

Modelling the Distortion produced by Cochlear Compression

Roy D. Patterson

Centre for the Neural Basis of Hearing, PDN, University of Cambridge, Downing Site,
Cambridge, CB2 3EG, UK

D. Timothy Ives

Ecole Normale Supérieure, Dept. d'Études Cognitives, 29 Rue d'Ulm, 75005, Paris, France

Thomas C. Walters and Richard F. Lyon

Google Inc., 1600 Amphitheatre Parkway, Mountain View CA 94043 USA

Corresponding author: Roy Patterson, rdp1@cam.ac.uk

Abstract

Lyon (2011) has described how a cascade of simple asymmetric resonators (CAR) can be used to simulate the filtering of the basilar membrane, and how the gain of the resonators can be manipulated by a feedback network to simulate the fast-acting compression (FAC) characteristic of cochlear processing. When the compression is applied to complex tones, each pair of primary components produces both quadratic and cubic distortion tones (DTs), and the cascade architecture of the CAR-FAC system propagates them down to their appropriate place along the basilar membrane, where they combine additively with each other and any primary components at that frequency. This suggests that CAR-FAC systems might be used to study the role of compressive distortion in the perception of complex sounds *and* that behavioural measurements of cochlear distortion data might be useful when tuning the parameters of CAR-FAC systems.

Key words: quadratic distortion tones, cochlear compression, propagation of distortion products

1. Introduction

In a classic paper, Goldstein (1967) used a cancellation-of-beats technique to measure the magnitude of the DTs produced by a pair of primary sinusoids. More recently Pressnitzer and Patterson (2001) used the same technique to measure the spectrum of low frequency, quadratic distortion tones (qDTs) produced by multi-harmonic complexes with consecutive harmonics of a 100-Hz fundamental (F_0), starting with the 15th harmonic. In their first experiment, the tones had 11 primaries in cosine-phase. The sound level was 54 dB SPL per component; the overall level was 65 dB SPL. They measured the level of the DTs at the first four harmonics of F_0 and showed that qDT magnitude decreased as harmonic number increased. In their third experiment, they demonstrated that qDT magnitude was strongly dependent on the relative phases of the primaries; when the phases of successive primaries were alternated between 0 and 90 degrees, the magnitude of the first and third qDTs decreased, whereas the magnitude of the second and fourth qDTs was unchanged, or increased slightly. The CAR-FAC system of Lyon (2011) was found to produce distortion tones with similar properties insofar as it produced DTs at the first four harmonics when the

primaries were in cosine phase and the magnitudes of the DTs at the first and third harmonics decreased when the phase of adjacent primaries was alternated between 0 and 90 degrees.

In their main experiment (the second), Pressnitzer and Patterson (2001) measured the magnitude of the qDT at 100 Hz as they increased the number of primaries (NP) in a cosine phase tone. Tones with 2, 3, 5, 9, or 17 harmonics were used to evoke, respectively, 1, 2, 4, 8, or 16 first-order difference tones between adjacent pairs of primaries in the complex tone. The magnitude of the qDT at 100 Hz was observed to increase by close to 3 dB per doubling of the number of first order difference tones, indicating an orderly summation of distortion components in the cochlea, as would be expected. Figure 1 shows the auditory spectra (AS) produced by the CAR-FAC system for the five complex tones. Each of the AS contains a band of activity in the region of the primaries above 1500 Hz and a band of activity in the region of the first five harmonics of 100 Hz. For convenience, these two regions of the AS will be referred to as the primary spectrum (PS) and the quadratic distortion spectrum (qDS).

When there are only two primaries, there is only one component in the qDS and it is at 100 Hz, as expected. As NP doubles, the magnitude of the qDT at 100 Hz increases and the range of components in the qDS increases; that is, primaries that are farther apart in frequency contribute qDTs at higher difference frequencies. In the CAR-FAC system, the magnitude of the qDT increases with qDT frequency from 100 to 300 Hz, above this frequency qDT magnitude decreases. In contrast, in the data of Pressnitzer and Patterson (2001), the qDT at 100 Hz has the greatest magnitude, and magnitude decreases monotonically as qDT frequency increases.

