# Binaural edge pitch

M. A. Klein and W. M. Hartmann

*Physics Department, Michigan State University, East Lansing, Michigan 48824*
(Received 9 July 1980; accepted for publication 13 April 1981)

The Huggins pitch effect is created by dichotic broadband noise with interaural phase varying from 0 to $2\pi$ over a narrow frequency region. The sensation of pitch, corresponding to the frequency of the phase shift region, is usually understood as the result of a binaural differencing operation. We report here a pitch effect created by dichotic broadband noise with interaural phase varying from 0 to $\pi$ over a narrow boundary region, creating an edge in a difference channel. We call this effect Binaural Edge Pitch. For experienced listeners the effect is similar in nature and strength to the Huggins pitch. It is strongest for boundary frequencies in the 350–800 Hz range. Pitch matching experiments in this range find that the spread of matches is 1%–2% of the boundary frequency and that the pitch is 4% higher or lower than the boundary frequency. This shift is identical to the shifts which we find for the pitch of high-pass and low-pass noise bands. The correspondence argues strongly for an explanation of the Binaural Edge Pitch in terms of the Equalization–Cancellation Model of binaural processing, and pitch derived from a central spectrum.

PACS numbers: 43.66.Hg, 43.66.Nm, 43.66.Pn, 43.66.Rq [FLW]

## INTRODUCTION

### A. Huggins pitch

The Huggins pitch (Cramer and Huggins, 1958) is a binaural effect in which a sensation of pitch is created from dichotic noise. The noise heard by a listener is identical in the two ears except that there is an interaural, frequency-dependent, phase shift from 0 to $2\pi$ over a narrow frequency range. Normally the phase shift is produced by passing the noise for one ear through a narrowly tuned two-pole all-pass filter. The unfiltered noise is sent to the other ear. The stimulus situation is shown in Fig. 1A and B. The fascinating nature of the effect is that there is no frequency-specific information present in either of the stimuli channels. Both channels must be combined centrally, with phase information intact, for a listener to extract the sensation of pitch. The pitch is most easily heard when the center of the phase shift region is between 200 and 1000 Hz and the width of the phase shift region is about ±5% of the center frequency. In this case subjects hear a pitch, within the noise, corresponding to the frequency at which the phase shift is $\pi$, the center of the region, and they can match the pitch with a standard error of about 3%.

In 1962 Guttman made further studies of the Huggins pitch effect. He found that an equally strong sense of pitch was created when the interaural phase shift varied from $-\pi$ to $+\pi$. Wightman *et al.* (1977) generated a digital version of the Huggins pitch using broadband noise made from many sine waves spaced 10-Hz apart, with an interaural phase shift of $\pi$ on only one sine component. If the frequency of the phase shifted component in this stimulus is varied then it is possible to play tunes, which a subject will recognize.

The Huggins effect can be interpreted as a manifestation of the same binaural mechanism that is responsible for the effect of the masking level difference. Durlach (1962, 1963, 1972) proposed a mechanism, the Equalization-Cancellation (E–C) Model, which offers an explanation for the MLD and, together with the concept of the central spectrum (cf. Bilsen, 1977), for the Huggins pitch phenomena. According to this model the binaural system is able to process signals to the two ears so as to achieve a maximum signal-to-noise ratio at a central processor.

Durlach's original model allows the binaural system an interchannel equalization consisting of a *frequency-independent* amplitude shift and a *frequency-independent* phase shift. The model presumes that the two equalization parameters are optimized so that when the difference between the two channels is taken (cancellation process) optimum signal-to-noise ratio is achieved. The E–C Model allows for errors in both the equalization and cancellation processes, but this refinement may be ignored for the application of the model in the present paper.

The Huggins pitch effect is understood within the context of the E–C Model as the result of a binaural cancellation of most of the noise components of the stimulus. Optimum cancellation leaves only a narrow band of noise centered at a frequency which is the center of the phase shift region. A simple experiment makes the E–C Model explanation of the Huggins pitch effect seem plausible. If the two channels of the Huggins stimulus are electronically subtracted to create a narrow spectral noise band the stimulus acquires a tonal character which is similar to that of the Huggins pitch. As the phase shift region is made wider in frequency the Huggins pitch eventually disappears. Similarly as the width of the *spectral* noise band is made greater the sensation of pitch disappears. Generally the disappearance of salient pitch occurs when the width of the phase shift region in the first case is equal to the width of the noise band in the second case.

### B. Noise band edge pitch

There is ample evidence that broadband noise with a sharp spectral edge creates a sensation of pitch (Bilsen, 1977). (Additional evidence is provided by experiment 2 of the present paper.) The pitch does not correspond exactly to the frequency of the edge but is shifted *into* the noise, i.e., the pitch is shifted up for high-pass noise and is shifted down for low-pass noise. (Fastl, 1971).

### C. Binaural edge pitch

It occurred to us to ask the following question. Given that the binaural system is apparently capable of interaural cancellation of portions of noise bands in the formation of a central spectrum, and given that the edge of a broadband noise elicits a sensation of pitch, is it then possible to create a sensation of pitch with a dichotic noise which allows the binaural system to produce a broadband central noise spectrum with a sharp edge?

Suppose that the two ears are presented with broadband noise with an interaural phase shift varying from 0 to $\pi$ across some transition region, which we will call the phase boundary, (instead of 0 to $2\pi$ as for Huggins pitch). All low-frequency components are in phase but all high-frequency components are exactly out of phase. The stimulus situation for this experiment is shown in Fig. 1A and C. According to the original E–C Model the ear has the opportunity to apply a single frequency-independent interaural phase shift to this stimulus. Therefore it is impossible for the binaural system to cancel both the high-and the low-frequency components.

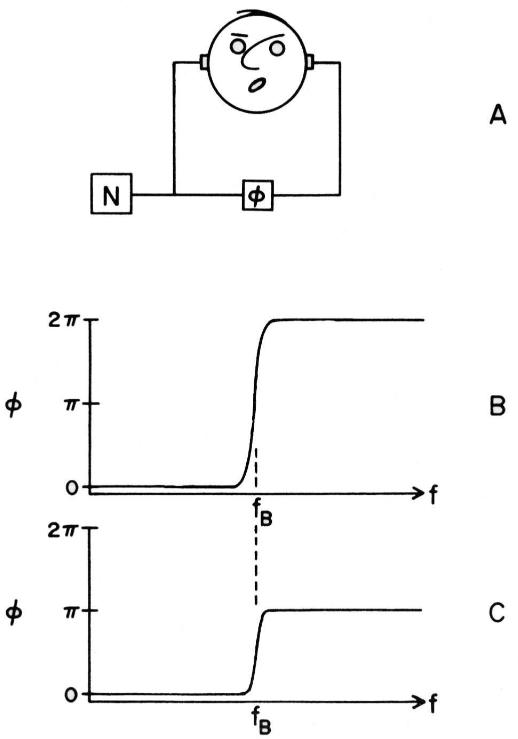


FIG. 1. A. Schematic representation of the signal paths for dichotic noise pitch phenomena. One ear receives the unaltered noise signal. The other ear receives the same noise signal after a phase manipulation. B. Interaural phase difference as a function of frequency for the Huggins pitch effect. Extreme high and low frequencies are in phase between the two ears. C. Interaural phase difference as a function of frequency for the Binaural Edge Pitch. Low frequencies are exactly in phase. High frequencies are exactly out of phase. There is a narrow boundary region about $f_B$.

