Additive Versus Multiplicative Combination of Differences of Interaural Time and Intensity

Samuel H. Tao

Abstract

We used the position-variable model of binaural interaction to describe concurrently the experimental results of the subjective lateralization of tones and bandpass noises. The model incorporated centroid-based multiplicative intensity weighting and a mechanism to weight trajectories of a stimulus' cross-correlation function (CCF) that were consistent across frequency (i.e. straight). The stimuli of the lateralization experiments we wished to describe were as follows:

- a tone with a range of frequencies, given a small interaural time difference (ITD) and no interaural intensity difference (IID)
- a 500-Hz tone over a range of ITDs and IIDs
- a bandpass noise with a 500-Hz center frequency over a range of bandwidths and IIDs, given one of four combinations of ITD and interaural phase difference (IPD)

The results of the latter two experiments were most important in our current research. We could not adequately describe those two sets of predictions at once by using the multiplicative intensity and the straightness weighting in our model. We obtained better results for the noise predictions using additive intensity weighting such that the predicted position with an intensity difference is a proportional offset of the IID from the predicted position with no IID. However, this type of intensity weighting alone cannot predict the presence of the cue-reversal phenomenon found in the tone/ITD/IID experiment.

Chapter 1
Introduction

The philosophical quotation about people having two ears and one mouth so that they should listen twice as much as they speak also contains a message relevant in the field of binaural hearing (Cherry, 1961).The presence of two ears greatly helps a person perform such tasks as detection, discrimination, and localization and lateralization. Detection refers to identification of the presence or absence of a masked stimulus. Discrimination refers to the process of distinguishing between two similar stimuli. Localization and lateralization refer to judgment of sound direction, the difference between the two being how sounds are presented to a subject. In the case of lateralization, acoustic stimuli are delivered via headphones so that acoustic images usually form along a line connecting the listener's two ears; with localization, acoustic stimuli are presented with loudspeakers so that the acoustic images appear in a free field around the listener and the speakers.

To test the binaural system's ability to judge locations of sounds, we wish to perform psychoacoustic experiments on test subjects. For example, one simple localization task might be to blindfold a person, have a trumpet player play his instrument at various places in a room, and have the person *point* to where he or she hears the sounds coming from. The test subject relies on both the differences in time and in intensity of the signals reaching each ear as primary cues for the procedure of localization. Unless the trumpet player is standing directly in front of the listener, the notes played will arrive at the listener's ears at slightly different times and with slightly different intensity levels, as shown in Figure 1-1. These cues of time and intensity differences are coupled together in this type of experiment. In order to isolate the effects of reflections and absorption from a room or from the head and body, we instead perform a lateralization experiment. That is, we would record a trumpet sound and manipulate it so that when presented over headphones to a subject, the sounds would appear at different locations inside his or her head.
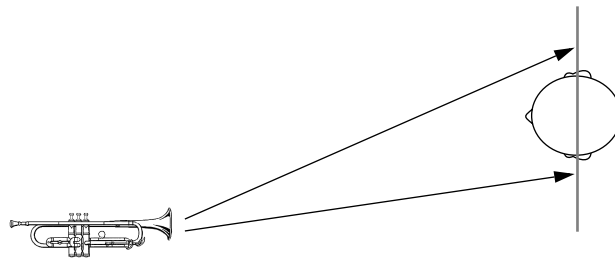

Figure 1-1: Examples of free-field time and intensity differences.

Our ultimate goal of modeling the binaural system is realized by being able to predict and replicate the results from psychoacoustic experiments. The experiments we focus on in this report are those of lateralization. Section 1.1 provides introductions to the process of acquiring these data and to the model we used for predictions. Section 1.2 describes the motivations behind the specific experiments used in this report and summarizes the report's content.

1.1 Background

An acoustic-pointing task is usually used to record the results of lateralization experiments, which is conducted as follows. The test subject is presented with a target stimulus that has a combination of interau-

ral intensity, phase, and time differences (IID, IPD, and ITD)[1], which causes an acoustic image to appear along a path between the ears of a listener. An ITD or an IPD is a respective time or phase difference that exists between the signals at each ear. Unless the stimulus is a pure tone, for a constant ITD the IPD will vary over the bandwidth of the stimulus. An IID is a difference in intensity level between the signals at the ears. The subject points to the intracranial position formed by manipulating a pointer stimulus such that the its location matches the target's location. Typically, the subject may adjust only the IID of the pointer.

The model of binaural interaction we use is the *position-variable model*, which was developed by Stern and Colburn (Colburn, 1973, 1977; Stern and Colburn, 1978) and later modified by Shear (Shear, 1987; Stern and Shear, 1992). It has successfully described the results of a wide variety of psychoacoustic experiments, including those of lateralization, discrimination, and binaural detection. The contents of this report deal exclusively with the results of lateralization experiments, so in this respect, the position-variable model of the binaural system can be viewed as a black box with two inputs and one output. The inputs represent the ears; the output, which lead to higher structures in the brain, is a position variable. The position variable is a signed number that represents a prediction of how a stimulus is subjectively lateralized; the magnitude represents the extent and the sign represents the side to which a stimulus is lateralized. The heart of the black box is a cross-correlation function (CCF) of the putative response of the auditory system to the two input signals. The cross-correlation is a function of frequency and interaural delay or lag, which represents the peripheral filtering of the auditory system and the interaural processing that is hypothesized to exist in the binaural system. We discuss these details further in Chapter 2.

1.2 Report Content

This report follows a chronological order to show what changes were required of the model in order to successfully predict the results of four experiments. All experiments involved the subjective lateralization of either tones or bandpass noise. A list of each experiment and its stimulus is listed below:

- Domnitz and Colburn (1977) – 500-Hz tone with combinations of ITD and IID
- Schiano, Trahiotis, and Bernstein (1986) – pure tones from 200 Hz to 2000 Hz with a small ITD but zero IID
- Buell and Trahiotis (1991) – noise with a 500-Hz center frequency and various combinations of ITD, IPD, and bandwidth
- Buell and Trahiotis (1991) – noise with a 500-Hz center frequency and various combinations of ITD, IPD, IID, and bandwidth

A broad description of the relevant results of these experiments are presented in the following paragraphs. The original position-variable model (Stern and Colburn, 1978) predicted results of stimuli from 500-Hz tones. The experimental results from Domnitz and Colburn showed that the subjective position depended jointly on ITD and IID. If a pure tone is presented with a small ITD and no other interaural differences, a subject would normally lateralize the sound to the ear or side of the head that received the signal that is leading in time. When pure tones are presented with large ITDs, increasing the ITD can under some circumstances cause the lateral position of the images to be perceived closer to the ear receiving the signal that

---

1. We reference all interaural differences with respect to the right ear. The following examples illustrate examples of interaural differences between x_L and x_R, which refer to signals presented to left and right ears, respectively:
IID=10 dB: x_L has an intensity 0.1 times that of x_R.
IPD=90 degrees of a 500 Hz pure tone: x_L is phase-delayed 90˚, or 500 µs, relative to x_R.
ITD=500 µs of a 500 Hz, 200-Hz wide, bandpass noise: x_L lags x_R by 500 µs from 400 to 600 Hz. The phase delay is again 90˚ at 500 Hz, but different at all other frequencies.

is lagging in time, rather than closer to the ear that is receiving the signal leading in time. This phenomenon has been called a "cue reversal". It is described by the position-variable model for reasons discussed in this report and in Stern and Colburn (1978). In 1987 Shear modified the position-variable model to operate on stimuli with a broader range of frequencies. He did this by constraining one function in the position-variable model by using the data from Schiano *et al.* (1986), which showed how the lateral position depended on the frequency of a tonal stimulus.

Stern, Zeiberg, and Trahiotis (1988a) observed lateralization results from a 500-Hz bandpass noise presented with no IID and various combinations of ITD, IPD, and bandwidth. Trahiotis and Stern (1989) expanded the number of test subjects for some of the experimental conditions, and Buell and Trahiotis (1991) performed a complete lateralization experiment to each of the experimental conditions and extended the data to include the effect of IIDs. When referring to the results from Stern *et al.* (1988a), we show the corresponding data from Buell and Trahiotis (1991). The results showed that for certain (ITD, IPD) combinations, lateral position was strongly dependent on bandwidth. The position-variable model was unable to describe these results, and to explain why the model was unsuccessful, Stern *et al.* (1988a) introduced a simplified version of the position-variable model called the weighted-image model. This model emphasized *straight* features of the cross-correlation, which was their term for consistency of interaural-delay information across frequency of the CCF. Straightness weighting allowed the weighted-image model to successfully predict the bandwidth dependence of each stimulus condition. In Stern *et al.* (1991a) the position-variable model incorporated straightness weighting to predict successfully the Buell and Trahiotis (1991) results.

Even while using straightness weighting, the position-variable model was unable to predict the complete results from Buell and Trahiotis (1991), which included the effects of IIDs. In fact, straightness weighting seemed to affect the predictions of each (ITD, IPD) combination differently. The objective of this report was to consider the interaction between the methods of weighting interaural intensity differences and straightness and to determine why the model was fundamentally unable to predict this experiment.

Chapter 2 reviews the position-variable model and its predictions for the tones experiments. Chapter 3 shows how the results of Buell and Trahiotis (1991) that corresponded to the results from Stern *et al.* (1988a) provided the motivation behind straightness weighting. Results of the tones experiments are compared with and without straightness. In Chapter 4 we find that the model cannot describe the data of Buell and Trahiotis (1991), which we remedy by implementing an additive- rather than the current multiplicative-based method of intensity weighting in the model. Chapter 5 shows the predictions of the additive model for the noise and for the tones experiments. We will find that this type of additive intensity weighting is unable to describe the cue-reversals seen in Domnitz and Colburn. Finally, in Chapter 6 we summarize our findings and offer suggestions for additional work.

Chapter 2
Review of Position-Variable Model

In this chapter we summarize the structure and predictions of the position-variable model as modified by (Shear, 1987; Stern and Shear, 1992). Section 2.1 describes the major stages used by the model to produce an estimate of lateralization. In Section 2.2 we review the model's predictions for the lateralization of tones experiments performed by Domnitz and Colburn (1977), and by Schiano *et al.* (1986)

2.1 The Position-Variable Model

We now characterize the position-variable model of binaural interaction as developed by Stern and Colburn (1978) and then modified by Shear (1987). The hypothesized events that take place in the auditory system during the process of lateralization are as follows:

- Filtering and processing of acoustic stimulus entering ears
- Cross-correlation of the filtered signals, forming what we term a binaural display
- Weighting of important features of the binaural display
- Extraction of lateralization information from the display using a centroid measure

Figure 2-1 displays the stages used to calculate the position-variable $\hat{P}$ that serves as the model's estimate of subjective lateral position. The relationship between the stages shown below and how each block is modeled is described in each of the subsections below.
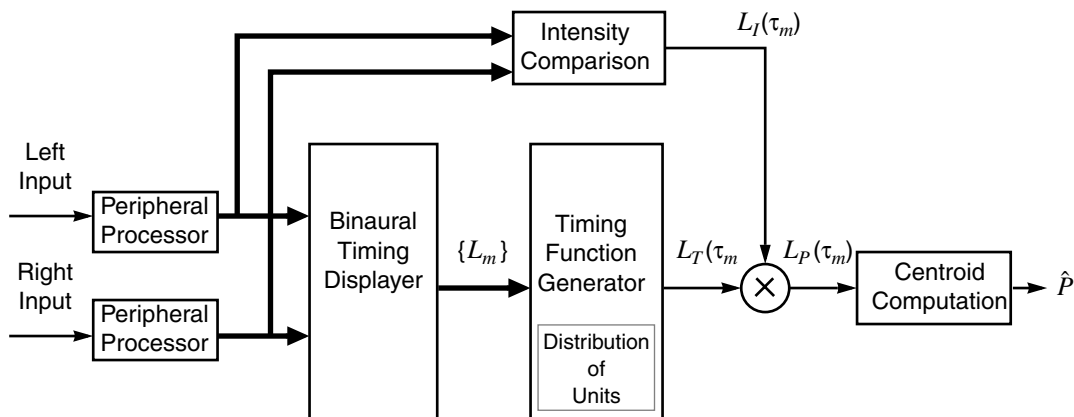

Figure 2-1: Block diagram of lateralization stages in the original position-variable model (from Stern and Colburn, 1978).

2.1.1 Peripheral Processing of the Auditory System

The periphery of the auditory system filters and processes acoustic stimuli. A stimulus reaches the ears via the pinnae and travels through the external auditory canal into the middle ear, where it reaches the tympanic membrane ("eardrum") and causes vibration of the three bony structures attached to the membrane. The chain of middle-ear ossicles amplifies the motions of the stimulus and transmits them to the cochlea. The basilar membrane (BM) partitions the cochlea, and its mechanical properties cause different frequen-

cies to stimulate different locations along its length. Rows of inner and outer hairs along the BM respond to the vibrations caused by the stimulus, and nerve fibers connected the base of each hair cell fire at a rate proportional to the intensity at the cell's location on the BM. Each nerve fiber has a particular firing threshold associated with it. Unstimulated, spontaneously firing fibers have a low average firing rate, and stimulated, active fibers fire at a higher rate. Each auditory-nerve fiber has a characteristic frequency (CF) at which nerve stimulation has the lowest threshold over the ranges of frequencies covered by the cell.

The position-variable model uses a number of stages to model the transformation of the acoustic pressure stimulus into its neural representation. The stages are shown in Figure 2-2, and are taken from Siebert (1968) and Colburn (1973). First, a bank of bandpass filters operating in parallel represents the frequency selectivity of the hair cells. The transfer function of each bandpass filter follows the shape of the threshold of the auditory-nerve fiber plotted against frequency, and each bandpass filter has a maximum at the CF of the nerve fiber it represents. The output of each bandpass filter passes through an automatic gain control
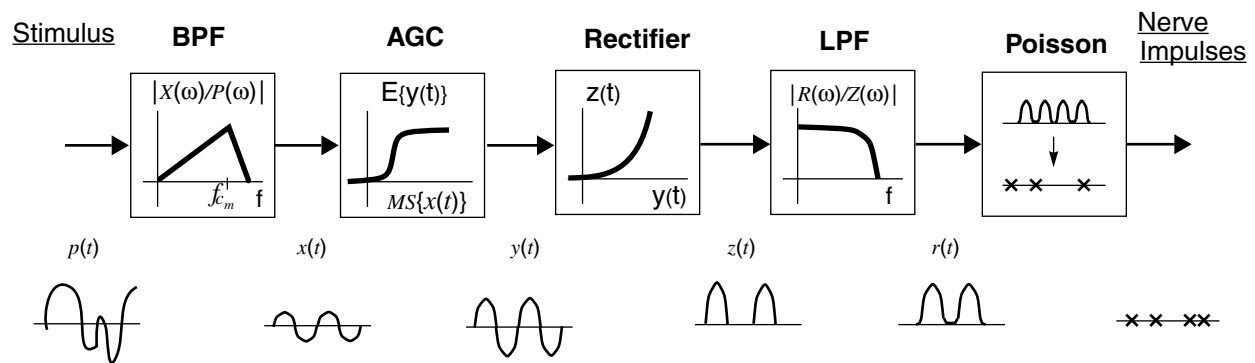

Figure 2-2: Stages of peripheral filtering in auditory-nerve model. The output of nerve impulses is from the mth (out of M) auditory-nerve fiber. Each block is labeled by its transfer characteristic; the x-axis represents either the frequency or the output of a previous stage, and the output of a block is shown by the y-axis.

(AGC), which models the hair cell's neural firing pattern. The AGC parameters adjust such that spontaneously firing and active fibers fire at their respective average rates of 50/sec and 200/sec. A lowpass filter that follows the AGC models the ear's inability to synchronize to high-frequency stimuli.

The half-wave rectification stage models a number of nonlinear effects of the auditory system. The combination of the rectifier and a Poisson-process generator converts the filtered stimulus into a stochastic process that represents the train of neural impulses propagating along the fibers of the auditory nerve. This stochastic point process is modeled as a nonhomogeneous Poisson process such that the amplitude of the filtered signal determines the instantaneous rate of events of the point process. As a result, the auditory-nerve firing times are synchronized to the peaks of the bandpass-filtered stimulus. The filtered and transformed signal is then transmitted to higher-level structures for binaural processing.

2.1.2 Binaural Display

The pressure waves caused by the acoustic stimulus have now been encoded as a train of stochastic neural impulses. In this section we describe a hypothetical mechanism that extracts timing information contained in the stimulus. This model, like similar theories, is an elaboration of a theory first proposed by Jeffress (1948), who postulated an "interaural-difference-detector model" that consisted of a network of

coincidence-counting units and delay lines. His proposed coincidence-counting unit was a nerve cell with two inputs and one output. The inputs were the series of neural pulse trains described above, one from each ear, with a fixed delay inserted in one of the inputs. This cell is assumed to produce an output pulse only when it receives a pulse from both inputs simultaneously (after the delay, in the fashion of a two-input Boolean AND gate). Thus, the firing rate of the unit is controlled by the number of coincidences it observes. The internal delay enables the units to respond maximally to inputs that are delayed relative to each other by an amount equal to the delay.

In 1969, Colburn (1973, 1977a, 1977b) extended Jeffress' work using an auditory-nerve-based model. His formulation includes a binaural timing display, which consists of a network of coincidence-counting units, each with their own internal or characteristic delay τ_m. Each unit's inputs are presumed to come from nerve fibers with the same characteristic frequency f_{c_m}, one from each ear. In other words, each unit, denoted by $L_m(\tau_m, f_{c_m})$, could detect interaural delays between the pulses of the nerve fibers with the same CF.

