An excitation-pattern algorithm for the estimation of \((2f_1 - f_2)\) and \((f_2 - f_1)\) cancellation level and phase

Larry E. Humes
Institute of Electroacoustics, Technical University Munich, D-8000 Munich, Federal Republic of Germany and The Division of Hearing and Speech Sciences, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

(Received 9 July 1984; accepted for publication 20 May 1985)

An excitation-pattern algorithm is described which provides an estimate of cancellation level and phase for the \((2f_1 - f_2)\) and \((f_2 - f_1)\) distortion products. An experiment is first conducted to demonstrate the need for such an algorithm for \((f_2 - f_1)\) level predictions. The results of this experiment, which employed three pairs of primaries having complementary input levels \((L_1 = 65, L_2 = 85 \text{ dB}; L_1 = 85, L_2 = 65 \text{ dB})\), do not agree with the predictions of another similar algorithm [E. Zwicker, J. Acoust. Soc. Am. 69, 1410–1413 (1981)]. A new excitation-pattern algorithm is then described. The predicted level behavior for \((f_2 - f_1)\) and \((2f_1 - f_2)\) is more accurate for the proposed algorithm. In addition, an accurate phase estimate is also provided by the new algorithm.

PACS numbers: 43.66.Ki, 43.66.Ba

INTRODUCTION

Probably the two most prominent two-tone distortion products perceptually are \((2f_1 - f_2)\) and \((f_2 - f_1)\) (where \(f_1\) and \(f_2\) are the two input frequencies, or primaries, with \(f_2 > f_1\)). Various models have been described previously by several investigators to account for the behavior of one or both of these distortion products in man through a description of the underlying mechanisms of distortion-product generation (e.g., Zwicker, 1955, 1979a; Goldstein, 1967; Smoorenburg, 1972; Hall, 1974; Kim, 1980). These models are undergoing further evaluation, development, and refinement.

In the interim, a different approach to the description of the behavior of these two distortion products was adopted by Zwicker (1981a). In this approach, formulas were developed to describe the psychoacoustic behavior of the \((2f_1 - f_2)\) and \((f_2 - f_1)\) distortion products. The formulas are comprised of variables thought to be important for the description of the data. Thus, for instance, the equation describing the dependence of the \((f_2 - f_1)\) level on input parameters contains three variables: the levels of the two primaries, \(L_1\) and \(L_2\), and the frequency of the lower primary \(f_1\). The formula for \((f_2 - f_1)\), moreover, assumes that the level of this distortion product follows "that produced in regular quadratic distortion by two primaries" (Zwicker, 1981a, p. 1410). The predictive formula for the \((2f_1 - f_2)\) level includes \(L_1\) and \(f_1\) as variables as well, but also includes a term, \(\Delta Z_{1,2}\) representing the distance between \(f_1\) and \(f_2\) within the cochlea. The latter variable is included in the predictive formula for \((2f_1 - f_2)\) because this distortion product is believed to originate within the cochlea and to depend critically on the proximity of the excitation patterns of the two primaries within the cochlea. However, \((f_2 - f_1)\) is believed by Zwicker to originate at a site peripheral to the cochlea (Zwicker, 1979b). Accordingly, it is not important to consider the separation of the excitation patterns of each primary within the cochlea (\(\Delta Z_{1,2}\)) when describing the behavior of \((f_2 - f_1)\).

Since the publication of the formulas for the calculation of \((f_2 - f_1)\) and \((2f_1 - f_2)\) level by Zwicker (1981a), two relatively large data bases have become available for each distortion product (Zwicker, 1981b; Humes, 1985). Median cancellation level and phase data are now available from six ears and for a variety of stimulus parameters for \((f_2 - f_1)\) (Zwicker, 1981b) and \((f_2 - f_1)\) (Humes, 1985). The median data from each study provide a good description of the general observations drawn from previous studies using restricted stimulus conditions and/or a very limited number of subjects.

The median data for \((f_2 - f_1)\) from Humes (1985) are plotted in Fig. 1 along with the \((f_2 - f_1)\) cancellation levels from other studies. In all cases, \(f_1\) was between 926 and 2000 Hz. The dependence of \((f_2 - f_1)\) level on \(L_1 + L_2\) predicted by the formula from Zwicker (1981a) for \(f_1 = 1500\) Hz is

![FIG. 1. The \((f_2 - f_1)\) cancellation level as a function of the sum of the two primary levels \((L_1 + L_2)\). The solid line running diagonally across the figure from top to bottom is the dependence predicted by the formula from Zwicker (1981a). The open symbols represent median data \((N = 6)\) from Humes (1985) with circles for \(L_1 = L_2\) and squares and triangles for \(L_1 = 75\) or \(80\) dB and \(L_2\) varied. G = Goldstein (1967), Fig. 21, \(f_2/f_1 > 1.1\); H = Hall (1972), \(f_1 = 926\) Hz, \(f_2/f_1 \approx 1.4\); and Z = Zwicker (1955), Figs. 5 and Zwicker (1979b), Figs. 2-4,7.](image-url)
represented by the solid line in this figure. The predicted dependence of the \(f_2 - f_1\) level on \(L_1 + L_2\) does not provide a good description of the data except at moderate levels \((L_1 + L_2 = 150 \text{ to } 160 \text{ dB})\).

