Distortion products and the perceived pitch of harmonic complex tones

D. Pressnitzer and R.D. Patterson

Centre for the Neural Basis of Hearing, Dept. of Physiology, Downing street, Cambridge CB2 3EG, U.K.

1 Introduction

Combination tones (CTs) produced by acoustic stimuli have been studied by many authors as an efficient way to learn about the numerous physiological nonlinearities involved in auditory processing. As far as perception is concerned, CTs have been proposed as a basis for the perception of the pitch of the missing fundamental (Helmholtz, 1877). It has since been shown that the missing fundamental can still be perceived when CTs are cancelled (Schouten, 1938). However, this does not preclude a possible *contribution* of CTs to the pitch of the missing fundamental in normal listening conditions.

To our knowledge, relatively few CT data are available which could shed direct light on the contribution of CTs to pitch perception. Most psychophysical distortion studies employed acoustic stimuli made of 2 tones and measured the different kinds of CTs produced (e.g. Goldstein, 1967). When a harmonic tone complex is used instead, one might predict that a complete distortion spectrum (DS) that reconstitutes the lower part of the harmonic series would be generated. Fletcher (1924) proposed the existence of such a DS but it is unclear how his data were collected and at which level (possibly more than 120 dB SPL). Greenwood (1972) showed that sounds with a continuous spectrum could produce noise distortion bands but generalisation of his results to harmonic sounds remains theoretical. In the current study, combination tones produced by harmonics 15 to 25 of a missing 100-Hz fundamental were measured psychophysically at a moderate sound level.

2 Experiment 1: Existence of the Distortion Spectrum

2.1 Method

Eleven pure tones between 1.5 to 2.5 kHz with a 100 Hz spacing were used as primaries. The spectrum level of each tone was 54 dB SPL, which gave an overall level of 65 dB SPL. All tones were in cosine phase (CPH). CTs at frequencies of 100, 200, 300 and 400 Hz were investigated (the first four components of the hypothetical DS).

For each CT, the "cancellation of beats" method was employed (Goldstein, 1967). A pure tone was added to the primaries at a frequency equal to that of the CT plus 3 Hz. Its amplitude was adjusted by each listener to produce clear beats. Failing the production of beats, the CT was considered not measurable. If beats could be heard, a second tone was

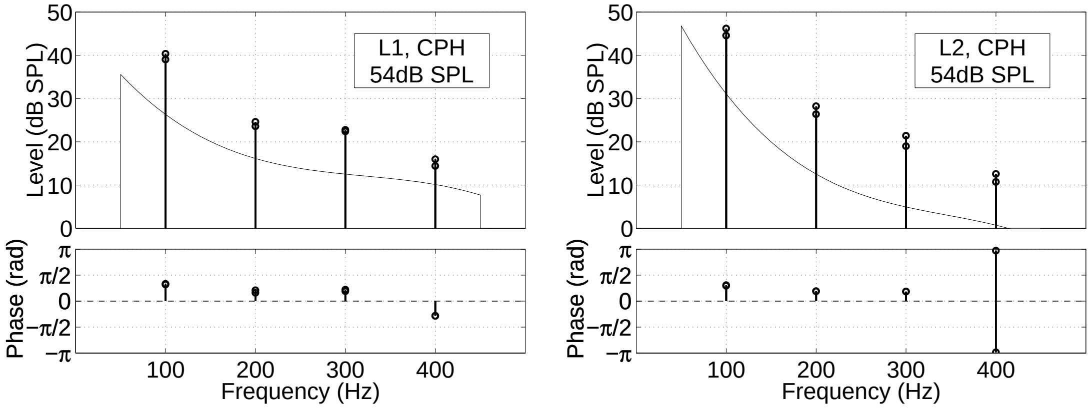

Figure 1: *Results for Experiment 1. L1 and L2 are the two listeners. Upper panels: Amplitude of the cancellation tones for each CT frequency. The white patch represents the hearing thresholds measured for each listener. Lower panels: phase of the cancellation tone. Each circle represents one adjustment.*

then added at the frequency of the CT. Its amplitude and phase could be adjusted to try to cancel the beats: if the cancellation tone has equal amplitude and opposite phase to the CT, beats should disappear. Listeners were asked to bracket the regions of amplitude and phase where the cancellation occured in order to determine its centre point. Amplitude and phase of the cancellation tone for each CT are the data presented.