A row of these units $L_m(\tau_m, f_{c_m})$ with the same CF f_{c_m}, plotted as a function of τ_m, produces a function proportional to the cross-correlation of the signals at that CF (after the peripheral filtering and other signal processing). The presence of rows of multiple coincidence-counting unit rows forms a matrix of coincidence-counting units whose outputs are proportional to the CCF over frequency.

A conceptual representation of the binaural displayer is shown in Figure 2-3. The counter $m \in 1, 2, \dots, M$ indexes each of the auditory nerve fibers. Each fiber has a characteristic frequency f_{c_m}, and associated with the left fibers are individual time delays τ_m. The mth coincidence-counting unit takes as input the pulse trains from the delayed left and non-delayed right nerve fibers, both with the same CF f_{c_m}. A fiber pair in which both fibers are actively firing is called "doubly-active", and a fiber pair in which only one fiber is actively firing is called "singly-active" (Shear, 1987). The range of frequencies that is useful for a given task varies according the stimulus presented. Shear (1987) used stimuli with no IID and performed calculations over what he termed the doubly-active region, the frequency range in which *all* fibers were doubly-active. In this report, we deal with stimuli with IIDs and assume that any coincidence-counting units that take input from pairs of fibers for which neither fiber is firing spontaneously provide useful information for lateralization. In this report we define the *doubly-active region* as the frequency range that includes singly- and doubly-active fibers. Predictions in this report were made using this region, and the differences in predictions using both definitions of the active region were slight. We also chose to include a greater number of inactive fibers to obtain greater compression of predicted position for larger IIDs, which is described in Section 2.2.3 .

The position-variable model implements this form of physiological cross-correlation by performing a mathematical cross-correlation of the signals after peripheral processing. The value of the cross-correlation function gives an indication of the average number of counts of each of the units in the display:

$$E\{L_m\} \equiv E\{L_m(\tau_m, f_{c_m})\} \propto R_{x_L(t)x_R(t)}(\tau_m) \tag{2.1}$$

where $E\{\cdot\}$ denotes the expected value; $x_L(t)$ and $x_R(t)$ are the filtered and rectified input signals to the left and right ears, respectively.

Examples of the model's representation of the binaural display is shown in Figure 2-3. The x-axis shows internal delay of the particular counting unit, the y-axis shows the CF of the counting unit, and the z-axis shows the relative strength or output firing rate of the associated unit. Noteworthy features are listed below:

- The left panel shows the response to a pure tone, and the right panel shows the response to a bandpass noise. Each stimulus was centered at 500 Hz with an ITD of 2000 μs.
- The CCF of the pure tone has straight peaks/modes because each of the bandpass filters extracts the same spectral component of 500 Hz. Conversely, the trajectories of the noise CCF are curved because the output of each filter is dominated by the center frequency of the filter (which is due to the stimulus bandwidth being wider than the filter bandwidth). The largest peak occurs at the value of the actual ITD, which occurs at 2000 μs in this example.
- These CCFs are shown over the doubly-active region. The noise CCF has straight modes starting at the CFs beyond the stimulus bandwidth. One may interpret this as counting units at these CFs picking up the fringe components of the stimulus bandwidth (this may not be readily apparent from the tone CCF). When interpreting the information contained in the CCFs of this report, one should concentrate attention to outputs from CFs within the stimulus bandwidth (i.e. the middle portion of each CCF).
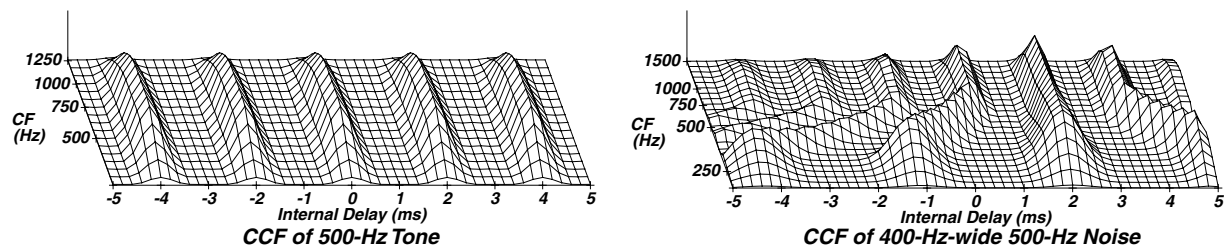

Figure 2-3: The position-variable model's representations of the interaural cross-correlation of the peripherally processed stimuli. The stimulus in the left panel is a 500-Hz tone with a 2000 μs ITD, and the stimulus in the right panel was a band of 500-Hz noise with a 400-Hz bandwidth and a 2000 μs ITD.

The mechanism we have described performs an effective cross-correlation between the peripherally-filtered signals from each ear. It is able to extract timing information about the signals at the ears. In the next section, we discuss some other features of the CCF that are considered important by the binaural system for lateralization, and how they are modeled.

2.1.3 Distribution of Coincidence-Counting Units

When listening to naturally-produced sounds, the geometry of the human head constrains the range of physically-possible ITDs to about ±600 μs (Woodworth, 1938). Hence, we would expect that a larger number of counting units operates on delays in that range. The position-variable model weights the output of the binaural display by a function that is assumed to represent the physical distribution of counting units over delay and frequency (Colburn, 1977). The weighting function is labeled $p(\tau_m, f_{c_m})$, and a typical function is depicted in Figure 2-4. At each CF, $p(\tau_m, f_{c_m})$ describes the distribution of counting units as being constant over a range of small delays and dropping sharply for larger delays. Note that as the CF increases, the function becomes somewhat steeper so that the presumed distribution of units becomes clustered to a smaller range of delays. In Section 3.3.2. we describe some of the reasons for this. We use the term *centrality* to refer to this emphasis on the information contained in the CCF at small time delays.
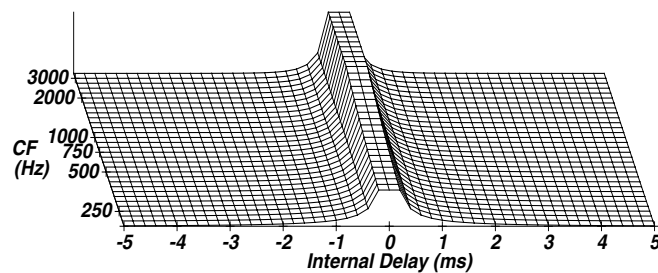
Figure 2-4: Distribution of counting units over internal delay and frequency.

2.1.4 Weighting of Interaural Intensity Differences

Since the cross-correlation function cannot encode amplitude differences between the inputs, the position-variable model applies an additional separate mechanism to model the effect of interaural intensity differences (IIDs). The position-variable model in its original form used a pulse-shaped multiplicative function denoted $L_I(\tau_m)$ to emphasize counting units with appropriate delays. The location of the weighting pulse is proportional to the IID, moving the region of emphasized delays toward one side or the other. The intensity pulse will negate and reinforce features of the CCF appropriate to modeling the effect of the IID. The presence of an intensity difference affects lateralization by manipulating the CCF and hence, the centroid. The type of shape of the function does not greatly affect lateralization (Stern and Colburn, 1977); the model uses a Gaussian function. The Gaussian pulse we use has two descriptive parameters: $L_I(\tau) \sim N(m_I(\alpha), w_I)$. The mean, or center of the pulse, is assumed to be a function of the intensity difference, and the width, or spread (standard deviation) of the pulse is assumed to be constant. Figure 2-5 shows a typical intensity-weighting function representing an IID of 0 dB.

The lateralization effect produced by a particular combination of ITD and IID is not always unique. For example, a tonal stimulus with a small ITD can appear to be centered at the midline by making the tone at the ear receiving the lagging signal more intense. The extent and effect of this trade between time and intensity varies greatly according to the stimulus parameters and conditions. The time-intensity *trading ratio* refers to the amount of the intensity difference necessary to produce the same lateralization as a certain time difference. For instance, a 25 μs/dB trading ratio for a given stimulus means that the stimulus with an ITD of 25 μs and zero IID would be lateralized the same as one with zero ITD and a 1-dB IID.
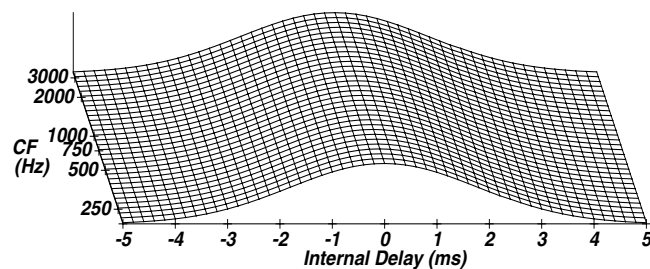

Figure 2-5: Intensity-weighting function. The function shown here follows a Gaussian function with zero mean and a standard deviation of 1.778 ms (the mean of zero would correspond to an IID of 0 dB).

2.1.5 Lateralization Using the Centroid Measure

The position-variable model computes lateralization estimates by calculating the centroid over internal delay of the CCF, weighted by $p(\tau_m, f_{c_m})$ and $L_I(\tau_m)$ (Stern and Colburn, 1978). This variable is, by definition, an average or center-of-mass, which we interpret as a prediction of subjective position. For a graphical view of the steps involved in calculating the centroid, refer to (Stern and Colburn, 1978).

A centroid evaluated to be at 0 µs ($\hat{P} = 0$) would correspond to an image lateralized at midline. With our conventions for the sign of ITD, IPD, and IID, positive centroids would represent images on the right side of the head, with the magnitude of the centroid determining the extent of laterality. A very negative centroid would indicate an image far from the midline and close to the left ear; e.g. a $\hat{P}$ of -20 would describe an image much closer to the left ear than a $\hat{P}$ of -5.

Position predictions can be scaled to correspond to a pointer IID by normalizing the position by the time-intensity trading ratio. This provides a linear conversion between the position variable and the actual perceived position if the relationship between pointer IID and image position is linear, which is not the case for large IIDs. Most of the experimental predictions in this report were scaled accordingly.

2.2 Comparison of Predictions with Experimental Data

We now describe the experiments dealing with the lateralization of tones and their corresponding predictions.

2.2.1 Review of Successful Model Predictions

The position-variable model is able to describe the results of a large set of psychoacoustic phenomena. A partial list of these experiments follows:

- interaural time-amplitude discrimination (Stern, 1976)
- detection thresholds of tones in noise (Stern, 1976; Shear, 1987)
- lateralization of AM tones (Stern *et al.*, 1988b)
- lateralization of bandpass rectangularly-modulated noise (Stern *et al.*, 1991b)
- the "dominant region" phenomenon described by Raatgever (1980, 1986) (Stern and Shear, 1992)

In the next two sections, we present results from two tonal lateralization experiments because they provide the most difficult lateralization constraints when considering those results in conjunction with results from the lateralization of bandpass noise, which are discussed in Chapter 3 and Chapter 4.

2.2.2 Lateralization of Tones Versus Frequency

Schiano *et al.* (1986) recorded the perceived laterality of tones that varied in frequency from 200 to 2000 Hz, presented with a 150-µs ITD and zero IID. The pointer stimulus used was a 500-Hz noise with a bandwidth of 200 Hz. His results are shown in Figure 2-6, which indicates that a pure tone was lateralized to the leading ear up to approximately 1200 Hz. At frequencies above 1200 Hz the subjective image decreased monotonically to the midline.
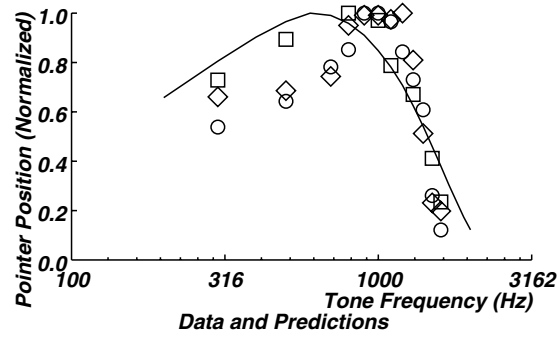
Figure 2-6: Results from Schiano *et al.* The different shapes (square, circle, diamond) represent separate experimental trials, and the position, which has been normalized to a maximum of one, represents the pointer IID required to match the position of the bandpass-noise target as a function of the target tone's frequency. Model predictions are shown by the solid line.

We explain this behavior by examining the shape of $p(\tau_m, f_{c_m})$, the internal distribution of coincidence counters (cf. Shear, 1987). Under the conditions of this experiment, the function $p(\tau_m, f_{c_m})$ attenuates all but the most central peaks of the CCF, and the position of these peaks determine the centroid and, hence, the lateralized position. Recall that the CCF of a tone is periodic with a primary peak at the ITD and peaks spaced multiples of the tonal frequency away from the primary peak. In this case, the peaks appear at $\tau_s \pm k\dfrac{1}{f_0}$=150 μs $\pm k$(2000 μs). For lower-frequency stimuli, the CCF after weighting by $p(\tau_m, f_{c_m})$ has a peak essentially only at the ITD of 1500 μs. This occurs because the lower frequency implies a relatively long period, and the non-central peaks are attenuated by the $p(\tau_m, f_{c_m})$ function. With higher-frequency stimuli, the peaks are spaced closer together and secondary peaks exist after weighting by $p(\tau_m, f_{c_m})$. The additional peaks bring the image closer to the midline.

Also shown in Figure 2-6 are the PV model's predictions for this experiment. The equation Shear used for $p(\tau_m, f_{c_m})$ was

$$p(\tau_m, f_{c_m}) = \left\{ \begin{array}{ll} C_{LF}(f_{c_m}), & |\tau| \le 200 \text{ μs} \\ \\ C_{LF}(f_{c_m})\dfrac{\left[e^{-2\pi f_l(f_{c_m})|\tau|} - e^{-2\pi f_h|\tau|} \right]}{|\tau|}, & \text{otherwise} \end{array} \right. \tag{2.2}$$

where

$$f_1(f_{c_m}) = \left\{ \begin{array}{ll} C_{LF}(f_{c_m}), & |\tau| \le 200 \text{μ}s \\ \\ C_{LF}(f_{c_m})\dfrac{\left[e^{-2\pi f_1(f_{c_m})|\tau|} - e^{-2\pi f_h|\tau|} \right]}{|\tau|} & \text{otherwise} \end{array} \right. \tag{2.3}$$

and f_h=3000 Hz, l_P=-1.1, and $C_{LF}(f_{c_m})$ is chosen so that $p(\tau_m, f_{c_m})$ is a valid density function. The parameter l_P controls the how steeply $p(\tau_m, f_{c_m})$ decays.

The equation specifying the frequency response of the lowpass filter $G(f)$ was

$$\tilde{G}(f) = \begin{cases} 1, & 0 \leq f < f_c \\ \dfrac{1 - f/f_s}{1 - f_c/f_s}, & f_c < f \leq f \\ 0, & f > f_s \end{cases} \tag{2.4}$$

where the cutoff- and stopband-frequencies f_c=1200 Hz and f_s=5600 Hz were chosen by Shear to fit the Schiano data.

As discussed in the beginning of this section, the narrowness of $p(\tau_m, f_{c_m})$ control the frequency at which the position is predicted to move to the midline. The lowpass filter $G(f)$ determines the shape of the rolloff to the midline.

2.2.3 Lateralization of Tones Versus ITD and IID

Domnitz and Colburn (1977) conducted a matching experiment using a 500-Hz tone target stimulus with various combinations of ITD and IID: ITDs ranged from -1 to +1 ms and IIDs ranged from -9 to +25 dB. The pointer was a 500-Hz *tone* with variable IID. (The fact that the other experiments presented in this report used pointers of bandpass noise will be mentioned again in the concluding comments of this report.) Their results are shown in the left panel of Figure 2-6. For a given IID, the subjective lateral position follows a sinusoidal trend with an excursion inversely proportional to IID. That is, as the IID increased in magnitude, listeners found less position dependence on the ITD, so the position curves compressed. This is a reasonable outcome, since for large IIDs the input stimulus approaches monaurality. Another feature of note is the migration of the locations of each cue-reversal point, which occur where a monotonic change in IID/ITD produces a subjective reversal in previous lateral estimations. In Figure 2-6 cue-reversal points occur where the position curves have zero slope with respect to ITD; two locations where this occurs are 600 ms and -500 ms for IID=3 dB.

The right panel of Figure 2-6 shows the model predictions. The multiplicative weighting function in the original formulation of the position-variable model adequately models the dependence of the location of the cue-reversal points on ITD and IID. The predictions also match the experimental data for small ITDs and an IID of 0 dB, but are overlateralized for other combinations of ITD and IID. The position curves from these predictions also exhibit much more swing than the original predictions from Stern and Colburn (1978), and the reason for this is described below.

