaling factors for multiple, pure-tone sinusoids, each of which is a component of an arbitrary sound waveform." i.e. sound spectrum data, analogous to the low-frequency domain information of the type presented in figure 8. The cochlea's logarithmic response is to send frequency-domain information (such as that in figure 9) to the brain.

The name "hearing" (or "auditory") nerve implies a single "wire" to the brain whereas the brain requires multiple frequency-specific signals relatable to information such as that in figure 9. Consequently the auditory "nerve" has the functionality of a loom of wires connecting the cochlea to the brain. Individual frequencies depend on which part of the cochlea a conductive component of the auditory nerve is connected to; with high frequencies at the start and low at the end.

The cochlea delivers constantly-changing, sound spectra to the brain.

The cochlea has a system of hairs designed to respond to different frequencies of sound vibration. These are connected to the auditory nerve via hair cells; which turn sound-induced vibrational energy into electrical signals.

These provide signals relatable to the Fourier-Coefficient, frequency-components of a sound. Individual frequencies are detected at different lengths along the cochlea, with high frequencies at the start and low at the end. I caution again that the true situation is more complex, but the net result are signals from which the brain can infer sound spectra, to compare with previously-memorised sound spectra, to recognise sounds.

Correcting Hearing

One common hearing problem is age-related hearing loss, where some of the 15000+ hair cells in the cochlea begin to fail. Figures 2 to 7 offer an analogy for a benchmark from which frequencies for some types of hearing loss can be examined.

In terms of this analogous visual demonstration: if frequencies above three were lost, the waveform detected would be that of Figure 4, which poorly defines its flat horizontal regions and vertical sides. Similarly in the frequency domain (Figure 9), all the bars after 3 would be missing. In the real-world case, an individual with high-frequency hearing loss would struggle to recognise consonants. Likewise, if any frequencies below three were lost there could be no indication of the basic form of the square wave. In practice, I suspect lower-frequency loss would lead to uncertainty in vowel sounds.

Providing the auditory nerve is intact the two main solutions are hearing aids and cochlea implants.

A hearing test tells an audiologist which frequency regions have impaired hearing. Providing there is residual usable hearing in those regions, they can prepare an aid which amplifies sound there, restoring some of the lost hearing. Nowadays, this involves a fast method for detecting the ever-changing sound spectra for incoming speech (Fast Fourier Transform) and applying appropriate, frequency-dependent amplification where needed.

As more hair cells are lost, the usefulness of aids declines and the next approach is usually a cochlear implant. The functionality of hair cells is replaced by an array of electrodes in contact with different regions of the auditory nerve. These replace the sound-spectra-related electrical inputs formerly provided by hair cells. Electrodes are wired to an aerial coil and magnet inside the skull above the ear. Externally, a microphone detects sound; sends it to a processor (which computes the constantly-changing spectra of incoming sound and electrical stimulation needed by the cochlea); sends information to a magnet-retained headpiece; which transmits it as radio waves to the aerial inside the skull, from where the current received in the aerial coil travels to appropriate electrodes; to stimulate the auditory nerve. One limitation of this is that electrodes leak current and can interfere with each other, so need to be well-spaced. Typically there are twelve to twenty-four electrodes attempting to do the job of around sixteen thousand hair cells. After implantation and switch-on, there is a lengthy period of rehabilitation. In the author's case, his implant provides an ability to hear a normal conversation in a quiet environment, but he struggles in cocktail parties or in other noisy environments. As an engineer I find it is amazing just how well just sixteen electrodes perform. A number of recent innovations such as Bluetooth are particularly useful, not least to facilitate the use of a mobile phone.

A cochlea implant can be normally be used to treat additional hearing problems when an auditory nerve is intact. One example is irreparable damage to a middle ear, preventing sound reaching the cochlea. In this case a cochlear implant can be used to by-pass the middle ear.

In terms of signal processing, correctional sound spectra can be added to account for any cochlear signals not mentioned so far and inbuilt programs designed to help in specific hearing situations such as echoes. Such programs are increasingly being invoked by artificial intelligence rather that the user.

Lived Experience

Hearing aids are "aids", not cures. They helped me in the earlier days of my hearing loss but never fully restored my hearing. Initially I was treated as a tinnitus patient and advised to wear an aid to reduce it, which it did. As my hair cells failed I went through a series of more powerful aids until arriving at a situation where there were too few hair cells left for them to be effective -- at which point a cochlear implant was the only viable alternative. One point about hearing aids is that they use the unchanged cochlea "wiring" from which sound-spectrum memories were created, which means that getting used to new aids is usually a relatively trivial matter.

A cochlear implant (CI) involves a meningitis vaccination, an operation, followed by an uncomfortable night in hospital and on-going rehabilitation. From memory, "switch on" was about three weeks after the operation and the start of rehabilitation. One difficulty with a CI is that the cochlea is, in effect, re-wired so that the memories of sound spectra have to be re-learned. Initial advice was that for a young person this would take about a year and for an older person, two. The optimal method of rehabilitation differs widely between patients so I caution that what I am about to describe may not be suitable for everybody. A close mathematical analogy to the CI rehabilitation problem is something called "targeted learning" in artificial neural networks. This involves trying to adjust numeric weights (analogous to chemical inhibitors in synapses that the brain uses to form memories) to match a known target. A CI-learning equivalent is to watch TV programs <u>of interest</u> where a digestible package of sub-titled words arrive <u>before</u> their sound. Documentary programs are usually good candidates. Subtitles are read, forthcoming words to be spoken are known before they are, and so <u>associations between phonemes and sounds are easily made</u>. An additional advantage is that lip-reading can now be used to help match sounds to known words, rather than to try to infer unknown words. By removing uncertainty and any need for acoustic closure, the effectiveness of learning new sound spectra is increased and fatigue greatly reduced, making longer periods of learning possible. In contrast, with news programs, subtitles appear in less-digestible packages and follow (not lead) speech. Post-speech subtitles create a mental overload situation of trying to match past speech to current subtitles, whilst at the same time having to remember a flood of current speech for future matching. This situation is much less effective, but also mirrored in some CI rehabilitation advice that was prevalent when I had to re-learn sound spectra. During my CI learning, TV sound received was via a radio aid, which cut out any corrupting environmental sounds. Currently, a Bluetooth TV connection would be a good way of avoiding environmental noise.

Another consideration was to try to rehabilitate as fast as possible. Once again artificial neural networks gave a pointer. Training requires a certain number of iterations but says nothing about the time over which it takes. This is one reason why the words "of interest" are underlined in the previous paragraph. If there is no interest, boredom sets in and effective training suffers. Equally, acquisition of words through sounds and lip-reading is much more difficult that reading pre-speech sub titles. With today's large range of TV channels, it is mostly possible to TV-train on what you want when you want. My rehabilitation at the age of seventy was both interesting and painless. In terms of the total amount of equivalent-effective training, I doubt it differed from the average, but it was possible to achieve it painlessly and in a much shorter time frame. An additional factor was the "snowball" effect of targeted learning. This advanced the acquisition of phoneme sound-spectra memories, which helped the inference of normal speech from sound and lip-reading, which in turn tended to create more opportunities for intelligible conversations; creating a virtuous learning circle. I suspect the effectiveness of targeted verses untargeted rehabilitation could be a useful research topic and my use of the term "equivalent-effective training" betrays my belief that targeted training is more effective, particularly in the early stages of rehabilitation when sound-spectra memories are least developed.

On the basis of speech hearing tests, after nine months, I was advised I was in the top five percent of CI users. I believe this speedy arrival was down to the quality and quantity of the retraining of my sound-spectra memories and not to any personal attributes.