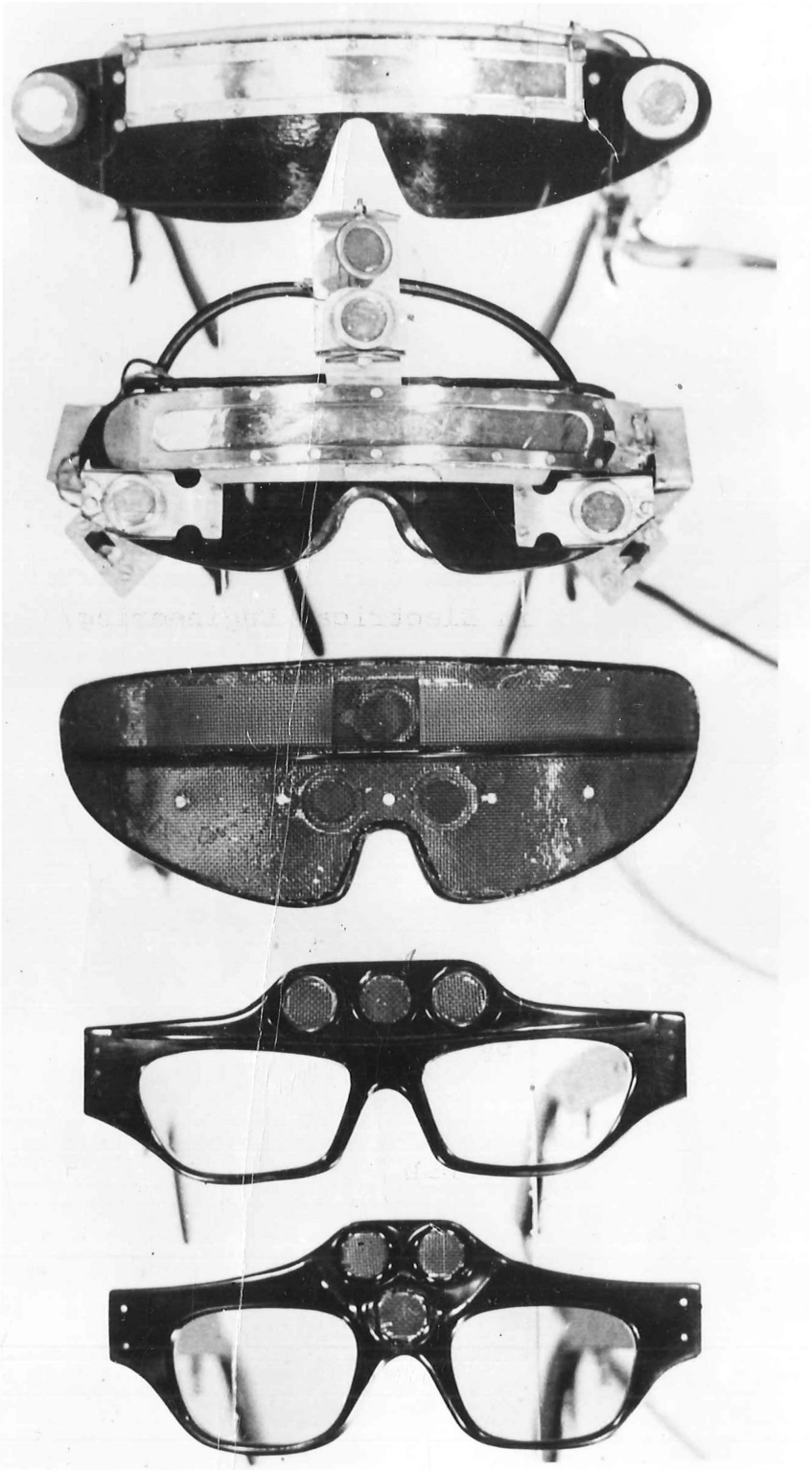
AUDITORY DISPLAY OF SPATIAL INFORMATION

A thesis presented for the degree of
Doctor of Philosophy in Electrical Engineering
in the University of Canterbury,
Christchurch, New Zealand.

by

D. ROWELL

1970



FRONTISPIECE: Evolution of a Mobility Aid for the Blind

"Perhaps for the first time, since the need for vision, I have become aware of the relationship between the environment and myself. A relationship characterized by the "Self-Concept" in which I have more certainty in realising the distance, speed and grandeur of objects, moving or otherwise. Due to this factor, I find myself concentrating more on my own physical plane when moving through crowds. Previously, it was necessary to concentrate on the position of people and what they might do. But with the use of the USG*, I can obtain this information well in advance and therefore, am able to consider which for me would be the most suitable mode of bypassing as it were, gracefully."

7th October, 1969.

from the daily training notes of I. Pivac, a blind university student, during training with the binaural mobility aid described in this thesis.

*ultra-sonic glasses

ABSTRACT

The problem of designing an auditory display for a binaural mobility aid for the blind is investigated. The basis for the project was the binaural linear F.M. echolocation mobility aid proposed by Prof. L. Kay. The sciences of engineering and psychophysics are freely intermixed to define the optimal form of the azimuthal dimension of an auditory spatial display.

It is shown that to take advantage of the frequency domain range coding it is necessary to use interaural amplitude difference (IAD) as the localization cue. A series of experiments confirmed this and suggested a modification to the system concept.

An extensive psychophysical investigation of auditory localization with IAD determined the parametric form of the localization function, from which the azimuthal dimension of the display was specified and analysed. The problem of generating this display from an echolocation system is discussed.

Measures of the resolution capability of the auditory system in a static two object environment are presented. A computer simulation of the auditory display was used to justify the use of data derived from dichotic experiments to specify a display for use with head movement. Adaptive strategies for estimating the localization function are described.

ACKNOWLEDGEMENTS

I gratefully acknowledge the support of my supervisor, Professor L. Kay, upon whose original idea this thesis is based.

I wish to thank Mr L. Elliott, who as a research assistant conducted the experimental sessions in Section 4.2; and I am grateful to all the subjects who gave of their time to participate in the psychophysical experiments.

The project was significantly aided by the technological expertise of Mr G.R.S. Clark and Mr J. Jongens who constructed much of the experimental apparatus.

I am appreciative of many helpful discussions with fellow postgraduate students and members of the University staff.

Financial support, in the form of Scholarships and Grants, during this project was received from the New Zealand Department of Education, the New Zealand Shipping Company Ltd, the New Zealand University Grants Committee, and the National Research and Development Corporation (England).

Finally, I am grateful to my wife for her forbearance and understanding throughout the project, and for her help during the preparation of this thesis.

TABLE OF CONTENTS

C H A P T E R 1

INTRODUCTION

1.1 The Nature of the Problem

The advent of modern technology, particularly the development of solid-state circuit elements, brought with it many proposals of electronic devices to aid the mobility problems of the blind[1,2,3]. Many of these devices were tried but then rejected by the blind population because they allowed only a marginal improvement in mobility. The emphasis in the development of many of the early devices was placed upon the engineering and technology of the aid itself, with little thought to the psychophysics of the mobility function or the sensory modality used for the display. Unfortunately, the state of the art in the science of mobility aids is still such that the designer must resort to intelligent try-it-and-see methods, with both laboratory and field evaluation of proto-type aids. Nevertheless the need for greater understanding of sensory processes and percept formation is now recognized as essential for a successful visual prosthesis in the human mobility function, and much current research is directed toward an understanding of the non-visual modalities.

It has commonly been stated[e.g. 4,5] that the engineer should take care in his design not to confuse the blind user with too much information. This type of reasoning has led to the proposal of many devices with simple binary state displays

of the 'stop-go' type, usually with tactile or auditory communication[e.g. 6,7]. However, the definition of information is "that which increases the organization of a system", which implies that too much information cannot, in fact, be given to the blind person. The informational content of the sensory display must be defined relative to the human observer, not an idealised signal analyzer. Careful consideration must be given to the system design to ensure that the signals at the display are matched to the behavioural characteristics of the sensory modality chosen. The dimensions of a display must be chosen with regard to absolute identification of stimulus magnitude, differential discrimination sensitivity of each dimension and masking effects in a complex stimulus situation.

The project described in this thesis was concerned with deriving the characteristics of an optimal binaural display of information from a two channel echolocation system. The basis of the research was the binaural ultrasonic mobility aid as proposed by Kay[8,9,10,11,12] and which is covered by British Patents Nos. 978471, 978472 and 3366922. This device, and its monaural counterpart the Kay-Ultra 'torch', differs from the majority of previous mobility aids in that they are true environmental sensors, with no information reduction before the auditory display. In the research described here the sciences of psychophysics and engineering were freely combined in order to use the behavioural characteristics of

binaural audition as the basis for the system specification.

Although extensive literature exists on the psychophysics of binaural hearing there has been an emphasis in the past on qualitative descriptions of localization phenomena. It was therefore necessary to re-examine experimentally some of the well known aspects of localization reported by early investigators (circa 1920). In these experiments the emphasis was on deriving quantitative data to be used in an engineering design; with estimates of the spread of individual differences of the localization coefficients over the population. This data was used to define the two channel echolocation system. Other experiments were designed to estimate the resolution capabilities of the auditory display in a static environment. Simulation studies using an idealized computer generated display were used to estimate the localization function in conditions of free head movement. In these psychophysical experiments extensive use was made of on-line computing techniques, using both analogue and hybrid computers. The particular forms of stimulus and response processing used are believed to be new and could be adapted for use in psychophysical studies related to other modalities.

1.2 Project History

The concept of a binaural mobility aid using a linearly frequency modulated echolocation waveform was developed by L. Kay in 1959, but active research on it was deferred for some years because of the system complexity. During the

intervening period research was concentrated upon a monaural hand-held version of the aid, which after preliminary evaluation proved to be promising enough for a commercial version to be produced by Ultra Electronics Ltd, England. Although it has been found that training is necessary to use this aid to its full advantage the 'torch' is in active use by blind people throughout the world and is without doubt the most successful of any electronic mobility aid to date.

In early papers[8,9] Kay hypothesized that the sonic navigation capabilities of some echolocating bats could be explained on similar principles to his proposed binaural mobility aid. In these papers the localization phenomena associated with such a display were stated to be dependent only upon propagation path length differences. In 1965[11] the additional use of the binaural intensity ratio, commonly denoted in logarithmic units as an interaural amplitude difference (IAD), was suggested. It is shown in this thesis that IAD is the only active localization cue in the display. Early experiments on a prototype system confirmed that there were localization phenomena associated with the display[12], but the first quantitative measures of these effects were carried out by the author (see Chapter 3).

Upon Professor Kay's arrival at the University of Canterbury in 1966 an extensive programme commenced with research based in three areas.

i) The requirements and control functions of human mobility

ii) An investigation of the binaural display of echolocation returns

iii) The technology of the device (ultrasonic transducers and electronic development).

The investigation into the difficult area of mobility did not reach fruition but the work of Martin[13], associated with the technology of the device, resulted in significantly improved prototype models that have been used in field evaluations. The second category of research is reported here.

A peripatologist, R.W. Pugh, was appointed to the programme in late 1968. Working with devices developed during the programme he has been able to show that the binaural aid can be of benefit to non-visual mobility.

1.3 The Role of the Mobility Aid

Little is known of human control functions in the environment. A working hypothesis on the role of the mobility aid was derived from a qualitative model which embodied the behavioural features of mobility. The model is shown in Fig. 1.1 and is briefly described here from an introspective viewpoint.

All actions of the mobile mechanism are defined as being based upon a mental image of the environment, derived from the perceptual processes and a-priori information. Mobility may be explained by the effect of a finite set of efferent motor actions upon the integrated set of re-afferent sensory

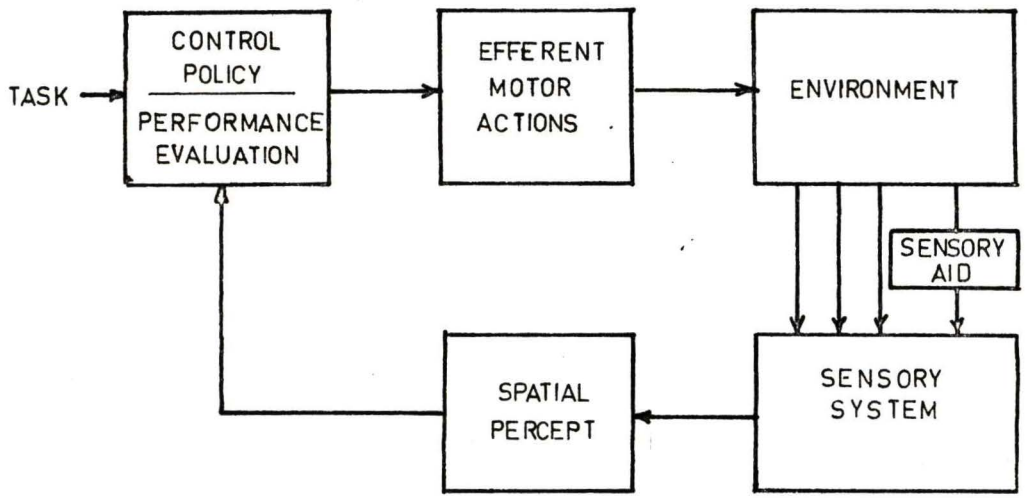


Fig. 1.1 Qualitative model showing the role of the sensory-motor feedback loop in mobility.

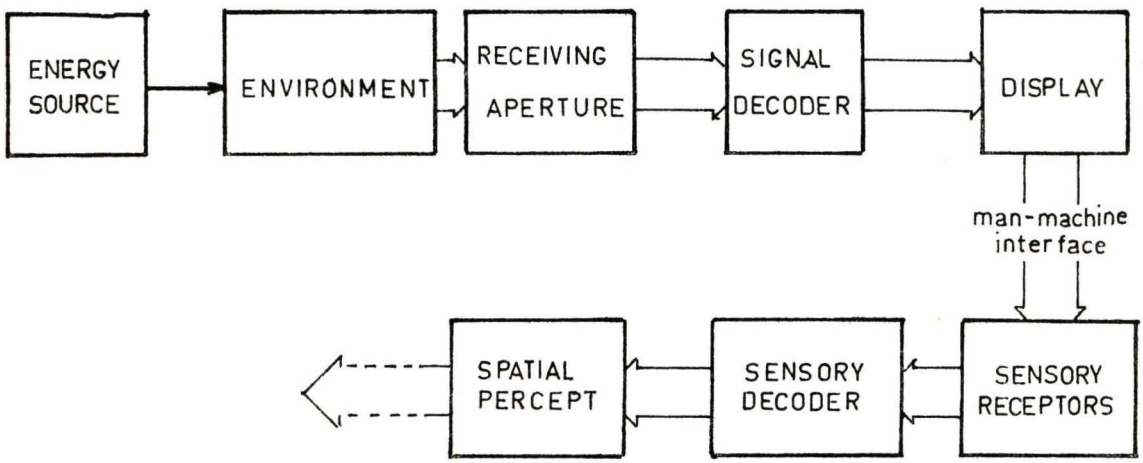


Fig. 1.2 Enhancement of perception with a sensory aid - the elements of the information processing.

stimuli across all modalities. The formulation and evaluation of a control policy in the execution of a task (or set of sub-tasks) is therefore dependent upon the efferent/re-afferent transformation of the sensory stimuli pattern by each action. Clearly the model resembles a basic feedback mechanism but it is defined to exhibit predictive and adaptive behaviour. The sensory stimuli need not be derived from a terrestrial environment; the role of a simulator in the training of complex control tasks is to substitute the re-afferent effects of a real environment by an artificial set so that new control policies may be developed.

The blind person's loss of re-afferent contact with the environment severely restricts the formulation and evaluation of control policies, and awkward movement results. Satisfaction, from interaction with the environment, and motivation are reduced. Furthermore the involuntary nodding, rocking or ungainly headmovements all classed as 'blindisms'[14] result directly from the severely limited efferent/re--afferent kinaesthetic feedback loop.

The role of the mobility aid is to restore the perception of spatial relations; that is to reconstruct upon some display the spatial nature of the immediate environment in a form that may be readily assimilated through some non-visual modality. The failure of the obstacle detector may be seen immediately. Although a measure of confidence may be

given the user, the basic problem has not been solved for the binary state display allows only two possible actions: "stop" or "go". The person using the obstacle detector is still forced to use his remaining senses to communicate with the environment.

The elements of the spatial perception process using a sensory aid are shown in Fig. 1.2. Objects in the environment perturb a spatial energy distribution. A receiving aperture is used to sample the perturbed field and after subsequent signal processing a multi-dimensional display is presented at the man-machine interface. Further sensory decoding is necessary to form a percept. Tanner[17], in a speculative paper, proposed that visual type of experience may be derived from stimulation of non-visual modalities. His arguments were based upon signal detection theory and he defined the requirements for spatial perception as a one-to-one mapping of the environment on to a non-visual sensory display. Optical imaging techniques have been used to generate a two dimensional tactile display[16] which allows the recognition of simple forms[15]. The present study is concerned with signal processing of the returns from an active echolocation system for coupling to human audition.

1.4 Thesis Organisation

In Chapter 2 characterization of the environment from echolation signals and the restrictions imposed upon the

display form by the auditory system, are discussed. It is shown that the display used in the FM aids is one of the few that will allow perception of multiple objects. In Chapter 3 a brief description of the prototype system is given and experimental studies of binaural localization phenomena associated with the display are described. The results of these experiments enabled the system to be defined and redundant binaural cues eliminated.

Chapter 4 is devoted to the psychophysics of the binaural localization of tones using IAD. Experiments are described which were designed to measure the effects of stimulus parameters upon the localization function; and to estimate the spread of individual differences in the blind and sighted populations.

In Chapter 5 the parametric form of the localization function is used to define the idealised display in engineering terms. In analyses of the performance of this display it was necessary to use simplified models of the azimuth estimation process. The emphasis in this particular chapter is upon properties of the display, rather than upon characteristics of the auditory localization process. The idealizations imposed in Chapter 5 are relaxed in Chapter 6 where the display of a practical system is discussed.

Chapter 7 describes psychophysical experiments designed to estimate the resolution capabilities of the auditory analyzer in a static multiple object environment.

In Chapter 8 a simulation of the display using an analogue computer is presented. The simulator is extended to include an adaptive strategy in an attempt to make an efficient estimate of the parameters of the localization function.

The results of the work described in this thesis are summarised in Chapter 9.

CHAPTER 2

AUDITORY DISPLAY OF ECHOLOCATION INFORMATION

2.1 Introduction

The waveform and signal processing to be used in an echolocation system is restricted by the behavioural characteristics of human audition. In this chapter it is shown that a temporal display of range, coded directly from propagation delay, will inhibit the perception of multiple objects within a range annulus of up to approximately 20 feet width. It is also shown that the two dimension (range x azimuth) display used in the FM mobility aids is one of the few that overcomes inhibition of all but the nearest object.

It is implicit in the following discussions that the echolocation medium is air-borne sonar. Although the analyses are general, the examples will assume that the velocity of propagation is that of sound in air, i.e. approximately 1130 ft/sec at room temperature.

The choice of sonar is dictated by both theoretical and practical constraints. The temporal resolution constant of Woodward[18] is a measure of the total signal ambiguity of a waveform and is proportional to the occupied bandwidth. This may be related directly to a similar range resolution constant by a factor proportional to the velocity of propagation. It can then be shown that spatial resolution in the immediate environment can be greatly enhanced by the use of sonar.

In fact, a sonar system with a bandwidth of 50 kHz is capable of the same range resolution as a radar system with a bandwidth of 45 GHz. Present day techniques will not allow a portable radar system with this bandwidth, but a bandwidth of 50 kHz of ultrasound in air is feasible with simple transistorized circuitry and electrostatic transducers.

2.2 Spatial Information in Wide-band Echolocation Signals

Classical radar signal analyses[18] concentrate upon the estimation of target range R and range rate $\frac{dR}{dt}$ from the propagation delay τ since

$$\tau = \frac{2R}{c} \qquad 2.1$$

where c is the velocity of propagation. The analyses assume a narrow fractional bandwidth for the transmitted waveform, so that no spatial information other than delay may be derived. The resolution in propagation delay is limited by the width of the autocorrelation function of the received waveform. Scanning techniques are normally used to reconstruct the complete target space.

With wide-band signals further spatial information is contained in the returns. Consider the environment to be a linear, passive time-invariant element. Then the environment is completely characterized by the impulse response $h_e(t)$ as measured at the transmitting and receiving apertures of the echolocation system. No information other than that contained in the impulse response may be derived from

observation of the received signal, but the structure of the back-scattered wide-band signals will contain further information that is ignored by the classical analyses. The received waveform r(t) will be given by convolution of the transmitted signal s(t) with the impulse response, i.e.

$$r(t) = \int_{-\infty}^{\infty} s(\sigma) h_e(t-\sigma)\, d\sigma \qquad 2.2$$

If the autocorrelation function of s(t) is a Dirac delta function $\delta(t)$ the impulse response may be immediately found by crosscorrelation of r(t) and s(t) because

$$R_{rs}(\tau) = \int_{-\infty}^{\infty} h_e(\sigma) R_{ss}(\tau-\sigma)\, d\sigma \qquad 2.3$$

or $$R_{rs}(\tau) = h_e(\tau) \qquad 2.4$$

by the sifting property of the delta function, where $R_{rs}(\tau)$ is the input-output cross correlation function and $R_{ss}(\tau) = \delta(\tau)$ is the autocorrelation function of the transmitted signal s(t).

Let the object space consist of n resolvable elements with angular coordinates specified by the vector θ_n, each with an impulse response $h_n(t)$ so that

$$h_e(t) = \sum_n h_n(t-\tau_n) \qquad 2.5$$

where τ_n is the propagation delay of the nth element. The narrow-band approximation assumes that $h_n(t) = \delta(t-\tau_n)$ for all n. Then if the transmitter/receiver aperture pair has an effective impulse response $h_a(t;\theta_n)$ which is uniquely defined by θ_n, the received signal will be

$$r(t) = s(t) \otimes \sum_n (h_n(t-\tau_n) \otimes h_a(t;\theta_n)) \qquad 2.6$$

where $\otimes$ denotes convolution, so that with a-priori knowledge of $h_n(t)$ and wide bandwidth $s(t)$, the coordinates θ_n may be found.

Cahlander[19] derived a similar model as a plausible explanation of monaural perception of elevation by echolocating bats. He defined the pinna (outer ear) as a 'gain-dispersive' antenna, and, with the implicit assumption that $h_n(t) = \delta(t)$, derived a signal function that is monotonically related to the probability density function of the best estimator of θ_n in noise.

Consider now a two channel system with two receiving apertures, denoted by the subscript $i = 1,2$ with impulse responses $h_{ai}(t;\theta)$. If $R_i(\omega)$ and $S(\omega)$ are the Fourier transforms of $r_i(t)$ and $s(t)$ respectively and $H_e(\omega)$ and $H_{ai}(\omega;\theta)$ are the Fourier transforms of $h_e(t)$ and $h_{ai}(t;\theta)$, then with suitable constraints upon $H_{ai}(\omega;\theta)$ a function $G(\omega;\theta)$ may be defined, e.g.

$$G(\omega;\theta) = \frac{|R_1(\omega)|}{|R_2(\omega)|}$$

$$\text{or} \quad G(\omega;\theta) = \frac{|H_{a1}(\omega;\theta)|}{|H_{a2}(\omega;\theta)|} \qquad 2.7$$

which uniquely defines θ. Note that for this estimator the assumption of wide bandwidth is not necessary. Even if $s(\omega)$ is a single frequency, θ may be specified in a single dimension.

It is interesting to note that a similar mechanism has been proposed to explain binaural perception of elevation in echolocating bats. Grinell and Grinell[20] suggested from neurophysiological studies that coding of this form was present in the responses of auditory units at the level of colliculus and that estimation of $G(\omega;\theta)$ from three spot frequencies should serve to locate an echo unambiguously in azimuth and elevation. A similar model to the one described above was developed by the author (Appendix 1) as a plausible explanation of the role of the pinnae in passive binaural sound localization. This is discussed further below.

2.3 Human Auditory Spatial Localization

The neural mechanisms and stimulus attributes that allow the perception of the spatial position of a sound source have yet to be quantitatively defined. It has been shown that the azimuth and elevation of a wide-band source can be determined binaurally, and to a limited degree monaurally, without head movement. This, together with the ability to discriminate front and back, cannot be explained by simple binaural differences.