To understand the appearance of these data, we first examine a position curve for 0-dB IID. For a 0-dB IID, the product of $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$ forms a weighting function that emphasizes information in the CCF located at internal delays of smaller magnitude. Because the CCF for each ITD is a copy of the CCF at 0-ms ITD shifted by the ITD along the τ_m-axis, the centroid of the CCF weighted by the functions $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$ will vary according to each ITD, and because the CCF of a pure tone is periodic, the position curve has a periodic shape.
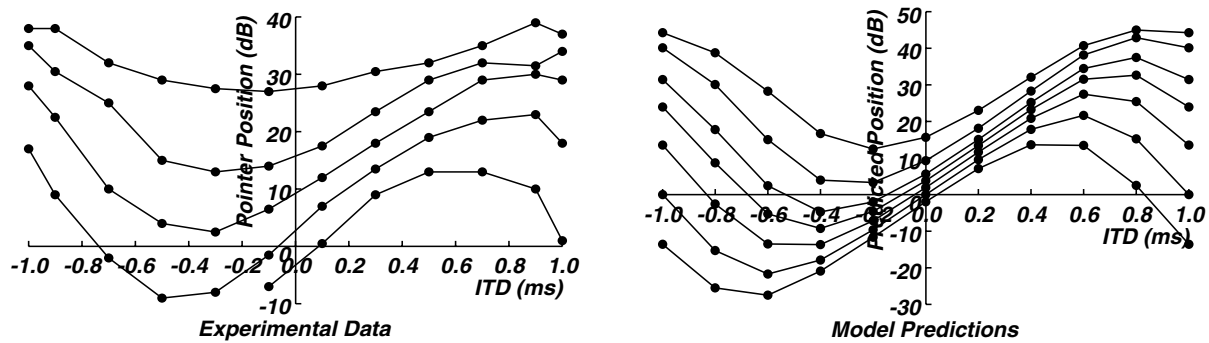
Figure 2-7 Experimental results from Domnitz and Colburn and predictions of the position-variable model. The left panel shows the experimental data, and the curves from bottom to top represent stimulus IIDs of -3, 3, 9, 15, and 25 dB, which can be seen from the pointer IIDs taken along a slice at 0-ms ITD. The right panel shows theoretical predictions for these data using a recent version of the model. From bottom to top, the curves were predicted using IID parameters of -3, 0, 3, 6, 9, 15, and 25 dB.

For non-zero IIDs, the intensity weighting function $L_I(\tau_m)$ will shift along the τ-axis, causing the locations of the cue-reversal points to shift. The movement of $L_I(\tau_m)$ will also cause the position curves themselves to be offset from one another. For large IIDs, the position curves of these predictions do not compress as those of the data because a small fraction of fibers is still synchronized to the input stimulus, causing a pronounced sinusoidal shape to the predictions. The predictions shown in Stern and Colburn (1978) exhibited this compression effect because they were performed using a much broader range of characteristic frequencies (CFs) than those present in the doubly-active region used in this report. The broad range of CFs included a greater number of unsynchronized fiber pairs, and the presence of these unsynchronized fiber pairs made the predictions more independent of ITD.

2.3 Summary

We described the operation of the position-variable model, and showed its predictions for two lateralization-of-tones experiments. We found that the shape of $p(\tau_m, f_{c_m})$ plays a primary role in determining the lateral position of high-frequency tones in Schiano *et al.* In the Domnitz and Colburn experiment, the lateral position of each IID-curve follows a sinusoidal trend over ITD, and the cue-points migrate with IID. In the next chapter introduce experimental bandpass-noise data of Stern, Zeiberg, and Trahiotis (Stern *et al.*). To predict this lateralization of the noise at various bandwidths, we will find that it was necessary to add another weighting mechanism called straightness to the model, and we will find that the addition of straightness does not significantly affect the Schiano and Domnitz and Colburn data.

Chapter 3
Incorporation of Straightness Weighting

Straightness refers to the consistency of interaural delay information over frequency in the auditory cross-correlation. The emphasis of straight features came about because without this weighting, the position-variable model was unable to describe the results of a class of bandpass-noise lateralization experiments. In Section 3.1   we describe the motivation for straightness weighting as well as how it was implemented. Using a more mathematically correct formulation of straightness weighting, we discuss in Section 3.2   the role of active and inactive fiber pairs to the lateralization process. In Section 3.3   we compare the predictions of a model that incorporates straightness weighting with the bandpass-noise data, and we will see that emphasizing straightness allows us to predict the lateralization of bandpass noise. Furthermore, as we shall see Section 3.4 , the use of straightness does not adversely affect the predicted tonal data of Schiano *et al.* (1986) or of Domnitz and Colburn (1977).

3.1 Motivation Behind Straightness Weighting

We now turn to the results of experiments describing the lateralization of bandpass noise presented with zero IID and various combinations of ITD, IPD, and bandwidth. Experimental results from a single subject was first presented in Stern *et al.* (1988a), and results from a greater number of subjects for some of the stimulus combinations were described in Trahiotis and Stern (1989). Buell and Trahiotis (1991) presented a more complete set of experimental data that also include the effects of IIDs. We show the subset of the Buell and Trahiotis data that corresponds to that found in Stern *et al.* (1988a).

The predictions of the original model, which are described in Section 3.1.1 , cannot describe the dependence of lateralization on bandwidth that is seen in the data. In Section 3.1.2   we explain this phenomenon in terms of the relationship between the straightness and the centrality of the data.

3.1.1 Lateralization of Bandpass Noise

Figure 3-6 shows the data from Buell *et al.* (1991) and predictions of the original form of the model. The experiment consisted of lateralization of a bandpass noise with a center frequency of 500 Hz and various combinations of ITD, IPD, and bandwidth. The ITD and IPD were chosen to produce modes of the interaural CCF at the same locations along the internal-delay axis at characteristic frequencies of 500 Hz, the center frequency of the noise. We denote each stimulus combination of ITD in μs and IPD in degrees as (ITD, IPD). The pointer was a 500-Hz noise with a 200-Hz bandwidth.

The lateral position of data from the (1500, 0) experiment was highly dependent on bandwidth: as the bandwidth increased from 0 Hz (a pure tone) to 400 Hz, the acoustic image moved from the left side to the right side of the head. As the ITD decreased while the IPD increased, the position becomes more and more independent of bandwidth: in the (0, 270) case, the position curve remained on the left side of the head. The predicted position does not move from the left to the right side of the head nearly enough as bandwidth increases.
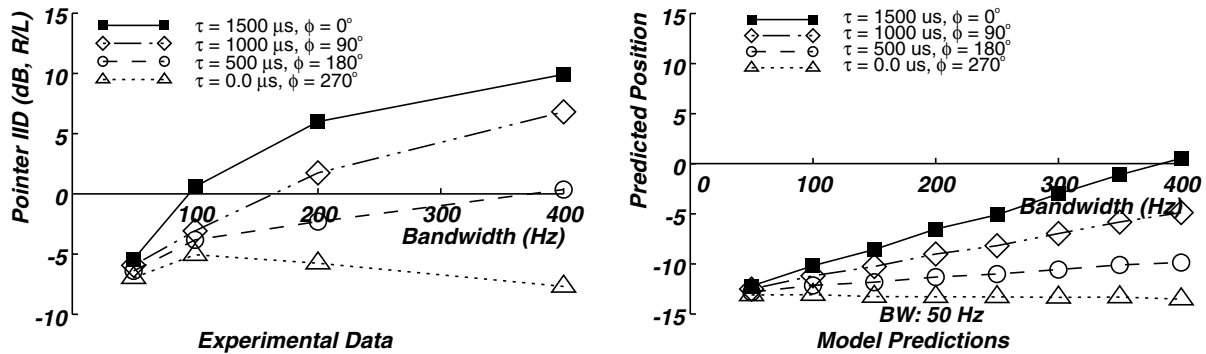
Figure 3-1: Experimental results and predictions for Buell and Trahiotis (1991). The left panel shows the average subjective position of four subjects in the experiment. The right panel shows predictions of the position-variable model.

3.1.2 Interaction Between Centrality and Straightness

In the following two subsections, we explain why the model does not describe the data of Stern *et al.* by examining the CCFs of the (1500, 0) and the (0, 270) and other stimulus conditions.

3.1.2.1 Analysis of (1500, 0) Stimulus Condition

To explain why the model does not describe the data of Stern *et al.*, we examine the (1500, 0) stimulus, which produced the widest change in subjective image position with bandwidth (refer to the position curve in Figure 3-6 marked by the solid squares). Recall that a 1500 μs ITD means that the signal to the right ear leads the one to the left ear by 1500 μs. Because of the 2000-ms periodicity of 500-Hz pure tones, a 1500 μs ITD is completely equivalent to an ITD of -500 μs. Thus, the stimulus presented leading strongly in time to one ear could appear as if it were leading in time to the contralateral ear (in this case, the left ear). As the bandwidth of the stimulus increases, the auditory system somehow uses the additional spectral cues to place the position to the ear corresponding to the actual time delay (the right ear). For further insight on how this happens, we examine the activity of the coincidence counters and interpret that activity in terms of an interaction between centrality and straightness.

Shown in Figure 3-2 are the CCFs of the (1500, 0) data with bandwidths of 50 Hz and 400 Hz. Each CCF is shown first as the output from the binaural display and then after weighting by $p(\tau_m, f_{c_m})$, the distribution of counting units.

Since noise with 500-Hz center frequency and a 50-Hz bandwidth is very similar to a 500-Hz pure tone, they share similar crosscorrelation properties. The CCF of that noise is periodic, with peaks at the ITD plus or minus multiples of the reciprocal of the center frequency. The peaks are straight because the output of each bandpass filter is dominated by the stimulus frequency. After weighting by the relative number of fiber pairs, only the most central peak at -500 μs remains, and so the centroid and position estimate would be dominated by this peak; *i.e.* by its effective centrality. (See left panels of Figure 3-2.)

On the other hand, the CCF of the 400-Hz-wide noise has only one straight mode at 1500 μs, and that mode corresponds to the actual value of ITD. The curved trajectories result because the stimulus bandwidth is greater than the bandpass filters' bandwidths, so the output of each filter is dominated by its center fre-
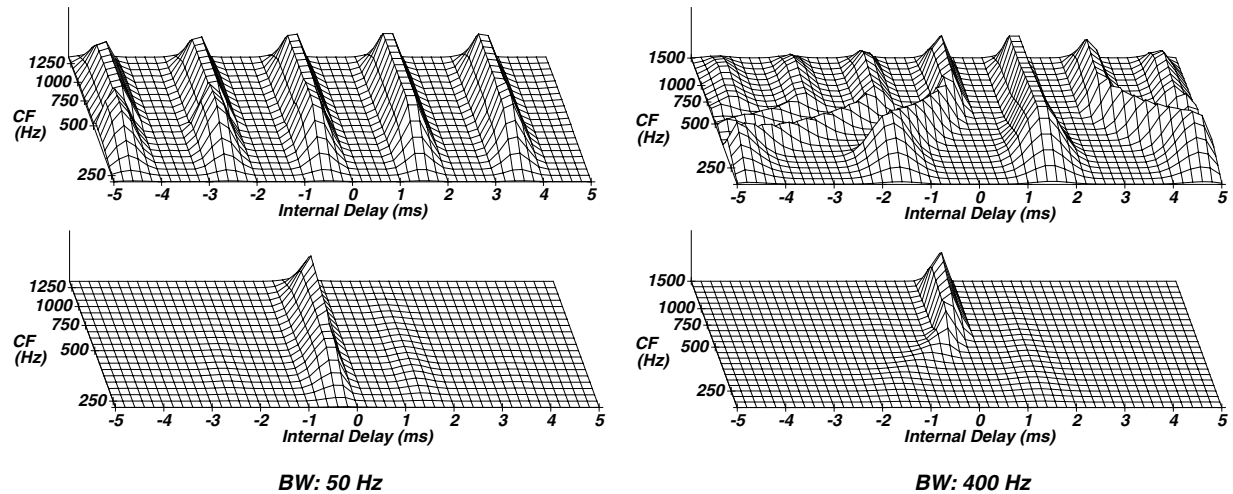
Figure 3-2: CCFs of (1500, 0) data with 50-Hz and 400-Hz bandwidths. The top panels show the output of the binaural display, and the bottom panels show the resulting CCF after weighting by $p(\tau_m, f_{c_m})$.

quency. This causes the periods of the CCFs at each CF to equal 1/CF. After weighting by $p(\tau_m, f_{c_m})$, the only significant activity occurs at the most central peak at 0 μs. Since there are relatively few coincidence-counting units with internal delays near 1500 μs, the straight peak that "encodes" the actual ITD is suppressed. (Refer to right panels of Figure 3-2.) From the results of this experiment, it appears that listeners use this straight trajectory to determine the actual time delay and the lateral position; however, the model predicts the centroid to be a midline position corresponding to the location of the central peak, explaining the discrepancy between data and predictions.

Enhancing the straight peaks of the CCF would improve the predictions for broadband noise since emphasizing the activity at 1500 μs would shift the centroid toward the right side of the head. Enhancing the straight peaks would not significantly affect the predictions for the narrowband noise, since *all* of the peaks are equally straight, so centrality would be the dominant factor in lateralization. As proposed in Stern *et al.* (1988a), lateralization may be determined by the interplay between straightness and centrality, where straightness dominates the lateral position of broadband stimuli, and centrality dominates the position of narrowband stimuli.

3.1.2.2 Analysis of (0, 270) and Other Stimulus Conditions

For further consideration, we turn to the (0, 270) data. In the limiting narrowband case, the combination of a 0-μs ITD and 270° IPD is functionally equivalent to a 1500-μs ITD and a 0° IPD. The CCFs of these two narrowband conditions are similar, resulting in almost identical lateral positions in the data. As the bandwidth increases, the peaks of the CCF diverge from straightness, as shown in Figure 3-3. The position remains lateralized on the left side of the head even for bandwidths of 400-Hz because the centroid is primarily derived from the peak at -500 μs, and this peak is the straightest as well as the most central. The model has no problem predicting the lateral position, and increasing the effect of straightness would not change these predictions.

Figure 3-4 shows the locations of the peaks of the CCF for each of the four stimulus configurations. The ITD and IPD jointly determine the location of the straightest trajectory, and the IPD controls the curvature of the trajectories. Referring to the experimental data in Figure 3-6, we note that the relative straightness of
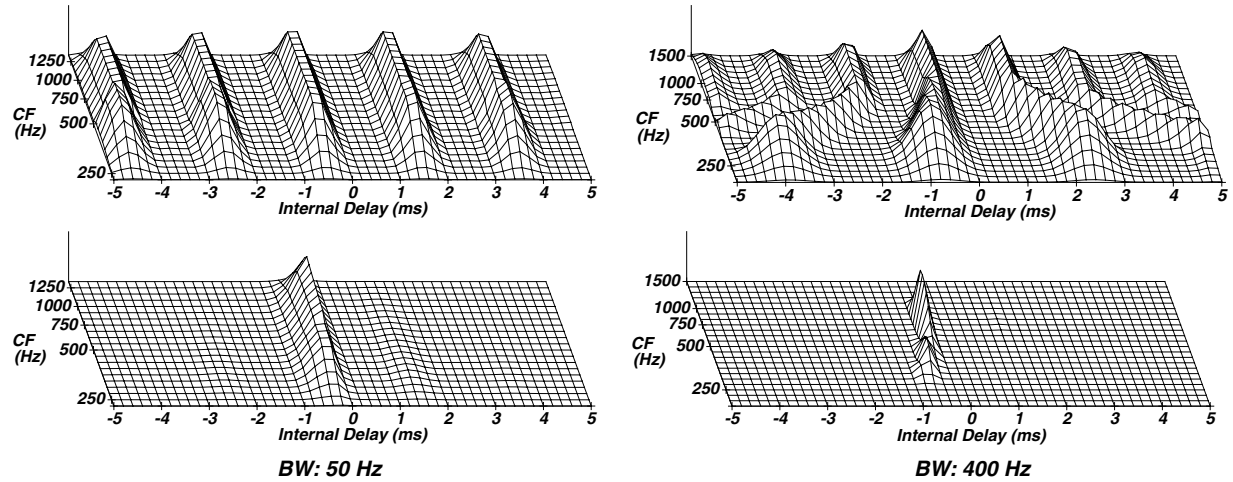
Figure 3-3: CCFs of (0, 270) data with 50-Hz and 400-Hz bandwidths. Each CCF is shown before and after weighting by $p(\tau_m, f_{c_m})$.
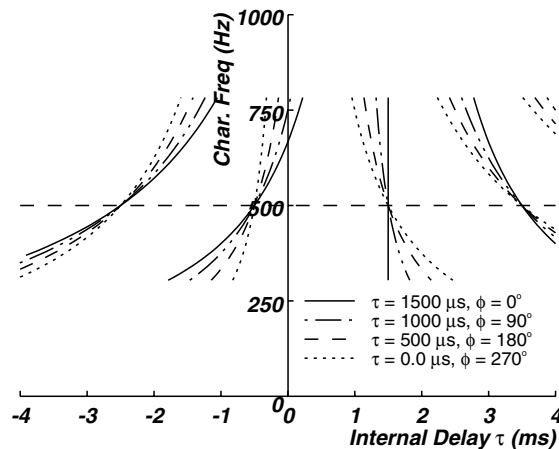

Figure 3-4: Trajectories of the *peaks* of the CCF for the (ITD, IPD) stimulus combinations used in Stern *et al.* (1988a) and Buell and Trahiotis (1991).

the peak at 1500 µs seems to control the lateral position observed for the stimuli with broader bandwidths, and the lateral positions observed for stimuli presented with intermediate combinations of ITD and IPD fall between the positions of data presented in the (1500, 0) and (0, 270) configurations. In the next section we discuss the introduction of straightness weighting into the position-variable model.

3.1.3 Implementation of Straightness Weighting

To implement a straightness weighting stage, the position-variable model uses a *second* layer of coincidence-counting units that detects consistency of interaural delay information across frequency of the response of units in the first (original) layer of coincidence-counting units. The second-level units record straightness by facilitating the outputs of first-level units that are consistent, relative to one another, over frequency. In an earlier study, Shear (1987) proposed a number of other procedures to implement straightness weighting, but none of them were effective. The description of the current method of straightness weighting first appeared in Stern *et al.* (1991).

Unlike the first-level units, which have only one input coming from each of the two ears, each second-level unit has a multiple number of inputs, N. A particular second-level unit takes input from N first-level units that have the same characteristic delay but characteristic frequencies that are assumed to be logarithmically spaced. Mathematically, this is shown as

$$S_m(\tau_m, f_{c_m}) \;=\; E\left\{ \prod_{i=-(N-1)/2}^{i=(N-1)/2} L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s) \right\} \tag{3.1}$$

where N is the "order" of straightness, Δf_s is the straightness step, and $E\{\cdot\}$ denotes the expected value. The order and straightness step refer to the number of first-level counting units and the frequency spacing between those first-level units, respectively, taken as input into a straightness unit. In the implementation of the straightness algorithm in this report, we have chosen the straightness steps to be multiples of the uniform log-frequency spacing of the first-level units.