In Sec. I an experiment is described which confirms the inadequacy of the predictive formula described by Zwicker (1981a) for the \((f_2 - f_1)\) distortion product. In Sec. II an algorithm to predict the cancellation level and phase of the \((f_2 - f_1)\) and \((2f_1 - f_2)\) distortion products, based upon the excitation patterns of Zwicker (Zwicker and Feldtkeller, 1967; Zwicker, 1982), is described. The cancellation phase is predicted in the present algorithm in order to simply provide a more complete description of the psychoacoustic data on distortion product behavior. The ability of the present algorithm and that of Zwicker (1981a) to describe the median data for \((f_2 - f_1)\) and \((2f_1 - f_2)\) is then evaluated in Sec. III.

I. TWO-TONE INPUT WITH COMPLEMENTARY INPUT LEVELS

A. Rationale and methods

For an \(f_1\) value of 1.5 kHz, the cancellation level of the \((f_2 - f_1)\) distortion product is predicted by Zwicker (1981a) to be \(L_1 + L_2 = 127 \text{ dB}\), where \(L_1\) and \(L_2\) correspond to the levels of the two primaries in dB SPL. Hence, the same \((f_2 - f_1)\) level should result for \(L_1 = 85 \text{ dB}, L_2 = 65 \text{ dB}\) as for \(L_1 = 65 \text{ dB}, L_2 = 85 \text{ dB}\). This was examined in the present experiment for \(f_1 = 1.5 \text{ kHz}\) and \(f_2/f_1 = 1.28, 1.44,\) and 1.60.

A detailed description of the equipment used in this study has been provided recently elsewhere (Humes, 1985). Briefly, the listener was presented with three tones simultaneously: \(f_1, f_2,\) and the cancellation tone at \(f_2 - f_1\). The phases of all three signals were set to 0° relative to the cosine prior to equalization and transduction. The levels of the primaries, \(L_1\) and \(L_2\), were either 65 and 85 or 85 and 65 dB SPL, respectively. The subject adjusted the level and phase of the cancellation tone to make the distortion product inaudible. Three estimates of the cancellation level and phase were obtained for each stimulus condition from each subject. Four normal-hearing young adults served as subjects. Both ears of one subject (the author) were tested. All listeners had several hours prior experience with the cancellation paradigm.

B. Results and discussion

Table I presents the median cancellation level and phase values for each subject for the \((f_2 - f_1)\) distortion product. It is clear that the predictions made by the formula of Zwicker (1981a) do not hold. Recall that the formula predicts an identical cancellation level for this distortion product for primaries having complementary input levels. The difference in the cancellation level, \(\Delta CT\) level, is plotted in Fig. 2, where the \(\Delta CT\) level equals the cancellation level for \(L_1 = 85 \text{ dB}, L_2 = 65 \text{ dB}\) minus that for \(L_1 = 65 \text{ dB}, L_2 = 85 \text{ dB}\). The difference in the cancellation phase is also plotted in the lower panel of Fig. 2.

Data similar to those in the upper portion of Fig. 2 have been obtained previously by other investigators. Greenwood (1972) provides data obtained with a simultaneous masking paradigm for complementary input levels \((L_1/L_2)\) of 90/70 and 70/90 dB. For an \(f_1\) value of 3000 Hz and averaged over several \(f_2/f_1\) values, the \(\Delta CT\) level observed is approximately + 7 dB. Similarly, Goldstein (1967) using the cancellation method, provides data which indicate a \(\Delta CT\) level of + 5 dB for one subject and + 4.5 dB for another \((L_1/L_2 = 70/50 \text{ vs } 50/70 \text{ dB})\). Finally, results obtained by Humes (1979) with a nonsimultaneous masking paradigm reveal an average \(\Delta CT\) level of + 4 dB based on median data \((N = 4)\) averaged over several \(f_1\) and \(f_2/f_1\) values \((L_1/L_2 = 75/60 \text{ vs } 60/75 \text{ dB} \text{ and } 90/75 \text{ vs } 75/90 \text{ dB})\). The \(\Delta CT\) levels are greater in the study by Humes (1979) for small \(f_2/f_1\) (1.16 vs 1.41). Only the results from one subject by Zwicker (1979b), which show essentially no difference in the \(\Delta CT\) level for complementary tone pairs, appear to contradict the present findings.

Median data on the \((2f_1 - f_2)\) distortion product obtained from six ears by Zwicker (1981b) under stimulus conditions similar to those of this experiment suggest a \(\Delta CT\) level of approximately + 12 dB \((f_1 = 1620 \text{ Hz}, f_2/f_1 = 1.11, 1.2)\) for that distortion product. These data result from complementary tone pairs having levels \((L_1/L_2)\) of 80/60 and 60/80 dB. The level data for \((2f_1 - f_2)\) generally agree, therefore, with the \((f_2 - f_1)\) level data. The \((2f_1 - f_2)\) data, however, exhibit a steeper increase in \(\Delta CT\) level as \((f_2/f_1)\) increases.

