# Ambient noise-cancellation for headphones and handsets

1. Introduction

Silence is a very precious commodity. Unfortunately, we cannot control it. During the course of the day, many of us become immersed in noisy environments, and we have grown accustomed to the presence of background noise as being an unavoidable element in our lives. At very low levels, ambient noise can be tolerable, but at higher levels, background noise can be distracting and annoying, or even physiologically damaging to our hearing mechanisms. Noise degrades our ability to talk and communicate with others, and it interferes with the ability to think clearly, and to relax, and to enjoy alternative sources of audio, such as recorded music or the radio. Wouldn't it be perfect if we could simply throw a switch and create our own personal zone of silence?

In principle, this should be possible. Perfect electronic noise-cancellation for earphone users has been a "holy grail" for many years now, in which ambient noise is detected and used to create an "inverse" noise signal at the ears of the listener, via the earphones themselves, thus causing destructive interference and a reduction in the perceived noise level. In practise, however, it is not so simple, because differences in the ear-related transfer functions between the noise signal and the cancellation signal restrict the frequency range and degree of noise-cancellation that can be achieved.

2. Principles and limitations

The basis of electronic ambient noise-cancellation is destructive wave interference between the unwanted noise signal and a synthesised "cancellation" signal that is created to be equal in magnitude to the noise signal, and opposite in phase to it. This is shown in Figure 1a, overleaf, which depicts a sine-wave noise signal and its cancellation counterpart, and also the remaining, residual noise that results from the interaction. Here, in this first Figure, the cancellation is perfect, and hence the residual signal is zero.
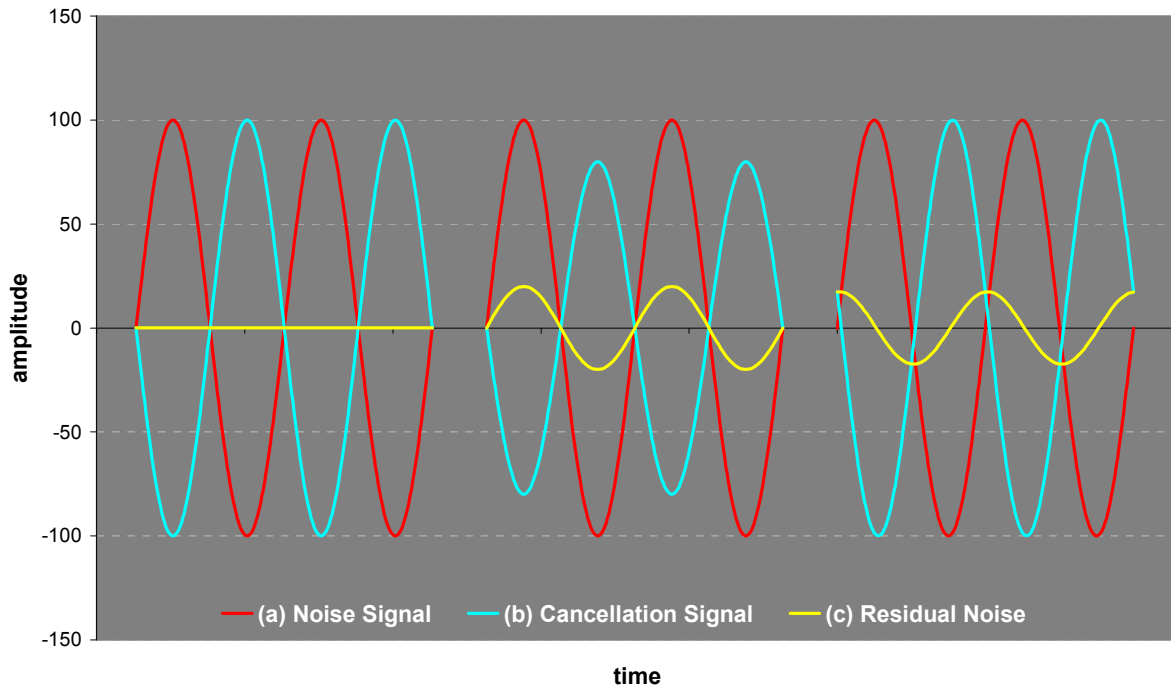