We rewrite Equation (3.1) to take advantage of the statistical independence of the firing rates of the fiber pairs:

$$S_m(\tau_m, f_{c_m}) \;=\; \prod_{i=-(N-1)/2}^{i=(N-1)/2} E\{L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} \tag{3.2}$$

Equation (3.2) states that each second-level straightness unit fires at a rate proportional to the product of the firing rates of the N first-level units. For example, Figure 3-5 depicts the output of a group of first-level units in the binaural display, and the output of two arbitrary second-level units; the size of each unit represents its firing rate. The two straightness units in this example each have an order of 3 and a straightness step of 2. The larger straightness unit in Figure 3-5 shows how a straight trajectory is emphasized, and the smaller unit shows how curved trajectories are effectively deemphasized relative to the straighter ones.
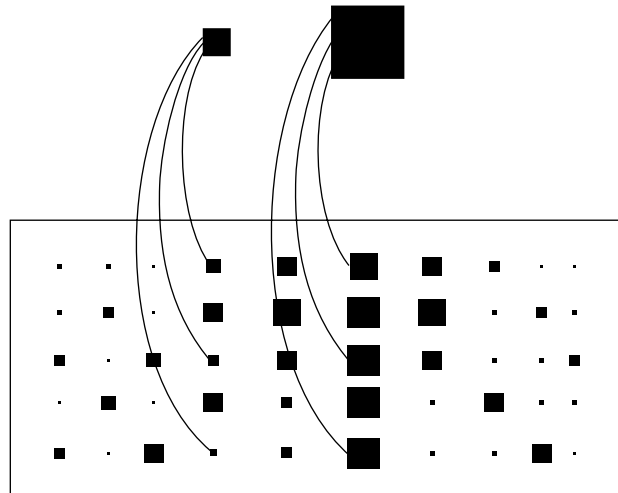

Figure 3-5: Relationship between output of first-level units from binaural display and a second-level straightness unit. Size of each unit indicates its firing rate.

The straightness unit responds proportionally to the geometric mean of its first-level units' firing rates. If the outputs of all of the N first-level units are high, then the second-level (straightness) unit will fire at a high rate. If the outputs of all of the N first-level units are low, then the straightness unit will fire at a low rate. Intermediate values of first-level firing-rate products result in a straightness (second-level) firing rate between the two extremes. The output of the network of second-level units will be the original CCF with its consistent trajectories emphasized; the graphical effect of straightness weighting is shown in the next section.

To control the degree of straightness weighting, we found that increasing the number of units (the order) or the spacing between the units (the step) places more emphasis on the straight trajectories. Increasing the order attenuates the non-straight modes by virtue of including more inputs to the straightness unit, and increasing the frequency step accomplishes the equivalent effect by spreading the first-level inputs farther apart. The specific effects of using various combinations of order and step over another are best determined through experiment.

3.2 Consideration of Active Fiber Proportions

The algorithm for weighting of straightness does not fully account for the presence of inactive fibers. An expression for the straightness algorithm is shown below:

$$S_m(\tau_m, f_{c_m}) = \prod_{i = -(N-1)/2}^{i = (N-1)/2} E\{L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} \tag{3.3}$$

If we know that *all* of the first-level inputs to a particular straightness unit $S_m(\tau_m, f_{c_m})$ are doubly-active (i.e. at the first-level units' CFs accessed by the straightness unit, all of the fiber pairs are doubly-active), then we can compute the expected number of counts for each unit $L_m(\tau_m, f_{c_m})$, and Equation (3.3) holds. As the CF of the straightness unit moves away from the bandwidth of the input stimulus and hence, the doubly-active region, then some of the fibers that enter the first-level units will become inactive. With the presence of inactive fibers, we know only in a statistical sense what the number of counts for an individual unit is. Section 3.2.1   shows how we modified the straightness routine to include properly the effects of inactive fibers. Section 3.2.2   presents a simplification of this more accurate implementation of the model.

3.2.1 Straightness and the Role of Active Fibers

To derive an expression for a form of straightness weighting that more accurately accounts for the permutations of doubly- and singly-active and doubly-inactive fiber pairs, we consider two simplified examples and extend the results to a general case. The simplest case of straightness occurs when the order N is one. In this case, no straightness weighting occurs since the second-level unit would merely pass the firing rate

of the first-level unit. We can distribute the expected number of counts L_m into three components, each of which depend on the number of active nerve fibers to each ear (cf. Stern, 1976):

$$E\{L_m(\tau_m, f_{c_m})\} \; = \; E\{L_m(\tau_m, f_{c_m})|\text{both fibers active}\}P\{\text{both fibers active}\} \tag{3.4}$$

$$+ \, E\{L_m(\tau_m, f_{c_m})|\text{one fiber active}\}P\{\text{one fiber active}\}$$

$$+ \; + \; \{L_m(\tau_m, f_{c_m})|\text{one fiber active}\}P\{\text{one fiber active}\}$$

$$= \; \mathcal{L}_2(\tau_m, f_{c_m})\eta_2(f_{c_m}) + \mathcal{L}_1(\tau_m, f_{c_m})\eta_1(f_{c_m}) + \mathcal{L}_0(\tau_m, f_{c_m})\eta_0(f_{c_m}) \tag{3.5}$$

The variables $\mathcal{L}_i(\tau_m, f_{c_m})$ refer to the expected number of counts given that i fibers are active. If both fibers are active, then the expected number of counts observed by that fiber pair is proportional to the interaural crosscorrelation function of the stimulus after peripheral bandpass filtering. If at least one of the fibers in the pair is inactive, the firing rates of each ear become statistically independent of each other, and the crosscorrelation function becomes the product of the average firing rates: $\mathcal{L}_1(\tau_m, f_{c_m}) = (200)(50)$ and $\mathcal{L}_0(\tau_m, f_{c_m}) = (50)(50)$. The variables $\eta_i(f_{c_m})$ represent the fraction of the total number of fiber pairs at f_{c_m} with i active fibers.

We can extend the results shown above by examining the output of a straightness unit that has an order $N=2$. The expression for the straightness unit is

$$E\{S_m(\tau_m, f_{c_m})\} \; = \; E\{L_m(\tau_m, f_{c_m} + \Delta f_s)\} \tag{3.6}$$

where $S_m(\tau_m, f_{c_m})$ is the output of the straightness unit and $L_m(\tau_m, f_{c_m})$ and $L_m(\tau_m, f_{c_m} + \Delta f_s)$ are the first-level firing rates from units with the same internal delay but adjacent CFs.

To find S_m correctly, we again expand the expression using conditional expectations to consider the proportions of active fibers. The final result is

$$E\{S_m(\tau_m, f_{c_m})\} \; = \; E\{L_m(\tau_m, f_{c_m})L_m(\tau_m, f_{c_m} + \Delta f_s)\} \tag{3.7}$$

$$= \; \sum_{j_0 = 0}^{2} \sum_{j_1 = 0}^{2} \left[\prod_{i = 0}^{1} E\left\{ L_m(\tau_m, f_{c_m} + i\Delta f_s) \Big|_{j_i \text{ is an active fiber}} \right\} \eta_{j_i}(f_{c_m} + \Delta f_s) \right]$$

$$= \; \sum_{j_0 = 0}^{2} \sum_{j_1 = 0}^{2} \left[\prod_{i = 0}^{1} \mathcal{L}_{j_i}(\tau_m, f_{c_m}) \eta_{j_i}(f_{c_m} + \Delta f_s) \right]$$

Equation (3.7) shows the expected number of counts from a straightness unit with two inputs. Now that we are concerned with more than one CF (as opposed to the first case of $N=1$), so we need to introduce slightly modified notation used in the above equation:

- $N \equiv$ order of straightness (number of fiber pairs in straightness group); here, $N=2$
- $\tau_m \equiv$ internal delay of counting unit (as defined in Chapter 2)
- $i \equiv$ index of fiber pair in straightness group
- $f_{c_m} + i \cdot \Delta f_s \equiv$ CF of ith fiber pair; Δf_s varies so that it is a constant over log-frequency units
- $j_i \equiv$ index for the number of active fibers (0, 1, or 2) into the ith fiber pair; j_i implicitly refers to fibers with CF $f_{c_m} + i \cdot \Delta f_s$
- $\mathcal{L}_{j_i}(\tau_m, f_{c_m} + i \cdot \Delta f_s) \equiv$ expected number of coincidences from first-level counting unit, given j_i active fibers
- $\eta_{j_i}(f_{c_m} + i \cdot \Delta f_s) \equiv$ fraction of total fiber pairs that have j_i active fibers

For the general case of N inputs (order N), the expression for the expected number of counts from a particular straightness unit is

$$E\{S_m(\tau_m, f_{c_m})\} = E\left\{ \prod_{i=-(N-1)/2}^{i=(N-1)/2} L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s) \right\} \tag{3.8}$$

It can be shown that by extending the number of summations and product terms in Equation (3.7) to cover the number of inputs N, the straightness expression in Equation (3.8) becomes

$$E\{S_m(\tau_m, f_{c_m})\} = \sum_{j_{-(N-1)/2}=0}^{2} \cdots \sum_{j_{(N-1)/2}=0}^{2} \left[\prod_{i=-(N-1)/2}^{i=(N-1)/2} \mathcal{L}_{j_i}(\tau_m, f_{c_m} + i \cdot \Delta f_s) \cdot \eta_{j_i}(f_{c_m} + i \cdot \Delta f_s) \right] \tag{3.9}$$

3.2.2 Simplifying the New Straightness Algorithm

To compute the output of each straightness unit using the new algorithm requires the order of 3^N times as many calculations as the old straightness algorithm of simply computing the product of N first-level firing rates. However, as shown below, we can simplify the new algorithm to a form resembling the old straightness algorithm.

Equation (3.7) consists of a large number of sums of product terms, and by rearranging its terms, it can be expressed as the following equivalent product of a summation[1]:

$$E\{S_m(\tau_m, f_{c_m})\} = \prod_{i=-(N-1)/2}^{i=(N-1)/2} \left[\sum_{j_i=0}^{2} \mathcal{L}_{j_i}(\tau_m, f_{c_m} + i \cdot \Delta f_s) \cdot \eta_{j_i}(f_{c_m} + i \cdot \Delta f_s) \right] \tag{3.10}$$

---

1. Equation (3.10) relies heavily on the distributive property to generate each product term in Equation (3.7). For example, compare the terms generated in each equation for the case $N=2$ by letting i vary from 0 to 1 and, for ease of notation and visualization, substituting a dummy variable $A_i^{j_i} \equiv \mathcal{L}_{j_i}(\tau_m, f_{c_m} + i \cdot \Delta f_s) \cdot \eta_{j_i}(f_{c_m} + i \cdot \Delta f_s)$.

The expression in brackets from Equation (3.10) resembles the bracketed part of the following expression for computing the timing function $L_T(\tau_m)$ (Stern and Colburn, 1978):

$$E\{L_T(\tau)\} = N \cdot p(\tau_m)[\mathcal{L}_2 \cdot \eta_2 + \mathcal{L}_1 \cdot \eta_1 + \mathcal{L}_0 \cdot \eta_0] \tag{3.11}$$

The terms in the brackets must be recomputed for each CF. The notation used in Equation (3.11) is as follows: N is the total number of fiber pairs, $p(\tau_m)$ is a frequency-independent form of $p(\tau_m, f_{c_m})$, and the $\mathcal{L}_i$ and η_i terms represent $\mathcal{L}_i(\tau_m, f_{c_m})$ and $\eta_i(f_{c_m})$, respectively. The expected value of the timing function $L_T(\tau_m)$ represents the total number of coincidences at each internal delay. In essence, the bracketed term in Equation (3.10) computes the relative number of coincidences at a particular CF. Defining the relative number of coincidences at a particular CF f_{c_m} as

$$E\{L'_T(\tau_m, f_{c_m})\} = \sum_{k=0}^{2} \mathcal{L}_k(\tau_m, f_{c_m}) \cdot \eta_{j_i}(f_{c_m}), \tag{3.12}$$

we can express Equation (3.10) as follows:

$$E\{S_m(\tau_m, f_{c_m})\} = \prod_{i=-(N-1)/2}^{i=(N-1)/2} E\{L'_T(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} \tag{3.13}$$

Equation (3.13) resembles Equation (3.8), the old straightness equation, except that instead of calculating a product of first-level firing rates, we calculate a product of expected relative number of coincidences at each CF in the straightness product. By forming a product of sums instead of a sum of products, we can save much computational overhead using this method for straightness weighting.

3.3 Comparison of Predictions to Bandpass Noise Experiment

In Figure 3-6 the predictions of the model incorporating the straightness weighting described above are compared to the data of Buell and Trahiotis (1991). This implementation of the model with straightness weighting clearly provides a better description of the perceived laterality of the stimuli for each (ITD, IPD) combination. Note how the predicted positions for stimuli in the (1500, 0) and (1000, 90) configurations now show the crossover effects and bandwidth dependence of the data while the (500, 180) and (0, 270) positions correctly remain near the left side of the head.

In Figure 3-7 the CCF of the (1500, 0) stimuli with 400-Hz bandwidth is shown with and without $p(\tau_m, f_{c_m})$. Note that the straightness function has facilitated the output of the counting units located near the straight trajectory, and that the straight peaks are more pronounced than before. Even after weighting by the $p(\tau_m, f_{c_m})$ function, the straight peak is prominent enough to move the centroid to the right side of the head.

A compression function was also introduced to make the range of predictions better match the lateralization data. The function we used was of the form $\hat{P}_c = \alpha\tan^{-1}(\beta \cdot \hat{P})$, where α and β are fitted compression parameters, and $\hat{P}_c$ and $\hat{P}$ are compressed and uncompressed position variables, respectively. The selectable parameters were chosen to produce predictions that were a best mean-square fit to the data.
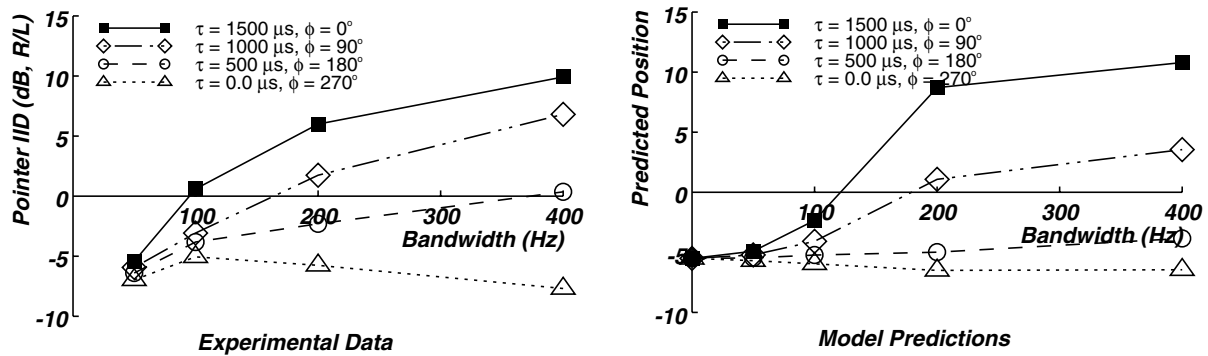
Figure 3-6: Experimental results and revised predictions for Buell and Trahiotis (1991). The left panel again shows the average subjective position of the four subjects in the experiment. The right panel shows predictions from the position-variable model that uses straightness weighting.
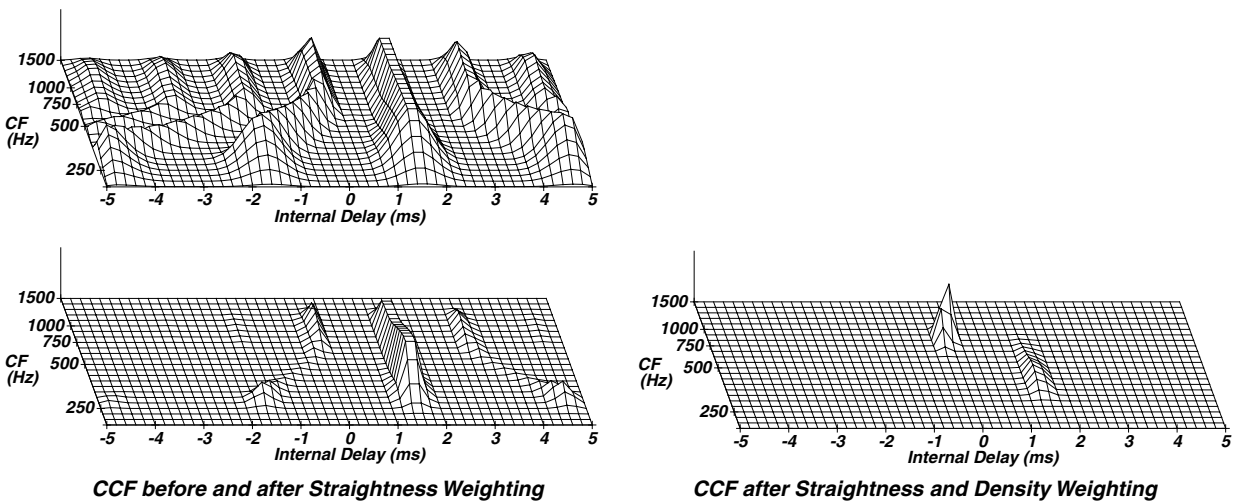

Figure 3-7: CCFs of (1500, 0) stimulus. The top panel shows the raw, unweighted CCF. The CCF immediately underneath shows the effect of straightness weighting, and the CCF on the bottom right shows the effect of straightness and $p(\tau_m, f_{c_m})$ weighting.