In this example, where the level of the primaries is fixed, the magnitude of the qDS grows with the number of primaries in absolute terms, and it grows even more with respect to the magnitude of the PS. This is because, locally, cochlear compression is driven by the overall level of the input. Thus, as NP increases and the PS broadens above 1500 Hz, the response at 1500 Hz decreases as the compression increases. Suppression grows within the PS as NP increases, and as a result, the PS develops edge tones when NC is 8 or more.

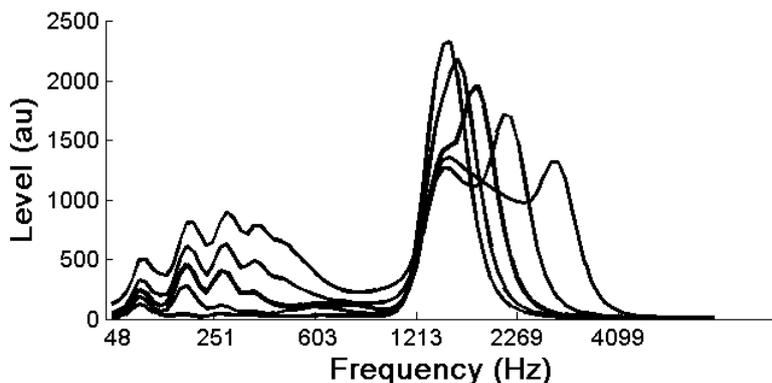


Figure 1. Auditory spectra for five complex tones with the same lowest component (15). The fundamental was 100 Hz and the tones had 2, 3, 5, 9 or 17 consecutive harmonics, all with the same level, 54 dB SPL.

These initial simulations indicate that the compression applied by CAR-FAC systems results in distortion spectra that are similar in many respects to those observed by Pressnitzer and Patterson (2001), although the magnitude of the qDT in the CAR-FAC system does not decrease monotonically with increasing frequency as it does in the behavioural data.

2. The Effects of Compressive Distortion in CAR-FAC Systems

The parameter values for the CAR-FAC system were those referred to as “fit 507” of the “Pole-Zero Filter Cascade” (Lyon, 2011, Fig. 6, PZFC). This version of the system provides a good fit to a wide range of notched-noise masking data and it produces pronounced quadratic distortion. The CAR-FAC system was used to generate sets of auditory spectra for a complex tone presented at levels varying from 40 to 80 dB SPL in 10 dB steps. The fundamental of the tone was 200 Hz and it had 11 adjacent harmonics. The lowest component, LC, of the tones was varied from 1 to 16 in doublings to survey the parameter space and locate typical distortion patterns. Broadly speaking, tones with LC equal 1, 2 or 4 produced similar patterns of AS, and high-frequency tones with LC equal 8 or 16 produced similar patterns of AS, so the results will be described for two tones, one with LC=2 and the other with LC=8.

2.1 Auditory Spectra for High-Frequency Complex Tones (LC 8)

Three sets of AS were generated for LC 8 (right-hand column of Figure 2). The AS for the standard CAR-FAC system are shown in the middle-right panel. The PS show that the primaries are not resolved; the DS reveal resolved peaks at the first four harmonics of the fundamental. The sequence of PS show that the system is compressive; in the range 40-60 dB SPL, each 10-dB step in stimulus level produces a roughly equal change in the level of the PS. As the level of the tone increases to 70 and then 80 dB SPL, the magnitude of the increase in PS increases. Nevertheless, the system remains compressive; the increases in AS level are far less than an order of magnitude in both cases. Over the same range of levels (60-80 dB SPL), the magnitude of the DS increases at a slower rate than it does at lower levels. Moreover, the pattern of PS activity changes above 60 dB SPL: there is an increase in suppression in the central region of the PS, leading to more pronounced edge tones; there is an increase in the magnitude of the fourth harmonic of F0; and there is an increase in the response in the region between the PS and the qDS. These are the characteristics of cubic distortion as they appear in the AS of the CAR-FAC system.