Figure 1: Effective noise-cancellation requires accurate alignment of the noise signal to the cancellation signal

At first, this seems a simple enough proposition, but it is interesting to consider what would happen to the residual noise signal if the amplitude and phase characteristics were *not* perfectly matched. This can be calculated readily from first principles, as described later in Appendix A, from which it is possible to quantify the amount of noise cancellation that could be achieved where the amplitude and phase properties depart from the ideal. For example, assuming that the phases were perfectly matched, then if the amplitude of the cancellation signal were 20% smaller than the amplitude of the noise signal, then the cancellation would be incomplete, and a considerably large residual noise signal would remain, as shown in Figure 1b. Conversely, if the amplitudes of the two signals were perfectly matched but there was a 10º *phase* mismatch, then the cancellation would, again, be imperfect, and consequently the residual noise signal, again, would be relatively large, as shown in Figure 1c.

Clearly, both the relative amplitude and phase of the cancellation signal (with respect to those of the noise signal) are critical elements in the noise-cancellation process. A simple equation can be derived (Appendix A) to define the magnitude of the residual signal, and its simultaneous dependence on variations in both amplitude and phase matching. The result, shown overleaf in Figure 2, is depicted as a three-dimensional surface graph, in which this functional dependence can be seen.

Here, the residual noise (y-axis, as a percentage of the original noise level) has been plotted as a function of both amplitude (x-axis) and phase (z-axis) differences between a noise signal and its cancellation counterpart, with the perfect match lying at the centre where the (negative) cancellation signal is equal to 100% of the noise signal and their relative phase difference is zero.
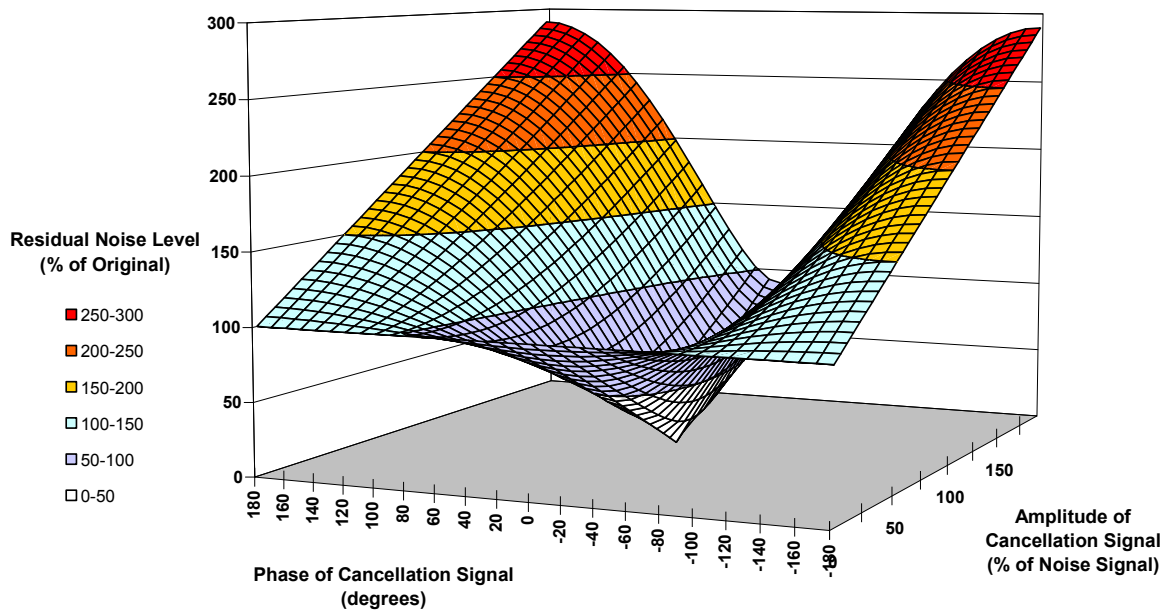
Figure 2: Noise-cancellation surface, showing sensitivity of cancellation on amplitude (z-axis) and phase (x-axis) alignment

The >50% cancellation region (-6 dB or better) is represented by the lowermost, white region of the very narrow funnel shape descending centrally to the floor of the plot. What becomes clear from this graph is that the noise-cancellation mechanism fails unless both the amplitude *and* phase properties of the signals are closely matched, simultaneously, and this condition must prevail throughout as wide a range of frequencies as possible. The graph also shows that, under certain conditions, the residual noise level can actually *exceed* the original noise level if the cancellation signal phase is sufficiently displaced.

