Lateralization of rectangularly modulated noise: Explanations for counterintuitive reversals

Richard M. Stern, Torsten Zeppenfeld, and Glenn D. Shear
Department of Electrical and Computer Engineering and Biomedical Engineering Program, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 1 May 1990; accepted for publication 4 June 1991)

A number of years ago Hafter and Shelton [J. Acoust. Soc. Am. Suppl. 1 **68**, S16 (1980), Hafter and Shelton, J. Acoust. Soc. Am. **90**, 1901–1907 (1991)] demonstrated that diotic bandpass noise that is multiplied by a dichotic pair of periodic rectangular gating functions could be strongly perceived toward the ear receiving the signal with the gating function that is *lagging* in time. These results were surprising because sounds presented with interaural time delays (ITDs) are normally perceived toward the ear receiving the signal that is *leading* in time. In order to better understand these counterintuitive phenomena an analytical expression for the cross-correlation function of these gated-noise stimuli is derived, and it is shown that many of the general properties of the data can be accounted for by considering only the properties of this function. The lateralization of the gated-noise stimuli is compared to predictions of three types of models of binaural signal processing: models based on the cross-correlation function of the entire signal, models based on the interaural group delay or phase delay at low frequencies, and the position-variable model, which is based on the response of the peripheral auditory system to the stimuli. It is shown that the observed lateralization results, including the unexpected reversals and various frequency effects, can be explained without further assumption by an extended implementation of the position-variable model [Stern and Colburn, J. Acoust. Soc. Am. **64**, 127–140 (1978)]. The cross-correlation functions of the original stimuli can also be related to the lateralization data, but in a much less natural fashion. It is argued that models based on the interaural group delay of the original stimuli will not describe the observed results, although low-frequency phase delay is qualitatively consistent with the observed lateralization phenomena.

PACS numbers: 43.66.Ba, 43.66.Nm, 43.66.Pn, 43.66.Qp [WAY]

INTRODUCTION

A number of years ago, Hafter and Ricard (1973) and Hafter *et al.* (1980) first described an interesting and unexpected phenomenon concerning the lateral position of diotic bandpass noise that is subsequently multiplied by a dichotic pair of periodic rectangular gating functions. They found that when the gating functions were presented with an interaural time delay (ITD), the gated-noise stimuli would be strongly perceived toward the ear receiving the signal with the gating function that was *lagging* in time, for many combinations of stimulus parameters. These results were surprising because one of the classic tenets of binaural hearing has been that sounds presented with ITDs are perceived toward the ear receiving the signal that is *leading* in time. In a more detailed presentation of these phenomena that appears as a companion paper in this issue of the journal, Hafter and Shelton (1991) describe several additional properties of the perception of the gated-noise stimuli: (1) The period of the gating function has virtually no effect on the subjective laterality of the stimuli; (2) the subjective laterality of the gated-noise stimuli is unaffected by changing the duration of the gating pulses and the center frequency of the bandpass filter (also referred to as the carrier frequency), provided that the product of these two quantities is held constant; and (3) the perceived lateralization of the stimuli is dominated by stimulus components below approximately 1500 Hz, even when all components of the bandpass noise are much higher in frequency.

Hafter and Shelton (1991) also describe predictions of a simple *ad hoc* model that uses a quantity closely related to the centroid of the cross-correlation function of the gated-noise stimuli to predict their subjective lateral position. The model uses cross-correlation functions that Hafter and Shelton approximate empirically using simulation techniques, and it describes the data quite well. When originally presented, this model had previously been criticized by Henning (1983), who believed that the binaural system was unable to make use of group-delay information at low frequencies, but subsequent experiments by Bernstein and Trahiotis (1985) demonstrated that these cues can in fact be used (at least to some extent) to lateralize sound.

In order to understand the mechanism that is likely to mediate these counterintuitive phenomena, an analytical expression for the cross-correlation function of the gated-noise stimuli is developed, which is expressed as the product of the autocorrelation function of the bandpass noise and the cross-correlation function of the gating functions. The lateralization data of Hafter and Shelton are then discussed in

terms of theoretical predictions produced by three types of binaural processing: models based on the cross-correlation function of the entire stimuli to the two ears (including that of Hafter and Shelton, 1991), models based on the interaural group delay and interaural phase delay of the stimuli, and an extended implementation of the position-variable model (Stern and Colburn, 1978). It is argued that of the three types of processing mechanisms considered, the extended position-variable model provides the best description of the lateralization data, including the unexpected reversals in lateralization and the dominant role played by low-frequency stimulus components.

The analytical expression for the gated-noise stimuli is developed in Sec. I, and it is shown that these expressions describe the empirically derived cross-correlation functions measured by Hafter and Shelton. It is shown in Sec. II that the dependence of perceived lateral position on the period of the gating pulses, and the joint dependence of lateralization on the center frequency of the noise and the duration of the pulses, are predicted by virtually *all* cross-correlation-based models, including the model of Hafter and Shelton (1991). In Sec. III the predictions of models based on the interaural group delay and phase delay of the stimuli are considered, and it is argued that group delay of the stimulus *per se* is unlikely to mediate the observed reversals in laterality. Finally, the predictions of the extended position-variable model are compared to the data in Sec. IV, and this model is compared to the original model proposed by Hafter and Shelton (1991).

I. CROSS-CORRELATION FUNCTIONS FOR RECTANGULARLY MODULATED NOISE

In this section we derive the cross-correlation functions for the rectangularly modulated noise stimuli used in the experiments of Hafter and Shelton (1991).

The signals to the two ears, $x_L(t)$ and $x_R(t)$, are generated by multiplying a bandpass noise process $n(t)$ by each of two periodic gating functions, $g_L(t)$ and $g_R(t)$, as shown in Fig. 1. The bandpass noise $n(t)$ is assumed to be created by passing a white Gaussian noise process $w(t)$ through an ideal bandpass filter. [The stimuli used in the experiments of Hafter and Shelton (1991) are actually generated by realizable filters with upper and lower skirts of 24 dB per octave. The differences between the ideal and practical bandpass filters do not appear to be important for this analysis.] Specifically,

$$x_L(t) = n(t)g_L(t) \tag{1a}$$

and

$$x_R(t) = n(t)g_R(t), \tag{1b}$$

where

$$n(t) = w(t) * h_{BP}(t). \tag{2}$$

The impulse response of the bandpass filter $h_{BP}(t)$ is the inverse Fourier transform of the transfer function

$$H_{BP}(f) = \begin{cases} 1, & f_c - W/2 \leqslant |f| < f_c + W/2, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

In the experiments of Hafter and Shelton (1991), the bandwidth W is fixed at 1000 Hz, and the center frequency f_c was varied as a parameter.

The periodic rectangular modulating (or gating) functions used by Hafter and Shelton (1991) are characterized by three parameters: modulation delay (which we represent by the symbol T_M), pulse duration (T_D), and gating-pulse period (T_P). In the experiments of Hafter and Shelton (1991), T_M is fixed at 37.5 μs, T_D is either 100, 200, or 400 μs, and T_P is either 1, 2, or 4 ms (but most commonly 2 ms). These gating functions can be modeled by the periodic stationary random processes

$$g_L(t) = \begin{cases} 1, & \theta < t \leqslant T_D + \theta, \\ 0, & T_D + \theta < t \leqslant T_P + \theta \end{cases} \tag{4a}$$

and

$$g_R(t) = \begin{cases} 1, & -T_M + \theta < t \leqslant T_D - T_M + \theta, \\ 0, & T_D - T_M + \theta < t \leqslant T_P - T_M + \theta, \end{cases} \tag{4b}$$

with $g_L(t) = g_L(t - T_P)$ and $g_R(t) = g_R(t - T_P)$ [i.e., $g_L(t)$ and $g_R(t)$ are periodic with period T_P]. The fixed random parameter θ is independent of the process $n(t)$ and uniformly distributed over the interval from 0 to T_P. This parameter has the effect of randomizing the absolute time origin, and it is needed to ensure that $x_L(t)$ and $x_R(t)$ are stationary. It has no direct effect on any of the theoretical predictions.