Stimuli were generated on a TDT System II (DD1, FT6, PA4, SM3, HB6). The primaries were continuously played on one channel. The beating and cancellation tones were generated on a second channel which had an extra 20 dB fixed attenuation (SM3). The channels were mixed and played to one ear of a AKG K-240-DF headset. The experimental apparatus was calibrated with a B&K 4153 artificial ear coupled to the headset, a 1/2” B&K 4134 microphone and a B&K 2610 measuring amplifier. Distortion of the physical signal introduced by the complete apparatus was not measurable at the SPL of the experiment and did not appear until 84 dB SPL (Stanford Research SR 780 spectral analyser).

Results are reported for two listeners and two independent repeats of each measurement. Listener 1 is the first author. Listener 2 was paid for her participation.

2.2 Results

Results are presented in Figure 1. The first four components of the DS could easily be measured for both listeners. There is some inter-subject variability, but the agreement between independent measures within each subject is very good.

The highest component of the DS is the first one, corresponding to f0. Its level is between 10 and 15 dB less than the primaries. This is quite high considering the moderate presentation level. The next three components of the DS have a decreasing amplitude but are still above hearing threshold. It is likely that more components of the DS would also be above threshold.

Thus, an harmonic complex tone in cosine phase with the lower part of the spectrum missing can produce a sizeable DS, even at moderate to low sound levels

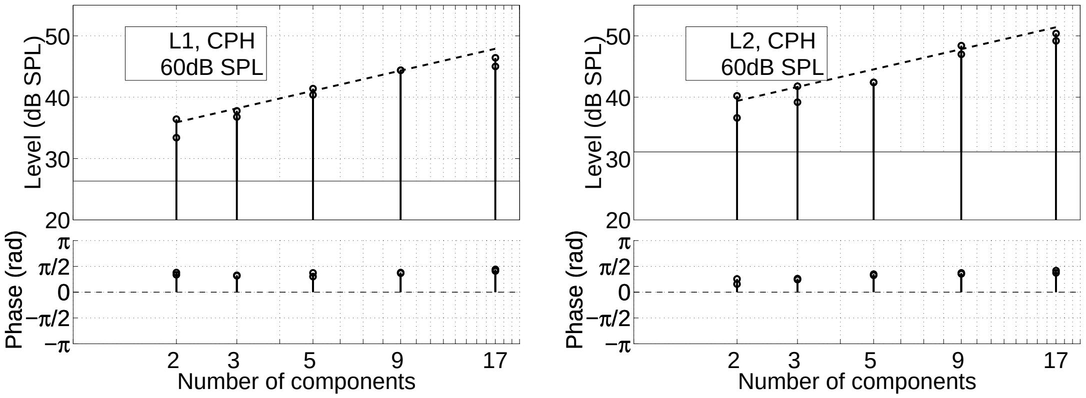

Figure 2: *Results for Experiment 2. Amplitude and phase of the cancellation tones at f0 for different number of primary components. The white patch represents the hearing threshold at f0 measured for each listener.*

3 Experiment 2: Build up of the Distortion Spectrum

3.1 Rationale

A simple conceptual way to explain the large DS observed in Experiment 1 is to consider the possible contribution of each pair of primary components to a given CT. For instance, the CT at f0 may contain the additive contribution of all pairs, f1 and f2, for which f2-f1 = f0. A quadratic distortion tone (QDT) is produced at f0 by such pairs.

To test this hypothesis, the number of components of the primary was manipulated. Primaries consisting of 2, 3, 5, 9 or 17 harmonics of a 100 Hz fundamental, starting at the 15th, were used. With these parameters, 1, 2, 4, 8 or 16 pairs of components may contribute to the CT at f0. The CT at f0 was measured with the method of Experiment 1, with the difference that the spectrum level was increased from 54 to 60 dB SPL in order to be able to measure the CT with 2 components. Cosine phase was used.