It is commonly accepted that the primary localization cues, interaural time difference (ITD) and the binaural intensity ratio (denoted as IAD) interact to provide cues for the estimation of a source's azimuth. Recent studies have shown the pinnae to be essential for the perception of

elevation. Roffler and Butler[24] reported that wide bandwidth signals with components above 7 kHz are necessary for the perception of elevation. They found that vertical localization was no longer possible if the pinnae were rendered ineffective. Batteau[25] found that spatial perception of a remote environment with dichotic stimuli was possible only if the two microphones were fitted with replicas of the pinnae. It was found that the high frequency components were essential for true spatial localization and that subjects fitted with reversed pinnae reported inverted vertical localization and reversed front-back discrimination.

Batteau concluded that the spatial information was coded upon the incident sound field as delays that were recognized per se by the central nervous system, whereas Roffler and Butler suggested that a comparison of the spectral modifications at the two ears was used in the evaluation of the vertical coordinate. This latter hypothesis is essentially the same as the method of estimating the angular coordinates of a reflecting source outlined in Section 2.2. If the pinnae are defined as having transfer functions $H_{ai}(\omega;\theta)$ where i = 1,2 as before, and the sound source is described in the frequency domain by $F(\omega)$ then some monotonic function of

$$G(\omega;\theta) = \frac{F(\omega)H_{a1}(\omega;\theta)}{F(\omega)H_{a2}(\omega;\theta)} \qquad 2.8$$

will serve to define θ as in Eq.(2.7). In Appendix 1 a model is developed showing that, for certain classes of signal,

peripheral binaural interaction does produce coding of this form.

Dynamic localization with head movement provides secondary kineasthetic cues. Wallach[26] created an illusion in which the position of the source reacted in the opposite direction as the head was moved. In this condition front-back discrimination was reversed. However, Pollack and Rose[27] found that subjects' responses to free field stimuli were more variable when free head movement was allowed. This conflicts with the finding that mechanically induced head movements reduce localization errors[28].

It has been shown that the localization mechanism can adapt, over a period of several days, to distortions in temporal and intensive binaural cues[29,30]. It was found that active interaction with the environment was necessary for adaption and that this process could be accelerated by training. Freedman and Zacks[31] found that the ability to discriminate small time differences was impaired after exposure to dichotic white noise, only if subjects moved actively through the environment. These experiments show that the development and reorientation of the auditory perceptual space is dependent upon efferent/re-afferent sensory feedback, in accord with the mobility model (see Section 1.3).

There are some subjective localization phenomena that are not related to the binaural stimuli. The 'proximity'

effect, reported by Gardner[21], states that in the absence of auditory distance cues vision will dominate, and the apparent source position will be the nearest plausible location. Another subjective effect is found with tonal stimuli, in which case the apparent elevation may depend on the frequency, with high frequencies apparently more elevated than low frequencies[22,23].

The ability of many blind people to use active or passive echolocation, sometimes at a subliminal level[32,33], demonstrates that auditory stimulation can result in a non-auditory (perhaps better described as 'semi-visual') type of percept. The common experience of those who use 'facial vision' is reported as "feeling something like a cloud on the face"[14], which is evidence supporting the hypothesis of Tanner[17]. There is no evidence of distance perception or multiple object resolution with the natural sonar, although it is found that wideband signals are essential for its operation.

2.4 Dichotic Binaural Phenomena

Much of the behavioural research on auditory localization has been designed around experiments using dichotic stimuli. The common experience of listeners in this situation is that the image no longer has the spatial attributes of a diotic source but appears to originate from within the head. Under these conditions some subjects report no angular projection

of the image and are able to judge only the 'sidedness' of the intracranial sound. This experience has been termed lateralization as opposed to true angular localization of the image in space. Jeffress and Taylor[40], in differentiating between the two phenomena, found that the type of report that subjects gave was dependent upon the task set them.

The presence of a single image from binaural stimulation implies interaction at some level. Several researchers have shown the auditory system to be extremely tolerant to binaural differences before fusion is destroyed and a separate sound heard at each ear. Ebata, Sone and Nimura[41] showed that the maximum interaural frequency difference for tones within which fusion occurred was dependent upon the stimulus duration. Deatherage[42] used clicks of different frequency content to show that fusion could be maintained provided the frequency difference was not too great. Dichotic white noise stimuli will give a fused image even when the correlation between the sources is as low as 0.1[43]. Cherry and Taylor[44] reported that to destroy fusion, speech signals required an ITD of approximately 15 msecs, which is about 20 times that experienced in normal hearing.

Woodworth[45] used, as a simple model of the human head, a sphere of diameter 8.75 cm to calculate the ITD resulting from a source at an azimuth angle θ as

$$t = 254(\theta + \sin\theta)\ \mu\text{secs}$$

which has been shown to be in agreement with measurements

made at the outer ear[46]. Objective measurements of the binaural intensity ratio are not so well behaved. Feddersen, Sandel, Teas and Jeffress[46] using probe microphones found results that were highly variable between subjects, with pronounced maxima and minima as the source was moved about the head. Their results, averaged over five subjects, are shown in Fig. 2.1. The variability in the measurements left Feddersen et al at a loss to explain how intensity cues could be used in auditory localization. They concluded that because the same IAD can be measured at many azimuth angles, the cue of intensity must be ambiguous. The results of Firestone[47] lead to similar conclusions.

Neurophysiological studies have shown that the peripheral neurons cannot maintain synchrony with individual cycles of the stimulus at frequencies greater than approximately 1500 Hz. At this frequency, the wavelength of sound in air is approximately twice the interaural distance so that ambiguous estimates of the source azimuth would result from comparison of phase differences at higher frequencies. For tonal stimuli, it is generally accepted that localization is derived from phase differences for frequencies below 1500 Hz and intensity differences at higher frequencies. In an elegant set of experiments Mills[48] compared the dichotic difference limens with objective measures of the interaural disparities resulting when a diotic source was just noticeably displaced from the median plane. The results, shown in

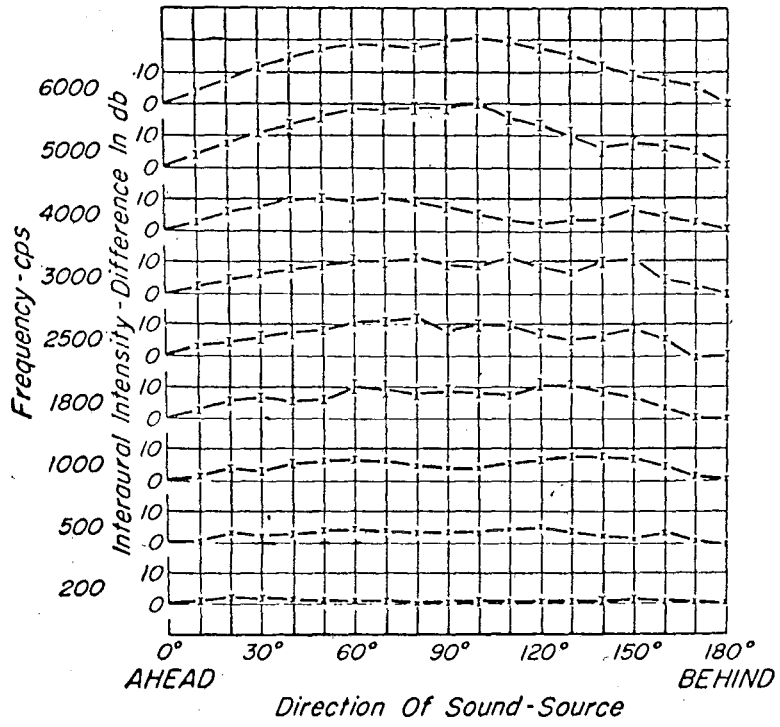


Fig. 2.1 IAD induced by the diffractive effects of the head and pinnae. The curves are the average of 5 subjects (From Feddersen, Sandel Teas and Jeffress, ref. 46).

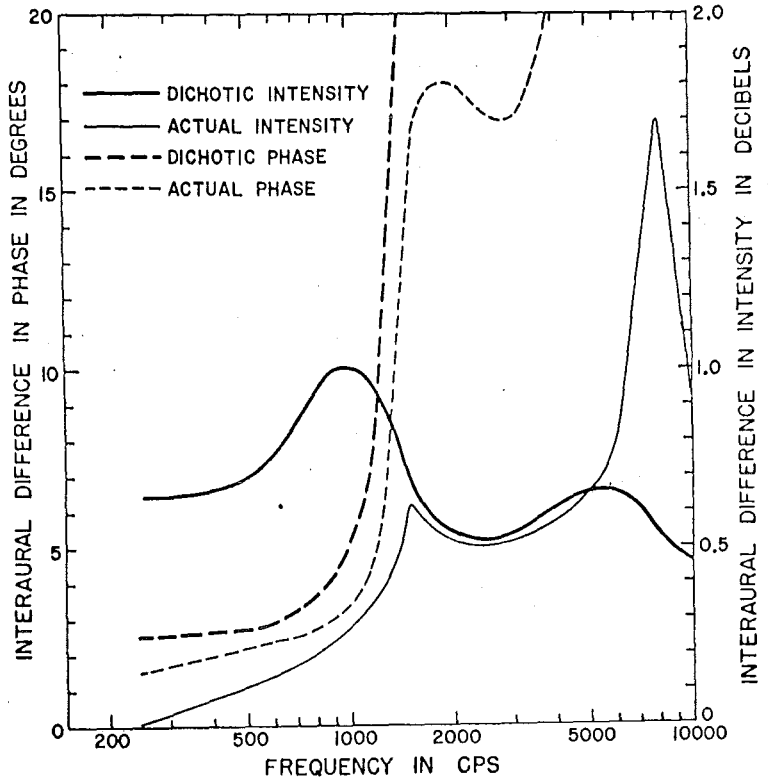


Fig. 2.2 Comparison of just noticeable dichotic interaural differences with the differences induced when a real sound source is just noticeably displaced from the median plane (From Mills ref. 48).

Fig. 2.2, suggest that the maximum frequency for which ITD is effective is 1500 Hz. Sandel, Teas, Feddersen and Jeffress[49] created an artificial situation in which the sound was loudest in the time lagging ear. Above 1500 Hz the image was localized toward the louder ear, while below this frequency the phase difference dominated. At approximately 1500 Hz listeners were confused.

Quantitative measures of binaural phenomena have shown that the stimulus parameters (intensity, frequency, bandwidth, duration, etc.) will affect results. Teas[50] found that the localization of a low frequency transient was more sensitive to ITD than that of a high frequency transient. Tobias and Zerlin[51] studied the effect of stimulus duration upon the difference limen for ITD and concluded that the localization mechanism included a temporal integration with a time constant of approximately 700 msecs. The dependence of the limen upon the form of the stimulus was studied by Klump and Eady[52] who estimated the limen as 9 μsecs for white noise, 11 μsecs for a 100 Hz tone and 28 μsecs for clicks.

The binaural localization cues may be dichotically traded against one another to give a resultant image direction. The interaction of opposing cues has been shown to be complex and not explainable by simple cue summation. In general, localization in the presence of opposing cues leads to reports of the image being indistinct, and higher variability in responses. Trading functions have been found to be highly

dependent upon the form of the stimulus with reported values between 1.7 μsec/db[53] and 150 μsec/db[54]. A typical set of trading functions, taken from reference 55, is shown in Fig. 2.3.

2.5 Perception of Multiple Objects: A Restriction Upon the Form of a Dichotic Display

In free-field listening, within a field of uncorrelated wideband sources, attention may be switched at will to a single source[34]. The 'cocktail party' effect, as it has been called, allows a listener to concentrate both temporally and spatially upon a single speaker in a distributed field of interfering sources[35].

However, the human auditory system shows a remarkable ability to inhibit conflicting localization information in a reverberant environment. Although secondary echoes may be perceived as a change in quality, the source localization is based almost entirely upon information contained in the direct path component of the incident sound. This phenomenon, known as the Haas or precedence effect[36], has been demonstrated by Wallach, Rosenweig and Newman[37], under both dichotic and free-field listening conditions. With dichotic stimuli a click pair will be heard as a single click if the delay between them is less than approximately 10 msecs, and the azimuthal image direction will be dominated by the interaural differences in the first click of the pair. In a field of

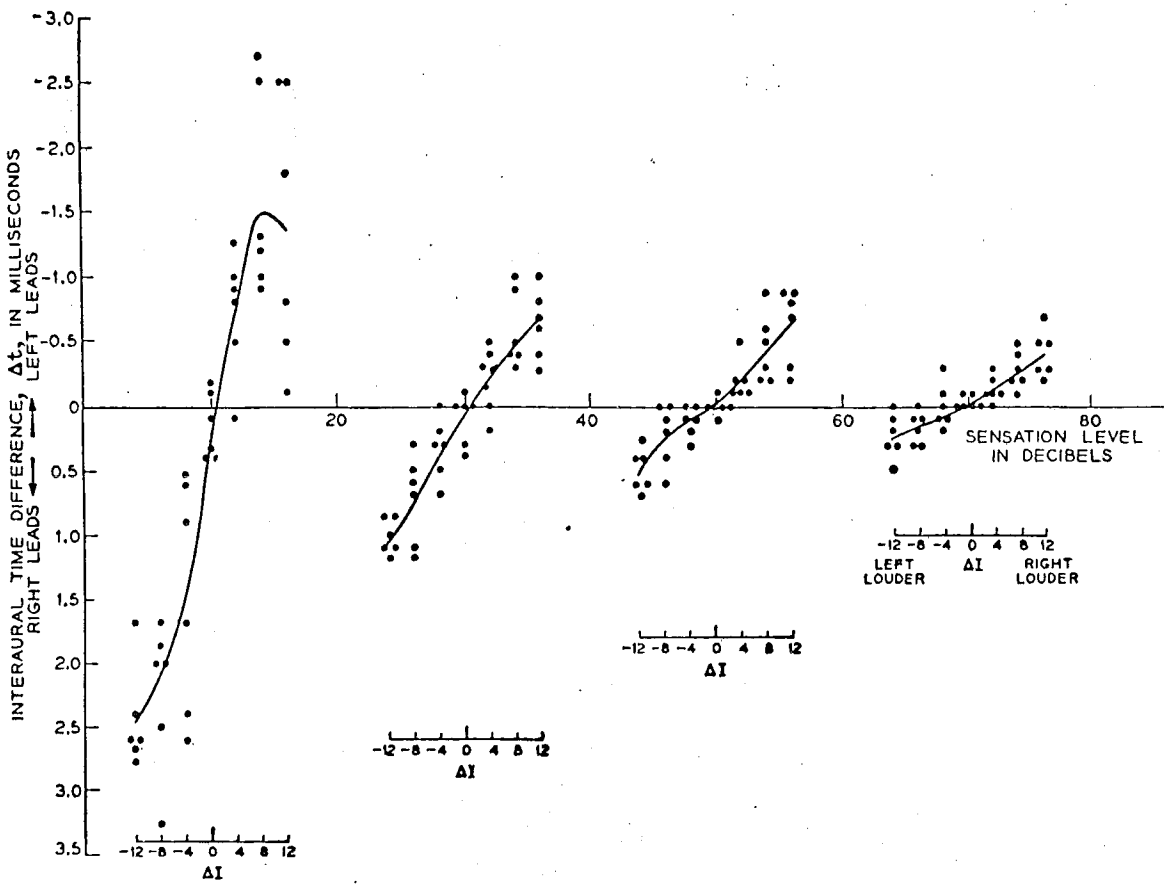


Fig. 2.3 Showing the "trading" effect between IAD and ITD. Note the dependence of the trading function (db/msec) upon the sensation level. (From David, Guttman and van Bergeijk ref.55).

multiple sound sources only the nearest will be perceived, and the perception of direction will be dominated by the nearest source. For some sounds the precedence effect may endure for as long as 40 msecs.

The precedence effect eliminates the possibility of using a real-time temporal display of echolocation returns in which object range is coded as delay.

It seems possible that a time expansion of the ultrasonic pulses, as reported by Rupf[64], might be used to overcome the precedence effect by increasing the delay differences; but the rate of sampling the environment must be reduced accordingly. The perception of multiple objects in real-time is dependent upon finding a form of range coding in which spatially separated objects are represented by uncorrelated display signals. The only other single channel display dimension is a frequency domain range coding. In other words the signal processing must code propagation delay as some monotonic function of frequency.

With this display form the bandwidth of the stimulus must be made as small as possible to attain the maximum range resolution and to minimise auditory masking effects. It was noted in the previous section that the perception of elevation required wide-band stimuli with coding impressed by the pinnae and head. The restriction of the display signal bandwidth therefore will inhibit any attempt to encode the required information. For this reason a dichotic display of this form

will be limited to two dimensions (range x azimuth) and the directional coding will not be as used in normal hearing. The factors affecting the display of azimuth are further discussed in Section 2.4.

Where the criterion is the perception of multiple objects, the idealised display is one that uses tonal stimuli, the frequency of which is a function of object range. The presence of combination tones[39] and threshold elevation by masking[38], will severely limit the attainable range resolution in a multiple object environment (see Chapter 7), but it has been found in practice that the frequency domain coding is far superior to a temporal display. Experiments by the author, using an ultrasonic echolocation system, have verified that pulse techniques allow the perception of only the nearest object.

This form of display affects the azimuthal coding to match the localization function because propagation path length differences, which in normal hearing give rise to phase differences, will be displayed as an interaural frequency difference (IFD). There is no suggestion in the literature of a lateralization with IFD, and in fact for small IFD the instantaneous phase difference may be recognized so that a laterally oscillating image results[68]. With this type of display only the envelope structure of the dichotic stimuli remains for localization using ITD.

The basis of this thesis is the specification of the dichotic azimuthal display. The literature reviewed indicates that in normal hearing the image position is dependent upon a complex interaction of binaural differences, but it is shown in Chapter 3 that the localization phenomena associated with the prototype system is based upon IAD alone. The rest of the thesis defines and examines this unnatural cue usage.

2.6 Range Coding in the Frequency Domain: The Linear FM Echolocation System

The mobility aids proposed by Kay[8-12] use a linearly frequency modulated transmitted waveform and multiplicative signal processing, which codes the range of a reflecting object as the frequency of the auditory stimulus. This form of echolocation display therefore overcomes the restrictions upon multiple object perception by the precedence effect. For this reason the FM echolocation system must be considered to be optimally matched to the behavioural characteristics of audition.

The derivation of the audio signal by the multiplicative signal processing is well documented[66,67] but is presented here for completeness. The linearly frequency modulated transmitted waveform s(t) may be written

$$\begin{aligned} s(t) &= \cos(\omega_0 t_n + \tfrac{1}{2} m t^2_n) \qquad 0 < t_n < T_a \\ &= 0 \qquad\qquad\qquad\qquad T_a < t_n < T_s \end{aligned} \tag{2.9}$$

where T_s is the repetition period, T_a is the pulse duration ($T_a < T_s$), $t_n = t - nT_s$, n is an integer 1, 2, 3..., ω_0 is the instantaneous angular frequency when $t_n = 0$ and m is the rate of change of angular frequency. The instantaneous angular frequency ω_i at time t_n is

$$\omega_i = \omega_0 + mt_n \qquad 2.10$$

Let the transmitting and receiving apertures be non-dispersive and isotropic and let the impulse response of the environment be

$$h_e(t) = \delta(t-\tau)$$

so that the received signal will be, from Eq. 2.2

$$r(t) = s(t-\tau). \qquad 2.11$$

The receiver signal processing consists of multiplication of s(t) and r(t) followed by low-pass filtering to retain only the difference component f(t) which is the audio signal i.e.,

$$\begin{aligned} f(t) &= \tfrac{1}{2}\cos(m\tau t_n + \omega_0\tau - \tfrac{1}{2}m\tau^2) && \tau < t_n < T_a \\ &= 0 && T_a < t_n < T_s \end{aligned} \qquad 2.12$$

The range coding is contained in the frequency ω_a of f(t) i.e.

$$\begin{aligned} \omega_a &= m\tau \\ &= \frac{2m}{c} R \end{aligned} \qquad 2.13$$

from Eq. 2.1. The frequency domain expression for f(t) may be found by evaluating the Fourier transform

$$F(\omega) = \tfrac{1}{2} \sum_{n=-\infty}^{\infty} e^{-jn\omega T_s} \int_{\tau}^{T_a} \cos(m\tau t_n + \omega_0 \tau - \tfrac{1}{2}m\tau^2) e^{-j\omega t_n} dt_n \qquad 2.14$$

which gives

$$F(\omega) = \tfrac{1}{2}\omega_s \delta(\omega - k\omega_s) \frac{T_a - \tau}{2} \frac{\sin(\omega - m\tau)\frac{T_a - \tau}{2}}{(\omega - m\tau)\frac{T_a - \tau}{2}} e^{j(\omega_0 \tau - \frac{1}{2}\omega\tau)} \qquad 2.15$$

where $\omega_s = \frac{2\pi}{T_s}$ and k is an integer $1, 2, 3 \cdots$. This line spectrum exhibits maxima in its envelope when

$$\omega = \pm m\tau$$

$$= \pm \frac{2m}{c} R$$

and zeroes at frequencies

$$\omega = \pm \left[\frac{2m}{c} R \pm \frac{2n\pi}{T_a - \tau} \right] \quad \text{for } n = 1, 2, 3 \cdots \qquad 2.16$$

The audio signal is therefore a repetitive burst of tone whose frequency is linearly proportional to the range of the object. This form of range display allows a restriction in the bandwidth of the analyzer without decreasing the system range resolution capability. In the normal temporal display the signal bandwidth must be maintained through the analyzer, if the full advantage of the transmitted waveform bandwidth is to be utilised. With the frequency domain coding a bandwidth restriction will only limit the region of the environment that can be examined, but the range resolution will be unaltered within this visible region. Therefore the echolocation system described above allows the

information gain derived from wide-band signals to be presented without loss at the auditory display.

Kay has shown that the absolute judgment of range is enhanced by the frequency domain range coding, compared to auditory estimates of delay using a pulse echolocation system[69]. It is shown in the next section that the frequency domain display gives an information increment of only 2.3 bits per stimulus presentation, however the auditory differential sensitivity to stimulus frequency has been shown to be very great (as low as 0.003 of the mean frequency[65]). This means that, when related to the requirements of mobility, this display form will provide efferent/re-afferent kineasthetic feedback that will be extremely sensitive to changes in range. In a practical echolocation system Doppler effects will affect the kineasthetic feedback; this is discussed further in Chapter 6.

2.7 Multidimensional Auditory Displays

An early attempt to substitute the auditory for the visual modality in an instrument monitoring task was the Flybar (Flying By Auditory Reference) display[56] developed during World War II. Multidimensional displays using frequency modulated tones and binaural differences were studied in flight simulators to investigate the possibility of auditory display of turn, bank and airspeed. The authors concluded that the amount of training for non-pilots to use

the auditory display was approximately the same as for the usual visual display.

Mudd[57] evaluated a three dimensional display (frequency, intensity and interaural difference) to determine the effectiveness of auditory cuing in a multiple instrument monitoring task. He concluded that the frequency and interaural difference dimensions were the most effective in reducing reaction time.

A study which is relevant to the present work is that of Black[58] who used a two dimensional display (frequency x interaural difference) to represent the spatial position of a pen on a square writing surface. The horizontal coordinate was displayed as an interaural difference and the vertical position was coded as the stimulus frequency. It was found that handwritten letters could be recognized at a rate of 20 letters/minute. Another experiment asked subjects to indicate the pen position on a sheet. The stimuli were derived from the nodes of an 11 x 11 square matrix presented in random order. The results of this experiment are reproduced in Fig. 2.4 where the mean responses are shown as white nodes and the true positions as black nodes. This study demonstrates that with training subjects are able to relate a two dimensional display to the generating spatial coordinates.