The functions $g_L(t)$ and $g_R(t)$ are sketched in Fig. 2(a), using an arbitrary value for the parameter θ.

The cross-correlation function of the stimuli is

$$R_x(\tau) = E[x_L(t)x_R(t-\tau)]. \tag{5}$$

Because the gating functions $g_L(t)$ and $g_R(t)$ are statistically independent of the bandpass noise process $n(t)$, the cross-correlation function of the stimulus factors into the product of the autocorrelation function of the bandpass noise and the cross-correlation function of the two rectangular gating functions:

$$\begin{aligned} R_x(\tau) &= E[x_L(t)x_R(t-\tau)] \\ &= E[n(t)g_L(t)n(t-\tau)g_R(t-\tau)] \\ &= E[n(t)n(t-\tau)]E[g_L(t)g_R(t-\tau)] \\ &= R_n(\tau)R_g(\tau), \end{aligned} \tag{6}$$

with the expectation in the latter case taken over the random parameter θ.

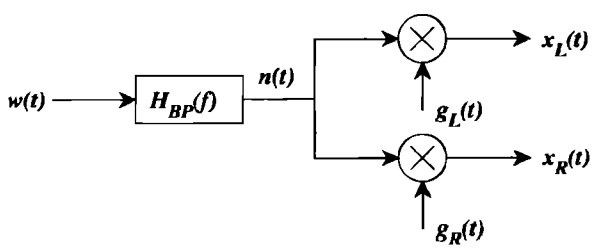


FIG. 1. Block diagram summarizing the generation of the stimuli to the two ears, $x_L(t)$ and $x_R(t)$, by multiplication of a bandpass noise process $n(t)$ by each of two periodic gating functions, $g_L(t)$ and $g_R(t)$.

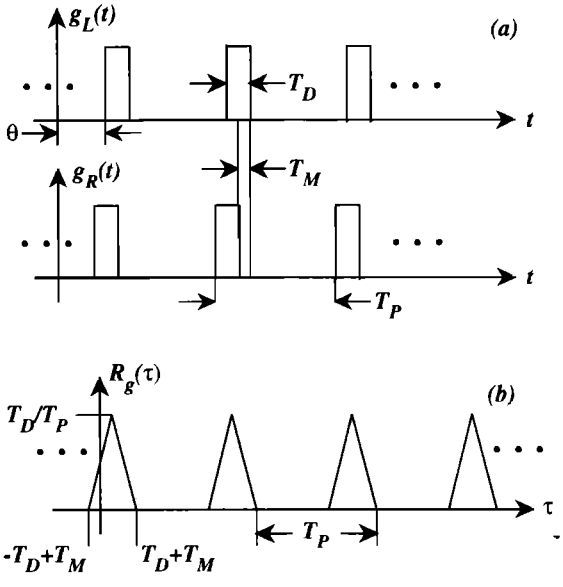


FIG. 2. (a) The gating functions used to generate the stimuli for the lateralization experiment. The parameter T_M refers to the modulation delay, T_D refers to the pulse duration, and T_P refers to the pulse period. (b) The cross-correlation function of the gating function $R_g(\tau)$.

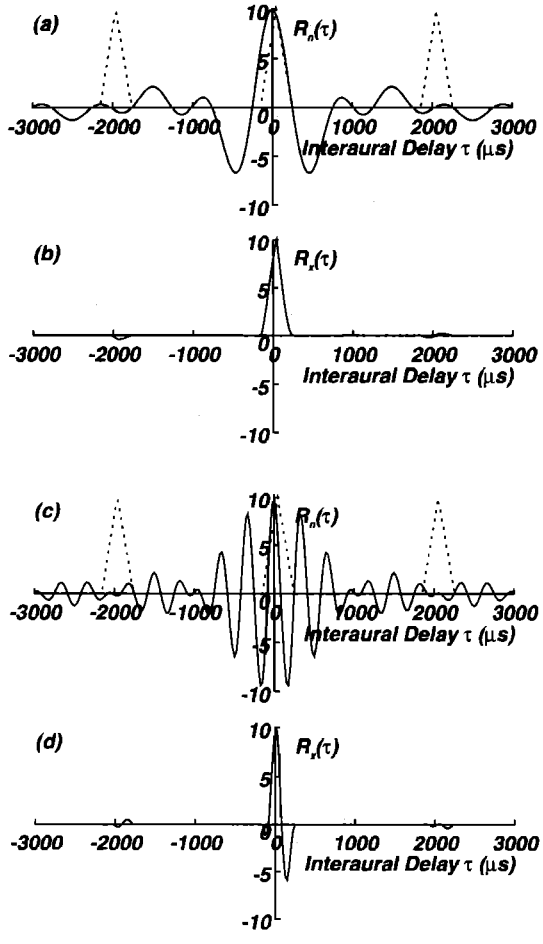


FIG. 3. Development of cross-correlation functions for gated-noise stimuli with a value of 200 μs for T_D, 2 ms for T_P, 50 μs for T_M, and two values of f_c, 1 kHz [panels (a) and (b)], and 3 kHz [panels (c) and (d)]. In panels (a) and (c), $R_n(\tau)$ and $R_g(\tau)$ are plotted as the solid and dotted curve, respectively. Here $R_x(\tau)$, the product of $R_n(\tau)$ and $R_g(\tau)$, is shown in panels (b) and (d) for each of the two center frequencies.

Here $R_n(\tau)$, the autocorrelation function of the bandpass noise before the modulation process, can be obtained from its Fourier transform, the corresponding power spectral density function,

$$S_n(f) = (N_0/2)|H_{BP}(f)|^2. \tag{7}$$

For the bandpass stimuli used in this experiment it is easy to show that

$$R_n(\tau) = N_0 W[\sin(\pi\tau W)/\pi\tau W]\cos(2\pi f_c\tau). \tag{8}$$

The function $R_g(\tau)$ is sketched in Fig. 2(b) for the gating functions used by Hafter and Shelton. As can be seen, $R_g(\tau)$ is periodic with respect to the variable τ with period T_P, and is zero for many values of τ. Consequently $R_x(\tau)$ is also zero for many values of τ.

For the stimuli in the present experiment, the nonzero values of $R_x(\tau)$ that are closest to τ equals zero are

$$R_x(\tau) = N_0 W \frac{\sin(\pi\tau W)}{\pi\tau W}\cos(2\pi f_c\tau) \times \frac{T_D}{T_P}\left(1 - \frac{|\tau - T_M|}{T_D}\right), \tag{9}$$

where N_0 is the power per Hz of the white noise before it is passed through the bandpass filter. Equation (9) is valid for τ in the range of $T_M - T_D \leqslant \tau \leqslant T_M + T_D$, and $R_x(\tau)$ is zero for other values of τ that are small in magnitude.