3.4 Comparison of Predictions to Tones Experiments

After successfully predicting the bandpass-noise data of Buell and Trahiotis (1991), we used the model with straightness weighting to repredict the data of Domnitz and Colburn, and Schiano *et al.* to determine the effect straightness had on those predictions. We did not expect a great change to occur for the tonal stimuli, because all of the peaks are equally straight and consistent over frequency, and straightness would affect the CCF's peaks equally. After straightness weighting, however, those straight peaks do tend to be narrower and more pronounced than before, which may cause changes from the predictions of the original model.

3.4.1 Effect of Straightness Weighting on Domnitz and Colburn Data

As expected, the addition of straightness weighting did not significantly affect the data of Domnitz and Colburn (1977). It caused a slight increase in the excursion of the predicted lateralization curves, which is a side effect of sharpening the peaks of the CCF.

3.4.2 Effect of Straightness Weighting on Schiano *et al.* Data

The modified model that incorporates straightness weighting predicts that the lateral image falls to the midline at a lower frequency than did the original model, as shown in Figure 3-8. The explanation is based on similar principles given in Section 2.2.2 on page 10. For a low-frequency tone, the CCF after weighting by $p(\tau_m, f_{c_m})$ has one main peak located at the ITD, which causes the stimulus to be lateralized according to that ITD. For higher-frequency tones, the CCF after $p(\tau_m, f_{c_m})$ exhibits not only that main peak, but also secondary peaks that tend to shift the centroid toward the midline.
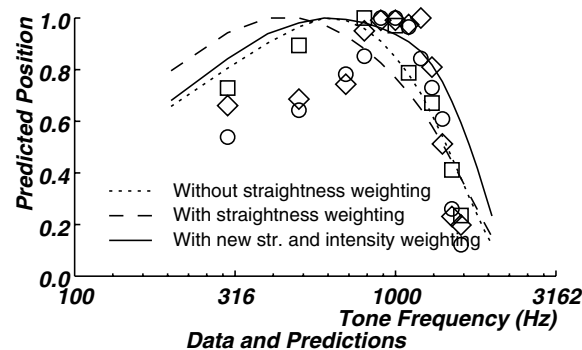

Figure 3-8: Predictions of Schiano *et al.* using the position-variable model. The dotted line shows the original model's predictions (without straightness weighting), and the dashed line shows the results after accounting for straightness. The solid line shows the predictions compensated for straightness weighting by modifying $p(\tau_m, f_{c_m})$ and $G(f)$. Geometric symbols represent the data. Vertical position has been normalized.

The addition of straightness weighting sharpens the peaks so that their amplitudes are much greater than the amplitudes of the valleys. In fact, after weighting by $p(\tau_m, f_{c_m})$, the secondary peaks of lower-frequency tones are much less attenuated than they were without explicit straightness weighting. These more powerful secondary peaks again move the image to the midline, but at a lower frequency than before. Therefore, as we increase the relative strength of straightness weighting, we must make the function describing $p(\tau_m, f_{c_m})$ narrower as a function of frequency to maintain the accuracy of the predictions. We accomplished this by changing the value of parameter l_P in Equation (2.3) on page 11 from -1.1 to -1.25, which had the effect of making $p(\tau_m, f_{c_m})$ narrower and steeper at each frequency.

Narrowing $p(\tau_m, f_{c_m})$ to prevent the higher-frequency predictions from returning to the midline at too low a frequency also necessitated changing the shape of the lowpass filter $G(f)$ in order to predict the Schiano *et al.* data at higher frequencies. We increased the steepness of the lowpass filters by lowering the stopband frequency from f_c = 5600 Hz to f_s=3200 Hz in Equation (2.4) on page 12. Predictions of the model without

straightness weighting, and predictions using straightness and the new $p(\tau_m, f_{c_m})$ and $G(f)$ parameters, are both shown in Figure 3-8. The revised predictions with straightness adequately model the trends of the data.

3.5 Summary

We found that in order to predict the results of the bandpass noise data at 0-dB IID of Buell and Trahiotis (1991), we must use straightness weighting to emphasize consistent ITDs over frequency. The reason could be seen by looking at the CCFs of the stimulus conditions: the listeners seem to use the relative straightness of the peaks as a cue for their perceived extent of laterality. Incorporating straightness weighting in the position-variable model results in a more accurate description of the experimental data. In general, centrality dominates the lateralization of narrowband stimuli and straightness dominates the lateralization of stimuli with wider bandwidths. The effect of straightness weighting is minimal on the predictions for the Domnitz and Colburn (1977) experiment, which was performed using 500-Hz tones. In order to predict the form of the data of Schiano *et al.* (1986), however, it was necessary to modify the distribution of fiber pairs $p(\tau_m, f_{c_m})$ and the lowpass filter function $G(f)$ because the straightness weighting altered the effect of centrality at each tone's frequency. Finally, we found it necessary to revise the straightness routine to account for inactive fibers, which led to a lengthy expression that was simplified to resemble the old form of straightness.

Chapter 4
Effects of Interaural Intensity
Differences

After showing that the position-variable model that incorporated straightness weighting was able to de-
scribe the Stern *et al.* (1988a) data, the next logical step was to expand the set of predictions to include
the effect of IIDs, and this effort is described in the remaining chapters of this report. In Section 4.1 we
examine data from Buell and Trahiotis (1991), which consisted of lateralization of bandpass noise for various
combinations of ITD, IPD, bandwidth, *and* IID. We will find that the model is unable to describe certain results
from that experiment because its method of intensity weighting does not interact properly with the effect of
straightness weighting. In Section 4.2 we discuss some unsuccessful modifications to the model that were
implemented in an attempt to make the predictions better fit the data. We eventually realized that the form
of intensity weighting used in the original formulation of the position-variable model is fundamentally unable
to describe the joint dependence of lateral position on ITD, IID, IPD, and bandwidth, and in the next chapter
we propose a remedy by using a new type of intensity weighting.

4.1   Interaction Between Straightness and IID

In this section we observe the effect of incorporating intensity differences to the bandpass noise stimuli
described in the previous chapter. In Section 4.1.1 we present the experimental data from Buell and Tra-
hiotis (1991) and we compare the data to the predictions of the model. In Section 4.1.2 we explain the
differences that appear between the data and the theoretical predictions.

4.1.1 Lateralization of Noise over ITD, IPD, Bandwidth, and IID

Subjects listened to bandpass noise with various IIDs and with one of the four combinations of ITD and
IPD used in Stern *et al.* (1988a): 1500 µs ITD, 0˚ IPD; 1000 µs, 90˚; 500 µs, 180˚; and 0 µs, 270˚. For each
(ITD, IPD) combination the bandwidth varied from 50 to 400 Hz; within each bandwidth the IID ranged from
-9 to +9 dB. The pointer was a band of 500-Hz noise with a 200-Hz bandwidth and adjustable IID. The results
of these experiments are shown in Figure 4-1. We focus again on the outermost conditions, which are shown
in the top left and bottom right of the figure, because the middle stimulus configurations represent interme-
diate cases of the extreme conditions.

Figure 4-2 shows the model's predictions for the (1500, 0) and (0, 270) data using straightness, inten-
sity, and compression parameters that produced a best visual fit to the data. The shape of $p(\tau_m, f_{c_m})$ and of
the lowpass filter $G(f)$ were still constrained by the Schiano *et al.* (1986) results. The figure shows that the
predicted positions of the (1500, 0) stimuli presented with intermediate bandwidths do not exhibit the extent
of laterality demonstrated by the listeners. More notable, however, is the departure of the shape of the pre-
dicted (0, 270) position curves from the data: for positive IIDs, many of the bandwidth positions remain in-
correctly lateralized on the left side of the head, and the curves diverge and spread in a fashion that is not
even hinted at by the experimental data. These results can be interpreted as a fundamental conflict of
straightness requirements between each set of stimulus conditions that is caused by the interaction be-
tween the type of straightness weighting and the type of intensity weighting used.
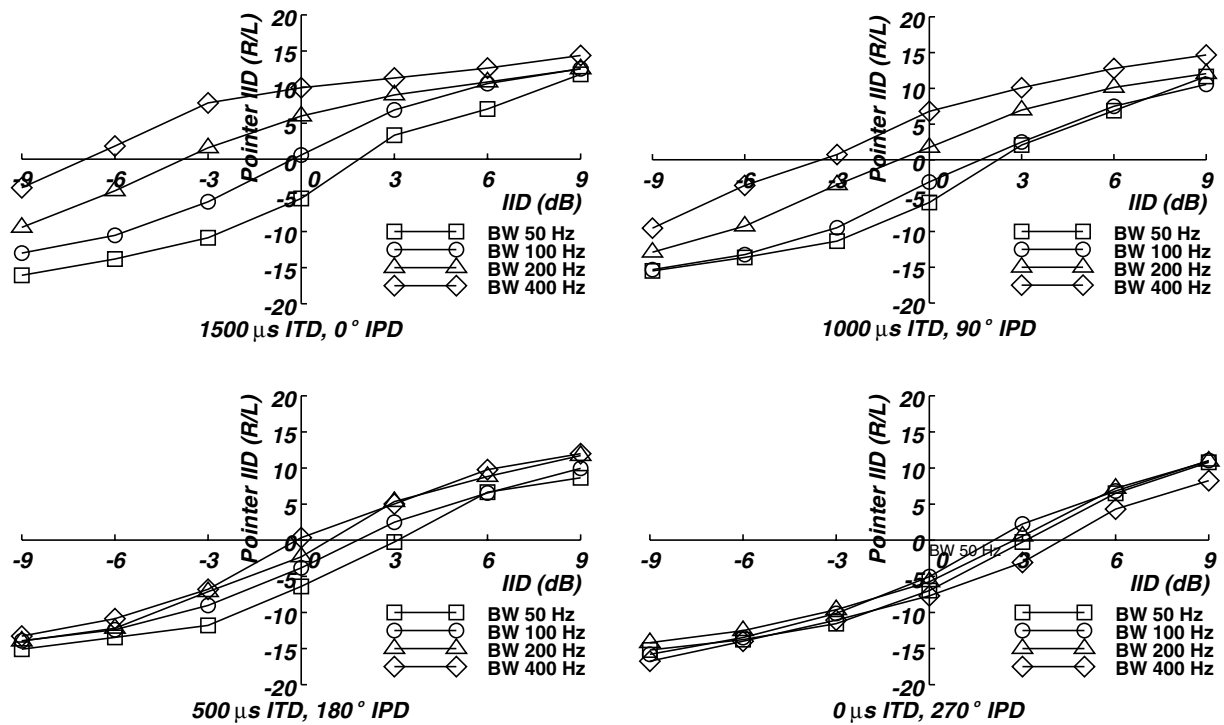
Figure 4-1: Experimental results from Buell and Trahiotis (1991). Starting from the top-left panel and moving counterclockwise, the IPD increases while the ITD decreases. Bandwidths are represented by different symbols.
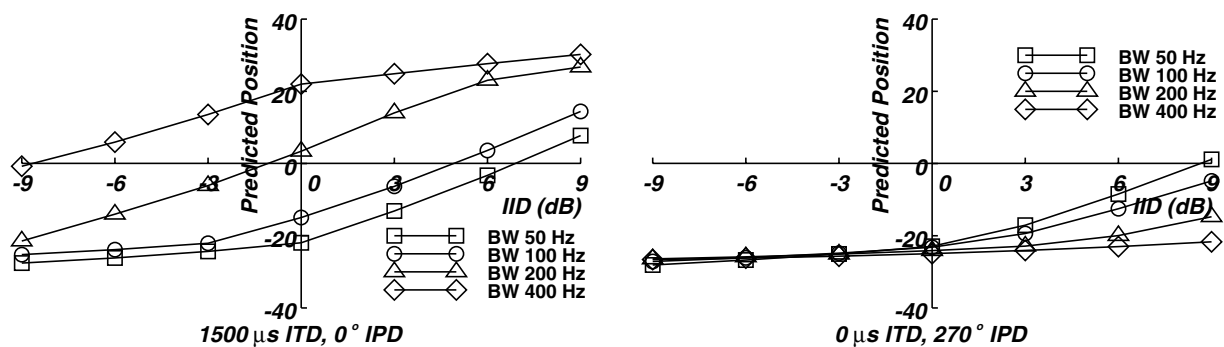

Figure 4-2: Predictions for Buell and Trahiotis (1991) experiment. The straightness and intensity parameters were chosen to produce a best set of predictions possible for the two outermost conditions using the current model.

4.1.2 Interaction Between Weighting of Straightness and IID

In the two subsections that follow, we analyze the results for each of the two outermost stimulus conditions presented above. We explain the discrepancies between predictions and data by examining the lateralization effects caused by weighting each CCF by the $p(\tau_m, f_{c_m})$, the function describing the density of fiber pairs, and the intensity weighting function $L_I(\tau_m)$. For each stimulus condition, we show the effects of a 0-dB IID, a negative IID, and a positive IID.

4.1.2.1 Analysis of (1500, 0) Stimulus Condition

The experimental stimuli at 0-dB IID were identical to those used in Stern *et al.* (1988a), which was discussed in Section 3.1.2.1   on page 15, with CCFs shown in Figure 3-2 on page 16. A summary of why the predictions for 50-Hz and 400-Hz bandwidths match the data follows. Before weighting by $p(\tau_m, f_{c_m})$, the CCF of the 50-Hz-wide noise has straight peaks located at 1500 μs $\pm\, k$(2000 μs), where k is an integer. After weighting for straightness and $p(\tau_m, f_{c_m})$, the peak that remains is located at -500 μs, causing the predicted position to be on the left side of the head. For the 400-Hz-wide data, the raw CCF has one straight peak at 1500 μs and curved modes elsewhere. After straightness and $p(\tau_m, f_{c_m})$ weighting, the straight trajectory located at 1500 μs causes the predicted position to move to the right side. The intensity function $L_I(\tau_m)$ does not provide a significant contribution in this 0-dB IID case because the product of $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$ is essentially $p(\tau_m, f_{c_m})$.

As the IID grows positive, the product of $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$ forms a function that has the greatest amplitude for some positive internal delay, which shifts the centroids, and hence the predicted positions, upward to the right side of the head. The combination of the two weighting functions effectively cause the centrality effects to be centered about a nonzero value of τ_m, as seen in the panel on the right side of Figure 4-3.
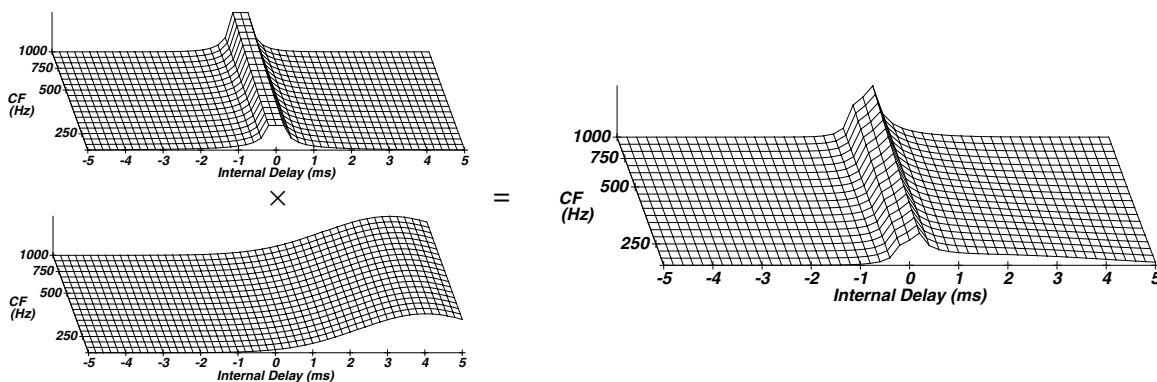
Figure 4-3: Combination of $p(\tau_m, f_{c_m})$ and $L_I(\tau_m)$ to form a new centrality function. The two smaller panels on the left show $p(\tau_m, f_{c_m})$ and $L_I(\tau_m)$, respectively. (The $L_I(\tau_m)$ shown here has a mean of 4 ms, which represents a particular value of positive IID.) The product of these two functions is shown in the right panel. A negative IID would result in mirror images of each of the CCFs.

As the IID becomes negative, the same concept holds true, except that now the multiplicative combination of the intensity function $L_I(\tau_m)$ and the density function $p(\tau_m, f_{c_m})$ forms a more spread-out function cen-

tered at more negative values of internal delay, which causes all of the predicted positions to shift downward to the left side of the head.

Overall, the data show that the lateral position is dependent on the bandwidth of the stimulus, and the model is able to predict this dependency through the interaction between the emphasis of straightness and the shifted effective centrality weighting of the CCF. The shift of centrality moves the position curves vertically while the straightness weighting properly emphasizes the role of the straight peaks to model the spread of the curves at each IID.

4.1.2.2 Analysis of (0, 270) Stimulus Condition

We now turn to the (0, 270) stimulus condition. Section 3.1.2.2 and Figure 3-3 (on pages 16 and 17, respectively) and discuss and show the CCFs at 0-dB IID. For narrow bandwidths, the 50-Hz-wide CCFs of the (0, 270) and the (1500, 0) stimuli are for all intensive purposes, identical. As explained in Section 4.1.2.1 , at 0-dB IID, the central peak at -500 μs causes the predicted position to be on the left side of the head. The CCFs diverge as the bandwidth of the noise increases, however. The CCF of the (0, 270) noise with 400-Hz bandwidth and 0-dB IID has its straightest mode at -500 μs. After weighting for straightness and $p(\tau_m, f_{c_m})$, the mode at -500 μs turns out to be both the straightest and most central peak, causing the predicted position to remain on the left side of the head close to the 50-Hz position. This straight-and-central mode at -500 μs will play an important role in explaining the predictions for non-zero IIDs.