The *cubic compressor* in the CAR-FAC system reduces the amplitude of the output of each section, by a small proportion of the cube of what would otherwise be the output, before the output passes to the next stage. In the model, it is associated with the operation of the outer hair cell. The input/output function for the compressor is radially symmetric, so it would be expected to introduce cubic distortion tones (cDTs) of the form $2f_1-f_2$ rather than qDTs of the form f_2-f_1 (where f_1 and f_2 , are the frequencies of two primaries). For a complex tone, cubic distortion appears as a spectrum of harmonics cascading down in frequency and magnitude from the low-frequency side of the PS. It will be referred to as the cubic portion of the distortion spectrum (cDS), and this is what appears in the AS at the higher tone levels (70 and 80 dB SPL). For comparison, the bottom-right panel of Fig. 3 shows the set of AS generated when the cubic compressor is turned off (CC off). The cDS that appears in the middle-right panel at high tone levels (70 and 80 dB SPL) disappears, and the suppression in the PS is greatly reduced, confirming that these are aspects of cubic compression. Note also that the qDS persists at low frequencies in the bottom-right panel when the cubic compressor is turned off. Indeed, the magnitude of the qDS *increases* at the higher tone levels (70 and 80 dB SPL) indicating that, in this model, the cubic distortion suppresses the qDS at high stimulus levels.

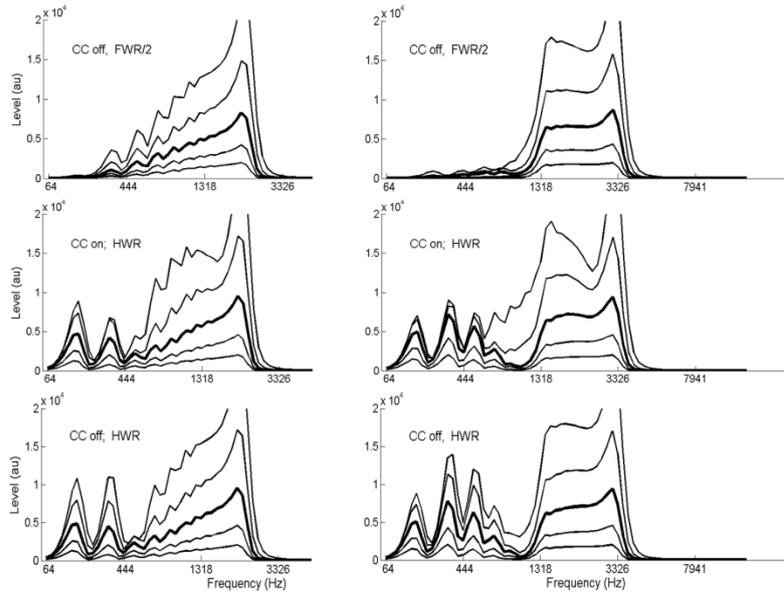


Figure 2. Auditory spectra for a 200-Hz complex tone with 11 harmonics at levels from 40 to 80 dB SPL in 10-dB steps; LC is 2 (left column) or 8 (right column). The middle row shows the AS produced with the CAR-FAC system. The lower row shows the AS when the cubic compressor is turned off; the upper row shows the AS when the cubic compressor is turned off *and* HWR is replaced with FWR/2.

The CAR-FAC system has an automatic-gain-control (AGC) network (Lyon, 2011, Fig. 3); it computes a continuous estimate of the magnitude at the output of each filter stage and, as the magnitude increases, the network reduces the gain of the filter. This is why the degree of PS compression is similar in the middle and bottom panels of the right-hand column.

The output of each filter stage is half-wave rectified to simulate the dominant aspect of inner hair-cell processing before it is passed to the AGC network, which consists of a cascade of four low-pass filters with time constants of about 3, 12, 48 and 200 ms. Their contributions are combined with increasing weights (1.0, 1.4, 2.0 and 2.8, respectively) to produce the current estimate of tone magnitude in a given channel. The top-right panel shows the AS generated for the same tones when the rectification is switched from half-wave to full-wave. This largely eliminates the rapid oscillation of gain associated with HWR, and it largely eliminates the quadratic distortion. The FWR signal has about double the magnitude of the HWR signal and so the FWR output was divided by 2 to ensure that the estimate going into the AGC network was about the same for the two forms of rectification. The compression range is about the same in the top and bottom panels, and the range is also similar to that in the middle-right panel.

In summary, the quadratic distortion produced by the default CAR-FAC system arises in the AGC network and is closely associated with the HWR that represents the operation of the inner hair cell in the model, rather than the cubic compressor that represents the operation of the outer hair cell in the model.