The binaural system has two choices: (1) It can introduce no phase shift, then, through the cancellation process the low-frequency components are suppressed and the upper frequencies remain to generate a high-pass noise in the difference channel. (2) It can introduce a $\pi$ phase shift, then the high-frequency components are equalized while the low-frequency components, forced out of phase, result in a low-pass noise in the difference channel.

We performed dichotic experiments with broadband noise which has an interaural phase shift varying from 0 to $\pi$. We found that the stimulus does indeed produce a sensation of pitch near the phase boundary. Because the pitch sensation is presumably caused by an edge in the central spectrum, produced by the equalization–cancellation process, we refer to the pitch effect as *Binaural Edge Pitch.*

After establishing the existence of the pitch effect we naturally wondered, what, actually, does the binaural system do with the ambiguity implicit in the stimulus? Does the system choose to create a central spectrum which is high-pass noise (no equalization phase shift) or a central spectrum that is low-pass noise (an equalization phase shift of $\pi$)? We found that by comparing Binaural Edge Pitches with the pitches of low-pass and high-pass noise bands we could investigate this question with considerable experimental reliability.

## I. EXPERIMENT 1

### A. Stimuli

The noise stimuli used to establish the existence of the Binaural Edge Pitch (BEP) were generated digitally.[1] Two noise signals were generated, each with 251 equally spaced equal-amplitude sine components at random initial phase angles. For the stimulus sent to the left ear all sine components below the phase-boundary frequency were the same phase as those components for the right ear. Above the phase boundary all components to the left ear were at $\pi$ phase relative to the corresponding components for the right ear. The noise to right and left ears is described by the following formulas:

$$s_R(t) = \sum_{n=1}^{251} A \sin(n2\pi f_0 t + \phi_n) , \tag{1}$$

$$s_L(t) = \sum_{n=1}^{n_B} A \sin(n2\pi f_0 t + \phi_n) - \sum_{n=n_B+1}^{251} A \sin(n2\pi f_0 t + \phi_n) . \tag{2}$$

Here $\phi_n$ is a random variable. The uniform spacing of the spectral components is $f_0$. The phase-boundary frequency is defined here as $f_B = (n_B + \frac{1}{2}) f_0$, halfway between the highest homophasic component and the lowest heterophasic component. (For less abrupt phase boundaries the phase-boundary frequency is defined as the frequency where the interaural phase shift is $\pi/2$.)

The two noise channels were created by two 12-bit DACS. (When the two channels were subtracted electronically the resulting spectrum showed a 30-dB drop at the phase boundary.) The stimuli were presented at 60 dB SPL through Beyer DT-48 headphones.

The phase boundary was varied from 126–2438 Hz. We performed the experiments with phase boundaries in three overlapping frequency ranges: low, middle, and high. For each range, there was a particular noise waveform characterized by the boundary number $n_B$. Within a range the phase-boundary frequency was varied by changing the component spacing $f_0$. This was done by changing the sampling rate. Details of the stimulus are given in Table I. Because the stimuli, within a given range, are simply time-scaled versions of each other the phase-boundary frequency is not the only temporal/spectral information provided to the subject. One must consider the possibility that information other than the phase-boundary frequency is responsible for the Binaural Edge Pitch. In Sec. IF we show that these other temporal/spectral cues are unlikely sources of the Binaural Edge Pitch.

## B. Procedure

Subjects matched a sine tone to the pitch that they perceived in the dichotic noise. The dichotic noise stimulus was part of a repeating four-segment presentation structure lasting 1.6 s. As shown in Fig. 2, the first segment contained the dichotic noise stimulus. In the second, third, and fourth segments identical noise was sent to both ears. During the third segment the sine tone matching signal was added to the diotic noise. This four-segment sequence repeated indefinitely until the subject was satisfied with the match between the sine tone and the BEP. By maintaining noise through the matching (third) interval we reduced the perceptual dissimilarity between the dichotic noise (first) interval and the matching interval. This procedure made the matching task easier. Two other procedural aspects facilitated the matching task. The subjects were allowed to vary the intensity of the matching tone so as to increase perceptual similarity with the dichotic noise stimulus, and subjects were able to turn off the matching tone altogether.

## C. Middle range experiment

The first experiment was performed in the middle range. In a single run subjects matched the pitches of

TABLE I. Parameters of the noise stimulus.

| | Range | | |
|---|---|---|---|
| | Low | Middle | High |
| $n_B$ | 40 | 100 | 180 |
| $n_B/251$ | 0.16 | 0.40 | 0.72 |
| Min $f_B$ (Hz) | 126 | 315 | 567 |
| Max $f_B$ (Hz) | 420 | 856 | 2438 |
| Min $f_0$ (Hz) | 3.15 | 3.15 | 3.15 |
| Max $f_0$ (Hz) | 10.5 | 8.56 | 13.5 |

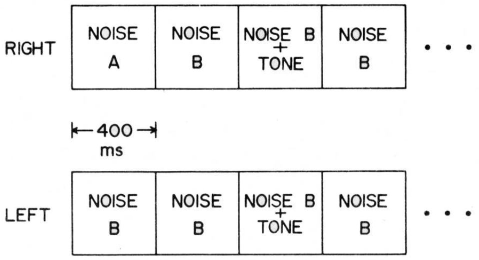


FIG. 2. The time structure of the stimuli presented to left and right ears during the method of adjustment task. This four-interval sequence repeats until the subject is satisfied that the pitch of the sine tone in interval 3 is the same as the pitch of the dichotic stimulus in interval 1.

dichotic noise with 12 different phase boundary frequencies. Five subjects participated in this experiment. Subjects G, M, and W were experienced listeners. Subjects M and W are the authors. Subjects D and R had not previously served in psychoacoustic experiments. Some subjects had practice in matching pitches of the Huggins stimulus before participating in this experiment.