3.2 Results

Results are presented in Figure 2. For 2 primaries, the QDT is measured at 20 to 25 dB below the primaries, which is roughly consistent with other studies (Goldstein, 1967). When the number of component is increased, the amplitude of the CT at f0 increases regularly. The dashed line superimposed on the results represents a slope of 3 dB per doubling of number of pairs. The experimental points seem to fall along this line, which indicates that each pair contributes approximately equally to the CT. It is noticeable that the phase of the CT changes very little in spite of the large variation in its amplitude.

One explanation for the large DS observed in Experiment 1 (15 components) could then be that the DS builds up with an increasing number of component.

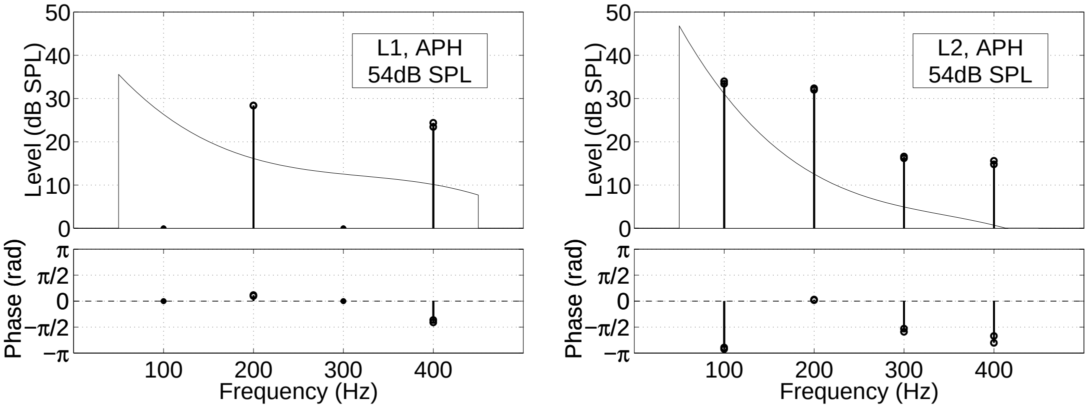

Figure 3: *Results for Experiment 3. As in Figure 1.*

4 Experiment 3: Alternating phase

4.1 Rationale

It is likely that the build up observed in Experiments 1 and 2 is dependent upon the phase relationship of the primaries. To test this hypothesis, the DS produced by a primary identical to that of Experiment 1 except for the phase between components was measured. An alternating phase (APH) relation was used: every other harmonic was shifted by $\pi/2$ rad. Stimuli for Experiments 1 and 3 had thus a same amplitude spectrum, but different temporal waveforms.

4.2 Results

The DS produced by the APH harmonic complex is different from the one obtained with the CPH condition. The first and third components of the DS are now either completely absent (L1), or substantially reduced (L2). When still present, these components have a very different phase than the one observed in the CPH condition. In contrast, the second and fourth components of the DS are almost identical to those observed in the CPH condition, both in amplitude and phase.

5 Qualitative Model

It is possible to account qualitatively for the behaviour of the measured DS of Experiment 1, 2 and 3 with a small number of simplifying hypotheses:

(1) the DS is mainly due to the vector sum of the quadratic distortion tones (QDTs) produced by all possible pairs of primaries. It is possible, to a first approximation, to neglect the contribution of cubic distortion tones and their interaction with acoustic primaries;

(2) the amplitude of the QDT produced by a given pair of primaries is only a function of the frequency difference between the primaries. The influence of absolute frequency can be neglected within the 1 kHz range of the harmonic complex;

(3) The phase of each QDT is the sum of two terms. The first term is the difference of the phases of the primaries. The second term is a constant that depends only on the frequency difference between primaries. The phase shift due for instance to different propagation lengths due to the site of generation is neglected.

The DS measured in Experiment 1 can be explained by (1), the superposition of QDTs at each differential frequency. Because of (3), the CPH produces a constructive build up. For Experiment 2, the regular increase in the amplitude of the f0 CT when the number of acoustic components is increased can be accounted by (2), and the absence of change in their phase by (3). For Experiment 3, it is possible to notice that because of the APH the predicted phase of the QDTs at f0 and 3f0 alternates between primary pairs. As a consequence, the vector sum should be destructive and the DS component very weak. The opposite is predicted for QDTs at 2f0 and 4f0, which should be identical in amplitude and phase to the CPH condition. This is roughly what is observed.