The present results and those available in the literature which made use of complementary input levels refute the simple predictive scheme for the \((f_2 - f_1)\) level described by Zwicker (1981a). The \(\Delta CT\) levels observed for complementary input levels are similar for both \((f_2 - f_1)\) and \((2f_1 - f_2)\). If one considers that both distortion products may originate in the cochlea, then an algorithm incorporating the excitation-pattern concept from the work of Zwicker (Zwicker and Feldtkeller, 1967; Zwicker, 1982) may provide a description of the behavior of both distortion products. Figure 3 provides a schematic representation of the excitation patterns underlying the complementary input levels used in this experiment. The upper panel depicts the situation for \(L_1 = 85 \text{ dB} \text{ and } L_2 = 65 \text{ dB}\) while the lower panel exhibits the complementary primary levels. The dashed patterns reflect the change in underlying excitation patterns as \(f_2\) increases in frequency. It is assumed here that the region of overlap of the
FIG. 2. Difference in the \((f_2 - f_1)\) CT level and phase for \(L_1 = 85\), \(L_2 = 65\) dB and \(L_1 = 65\), \(L_2 = 85\) dB plotted for three \(f_2/f_1\) values. Open circles represent group medians while filled circles depict individual data.

excitation patterns of \(f_1\) and \(f_2\) is an approximation of the input to the nonlinearity (Goldstein, 1967; Hall, 1974; Zwicker, 1980). In the upper panel, that region of overlap is the entire excitation pattern associated with \(f_1\). It is constant for all values of \(f_2\). In the lower panel, the region of overlap between \(f_1\) and \(f_2\) is never as great as that for the complementary situation in the upper panel. The region of overlap in the lower panel, moreover, decreases as \(f_2\) increases. The input to the nonlinearity for \(L_1 = 65\) dB and \(L_2 = 85\) dB, therefore, is always less than that for \(L_1 = 85\) dB, \(L_2 = 65\) dB and this difference increases as \(f_2\) increases. This interpretation is consistent with the data from the present experiment and the results of other investigators in which complementary primary levels were employed.

In the next section, an excitation-pattern algorithm based upon the excitation patterns of Zwicker is described. The algorithm is shown to provide an accurate description of the median cancellation data for both \((f_2 - f_1)\) and \((2f_1 - f_2)\).

II. THE EXCITATION-PATTERN ALGORITHM

The general scheme for the excitation-pattern algorithm proposed here is illustrated schematically in Fig. 4. Shown in this figure are the excitation patterns of the primaries expressed in excitation level \((L_e)\) as a function of critical-band rate \((z)\) in Barks. The excitation level associated with the peak of each pattern is taken to be its level in dB SPL. The frequency scale has been transformed to the critical-band-rate scale using the following formula from Zwicker and Terhardt (1980):

\[
\frac{Z_x}{\text{Bark}} = 13 \arctan \left( \frac{0.76 f_x}{\text{kHz}} \right) + 3.5 \arctan \left( \frac{f_x^2}{7.5 \text{kHz}} \right),
\]

where \(Z_x\) is the critical-band-rate value for frequency \(f_x\), expressed in kHz. This transformation is applied for \(f_1\), \(f_2\), \(f_2 - f_1\), and \(2f_1 - f_2\).

The slopes of the excitation pattern are calculated using formula described previously by Terhardt (1979). The formulas for the low-frequency slope, \(S_1\) and the high-frequency slope, \(S_2\), are as follows:

\[
S_1 = 27 \text{ dB/Bark},
\]

\[
S_2 = (24 - 0.2 L) \text{ dB/Bark},
\]

where \(L_e\) is either \(L_1\) or \(L_2\) for the present applications.

Once the excitation patterns of both primaries are plotted, the region of excitation-pattern overlap can be found. This region is represented as the cross-hatched area in Fig. 4. The basis of the simple scheme for estimating the \((f_2 - f_1)\) and \((2f_1 - f_2)\) levels is to first locate the peak of the region of overlap along the \(Z\) scale. The critical-band-rate difference, \(\Delta Z\), between the peak of the excitation-pattern overlap \((\Delta Z_p)\) and the distortion product \([Z_{dp}]; \text{either } (f_2 - f_1)\text{ or } (2f_1 - f_2)\] is determined. One then multiplies this \(\Delta Z\) value by an attenuation factor of 5 dB/Bark and subtracts the product from the excitation level associated with the peak of the region of overlap \((L_p)\). More formally,

\[
L_{dp} = L_p - [5(\Delta Z)].
\]

The dashed line labeled \(- 5 \text{ dB/Bark}\) depicts this process graphically. A slight correction factor is then applied to the
A computer program written in BASIC which calculates the cancellation level and phase for \((f_2 - f_1)\) and \((2f_1 - f_1)\) in accord with the current algorithm for a specified \(L_1\), \(L_2\), \(f_1\), and \(f_2\) is provided in the Appendix. In the next section the accuracy of the excitation-pattern algorithm described here, that suppression of the signal at \(f_1\) by the signal at \(f_1\) begins when \(L_1 > L_2 + 10\) dB. Suppression reduces the effective level of \(L_2\) as follows:

\[ L_{2^*} = L_2 - 0.65(L_1 - L_2 + 10) \]  

It has been assumed, moreover, that suppression results in a reduction of the entire excitation pattern associated with the suppressed primary. The literature is mixed in support of this assumption (Moore, 1980; Duijnhuis, 1980b; Verschuure and Brocaar, 1980). The suppressed \(L_1\) and \(L_2\) values (i.e., \(L_1^*\) and \(L_2^*\)) must be used in Eqs. (4)-(10) whenever \(\Delta Z_{1,2} < 2\).