For negative IIDs, the intensity-weighting pulse is located at negative values of internal delay, and the spread-out combination of $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$ brings the centroid and the predicted positions to the left side of the head. Since the peak at -500 μs (after weighting by $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$) exhibits the greatest amount of straightness and centrality for stimuli with negative IIDs, it dominates the predictions, and the position curves show very little dependence on IID.

A significant divergence of the predictions from the data occurs for positive IIDs: the (0, 270) lateralization stimuli are lateralized the right side of the head in a fashion that is also relatively independent of bandwidth. The predictions of the model for these data, however, are spread apart and remain on the left side of the head for most bandwidths. This can be explained by examining Figure 4-3, which shows the CCFs of the 50-Hz and 400-Hz bandwidths after weighting for straightness, $p(\tau_m, f_{c_m})$, and $L_I(\tau_m)$ (for a positive IID). The figure shows that the effect of a positive IID on the product of $L_I(\tau_m)$ and $p(\tau_m, f_{c_m})$ was to emphasize features of the CCF at positive delays. For the 50-Hz-wide stimulus, all of the peaks of the CCF are equally weighted by straightness, but for positive IIDs the effective centrality (including the effects of IID) emphasizes the peak at 1500 μs, causing the predicted position to move to the right side. However, as the stimulus bandwidth increases, the straight peak at -500 μs becomes weighted more heavily, causing the predicted position to shift toward the left side of the head..

Unfortunately, these predictions are clearly inconsistent with the trends of the data, so this type of intensity weighting seems to be fundamentally unable to describe the independence of the (0, 270) positions with respect to bandwidth. In the following sections we describe several relatively minor modifications to this type of intensity weighting that were introduced in an attempt to describe more accurately the (0, 270) data without adversely affecting the ability of the model to describe the (1500, 0) data.
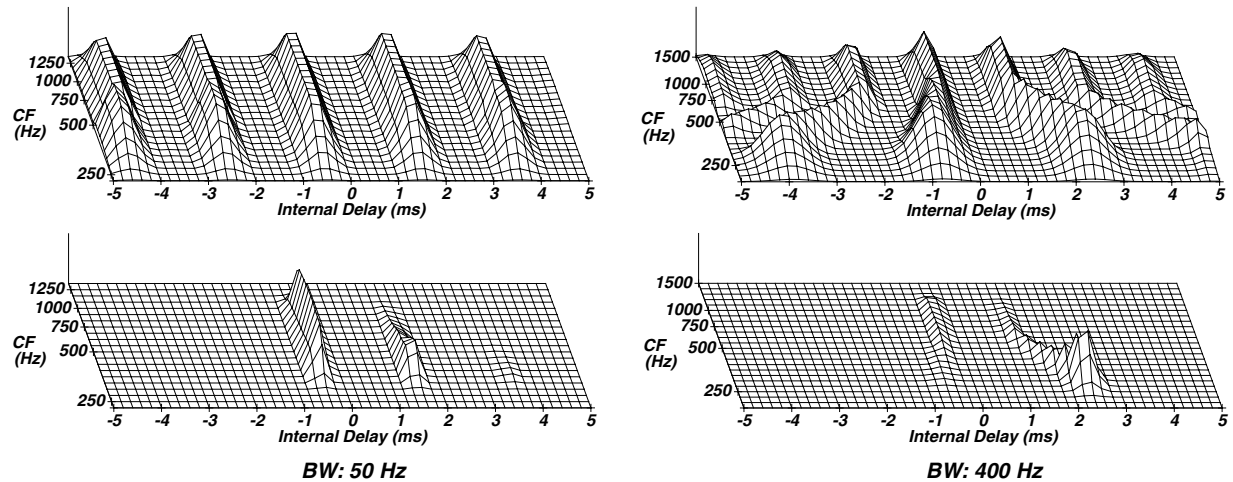
Figure 4-4: CCFs of (0, 270) stimulus with a positive IID. The left and right panels show the CCFs for bandwidths of 50 Hz and 400 Hz, before and after weighting for straightness, $p(\tau_m, f_{cm})$, and $L_I(\tau_m)$.

4.2 Overcoming the Straightness/Centrality Conflict

We have found that the (1500, 0) data and the (0, 270) data exhibit differing sensitivities to bandwidth: the (1500, 0) data depend heavily on bandwidth, but the (0, 270) data do not. In terms of straightness weighting, this would imply that the lateralization of the (1500, 0) stimulus is highly dependent on the additional cues provided by straightness, while the (0, 270) stimulus is not. Hence it is not surprising that any modification to the model that is based strictly on adjusting the overall potency of the straightness weighting would always be beneficial to one set of the data and detrimental to the other.

Because the dependence of the data on bandwidth is different for the (1500, 0) and (0, 270) configurations, we attempted to ways of altering the intensity weighting so that it would have more effect on one stimulus configuration than the other. We examined several techniques to isolate and reduce the effect of this straight peak in the (0, 270) data. The main techniques we tried are listed in the sections below. Unfortunately, none of these techniques proved successful, so we began to investigate other methods of introducing the weighting of IID without producing an adverse interaction with straightness weighting.

4.2.1 Narrowing the Width of the Intensity-Weighting Pulse

Recall that the intensity-weighting function proposed by Stern and Colburn (1978) is a Gaussian function with the form $L_I(\tau) \sim N(m_I(\alpha), w_I)$. The location of the pulse, $m_I(\alpha)$, is assumed to be a function of intensity difference: As the IID increases, the intensity function becomes situated further and further from the midline position. (For each IID α we fit the mean of $L_I(\tau_m)$ to produce a normalized predicted position of α dB for a tonal stimulus with no ITD or IPD.) This form of the intensity function uses a constant width w_I, which controls the spread of the curve from its mean position. Our first approach was to reduce the width of this pulse. A narrower pulse would effectively dampen out the peak at -500 μs for the (0, 270) experiment, reducing its unwanted "pull" for positive IIDs. The main problem with this approach was that the locations of the (narrow) intensity pulses become located at unreasonably large values of internal delay as the width becomes more narrow.

4.2.2 Changing the Shape of the Intensity-Weighting Pulse

The second change we made was in the shape of the intensity-weighting $L_I(\tau_m)$. The Gaussian function was originally chosen because it is a relatively simple, smooth, closed-form function. We considered the possibility that the infinite tails of the Gaussian may be emphasizing too much of the straight peak in the (0, 270) data. As an alternative method of damping out the unwanted straight trajectory in the (0, 270) CCF, we considered the finite-duration raised-cosine function of the form

$$
L_I(\tau) \;=\; \left\{ \begin{array}{cc} \dfrac{1}{2}\left[1 - \cos\!\left(\dfrac{2\pi(\tau - m_I(\alpha) - w_I)}{2w_I}\right)\right], & |\tau| < w_I \\[2ex] 0, & \text{O.W.} \end{array} \right.
\tag{4.1}
$$

where $m_I(\alpha)$ is the center, and w_I is the width of the raised-cosine function. This type of intensity function is zero for magnitudes of internal delay τ that are larger than the size of the window; i.e. it has finite duration.

Using the raised-cosine pulse resulted in position predictions for the (0, 270) case that resembled the data somewhat more closely. Specifically, the lateralization curves for different bandwidths were not as spread as before because the finite-duration $L_I(\tau_m)$ served to de-emphasize the contribution of the straight peak at -500 µs in the (0, 270) CCF. However, predictions for the (1500, 0) stimuli described the lateralization data more poorly because the straight peak at 1500 µs was similarly deemphasized.

4.2.3 Making Width of the Intensity-Weighting Pulse Dependent on Location

The final significant change was making the width of the function inversely proportional to IID. The two motivations behind this change were as follows:

1. In the (0, 270) condition the predictions diverge more and more from the data for large IIDs. As the IID grows larger, the location of the weighting function moves further along the time axis. The narrowband predictions move to the right side of the head in correspondence to the IID, but the wideband predictions get pulled to the left side due to the straight peak on the left-side of the CCF. Using a variable-width intensity function would affect the predictions when they are the most incorrect.

2. The predicted position of Domnitz and Colburn experiment of lateralizing tones according to ITD and IID does not narrow in range for large IIDs. Using an intensity function that narrows as IID increases should restrict the range of predictions, because in the limiting case of modeling a huge IID by an extremely narrow pulse, the location of that pulse would determine the predicted position independent of the ITD and shape of the CCF.

We tried many different techniques to fit the locations and widths of the intensity weighting function. Unfortunately, in all cases the variable-width function became too narrow and we had the same problem when we just tried narrowing the pulse for all IIDs, i.e. the means (locations) of the functions were located at unreasonably large values along the internal-delay axis. In the instances when we fixed the means and fitted the widths of the intensity function, we ended up with some very narrow widths that caused predictions became very erratic and anomalous. For example, the cue-reversal points of Domnitz and Colburn shifted radically.

4.2.4 Decreasing the Region of Calculation

Figure 2-3 showed some sample CCFs over the entire doubly-active region. Those CCFs showed that each trajectory of peaks has a straight "tail" associated with it; these tails are caused by the bandpass filters with large or small CFs picking up activity from the fringe bandwidth components of the stimulus. We thought that perhaps these straight remnants were adversely affecting the data we wished to predict, so we applied a fixed region over which the CCF was calculated. Using the fixed region effectively truncated the straight tails, but had no significant improvement on the predictions. Possible reasons why little change occurred are (1) the tails in the CCF are present in all stimuli, and this may have equalized any individual, stimulus-specific effects; and (2) the amplitudes—hence, the effect—of the tails decrease by moving away from the active region. The presence of the tails, however, was a more peripheral problem than the one caused by the interaction of the straightness and the intensity weighting functions.

4.2.5 Reversing the Weighting Operations

Currently, we apply weighting of straightness before weighting of intensity differences. The order in which we apply these operations does not affect the results of predictions. For example, the output of a straightness unit is as follows:

$$\prod_{i = -(N-1)/2}^{i = (N-1)/2} E\{L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} \tag{4.2}$$

The output of a straightness unit after weighting by the intensity function $L_I(\tau_m)$ is as follows:

$$L_I(\tau_m) \prod_{i = -(N-1)/2}^{i = (N-1)/2} E\{L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} \tag{4.3}$$

Interchanging the order of straightness and intensity weighting results in a function

$$\prod_{i = -(N-1)/2}^{i = (N-1)/2} L_I(\tau_m) E\{L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} = [L_I(\tau_m)]^N \prod_{i = -(N-1)/2}^{i = (N-1)/2} E\{L_m(\tau_m, f_{c_m} + i \cdot \Delta f_s)\} \tag{4.4}$$

Thus, reversing the two operations results in changing the shape of the intensity-weighting function (which was described in the sections above), but it does not affect the fundamental problem.

4.3 Summary

In summary, we found that various combinations of the current straightness and intensity parameters could not properly predict the Buell and Trahiotis (1991) experiment. We concluded that because the current method of multiplicatively-weighting an intensity function on top of the CCF cannot simultaneously describe the straightness and centrality requirements required by the various conditions of the Buell and Trahiotis data. The next chapter describes a different form of intensity weighting: Rather than multiply an intensity function to the CCF to account for a difference of intensity, we instead take the centroid computed without any intensity weighting and *add* an offset proportional to the intensity difference. The model that uses the latter method successfully predicts the Buell and Trahiotis experiment. However, it cannot by itself describe the cue-reversal points observed in the Domnitz and Colburn experiment.

Chapter 5
Implementation of the Model Using
Additive Intensity Weighting

The results of Chapter 4 showed that the interaction of the forms of straightness and intensity weighting used in the current form of the model cannot describe the Buell *et al.* data. In this chapter we consider an alternative form of intensity weighting, which enables the model to describe the Buell *et al.* data for all values of IID. In Section 5.1. we discuss the rationale behind using a model with an *additive* form of intensity weighting. In Section 5.2. we present predictions using this type of model. Although the type of time-intensity interaction described in this chapter enables the model to describe the lateralization data of Buell *et al.* (1991), it still appears to be unable to describe some aspects of the cue-reversal phenomena reported by Domnitz and Colburn (1977).

5.1 Alternative Form of Intensity Weighting

The inability of the position-variable model to describe the joint dependence of the lateralization of band-pass noise on bandwidth, ITD, IID, and IPD motivated us to consider other ways of combining interaural timing and intensity information. One promising alternative to the multiplicative time-intensity interaction of the original model is to weight *additively* the centroid of the timing function $L_T(\tau_m)$ by an amount proportional to the IID in dB. Figure 5-1 compares block diagrams of the original and the proposed method of weighting intensity differences.
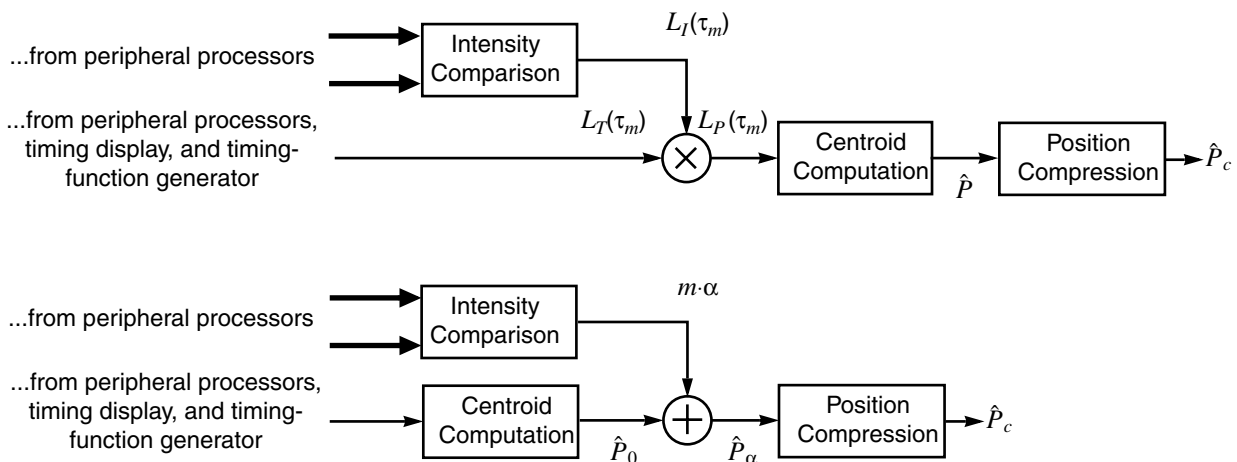
Figure 5-1: Block diagrams comparing original method of intensity weighting (top panel) to alternative mechanism (bottom panel) for weighting differences of intensity in the centroid computation. In the lower panel, α is the intensity difference, and m is the position/IID slope.

The top panel in Figure 5-1 summarizes the steps for multiplicatively applying the intensity function to the timing function (the current method of weighting IIDs). The bottom figure summarizes the proposed additive method of displacing the position-variable by an amount proportional to the IID. The centroid is calcu-

lated directly from the timing function, and a displacement that depends on IID is added to the centroid. The following equation shows show the additive model calculates position:

$$\hat{P}_c \;=\; C\{P_\alpha\} \;=\; C\{\hat{P}_0 + m \cdot \alpha\} \tag{5.1}$$

where $C\{\cdot\}$ is the compression function, α is the IID in dB, $\hat{P}_a$ is the predicted position given the α-dB IID, and m is the position displacement per IID. In one sense, this position function linearly extrapolates nonzero-IID positions from the positions with 0-dB IID. The extrapolation factor is proportional to the IID, hence the term $m\alpha$. A compression function at the output allows a non-linear shaping of the output positions.

We chose this specific ordering of operations after reconsidering the experimental data in Figure 4-1, which show that the dependence of lateral position on bandwidth for each stimulus configuration remains relatively constant for most small and moderate IIDs. To describe the changes in spacing between predictions for larger intensity differences, we apply a nonlinear limiting function to the output of the model. This function models the saturation of the position range as IID increases in this experiment, and a properly chosen function could describe possible non-linear behavior in other psychoacoustic tasks.

We now describe the results of the predictions obtained with the position-variable model using the additive-intensity-based mechanism. In this chapter the we refer to the position-variable model using an additive-form of intensity weighting as the additive model, and we call the model using the original type of intensity weighting the multiplicative model.

5.2 Predictions Using the Additive Position-Variable Model

The sections below describe the results from predictions for the Buell and Trahiotis, Domnitz and Colburn, and Schiano *et al.* experiments.

5.2.1 Predictions for the Buell and Trahiotis Experiment

Predictions for each of the stimulus conditions of the BBT experiment are shown in Figure 5-2. For straightness parameters, we chose an order N=11 and a ratio r=3. Appendix A outlines the procedure for choosing parameters for the straightness, additive-weighting, and compression functions to generate predictions.

These results describe the actual data shown in Figure 4-1 much better than the results of the multiplicative model in Figure 4-2, because the lateral positions now exhibit the same dependence on bandwidth at all small and moderate IIDs that they exhibit at zero IID. The (1500, 0) predicted positions show a significant dependence on the bandwidth; the predictions at 0 dB IID were obtained by using parameters that produced a relatively strong emphasis on straightness in order to have the necessary laterality of the wide-bandwidth stimuli. Using the same straightness conditions under the (0, 270) stimulus condition, we obtain little bandwidth dependence at 0 dB and, consequently, for all other predictions, for the reasons discussed in the previous chapter.