Very tight signal matching is needed for even a modest amount of noise-cancellation. If, say, 65% cancellation (-9 dB) is to be achieved (that is, residual noise signal would be 35% of the original amplitude), then, assuming perfect phase matching, the amplitude of the cancellation signal must be matched to that of the noise signal within ±3 dB. Similarly, even if the amplitudes were perfectly matched, the relative phase of the signals must lie within ±20° for -9 dB cancellation, which, at 2 kHz, corresponds to a time period of only 28 μs, and represents an acoustic path length of only 10 mm. Consequently, it will be appreciated that even very small discrepancies in acoustic signal path lengths (10 mm or so) can impair the noise-cancellation process significantly.

3. Feedback and feedforward noise-cancelling systems

There are two different approaches for providing ambient noise-cancellation for headphones and earphones, based respectively on a "feedback" method, and on a "feedforward" technique, depicted here in much simplified form to illustrate the principles, in Figures 3a and 3b, respectively. Both of these methods have their merits and drawbacks.
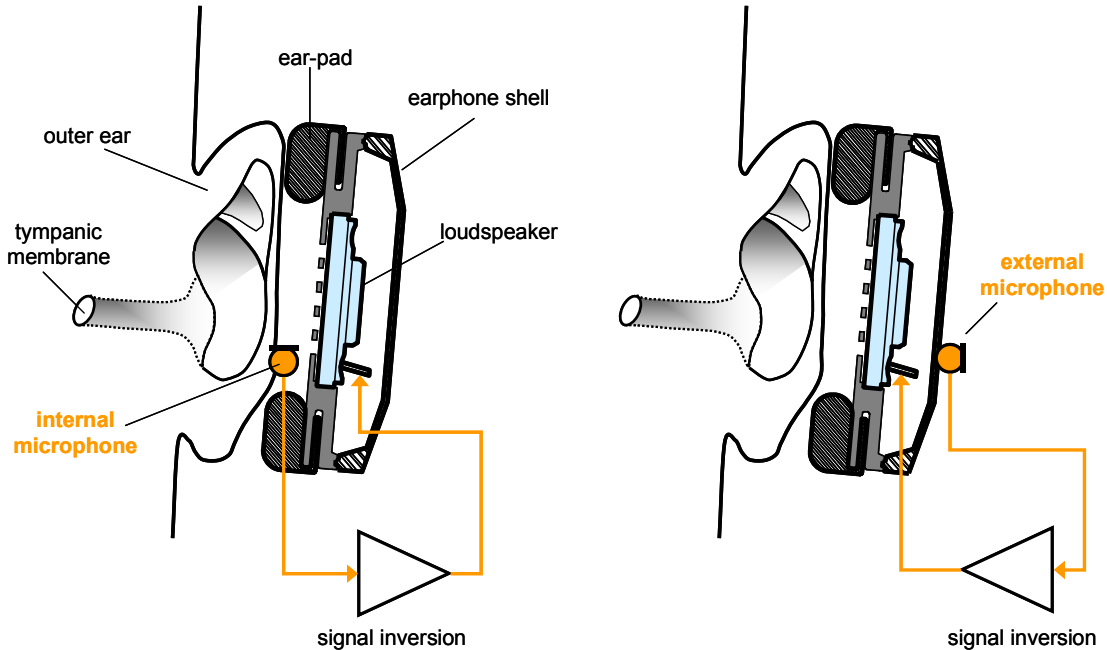
Figure 3: (a, left) Feedback, and (b, right) feedforward noise-cancellation configurations (simplified)