The proposed hypotheses are naturally gross over-simplifications. For instance, the predicted cancellation of the first and third component of the APH DS is not totally observed for L2 (Figure 3), which indicates that other mechanisms should be taken into account. The method of superposing linearly the effects of non-linear phenomena is also questionable. Nevertheless, it is surprising how far the qualitative behavior of the DS can be understood with such simple hypotheses. Most non-linearities present in proper models of auditory non-linearity would satisfy these hypotheses. On the other hand, even a square-root law would display the correct behavior.

6 Experiment 4: Minimisation of the Distortion Spectrum

The APH condition indicates that it is possible to cancel or diminish certain components of the DS. However, in the APH case, one component out of two of the DS is as large as in the CPH condition. Based on hypotheses (2) and (3) of the model, it is possible to derive phase relationships that should cancel exactly the component at f0 (which is the largest one for CPH) but that does not produce a constructive build up for other DS components.

One solution is to spread regularly within 2π rad the phase of the QDT produced by all pairs of adjacent primaries, so that their vector sum amounts to zero. If N is the number of acoustic primaries and i their index from 1 to N, this phase can be derived as $\phi_i = \pi i(i-1)/(N-1)$. The expression is similar the the phase relation proposed by Schroeder (1970) to reduce the crest factor of harmonic complexes.

Measurements of the DS produced by such a “circular” phase relation were made with the methods and listeners of Experiment 1. Results showed that for L1, the first 3 components of the DS were absent and the 4th one was just at hearing threshold. L2 displayed the first 3 components just at hearing threshold, whereas the 4th one was as large as in the CPH case.

These results show that the circular phase reduces the DS and generally confirm the qualitative model predictions.

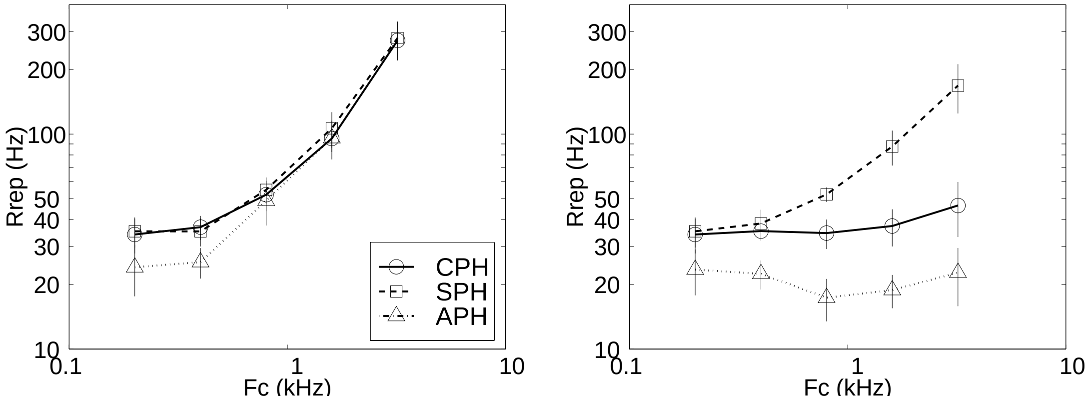

Figure 4: *Results for Lower Limit of Melodic Pitch. Mean and standard deviations for three listeners with (left) and without (right) lowpass masking noise.*

7 Influence of the DS on the Lower Limit of Melodic Pitch

7.1 The Lower Limit of Melodic Pitch

In a former study, we investigated the Lower Limit of Melodic Pitch (LLMP) with bandpass filtered harmonic complexes (Pressnitzer et al., 1999). A four-note, random melody was presented to listeners. It was immediately repeated with a semi-tone change introduced at random on one of the note. The listeners' task was to report on which note the change had been introduced. Melodies were drawn from the chromatic scale within a major third (4 semi-tones) of a given base note. The repetition rate of the base note was adaptively lowered (3-down 1-up) until threshold was reached.