## D. Results: Middle range

Four of the five subjects succeeded in achieving reproducible pitch matches to the binaural edge stimulus. The data for these subjects, based on four or more runs, were remarkably similar. The pitch associated with a phase boundary at frequency $f_B$ was reliably matched by a sine tone with frequency $f_m$, about a quarter-tone away from the boundary frequency. This result showed no dependence upon the phase-boundary frequency. Therefore it is possible to average the data for the 12 boundary frequencies. The average results are given in Table II. The table shows the percentage deviation of the matching frequency from the boundary frequency, $(f_m - f_B)/f_B$. The data for subjects G and M strongly suggested a bimodal distribution. For subject M the data were bimodal at the $p < 0.001$ level. For subject G the data were even more significantly bimodal. The data for subject W were unimodal, with a peak nearly coincident with the major peaks of the distributions for subjects M and G. The data for the less ex-

TABLE II. The table describes the distribution of pitch matches in the middle range of experiment 1 for four subjects. There are two peaks in the distribution. The means, widths, and relative weights of the peaks are given in the columns. The mean values and standard errors are expressed as percent deviation from the phase-boundary frequency.

| | Lower peak | | | Upper peak | | |
|---|---|---|---|---|---|---|
| | $(f_m/f_B - 1)$ | $\sigma$ | $P$ | $(f_m/f_B - 1)$ | $\sigma$ | $P$ |
| Subject | % | % | % | % | % | % |
| $W$ | ... | ... | 0 | 3.5 | 0.9 | 100 |
| $M$ | −4.1 | 3.9 | 33 | 3.6 | 1.5 | 67 |
| $G$ | −5.2 | 1.4 | 40 | 3.1 | 1.1 | 60 |
| $D$ | −4.5 | 4.8 | 16 | 5.9 | 6.9 | 84 |

perienced subject D showed larger errors, but also appeared to be bimodal, and we have analyzed them as bimodal. Columns P in Table II indicate the fractional weights of the two peaks of the distributions.

The actual data for the subjects are plotted with circles in Figs. 3–6, which show the relative match, $f_m/f_B$, as a function of the boundary frequency $f_B$. Two data points are shown for each value of $f_B$ where a bimodal distribution was indicated by the results. The sizes of the circles reflect the relative fractions of data in the upper and lower peaks of the distribution for each value of $f_B$. Error bars are two standard deviations in overall length. Where no error bars are shown the errors are smaller than the points. Error bars shown to the left of the circles show the median error in the middle range. The values of the means and errors show that we found no difficulty in deciding whether any particular point fell within the upper or the lower peak. The second inexperienced subject, R, never learned to perform the pitch matching task consistently, even after 20 runs of the 12-stimuli set.

## E. Low and high range experiments

After the completion of the middle range experiments, subjects matched pitches to dichotic noise with phase boundaries in the low and high ranges, respectively, 126–420 Hz and 567–2438 Hz. The latter range is especially interesting because it extends above the frequency region where significant binaural effects are expected. Each of these frequency ranges included seven different phase-boundary frequencies. The experiment was otherwise identical to the middle range experiment. Only the experienced subjects, G, M, and W participated in this experiment.

## F. Results, low and high ranges

The results of the dichotic experiment in the low- and high-frequency ranges are shown in Figs. 3–6 by squares and triangles, respectively. Again, the size of the symbols indicates the fractional weight in upper and lower peaks of a bimodal distribution. Comparison with the results of the middle range experiment leads to the following conclusions.

(1) Errors are larger in the low and high ranges than in the middle range. At the lowest boundary frequencies errors can be very large.

(2) Subject W exhibits bimodal behavior in the high and low ranges, similar to that exhibited by G and M in all ranges.

(3) As the boundary frequency decreases in the low range the pitch matches on the high side of the boundary become very high, exceeding the boundary frequency by 30%. Pitch matches on the low side of the boundary tend to become very low, again by 30%.

(4) In the high range most of the pitch matches are on the low side of the boundary. By contrast, in the middle range most of the matches were on the high side of the boundary.

(5) At the highest frequencies pitch matches become

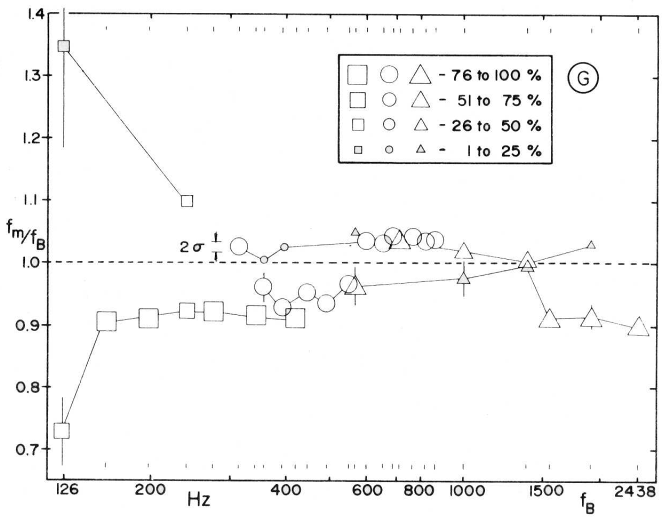


FIG. 3. The figure shows data for subject G in experiment 1 in three frequency ranges. Frequencies matched to the BEP, relative to $f_B$, are plotted as a function $f_B$. Overall length of the error bars represents twice the standard deviation. Points with no error bars have errors less than the size of the point; the median length of those errors is shown to the left of the points in the middle range. Circles represent the values of $f_B$ for the middle range, squares mark the low range, and triangles mark the upper range of $f_B$. The size of each point represents the proportion of responses on each side of $f_M = f_B$. The inset shows the point size scale.

uncharacteristically low, about 10% below the boundary frequency.

(6) It is of interest to know whether the relationship between the phase boundary frequency and the top frequency of the noise band affects the results. This question can best be studied at those frequencies where the low and high ranges overlap the middle range. The comparison shows that the results are qualitatively the same in all ranges, but there appear to be some quantitative differences between ranges. Although there are occasional individual differences, the following trends appear. When the ratio of boundary frequency to top

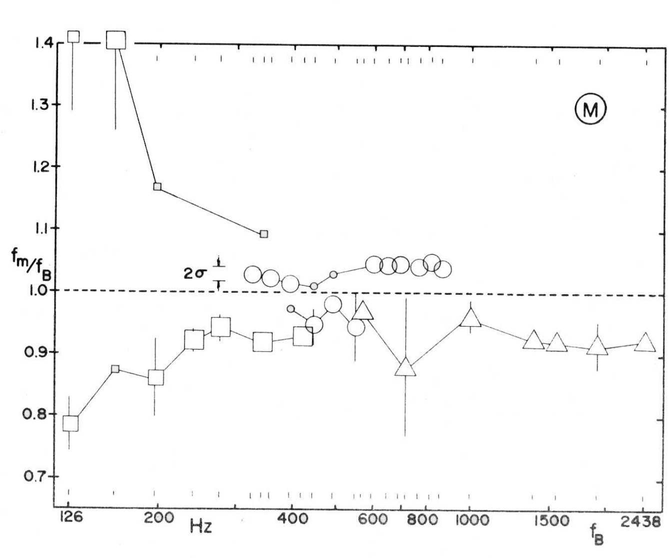


FIG. 4. Same as Fig. 3 for subject M.