The informational content of multidimensional displays has been investigated by several groups. Although mainly of

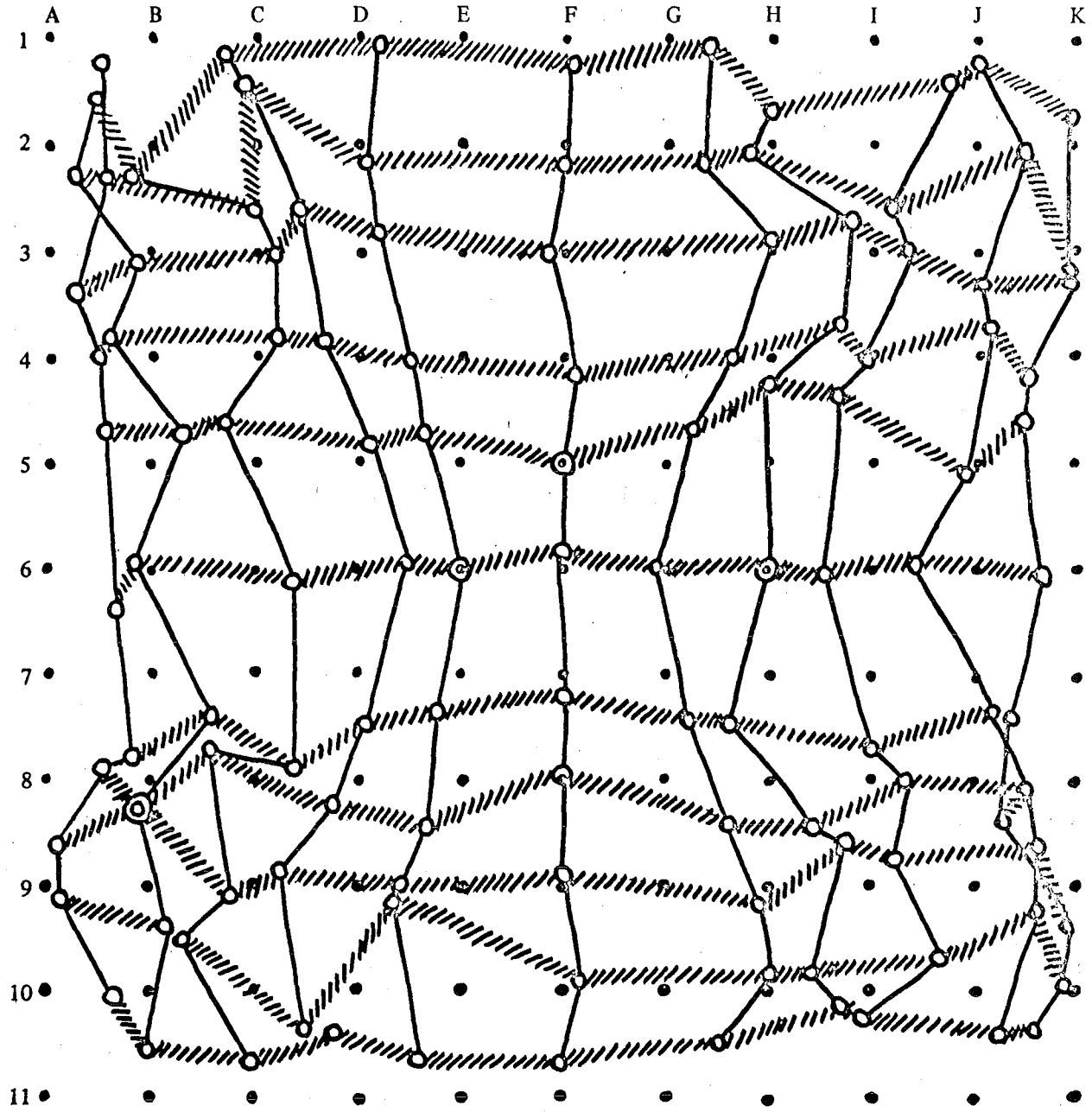


Fig. 2.4 The spatial perception afforded by a simple two dimensional (frequency x interaural difference) display representing the position of a pen on a writing sheet. The mean response at each black node is shown as a white node (From Black ref.58).

academic interest, these studies have shown that the auditory system is limited in channel capacity. The original work of Pollack[59] showed that the vast majority of human subjects cannot, in a task requiring absolute identification, discriminate more than five logarithmically distributed isolated frequencies. Similarly, Garner[60] has shown that only five or six states of absolute intensity can be recognized. These studies show that in a recognition task a single dimension display allows an increment of only 2.3 bits per stimulus presentation.

Pollack and Ficks[61] extended the dimensionality of the display to include frequency and intensity with other independent stimulus parameters (direction, noise content, duration etc.) and concluded that the maximum information per stimulus presentation could be increased to approximately 7 bits. This finding has been supported by other workers[62,63].

CHAPTER 3

EXPERIMENTS IN AUDITORY LOCALIZATION USING A PROTOTYPE BINAURAL ECHOLOCATION SYSTEM

3.1 Introduction

In an early set of experiments Kay reported that he was able to localize the image and maintain fusion when using a prototype model of the binaural mobility aid[12]. He concluded that the localization effect was the result of a complex interaction between the binaural differences generated by differential propagation delays (ITD and IFD) and differences in the angular response of the two receiving apertures (IAD).

The experiments reported in this chapter were the first quantitative measurements of the auditory localization phenomena associated with this auditory display. Two series of experiments were conducted to estimate the effect of the inherent binaural differences upon untrained subjects. The first of these experiments was designed to verify the existence of localization cues in the display and to attempt to map the azimuthal dimension on to the auditory space. The subsequent experiments were designed to study the activity of the three binaural differences in the localization process.

From the results of these experiments it was possible to modify the concept of the binaural system and design a

modified echolocation system that overcame several disadvantages in the prototype display.

3.2 The Binaural Linear FM Echolocation System

The mobility aids proposed by Kay are designed around the linearly frequency modulated waveform and multiplicative signal processing described in Section 2.6. The elements of the binaural system are shown in Fig. 3.1a.

3.2.1. The Prototype System: The electronic circuitry of the device used in the experiments was similar to that of the 'Kay Sonic Aid' manufactured by Ultra Electronics Ltd, and described by Martin[13], with the addition of a second receiver channel. Slight modifications were made to the receiver preamplifiers to reduce capacitive loading of the electrostatic transducers by the connecting cables.

The transducer geometry of the prototype headset is shown in Fig. 3.1b; the actual headset used in the experiments is shown at the top of the Frontispiece. The radiator was a solid-dielectric electrostatic transducer[13] in the form of a 60^{o} circular arc because of the theoretical constancy of the beamwidth of this aperture configuration over a wide bandwidth[77]. In practice these circular arc transducers proved extremely difficult to make and they were later abandoned in favour of a simple disc radiator (see Frontispiece). Two 1 cm diameter transducers, with a beamwidth of approximately 50^{o}, were used as receivers.

The parameters of the echolocation waveform s(t) were

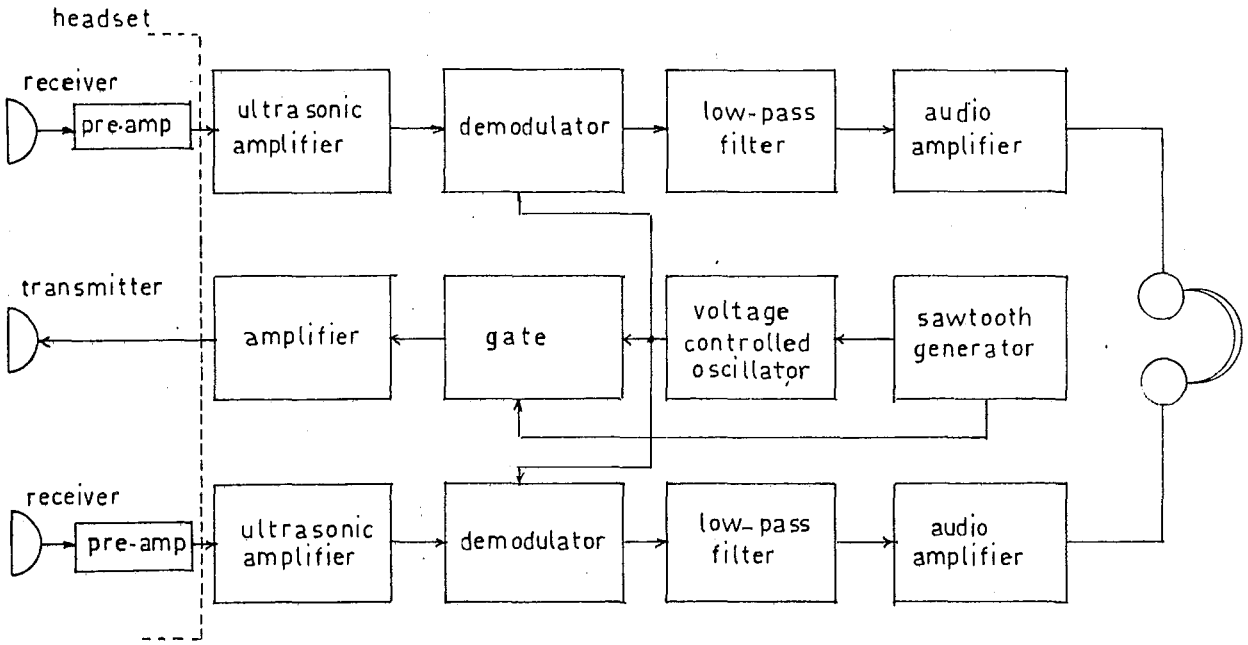


Fig. 3.1a Block diagram showing the elements of the prototype binaural system.

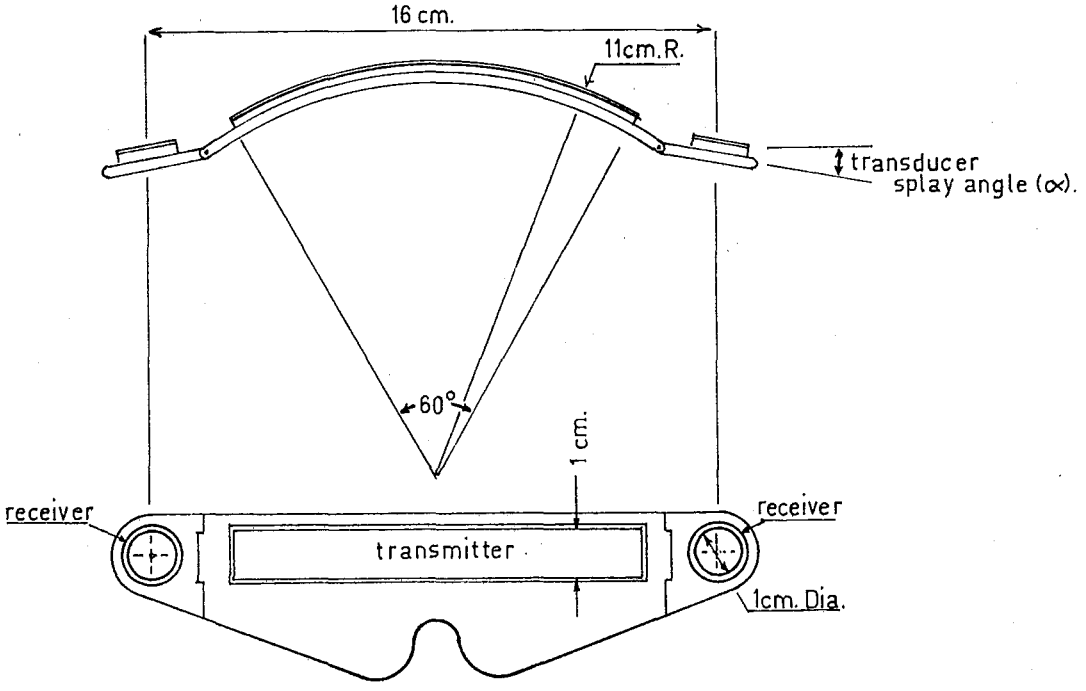


Fig. 3.1b Geometry of the electrostatic transducers of the prototype binaural system.

Repetition period (T_s)	250 msecs
Sweep duration (T_a)	220 msecs
Sweep limits	90-50 kHz
Envelope rise time	20 msecs
Sweep direction	Downward

which gave a range coding at the display of 320 Hz/ft. (In practice it is necessary to limit the bandwidth of s(t) to slightly less than one octave to prevent harmonic intermodulation products from appearing as spurious display signals.)

A spectrum analyzer capable of resolving spectral lines 2Hz apart was developed by the author[78] (described briefly in Appendix 2) to examine the nature of the audio signals. The measured spectra were found to bear little resemblance to the theoretical $\frac{\sin x}{x}$ form derived in Section 2.6. All measured spectra had a pronounced 'sidelobe' on the high frequency side of the main peak. This distortion was found to be present in all of the monaural mobility aids from Ultra Electronics Ltd, and can be explained by the presence of higher order terms in the frequency modulation of the transmitted waveform[79]. The envelopes and amplitude spectra of typical auditory stimuli in the experimental situation are shown in Fig. 3.2.

3.2.2. Binaural Differences in the Prototype System: The two receiver transducers were separated by 16 cm, which is approximately the normal interaural spacing, so as to

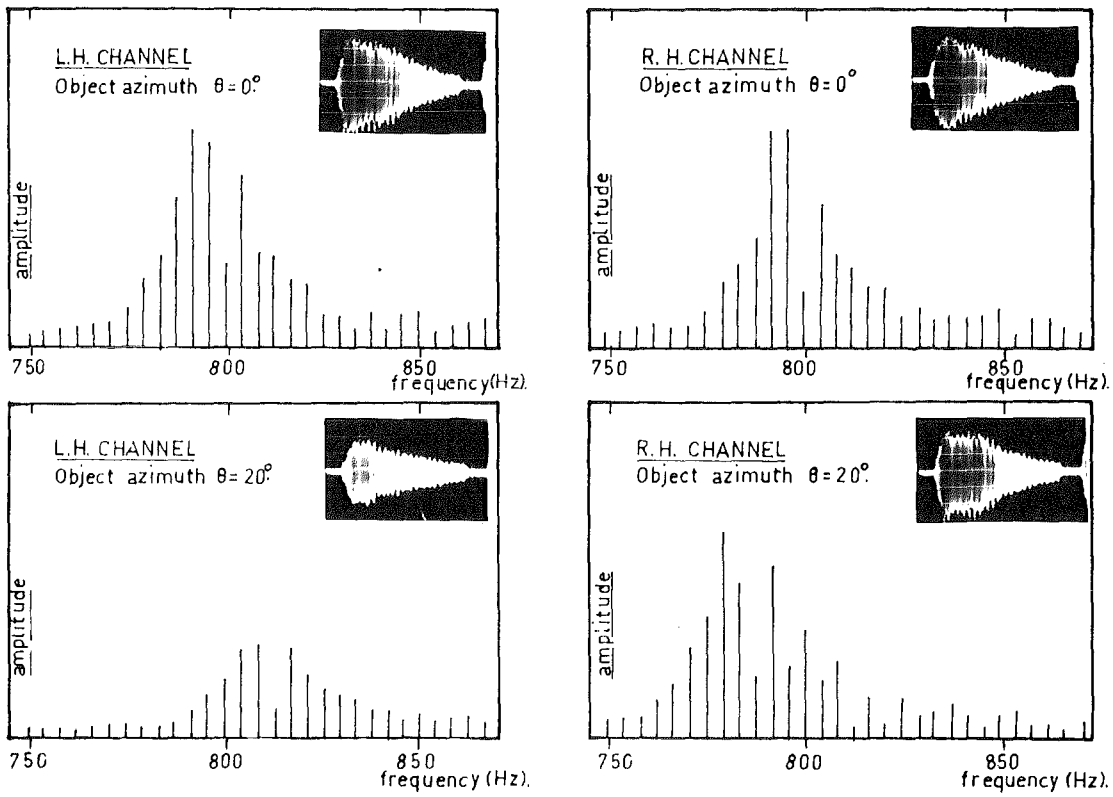


Fig. 3.2 Typical amplitude spectra and waveform envelopes from the prototype binaural system in the experimental environment.

preserve propagation path length differences. From the geometry shown in Fig. 3.3 the path length difference ΔR from an object at range R and azimuth θ will be

$$\Delta R = (R^2 + d^2 + 2Rd\sin\theta)^{\frac{1}{2}} - (R^2 + d^2 - 2Rd\sin\theta)^{\frac{1}{2}} \quad 3.1$$

where 2d is the receiver spacing. The ITD (Δt) and IFD (Δf) are given by

$$\Delta t = \frac{\Delta R}{C}$$

$$\Delta f = \frac{m\Delta R}{2\pi C} \quad 3.2$$

where C is the velocity of propagation, and m is the rate of change of angular frequency as defined in Section 2.6. Calculated values of these binaural differences are plotted in Fig. 3.4. Because the dependence of ΔR upon R is small, the limiting form of Eq. 3.1, i.e.

$$\lim_{R\to\infty} \Delta R = 2d \sin\theta$$

is a satisfactory approximation over most of the working range of the device.

It should be noted that in a two channel system with spaced receivers the estimation of object bearing from path length differences will introduce multiple target ambiguity. If n resolvable objects are present within a range annulus of width equal to the receiver spacing, estimation of both range and direction from propagation delays will give n^2 possible object positions[80]. Such ambiguities may be resolved only by a series of independent measurements on

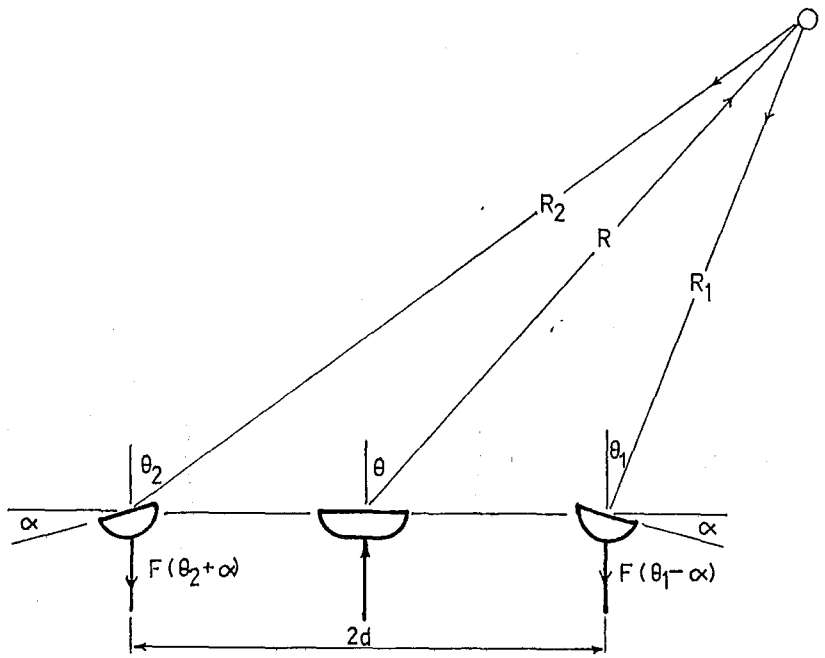


Fig. 3.3 Geometry of a point reflector at a distance R and azimuth θ for deriving the interaural differences in the prototype display.

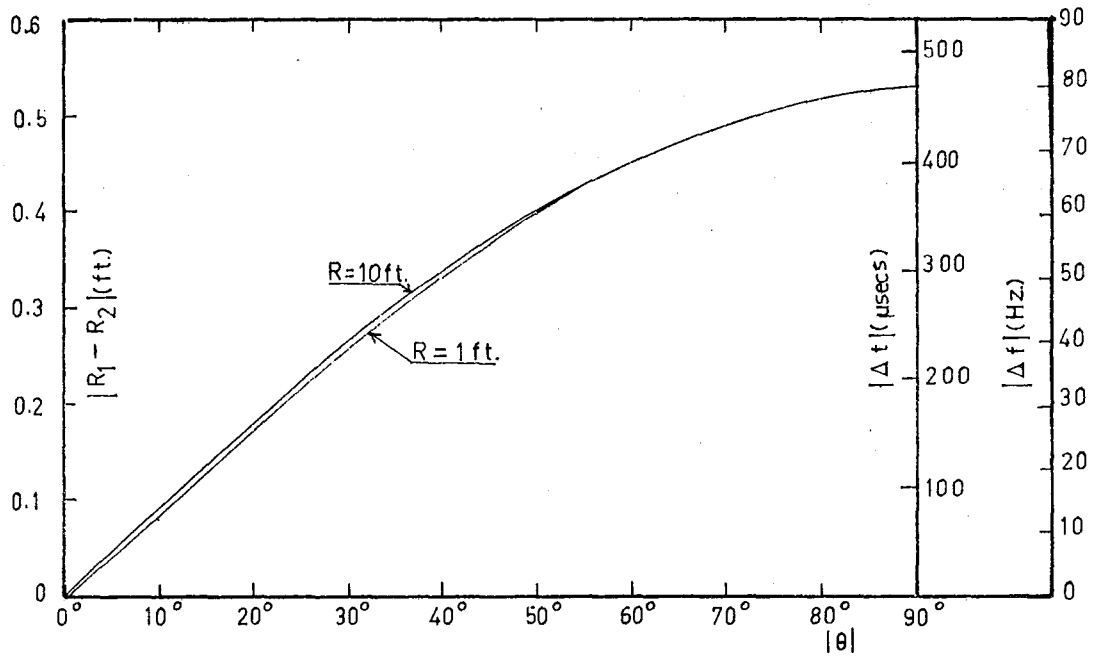


Fig. 3.4 Calculated interaural differences arising from propagation path length differences.

different baselines.

If the azimuthal polar response functions of the two receiving channels are denoted $F_i(\theta;R)$, where the subscript i denotes the channel (i = 1,2), then the IAD (ΔI), expressed in decibels, at the display will be

$$\Delta I = 20 \log_{10} \frac{F_1(\theta;R)}{F_2(\theta;R)} . \qquad 3.3$$

If the two receiving apertures are identical with an azimuthal response $F(\theta_i)$, where θ_i is the azimuth subtended at each aperture, the IAD will be

$$\Delta I = 20 \log_{10} \frac{F(\theta_1)}{F(\theta_2)} . \qquad 3.4$$

From the geometry shown in Fig. 3.3 the azimuth angles will be

$$\theta_i = \cos^{-1} \left[\frac{[R^2 + d^2 - 2Rd \cos(\frac{\pi}{2} + (-1)^i \theta)]^{\frac{1}{2}}}{R \sin(\frac{\pi}{2} + (-1)^i \theta)} \right] \qquad 3.5$$

The range dependence of ΔI may be shown by substituting an analytic form for $F(\theta_i)$: Fig. 3.5a shows the variation of ΔI for

$$F(\theta_i) = \exp [-\gamma \theta_i^2] \qquad 3.6$$

where the constant γ is chosen to specify the beamwidth of the aperture response. The variation of ΔI in the figure was computed with $\gamma = 4.49$ which corresponds to a 6db beamwidth of 45°. The rationale for choosing this expression for $F(\theta_i)$ is discussed in Chapter 5, suffice it to say at this

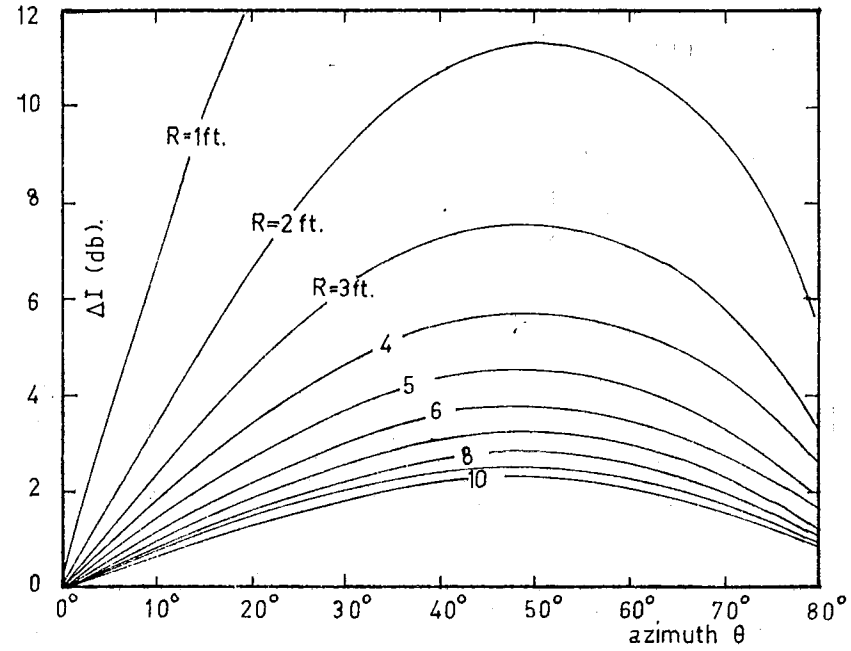


Fig. 3.5a Calculated IAD for receiver azimuthal response given by Eq.3.5 and a transducer separation of 16 cm. The splay angle α is zero for these curves. Note the dependence upon the range R.