The straightness parameters used to generate the predictions in Figure 5-2 can be interpreted as follows: Each straightness unit takes as input 11 first-level counting units, and the CFs of those first-level inputs are spaced apart 3 times the spacing of the first-level units (which are uniformly distributed over log-frequency). Since such a large number of first-level input units may not be physiologically plausible, we generated
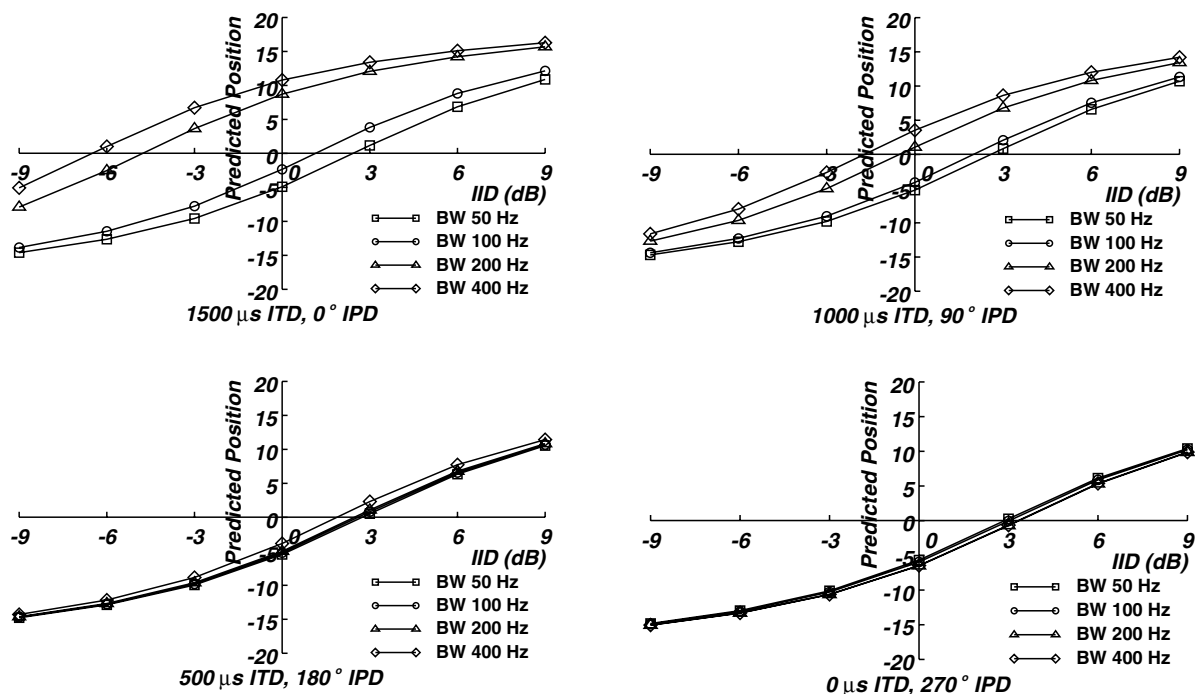
Figure 5-2: Lateralization predictions using the additive model, for the data of Buell and Trahiotis (1991). The predictions were obtained using straightness parameters of order N=11 and a ratio r=3. The stimulus combinations are listed below each figure. Moving counterclockwise from the upper-left panel, the ITD increases and the IPD decreases. Bandwidths are shown by the different symbols. The corresponding data are in presented in Figure 4-1 on page 24.

another set of predictions smaller value of the parameter specifying the straightness order, increasing the straightness ratio until we achieved the same degree of straightness as before. These results are shown in Figure 5-2. This combination of model parameters appears to produce underestimates of lateral position for bandwidths of 100 and 200 Hz. These results show that changing the straightness parameters can affect the CCFs in subtle ways, causing isolated effects on the predictions.

We note that the (1500, 0) prediction of the 100-Hz bandwidth stimulus at 0 dB, in general, does not agree with the data for either set of model parameters. The data show a position near the midline, but the model consistently predicts a position to the left of the midline, and this position is similar to that given for a 50-Hz bandwidth. The reason the two predicted positions are quite close because the $p(\tau_m, f_{cm})$ function greatly diminishes the difference between the CCFs at 50- and 100-Hz bandwidths. The wider bandwidth has trajectories that are slightly more curved, but this information is suppressed by the narrowness of the distribution function, which was modified to describe the results of Schiano *et al.* to further emphasize information at more central delays (cf. Section 3.4.2 on page 24). However, the 100-Hz data from individual subjects shows a relatively high *inter*subject standard deviation (of 5 to 6 dB), with subjects lateralizing on both sides near the midline position.
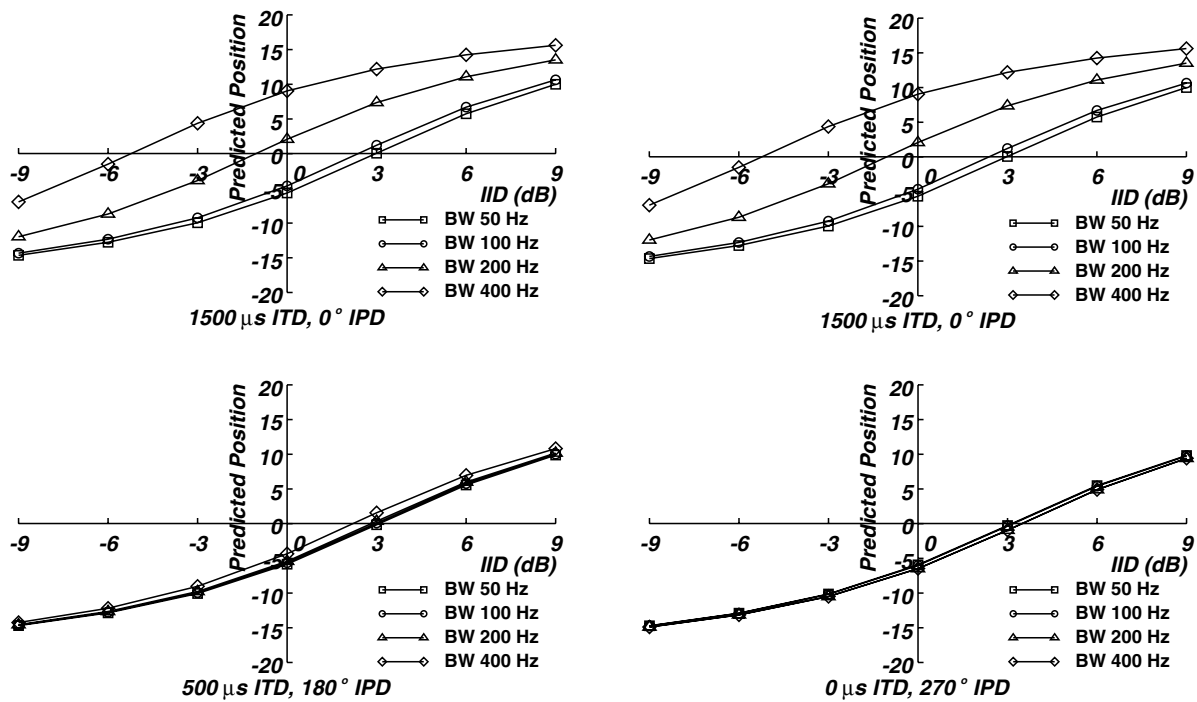
Figure 5-3: Predictions using the additive-model, of the data of Buell and Trahiotis (1991). The predictions were made using straightness parameters of order *N*=5 and a ratio *r*=10. The stimulus combinations are listed below each figure. Moving counterclockwise from the upper-left panel, the ITD increases and the IPD decreases. Bandwidths are shown by the different symbols. The data are shown in Figure 4-1 on page 27.

5.2.2 Predictions for the Schiano *et al.* Experiment

We found that the additive model produced virtually the same predictions as the multiplicative model did for the Schiano *et al.* data. Since the Schiano experiment was performed with no intensity difference, the multiplicative model would apply to the timing function an intensity function centered at an internal delay of 0 ms, but we omit this function for additive-weighting. This change did not affect the original predictions because the product of the functions $p(\tau_m, f_{cm})$ and $L_I(\tau)|_{\text{IID}=0}$ is approximately $p(\tau_m, f_{cm})$, so removing the intensity weighting function has little effect on the predictions. Compressing the output also does not significantly alter the shape of the Schiano *et al.* experiment after normalization.

5.2.3 Predictions for the Domnitz and Colburn Experiment

Figure 5-4 shows the additive model's predictions for the Domnitz and Colburn experiment. We note the following points:

- The compression function does reduce the dependence of predictions on ITD for larger IIDs, as is seen in the data.
- The additive model cannot describe the locations of the cue-reversal points seen in the data because the curves are vertically-offset copies of the curve obtained using the 0-dB IID. The compression function only affects the vertical scaling of the curves.
- The data of Domnitz and Colburn exhibit excursions of greater than ±30 dB. The lateralization-compression function of the additive model was fitted to describe the data of Buell *et al.,* which appear to saturate at about ±15 dB. This means that a model that was fitted to describe the data of Buell *et al.* will be fundamentally unable to describe the excursions in the data of Domnitz and Colburn that are seen for large IIDs.
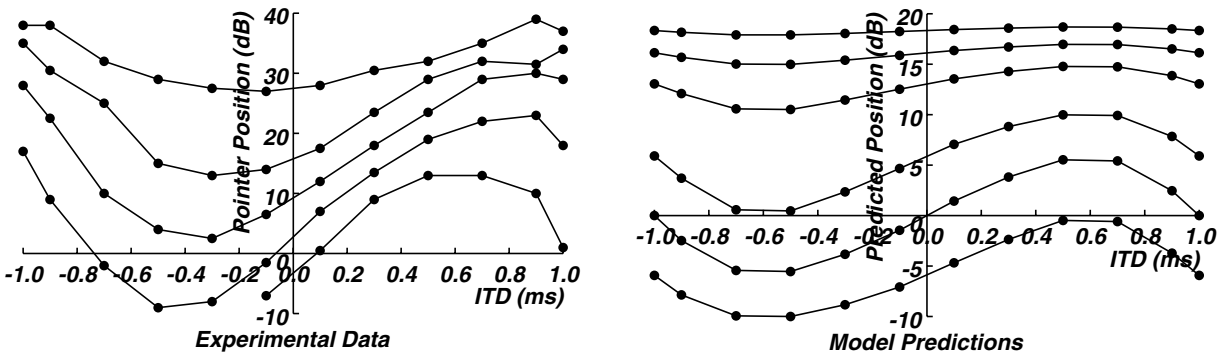

Figure 5-4: (Left panel) Experimental results of the Domnitz and Colburn experiment. (Right panel) Predictions using the additive-based model. The compression function parameters were set from the predictions of the data from Buell and Trahiotis.

We are concerned about the discrepancies between the predictions and the data. These differences (listed in the latter two items above) are probably at least in part a consequence of the differing conditions under which the data were obtained, but addressing these differences goes beyond the scope of this work. Preliminary results of reperforming the Domnitz and Colburn experiment using the experimental procedure of Buell and Trahiotis shows a much more subtle effect of the migration of cue-reversal points. In this report we form conclusions based on the Buell and Trahiotis data, which is more stable and salient than the cue-reversal phenomenon of Domnitz and Colburn. We will continue to explore the results and relationships between the two experiments.

5.3 Summary

The form of intensity weighting in the position-variable model was modified to better describe the data of Buell *et al.* Instead of multiplying a weighting function $L_I(\tau_m)$ by the timing function, we now calculate the centroid of the timing function directly, and add to the centroid a displacement proportional to the IID. This type of additive intensity weighting better describes the Buell *et al.* experiment. It also does not affect the multiplicative-based model's predictions for the Schiano *et al.* experiment, since those data were obtained with zero IID. However, using this type of intensity weighting inherently cannot model the migration of cue-reversal points seen in the data of Domnitz and Colburn.

Chapter 6
Summary

In Section 6.1 we summarize the major findings of this report, and in Section 6.2 we offer suggestions for future work.

6.1 Summary of Findings

The original formulation of the position-variable model of binaural interaction was able to predict successfully how the lateralization of pure tones depended on their ITD and IID (as seen in the data of Domnitz and Colburn, 1977), and frequency (as seen in the data of Schiano *et al.,* 1986).

In this work the position-variable model was extended to describe the "straightness" phenomena that characterizes the dependence of the lateralization of bandpass noise on ITD, IPD, and bandwidth [as seen in the data of Stern *et al.* (1988a), Trahiotis and Stern (1989), and Buell and Trahiotis (1991)]. Nevertheless, even the extended model is unable to describe the effect of the interaction of these stimulus parameters with IID (as seen in the data of Buell *et al.,* 1991), because of the type of multiplicative weighting of the effects of IID that had been a part of the model since its initial formulation (Stern and Colburn, 1978). In this work we propose a new additive central combination of the effects of ITD and IID that is better able to describe the data of the Buell and Trahiotis (1991) experiment. However, it does not predict the cue-reversal phenomena seen in the data of Domnitz and Colburn (1977).

6.2 Future Work

The most obvious extension to the findings presented is to redo the Domnitz and Colburn experiment using stimulus conditions that correspond to those used in the measurements of Buell and Trahiotis. Very preliminary results indicate that such data obtained under conditions more closely matching those of Buell and Trahiotis would also more closely be described by the predictions of the model. However, if this proves not to be the case, a model that implements a combination of additive and multiplicative intensity weighting my be desirable.

Because weighting the CCF for straightness tends to sharpen and narrow the peaks in the CCF, we needed a narrower $p(\tau_m, f_{c_m})$ function and a lowpass filter with a lower stopband to predict the results of the experiment of Schiano *et al.* (1986). We also need to reevaluate the results of binaural masking-level difference experiments (measuring the detectability of tones in noise) with these revised constraints.

It would also be worthwhile to determine why the position-variable model lateralizes bandpass noise with a 500-Hz center frequency and 50-Hz and 100-Hz bandwidths to almost the same position, which is contrary to the trends of the corresponding data. The problem stems from these two conditions having similar CCFs, which was further intensified by using a narrower $p(\tau_m, f_{c_m})$ function.

Appendix A
Generating Predictions Using the
Additive Model

PVModel is a computer program that implements the position-variable model. This chapter shows how to use the additive-IID weighting module to generate predictions for Buell and Trahiotis (1991). Section A.1 describes the procedure followed to make additive-model predictions for Buell and Trahiotis Section A.2 shows how to use the *PVModel* program to make predictions.

A.1 Choosing Weighting Parameters

Described in each section below is the procedure we followed to fit the free parameters in Buell and Trahiotis: straightness, intensity, and compression. The goal was to choose parameters such that applying a compression function onto the predictions would closely describe the Buell and Trahiotis (1991) data.

A.1.1 Straightness Parameters

We chose the straightness order N and ratio r to describe the (1500, 0) data at 0-dB IID and 50-Hz and 400-Hz bandwidths; these positions had the most bandwidth dependence of the data. The two positions roughly follow the relationship $\hat{P}_0\big|_{400\ \text{Hz}} = -2\hat{P}_0\big|_{50\ \text{Hz}}$, and we found that combinations of $N = 11$, $r = 3$ and of $N = 5$, $r = 10$ were able to describe this spread (anticipating slight compression of the predictions in the steps ahead).

A.1.2 Intensity Parameters

The model calculates positions for non-zero IIDs by using the following equation:

$$\hat{P}_\alpha = \frac{\hat{P}_0}{k} + m\alpha \tag{A.1}$$

where k is a scaling factor, and m is the slope parameter relating position to the IID. The variables k and m must be properly chosen in order for the compression function to accurately describe the data. We would like a plot of the data versus the predictions to be smooth, compact and unclustered, and as monotonic as possible. For example, Figure A-2 shows input-output plots using $k = 0.0009$, $m = 150$ and $k = 0.001$, $m = 180$; clustering shows up in the former case, but the latter, better choice of parameters, describes a sigmoidal characteristic that can be described using a nonlinear compression function.

A.1.3 Compression Parameters

The compression function has the form

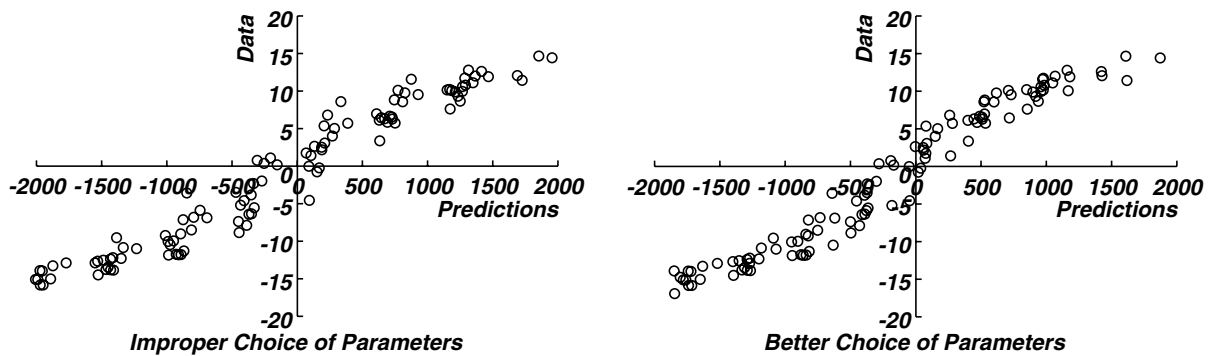$$\hat{P}_c = \alpha\tan^{-1}(\beta\hat{P}) \tag{A.2}$$

Figure A-1: Plots of data versus predictions. The left and right panels show incorrect and better choices of intensity parameters, respectively.

where $\hat{P}_c$ and $\hat{P}$ refer to compressed and uncompressed predictions, respectively, and α and β are user-chosen compression parameters. We computed α and β that would provide a best least-squares fit of the compressed positions to the data.