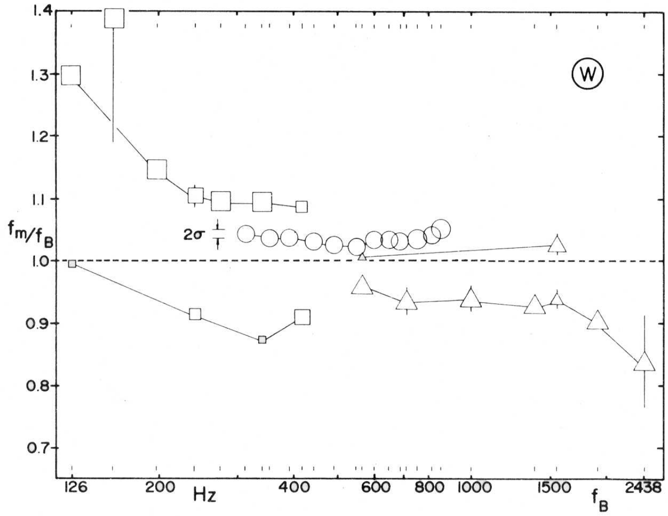


FIG. 5. Same as Fig. 3 for subject W.

frequency is made smaller (low range) the peaks of the distribution of pitch matches become more widely separated. There is an increasing tendency to match on the low side of the boundary. Data in the high range (larger ratio of phase boundary frequency to top frequency) fit more smoothly onto the middle range data. High range matches tend to fall on the lower side of the boundary somewhat more than do middle range matches. Overall, it seems that the relationship between phase-boundary frequency and top frequency does have a small effect on the BEP. However, different values of the ratio of boundary frequency to top frequency give qualitatively similar results. When the top frequency and the intercomponent spacing change by more than an octave, the BEP changes by only a few percent.

## G. Qualitative results

Because of the novelty of the BEP effect it seems worthwhile to note some qualitative results from our experiment. The pitch matching task was initially difficult for most subjects. To complete a run of 12 matches initially required about 45 min. As subjects learned the task, this time was reduced to about 15 min. On an informal basis, six other inexperienced listeners attempted the matching task. Of these, two produced matches near the phase-boundary frequencies, while the others apparently responded randomly. The experience of subject R suggests that some subjects may never be able to learn to perform the task.

Those subjects that could reliably match pitches of the dichotic noise stimulus reported that the pitch sensation sounded like a very narrow noise band added to a wideband noise. The pitch sensation was similar in strength and character to the Huggins pitch. The BEP sensation was located within the head, somewhere between the center of the head and the right ear. Presumably the asymmetry occurs because it is the right channel which changes. When the headphones were reversed, the sensation appeared between the center and the left ear.

The process involved in matching the sine tone to the BEP sensation varied in difficulty and character from trial to trial. Sometimes the pitch sensation seemed stronger. This made the task of tuning the matching tone much easier. Sometimes the pitch sensation was perceived immediately at the beginning of the trial while at other times random "searching" with the matching tone aided performance. As the matching tone approached the vicinity of the boundary frequency, the edge pitch suddenly would "pop out" and then be quite easily matched. At times the edge pitch would be quite "elusive" as the final adjustments of the matching tone were being made. As the subject was slowly increasing the frequency of a matching tone that seemed flat, the matching tone would suddenly sound sharp. This evasion would occur as the matching tone approached the BEP from either direction and continue until the subject finally gave up and settled for a less precise tuning.

## H. Discussion

The results of experiment 1 showed that subjects can reliably match a sine tone to the pitch of the dichotic noise. Pitch matches were consistently higher or lower than the phase-boundary frequency. The shift was ±4% in the standard range and increased as the phase-boundary frequency approached the upper and lower extremes.

A matching frequency shifted from the phase boundary is consistent with the hypothesis that the Equalization–Cancellation process creates a high-pass or a low-pass central spectrum with a cutoff at the phase boundary frequency. Because a noise band with a sharp edge produces a sensation of pitch which is shifted into the noise one would expect the BEP to be shifted above or below the phase-boundary frequency. It is, therefore, interesting to check the detailed behavior of the pitch of high-pass and low-pass noise to see whether it correlates with the detailed behavior observed for the BEP. This check was made in experiment 2.

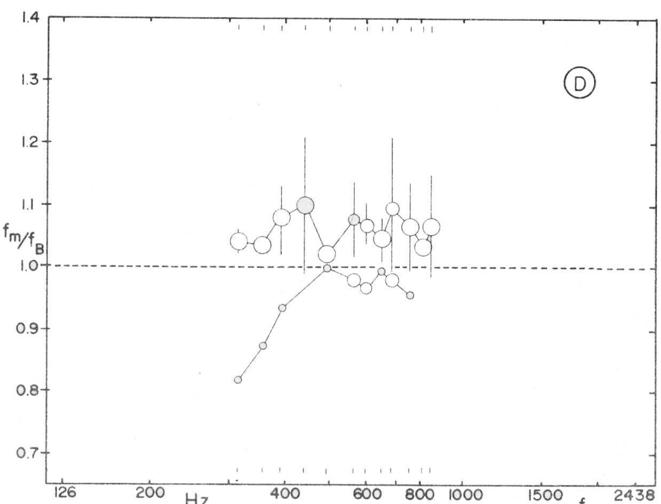


FIG. 6. Same as Fig. 3 for subject D. Subject D matched pitches only in middle range.

## II. EXPERIMENT 2

The second experiment determined the pitch elicited by high-pass and low-pass noise bands.

### A. Method

The procedure for this experiment was identical to the procedure for experiment 1. The stimulus was changed only in that the first interval did not contain the dichotic stimulus but instead contained a high-pass or low-pass noise band. These bands were created by electronically subtracting or adding the two channels of noise from experiment 1. The resulting spectra showed a 30-dB drop in level at the band edge. This combined noise stimulus was sent in phase to both ears during the first interval. The other intervals of the sequence remained the same as in experiment 1. Subjects again matched a sine tone in noise to the pitch of the edge stimulus. The experiment was done with spectral edges in the same three overlapping ranges.

### B. Results

Figures 7–9 show the results for the three participating subjects, G, M, and W. As can be seen, the frequency of the matching sine tone always deviated from the band edge and generally shifted *into* the noise band. For low-pass noise the matched pitch was below the edge frequency and for high-pass noise the matching tone was above the edge frequency. Subjects W and G commented that the pitch of the high-pass noise was easier to match than the pitch of the low-pass noise. Subject M had no preference.