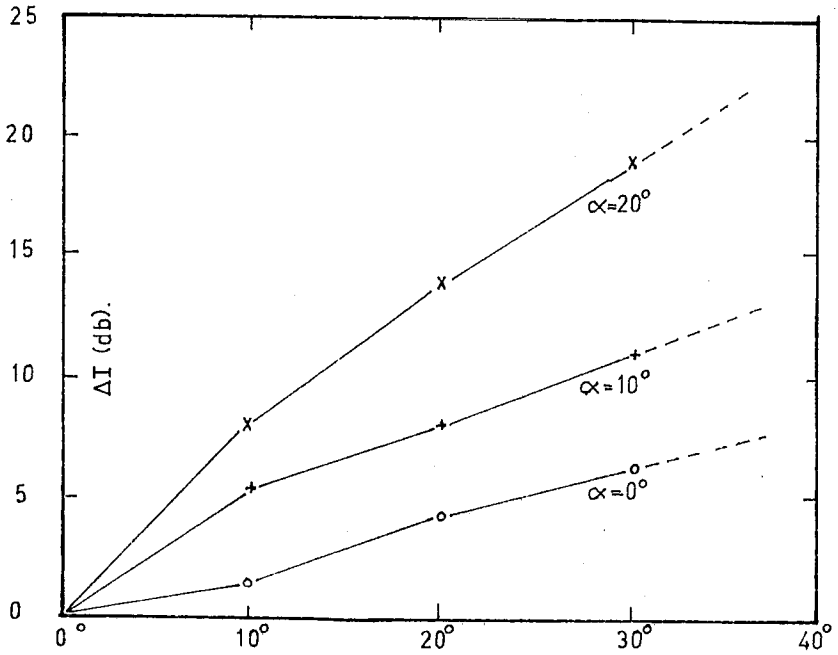


Fig. 3.5b Measured variation of IAD in the prototype binaural system for three values of α.

point that this form is an approximation to the measured response of the electrostatic transducers (Chapter 6).

Additional IAD may be introduced by splaying the two receiver apertures outward at an angle α as shown in Fig. 3.1b. The expression for IAD then becomes

$$\Delta I = 20 \log_{10} \frac{F(\theta_1-\alpha)}{F(\theta_2+\alpha)} \qquad 3.7$$

The measured variation of IAD with object azimuth in the prototype system, within the environment used in the experiments described in this chapter (R = 3'1"), is shown in Fig. 3.5b for values of α of 0^o, 10^o, 20^o.

3.3 Experimental Methods

All experiments were carried out using dichotic stimuli derived from a static environment. Although this experimental situation does not simulate the use of a mobility aid in a dynamic environment, it does enable a study of the binaural cue usage without interference from secondary cues such as head movement and kinaesthesis.

The environment for all but one of the experiments consisted of a single 10" diameter precision brass sphere suspended 3'6" in front of the transducer assembly in an anechoic chamber. This range was chosen because it was one half of the specified working range of the 'Kay Sonic Aid', and because the signal to noise ratio at the display was much reduced at appreciably greater ranges. The headset was

Plate 3.1 The prototype binaural echolocation system in the experimental environment for localization experiments.

pivoted so that the effective azimuth angle of the sphere could be altered. The subjects were seated in an adjacent room and listened to the signals without knowledge of the object azimuth angle.

In the localization experiments an objective measurement of the azimuthal image direction was required. The common psycho-acoustic techniques of matching[81], centring[82], and nullity[83] were unsuitable because they do not record the image placement. Furthermore, it was desirable that the image indication method should be suitable for both blind and sighted subjects. Therefore, methods that involved pointing in the direction of the image, or noting a position on a horizontal chart, were not satisfactory because of their dependence on vision and the requirement that the head position be fixed throughout each experimental session.

The method used was that described by von Bekesy[84]. A pulsating jet of air was directed on to the subject's forehead as shown in Fig. 3.6. The subject's task was to align the tactile sensation with the auditory image so that the image direction could then be read directly from a protractor, or monitored externally. Because the apparatus was affixed to the subject's headset, all measurements were made with respect to the auditory egocentre and there was no need to restrict head movement. This method forces the subject to localize the image in azimuth and therefore largely suppresses reports of pure lateralization[40].

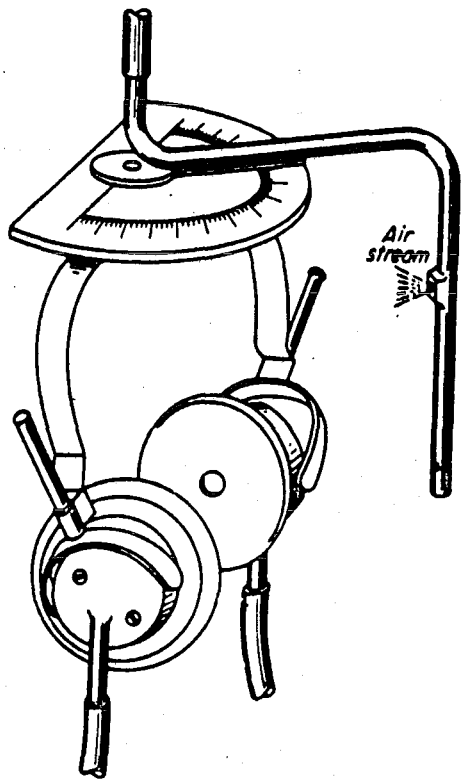


Fig. 3.6 Schematic representation of the method of image direction indication used in localization studies (from von Bekesy, ref. 84).

All subjects who took part in the following experiments were undergraduate or graduate male students in Electrical Engineering, none of whom had previously participated in psychophysical experiments. Audiometric tests were given to all subjects to ensure the absence of any threshold abnormalities.

3.4 The Localization Ability of Untrained Subjects using the Prototype Auditory Display

This experiment was designed to confirm the localization phenomena reported by Kay[12] and to estimate the quantitative mapping of the azimuthal dimension of the display on to each subject's auditory space. The observation of Kay that the localization ability was affected by the splay angle α was examined by measuring localization curves with α set at 0°, 10°, 20°.

The object azimuth angle θ was randomised within and between sessions, with stimulus presentations at 30°, 20°, 10°, 0° to the left and right of the median plane. At each stimulus presentation the subject's task was to indicate the image direction using the apparatus described in Section 3.4. To prevent possible adaptation of the auditory space no corrective feedback was given to subjects at any time. Several sessions were devoted to familiarisation with the task, and during one of these sessions all subjects were informed verbally that the object would always lie within 30° of the median plane, although they should feel free to respond in any direction if appropriate. This was done for

reasons similar to those given by Sayers[85], i.e. to indicate to subjects the expected range of responses. It was realized later that this instruction had significantly influenced the responses.

The stimulus frequency was centred about 980 Hz. In retrospect it would appear that this frequency may have been an unfortunate choice, for later experiments (Chapter 4) showed that the localization function exhibited minimum sensitivity at approximately 1 kHz. The sensation level in each headphone (Akai ASE-9) was adjusted to 50 db when the object was on the median plane.

Seven subjects were used, although it was necessary to discard two of them as reported below. Ten observations at each object azimuth were taken over a total of ten sessions per subject. No attempt was made to hasten responses; subjects reported verbally when they had made their judgement. The stimulus was removed after each response to eliminate possible cues from movement.

Results: Two subject were unable to assign any directional quality to the image and reported that they either heard two separate sounds at the two ears or that the sounds appeared to surround them. There was no noticeable improvement after several periods of explicit corrective feedback, so it was therefore necessary to discard both of these subjects.

The other five subjects were able to localize the image immediately, although one (RJA) reported from time to time

that fusion had broken down with the result that he heard two distinct tones; one at each ear. With practice he was able to maintain fusion throughout the experimental sessions (c.f. Thurlow and Elfner[86]). Two other subjects (DB and JE) expressed difficulty at times because the image appeared to originate from behind them. They found that with practice it was possible to project the image forward.

Localization curves, relating image and object azimuth angles, were plotted for the five subjects and were found to be similar in form. Mean localization curves for the data from all subjects are presented in Fig. 3.7a and the rms error curves are shown in Fig. 3.7b. The localization curves for a single subject are shown, for comparison, in Fig. 3.8.

Discussion of Results: It was found from discussion with the five subjects, at the completion of the experiment, that they had learned to recognize the extreme angular positions (30° either side of the median plane), by the reduced stimulus intensity at the fringes of the polar response of the transmitting transducer, and were confining their responses to within the 60° sector mentioned to them in the instructions. The activity of the binaural differences at the extreme angles was therefore masked and doubt exists as to the true effect of the angle α upon the quantitative mapping of the binaural differences on to the auditory space.

Nevertheless, the experiment proved that the prototype display contained localization information that could be

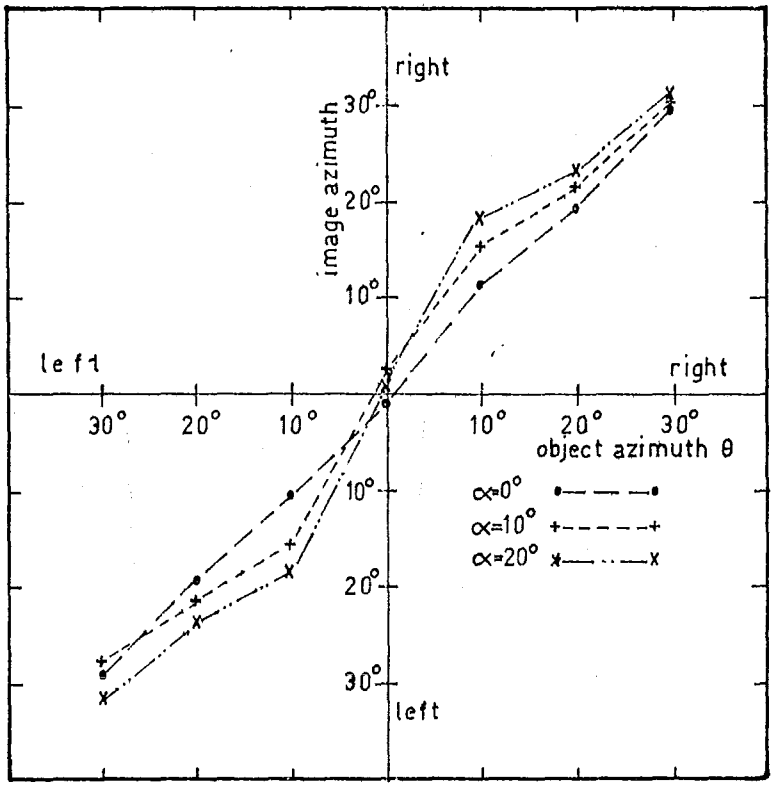


Fig. 3.7a Mean localization curves for all subjects.

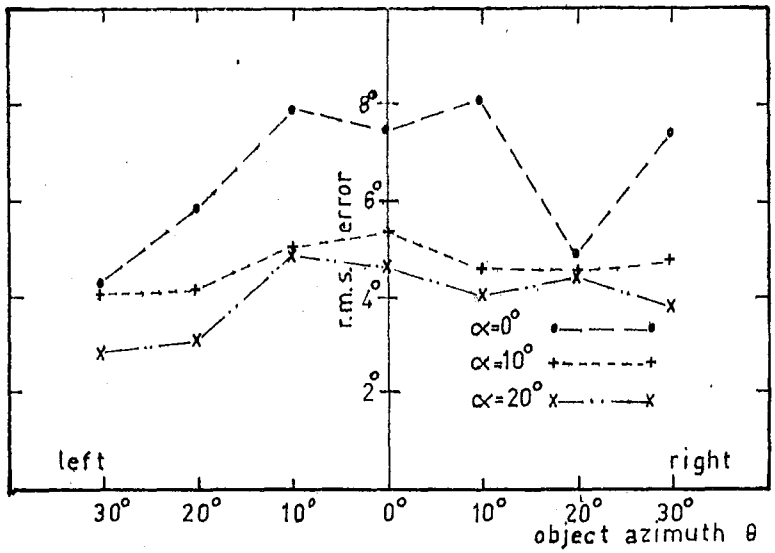


Fig. 3.7b RMS localization error curves for all subjects.

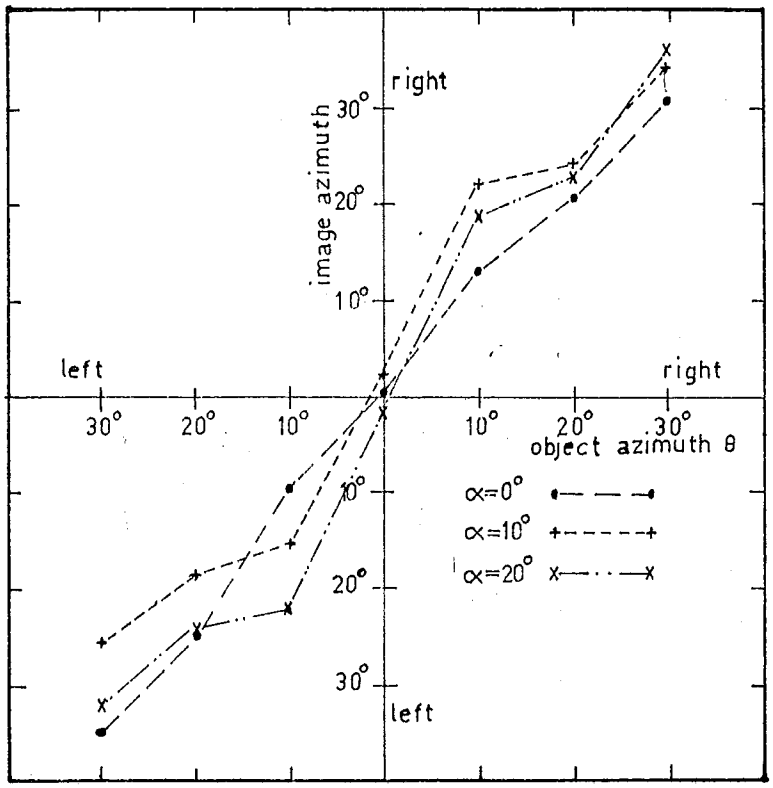


Fig. 3.8a Localization curves for subject TPF.

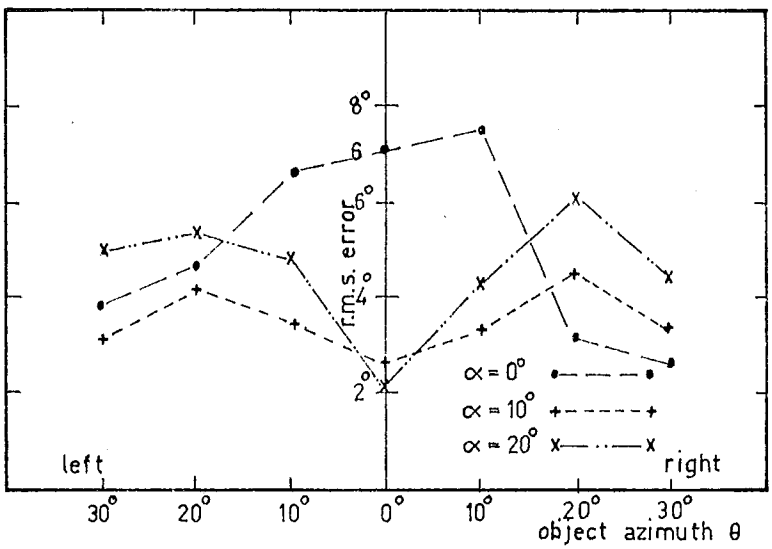


Fig. 3.8b Rms localization error curves for subject TPF.

used by the majority of subjects without training. The fact that for some subjects binaural fusion was not present at all times suggested that the azimuthal display might be optimal. This is discussed further later in this chapter.

3.5 The Role of Path Length Differences in the Localization of the Image from the Prototype Display

3.5.1. The Role of ITD: The experiment to measure the activity of ITD in the localization mechanism was based upon the argument that if ITD was being used, the addition of a further delay in either channel would cause a detectable shift in the image position. The experiment proceeded in three parts as follows:

a) It was first established that all subjects were able to use ITD in the localization of dichotically identical stimuli by using the signal from a single channel of the prototype display applied to the two ears with a differential delay. After two familiarisation sessions the image placement was measured for delays in both channels of 0, 100, 200, 300 μsecs. Twenty-five observations at each delay were recorded for each subject and localization curves were plotted. For this part of the experiment the sphere was located on the median plane ($\theta = 0^{\circ}$) and the splay angle α was set to 0°. The sensation level in each headphone was 50 db.

b) In the main part of the experiment the subject's task was to track the image position as the differential delay was continuously varied between ± 300 μsecs. Stimulus

presentations were randomised between the full binaural display and that derived from a single channel as in part (a) above. The subject was required to give a "movement" or "no movement" response. This part of the experiment consisted of 14 sessions, with two sessions devoted to each of the seven object azimuth angles $\theta = 30^{\circ}$, 20°, 10° to the left and right, and $\theta = 0^{\circ}$. The splay angle α was zero and the sensation level in each headphone was 50 db when $\theta = 0^{\circ}$.

c) The final part was designed to test an apparent anomaly in the results of part (b). The experimental method was the same but the object was carefully aligned on to the median plane so that the audio signals from the two channels remained in phase throughout the pulse.

The additional time delay was introduced by using telescoping sections of acoustic waveguide to insert acoustic paths of different lengths in the two channels as described by von Békésy[87]. The unit is shown schematically in Fig. 3.9. If the difference in path length is Δl then the differential delay Δt will be

$$\Delta t = \frac{\Delta l}{C}. \qquad 3.4$$

Substitution for the acoustic velocity of propagation in air gives $\Delta t/\Delta l = 29.2$ μsecs/cm. The end 10 ft of each tube was filled with acoustic damping material (wool) to reduce

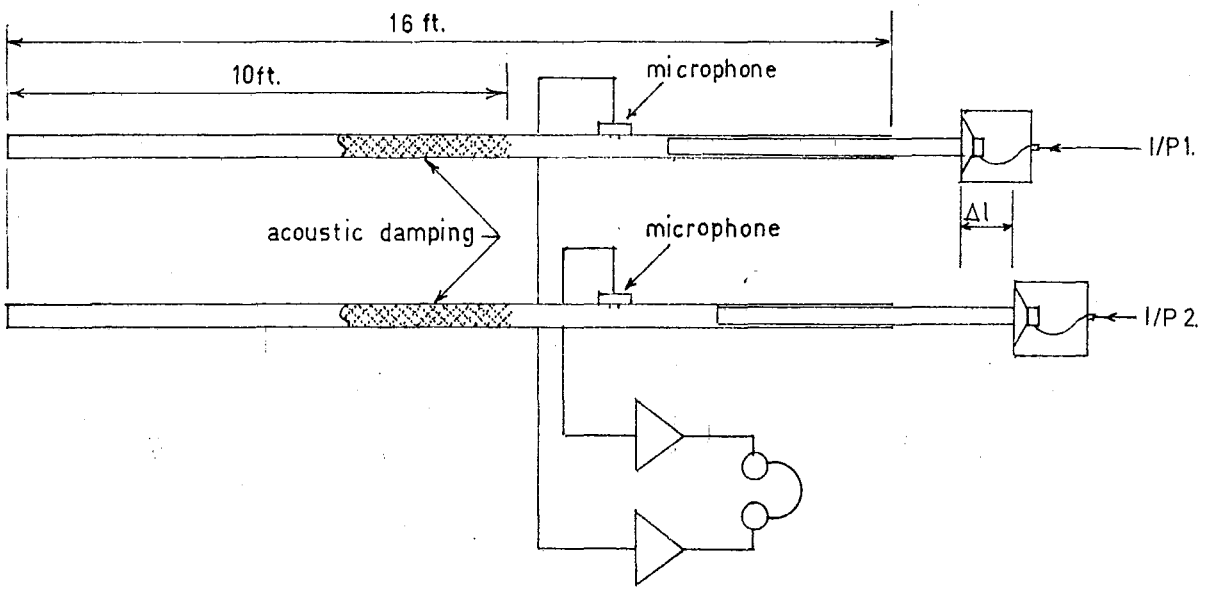


Fig. 3.9 Schematic representation of acoustic delay lines used to generate ITD.

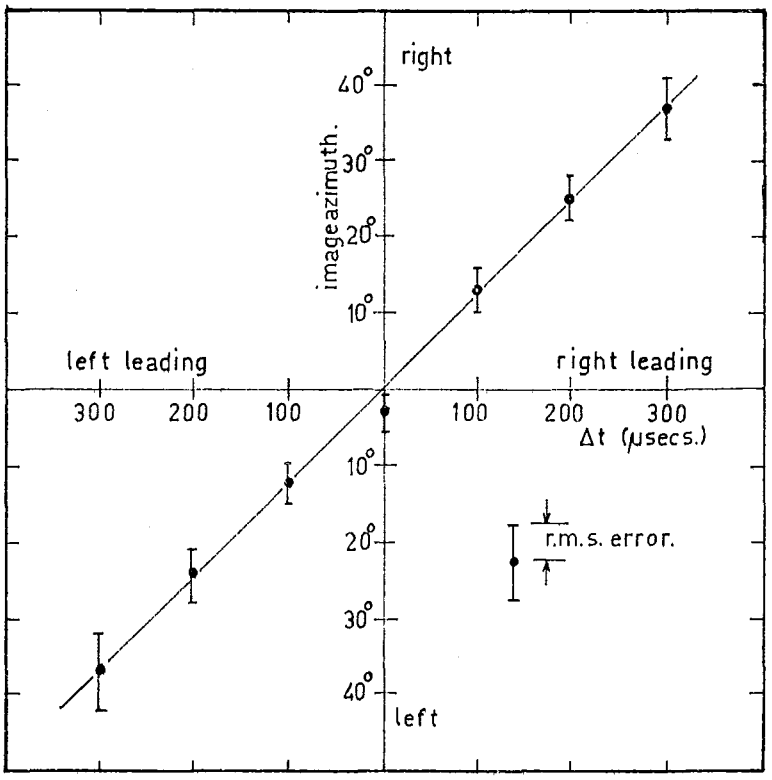


Fig. 3.10 Mean ITD localization curves for three subjects with the stimulus derived from a single channel of the display.

the amplitude of standing waves. The measured variation of amplitude with delay was less than 0.5 db over the full range of delay (± 1 msec). The frequency response of the unit was within 3 db between 500 Hz and 2500 Hz, which was adequate for the narrow band stimuli used in the experiments.

Results:

a) All three subjects were able to localize the image immediately. The least squares line of best fit for individual localization curves had slopes of: DWB 0.13°/μsec, PMC 0.11°/μsec, JDH 0.14°/μsec. The regression coefficients of the three curves were highly significant ($p < 0.01$). The data of the three subjects were pooled to give the localization curve shown in Fig. 3.10. The slope of this curve (0.127°/μsec) shows good agreement with the objective measurements of ITD resulting from a displaced source[46], which indicates that temporal differences would probably be used for the localization of the experimental stimulus in a free field situation.

b) The second part of the experiment gave conclusive results. No subject reported any movement of the image over the total range of interaural delay (± 300 μsecs) with the full binaural display, and all reported image movement for every stimulus derived from a single channel of the system. There were no reports of breakdown in binaural fusion during this experiment.

c) When the object was carefully aligned on to the median plane, so that the signals from the two channels remained in phase throughout the whole pulse, the ability of all subjects to detect an image shift as the ITD was varied, was restored. However, it was found that a very slight displacement of the azimuth angle ($< 0.5^{\circ}$) was sufficient to destroy this ability.

Discussion of Results: The first part of the experiment proved the stimulus structure was such that ITD could be used for localization by all subjects if the dichotic stimuli were identical. This suggested that binaural mechanism would use this cue if it was available in the display. However, when the full binaural display was used, the difference limen (DL) for ITD was eliminated (or at least degraded to greater than the maximum delay of 300 μsecs used in this experiment). This was found to be so even when the object was on the median plane. Therefore the third part of the experiment was undertaken to find the condition that would restore the limen. It was found that very careful alignment of the object on to the median plane, using an oscilloscope to check the phase relationship between channels, was necessary to restore the limen. In part (b) of the experiment the alignment had not been careful enough and the small IFD was sufficient to prevent the use of ITD. It was therefore concluded that the IFD destroyed the ability to use ITD.