The <u>feedback</u> method is based upon a closed-back, circumaural-type headphone system. Inside the cavity that is formed between the ear and the inner, frontal area of the headphone unit, a miniature microphone is placed directly in front of the headphone loudspeaker (Figure 3a), and it is coupled back to the loudspeaker using a negative feedback loop such that it forms a simple servo system in which the loudspeaker is constantly attempting to create a null sound pressure level at the microphone. Although this principle is simple, there are some practical problems relating to the intrinsic phase response of the loudspeaker and to the propagation delay between the speaker and microphone, which both introduce a phase lag at higher frequencies. Consequently, high-frequency filtering must be introduced into the feedback loop, which tends to restrict the upper frequency of operation to around 1 kHz or below. This means also that effective passive acoustic attenuation must be provided to prevent the ingress of ambient noise above this 1 kHz limit, and this is done by ensuring that the ear-enclosing headphone seal blocks these frequencies. Furthermore, when music or speech is fed to the user's earphone, then provision must be made to avoid these particular signals being cancelled out by the feedback system, and this process can introduce "colouration" of the music signal.

Feedback systems can be effective at cancelling low-frequency noise, but the user can feel isolated and blocked from the environment by the acoustic seal, which also amplifies "self-noises" perceived by the user – their voice is heard as loud and distorted, and the noises of chewing and eating are no longer quiet. The requirement for an effective ear-seal also tends to make feedback systems bulky.

These attributes do not lend themselves well to mobile device applications, where earphones must be lightweight and small, and permit user awareness of the environment, and where electronic activation and control of the noise-reduction system can be provided. These attributes can be conferred by the use of the feedforward noise-cancellation method, which uses a different principle, and is depicted in basic, simplified form in Figure 3b. In contrast to the feedback system, a microphone is placed at the exterior of the headphone shell in order to detect the incoming, ambient noise signal, which is then inverted and added to the headphone drive signal, thus creating the cancellation signal. As a consequence, destructive wave-cancellation occurs between the cancellation signal and the incoming ambient acoustic noise signal, adjacent to the headphone loudspeaker outlet port within the cavity between the headphone shell and the outer ear.

The feedforward system does not require such a well-sealed cavity around the ear, and so it can be arranged that the user "hears through" the 'phones, and retains some degree of spatial hearing[1], such that there is no "blocked off" feeling – the headphones appear more acoustically "transparent". Also, music playback quality is not impaired because there is no enclosed cavity around the ear, and hence there are no voice or eating artefacts, either. The headphones can be designed to be lightweight and thin. In short, the feedforward system can provide a more natural user experience, well suited for mobile applications and for use whilst travelling.

The feedforward method of ambient noise cancellation is very simple to implement in a basic form, but current systems are far from perfect and generally limited in their effectiveness for several reasons. One fundamental difficulty is caused by the differing acoustic pathways that the noise [N] and cancellation [C] signals travel, as shown in Figure 4, below, which confer differing frequency-dependent amplitude and phase variations on the two signals. This requires careful signal-processing to re-balance and match the two signals insofar as is possible. Unfortunately, the use of conventional digital methods is ruled out because even a single A-to-D sampling period might introduce, say, a 30 µs delay, and this alone would disturb the phase alignment of the signals very significantly.