In order to provide a more intuitive feel for how the functions $R_n(\tau)$ and $R_g(\tau)$ interact to produce $R_x(\tau)$, these three functions are plotted in Fig. 3, using values of 200 μs for T_D, 2 ms for T_P, 50 μs for T_M, and two values of f_c, 1000 and 3000 Hz. We plot $R_n(\tau)$ and $R_g(\tau)$ as the solid and dotted curve, respectively, in Fig. 3(a), for a value of 1000 Hz for f_c; their product $R_x(\tau)$ is shown in Fig. 3(b). These curves are repeated in Fig. 4(c) and (d), but with a value of 3000 Hz for f_c. It can be seen in Fig. 3 that as f_c increases, the function $R_x(\tau)$ changes shape as different positive and negative modes of the factor $\cos(2\pi f_c\tau)$ fall within the envelope produced by the product of $R_g(\tau)$ and the $\sin(\pi\tau W)/\pi\tau W$ factor of $R_n(\tau)$. These functions closely resemble the corresponding empirically derived cross-correlation functions shown in Fig. 8 of Hafter and Shelton (1991).

II. LATERALIZATION BASED ON THE CROSS-CORRELATION FUNCTION OF THE ENTIRE STIMULUS

Hafter and Shelton (1991) describe the lateralization of rectangularly modulated noise, including the reversal phenomena, in terms of a model that is based on the cross-correlation function $R_x(\tau)$ of the entire stimuli to the two ears. In this section we consider the observed dependence of their data on the gating-function period T_P, and on the product of the gating-pulse duration T_D and center frequency f_c, and we show that these dependencies can easily be related to corresponding dependencies of the cross-correlation function $R_x(\tau)$. As a consequence, it is argued that these proper-

ties of the data would be predicted by *any* model that is based on $R_x(\tau)$, including the model of Hafter and Shelton. This section is concluded with a brief comparison of the Hafter and Shelton model to the more detailed position-variable model that is discussed later in the paper.

A. Dependence on T_P

Hafter and Shelton (1991) report that the specific value of the pulse period T_P has very little effect on lateralization performance.

Figure 4 shows plots of the cross-correlation function of the stimuli using values of 200 μs for T_D, 1000 Hz for f_c, and 1 ms for T_P. As in Fig. 3, we plot the functions $R_n(\tau)$ and $R_g(\tau)$ in Fig. 4(a), and their product $R_x(\tau)$ in Fig. 4(b). By comparing the plots of Fig. 4 to Fig. 3(a) and (b), it can be seen that the specific value of the parameter T_P has no effect on the shape of $R_x(\tau)$ for values of τ near zero, provided that T_P is greater than twice T_D. Here T_P only affects the locations of the secondary modes of $R_g(\tau)$ that occur near $|\tau| = \pm 1$ ms for the functions shown in Fig. 4. These secondary modes of $R_g(\tau)$ are believed to have little effect on perception for two reasons. First, the modes of $R_x(\tau)$ tend to become attenuated as $|\tau|$ increases because of the damping nature of the tails of the $\sin(\pi\tau W)/\pi\tau W$ component of $R_n(\tau)$. Second, both psychophysical and physiological evidence supports the contention that there are relatively few processing units in the binaural system that are specifically sensitive to stimulus time delays of large magnitude (e.g., Kuwada and Yin, 1983; Kuwada *et al.*, 1987), so the major contribution of $R_x(\tau)$ to the lateralization process is likely to come from values of τ that are close to the origin (Hafter and DeMaio, 1975; Shear, 1987; Stern *et al.*, 1988).

B. Joint dependence on T_D and f_c

Hafter and Shelton (1991) show that the dependence of the perceived lateral position on center frequency f_c and

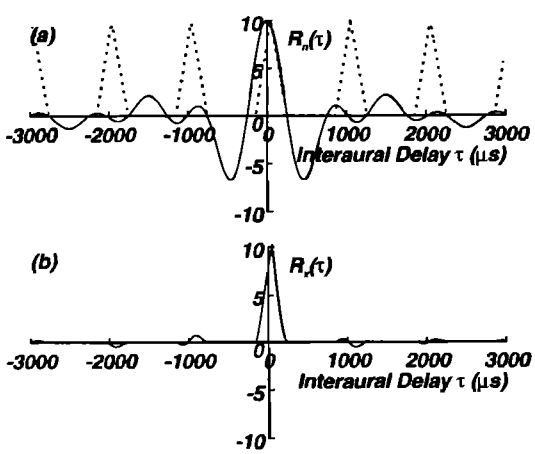


FIG. 4. Development of cross-correlation functions for gated-noise stimuli with a value of 200 μs for T_D, 1 ms for T_P, 50 μs for T_M, and 1 kHz for f_c. Here $R_n(\tau)$ and $R_g(\tau)$ are plotted as the solid and dotted curve, respectively, in panel (a) (as in Fig. 3), and their product $R_x(\tau)$ is shown in panel (b). It can be seen by comparing these curves with those in Fig. 3(a) and 3(b) that the value of T_P has little effect on $R_x(\tau)$.

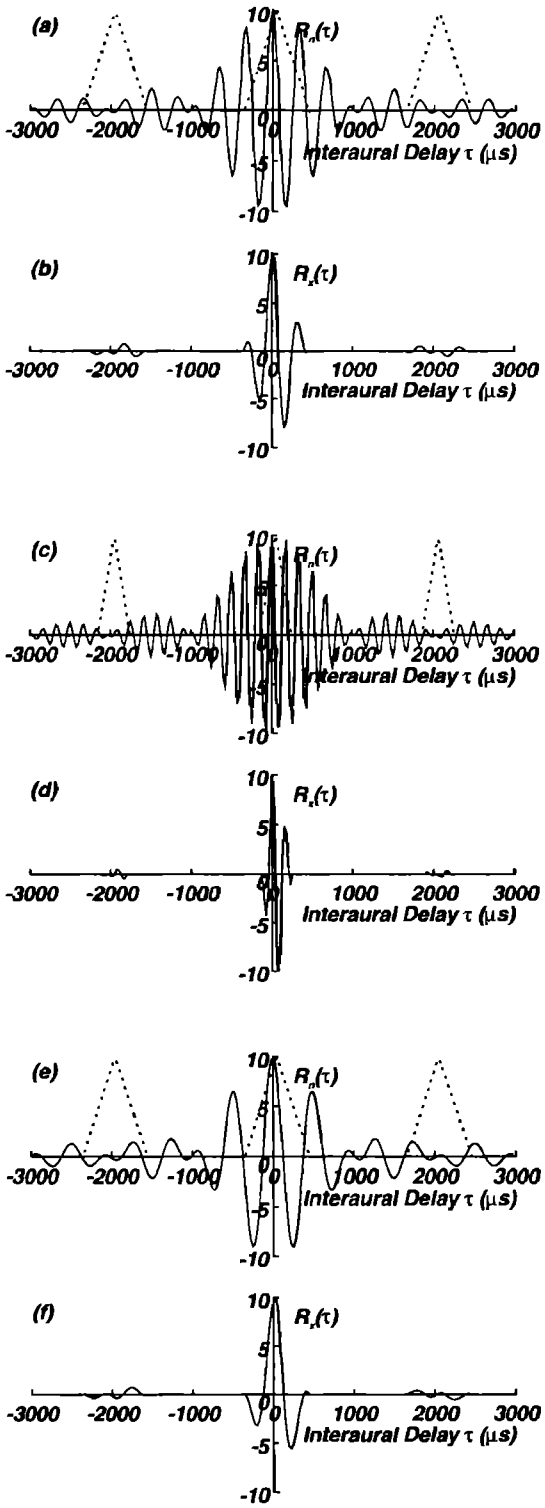


FIG. 5. Development of cross-correlation functions for gated-noise stimuli with a value of 2 ms for T_P, 50 μs for T_M, and several combinations of f_c and T_D: 3 kHz and 400 μs [panels (a) and (b)], 6 kHz and 200 μs [panels (c) and (d)], and 2 kHz and 400 μs [panels (e) and (f)]. As in Figs. 3 and 4, $R_n(\tau)$ and $R_g(\tau)$ are plotted as the solid and dotted curve, respectively, of panels (a), (c), and (e), and their product $R_x(\tau)$ is plotted in panels (b), (d), and (f) for each combination of f_c and T_D. Comparing the curves, it can be seen that the noise bandwidth W has little effect on the shape of $R_x(\tau)$, and combinations of T_D and f_c that maintain a constant value of the product $f_c T_D$ will produce families of cross-correlation functions $R_x(\tau)$ that are expanded or contracted along the τ axis but that maintain approximately the same shape.

pulse duration T_D is only through the product of these two quantities.