Aside from the estimation of distortion-product level, the present algorithm, unlike that of Zwicker (1981a), also provides formulas for the estimation of the cancellation phase. It is assumed that the primaries and the cancellation tone are all in 0° phase relative to the cosine as measured prior to transduction. The equation for the \((f_2 - f_1)\) cancellation phase in degrees is

\[ \phi(f_2 - f_1) = 440 - [30(\Delta Z)] \]  

where \(\Delta Z = Z - Z(f - f'_1)\). This equation has not been generalized to apply to a wide range of \(f_1\) values because the only sufficiently large data base for this distortion product (Humes, 1985) is restricted almost exclusively to a single \(f_1\) value (1.5 kHz).

The equation to calculate the \((2f_1 - f_1)\) cancellation phase in degrees is as follows:

\[ \phi(2f_1 - f_1) = \begin{cases} [Z_1 - Z \{(u - f_1)\}0.12f_1]360 \) & \{(60/Z_1) - 1\}0.22f_1]360 \) \\ + [50(1.62/f_1) - 1] \end{cases} \]  

The first term in this equation relates the cancellation phase to the critical-band-rate difference between the lower primary and the distortion product. This is similar in concept to the \(\Delta Z\) parameter described previously, except that \(Z_1\) has been substituted for \(Z_2\). The second term describes the dependence of the \((2f_1 - f_1)\) cancellation phase on the level of the lower primary, \(L_1\). The third term is an adjustment to improve the fit for different \(f_1\) values. Just as in the previous equations, the \(L_1\) and \(L_2\) values that appear in Eqs. (13) and (14), as well as the \(Z_2\) values determined from \(L_1\) and \(L_2\), are the suppressed \(L_1\) and \(L_2\) values.

III. COMPARISON OF PREDICTIONS TO DATA

As mentioned previously the two most extensive data bases for the \((f_2 - f_1)\) and \((2f_1 - f_1)\) cancellation level and phase are Humes (1985) and Zwicker (1981b), respectively. It is important to note, however, that the median data of these two studies are generally consistent with data from other studies using the cancellation method. With this in mind, the ability of the present excitation-pattern algorithm and the
TABLE II. The rms error between the predicted and observed \((2f_i - f_2)\) CT level and phase. Observed values are from Humes (1985).

<table>
<thead>
<tr>
<th>f_1[kHz]</th>
<th>f_2[kHz]</th>
<th>Input levels [dB SPL]</th>
<th>Algorithm</th>
<th>rms error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Zwicker (1981a)</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>1.74</td>
<td>(L_1 = L_2 = 70,75,80,90)</td>
<td>1.1</td>
<td>40</td>
</tr>
<tr>
<td>1.5</td>
<td>1.98</td>
<td>(L_1 = L_2 = 60,70,80,90)</td>
<td>3.8</td>
<td>174</td>
</tr>
<tr>
<td>1.5</td>
<td>2.16</td>
<td>(L_1 = L_2 = 70,75,80,90)</td>
<td>3.8</td>
<td>174</td>
</tr>
<tr>
<td>1.5</td>
<td>2.52</td>
<td>(L_1 = L_2 = 70,80,90)</td>
<td>4.0</td>
<td>18</td>
</tr>
<tr>
<td>1.5</td>
<td>2.74</td>
<td>(L_1 = 80; L_2 = 65 - 95)</td>
<td>4.5</td>
<td>26</td>
</tr>
<tr>
<td>1.5</td>
<td>2.16</td>
<td>(L_1 = 75; L_2 = 60 - 95)</td>
<td>1.1</td>
<td>76</td>
</tr>
</tbody>
</table>

formulas of Zwicker (1981a) to describe the median data for each distortion product will be examined.

A. \((f_2 - f_1)\)

Table II provides the rms error of the predictions made by the current algorithm and that of Zwicker (1981a) for the dependence of the \((f_2 - f_1)\) level and phase on the stimulus parameters employed in the study by Humes (1985). As shown in the table, \(f_1\) was constant in that study and four values of \(f_2/f_1\) were employed \((1.16, 1.32, 1.41, 1.68)\). At each \(f_2/f_1\) value a function was obtained relating the \((f_2 - f_1)\) cancellation level and phase to \(L_1 = L_2\). In addition, for \(f_2/f_1 = 1.16\) and 1.41, the \((f_2 - f_1)\) cancellation data were obtained for \(L_1\) fixed at either 75 or 80 dB and \(L_2\) varied.

The algorithm by Zwicker (1981a) produces an estimate of distortion product level only. The cancellation phase is not considered. Comparing the rms error between the predicted and observed \((f_2 - f_1)\) level for each algorithm, it is apparent that the current algorithm provides a more accurate description of \((f_2 - f_1)\) behavior. The rms error is typically 2-3 times larger for the Zwicker (1981a) formula than for the proposed algorithm.

The conditions for which the rms error of the predictions made by the excitation-pattern algorithm is largest are \(f_2/f_1 = 1.16, L_1 = 80\) dB, and \(L_2\) varied. Under these conditions, the rms error of the present algorithm was only 2 dB lower than that for the predictions made by Zwicker (1981a). Figure 5 illustrates the agreement between the data and the two sets of predictions for these stimulus conditions. Note that although the excitation-pattern algorithm proposed here underestimates the \((f_2 - f_1)\) level by 5-6 dB for \(L_2 > L_1\), the slopes of the function for this algorithm are more consistent with the data than those predicted by the regular quadratic distortion incorporated in Zwicker’s formula.