When considered as a whole, the three parts of this experiment show that the form of the binaural display renders the natural localization cue of ITD unavailable to the auditory system. The lack of cross-correlation between the microstructure of the dichotic stimuli deprives the localization mechanism of the delay information that would normally be used, and the envelope structure of the tonal stimuli is unsuitable for the detection of ITD from an envelope comparison process.

3.5.2. The Role of IFD: The following experiment was performed to study the effect of IFD upon the perceived image position.

Pairs of stimuli were presented, with the object at a fixed azimuth angle, in which the IFD in the first and second stimuli differed by a factor of approximately 2:1 while the other binaural differences were maintained constant. The subject's task was to localize the image on the first presentation using the inherent localization cues, and to indicate the direction of any image shift ('left', 'right' or 'no change') upon the second presentation of the pair.

The IFD-azimuth angle relationship was altered between presentations by switching a resistive attenuator into the control voltage circuit of the FM oscillator (Fig. 3.1a) thus reducing the sweep limits to approximately 70-50 kHz. The sweep rate m and hence the range coding were therefore reduced accordingly. The change in frequency of the dichotic

stimuli was compensated by changing the effective object range. It was found in a pilot experiment that it was impracticable to move the 10" diameter sphere between stimuli, therefore the reflector was simulated by switched remote transmitters at distances of 7'0" and 14'0". These transducers were fed from pre-set attenuators to equalize the stimulus intensity at the two ranges. The differences in the other two binaural cues between the two IFD conditions with the object azimuth at 30°, which was the worst case used in the experiment, were subliminal (calculated difference in ITD < 5 μsecs, measured difference in IAD < 0.5 db).

The order of IFD presentation in the stimulus pair and the object azimuth angle (30°, 20°, and 10° either side of the median plane) were randomised between and within sessions. The splay angle α was zero and the sensation level in each headphone was 50 db with the object upon the median plane. The same three subjects who participated in the experiment on the role of ITD were used in this experiment; they were experienced in detecting any lateral image shift.

Results: The three subjects were able to use the inherent binaural cues to localize the image on the first presentation of each pair, irrespective of the order of IFD presentation, with no reports of breakdown in binaural fusion. Although they noted a slight change in pitch between the two IFD conditions, no subject was able to detect any change in the azimuthal image direction. The experiment was discontinued

after each subject had spent three half-hour sessions attempting to detect any image shift.

Discussion of Results: The experiment failed to confirm any lateralization of the image when IFD alone is changed, as was reported by Kay[12]. The subjects were not able to detect the frequency difference or any change in it, therefore it could not be used for the detection of azimuth as hypothesized in the earlier papers[8,9,10]. However, this and the previous experiments demonstrate that binaural fusion and localization are not dependent upon correlation between the microstructure of dichotic stimuli[88].

3.6 The Role of Interaural Amplitude Differences in the Localization of the Image in the Prototype System

The centring technique was used to estimate the IAD required to align a displaced image, for various object azimuth angles, back on to the median plane. The resultant trading function estimated the IAD required to compensate for the summed effect of all active localization cues. Conclusions as to cue usage were inferred from the previous experiments described in this chapter.

Experimental Method: A continuously variable differential attenuator was inserted in the binaural audio channels as shown in Fig. 3.11. The potentiometer was under the subject's control but was designed so that only the experimenter could see the calibrations.

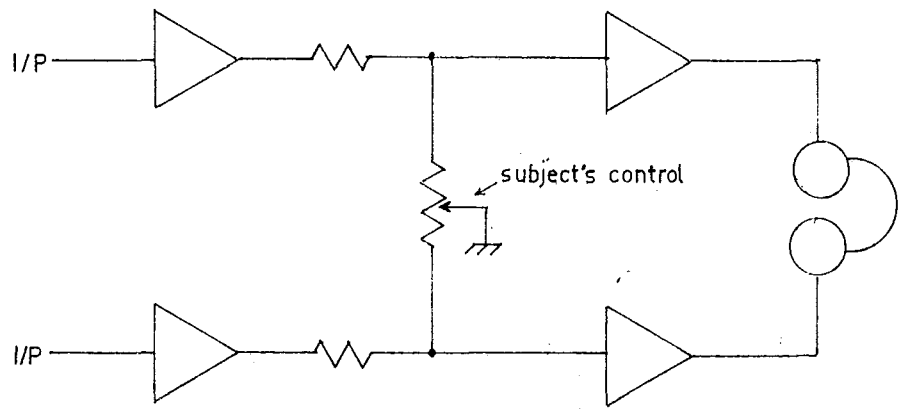


Fig. 3.11 Continuously variable differential attenuator used in centring experiment. The subject was not aware of the absolute setting of the potentiometer.

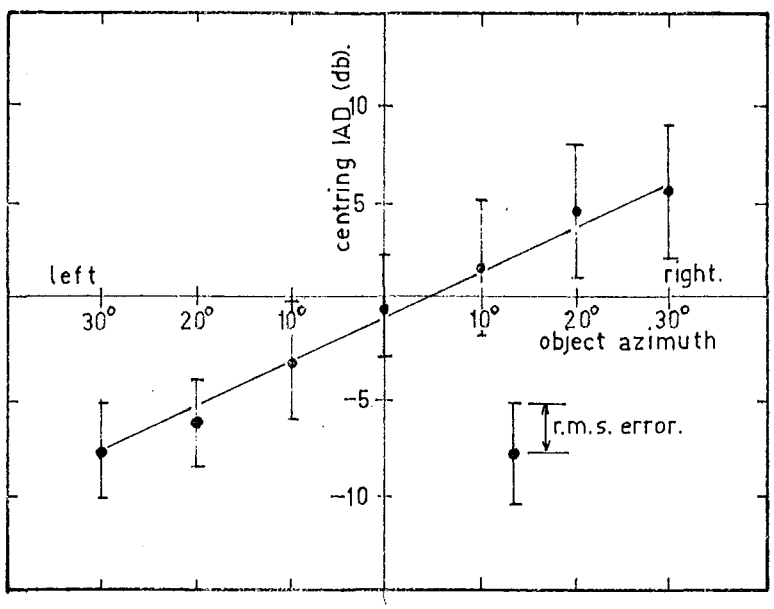


Fig. 3.12 Mean centring IAD-azimuth trading function for five subjects. The slope is 0.23 db/deg.

At each stimulus presentation the subject's task was to adjust the potentiometer so that the auditory image was aligned on to the median plane. The object azimuth angle was randomised within and between sessions and was presented at 30°, 20°, 10° left and right and 0° (straight ahead). The trading function was estimated from twenty observations taken at each azimuth angle. All centring judgements were started with the attenuator set for 0 db differential attenuation and no attempt was made to hasten judgments. The stimulus was removed after each response.

Five subjects were employed in this experiment; two (PMC and JDH) had participated in the experiments on the role of ITD and IFD, and the others had not previously taken part in psycho-physical experiments. Two half-hour familiarisation sessions were given to all subjects before results were recorded.

The auditory display was as used in previous experiments. The reflecting object was the 10" diameter sphere at a range of 3'6", the receiver transducers were set with $\alpha = 0^{\circ}$, and the sensation level was adjusted to 50 db in both channels when the object was located on the median plane and the potentiometer set for 0 db differential attenuation.

Results: All subjects were able to centre the auditory image at all object azimuth angles; although some reported difficulty at times because of the diffuse nature of the image. Trading functions were calculated from the regression

coefficients of the object-azimuth centring-IAD data; the results for individual subjects are presented in Table 3.1. All regression coefficients were highly significant ($p < 0.01$).

Table 3.1. Individual Trading Functions for IAD Centring Experiment

Subject	Trading Function db/deg	St.Dev. of Estimate db/deg
PMC	.29	.07
JDH	.30	.10
AER	.23	.06
DR	.27	.06
RPS	.17	.07

The results of individual subjects were compared in an analysis of variance (similar to the one developed in Appendix 3 for a factorial design) to test the significance of differences in the slopes of the regression lines in a one-way classification. It was not possible to reject the null hypothesis of a common trading function for all subjects ($p > 0.1$) therefore all results were pooled to derive a mean estimate of the trading function of 0.23 db/deg. of object azimuth.

The trading function derived from the pooled data is plotted in Fig. 3.12. The correlation coefficient between the mean centring IAD at each azimuth angle and the measured IAD at that angle was calculated as r = 0.97, which suggested that subjects were compensating only for the IAD in the display.

Discussion of Results: The activity of IAD in the localization process must be inferred from the results of the previous experiments, and the correlation between centring IAD and measured IAD at each object azimuth. This experiment showed that, within sampling errors, subjects were compensating for the summed effect of all cues with an opposing IAD which was almost equal to the IAD in the display. Because it was shown in the previous experiments that the other binaural differences had no localization effect, it is reasonable to conclude that the only active localization cue in the display is IAD.

3.7 Conclusions as to Cue Usage

The conclusion drawn from the above series of experiments was that IAD was the only binaural difference recognized by the localization mechanism. This finding is at variance with the conclusions drawn by Kay[12] who described a series of experiments on the role of the binaural differences. It was stated that with IFD alone and in combination with ITD a definite lateralization of the image resulted; experiments described in this chapter failed to confirm these phenomena.

There was an implicit assumption in Kay's experiments that when the transducer splay angle α was zero there was no IAD in the display; the author has shown this to be incorrect, see Figs 3.5a and 3.5b. The binaural cues in these experiments were therefore similar to those of the experiment described in Section 3.4 with $\alpha = 0^{o}$. If the IAD is considered, the qualitative results of Kay's experiments are reasonable and agree with the conclusions of this chapter.

In the same paper Kay found that with ITD alone there was only a vague lateralization effect. The results of the experiment on ITD in this chapter showed a definite and sensitive localization effect when the dichotic stimuli were correlated. However, the stimulus frequency used by Kay was 3 kHz, which is well above the accepted maximum frequency of 1500Hz for the recognition of phase differences by the localization mechanism. The author considers that the vague lateralization effects reported by Kay were caused by the IAD in the display.

3.8 The Effect of Object Orientation upon Auditory Localization with the Prototype System

In localization experiments using objects with a directional back-scattering response it was found that the image direction was critically dependent upon the object orientation. For example, with a metal strip 2" wide on the median plane at a range of 5' a small rotation (approximately 10^{o}) was

sufficient to completely lateralize the image in either direction. This effect is not the same as that induced by head movement, when the azimuth angle of the object is changed.

Consider the geometry shown in Fig. 3.13. Let the scattering object at P have range R, azimuth θ and orientation ρ as shown. Then the amplitude of the back-scattered energy may be described by a function $B(\theta+\rho,\psi)$ where $\theta+\rho$ is the angle of the incident energy and ψ is the scattering angle. In this situation the IAD $\Delta I'$ will be

$$\Delta I' = 20 \log_{10} \frac{F(\theta_1-\alpha)B(\theta+\rho,\theta_1)}{F(\theta_2+\alpha)B(\theta+\rho,\theta_2)}$$

$$= \Delta I + 20 \log_{10} \frac{B(\theta+\rho,\theta_1)}{B(\theta+\rho,\theta_2)}$$

where ΔI is given by Eq. 3.6. The condition that there will be no detectable change in image direction with orientation is

$$20 \log_{10} \frac{B(\theta+\rho,\theta_1)}{B(\theta+\rho,\theta_2)} < \delta I \quad \text{for all } \rho,$$

where δI is the psychophysical difference limen in IAD. For a given object the function $B(\theta+\rho,\psi)$ will be fixed so that the design of the echolocation system so as to overcome this effect will be dependent upon the relationship between θ_1 and θ_2. The localization experiments with finite reflectors showed that the prototype display was not satisfactory in that it did not preserve a one-to-one mapping of the environment on to the auditory space.

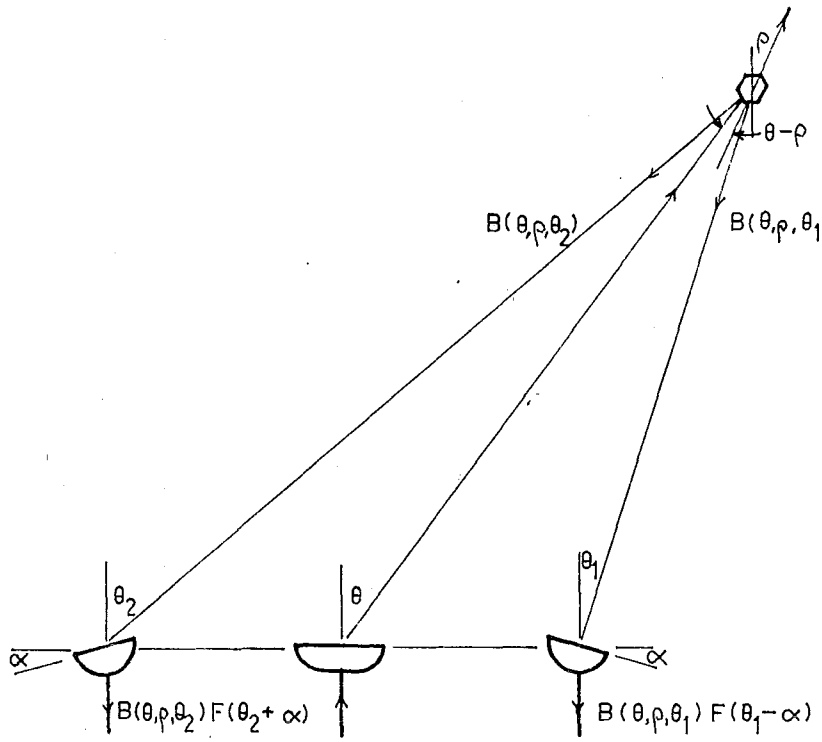


Fig. 3.13 Geometry showing addition IAD impressed upon the display by a directional scatterer.

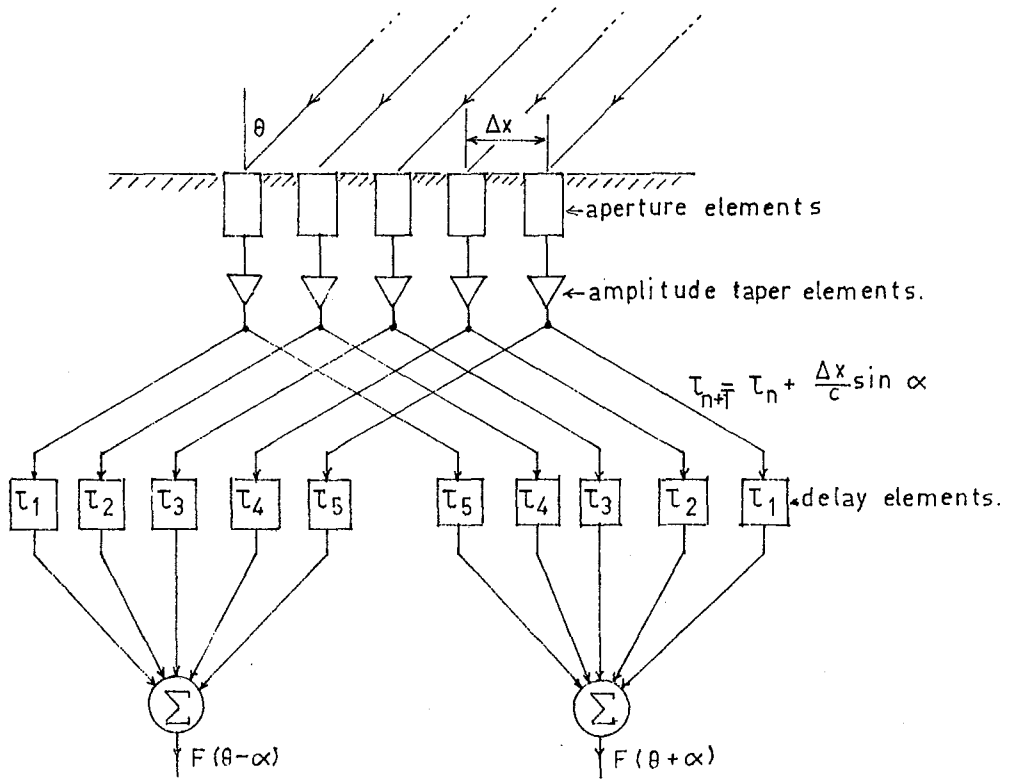


Fig. 3.14 A possible method of eliminating path length and scattering angle differences by synthesizing the angular displacement of two receivers using a single aperture.

3.9 The Modified Binaural Echolocation System

The experimental work described in this chapter showed that the auditory display from the prototype system was not optimum for the following reasons:

i Some subjects were able to detect the dichotic frequency difference by a breakdown in binaural fusion (Section 3.4).

ii The display contained redundant binaural differences that were shown to have no effect upon localization but which could have introduced additional variability into the subjects' responses (Section 3.5).

iii For some environmental objects the apparent azimuth angle was critically dependent upon the orientation of the object (Section 3.8).

iv The active localization cue, IAD, was dependent upon the range of the reflecting object (Section 3.2.2).

These four factors have a common cause; the spatial separation of the two receiving apertures.

The author proposed that path length differences should be eliminated by physically mounting the two receiving apertures as close together as possible. In this modified system the binaural differences caused by propagation delays are eliminated (or made as small as practicable) and θ_1 is made equal to θ_2 so that the range dependence of IAD and the effects of directional scatterers are eliminated. The only remaining binaural difference is IAD, which is a function only of the system parameters, namely $F(\theta)$ and α. If $\theta_1 = \theta_2$

Eq. 3.6 may be written

$$\Delta I = 20 \log_{10} \frac{F(\theta-\alpha)}{F(\theta+\alpha)}$$

In the following chapters the form of $F(\theta)$ is derived so that the display IAD is matched to the behavioural characteristics of the localization mechanism.

The idealized echolocation system based on this concept would use a single receiving aperture to derive the two dichotic signals, as shown in Fig. 3.14. With a suitable linear delay taper across the two outputs from the array the signals in the two channels will be the same as that derived from two separate but coincident apertures. If the spatial separation of the elements is δx the delay $\Delta\tau$ between adjacent elements to simulate an angular displacement α will be

$$\Delta\tau = \frac{\delta x}{c} \sin\alpha .$$

The form of $F(\theta)$ can be controlled by further amplitude and delay tapers across the aperture.

The present state of air-borne ultrasonic transducer technology has required the use of separate but adjacent transducers. The evolution of a suitable transducer geometry for the mobility aid can be traced through the Frontispiece.

Extensive tests were carried out to compare the original and modified displays in both static and dynamic environments. A model of the mobility aid was made in which the two displays could be compared by switching between two

sets of receiving transducers (see the second model in the Frontispiece), the first pair spaced at the normal interaural distance, the second pair mounted vertically above one another. The modified display was found to be superior and has been used on all subsequent models of the mobility aid.

In the development of the specification, and the estimates of the performance of the binaural echolocation system in the following chapters, the modified system is assumed and all analyses are based on the estimation of azimuth from IAD alone.

C H A P T E R 4

THE LOCALIZATION OF TONES USING INTERAURAL AMPLITUDE DIFFERENCE

4.1 Introduction

The experimental results of Chapter 3 showed that in a display using a frequency domain range coding, IAD is the only binaural difference that can be used as the azimuthal dimension. A specification of the analytic form of the azimuthal response of the receiving apertures, to map the environment on to the auditory space, is therefore dependent upon a quantitative knowledge of the IAD localization curve. In order to determine the range of system parameters required to match the system to all individuals it is also necessary to know the spread of the coefficients of the localization curve throughout the blind and sighted populations.

Two experiments were conducted with dichotic tonal stimuli to derive quantitative data on the effect of IAD upon the auditory image direction. The first experiment was a detailed study of the 'within subject' dependence of the localization function upon the stimulus parameters. The second experiment used a larger sample, containing both blind and sighted subjects, to estimate the spread of the coefficients of the localization function. A third experiment investigated localization phenomena when the IAD

was time varying, as is the case with the dichotic signals from an echolocation system.

The object of these experiments was to measure the perceptual effect of IAD and to gauge the reliability of this display dimension. Some of the questions that were implicit in the experimental designs were:

i What is the required analytic form of the IAD-azimuth angle relationship to map the environment on to the true auditory space, and will the coefficients of this function have to be determined for each individual?

ii Are there any significant differences in the localization functions of the blind and sighted populations?

iii Is the localization function frequency dependent, i.e. will there be interaction between the range and azimuth dimensions?

iv Is the localization function intensity dependent, i.e. will the apparent direction of an object depend upon the scattering cross-section and the system power gain?

v Is the localization function dependent upon the stimulus duration (as it appears to be when ITD is the active cue[51]), and the repetition period?

4.2 Literature Review

Despite the volume of the literature[89,90] on dichotic phenomena there are few references of direct relevance to the present study. The majority of works that quantitatively

define the effects of IAD were published in the period circa 1925, when the available apparatus did not accurately specify or control the stimulus parameters. More recent studies have tended to concentrate upon defining the role of the binaural differences in normal hearing, and the experimental techniques devised for this purpose have not given quantitative measures of the image direction.

In an early paper Stewart and Hovda[91] reported that the azimuthal direction of a low frequency tonal image could be expressed in the form

$$\hat{\theta} = K \log \left[\frac{I_R}{I_L}\right] + \theta_0 \qquad 4.1$$

where $\hat{\theta}$ is the estimated azimuth, I_R and I_L are the stimulus intensities in the right and left ears respectively, K is a constant defining the activity of the cue, and θ_0 is the value of $\hat{\theta}$ when the IAD is zero. In a later paper Stewart[92] reported that IAD was ineffective in displacing the image over a range of frequencies between 200 and 4000Hz. Other workers[93,94] failed to confirm this effect and concluded that IAD was effective over the whole audio spectrum.

Recent papers relating the image displacement to dichotic differences are rare. Researchers have concentrated upon defining the physiological sites of binaural interaction, and measuring the interrelation between dichotic phenomena and objective binaural differences at the outer ear in free field listening. In these studies, two experimental methods

have commonly been used, both of which are variants upon the psychophysical method of average error[95]. The interrelation between cues has been studied by the 'centring' method, in which a displaced image is realigned on to the median plane by another cue acting in opposition[82]. (This method was used to investigate the role of IAD in the prototype display, see Section 3.6.) The 'matching' or 'pointer' method uses a second dichotic image as direction indicator. Observed localization phenomena may then be related to known binaural differences in the pointer stimulus which is usually wide-band noise[81]. Neither of these methods gives a direct estimate of the effect of a single cue acting alone, and have led to different conclusions as to the physiological processing of IAD. For example, Deatherage and Hirsh[96], using centring techniques, concluded that IAD was effective only in varying the latency of peripheral neurons, while Hanson[97], and Moushegain and Jeffress[98], using matching, inferred that IAD must be recognized per se by the central nervous system.

Many experimenters have reported that, under conditions of unnatural binaural stimulation, discerning subjects are able to report the presence of multiple images[99,100,101,102]. In centring experiments two images may be reported; one of which will be dependent upon time differences for its lateral position, the other dependent upon amplitude differences. The trading functions for the two images differ greatly in magnitude. In reporting the presence of dual images in dichotic localization experiments Trimble[99] stated that IAD

alone did not give rise to a splitting of the image, whereas phase differences alone did so. He reported that the displaced image was sharpest and best defined when IAD was the only localization cue. Other workers have contradicted this finding by reporting the image to be diffuse and ill-defined under impressed IAD[103,104]. Sayers[105] reported difficulty in conducting localization experiments using IAD because of the imprecise nature of the image.

There is further conflict in papers defining the role of IAD in free-field localization. As pointed out in Chapter 2, Mills[48] concluded that IAD was actively used for localization of tones above 1500Hz, while Feddersen et al[46] found that there was insufficient IAD induced from an offset source to explain localization at frequencies below 5kHz.

The available literature on the effect of dichotic IAD must therefore be described as contradictory. It would appear that the conclusions drawn are dependent upon the experimental techniques used. Even between recent papers there exists such differences in the results and conclusions that it is not possible to collate the available data and produce a unified theory of auditory localization[106].

4.3 The Effect of IAD upon the Azimuthal Direction of a Tonal Image

This experiment was undertaken in an attempt to resolve the contradictions in the literature on the quantitative effect of IAD, and to investigate the effect of the stimulus

parameters (frequency, sensation level, duration) upon the localization function. The experiment was constructed in a three way factorial design with seven levels of frequency, five levels of pulse duration, and three levels of sensation level. In each cell of the design, seven levels of IAD were applied, each a total of ten times, to derive the coefficients of the localization function. The activity of the cue was estimated by the slope of the localization curve.