Figure 5 shows the cross-correlation function of the stimuli using a value of 2 ms for T_P, 50 μs for T_M, and several combinations of the modulating pulse duration T_D and the center frequency f_c. Again, we plot separately the functions $R_n(\tau)$, $R_g(\tau)$, and their product $R_x(\tau)$. We compare in these figures the correlation functions that result from three combinations of f_c and T_D: 3000 Hz and 400 μs [Fig. 5(a) and (b)], 6000 Hz and 200 μs [Fig. 5(c) and (d)], and 2000 Hz and 400 μs [Fig. 5(e) and (f)]. The first two pairs of conditions have the same $f_c T_D$ product, but the third pair does not. Comparing the resulting cross-correlation functions, it can be seen that although the function in Fig. 5(d) is contracted in time relative to that in Fig. 5(b), it contains the same ensemble of positive and negative peaks for values of $|\tau|$ that are close to zero. The cross-correlation function in Fig. 5(f) is different in shape, lacking, for example, the positive mode to the right of the major central mode. Hafter and Shelton (1991) argue convincingly that the presence or absence of reversals in lateralization can be related to the overall shape of the cross-correlation function of the stimuli, and specifically the presence of negative modes and asymmetries in this function.

The plots in Fig. 5 show that the shape of $R_x(\tau)$ for small values of $|\tau|$ is dominated by two factors: the duration of the modulating pulses T_D [which controls the width of the triangular pulses of $R_g(\tau)$] and the center frequency f_c of the bandpass noise [which controls the locations of the modes of $R_n(\tau)$]. The bandwidth W has relatively little effect on the shape of $R_x(\tau)$ because the $\sin(\pi\tau W)/\pi\tau W$ component of $R_n(\tau)$ is much broader than the modes of $R_g(\tau)$ (as long as $1/W$ is much greater than T_D). Consequently, manipulations of the values of T_D and f_c that maintain a constant value of the product $f_c T_D$ will produce families of cross-correlation functions $R_x(\tau)$ that are expanded or contracted along the τ axis but maintain approximately the same shape. (The cross-correlation functions of the stimuli would have maintained exactly the same shape if the bandwidth W had also been varied in proportion with f_c and the modulator delay T_M had also been varied in proportion with T_D.) This is confirmed in our examples by noting that the cross-correlation functions in Fig. 5(b) and (d) have approximately the same shape, but that the function shown in Fig. 5(f) has a different shape, as noted in the previous paragraph. Hence, the major dependence of lateralization on f_c and T_D is through their product, as demonstrated by the results shown in Fig. 4 of Hafter and Shelton.

C. Lateralization predictions based on the cross-correlation function of the entire stimuli

Hafter and Shelton (1991) demonstrate that most aspects of their data can be described by the simple equation

$$\hat{P} = \int_{-\infty}^{\infty} \frac{\tau R_x(\tau)}{|R_x(\tau)|}\, d\tau. \tag{10}$$

This equation was adopted because it provided the best description of the data, as other plausible operations on $R_x(\tau)$ including the centroid along the τ axis did not work as well.

While Eq. (10) provides a good description of most of the observed data, it does have two major shortcomings, which are acknowledged by Hafter and Shelton. First, the model is based on the cross correlation of the entire stimulus. The overwhelming body of accumulated physiological and psychophysical evidence indicates that sounds are processed by parallel channels of the peripheral auditory system, with the signals undergoing bandpass filtering and nonlinear rectification in each channel. We believe that it is especially appropriate to discuss the data in terms of models that specifically include frequency analysis, because Hafter and Shelton demonstrate that the phenomena themselves appear to be mediated by a low-frequency mechanism. Second, the arbitrary nature of the processing implied by this model is unsatisfying. In fact, Eq. (10) computes a quantity that is related in a general way to the center of mass of $R_x(\tau)$, while avoiding the difficulties that develop when one attempts to compute the actual centroid of $R_x(\tau)$ when the integral of $R_x(\tau)$ over all τ is zero. In Sec. IV we argue that the position-variable model performs a rather similar computation to that of the model of Hafter and Shelton, but without the need for an arbitrarily imposed rectification operation, since natural rectification of the peripheral auditory system ensures that the cross-correlation function will be positive. Despite the shortcomings of the Hafter–Shelton model, its success in describing the gated-noise data motivated us to develop the corresponding predictions for the centroid-based position-variable model with more plausible frequency-specific peripheral processing.

III. LATERALIZATION BASED ON INFORMATION IN THE CROSS-SPECTRAL DENSITY FUNCTION

We now consider the extent to which the lateralization of the gated-noise stimuli can be accounted for by the form of the interaural group delay and interaural phase delay of the stimuli. Both of these attributes have been implicated in the lateralization of complex binaural stimuli: the interaural phase delay is the dominant lateralization cue at low frequencies, and the lateralization of high-frequency stimuli with low-frequency envelopes appears to depend on the interaural group delay of these stimuli. (See Henning, 1980, for a particularly elegant presentation of this hypothesis.)

Interaural differences in group delay and phase delay can be derived from the cross-spectral density function $S_x(f)$, which is the Fourier transform of $R_x(\tau)$. We show in the Appendix that the cross-spectral density function for the stimuli used in the experiment of Hafter and Shelton can be expressed in terms of the summation

$$S_x(f) = \frac{N_0}{2} \sum_k \frac{\sin^2(k\pi T_D/T_P)}{(k\pi)^2}\, e^{-j(2\pi k T_M/T_P)}, \tag{11a}$$