The rms error for the cancellation phase is unacceptably large for \(f_2/f_1 = 1.32\) \((174^\circ)\) and possibly for \(f_2/f_1 = 1.41\) \((90^\circ,76^\circ)\). These discrepancies are illustrated graphically in Fig. 6. In this figure, the \((f_2 - f_1)\) cancellation level and phase are plotted as a function of \(f_2/f_1\) for \(L_1 = L_2 = 80\) dB. The inaccuracy of the present phase predictions at \(f_2/f_1 = 1.32\) is readily apparent. A good fit is obtained, however, at \(f_2/f_1 = 1.16\) and 1.68 and a reasonable fit to the data results for \(f_2/f_1 = 1.41\). The dependence of the \((f_2 - f_1)\) cancellation level on \(f_2/f_1\), as illustrated in the upper panel of Fig. 6, is described equally well by the predictive equations proposed here and those of Zwicker (1981a). It is noteworthy that a cochlear-based algorithm, such as the one proposed here, can predict an \((f_2 - f_1)\) level that is almost independent of \(f_2/f_1\). This has also been demonstrated in the cochlear model of nonlinearity developed by Hall (1974). Thus the lack of a sharp dependence of the \((f_2 - f_1)\) level on \(f_2/f_1\) does not imply a noncochlear origin for this distortion product.

The present excitation-pattern algorithm provides an accurate estimate of the \((f_2 - f_1)\) cancellation level and a reasonable estimate of the \((f_2 - f_1)\) cancellation phase under most stimulus conditions. The description of the level behavior of this distortion product by the present algorithm is better than that provided by the formula from Zwicker (1981a).

The present algorithm makes predictions for conditions not examined in detail in the study by Humes (1985). The excitation-pattern algorithm, for instance, predicts that the \((f_2 - f_1)\) level should decrease with increase in \(f_1\) and the precise manner in which it decreases with \(f_1\) should vary with \(f_2/f_1\). These trends appear in the literature, although the limited data make it difficult to evaluate quantitatively (Goldstein,
These effects were evaluated for a level increase of 4 dB and the predictions were compared to those for the 0 dB level. Here, $S_2$ was de-

circles = excitation-pattern algorithm; " + " = formula from Zwicker (1981a).

An increase in the slope of the function relating the $(f_2 - f_1)$ level to $L_1 = L_2$ has been observed frequently in some individuals at high sound levels and/or wide frequency separations of the two primaries. This has been observed in psychoacoustic studies (Wenner, 1968; Greenwood, 1972; Humes, 1979, 1980, 1983, 1985) and physiological investigations (Worthington and Dallos, 1970). Such an increase in slope is explainable within the present excitation-pattern algorithm. Recall that one of the main determiners of the excitation level at the peak of the region of excitation-pattern overlap ($E_p$), and therefore the $(f_2 - f_1)$ level, is the upper slope of the excitation pattern associated with $f_1$ ($S_2$; see Eq. (10)). This slope is known to vary widely across individuals, much more so than the lower slope ($S_1$) (e.g., Zwicker and Schorn, 1978). By modifying Eq. (3), which is used to calculate $S_2$, the effects of increased upward spread of excitation on the predictions of the excitation-pattern algorithm for the $(f_2 - f_1)$ level could be evaluated. Here, $S_2$ was decreased 4 dB and the predictions were compared to those obtained under conditions of normal upward spread of excitation. The two major consequences of this change were: (1) slopes relating $(f_2 - f_1)$ level to $L_1 = L_2$ increased by approximately 0.4 dB/dB when calculated across the entire range of input levels (70 to 90 dB); and (2) within the range of input levels evaluated the slope increased by approximately 0.45 dB/dB from low to high levels ($L_1 = L_2 > 80$ dB). These effects were evaluated for $f_1 = 250$ to 8000 Hz and $f_2/f_1 = 1.1$ to 1.4. Within the present framework much of the individual variability in slopes of the function relating the $(f_2 - f_1)$ level to $L_1 = L_2$ observed previously can be ascribed to individual variability in the upper slope of the excitation pattern associated with $f_1$ ($S_2$). It is also possible, however, that the change in slope at high sound levels can arise from a different type and/or source of nonlinearity.

Finally, the $\Delta$ CT levels and $\Delta$ CT phase values predicted by the present algorithm for the stimulus conditions of the experiment described in Sec. I are 5.2, 8.1, and 11.4 dB and $-40^\circ$, $-37^\circ$, and $-90^\circ$ for $f_2/f_1 = 1.28, 1.44, \text{and} 1.60$, respectively. These values are in reasonable agreement with the results of that experiment (Fig. 2).