The levels of the factors in the design are defined in Table 4.1. A table of random numbers was used to divide the complete block of 735 stimulus conditions into seven sub-blocks each containing 105 stimulus conditions.

TABLE 4.1 DEFINITION OF LEVELS IN THE FACTORIAL DESIGN

Factor	Units	Level 1	2	3	4	5	6	7
Frequency	Hz	100	300	500	1000	1500	3000	5000
Duration	msec.	50	100	250	500	*	-	-
Sensation level	db	40	60	80	-	-	-	-
IAD	db	8	4	2	0	-2	-4	-8

*Continuous tone

These sub-blocks were presented in randomised direction and order, between and within the ten repetitions of the complete

block. In a pilot study run before the main experiment, the localization function for two stimulus conditions (300 and 3000Hz at 60db sensation level and 250 msecs duration) was estimated. The purpose of this part of the experiment was twofold: to familiarise subjects with the experimental technique and localization task, and to establish the statistical significance and the form of the localization curve. All subjects spent at least 15 sessions on the pilot study and did not proceed to the main experiment until an analysis of variance for a linear regression had established that the slope of the localization curve was significant at $p < 0.01$.

Experimental Apparatus: The stimulus generator was built as a special unit with push button selection of the stimulus parameters. The block diagram of the apparatus is shown in Fig. 4.1.

The tone generator was an amplitude stabilised Wein bridge oscillator with harmonic content of less than 1%. The pulse duration and repetition period were controlled by a synchronous gating circuit that was triggered by positive-going zero-crossings of the oscillator. All pulses were therefore an integral number of complete cycles in duration. No attempt was made to shape the rise-fall characteristics of the envelope, other than the inclusion of a small time constant (250 μsecs) to eliminate the switching transients from the FET gate. The breakthrough from the gate in the 'off' condition was greater than -80db with respect to the 'on' condition.

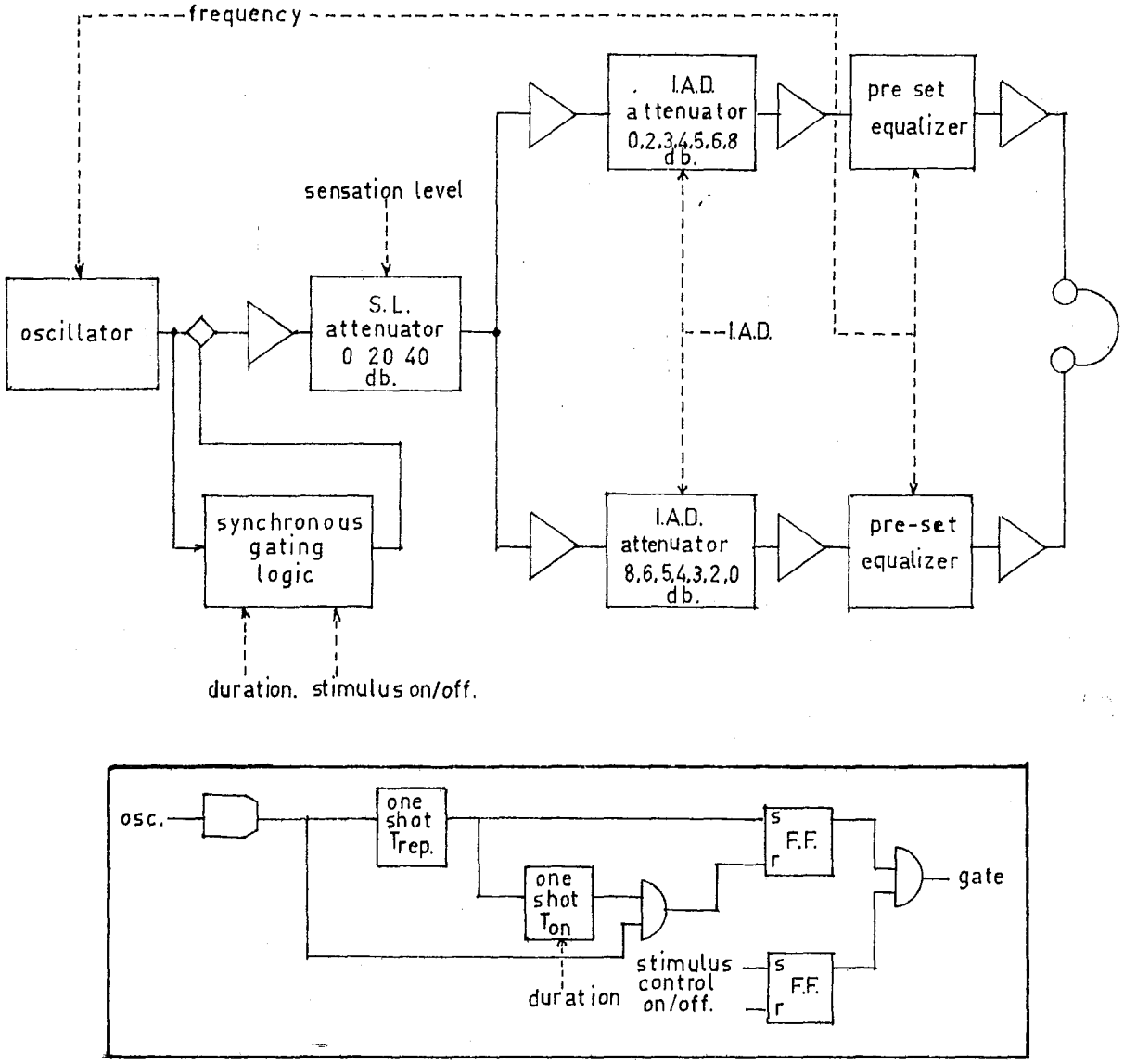


Fig. 4.1 Block diagram of stimulus generator. The apparatus is shown in Plate 4.1 together with the image direction indicator.

Inset: The elements of the synchronous gating logic. See Appendix 5 for definition of the symbols.

The switched attenuators (sensation level and IAD) were constructed from 0.1% tolerance resistors. Pre-set equalizers were used to compensate for the headphone response and the frequency dependence of the audiometric threshold. The buffer amplifiers and headphone drivers were compound emitter followers with an input impedance greater than 4mΩ and an output impedance of less than 50Ω.

The choice of headphone type for this experiment was based upon a series of tests of the repeatability of the measured sound pressure level (SPL) on successive fittings of various headphones to a synthetic rubber mould of the human pinna. It was found that large amplitude standing waves were set up in the aural cavity at frequencies above 1500 Hz for circumaural and supra-aural headphones; at high frequencies the measured SPL was critically dependent upon the headphone position. The most repeatable results were obtained with an insert type of headphone with moulded polystyrene inserts. The units used in the experiment were constructed from dynamic microphone inserts, (Akai DM261) fitted in turned brass cases. These were acoustically coupled to the ear insert by polystyrene tubing (inside diameter 2 mm, length 10 cm) that was loosely filled with acoustic damping material to reduce standing waves.

The image direction indication apparatus was basically the same as used for the experiments in Chapter 3. The subjects generated the pulsating air jet with a rubber

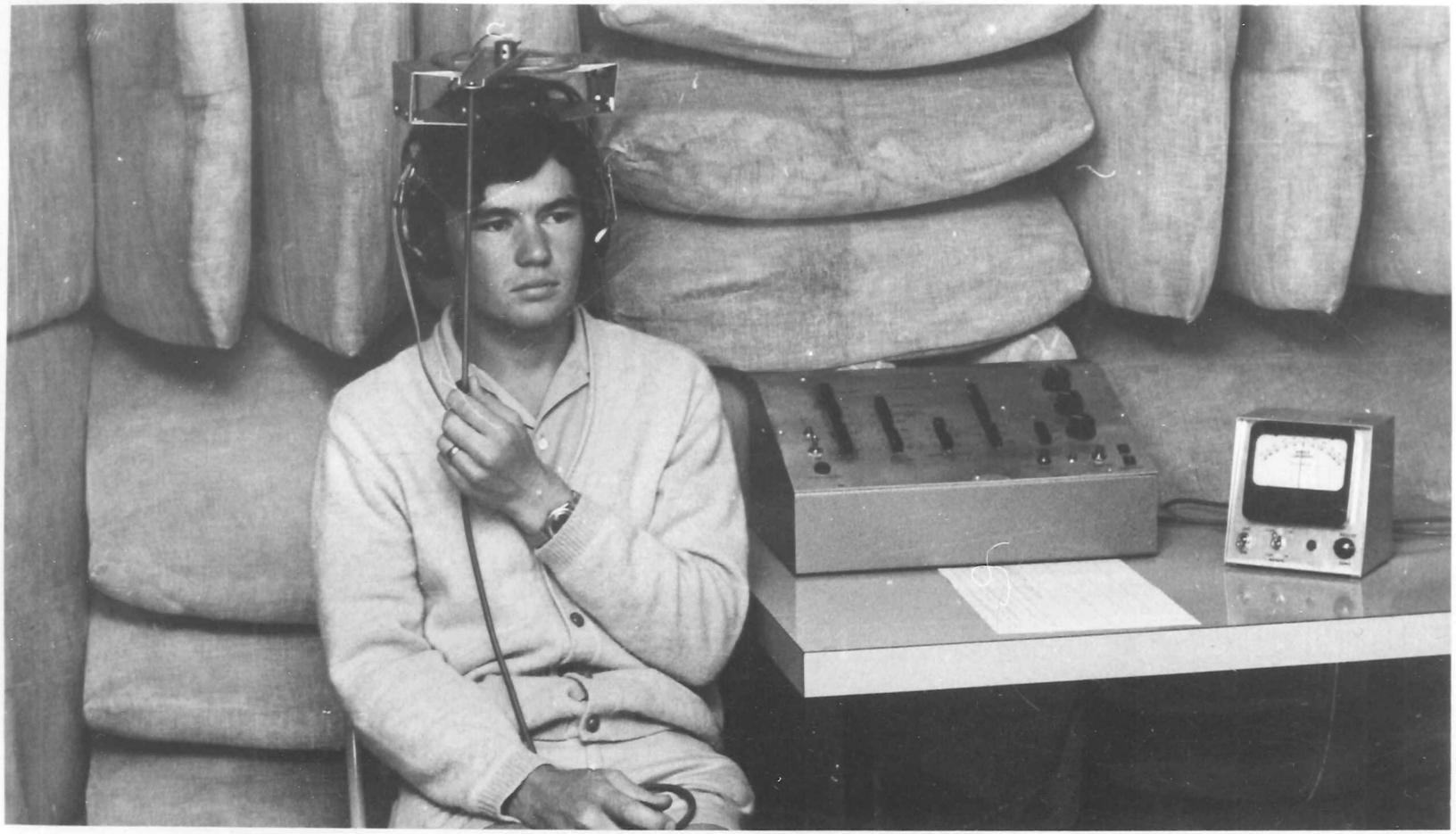


Plate 4.1 Equipment used in dichotic IAD localization experiment. The subject is holding the image direction indicator; the meter at right provides a remote readout.

squeeze bulb. The indicator position was taken from a potentiometer fitted to the headset and was monitored externally by the experimenter.

Calibration: The SPL from the two headphones was calibrated in an ASA 2 cc coupler using a capacitor microphone (GR1551-P1) and sound level meter (GR1551-A). The calibration was done with the ear inserts and connecting tubes attached. The SPL from both channels was checked throughout the experiment and was held to within 1 db of the specified level.

The oscillator frequencies were checked at the start and finish of the experiment and were found to be within 2% of the nominal value specified in Table 4.1.

Experimental Method: At the start of each session the median plane was established by taking a series of readings of the subjects' judgment of 'straight ahead'. The mean of these angles was defined as 0° azimuth. The variance of these estimates was much smaller than that of the localization judgments and was not compensated for in the data analysis.

Subjects were instructed to localise the image with the indicator and to respond verbally when they had made their judgment. The stimulus was then removed and the response was recorded. No attempt was made to hasten responses; the normal response rate during the main experiment was between 100 and 150 per half-hour session. After each response the indicator was returned to the median plane.

Five male university students were used as subjects. All were given audiometric tests to ensure that there were no threshold abnormalities. No subject had previously participated in psychophysical experiments. The duration of the experiment was approximately 20 weeks during which time subjects reported daily, except for weekends. The sessions were postponed if a subject reported any sign of a common cold.

Analysis: Polynomial regression analysis was used to define the form of the localization curve in each cell of the factorial design. It was established that a linear curve described the localization process as

$$\theta' = \beta_{ijk}\Delta I + \alpha_{ijk} + \xi \qquad 4.2$$

where θ' is the azimuthal response, ΔI is the impressed IAD, α_{ijk} and β_{ijk} are the regression coefficients under the ith level of frequency, the jth level of duration and the kth level of sensation level. The variable ξ is zero-mean normal random process. Then the statistical expectation of θ' is

$$\hat{\theta} = E\{\theta'\}$$

$$= \beta_{ijk}\Delta I + \alpha_{ijk}. \qquad 4.3$$

Then β_{ijk} and α_{ijk} are estimators of K and θ_0 respectively in Eq. 4.1.

Cochran's test for homoscedasticity[107] (constancy of variance) in all cells was applied to justify the use of analysis of variance techniques. Least squares estimators of α_{ijk} and β_{ijk} were calculated and the significance of the

slope β_{ijk} was established from the analysis of variance for a linear regression[108].

The slope β_{ijk} of the localization curve was taken as a measure of the cue activity under the prescribed stimulus conditions. The analysis of variance technique was extended and used to estimate the statistical significance of differences in the estimates of β_{ijk} for the main effects and interactions in the factorial design. The derivation of the analysis used is given in Appendix 3.

Results: All subjects were able to localize the auditory image; there were no reports of pure lateralization and a consequent inability to assign an image direction. Similarly there were no reports of the image possessing elevation[22], or being "in the back of the head" as was found in the experiments described in Chapter 3.

As stated above, the polynomial regression analysis verified the suitability of Eq. 4.2 as an empirical description of the localization process. A few of the localization curves showed a marginal improvement in the F-ratio when a 3rd, or even 5th, order polynomial was fitted to the data, but these instances were not widespread. All of the linear regression curves had a highly significant slope ($p < 0.001$); therefore the null hypothesis that all responses were drawn from a common population was rejected.

Fully tabulated results giving the 'least squares' estimators of α_{ijk} and β_{ijk} are given elsewhere[109]. The

variation of the cue activity β_{ijk} over the complete experimental block is summarised in Fig. 4.2a-e where the results for each subject are presented against an abscissa of the stimulus frequency; the pulse duration is represented by different lines within each graph, and the sensation levels divide the results into three plots. It should be noted that in these, and all subsequent graphs, it has been necessary to represent the results for subject ARJ on a different ordinate scale to that used for the other four subjects.

The results of the analysis of variance to test the significance of 'within subject' differences in the estimate β_{ijk} are summarised in Table 4.2. The limits of the available computer storage prevented the inclusion of a 'between subjects' factor in the analysis, however the results of Fig. 4.2a-e suggest that there is a significant main effect difference between subjects.

The only highly significant main effect for all subjects was frequency ($p < 0.005$); the other two main effects were significant in three of the five subjects. The only highly significant interaction for all subjects was frequency x sensation level.

The results of the regression analysis are further summarised in Fig. 4.3a-c where the main effects $\overline{\beta}_{i..}$, $\overline{\beta}_{.j.}$, $\overline{\beta}_{..k}$ are presented for all subjects. These curves should be interpreted with care, for although they indicate trends in the responses the significant interactions are masked.

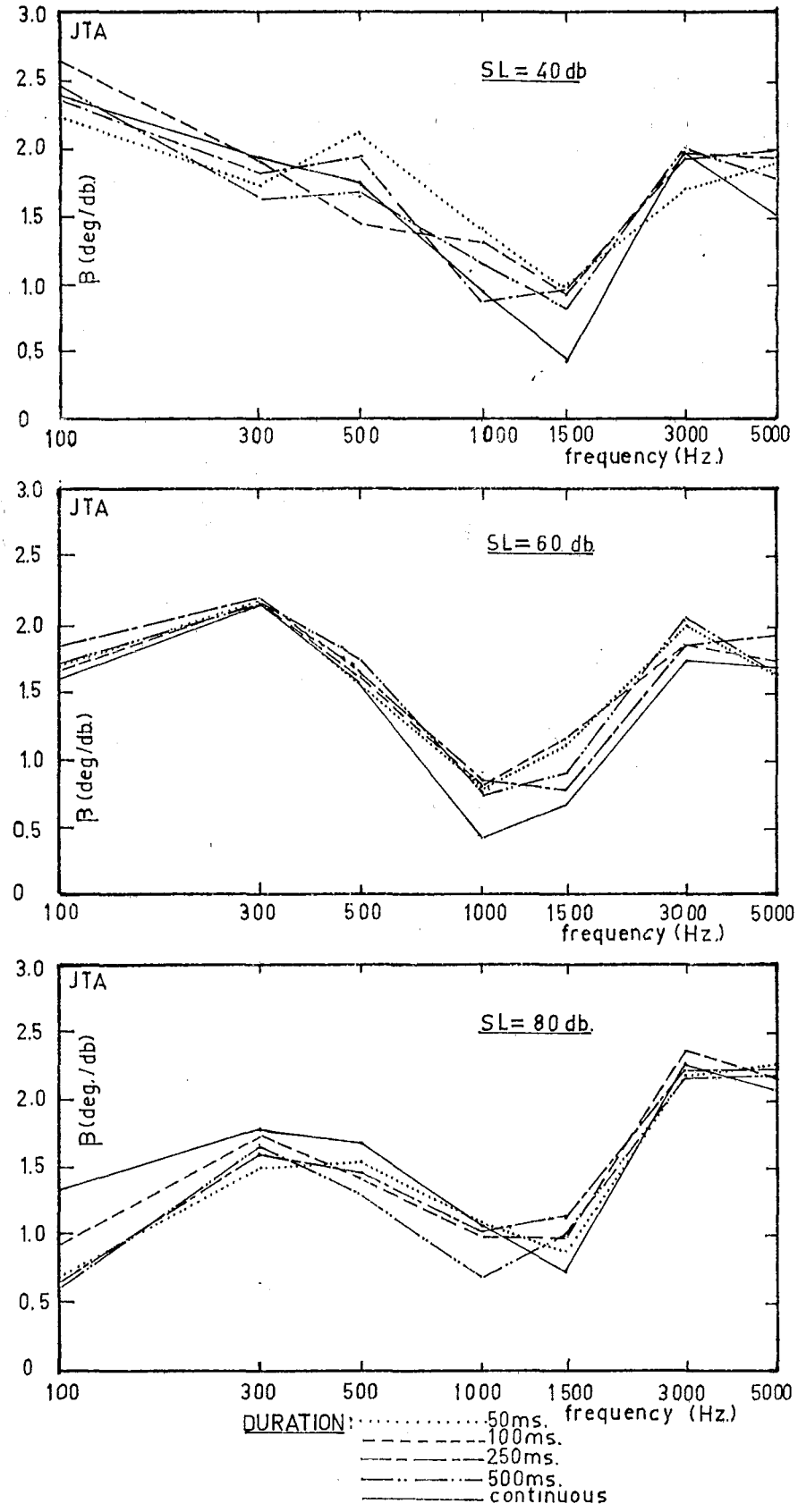


Fig. 4.2a Variation of linear regression coefficient with experimental conditions. Subject JTA.

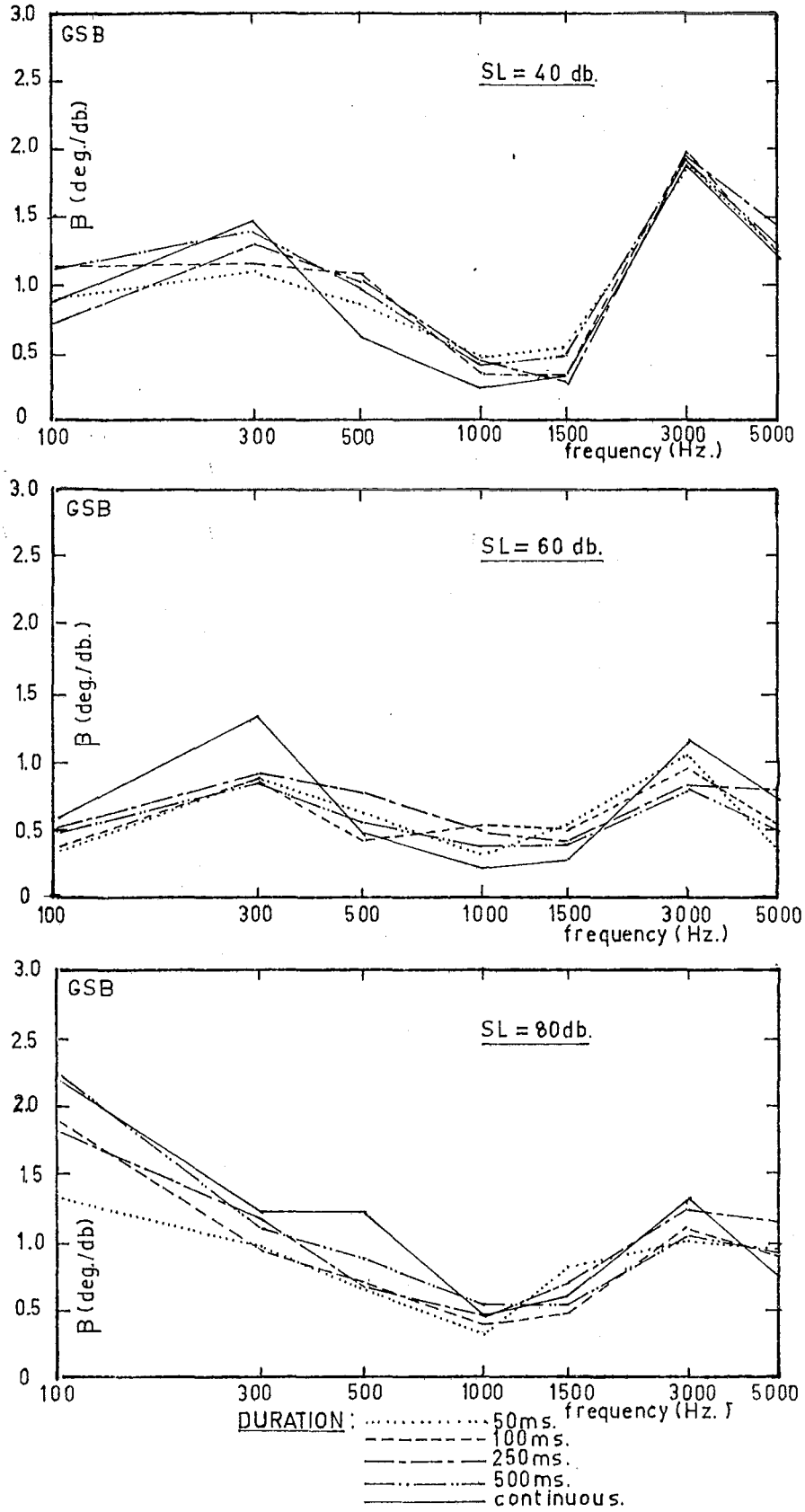


Fig. 4.2b Variation of linear regression coefficient with experimental conditions. Subject GSB.