where the summation is evaluated over those values of k for which

$$(f+f_c)T_P - T_P W/2 < k \leqslant (f+f_c)T_P + T_P W/2 \tag{11b}$$

or

$$(f-f_c)T_P - T_P W/2 < k \leqslant (f-f_c)T_P + T_P W/2. \tag{11c}$$

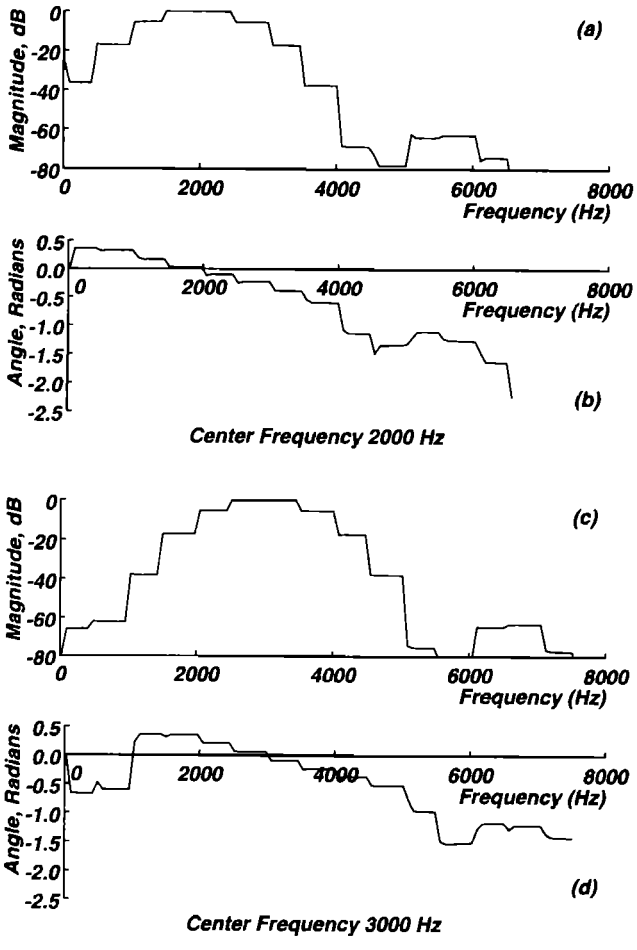


FIG. 6. Magnitude and phase of the cross-spectral density function $S_x(f)$ for gated-noise stimuli with T_D equal to 400 μs, T_P equal to 2 ms, T_M equal to 50 μs, and two values of f_c: 2000 Hz [panels (a) and (b)] and 3000 Hz [panels (c) and (d)]. [The cross-correlation functions for these stimuli are shown in Fig. 5(b) and (f).] The curves are plotted only for those frequencies for which $|S_x(f)|$ is within 80 dB of its maximum value.

Note that the frequency dependence in the above expression for $S_x(f)$ enters through the values of the index k.

We show plots of the relative magnitude and the phase of the cross-spectral density functions $S_x(f)$ in Fig. 6 for T_D equals 400 μs, T_P equals 2 ms, T_M equals 50 μs, and two values of f_c: 2000 Hz [Fig. 6(a) and (b)] and 3000 Hz [Fig. 6(c) and (d). [The cross-correlation functions for these stimulus conditions are shown in Fig. 5(b) and (f).] With a value of T_D equal to 400 μs, a stimulus with f_c equal to 2000 Hz is lateralized toward the "incorrect" side of the head (i.e., the side with the ear receiving the signal with the gating function that is lagging in time), while a stimulus with f_c equal to 3000 Hz is heard toward the opposite or "correct" side of the head. The curves shown in Fig. 6 are typical of those that are observed for stimulus parameters producing expected and anomalous lateralizations.

It can be seen from Fig. 6 that the actual magnitude and phase functions of the cross-spectral density function for these unusual stimuli consist of a series of "steps" that are approximately constant over successive ranges of 500 Hz, as is indicated by Eq. (11). At low frequencies, the interaural phase function for the 3000-Hz stimulus that is lateralized on the "correct" side of the head is negative, while the interaural phase function for the 2000-Hz stimulus that produces "anomalous" lateralizations exhibits a positive phase shift for low frequencies (i.e., a lead in phase). The "steps" of each of these two phase functions tend to become more negative as frequency increases.

We believe that the experimental data and the form of the interaural phase spectra argue strongly against the hypothesis that the counterintuitive reversals in the lateralization data of Hafter and Shelton are mediated by a lateralization mechanism that makes direct use of interaural group delay of the stimuli themselves. Since the interaural phase spectra are flat for most frequencies, the group delay of the stimuli, which is the derivative with respect to frequency of the interaural phase spectrum, would also be zero for most frequencies, for all stimulus configurations. It would probably be more reasonable to consider the overall trend of the group delay as a function of frequency, but again it can be seen that (except for one step at 1000 Hz for the 3000-Hz stimulus) the slope of the interaural phase difference is generally negative for both values of f_c. Hence, any model that is based directly on the interaural differences in group delay of the stimuli used in the experiments of Hafter and Shelton (1991) would be expected to predict that the stimuli are lateralized toward the same side of the head for all values of f_c, contrary to the reversals in lateralization that were observed in the actual data.

Predictions based on the interaural phase delay at low frequencies, on the other hand, are likely to be more successful, as it can be seen that the low-frequency phase delay is positive for f_c equals 2000 Hz and negative for f_c equals 3000 Hz. Similarly, if the interaural group delay were smoothed over frequency it would become a better predictor of the reversals in lateralization. Nevertheless, no model has yet been developed that provides quantitative lateralization predictions by directly using either interaural group delay or interaural phase delay.

IV. LATERALIZATION PREDICTIONS FROM THE POSITION-VARIABLE MODEL

The position-variable model (Colburn, 1973; Stern and Colburn, 1978) describes the subjective lateral position of simple stimuli in terms of putative discharge patterns of fibers of the auditory nerve. These firing times are modeled as sample functions of nonhomogeneous Poisson processes with rate functions that depend on the stimulus. Information pertaining to the interaural time difference of a stimulus is characterized by the outputs of a display of binaural information containing a network of units that, effectively, estimates the interaural cross-correlation function of binaural stimuli after peripheral frequency analysis. Specifically, each unit is assumed to record coincidences in firing times from auditory-nerve units of comparable characteristic frequency from the two ears, after a small fixed internal interaural time delay. This display may be regarded as an implementation of the running cross-correlation operation proposed by Sayers and Cherry (1957), using a neural mechanism inspired by Jeffress (1948).

Predictions of the subjective lateral position of a stimulus are obtained by computing the center of mass along the internal-delay axis of the response of the coincidence-counting units, as described in detail in Stern and Colburn (1978). Predictions for "objective" discrimination and detection experiments can also be obtained by assuming that judgments are based on the subjective position of the stimuli, and that performance is limited by the intrinsic variability of the auditory-nerve response to the sounds (Stern, 1976; Stern and Colburn, 1985).

While the position-variable model was originally developed to describe the laterality of stimuli at 500 Hz, we have recently extended the model to describe the perception of higher-frequency stimuli as well (Shear, 1987; Stern *et al.*, 1988). The revisions to the model that enable us to describe data at higher frequencies will be presented and discussed in detail in a future paper (Stern and Shear, 1992). In brief, they consist of the following.

(i) The exponential rectifier was replaced by a half-wave cube-law rectifier. This was done to obtain a more accurate representation of the response of the model to noise stimuli.

(ii) The single-pole low-pass filter used by Colburn (1973) was replaced by a low-pass filter with a transfer function that has a steeper high-frequency slope, suggested by the physiologial measurements of Johnson (1980). This low-pass filter limits the model's ability to develop a synchronous response to the fine structure of high-frequency stimuli.

(iii) The nonlinear rectifier and the low-pass filter were interchanged, so that the auditory-nerve model consists of (in part) the cascade of a bandpass filter followed by a nonlinear rectifier and a low-pass filter. This reordering of the elements of the auditory-nerve model enables the model to synchronize to the low-frequency envelopes of high-frequency stimuli.

(iv) The density function that describes the relative number of coincidence-counting units as a function of their internal interaural delay [which is referred to as $p(\tau)$ in the papers of Colburn (1977) and Stern and Colburn (1978)] was replaced by a new frequency-dependent density function $p(\tau|f_c)$, for reasons discussed at length in Stern and Shear (1992).