B. $(2f_1 - f_2)$

The data base obtained for $(2f_1 - f_2)$ by Zwicker (1981b) is more extensive than that for $(f_2 - f_1)$. A wider range of stimulus conditions was employed. Here, $f_1$ was either 0.54, 1.62, 4.8, or 10.0 kHz. Because of the extremely large between-subject variability and the limited number of stimulus conditions for the data obtained at $f_1 = 10$ kHz, prediction of those data was not attempted. A summary of the remaining stimulus conditions for which median data were available is provided in Table III. The rms error of the predictions made by the excitation-pattern algorithm and Zwicker's formula for the $(2f_1 - f_2)$ level are also shown in this table. A comparison between the rms error for the $(2f_1 - f_2)$ level for the two predictive schemes is also provided graphically in the upper portion of Fig. 7. The average rms error across all conditions is 1.25 dB smaller for the current excitation-pattern algorithm. The largest disparity between the two sets of predictions occurs for the data obtained at $f_1 = 540$ and $f_2 = 780$ Hz. For these conditions the rms error associated with the formula from Zwicker (1981a) is typically 4–5 times greater than that of the present algorithm.

For the excitation-pattern algorithm described here the poorest prediction of the $(2f_1 - f_2)$ cancellation level and phase occurs for $f_2 = 4.8$ kHz. A comparison of the data to the predictions for these conditions is provided in Fig. 8. The predicted $(2f_1 - f_2)$ levels are consistently 5–6 dB above those observed by Zwicker (1981b). The slopes of both the observed and predicted function, however, are very similar. In light of the 20-dB range of the $(2f_1 - f_2)$ levels observed across subjects by Zwicker (1981b) for these particular stimulus conditions, the error is not considered critical. One could obtain a closer approximation of median values by simply subtracting 5 dB from the estimated levels for 4.8 kHz.

The rms error associated with the prediction of the $(2f_1 - f_2)$ cancellation phase is also largest for the conditions depicted in Fig. 8. The difference between predicted and observed $(2f_1 - f_2)$ cancellation phase is illustrated in the lower portion of this figure. The sizable rms error that results for these conditions ($69^\circ, 93^\circ$) arises from the poor fit for $L_1 < 50$ dB. Again, this error in predicting median values must be evaluated in light of the $150^\circ-200^\circ$ range observed across subjects for these conditions by Zwicker (1981b).

For most of the remaining stimulus conditions, the predicted $(2f_1 - f_2)$ cancellation phase was within approximately $30^\circ$ of the observed median value at all input levels. Aside from the data depicted in Fig. 8, the only other exceptions to this generalization are the three conditions in Table III and Fig. 7 having a rms error greater than $35^\circ$. In these three
TABLE III. The rms error between the predicted and observed \((2f_t - f_i)\) CT level and phase observed values are from Zwicker (1981b).

<table>
<thead>
<tr>
<th>Stimulus condition</th>
<th>(f_t) (kHz)</th>
<th>(f_i) (kHz)</th>
<th>Input levels (dB SPL)</th>
<th>Algorithm</th>
<th>Zwicker (1981a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.62</td>
<td>1.8</td>
<td>(L_1 = 60, 70, 80)</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>2</td>
<td>1.62</td>
<td>1.8</td>
<td>(L_1 = 40, L_2 = 30 - 60)</td>
<td>4.6</td>
<td>6 (38)</td>
</tr>
<tr>
<td>3</td>
<td>1.62</td>
<td>1.8</td>
<td>(L_2 = 60, L_1 = 45 - 80)</td>
<td>3.5</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>1.62</td>
<td>1.8</td>
<td>(L_2 = 80, L_1 = 60 - 80)</td>
<td>1.4</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>1.62</td>
<td>1.944</td>
<td>(L_1 = L_2 = 60, 70, 80)</td>
<td>1.6</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>1.62</td>
<td>1.944</td>
<td>(L_1 = 40, L_2 = 40 - 70)</td>
<td>3.9</td>
<td>66</td>
</tr>
<tr>
<td>7</td>
<td>1.62</td>
<td>1.944</td>
<td>(L_2 = 60, L_1 = 50 - 80)</td>
<td>3.1</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>1.62</td>
<td>1.944</td>
<td>(L_2 = 80, L_1 = 60 - 80)</td>
<td>2.6</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>1.62</td>
<td>2.192</td>
<td>(L_2 = 60, L_1 = 60 - 80)</td>
<td>4.2</td>
<td>31</td>
</tr>
<tr>
<td>10</td>
<td>1.62</td>
<td>2.192</td>
<td>(L_2 = 80, L_1 = 65 - 80)</td>
<td>2.9</td>
<td>60</td>
</tr>
<tr>
<td>11</td>
<td>1.62</td>
<td>2.192</td>
<td>(L_2 = 60, L_1 = 60 - 80)</td>
<td>3.9</td>
<td>33</td>
</tr>
<tr>
<td>12</td>
<td>0.54</td>
<td>0.702</td>
<td>(L_2 = 40, L_1 = 50 - 65)</td>
<td>2.9</td>
<td>23</td>
</tr>
<tr>
<td>13</td>
<td>0.54</td>
<td>0.702</td>
<td>(L_2 = 50, L_1 = 50 - 75)</td>
<td>4.6</td>
<td>23</td>
</tr>
<tr>
<td>14</td>
<td>0.54</td>
<td>0.702</td>
<td>(L_2 = 60, L_1 = 55 - 80)</td>
<td>6.0</td>
<td>24</td>
</tr>
<tr>
<td>15</td>
<td>0.54</td>
<td>0.702</td>
<td>(L_2 = 70, L_1 = 60 - 90)</td>
<td>6.0</td>
<td>24</td>
</tr>
<tr>
<td>16</td>
<td>0.54</td>
<td>0.702</td>
<td>(L_2 = 80, L_1 = 70 - 90)</td>
<td>5.7</td>
<td>27</td>
</tr>
<tr>
<td>17</td>
<td>0.54</td>
<td>0.78</td>
<td>(L_2 = 50, L_1 = 60 - 80)</td>
<td>1.0</td>
<td>19</td>
</tr>
<tr>
<td>18</td>
<td>0.54</td>
<td>0.78</td>
<td>(L_2 = 60, L_1 = 60 - 85)</td>
<td>1.6</td>
<td>17</td>
</tr>
<tr>
<td>19</td>
<td>0.54</td>
<td>0.78</td>
<td>(L_2 = 70, L_1 = 65 - 90)</td>
<td>1.9</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>0.54</td>
<td>0.78</td>
<td>(L_2 = 80, L_1 = 70 - 90)</td>
<td>2.3</td>
<td>16</td>
</tr>
<tr>
<td>21</td>
<td>4.8</td>
<td>5.365</td>
<td>(L_2 = 40, L_1 = 35 - 65)</td>
<td>5.8</td>
<td>69</td>
</tr>
<tr>
<td>22</td>
<td>4.8</td>
<td>5.365</td>
<td>(L_2 = 60, L_1 = 45 - 80)</td>
<td>7.0</td>
<td>93</td>
</tr>
</tbody>
</table>