The notes were bandpass-filtered harmonic complex tones. The passband had an equivalent rectangular bandwidth of 1.6 kHz. The lower filter cut-off, noted Fc, was a parameter of the study. Another parameter was the phase relation between components. Three phase conditions were used: Cosine, Alternating and Schroeder phase (CPH, APH and SPH). The overall level of presentation was 55 dB SPL. Lowpass continuous pink noise was added to the stimuli before presentation.

Mean results for three listeners of this previous study are reproduced in the left panel of Figure 4. Overall, there is a large influence of frequency region on the LLMP. Results for CPH and SPH are similar: for low Fcs, the LLMP is found to be around 30 Hz but it increases rapidly for higher Fcs. For the APH, the LLMP is lower in low frequency regions but increases rapidly and becomes impossible to measure in the highest Fc. These results could be modelled by a modified autocorrelation-based model of pitch perception (Meddis and Hewitt, 1991; Pressnitzer et al., in preparation).

7.2 The LLMP without lowpass masking noise

The LLMP experiment was repeated without lowpass masking noise. Results, averaged for the same 3 listeners, are presented in the left panel of Figure 4. It is obvious that the omission of the lowpass noise had a dramatic influence on the results. Whereas the SPH measures are relatively unaffected, the CPH LLMP is now almost constant around

30 Hz. The APH LLMP is also almost constant, up to the highest filter condition, and approximately one octave lower than the CPH condition.

The large influence of masking noise can be interpreted in terms of distortion spectra. The stimuli of the present experiments are actually similar to the ones found in the LLMP task for Fc=1.6 kHz and a repetition rate of 100 Hz (except that they were 12 dB louder). According to the results of Experiment 1, the CPH harmonic complex should produce a DS starting at f0. The DS creates energy in the low frequency region, where listeners are good to perform the melody task, regardless of Fc. This could explain why the CPH LLMP without lowpass masking noise is constant across all Fc conditions. For APH, Experiment 3 suggests that the DS will have one component out of two missing, i.e. it will be shifted by one octave. The APH LLMP results without noise are indeed constant and improved by approximately one octave compared to CPH. Finally, the SPH condition resembles the circular phase condition of Experiment 4 and should produces little or no DS. LLMP results with and without masking noise did not differ much for SPH.

The comparison of the LLMP with and without lowpass masking noise shows that distortion spectra can influence pitch perception. Listeners can use longer periodicities to perform a pitch task when energy is present in lower frequency regions (Ritsma, 1962). For high-pass filtered complexes, DS introduce energy in low frequency regions and improve pitch tasks performance for long periodicities. Note that this does not mean that the task is then based on the fundamental alone of the DS. As more than one component is present, it is also possible that the temporal information provided by the DS serves as the basis for the pitch task.

8 References

Fletcher, H. F. (1924). The physical criterion for determining the pitch of a musical tone. Phys. Rev. (23), pp. 427-437.

Goldstein, J. L. (1967). Auditory nonlinearity. J. Acoust. Soc. Am (41), pp. 676-689.

Greenwood, D. D. (1972). Combination bands of even order: masking effects and estimation of level of the difference bands $(f_h - f_l)$ and $2(f_h - f_l)$. J. Acoust. Soc. Am. (52), pp. 1155-1167.

Helmholtz, H. L. F. von (1877). On the Sensations of Tone as the Physiological Basis for the Theory of Music. 2nd. Ed. trans. A. J. Ellis (1885), from German 4th Ed., Dover, New York (1954).

Meddis, R. and Hewitt, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. J. Acoust. Soc. Am (89), pp. 2866-2882.

Pressnitzer, D. and Patterson, R. D. and Krumbholz, K. (1999). The Lower Limit of Melodic Pitch with filtered harmonic complexes. J. Acoust. Soc. Am. (105), p. 1152 (A).

Schouten, J.F. (1938). The perception of subjective tones. K. ned. Akad. Wet. Proc. (41), pp. 1086-1093.

Ritsma, R. J. (1962). Existence region of the tonal residue I. J. Acoust. Soc. Am. (34), pp. 1224-1229.

Schroeder, M. R. (1970). Synthesis of low peak-factor signals and binary sequences with low autocorrelation. IEEE Trans. Inf. Theory (16), pp. 85-89.