The output of the extended position-variable model is a random variable representing the position estimate, which we refer to as $\hat{P}$. Its mean is approximated by

$$E[\hat{P}] \approx \frac{\int_{\mathbf{R}_{lat}} p(f_c) \int_{-\infty}^{\infty} \tau E\left[L_m(\tau_m, f_{c_m})\right] p(\tau|f_c)\, d\tau\, df_c}{\int_{\mathbf{R}_{lat}} p(f_c) \int_{-\infty}^{\infty} E\left[L_m(\tau_m, f_{c_m})\right] p(\tau|f_c)\, d\tau\, df_c}, \tag{12}$$

where $E\left[L_m(\tau_m, f_{c_m})\right]$ refers to the expected number of coincidences of a fiber pair with internal delay τ_m and characteristic frequency f_{c_m}, and $p(f_c)$ describes the distribution of fiber pairs with respect to characteristic frequency. The term $\mathbf{R}_{lat}$ refers to the range of characteristic frequencies over which predictions are computed. In this paper that range is the set of frequencies over which auditory-nerve fibers from both ears are all firing at a rate above their spontaneous rate, as discussed in Shear (1987).

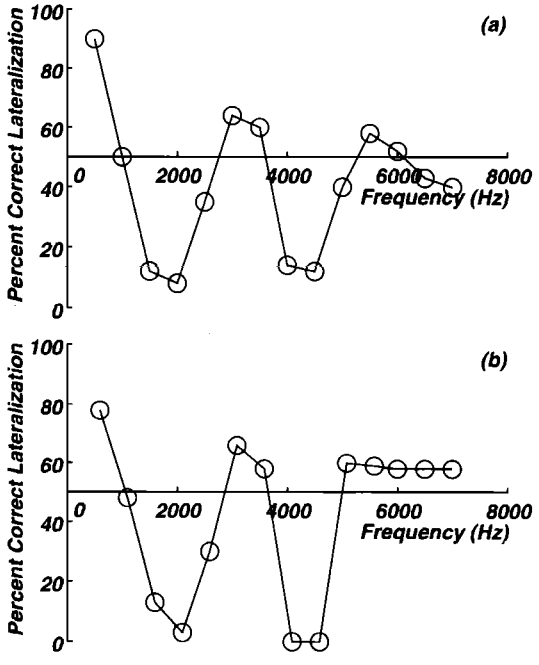


FIG. 7. (a) Typical lateralization results by Hafter and Shelton (1991) for gated-noise stimuli as a function of center frequency f_c, with T_D equal to 400 μs, T_P equal to 2 ms, W equal to 1000 Hz, T_M equal to 37.5 μs, and a sound pressure level of 70 dB SPL. Lateralization percentages below 50% are the counterintuitive "illusory" reversals. (b) Predictions of the extended position-variable model for the same stimulus conditions, except that T_M equals 50 μs for computational reasons.

Figure 7 compares predictions and data using the position-variable model for a typical set of experimental results from Hafter and Shelton (1991). Predictions are shown for T_D equals 400 μs, T_P equals 2 ms, W equals 1000 Hz, and a stimulus level of 70 dB SPL. The center frequency of the bandpass noise f_c is varied as a parameter. The ITD of the gating function T_M was set to 50 μs, which was the closest possible approximation to the 37.5-μs delay used in the experiments, given our sampling rates. The center frequencies that produce the counterintuitive reversals in lateralization are indicated by lateralization percentages of less than 50% in the data and predictions. Since predictions are obtained in terms of laterality estimates (which are both positive and negative) while data are described in terms of percentage "correct" responses (which vary from 0%–100%), the predictions were renormalized by a linear transformation that maps a predicted position of zero (i.e., the center of the head) to correspond to 50% correct discrimination performance, and vertically normalizes the data to best describe the observed data (while restricting the predictions to lie between 0% and 100% correct). The predictions describe the data quite well for center frequencies up to about 5000 Hz. Although they are not provided here, predictions for other combinations of stimulus parameters describe the data equally well, primarily for the reasons discussed in Sec. II.

Figure 8 shows how predictions of the position-variable model depend on the frequency components of the stimuli. The solid curve with the circular symbols describes predic-

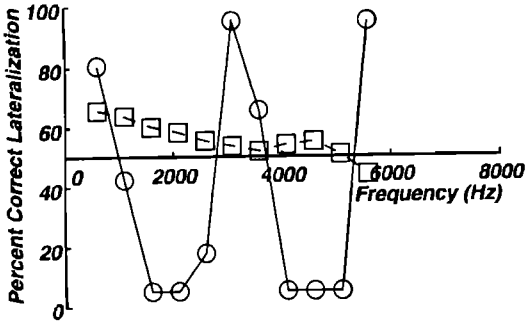
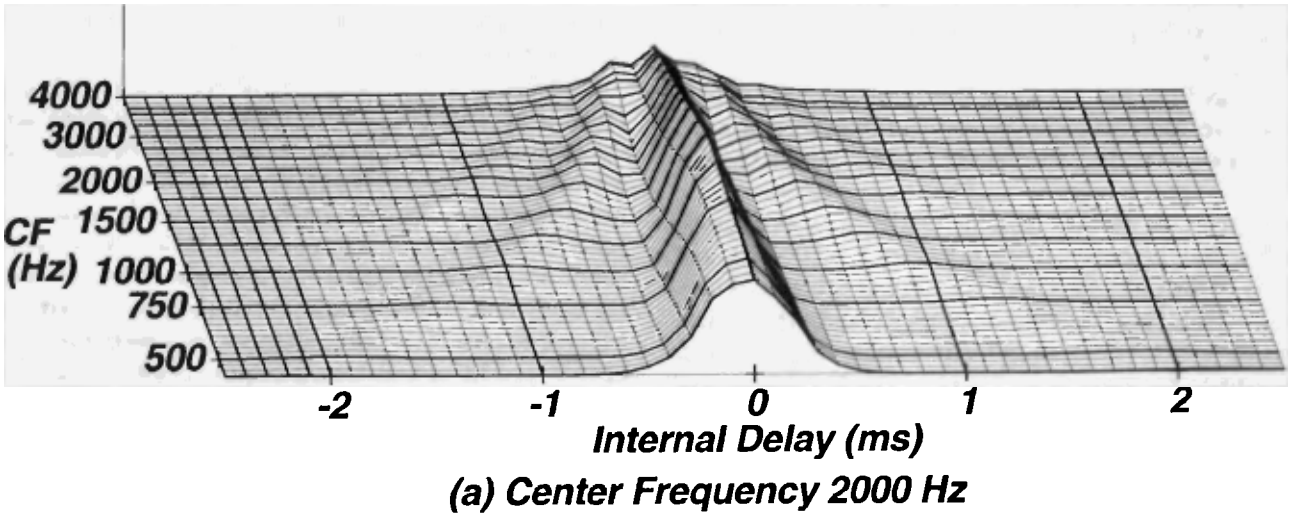


FIG. 8. Predictions of the position-variable model using only components of the cross-correlation function above 1500 Hz (squares) and below 1500 Hz (circles). These predictions closely approximate the predictions that would have been obtained if the stimuli were passed through a 1500-Hz high- and low-pass filter, respectively. Stimulus parameters are as in Fig. 7.

tions obtained when the model ignores frequency components above 1500 Hz. This curve approximates the response that would be obtained when the stimuli are low-pass filtered with a cutoff of 1500 Hz. The dashed curve with the square symbols shows predictions obtained when the model ignores frequency components below 1500 Hz (approximating the response to 1500-Hz high-pass-filtered stimuli). These curves show that, except when $H_{BP}(f)$ has a very low center frequency, predictions of the model depend almost entirely on stimulus components of the stimuli that are below 1500 Hz, which corresponds precisely to what Hafter and Shelton (1991) found in their empirical observations.