\(\bar{\chi} = 3.55, 29.5, 4.8\)

additional cases the error was similar to that depicted in Fig. 8 for \(f_i = 4.8\) kHz. In particular, an accurate estimate of the \((2f_t - f_i)\) phase was obtained at high \(L_1\) values, but the phase estimates continued to overestimate the median values as \(L_1\) decreased.

In conclusion, the excitation-pattern algorithm proposed here provides a more accurate description of the \((2f_t - f_i)\) cancellation level than does the algorithm of Zwicker (1981b). In addition, it provides a good description of the \((2f_t - f_i)\) cancellation phase. The algorithm has also been applied to other less conventional stimulus configurations and was found to provide an excellent description of the data. Zwicker (1980), for example, reported data from two subjects for an iso-\((2f_t - f_i)\)-place experiment in which \((2f_t - f_i)\) was fixed in frequency at 1400 Hz, and \(f_i\) and \(f_t\) were increased in several steps from 1600 and 1800 to 2600 and 3800 Hz, respectively. The present algorithm provides a

FIG. 7. The rms error (from Table III between median data of Zwicker (1981b) and the predictions made by the current algorithm (open circles) and the formula from Zwicker (1981a). "Stimulus condition" along the abscissa refers to the conditions listed in Table III. The upper panel depicts the rms error for the CT level while the lower one illustrated that for the CT phase.

FIG. 8. Comparison of median the \((2f_t - f_i)\) cancellation level and phase from Zwicker (1981b) to predictions made by the proposed algorithm (open circles) and the formula from Zwicker (1981a) (+ " symbols). Data are for \(f_i = 4.8\) kHz, \(f_t = 5.365\) kHz, and \(L_1 = 40\) or 60 dB.

good description of these data as well. The present algorithm, however, does not describe the nonmonotonic behavior of \((f_2 - f_1)\) that has been observed by many investigators [see Zwicker (1981b), for review]. The specific stimulus conditions necessary to produce a minimum in the dependence of the \((f_2 - f_1)\) level on input level vary widely across subjects and are not present in median data (Zwicker, 1981b).

### IV. SUMMARY

The formulas for calculating the levels of the \((f_2 - f_1)\) and \((f_1 - f_2)\) distortion products proposed by Zwicker (1981a) were demonstrated to be inadequate for \((f_2 - f_1)\). An excitation-pattern algorithm is described which provides a better prediction of \((f_2 - f_1)\) level for a variety of stimulus conditions. This same algorithm yielded slightly better estimates of the \((f_1 - f_2)\) cancellation level. Finally, the present algorithm has the added advantage of providing an accurate estimate of distortion-product cancellation phase.

### ACKNOWLEDGMENTS

This work was performed while the author was on a fellowship from the Alexander von Humboldt Foundation. The author expresses his gratitude to Professor Eberhard Zwicker for making facilities and equipment available for this research and to Dr. Fred H. Bess for granting me the leave time to pursue this investigation. Finally, I thank Judy Haggard for typing the manuscript.