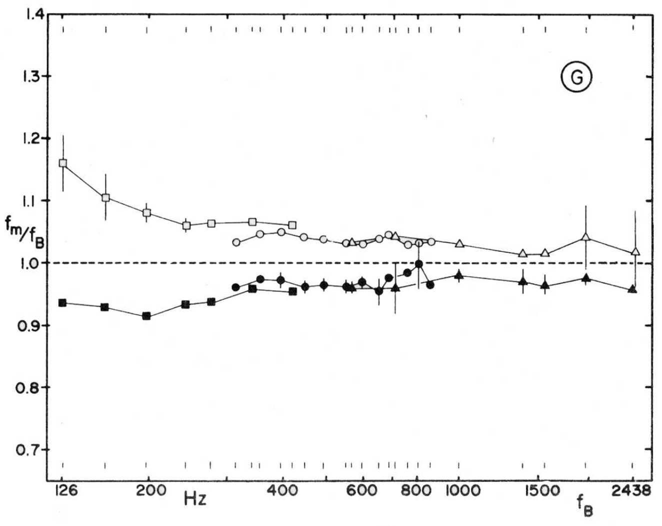


FIG. 7. The figure shows the data for subject G in experiment 2 in all frequency ranges. Open symbols represent pitch matches relative to the cutoff frequency $f_B$ for high-pass noise. Filled symbols represent pitch matches relative to the cutoff frequency $f_B$ for low-pass noise. Squares represent values of $f_B$ in the low-frequency range; circles, the middle range; and triangles the high-frequency range. Error bars show twice the standard deviation. Points with no error bars have errors smaller than the size of point. Each point represents a minimum of four trials.

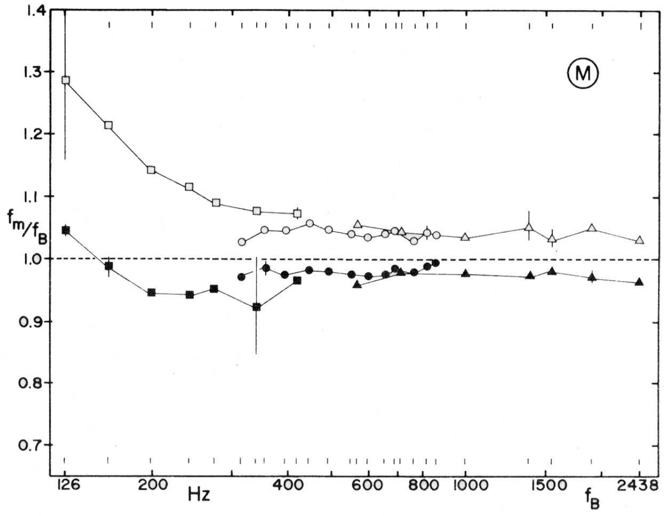


FIG. 8. Same as Fig. 7 for subject M.

### C. Discussion

The relative shifts observed in experiment 2 are, within the errors, identical to the shifts observed in experiment 1 in the middle range. The discontinuity in the data in experiment 1, where the middle range overlaps the low range, appears in the experiment 2 data as well. The rapid increase of the pitch with decreasing phase-boundary frequency in experiment 1 is similar to the increasing pitch with decreasing cutoff frequency for high-pass noise in experiment 2. The tendency for the BEP to drop at the highest frequencies of experiment 1 is not reproduced in experiment 2. As in the case of experiment 1 the pitch matches *below* the edge frequencies at the lowest frequencies in experiment 2 show considerable individual differences.

Overall, the close correspondence between the results of experiments 1 and 2 would seem to make the E–C Model a most attractive explanation for the dichotic noise pitch effect observed in experiment 1. The correspondence tends to justify the term, Binaural

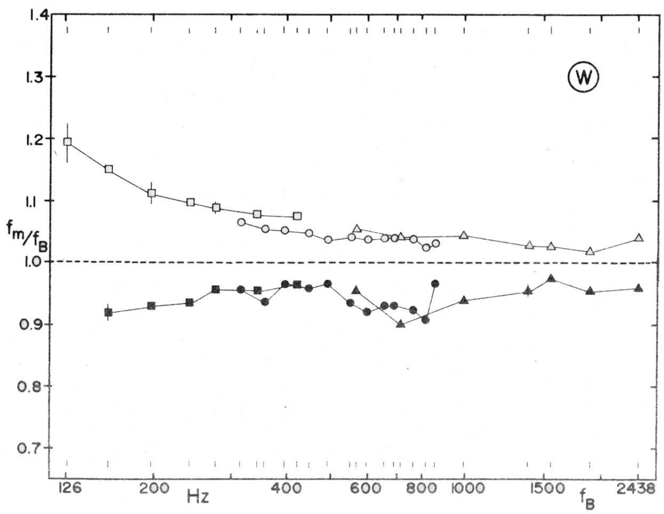


FIG. 9. Same as Fig. 7 for subject W.

Edge Pitch, for the effect. According to the E–C Model the binaural system can deal with the dichotic noise stimulus of experiment 1 by presenting either a low-pass or a high-pass noise band to the central processor. The pitch matches to the dichotic noise then simply reflect the shifts associated with low-pass or high-pass noise. The shifts themselves present an interesting challenge to auditory theory. The shifts can be understood from a place model of pitch perception in which neural excitation patterns, established by an initial spectral analysis, are subsequently modified by lateral inhibition. (See the references in the review by Bilsen, 1977.) The similarity of shifts found in the BEP and the noise band edge pitch suggests a lateral inhibition *operative on the central spectrum.*

## III. EXPERIMENT 3

In 1962 Guttman studied the strength of the Huggins pitch. His subjects adjusted a sine tone frequency and amplitude to match a Huggins stimulus for both pitch and pitch strength. Subjects found the Huggins pitch to be 4.6 dB above masked threshold for the sine tone at the matching frequency.

Experiment 3 was designed to determine the strength of the BEP relative to masked threshold and relative to the Huggins pitch. Cramer and Huggins found the existance region for Huggins pitch to be 200–1600 Hz. Our experiment 1 found that pitch matches could be made to the binaural edge stimulus up to 2438 Hz. Therefore a comparison of pitch strength in the high-frequency range is of special interest. Two of the phase boundaries in our high range occur at frequencies above 1600 Hz.

### A. Method

To make the stimuli as similar as possible both the BEP stimuli and the Huggins stimuli were generated digitally. The binaural edge generated for experiment 1 involves a discontinuous transition and so we chose to use the Wightman–Grantham–Fowler (1977) version of the Huggins stimulus as our comparison. The sensations produced by these two stimuli are quite similar and so comparison is reasonable.

BEP stimulus was produced as for experiment 1. The same spectral components were used to produce the Huggins stimulus. The one component chosen to be out of phase corresponded exactly to the phase transition boundary of the BEP stimulus. The strengths of the BEP and of the Huggins pitch were determined by having subjects match the sine tone of the third interval for both pitch and loudness. Because these dichotic noise pitch effects sound rather like a sine tone in noise the experiment is not hard to do. Masked thresholds for sine tones of the relevant frequencies were determined by a method of adjustment task run concurrently with the BEP and the Huggins matching experiments.