We believe that the easiest way to understand the likely mechanism for the counterintuitive reversals, and specifically the predictions of Figs. 7 and 8, is by considering how the



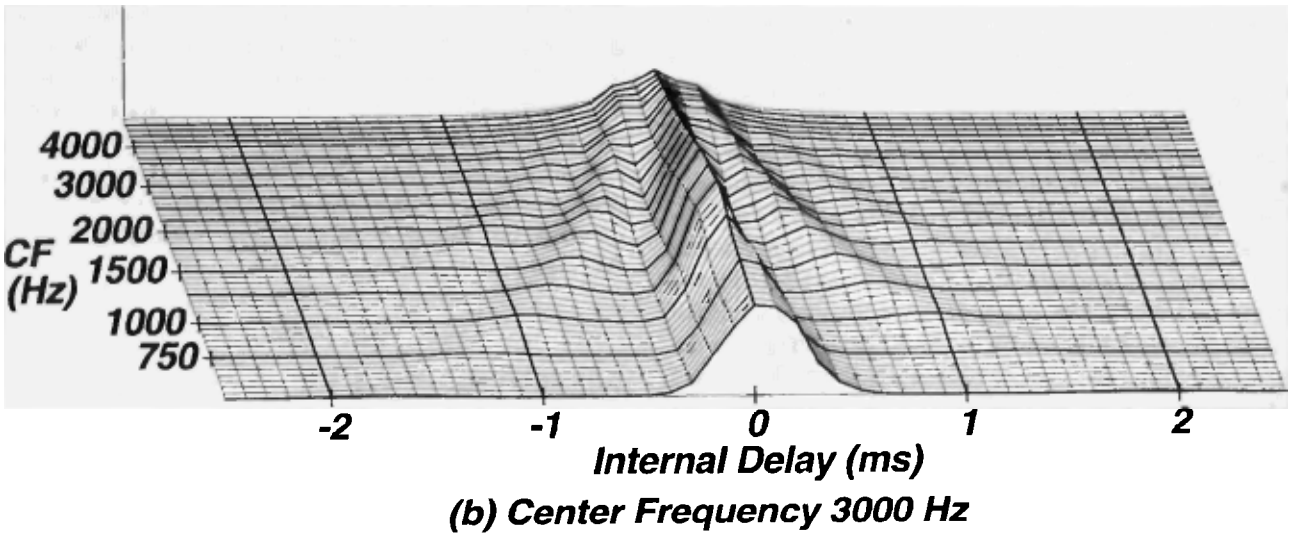


FIG. 9. Cross-correlation patterns showing the response of an ensemble of binaural fiber pairs to a gated-noise stimulus with T_D equal to 400 μs and f_c equal to 2000 Hz [panel (a)] and 3000 Hz [panel (b)].

gated-noise stimuli are likely to be represented by the binaural system. Figure 9 shows the product $L(\tau, f_c)p(\tau|f_c)$, which represents the relative response of an ensemble of binaural fiber pairs to the gated-noise stimuli with T_D equal to 400 μs, T_M equal to 50 μs, T_P equal to 2 ms, W equal to 1000 Hz, a stimulus level of 70 dB SPL, and two values of f_c, 2000 and 3000 Hz. These combinations of stimulus parameters were chosen because they illustrate center frequencies that produce normal and counterintuitive subjective lateral positions, respectively, and because they were included in other examples in Figs. 5–8. The horizontal axis indicates the "characteristic" internal delay parameter of binaural fiber pairs, as described in Colburn (1973, 1977) and Stern and Colburn (1978), and the oblique axis indicates the characteristic frequency of the auditory-nerve fibers in Hz. The shape of the two-dimensional function specifies the position-variable model's representation of the internal response to the gated-noise stimuli, including the implicit weighting by the function $p(\tau|f_c)$, which describes the relative number of fiber pairs with a given characteristic frequency and characteristic delay.

Figure 9(a) and (b) show that the gated-noise stimuli produce an "internal cross-correlation function" that has a broad mode near the center of the internal-delay axis. The differences in predicted position for these two values of f_c are clearly reflected in the locations of the dominant mode, which is slightly to the left of $\tau = 0$ when f_c is 2000 Hz, and slightly to the right when f_c is 3000 Hz. It can also be seen that $L(\tau, f_c)p(\tau|f_c)$ is almost identical (and symmetrical about the internal delay axis) for fiber pairs with characteristic frequencies above about 1500 Hz. This is a reflection of the fact that fibers of the auditory nerve are not able to respond in sychrony to stimulus fine structure to components of the stimulus above that frequency. Hence, the theoretical predictions (and the data) depend primarily on stimulus components below 1500 Hz.

The major difference between the product $L(\tau, f_c)p(\tau|f_c)$ and the function $R_x(\tau)$ [from which predictions are obtained by Hafter and Shelton (1991)] is that $R_x(\tau)$ is the cross correlation of the stimuli themselves while $L(\tau, f_c)p(\tau|f_c)$ represents the cross correlation of the putative response of the peripheral auditory system to the same stimuli. The centroid of $L(\tau, f_c)p(\tau|f_c)$ differs from the centroid of $R_x(\tau)$ because of the nonlinear rectification that is part of the model of auditory-nerve activity and (to a lesser extent) because of the logarithmic distribution of the characteristic frequencies of the fibers. When Hafter and Shelton (1991) found that the centroid of $R_x(\tau)$ did not describe the data [in part because $R_x(\tau)$ is negative for some values of τ], they introduced the absolute value operation in the denominator of Eq. (10) to overcome this problem. Because of the rectification of the peripheral auditory system, $L(\tau, f_c)$ is always positive, and the mean of the position estimate is obtained directly from the centroid computation, without further approximation. The location (or at least the sidedness) of the major mode of $L(\tau, f_c)p(\tau|f_c)$ does appear to correspond to the values of the phase of the cross-spectral density function $S_x(f)$ at frequencies, as can be seen by comparing Figs. 6 and 9.

V. SUMMARY

In this paper we examined the mechanisms underlying the anomalous reversals in lateralization of the binaural gated bandpass-noise stimuli described by Hafter and Shelton (1991).

We obtained an analytical characterization of the cross-correlation functions of these stimuli, and found that these functions could be expressed as the product of the cross-correlation function of the bandpass noise and the cross-correlation function of the gating functions. Like the experimental lateralization data, the shapes of these cross-correlation functions are largely independent of the period of the gating function. The shapes of the cross-correlation functions are also independent of the center frequency of the bandpass filter and the duration of the gating pulses, provided that the product of these two quantities is held fixed. Hence, *any* model that bases lateralization predictions on the shapes of these cross-correlation functions of the stimuli should produce lateralization predictions that describe these trends of the data.

As Hafter and Shelton report, accurate lateralization predictions cannot be directly obtained from the centroid of the cross-correlation function of the stimuli because of the destabilizing effects of the negative modes of the cross-correlation functions. The position-variable model, on the other hand, bases its lateralization predictions on the centroid of the cross-correlation function of the stimuli after undergoing the type of bandpass filtering and nonlinear rectification believed to take place in the auditory periphery. These predictions provide an excellent description of the lateralization data, including the observed reversals and the dominant role played by low-frequency stimulus components. We also obtained analytical expressions for the cross-spectral-density functions of the stimuli. On the basis of the forms of these functions, we argue that models based on the interaural group delay of the stimuli *per se* are not likely to be able to describe the anomalous reversals in the data, although phase delay may be a more useful cue. Considering all of the possible mechanisms that mediate the counterintuitive reversals described by Hafter and Shelton, we believe that the type of binaural processing assumed by the position-variable model, and specifically consideration of the auditory system's "internal cross-correlation function" rather than the cross-correlation function of the entire stimuli to the two ears, not only provides more accurate quantitative lateralization predictions, but also does so in a more natural and intuitive fashion.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant No. BNS-87099349). We thank E. R. Hafter and B. R. Shelton for sharing their interesting experimental data with us some years ago, and we are grateful to L. Bernstein, H. S. Colburn, E. R. Hafter, C. Trahiotis, and W. S. Woods for their comments on earlier drafts of this manuscript.