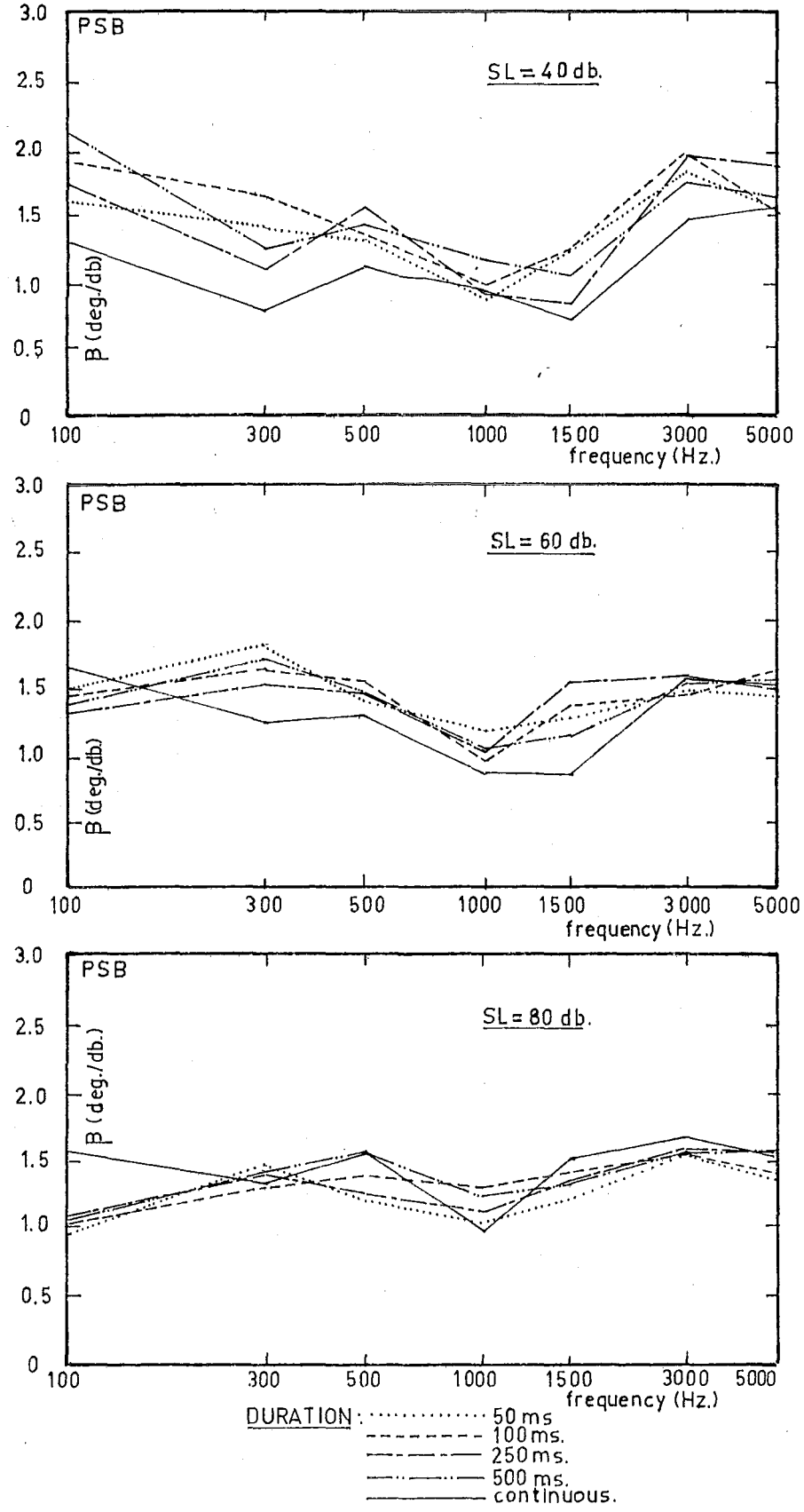


Fig. 4.2c Variation of linear regression coefficient with experimental conditions. Subject PSB.

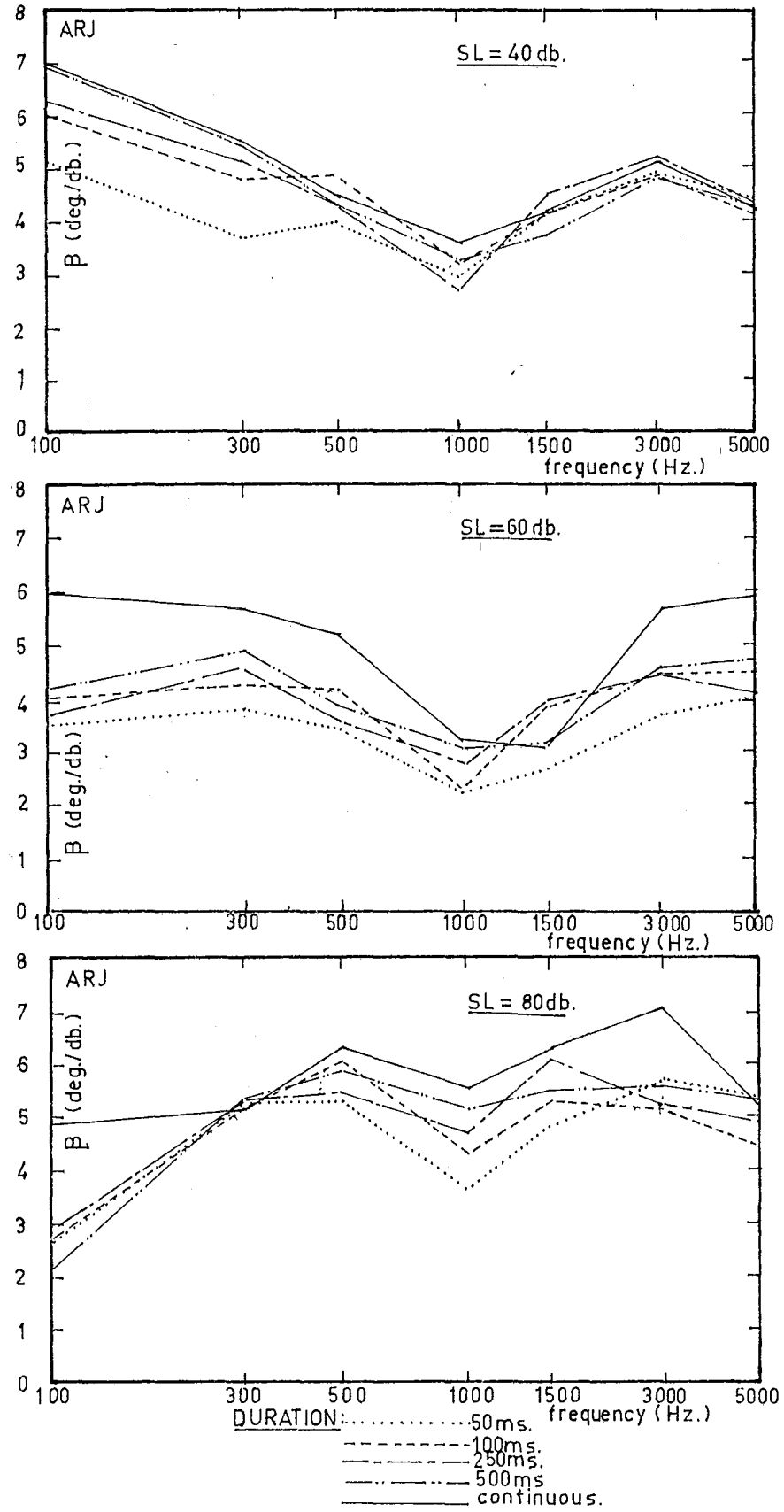


Fig. 4.2d Variation of linear regression coefficient with experimental conditions. Subject ARJ.

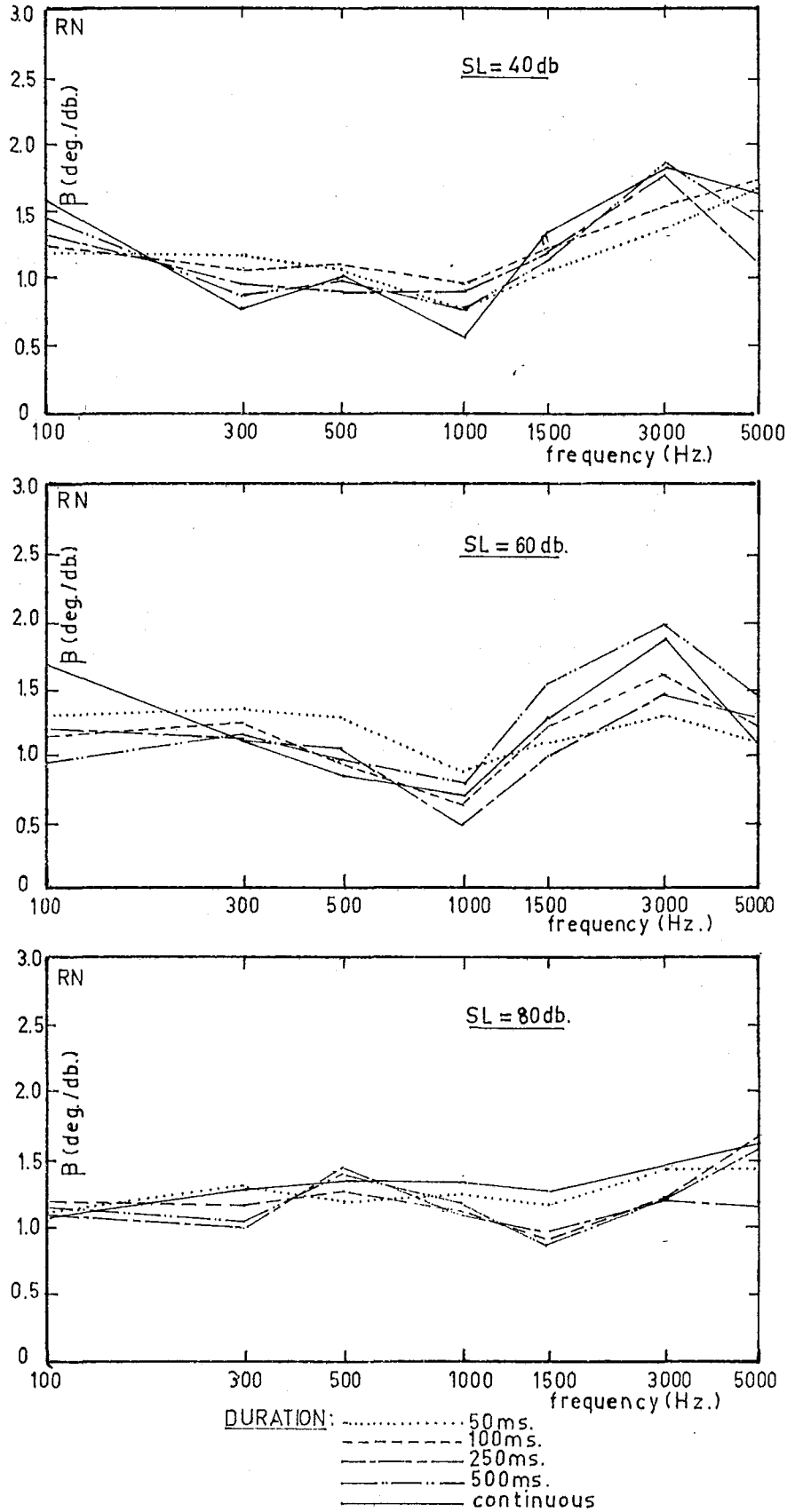


Fig. 4.2e Variation of linear regression coefficient with experimental conditions. Subject RN.

TABLE 4.2 ANALYSIS OF VARIANCE: TO TEST THE SIGNIFICANCE OF DIFFERENCES IN THE REGRESSION SLOPES (see Appendix 3)

Source of Variance	df	F-ratios				
		Subject JTA	Subject PSB	Subject GSB	Subject ARJ	Subject RN
Common Slope	1	11682.383***	12972.707***	5212.297***	16487.949***	6252.941***
Between Group Means	104	49.478***	69.856***	28.381***	77.479***	30.025***
Between Group Slopes:						
Frequency (F)	6	144.070***	38.383***	115.349***	31.607***	29.424***
Duration (D)	4	1.695	5.238***	2.474***	27.441***	0.843
Sensation Level	2	13.522***	0.450	110.801***	59.980***	2.498*
F x D	24	1.493*	1.640*	2.335***	2.140***	0.953
D x SL	8	1.694*	4.919***	2.641**	2.660**	0.996
F x SL	12	32.306***	9.808***	33.112***	28.402***	6.998***
F x SL x D	48	0.806	1.581**	1.351*	1.354*	0.995
		Mean Square				
Error	7140	36.315	25.876	25.654	223.075	40.912

Significance of F ratios:

* $p < 0.1$

** $p < 0.05$

*** $p < 0.005$

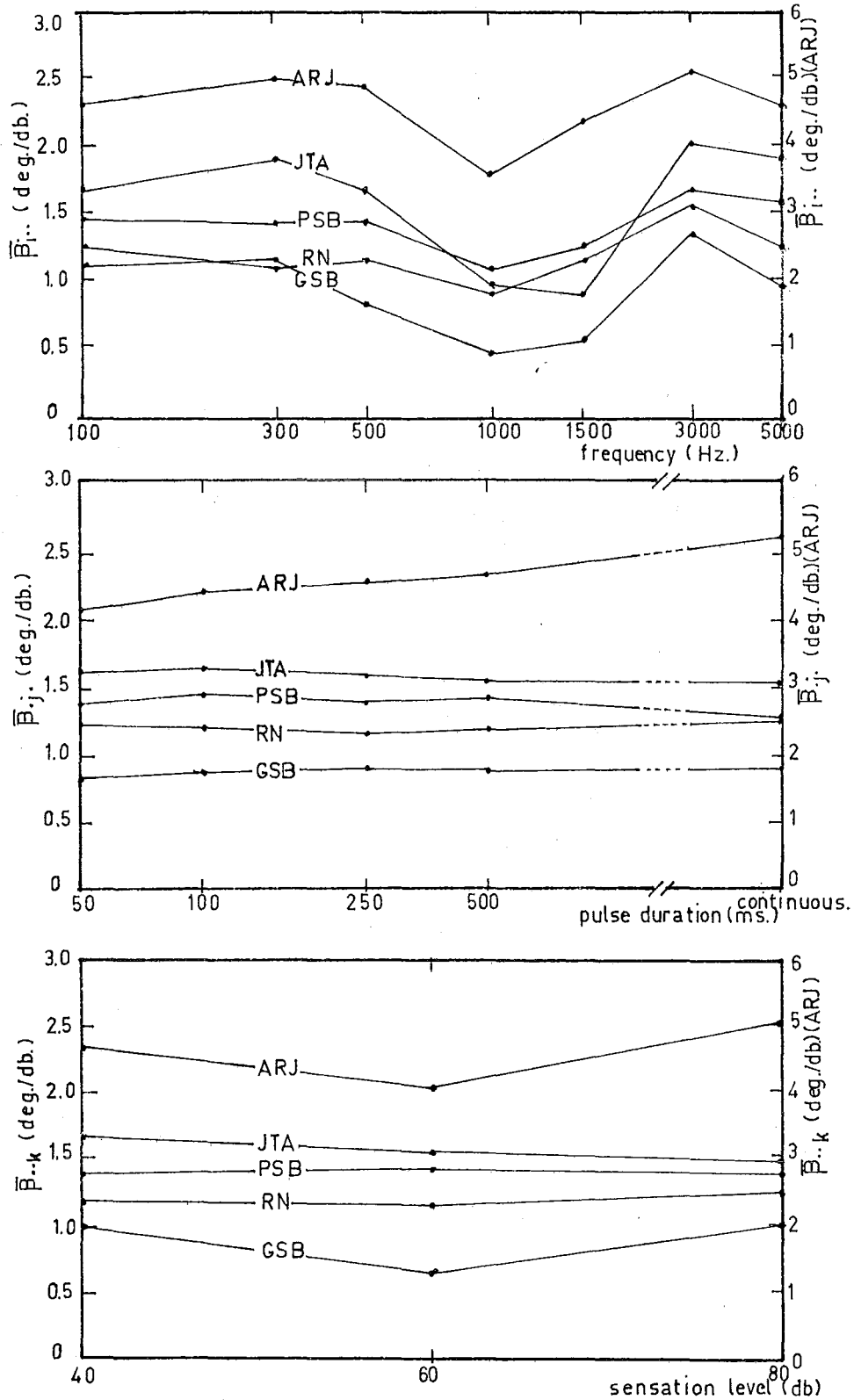


Fig. 4.3a-c Main effects variation of the linear regression coefficients for each of the five subjects
Top Fig 4.3a: Frequency
Centre Fig 4.3b: Pulse duration
Bottom Fig 4.3c: Sensation level

Discussion: The linear relationship between IAD and image azimuth angle, as proposed by Stewart and Hovda[91], was verified. The statistical analysis of the results showed that the constant of proportionality in Eq. 4.1 was dependent upon the stimulus parameters.

The frequency dependence of the localization function is of particular interest because it shows that the two display dimensions (frequency x IAD) are not independent. It is shown in Chapter 5 that the azimuthal dimension of the display can be matched to any localization function described by Eq. 4.3 but that the 'within subject variation' cannot be compensated. Therefore assume that the display is matched to the overall mean localization function, $\bar{\beta}_{...}$, then the expectation of the bias in the estimation of azimuth may be expressed as

$$E\{\theta'-\theta\} = \frac{\bar{\beta}_{i..}-\bar{\beta}_{...}}{\bar{\beta}_{...}}\,\theta$$

The expectation of the fractional error is plotted for all subjects in Fig. 4.4. The similarity in all curves is evident, with a pronounced underestimation of azimuth at frequencies of about 1000 Hz. It is interesting to conjecture that this dip in response magnitude may have been the reason for the report by Stewart[92] that IAD was ineffective over a region of the audio spectrum.

A comparison may be made between an extension of these results and the frequency dependence of the DL in IAD as

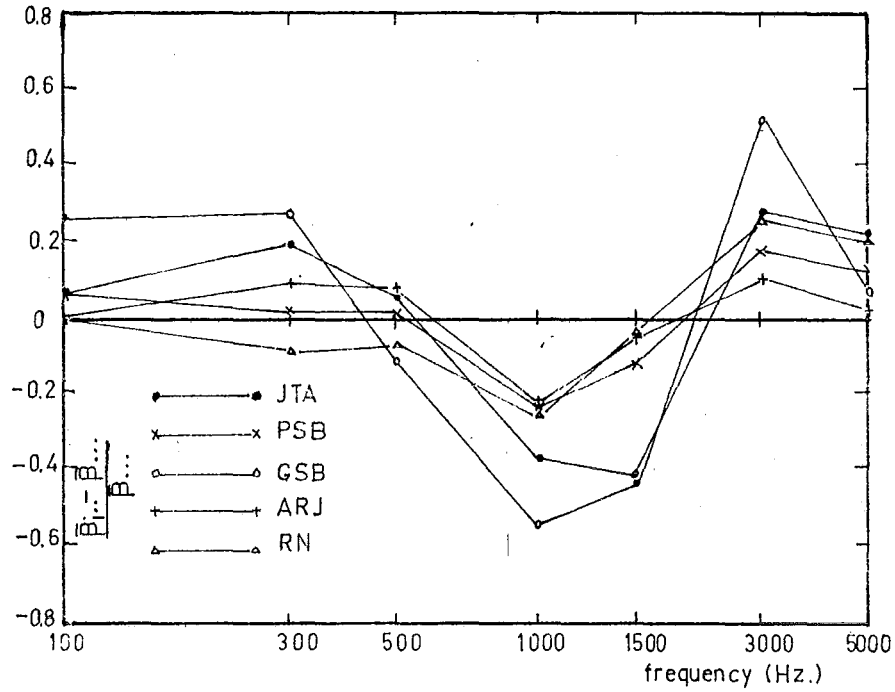


Fig. 4.4 Fractional error in the azimuthal estimator as a function of frequency for a display based on the overall mean localization function.

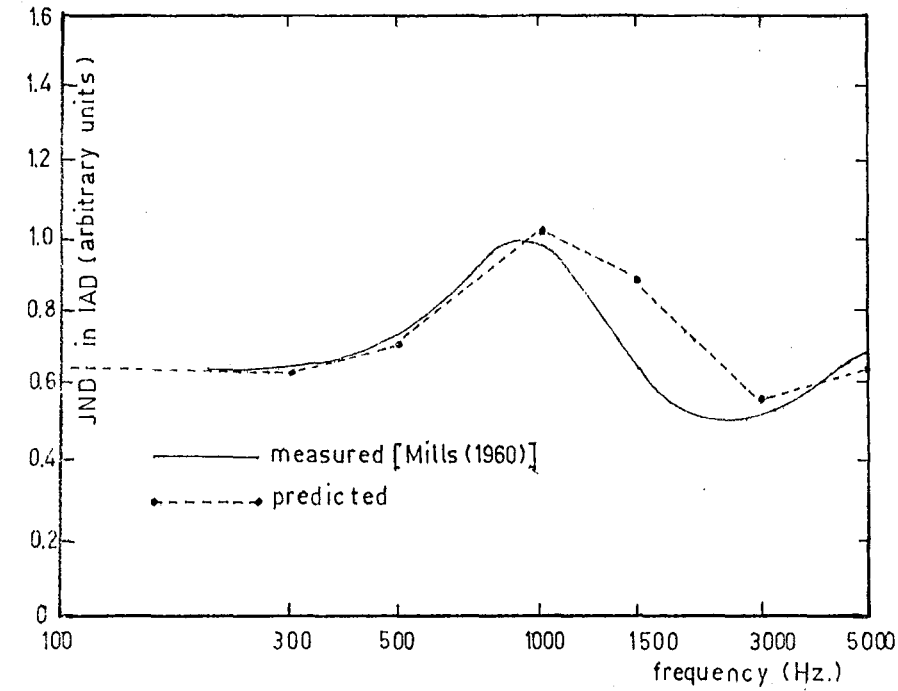


Fig. 4.5 Comparison of the predicted form of the frequency dependence of the just noticeable IAD with the results of Mills ref.48.

reported by Mills[48]. If it is hypothesized that the DL in a two-choice experiment results from a constant angular displacement of the auditory image, then the DL will be inversely proportional to the slope of the localization function. In Fig. 4.5 the normalised results of Mills are compared with those predicted on the above hypothesis using the results of the present experiment. The two curves are similar in form.

The other two main effects (pulse duration and sensation level), although both statistically significant in three of the five subjects, show no general trends and therefore do not enable specific conclusions to be drawn. There is no evidence of a time constant of the order of 700 msecs in the processing of IAD as reported by Tobias and Zerlin[51] on the DL for ITD.

The only significant two way interaction, frequency x sensation level, occurred principally in the low frequency region (Fig. 4.2a-e). At these frequencies four of the five subjects showed a decrease in $\overline{\beta}_{i \cdot k}$ with increasing stimulus intensity. This could be interpreted as support for the hypothesis that IAD is processed in the latency of peripheral neurons[96,110]; for the IAD-ITD trading function (typified by Fig. 2.3) is decreased in magnitude with sensation level.

The magnitude of $\overline{\beta}_{\cdot jk}$ is in general agreement with results published by earlier workers[46]. In Fig. 4.6 the mean

value of $\overline{\beta}_{i..}$ is compared with the required localization function to use the IAD existing at the outer ear in free-field listening. The latter curve was derived from the data of Sivian and White[111]. Below approximately 5 kHz there is considerable difference between the two curves; it would appear that of the five subjects who took part in the experiment, only ARJ could use IAD in the free-field localization of tones. This finding supports the conclusion of Feddersen et al[46] who also found that the IAD localization function was not sufficient for free-field localization below 5 kHz.

4.4 Estimation of Population Parameters

Whereas the previous experiment examined the 'within subject' variation of the IAD localization function in detail, the experiment described here was designed to estimate the range of variation in the localization function throughout the population. A larger sample, containing both blind and sighted subjects, was used and the distribution of the slopes of the localization curve at three frequencies was derived. The blind and sighted samples were compared in order to justify the use of data from sighted subjects, such as in Section 4.3, to specify the auditory display for a mobility aid for the blind.

The second part of the experiment was designed to evaluate the centring method as a means of estimating the

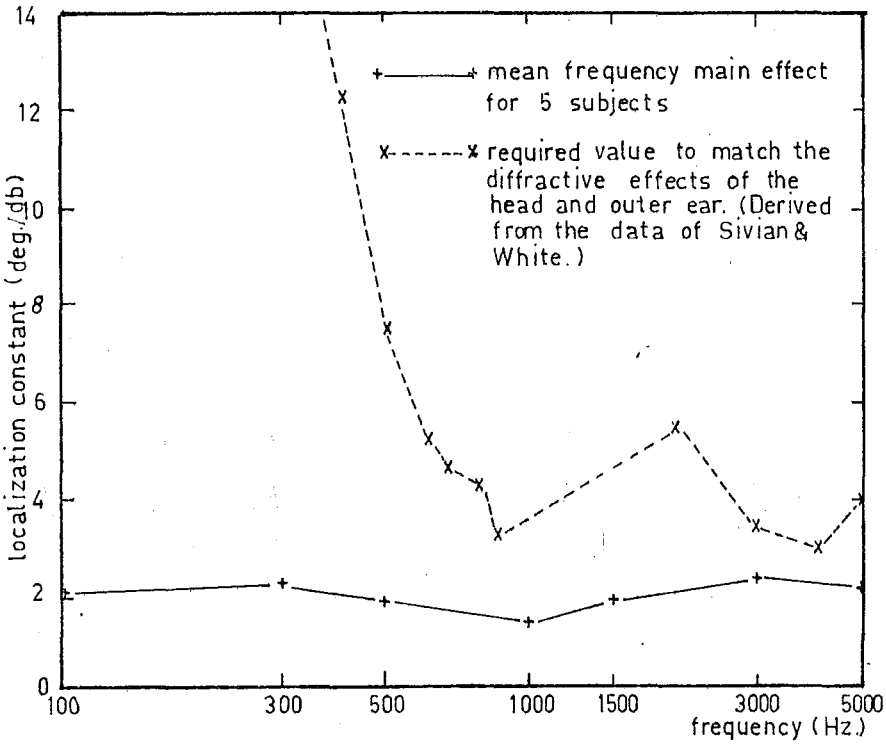


Fig. 4.6 Comparison of the overall mean frequency dependence of the linear regression coefficient with that necessary for free field localization with IAD derived from diffraction by the head and pinnae.