There are two further major difficulties in trying to time-align the signals. Firstly, the acoustic path length difference between the ambient-to-ear and ambient-to-microphone varies according to the direction of the noise source. For example, if the microphone inlet lies at the centre of the outer shell, then a wave-front arriving from a frontal (or rearward) source would arrive at the microphone and at the ear more or less at the same time, whereas a wave-front arriving from a lateral source would arrive at the microphone first, but then has an additional path of three or four centimetres to travel to the ear, corresponding to time periods of 87 µs to 117 µs, and this would cause large phase differences that would prevent effective noise-cancellation occurring. This alone is bad enough, but it is compounded by the second problem: the response time of the loudspeaker, which is likely to be 90 µs or more.
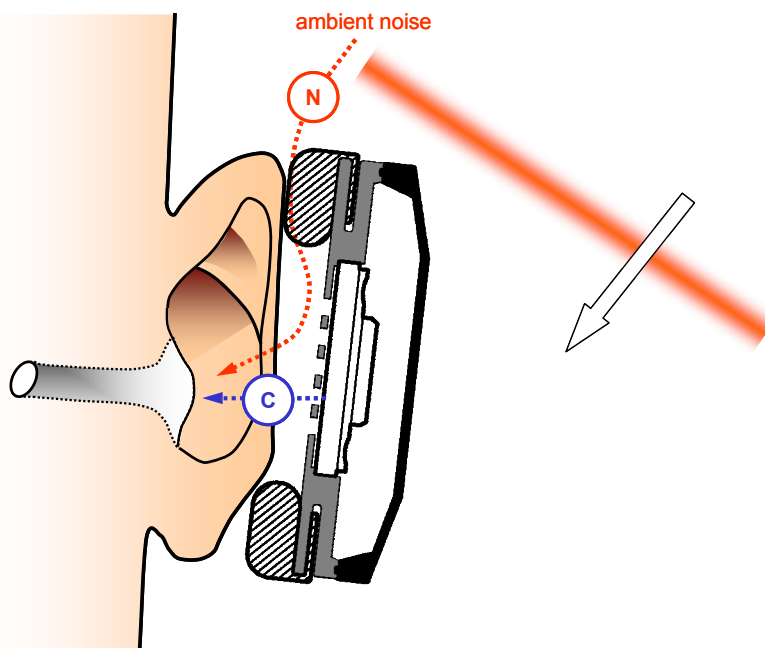
**Figure 4: Acoustic pathways to the ear of the noise signal, N,
and the cancellation signal, C**

4. Wolfson time-aligned noise-cancellation

A new method has been developed of overcoming these timing problems so as to provide a time-aligned feedforward noise-cancellation system. The invention is the use of a distributed array of microphones around the perimeter of the headphone shell – a "ring microphone" array[2], as shown in Figure 5, overleaf. Each microphone represents a notional noise-leakage entry point around the shell, and hence the sum of their signals represents the summed sound pressure level around the shell, which is the driving force behind the incoming leakage. Because the headphone shell acts as a baffle, the acoustic leakage pathway from ambient-to-eardrum is forced to traverse one-half of the diameter of the earphone assembly before reaching the auditory canal axis. Accordingly, by placing microphones at the rim of the headphone, the ambient noise signal can be acquired and driven to the loudspeaker *in advance* of its arrival at the eardrum, thus compensating for the intrinsic response time of the loudspeaker. Furthermore, because of the distributed geometry of the microphone array, this applies to wave-fronts arriving *from all directions*, and hence the system is largely direction independent.
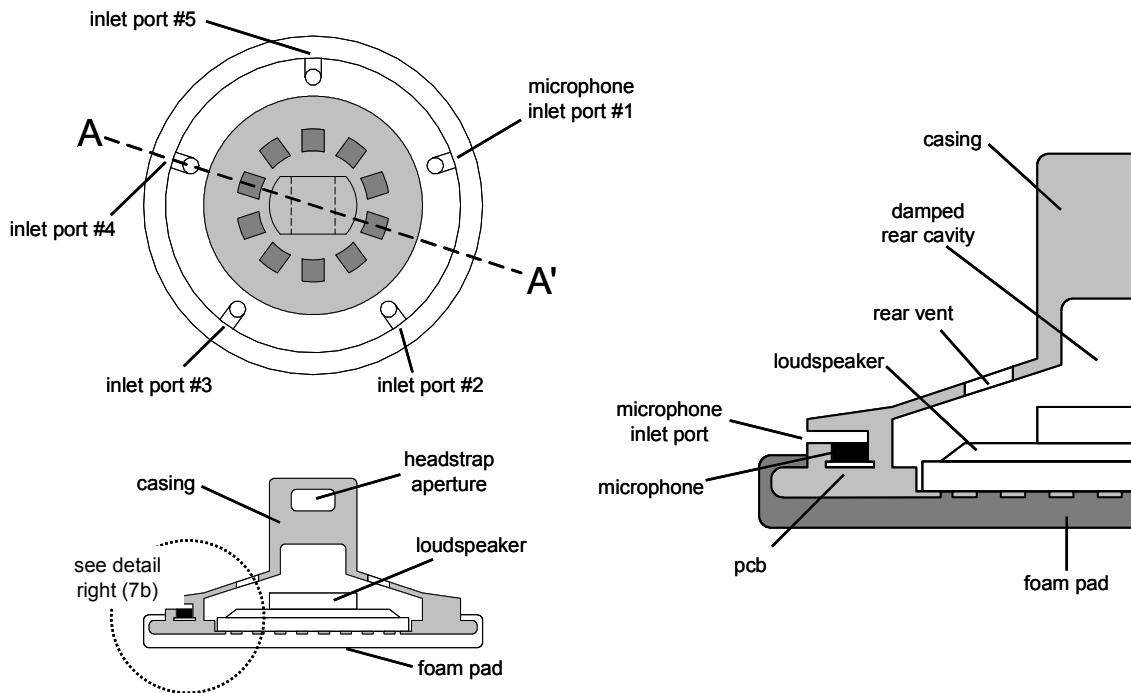
Figure 5: Headphone with distributed-microphone array