APPENDIX: DERIVATION OF THE CROSS-SPECTRAL DENSITY FUNCTION

In this appendix we develop an analytical expression for the cross-spectral density function $S_x(f)$, which is the Fourier transform of the cross-correlation function $R_x(\tau)$.

Since $R_x(\tau)$ is the product of the autocorrelation function of the bandpass noise and the cross-correlation function of the gating functions

$$R_x(\tau) = R_n(\tau)R_g(\tau), \tag{A1}$$

its Fourier transform is equal to the convolution

$$S_x(f) = S_n(f) * S_g(f), \tag{A2}$$

where the power spectral density function $S_n(f)$ and the cross-spectral density function $S_g(f)$ are the Fourier transforms of $R_n(\tau)$ and $R_g(\tau)$, respectively.

The power spectral density function of the bandpass noise $n(t)$ is equal to

$$S_n(f) = \begin{cases} N_0/2, & f_c - W/2 \leqslant |f| < f_c + W/2, \\ 0, & \text{otherwise}, \end{cases} \tag{A3}$$

as defined in Sec. I.

It is convenient to represent $R_g(\tau)$ as the convolution of three functions:

$$R_g(\tau) = R_{g_0}(\tau) * \delta(\tau - T_M) * \sum_{k=-\infty}^{\infty} \delta(\tau - kT_P), \tag{A4}$$

where

$$R_{g_0}(\tau) = (T_D/T_P)(1 - |\tau|/T_D), \quad \text{for } |\tau| \leqslant T_D,$$
$$= 0, \quad \text{otherwise.}$$

The Fourier transform of $R_g(\tau)$, then, is

$$S_g(f) = \frac{1}{T_P} \frac{\sin^2(\pi f T_D)}{(\pi f)^2} e^{-j(2\pi f T_M)} \frac{1}{T_P} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T_P}\right)$$
$$= \sum_{k=-\infty}^{\infty} \frac{\sin^2(k\pi T_D/T_P)}{(k\pi)^2} e^{-j(2\pi k T_M/T_P)} \delta\left(f - \frac{k}{T_P}\right). \tag{A5}$$

Convolving with $S_n(f)$ produces

$$S_x(f) = \sum_{k=-\infty}^{\infty} \frac{\sin^2(k\pi T_D/T_P)}{(k\pi)^2}$$
$$\times e^{-j(2\pi k T_M/T_P)} S_n\left(f - \frac{k}{T_P}\right). \tag{A6}$$

Because of the ideal bandpass nature of $S_n(f)$, $S_n(f - k/T_P)$ will be nonzero for only a small number of values of k. As a result, the expression for $S_x(f)$ can be rewritten as

$$S_x(f) = \frac{N_0}{2} \sum_k \frac{\sin^2(k\pi T_D/T_P)}{(k\pi)^2} e^{-j(2\pi k T_M/T_P)}, \tag{A7a}$$

where the summation is evaluated over those values of k for which

$$T_P(f + f_c) - T_P W/2 < k \leqslant T_P(f + f_c) + T_P W/2 \tag{A7b}$$

or

$$T_P(f - f_c) - T_P W/2 < k \leqslant T_P(f - f_c) + T_P W/2. \tag{A7c}$$

Bernstein, L. R., and Trahiotis, C. (1985). "Lateralization of Low-Frequency Complex Waveforms: The Use of Envelope-Based Temporal Disparities," J. Acoust. Soc. Am. 77, 1868–1880.

Colburn, H. S. (1973). "Theory of Binaural Interaction Based on Auditory-Nerve Data. I. General Strategy and Preliminary Results on Interaural Discrimination," J. Acoust. Soc. Am. 54, 1458–1470.

Colburn, H. S. (1977). "Theory of Binaural Interaction Based on Auditory-Nerve Data. II. Detection of Tones in Noise," J. Acoust. Soc. Am. 61, 525–533.

Hafter, E. R., and Ricard, G. L. (1973). "Binaural Interaction with Stimuli that Produce Periodicity Pitch," J. Acoust. Soc. Am. 53, S34 (A).

Hafter, E. R., and De Maio, J. (1975). "Difference Thresholds for Interaural Delay," J. Acoust. Soc. Am. 57, 181–187.

Hafter, E. R., Shelton, B. R., and Green, D. M. (1980). "A Reversal in Lateralization, with Images Appearing on the Side of the Delay," J. Acoust. Soc. Am. Suppl. 1 68, S16 (A).

Hafter, E. R., and Shelton, B. R. (1991). "Counterintuitive Reversals in Lateralization Using Rectangularly Modulated Noise," J. Acoust. Soc. Am. 90, 1901–1907.

Henning, G. B. (1980). "Some Observations on the Lateralization of Complex Waveforms," J. Acoust. Soc. Am. 68, 446–453.

Henning, G. B. (1983). "Lateralization of Low-Frequency Transients," Hearing Res. 9, 153–172.

Jeffress, L. A. (1948). "A Place Theory of Sound Localization," J. Comput. Physiol. Psychol. 41, 35–39.

Johnson, D. H. (1980). "The Relationship Between Spike Rate and Synchrony in Responses of Auditory-Nerve Fibers to Single Tones," J. Acoust. Soc. Am. 68, 1115–1122.

Kuwada, S., and Yin, T. C. T. (1983). "Binaural Interaction in Low-Frequency Neurons in Inferior Colliculus of the Cat. I. Effects of Long Interaural Delays, Intensity, and Epetition Rate on Interaural Delay Function," J. Neurophysiol. 50, 981–999.

Kuwada, S., Stanford, T. R., and Batra, R. (1987). "Interaural Phase-Sensitive Units in the Inferior Colliculus of the Unanaesthetized Rabbit: Effects of Changing Frequency," J. Neurophysiol. 57, 1338–1360.

Sayers, B. McA., and Cherry, E. C. (1957). "Mechanism of Binaural Fusion in the Hearing of Speech," J. Acoust. Soc. Am. 61, 973–987.

Shear, G. D. (1987). "Modeling the Dependence of Auditory Lateralization on Frequency and Bandwidth," Master's thesis, Electrical and Computer Engineering Department, Carnegie Mellon Univ.

Stern, R. M., Jr. (1976). "Lateralization, Discrimination, and Detection of Binaural Pure Tones," Doctoral dissertation, MIT.

Stern, R. M., Jr., and Colburn, H. S. (1978). "Theory of Binaural Interaction Based on Auditory-Nerve Data. IV. A Model for Subjective Lateral Position," J. Acoust. Soc. Am. 64, 127–140.

Stern, R. M. and Colburn, H. S. (1985). "Lateral-Position-Based Models of Interaural Discrimination," J. Acoust. Soc. Am. 77, 753–755.

Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (1988). "Lateralization of Complex Binaural Stimuli: A Weighted Image Model," J. Acoust. Soc. Am. 84, 156–165.

Stern, R. M., and Shear, G. D. (1992). "Lateralization and Detection of Low-Frequency Binaural Stimuli: Effects of Distribution of Internal Delay," J. Acoust. Soc. Am. (in preparation).