---

**APPENDIX: BASIC PROGRAM**

```basic
10 INPUT "ENTER L1, L2 (dB SPL)"; L1, L2
20 INPUT "ENTER F1, F2 (kHz)"; F1, F2
30 Z1<(I3*ATN(.75*F1))+(3.5*ATN((F1/7.5)^2)))  "ZWICKER & TERHARDT (1980)
40 Z2<(I3*ATN(.75*F2)+(3.5*ATN((F2/7.5)^2)))
50 DELTA2=Z2-Z1
60 IF DELTA2=2 THEN GOTO 90
70 IF L2>L1 THEN L1=(L2-L1)*.65+"SUPERI0F F1 BY F2"
80 IF L1>(L2+10) THEN L2=(L1-L2-10)*.65+"SUPPRESSION OF F2 BY F1"
90 S2=24-<(.2L1);"UPPER SLOPE OF L1 EXCITATION PATTERN (TERHARDT,1979)
100 IF (L1-(DELTA2S32)>L1 THEN GOTO 150
110 IF (L2-(DELTA2S32)>L1 THEN GOTO 180
120 ZPEAK=Z1+<(<Z2-DELTA2)+L1-L2)>s<Z1+Z2>)  "Z FOR PEAK OF EX.PATT. OVERLAP
130 LPKEAK=Z2-<(<Z2-ZPEAK)X5>)  "L FOR PEAK OF EXCITATION PATTERN OVERLAP
140 GOTO 200
150 ZPEAK=Z2
160 LPEAK=L2
170 GOTO 200
180 ZPEAK=Z1
190 LPEAK=L1
200 FCDT=<(2F1)-F2>
210 ZCDT=<(I3*ATN(.75*FCDT)+(3.5*ATN((FCDT/7.5)^2)))
220 FCDT=F2-F1
230 ZSDT=<(I3*ATN(.75*FSDT)+(3.5*ATN((FSDT/7.5)^2)))
240 LCDT=LPEAK-<(<Z2-ZCDT)X5>-<(<L2/40)-1>12>)
250 LSDT=LPEAK-<(<Z2-ZSDT)X5>-<(<L2/40)-1>12>)
260 PHASECDT=<(12F1*ZCDT)+<(12F1*ZCDT)+12>)
270 PHASESDT=<(12F1*ZSDT)-(12F1*LSDT)+12>)
280 PHASECDT=<(12F1*ZCDT)+(12F1*ZCDT)+12>)
290 PRINT"CDT (dB SPL) SDT (dB SPL) CDT PHASE SDT PHASE(DEG)"
300 PRINT"-------------------------"
310 PRINT LCDT, LSDT, PHASECDT, PHASESDT
320 END
```

1Zwicker (1981a) published a figure similar to Fig. 1 (Fig. 1, p. 1411). There were two errors in the data plotted in that figure regarding the data from Hall (1972) and Goldstein (1967). First, the data from Hall (1972) obtained at the lowest level \((L_1 = L_2 = 58 \text{ dB SPL})\) were not plotted while those for \(L_1 = 68 \text{ and } 78 \text{ dB SPL} \) were included. All of the data from Hall (1972) are included in the present Fig. 1. Second, the input levels and cancellation levels for the results of Goldstein (1967) were treated as sound pressure levels by Zwicker (1981a). Goldstein (1967), however, reported all his \((f_2 - f_1)\) data in "DBSL," which represents dB "relative to the threshold at the lower tone" (p. 678). Thus, for example, a median cancellation level of 25 dB SL, for \(L_1 = L_2 = 60 \text{ dB SL} \) at \(f_1 = 2000 \text{ Hz} \) would be appropriately plotted in a figure such as Fig. 1 as a cancellation level of 31.5 dB SPL for \(L_1 = L_2 = 153 \text{ dB SPL} \) (reference threshold for normal ears at 2000 Hz for the PDR-10 earphone used by Goldstein is 6.5 dB SPL, ANSI S3.6-1969). Zwicker (1981a) did not make the conversion from dB SL to dB SPL when plotting Goldstein's data.

2As published in Zwicker (1981a), the formula for estimating the \((f_2 - f_1)\) cancellation level [Eq. (2), p. 1412] is incorrect. The final term in that equation appears as an exponential value of \(-2\). Rather than raising the bracketed quantity by a power of \(-2\), the term should be subtracted from the quantity.

That is, \(-2\) is not an exponent as it appears. It should also be noted here that Fig. 3 (p. 1412) from Zwicker (1981a) which compares his predicted \((f_2 - f_1)\) levels for several values of \(f_1\) to cancellation data from Goldstein (1967) contains the same dB reference problem described in footnote 1. When conversions for \((f_2 - f_1)\) are made to dB SPL the fit of the data to predictions is actually much better than Fig. 3 of Zwicker (1981a) would indicate.

---


Greenwood, D. D. (1972). "Combination bands of even order: Masking ef-
fects and estimations of level of difference bands \( f_2 - f_1 \) and \( 2(f_2 - f_1) \)." 

Hall, J. L. (1972). "Auditory distortion products \( f_2 - f_1 \) and \( 2f_1 - f_2 \)." 


(Academische Pers B. V., Amsterdam).

Humes, L. E. (1979). "Perception of the simple difference tone, \( f_2 - f_1 \)." 

Am. 67, 2073–2083.

and distortion products \( (f_2 - f_1) \) and \( (f_3 - f_1) \)." J. Acoust. Soc. Am. 73, 
930–950.

Humes, L. E. (1985). "Cancellation level and phase of the \( f_2 - f_1 \) distortion 


Moore, B. C. J. (1980). "Detection cues in forward masking," Psychophy-

cal, Physiological, and Behavioral Studies in Hearing, edited by G. van 

Zwicker, E. (1981b). "Dependence of level and phase of the \( (2f_1 - f_2) \) can-
cellation tone on frequency range, frequency difference, level of primaries 


stimuli," in Psychophysical, Physiological and Behavioral Studies in 
Hearing, edited by G. van den Brink and F. A. Bilsen (Delft U. P., Delft, 
The Netherlands).

band rate and critical bandwidth as a function of frequency," J. Acoust. 
Soc. Am. 68, 1523–1527.