Figure 5 shows a lightweight headphone design based on this invention, in which a five-microphone ring array can be seen around the rim of each headphone shell assembly. Products based on this technology will be available in 2008, featuring effective noise-cancellation over a wide frequency range, typically 20 dB cancellation in the range 80 Hz to 3 kHz.

5. Handset time-aligned noise-cancellation

This time-alignment noise-cancellation technology has also been applied to mobile phone handsets in order to generate a noise-cancellation signal via the handset's inbuilt earphone, thus creating a zone of relative silence at the listener's ear, and improving the intelligibility of conversation. Figure 6 shows a dual-microphone array built into the lateral edges of a handset. This has proved to be remarkably effective in use, and enables the user to understand an incoming voice call in a noisy listening environment, such as an airport departure area.
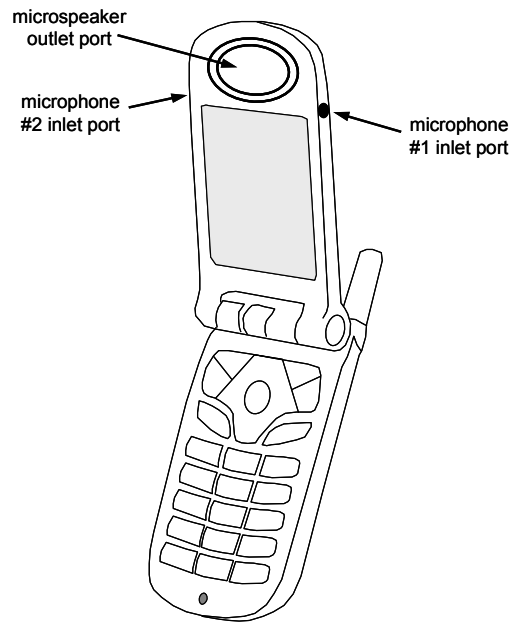
Figure 6: Cellular phone handset featuring multi-microphone, time-aligned noise cancellation

The improved intelligibility can be quantified as an "Articulation Index" factor (AI), representing the percentage of words that a listener would understand. When the external noise level is loud, and similar to the voice level at the listener's ear, an articulation index improvement of up to 28% has been achieved, with more than 20% improvement in AI over a wide range of ambient noise levels.

REFERENCES

1. Hearing in three dimensions
    A Sibbald
    Ingenia, August 2000, 5, pp. 47-50

2. UK Patent GB 2,434,708 B
    Ambient noise reduction arrangement
    A Sibbald

Appendix A

Sensitivity of Noise-Cancelling Signals to Variations in Amplitude and Phase

The principle of noise cancellation requires the addition of a cancellation signal, $C_{(t)}$, to a signal containing the noise waveform, $N_{(t)}$, such that the cancellation signal is equal in magnitude, inverted, and in-phase with the unwanted noise signal (that is, equal in magnitude and with a 180° phase difference). When this occurs exactly, superposition of the waveforms results in destructive, total cancellation of the noise signal. When this condition is not met, because the cancellation signal has a different magnitude or phase value (or both of these), then the cancellation is incomplete. Indeed, under certain conditions, the noise signal can actually be increased, rather than reduced or cancelled. In order to achieve a specific value of noise reduction, then limits on the deviations from the ideal values of amplitude and phase can be calculated as follows.