2.2 Auditory Spectra for Low-Frequency Complex Tones (LC 2)

When LC is 8 or more, the quadratic and cubic distortion products appear in separate frequency regions and their interaction is minimal. Pressnitzer and Patterson (2001) chose a relatively high LC value (15) to minimize the interaction of quadratic and cubic distortion, and so provide uncontaminated measurements of the quadratic distortion. However, the complex tones of speech and music typically have prominent low-order harmonics where

quadratic and cubic effects of compression might be expected to interact. CAR-FAC systems have the advantage that all of the effects of compression appear in the AS for any sound, and thus, they make it possible to investigate the complicated interaction of primaries and distortion tones in everyday sounds. The left-hand column of Figure 2 shows sets of AS like those in the right-hand column but for tones in which LC is 2 rather than 8. As noted above, similar AS are generated by tones with lowest components in the range 1-4.

The AS in the top-left panel were produced with the cubic compressor turned off (CC off) and the HWR replaced by FWR/2, so these AS show the effects of compression with the minimum of distortion. For the lower levels, 40-60 dB SPL, the AS largely reflect the energy in the stimulus at the output of the filterbank. The lowest harmonics appear as separate peaks and there is minimal activity at the fundamental. Then, as frequency increases and resolution decreases, multiple components pass through the filters and the magnitude of the AS rises. The upper edge tone does not appear until about 60 dB SPL indicating that suppression does not play a large role in determining the shape of these AS at low to moderate levels. At higher levels, 70 and 80 dB SPL, the slope of the spectrum increases because the bandwidths of the filters increase at the higher levels. At the same time, the relative magnitude of the upper edge tone increases indicating that suppression may affect the slope of the PS at higher levels.

The middle-left panel shows the AS with cubic compression (CC on) and HWR. At the lower levels, 40-60 dB SPL, the high-frequency end of the PS exhibits slightly less compression (the displacement of the AS increases slightly relative to that in the top-left panel), but the AS is still largely determined by the energy in the primaries. At higher levels, the rate of increase in PS level increases and the relative strength of the upper edge tone increases. However, the main difference, relative to the AS with CC off and FWR/2, appears in the region of harmonics 1-3: A prominent distortion tone appears at the fundamental (200 Hz) where the stimulus has no energy, and there is an increase in level of activity at the second harmonic (400 Hz). In contrast, the level of the third harmonic (600 Hz) in the middle-left panel *decreases* with respect to that in the top-left panel where the distortion was minimized. This means that, in this CAR-FAC system, primaries in cosine phase can generate distortion tones whose aggregate magnitude at a specific frequency is greater than that of the primary at that frequency, and the phase of the distortion tone is such that it partially cancels the primary. There are no perceptual measurements that might confirm or refute the prediction that the internal magnitude of the third harmonic is less than that of the adjacent harmonics for multi-harmonic tones of this form.

Finally, compare the AS in the middle-left panel with those in the bottom-left panel where the cubic compressor is turned off. The rectification mode (HWR) is the same in these two versions of the system. At the lower levels, 40-60 dB SPL, the PS and DS are essentially the same, which confirms that it is the AGC network that determines the compression in the system. At the higher levels, there is no evidence of cubic distortion on the lower flank of the PS in the bottom-left panel, as expected. At the same time, the levels of the fundamental and second harmonic *increase*, indicating that cubic distortion interferes with quadratic distortion in the full system (middle-left panel) when the stimulus contains low order harmonics.

3. Conclusion

CAR-FAC systems provide a means of investigating the plethora of distortion components produced by cochlear compression when the stimulus is a multi-harmonic tone, and the interaction of the distortion components on the basilar membrane. Perceptual measures of cochlear distortion are sufficiently accurate to assist in determining which of the many forms of CAR-FAC system might provide the most promising models of cochlear processing.

References

- Goldstein, J. L., 1967. Auditory nonlinearity. *J. Acoust. Soc. Am.* 41, 676-689.
- Lyon, R. F., 2011. Cascades of two-pole–two-zero asymmetric resonators are good models of peripheral auditory function. *J. Acoust. Soc. Am.* 130, 3893–3904.
- Pressnitzer, D., Patterson, R. D., 2001. Distortion products and the perceived pitch of harmonic complex tones. In: Breebaart, D., Houtsma, A., Kohlrausch, A., Prijs, V., Schoonhoven, R. (Eds), *Physiological and Psychophysical Bases of Auditory Function*, Shaker BV, Maastricht, pp 97-104.