Two subjects, M and W from experiments 1 and 2, participated in this experiment. They matched pitch and loudness and adjusted threshold for the 19 stimuli of the standard and high ranges. Each subject performed two runs for both BEP and Huggins conditions.

### B. Results and discussion

The results of experiment 3 are shown in Fig. 10. The loudness matches and thresholds are shown on a vertical dB scale relative to the power density of the noise, i.e., the quantity ($60-10\log\Delta f$) has been subtracted from each of the measured values, where $\Delta f$ is the noise bandwidth in Hz.

Within the middle range, 315–856 Hz, the Huggins pitch and BEP are both 4–9 dB above masked threshold, though the BEP is slightly stronger than the Huggins pitch, over most of the range. Changing to the high range produces a measureable drop in strength for both pitch effects. Above 800 Hz individual differences are significant. For subject M the strengths of both pitch effects were similar and decreased slowly, relative to masked threshold, for increasing frequency. For subject W the strength of the Huggins pitch fell rapidly near 800 Hz whereas the BEP strength decreased only slowly.

Our observation that both BEP and Huggins pitch can be matched for frequencies above 2 kHz was surprising. The persistence of these pitch effects at high fre-

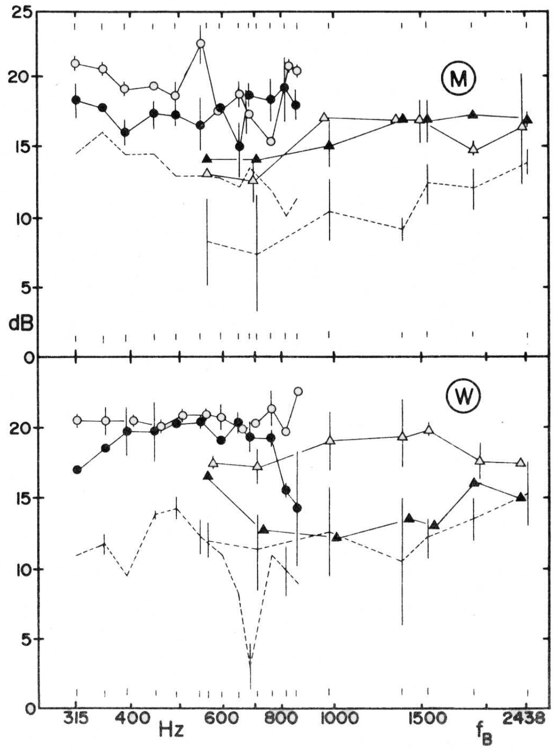


FIG. 10. The strength of the pitch sensations from the Huggins effect (filled symbols) and the BEP (open symbols) in the middle (circles) and high (triangles) frequency ranges. The dashed line shows masked threshold for a sine tone of corresponding frequency in the two ranges. Error bars show twice the standard deviation. All points are plotted relative to the sound power density of the noise. The upper half of the figure shows the data for subject M. The lower half shows the data for subject W.

quency seems inconsistent with other studies of Huggins pitch. (Note that Bilsen quotes an upper limit of 2 kHz for dichotic noise pitch effects of this kind.) The discovery of the high-frequency effect prompted a renewed and unsuccessful search for stimulus artifacts. We concluded that our observation was the result of particularly favorable circumstances, present in our experiment and not present in previous dichotic pitch experiments. These are as follows:

(1) Subjects were allowed unlimited exposure to the stimulus before making a pitch match.

(2) The matching tone, used to indicate pitch, also served as a valuable aid to finding the anomalous frequency within the noise. The subject could use the matching tone as a (nonsimultaneous) probe to focus his attention on one spectral region.

(3) The dichotic stimulus was immediately preceded and followed by diotic noise with a power spectrum which was identical to the power spectrum for the signals in the two ears on the dichotic interval. Changes resulting from dichotic presentation were emphasized by this procedure. The subject could hear the pitch of the dichotic stimulus turn on and off. The subject knew when to expect the dichotic interval because of an indicator light.

(4) The diotic noise was continued through the matching interval, which made the matching interval sound rather similar to the dichotic interval.

## IV. FURTHER EXPERIMENTS

There are a number of possible variations of experiment 1, the standard BEP experiment. Below we describe variations which produced *no change* in the basic result. Under all the following experimental conditions subjects continued to match the BEP at approximately plus or minus 4% deviation from the phase-boundary frequency.

The experiments of this section were confined to the 12 phase-boundary frequencies of the middle range, 315–856 Hz. Two subjects, M and W, performed at least two runs for each experiment. Because the Equalization–Cancellation Model of binaural phenomena seems to account well for the BEP, we note the implications of experiments which follow for that particular model.

(1) *Reversed discontinuous binaural edge*: This experiment is identical to experiment 1 except that the phase relations are opposite to those in experiment 1, i.e., the noise components with frequency below the transition are out of phase and components above transition are in phase. The fact that no changes were found in this experiment shows that the binaural system is capable of introducing phase shifts before cancellation to create an individual optimum central spectrum.

(2) *Quadrature discontinuous binaural edge*: In this experiment the interaural phase shift is $\pi/2$ for all components below the transition frequency and is $3\pi/2$ for all components above that frequency. An entertaining feature of this stimulus is that we cannot observe any transition by power spectrum measurements on the sum and difference of the two channels. To observe this phase boundary we need to put on the headphones and listen to the dichotic noise. It sounds no different from the stimulus of experiment 1. This result suggests that the binaural equalization process is able to introduce an arbitrary phase shift and is not limited to phase shifts of 0 or $\pi$.

(3) *Standard discontinuous binaural edge superimposed on a gradual phase shift*: In this experiment we added a gradual frequency-dependent phase shift to one of the stimulus channels. The gradual phase shift varied from $\pi$ to $-\pi$. The phase shift was 0 at 500 Hz, and the $\pi/2$ and $-\pi/2$ points occurred at 200 and 1400 Hz. This smooth phase variation closely resembles an interaural time delay of 1 ms for components below 500 Hz. Whereas our other experiments use a constant (or zero) interaural phase change plus a discontinuity, the present experiment approximates a constant interaural time delay plus a discontinuity. The BEP is unaffected by this change.

The significance of our null result is this. In his review of the E–C Model, Durlach (1972) complained that the frequency-independent internal phase shift presumed by the model equalization process seems unrealistic. Auditory theory makes it seem more plausible that the binaural system introduces a constant internal delay. For our standard experiments the binaural system cannot completely cancel a band of noise by a constant internal time delay. For the present experiment the binaural system cannot completely cancel a band of noise by a constant phase shift. Either the binaural system can apply both time delay and phase shift or else the E–C process is preceded in the auditory system by a frequency analysis. (If the E–C operation is restricted to a small enough frequency bandwidth then there is no meaningful difference between time delay and phase shift.) The latter hypothesis is consistent with the contemporary view of the auditory system.