At any particular frequency, the noise waveform, $N_{(t)}$, having modulus N, can be represented by the familiar equation:

$$N_{(t)} = N.\sin \omega t \qquad\qquad \text{(A1)}$$

Similarly, the noise cancelling waveform, $C_{(t)}$, having modulus C, can be represented by:

$$C_{(t)} = -C.\sin \omega t \qquad\qquad \text{(A2)}$$

(where the negative term represents the required signal inversion of the cancellation signal).

The addition of these two waveforms creates a resultant, residual signal, $R_{(t)}$.

$$R_{(t)} = N.\sin \omega t - C.\sin \omega t \qquad\qquad \text{(A3)}$$

For perfect noise cancellation, the value of C must be equal to N, and, consequently, $R_{(t)} = 0$. This represents an ideal condition. In practise, however, the determination of a value for C at any particular frequency will involve experimental errors in both amplitude and phase. In order to assess their influence, let the phase of the cancellation signal incorporate an error component, ϕ. Expression (2) then becomes:

$$C_{(t)} = -C.\sin(\omega t + \phi) \qquad\qquad \text{(A4)}$$

This expression allows the amount of noise suppression to be quantified when there are phase and amplitude errors present in the cancellation signal, compared to the ideal values that would otherwise result in total cancellation. The residual noise signal, expressed as a fraction, F, of the original noise amplitude, is:

$$\frac{R_{(t)}}{N} = F = \sin \omega t - \left\{ \frac{C}{N}.\sin(\omega t \pm \phi) \right\} \qquad \text{(A5)}$$

It is the magnitude (modulus) of the residual signal that is perceived by the listener, and not its phase, which, therefore, is irrelevant here. The modulus must be calculated by a vector addition of the two combining waves. This can be used to create a surface plot of the residual noise fraction as a function of both C and the phase error ϕ. (Note that for any given frequency, perfect cancellation is achieved only when E = 1 and ϕ = 0 simultaneously.)
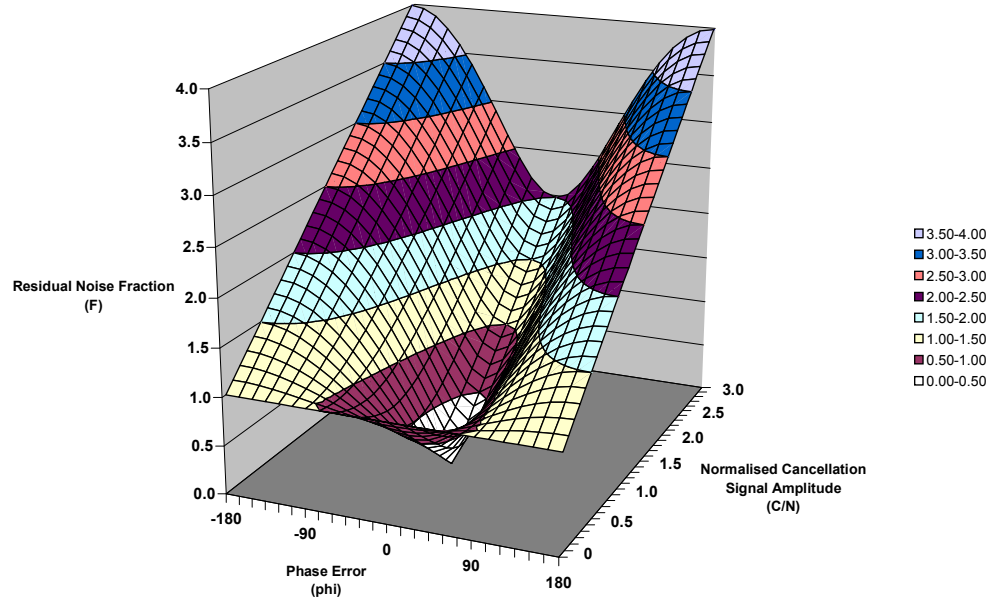
$$F_{in-phase} = 1 - \left\{ \frac{C}{N}.\cos \phi \right\} \qquad \text{(A6)}$$

Addition of the quadrature signal components yields the quadrature amplitude of the residual:

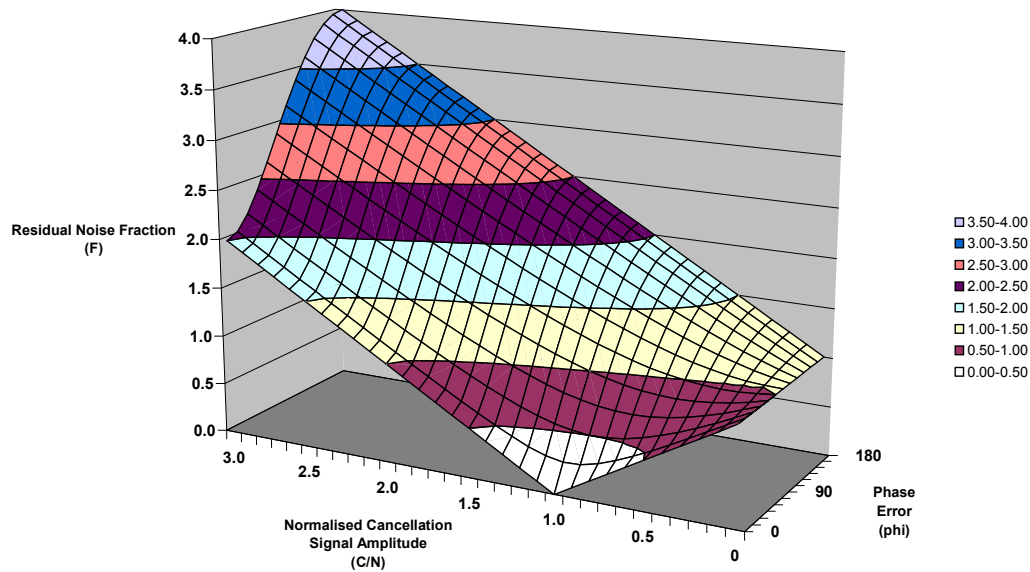$$F_{quadrature} = - \left\{ \frac{C}{N}.\sin \phi \right\} \qquad \text{(A7)}$$

…and this allows the modulus of the residual fraction, F, to be calculated:

$$F = \sqrt{ \left(1 - \left\{ \frac{C}{N}.\cos \phi \right\} \right)^2 + \left\{ -\frac{C}{N}.\sin \phi \right\}^2 } \qquad \text{(A8)}$$

Figure A1: Influence of amplitude and phase variations on residual noise

This now makes it possible to plot F as a function of both C (expressed as a fraction of the ideal value, N) and ϕ, as shown in Figures A1 and A2, below.

Figure A2: Influence of amplitude and phase variations (detail)

The important information that these plots reveal is the critical nature of the accuracy of the noise cancellation signal. If even a modest amount of noise cancellation is to be achieved, say at least 50% of the noise must be suppressed, then Figures A1 and A2 show the limited area in which this can be achieved, indicated by the white area at the very base of the inverted cone surface.

For 50% cancellation, if there were no phase error whatsoever in the cancellation signal, then the amplitude must lie within ±50% of the ideal value at all times (from -6 dB to +3 dB). Similarly, if there were no amplitude error, then the phase must lie within the range ±30°.

If a more substantive amount of noise cancellation were to be achieved, say a 9 dB reduction (which still represents a residual noise fraction of 35%), then, for perfect phase alignment, the cancellation signal amplitude must lie within the range –3.7 dB to +2.6 dB of the ideal value at all times, and, simultaneously, the its (assuming perfect amplitude values) must lie within the range ±20° (0.35 radians) of the ideal value.

In practise, of course, both amplitude and phase errors are likely to be present simultaneously, which compounds the problem.

It is important to note, in Figures A1 and A2, that there are areas where the residual noise index can actually exceed unity. For example, where the cancellation signal is too large and it is being added to the noise signal with a large phase error, such that it is more in-phase than out-of-phase. In these cases, the resultant noise signal is actually _greater_ than the original noise itself.

In summary, substantive noise cancellation (better than -9 dB) requires that the cancellation signal is matched to the noise signal with an amplitude accuracy of about ±3 dB, and, simultaneously, a phase accuracy of better than ±20° (0.35 radians).