A.2 Using the *PVModel* Program

This section explains what settings to use in *PVModel* program to perform the additive-IID type of predictions for the Buell and Trahiotis (1991) experiment. At the time of this report, *PVModel* was implemented on the NeXT computer, and it is assumed that some familiarity of that computer and of the program is known.

1. Start the *PVModel* program.

2. Set the bounds of calculation as shown in Figure A-2. These settings on the **Region** panel choose calculation over the doubly-active region. The settings on the **CCF Parameters** panel choose the ranges of the binaural display: CFs are distributed uniformly over log-frequency every .01 Hz, and internal delays exist every .025 ms in the range ±5 ms. For the predictions in this report, we used the ±5 ms range after obtaining similar predictions over a larger (±12 ms) range.

3. Set the parameters of $p(\tau_m, f_{cm})$ and $G(f)$, the distribution and lowpass-filter functions, according to the appropriate panels in Figure A-2. The variables in the **Ptau Parameters** panel describe the shape of $p(\tau_m, f_{cm})$, and are explained in Shear (1987) and in the *PVModel* documentation. The variable $l_P = -1.23$ directly affects the narrowness of $p(\tau_m, f_{cm})$. The **Lowpass** panel sets the cutoff and stopband frequencies, f_c and f_s, of the lowpass filter $G(f)$ of the auditory-nerve model. In this example the values are f_c=1200 Hz and f_s=3200 Hz. :

4. Set the parameters for the position output and the straightness and intensity weighting routines as shown in Figure A-2. The current **Position Output** options select the centroid method for calculating the position variable, no compression of the output position, no emphasis on Raatgever dominant region, and a scaling of the output by a factor k=.0009: The **Straightness & Intensity** panel has the old method of straightness weighting selected, with order and step (ratio) parameters of 11 and .3 (3), respectively. The additive-IID
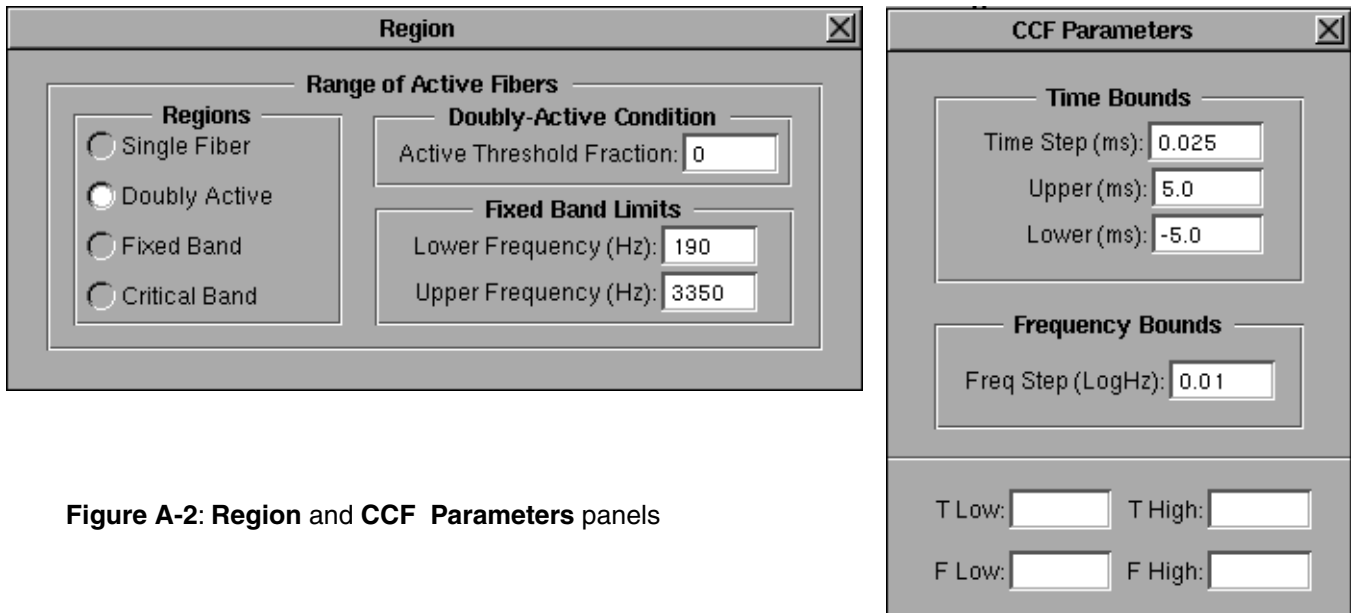
Figure A-2: **Region** and **CCF Parameters** panels

weighting is selected, so the Mulltiplicative-IID parameters are disregarded. The intensity-slope function m is chosen to be 150.
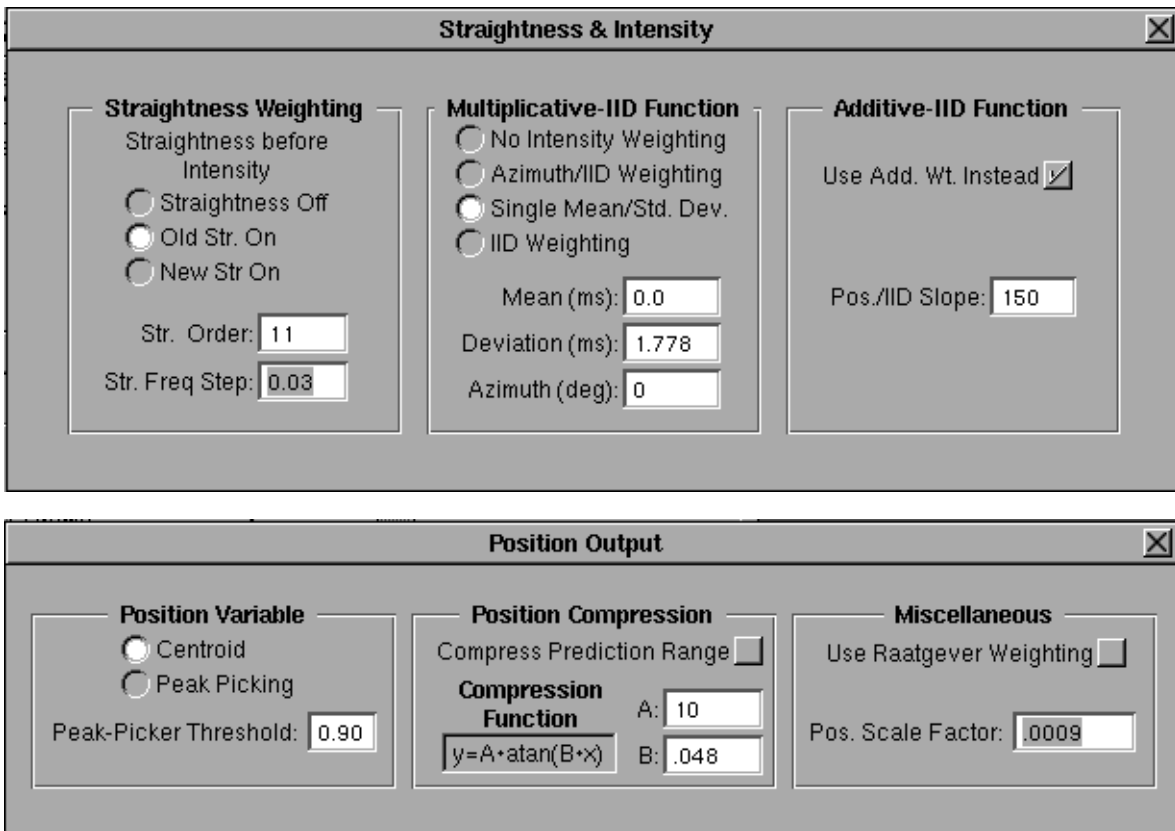
Figure A-4: **Straightness & Intensity** and **Position Output** panels.

Ptau Parameters ☒

IF (Hz): `0.1`

IP : `-1.23`

hF (Hz): `3000.0`

hP : `0.0`

MuF: `0.2e-3`

MuP: `0.0`

F Thresh (Hz): `1200.0`

Lowpass ☒

Parameters

fc (Hz): `1200.0`

fs (Hz): `3200.0`
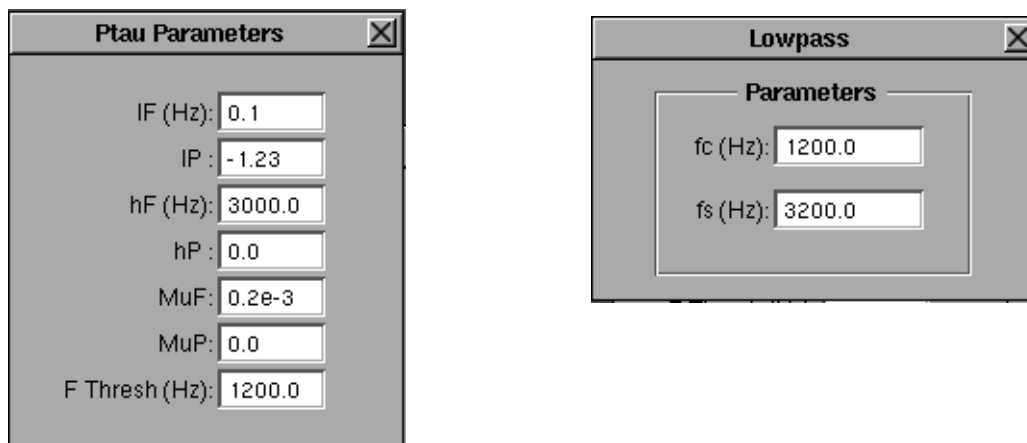
Figure A-3: **Ptau Parameters** and **Lowpass** panels

5. Figure A-2 shows the parameter settings used in this example, and set in the **Noise** panel: a 500-Hz bandpass noise with 400-Hz bandwidth at 70-dB SPL, an ITD of 1500 μs, and an IPD of 0 deg. Since we are using the additive-form of IID weighting, the value in the IID form does not matter.

6. Click the Lateralize button. As illustrated in Figure A-2, predictions at IIDs of 0, ±3, and ±9dB will be shown in the **Text** panel.

```
Noise                                    ⊠          □            Text Window                         ⊠

              Parameters                                  ;text window enabled...
                                                          ;ts=1.50 ps=0.0 fc=500.0 bw=400.0 iid=0.0 SPL=70.0
    ts (ms): 1.5        inc (ms): 0.0                     ;order=11 str_fstep=0.0300
                                                          ;x = iid
   ps (deg): 0         inc (deg): 0.0                     ;P/=0.000900; mI=0.000, wI=1.778
                                                                   -9        -377.368382
    fc (Hz): 500        inc (Hz): 0.0                              -6          72.631618
                                                                   -3         522.631618
   BW (Hz): 400         inc (Hz): 0.0                               0         972.631618
                                                                    3        1422.631618
   IID (dB): 0.0        inc (dB): 0.0                               6        1872.631618
                                                                    9        2322.631618
  SPL (dB): 70.0

           Number of Iterations: 1

   Calc CCF                   Lateralize                       Save Name: text.svd

  Messages:
```
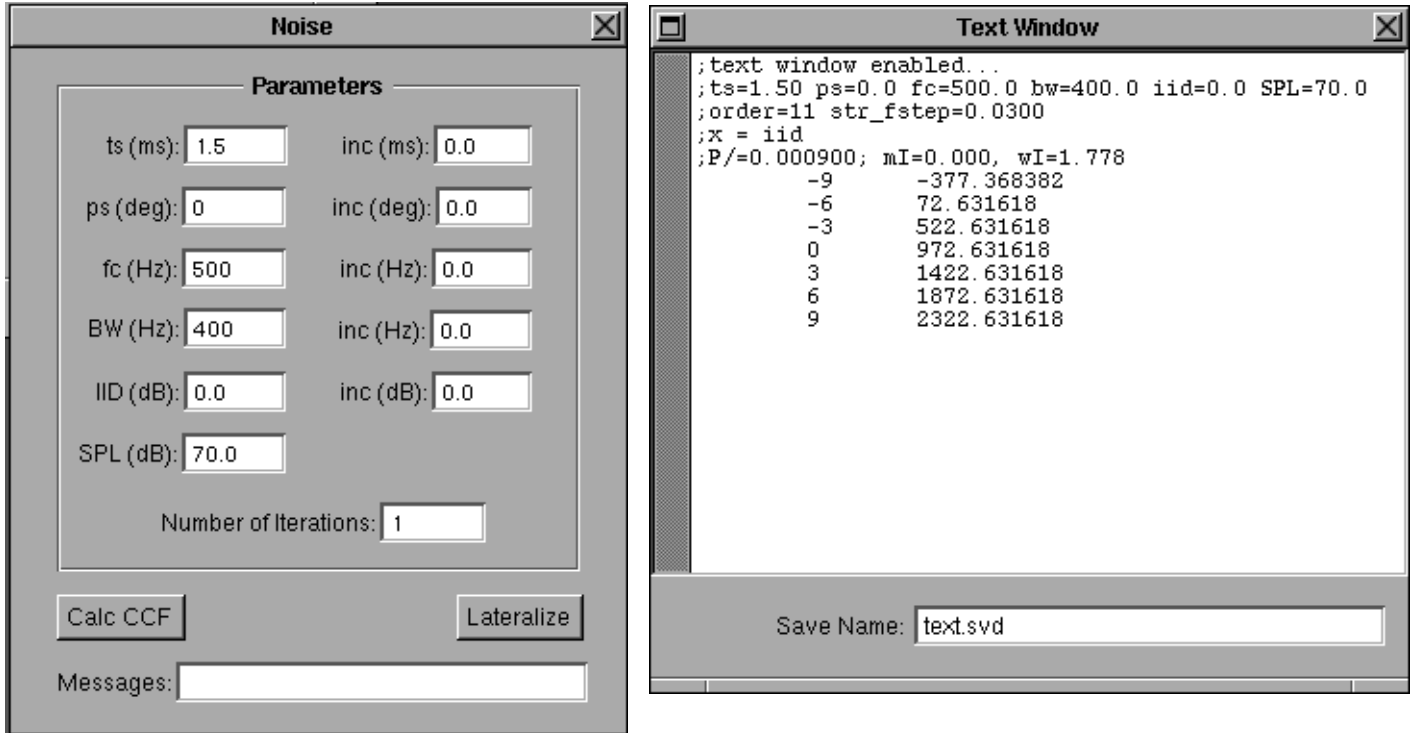
Figure A-5: **Noise** and **Text** panels.

References

Bilsen, F.A. (1976). Pronounced Binaural Pitch Phenomenon. *J. Acoust. Soc. Am.* 59, 467–468.

Buell, T.N. and Trahiotis, C. (1991). Potency of Interaural Differences of Intensity on the Lateralization of Interaurally Delayed Bands of Noise, *J. Acoust. Soc. Am.* 90, 2266.

Cherry, C. (1961). Two Ears—But One World. In W. A. Rosenblith(Ed.), *Sensory Communication*. Cambridge, MA: MIT Press.

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. (1977a). Theory of Binaural Interaction Based on Auditory-Nerve Data. II. Detection of Tones in Noise, *J. Acoust. Soc. Am.* 61, 525-533.

Colburn, H.S. (1977b). Theory of Binaural Interaction Based on Auditory-Nerve Data. II. Detection of Tones in Noise. Supplementary Material. *AIP Document No. PAPS JASMA-61-525-98*, 61, 1–81.

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

Green, D.M. (1976). *An Introduction to Hearing*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Schiano, J.L. , Trahiotis, C., and Bernstein, L.R. (1986). Lateralization of Low-frequency Tones and Narrow Bands of Noise. *J. Acoust. Soc. Am.* 79, 1563–1570.

Raatgever, J. and Bilsen, F.A. (1986). A Central Spectrum Theory of Binaural Processing. Evidence from Dichotic Pitch. *J. Acoust. Soc. Am.* 80, 429–441.

Raatgever, J. (1980). On the Binaural Processing of Stimuli with Different Interaural Phase Relations. Doctoral dissertation, Technische Hogeschool Delft.

Shear, G.D. (1987). *Modeling the Dependence of Auditory Lateralization on Frequency and Bandwidth*. Master's thesis, Elec. and Comp. Eng. Dept., CMU.

Siebert, W.M. (1968). Stimulus transformations in the peripheral auditory system. In P. Kolers and M. Eden (Ed.), *Recognizing Patterns*. Cambridge, MA: MIT Press.

Stern, R.M. (1976). *Lateralization, Discrimination, and Detection of Binaural Pure Tones*. Ph.D. thesis, Elec. Eng. Dept., MIT, December 1976.

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

Stern, R.M., Shear, G.D., and Zeppenfeld, T. (1988b). Lateralization Predictions for High-Frequency Binaural Stimuli. *J. Acoust. Soc. Am.* 84, S60 (A).

Stern, R.M., Xu, X., and Tao, S. (1991a) A Coincidence-Based Model That Describes "Straightness" Weighting in Binaural Perception. 14th Midwinter Research Meeting, Association for Research in Otolaryngology.

Stern, R.M., Zeppenfeld, T., and Shear, G.D. (1991b). "Lateralization of Rectangularly Modulated Noise: Explanations for Counterintuitive Reversals," *J. Acoust. Soc. Am.* 90, 1908–1917.

Trahiotis, C. and Stern, R.M., Jr. (1989). "Lateralization of Bands of Noise: Effects of Bandwidth and Differences of Interaural Time and Phase," *J. Acoust. Soc. Am.* 86, 1285–1293.

Yost, W.A. and Nielsen, D.W. (1977). *Fundamentals of Hearing*. Holt, Rinehart and Winston.