(4) *Diffuse binaural edge*: Experiments were done in which the interaural phase shift varied linearly from 0 to $\pi$ or from $\pi$ to 0 (reversed edge) over a frequency range which was 10% or was 20% of the center frequency. These conditions are then somewhat analogous to the original Huggins experiment with an analog all-pass filter. It was interesting to note that the errors did not significantly increase in this condition. Experimental pitch matches, as a fraction of the center frequency of the boundary region, were the same as in experiment 1. This invariance could not have been anticipated because the central spectrum predicted by the E–C Model is considerably widened by the diffuse edge.

(5) *Standard discontinuous binaural edge at reduced intensity*: We performed experiment 1 at 40 dB SPL and at 30 dB SPL. Errors at 40 dB were larger than those at the standard 60 dB by a factor of about 2. At 30 dB the errors were twice those at 40 dB. We found it impossible to perform the experiment at 20 dB. Like the Huggins pitch, the BEP is best heard at intermediate sound pressure levels.

(6) *Standard discontinuous binaural edge with restricted noise band*: As one of our tests for artifacts we reduced the upper frequency range of the noise band. We did this by two different techniques. We removed sine components from the digitally generated signal or we low-pass filtered the noise bands with an analog low-pass filter with a slope of 48 dB/oct. We observed only small changes in pitch, which confirms that the BEP is not very sensitive to the details of the noise spectrum.

(7) *Standard discontinuous binaural edge with matching tone in quiet*: As is well known, the pitch of a sine tone is raised by the addition of noise. In experiment 1 the matching sine tone was presented in a noise background to increase perceptual similarity with the dichotic noise interval. It is possible, however, to perform the matching task with the matching tone in quiet, and we did this experiment to estimate the effect which the background noise might have on the matching experiment. As expected the frequency of the sine tone matched in quiet was slightly higher than that obtained in the standard experiment. But it was only 1% higher, and the errors were rather large.

## V. IMPLICATIONS FOR AUDITORY THEORY

The data presented in this paper, on the pitch of diotic noise bands and the BEP, have implications for auditory theory. Thus far we have couched the presentation in terms of a central place model of pitch perception, a model in which pitch is derived from a central spectrum, created in the case of dichotic stimuli by the E–C process. But there are alternative models of pitch perception and binaural interaction. Below we examine several alternative models in the light of our experiments.

### A. Pitch of noise bands

High-pass and low-pass noise bands have a pitch which is close to the edge of the band but shifted into the noise. It is conceivable that temporal aspects of the noise waveforms are responsible for the pitch of the noise bands. It is well known that a discontinuity in the spectrum is associated with oscillations in the waveform, with a frequency close to that of the spectral discontinuity. To investigate this conjecture we made strip chart recordings of our low-pass and high-pass noise stimuli for the three experimental ranges. We gauged the periodicity by eye and determined the period of each "cycle." We calculated a mean $\overline{T}$ and standard error. The quantity $R \equiv (f_B \overline{T})^{-1}$ is a plausible estimate for the prediction of a timing model for our experimental quantity, $f_m/f_B$, plotted in the figures.

The predictions for the low-pass noise bands were as follows: for $n_B = 40$, $R = 0.86$ (23%); for $n_B = 100$, $R = 0.78$ (28%); for $n_B = 180$, $R = 0.61$ (38%). The numbers in parenthesis are the standard deviations as a percentage of the mean. The values of $R$ are considerably smaller than the experimental values of $f_m/f_B$. Further, the dependence of $R$ upon $n_B$ is not in agreement with the dependence of the experimental values of $f_m/f_B$. The standard deviation is large. Clearly a temporally based pitch perception mechanism would have to cope somehow with a very noisy record.

We were unable to determine reasonable periodicity estimates for high-pass noise stimuli. A study of waveform oscillations produced values of $R$ greatly in excess of 1.0, and the waveform was too ragged to fit a plausible envelope. By contrast there is no important difference between high-pass and low-pass noise in the strength of pitch sensations. In sum, there is information in the temporal fine structure and the envelope which could conceivably code for pitch, but any practical model for a temporal pitch perception process must contend with some very serious problems in the case of noise band edge pitch.

An alternative temporal model for pitch perception is based upon the autocorrelation function of the stimulus. Neural autocorrelator models have had a widespread appeal (cf. Licklider, 1959). We have derived an expression for the autocorrelation function for digital noise. The result, given in the Appendix, is that the autocorrelation function oscillates with the frequency of the spectral edge. Therefore an ideal autocorrelator cannot explain the pitch shifts away from the edge frequency observed experimentally. A neural autocorrelator, of course, might not be perfect. One can imagine ways in which a neural autocorrelator might be vulnerable to certain pitch shift effects such as diplacusis or dependence on intensity. It is considerably more difficult, though obviously not impossible, to understand how a neural autocorrelator could produce the bidirectional shifts observed in the noise band edge experiments.

### B. Binaural edge pitch

The BEP data provide evidence against models of binaural pitch perception in which pitch is extracted from the cross-correlation function. This can be seen from the following argument. Consider the noise bands used in the diotic experiment discussed above. These are created by electronically adding or subtracting the two channels of the BEP stimulus. By expanding the product in the autocorrelation integrand for the diotic noise bands one can see that the oscillations at the phase-boundary frequency are identically those of the cross-correlation function for the BEP stimulus. Therefore the cross-correlation function must oscillate at the phase boundary frequency. Thus the cross-correlation model fails to account for the shifts in the BEP for the same reason that the autocorrelation function fails to account for the pitch shifts associated with noise band edges.

The cross-correlation process was suggested by Licklider in 1956. He added the cross-correlation process to his duplex model of pitch perception to create the triplex model, an expansion which was required by the discovery of the Huggins pitch effect a few years earlier. As many authors have noted, the Huggins pitch effect is equally well explained by Licklider's triplex theory or by the E–C Model. This is not true of the BEP, and this is the special value of the BEP for auditory theory. It seems to us that the shifts found in the BEP make the E–C Model a plausible explanation for binaural pitch effects and make the triplex model an implausible explanation.

The neural cross-correlation mechanism proposed by Bilsen (1977) and by Raatgever and Bilsen (1977) is entirely different from Licklider's mechanism. Bilsen's mechanism is essentially a neural process which performs the equalization function required in the E–C Model, but with this added feature. The equalization operation results in lateralization of the pitch sensation according to the neural delay required. Most of our experimental stimuli would not seem to require internal equalization. They are presented in optimum form for cancellation and, therefore, according to Bilsen, should be centered in the head. We have not as yet sufficiently explored the lateralization of the BEP sensation to be able to say whether the positive and negative shifts are related to different locations in the head. Nor have we noted any particular lateralization effect in the two conditions (experiments 2 and 3 in Sec. IV) which require internal equalization.

## VI. CONCLUSION

We have found a dichotic noise pitch effect which has not been previously reported. A sensation of pitch is created by dichotic noise with an interaural phase variation from 0 to $\pi$ over a narrow frequency range. We called this effect, Binaural Edge Pitch (BEP). The effect is strongest (4–9 dB above masked threshold) for phase boundary frequencies between 300 and 800 Hz. With favorable experimental conditions the effect can be heard for boundary frequencies as low as 125 Hz and as high as 2400 Hz, though it is weaker and pitch matching errors are larger. For experienced listeners the BEP is similar in nature and strength to the Huggins pitch (overall phase variation from 0 to $2\pi$). The BEP is present for noise at 60 and 30 dB and for phase boundary widths which are $\frac{1}{2}\%$, 10%, and 20% of the phase boundary frequency. Pitch matching experiments find that the spread of matches is 1%–2% of the phase boundary frequency, as for the Huggins pitch, and that the BEP does not depend upon the direction of the phase change at the boundary.

Like the Huggins pitch effect the BEP finds a natural explanation in terms of the Equalization–Cancellation Model of binaural processing. According to the model the binaural system processes the dichotic noise by changing the interaural phase and subtracting left and right channels so as to produce a central spectrum with a sharp edge. The central spectrum created by the binaural system may be either high-pass or low-pass. The strongest evidence in favor of this explanation is that the BEP is shifted above or below the phase-boundary frequency. The upward and downward shifted BEP's correspond well with the pitches of high-pass and low-pass noise bands, respectively, with edges in the *physical* spectrum.

If one accepts this explanation it is naturally of interest to try to find a pattern in the distribution of pitch matches above and below the phase-boundary frequency. Under what conditions does the binaural system output a high-pass noise or a low-pass noise? Our experiments provide no clear answer to this question. Often subjects switch from one mode to the other, though in the middle range one subject consistently chose the high-pass central spectrum. There is a clear tendency to prefer the low-pass central spectrum when phase boundary frequencies are high and within 30% of the top of the noise band. Evidently the binaural system does *not* choose to minimize the noise power. Analysis of the distribution for low phase-boundary frequencies reveals only individual differences.

An overall view of the pitch matching data shows some aspects of a regression effect in the high and low phase boundary ranges. This explanation for the phase-boundary dependence of the pitch, however, disagrees with most of the low-frequency data for two out of three subjects. We believe that the frequency dependence of the distribution provides potentially useful information on binaural processing; but we do not yet know how to deal with the individual differences.

Both the Huggins pitch and the BEP tend to support the E–C Model. However, the BEP provides the stronger support because the Huggins pitch effect can equally well be explained by a cross-correlation model. By contrast, the frequency shifts observed in the BEP effect suggest that the E–C process and the resulting central spectrum are, in fact, responsible for dichotic noise pitch phenomena. The supplementary experiments of Sec. IV indicate that the E–C process is a flexible one, with parameters which a subject can tune to his best advantage.


## ACKNOWLEDGMENT

Supported by the NSF, grant, BNS 79–14155.


## APPENDIX: AUTOCORRELATION FUNCTION FOR NOISE

The autocorrelation function is the Fourier transform of the power spectrum. We consider the autocorrelation function for two kinds of noise.

(1) *Constant spectral density over a continuous frequency range*. Suppose that the power spectrum is a continuous rectangular density, finite from a low angular frequency $\omega_L$ to a high frequency $\omega_H$ and otherwise zero. Then the autocorrelation function is given by

$$C(T) = [\sin(\omega_H T) - \sin(\omega_L T)]/[(\omega_H - \omega_L)T]. \qquad \text{(A1)}$$

It is a function with oscillations at the frequencies of the spectral boundaries.

(2) *Uniform but discrete spectral density.* When noise is generated digitally by adding sine waves of equal amplitude and random phase angles the autocorrection function is slightly different. We consider the case where the spectral components have equal spacing $\omega_0$. Suppose that the lowest spectral component has angular frequency $(N+1)\omega_0$ and that the highest spectral component has frequency $M\omega_0$. Then the autocorrelation function is given by

$$C(T)=\frac{1}{M-N}\,\mathrm{Re}\sum_{n=N+1}^{M} e^{i\omega_0 nt}, \qquad (A2)$$

$$C(T)=\frac{1}{2(M-N)}\left(\frac{\sin\omega_0 T(M+\frac{1}{2})}{\sin(\omega_0 T/2)}-\frac{\sin\omega_0 T(N+\frac{1}{2})}{\sin(\omega_0 T/2)}\right). \qquad (A3)$$

Note that for the standard low-passed noise index $N=0$, and the second term above equals 1.0. The autocorrelation function does not oscillate at the frequencies of the extreme spectral components. Instead it oscillates with frequencies which are above the highest component and below the lowest component and below the lowest component by half a spectral line spacing. An alternative form of Eq. (A3) is the envelope and sum form

$$C(T)=\frac{\cos[\omega_0 T(M+N+1)/2]\sin[\omega_0 T(M-N)/2]}{(M-N)\sin(\omega_0 T/2)}. \qquad (A4)$$

[1]An analog all-pass filter cannot produce the required overall $\pi$ phase shift within a narrow frequency region.


Bilsen, F. A. **(1977)**. "Pitch of noise signals: Evidence for a central spectrum," J. Acoust. Soc. Am. **61**, 150–161.

Cramer, E. M., and Huggins, W. H. **(1958)**. "Creation of pitch through binaural interaction." J. Acoust. Soc. Am. **30**, 413–417.

Durlach, N. I. **(1962)**. "Note on the creation of pitch through binaural interaction," J. Acoust. Soc. Am. **34**, 1096–1099.

Durlach, N. I. (1963). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. 35, 1206–1218.

Durlach, N. I. **(1972)**. "Binaural signal detection: Equalization and cancellation theory," in *Foundations of Modern Auditory Theory*, Vol. II, edited by J. V. Tobias (Academic, New York).

Fastl, H. **(1971)**. "Uber Tonhohenmpfindungen bei Rauschen," Acustica 25, 350–354.

Guttman, N. **(1962)**. "Pitch and loudness of a binaural subjective tone," J. Acoust. Soc. Am. **34**, 1996, and Bell Telephone Laboratories Memo.

Licklider, J. C. R. **(1956)**. "Auditory Frequency Analysis," in *Proceedings Third Symposium on Information Theory*, edited by C. Cherry (Academic, New York), pp. 253–268.

Licklider, J. C. R. **(1959)**. "Three Auditory Theories," in *Psychology, A Study of a Science*, edited by S. Koch (McGraw-Hill, New York), pp. 41–144.

Raatgever, J., and Bilsen, F. A. **(1977)**. "Lateralization and dichotic pitch as a result of spectral pattern recognition," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 443–453.

Wightman, F. L., Grantham, D. W., and Fowler, C. **(1977)**. "Existence region of dichotic noise pitch," J. Acoust. Soc. Am. Suppl. 1, **62**, S24.