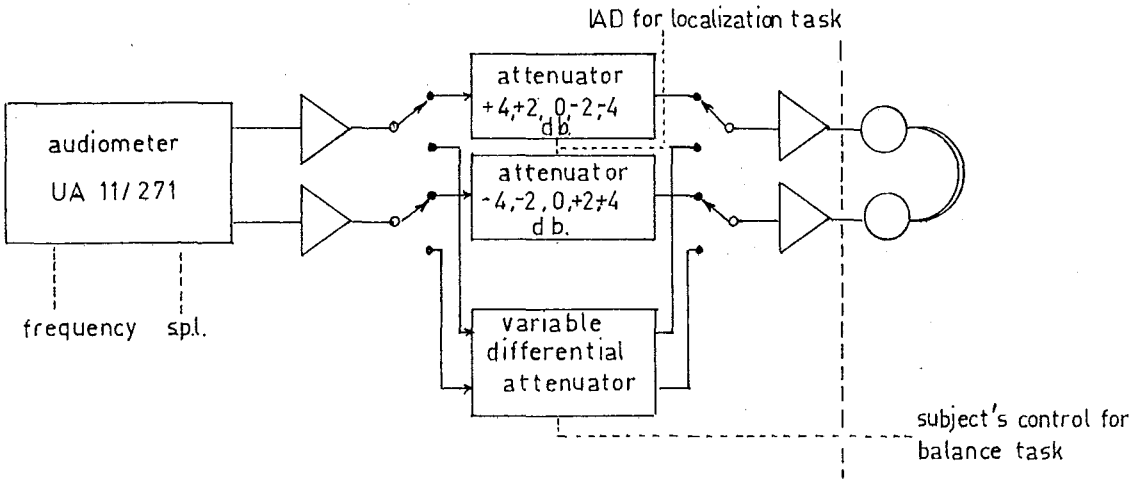


Fig. 4.7 Experimental apparatus for localization and centring experiment.

intercept α of the localization curve. The value of IAD required to centre the image was correlated with that predicted from the localization curve.

Experimental Apparatus: A clinical audiometer (British Hearing Aids (NZ) Ltd, Model UAL 11/271) was used as the tone generator. An additional unit containing stepped and continuously variable differential attenuators, and buffer amplifiers (Fig. 4.7) was used to impress the IAD on to the two channels. The headphones were Telephonics type TDH39 with MX-41/AR ear cushions, calibrated for use with the audiometer. The image indicating apparatus was the same as used in the previous experiment.

Experimental Method: Subjects were chosen to cover a wide range of ages (from teenagers to pensioners) of both sexes. Of the 24 subjects, 16 were sighted and 8 were blind. Each subject attended for three half-hour sessions on consecutive week days. The first session was used for audiometric threshold tests, loudness balance tests, and familiarization with the localization and centring tasks. The data from the localization and centring experiments were recorded on the second and third days.

The localization curves were measured at frequencies of 500, 1500, 3000 Hz with the following parameters:

Sensation level : 60 db

Pulse duration : 1 sec.

Repetition period: 2 sec.

Envelope rise-fall time: 50 msecs.

The localization experiment was conducted in a similar manner to that used in Section 4.3. The median plane was estimated from a series of judgments of "straight ahead" before each session. The stimulus presentation order was completely randomised, with a total of 15 observations at each IAD level (+8, 0, -8 db re left ear). Least squares estimators were used to derive the coefficients of the localization function.

The centring experiment consisted of 10 judgments of the IAD required to centre the image at each frequency. At each judgment the subject aligned the image to the "straight ahead" position with the continuously variable attenuator. The initial position of the attenuator was randomly set by the experimenter between judgments.

No attempt was made to limit the time taken for responses in either part of the experiment. The stimulus was removed after each judgment to remove possible cues from changes.

Results: The sample distributions of the least squares regression slopes are summarised in Fig. 4.8 as histograms and fractile diagrams.

The hypothesis of a common parent population for the blind and sighted samples was tested with a distribution-free statistic of the Kolmogorov-Smirnov type[112,113], based upon the maximum difference between the two sample cumulative frequency distributions. Further statistics, under a hypo-

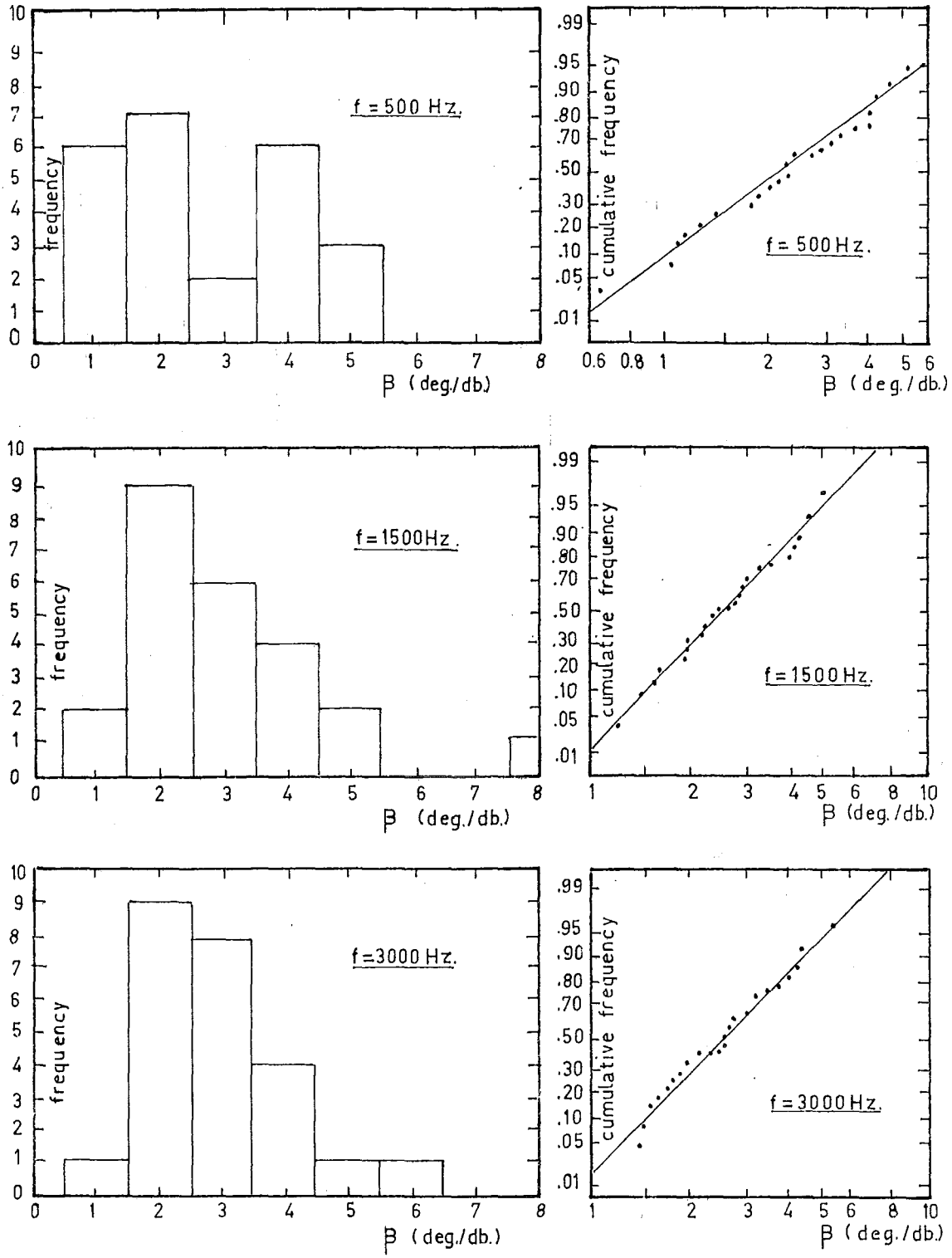


Fig. 4.8 Distribution of the linear regression coefficient.

thesis of normality, (the t and F statistics) were computed to test null hypotheses of a common mean, and a common variance. The results of these tests are summarised in Table 4.3; it was not possible to reject any of the null hypotheses so there was no evidence of any difference between the blind and sighted samples.

TABLE 4.3 SUMMARY OF COMPARISONS BETWEEN BLIND AND SIGHTED SAMPLES

Assumed Distribution:	Distribution-Free		Normal		Normal	
Test Hypothesis:	$H_0:(F_1(x)=F_2(x)$		$H_1: (\sigma_1 = \sigma_2)$		$H_2:(\mu_1 = \mu_2)$	
Statistic:	Kolmogorov-Smirnov		F-ratio		Student's t	
Frequency	$d(n_1n_2)$*	P	$F(v_1,v_2)$*	P	t(v)*	P
500	.25(8,16)	0.5	1.69(15,7)	0.1	.15(22)	0.5
1500	.31(8,16)	0.4	2.24(15,7)	0.1	.16(22)	0.5
3000	.31(8,16)	0.4	1.11(7,15)	0.1	.43(22)	0.5

*Quantities in parentheses are sample size (Kolmogorov-Smirnov) or degrees of freedom (F,t).

The small sample size precludes the reliable use of the χ^2 statistic in estimating the goodness of fit of the data to a hypothetical distribution. Instead, the Kolmogorov statistic[112], which is based upon the cumulative distribution,

was used to test null hypotheses of a normal distribution, or alternatively a logarithmic-normal distribution. The latter test was prompted by an apparent curvature in the cumulative distributions (Fig. 4.8); this curvature was eliminated when the slope β was plotted to a logarithmic scale. It was not possible to reject either hypothesis; $p > 0.4$ for all frequencies on both tests. In the absence of further evidence the simpler hypothesis of a truncated normal distribution for β at each frequency was assumed.

The estimated population parameters are summarised in Table 4.4. An analysis of variance in a one-way classification showed no significant differences between the population means for the three frequencies ($p > 0.5$). This does not mean that there were no "within subject" differences in the β for individual subjects, it simply means that trends in individual differences were masked by the averaging process.

TABLE 4.4 ESTIMATED POPULATION PARAMETERS

Frequency	Mean ($\bar{\beta}$)	Standard Dev. (σ)	90% Confidence Range
500	2.78	1.41	0.46 - 5.10
1500	3.05	1.45	0.66 - 5.44
3000	2.79	1.21	0.80 - 4.78

The mean values of β at each of the three frequencies were compared with the corresponding frequency main effects $\overline{\beta}_{i\circ\circ}$ from the experiment in Section 4.3. The same set of statistics (Kolmogorov-Smirnov, F, t) as were used in the blind-sighted comparison were computed and are presented in Table 4.5. The Kolmogorov-Smirnov statistic requires rejection of the hypothesis of a common distribution at 1500 Hz ($0.05 > p > 0.01$).

TABLE 4.5 COMPARISON BETWEEN RESULTS OF EXPTS.

Assumed Distribution:	Distribution Invariant		Normal		Normal	
Null Hypothesis:	$H_0:(F_1(x)=F_2(x)$		$H_0:(\sigma_1=\sigma_2)$		$H_0(\mu_1 = \mu_2)$	
Statistic:	Kolmogorov-Smirnov		F-ratio		Student's t	
Frequency	$d(n_1 m n_2)$*	P	$F(v_1, v_2)$*	P	t(v)*	P
500	.56(24,5)	0.05	1.39(4,23)	0.1	1.10(27)	0.2
1500	0.76(24,5)	0.01	1.14(4,23)	0.1	1.91(27)	0.05
3000	.46(24,5)	0.1	1.67(4,23)	0.1	0.73(27)	0.4

*Quantities in parentheses are sample sizes (K-S statistic), and degrees of freedom (F, t statistics).

This requires the rejection of the hypothesis of common means on the t-test ($0.1 > p > 0.05$) and acceptance of the

alternative hypothesis that at 1500 Hz the mean β as estimated from this experiment is greater than that from the previous experiment.

The data from the centring experiment showed that there was little correlation between estimates of θ_0 in Eq. 4.1 from the localization intercept and centring IAD. Although there was a small positive correlation ($r = 0.101$), this was too small to indicate a reliable testing procedure.

Discussion of Results: The blind and sighted samples were compared to determine whether the reduced sensory feedback from the environment in the case of the blind had affected the formation and organization of the auditory space. There was no evidence in the results of this experiment of any difference between the blind and sighted subjects, therefore there was no evidence suggesting that the design of a mobility aid should not be based upon data obtained from sighted subjects.

The statistical frequency distributions of the slopes of the localization curves (Fig. 4.8) show that a binaural echolocation system must be capable of being matched to a wide range of individual localization functions. The 90% confidence intervals in Table 4.4, which were derived by assuming normality of the distributions, show that this range will be greater than a factor of approximately 10:1. The method of matching the system to an individual is described in Chapter 5.

The reasons for the difference in the distributions of β between this and the previous experiment at 1500 Hz are not understood. There was no significant interaction with sensation level or duration that would explain the difference. The only difference in the stimuli was in the envelope rise times, and although this is accepted as influencing responses[114], Perrott[115] found a decrease in the minimum audible angle (MAA) with increasing rise time at 2000 Hz; which is the opposite to the results observed here.

The sample size was found to be too small to allow generalizations as to the form of the statistical distributions of β. Although a normal distribution was assumed, a logarithmic normal distribution could not be rejected. The latter distribution eliminated the apparent curvature in the cumulative distributions when plotted on probability paper, and has the advantage of eliminating the finite probability that $\beta < 0$ (i.e. the sound being localized toward the ear with the lowest intensity) asserted by an untruncated normal distribution. An experiment using a much larger sample will be necessary to determine the form of the distributions.

4.5 Localization with Time-Varying IAD

The exploratory experiment reported here was prompted by the observation that although differences in the binaural envelope structures of the mobility aid were observed (see Fig. 3.2), no subject had reported any movement of the image

caused by the time-varying IAD. The experiment demonstrated a new phenomenon that cannot be explained by simple linear processing of IAD.

Experimental Method: The binaural stimuli consisted of equal energy, linearly amplitude modulated pulsed tones, in which the direction of modulation in the two channels was opposite. Four subjects listened to the tones on two occasions and reported the perceptual effect of the modulation parameters.

Apparatus: The generation and control of the stimuli was implemented through an EAI 580 analogue computer patched as shown in Fig. 4.9a. The two modulation functions were generated from an astable pair of integrators with constant voltage inputs. Independent control of the modulation amplitude, rate, and 'off' period was provided. The envelopes of the stimuli are shown in Fig. 4.9b and were described by the equations

$$M_i(t_n) = \tfrac{1}{2}(1+(-1)^i a) - (-1)^i b t_n \qquad 0 < t_n < \frac{2a}{b}$$

$$= 0 \qquad \frac{2a}{b} < t_n < T_s$$

where $t_n = t - nT_s$, $n = 1, 2, 3 \cdots$,

$M_i(t_n)$ is the envelope in the ith channel ($i = 1,2$), $2a$ is the peak-to-peak modulation amplitude ($0 < a < 0.5$), b is the modulation rate, $T_s = \frac{2a}{b} + t_{off}$ is the repetition period.

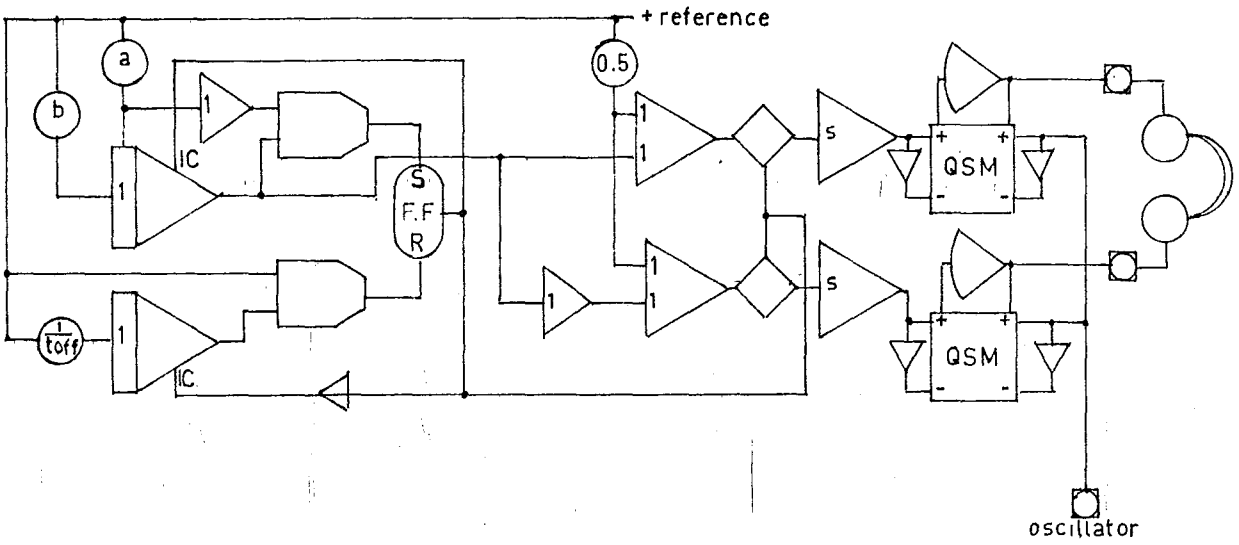


Fig. 4.9a Analogue computer patching to generate time-varying IAD (see Appendix 5 for analogue computer symbols).

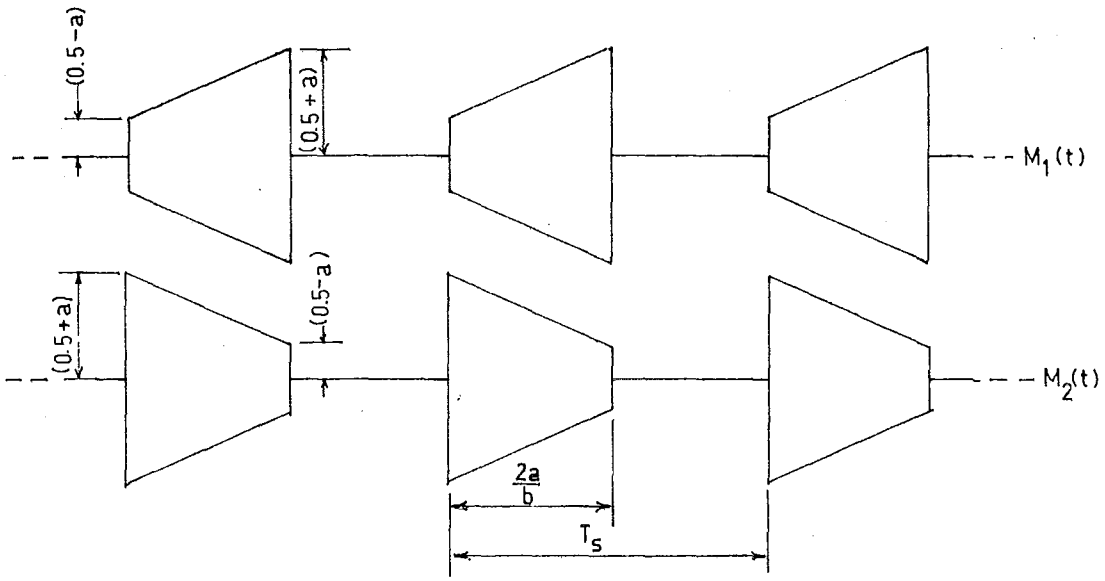


Fig. 4.9b Dichotic signal envelopes for time-varying IAD experiment.

The amplitude modulation of the externally generated tone (Hewlett Packard 3300A function generator) was accomplished through quarter-square multipliers with a static accuracy of 0.01%. The sensation level in the two headphones was 60 db at the mean signal amplitude, i.e. when $t_n = \frac{a}{b}$.

Results: Only qualitative results were obtained. The experience of all subjects was similar. For isolated pulses ($T_s >> \frac{2a}{b}$) the report was of a laterally moving image, with the apparent motion toward the ear stimulated by $M_2(t)$. The magnitude of this perceived movement was dependent upon the modulation amplitude. Motion could be perceived until the duration was reduced to approximately 30 msecs. However, as t_{off} was reduced the magnitude of the apparent movement was reduced, until some critical point was reached when the image was stationary. At low modulation amplitudes the image was then localised as straight ahead, but at higher levels (e.g. a = 0.4) of modulation the image was localised toward the side stimulated by $M_1(t)$ (decreasing amplitude). This effect was observed over a range of pulse durations from 30 msec to 500 msec.

Discussion: This experiment showed that for tonal stimuli the intra-pulse structure is not the only factor affecting its localization; the relation of the pulse to others in the stimulus sequence is important. The 'repetition' effect described above is another example demonstrating the role of the signal onset in sound localization, because although the

stimuli contained equal energy the image was localized toward the direction indicated by the initial IAD.

The magnitude of the differences in the modulation functions used in this experiment were much greater than those observed in the echolocation system. It is considered that the waveform parameters (duration and repetition period) in the echolocation display were such that the 'repetition' effect prevented motion of the image, but that the differences in the envelope modulation were not sufficient to cause a noticeable image displacement.

This effect should be investigated in more detail with quantitative measures of the image displacement and the repetition period required to produce a stationary image.

C H A P T E R 5

THE SPECIFICATION AND PERFORMANCE OF THE TWO DIMENSIONAL BINAURAL DISPLAY

5.1 Introduction

In this and the following chapter, attention is diverted from the psychophysics of binaural hearing to the specification of the auditory display of the modified echolocation system described in Chapter 3. This chapter is concerned with the derivation and performance of the azimuthal dimension of the display, using the results of the studies in dichotic localization with IAD discussed in Chapter 4.

In the absence of a complete knowledge of the effects of stimulus parameters upon binaural phenomena the following simplifying assumptions have been made upon the display form:

i It is assumed that the binaural stimuli are repetitive bursts of sinusoids of the same frequency with no amplitude modulation impressed upon the envelopes.

ii It is further assumed that there is no interaural phase difference; or alternatively that phase differences do not influence the estimation of azimuth.

The relevance of the first assumption is discussed in Chapter 6; the second is based upon the results of the experimental work in Chapter 3 where it was shown that a small IFD in the

stimuli destroyed the ability to recognize or use phase or time differences as a display dimension.

5.2 The Idealized Azimuthal Variation of Stimulus Amplitude in the Two Echolocation Channels

The empirical description of a subject's response θ' in a dichotic localization experiment (Eq. 4.2) may be rewritten in the form

$$\theta' = K(s) \log_e \left(\frac{A_1}{A_2}\right) + \theta_0(s) + \xi \qquad 5.1$$

where $K(s)$ and $\theta_0(s)$ are descriptions of the localization function for the given parameter set s, A_1 and A_2 are the amplitudes of the sinusoidal stimuli in the two channels and ξ is a zero-mean normal random process. Then using this notation an azimuth estimator $\hat{\theta}$ may be defined as the expectation of θ', i.e.

$$\hat{\theta} = E\{\theta'\}$$

$$= K(s) \log_e \left(\frac{A_1}{A_2}\right) + \theta_0(s) \qquad 5.2$$

Under the assumptions stated in Section 5.1 the auditory stimulus from each channel of the echolocation system may be written

$$f_i(t|R,\theta,\phi) = A_i(\theta,\phi)D(R) \cos(\omega(R)t+\eta_i) \qquad 0 < t_n < T_a \qquad 5.3$$

$$= 0 \qquad T_a < t_n < T_s$$

for $i = 1,2$; where $t_n = t-nT_s$ $(n = 1,2,3,\cdots)$, R, θ, ϕ are the range, azimuth and elevation of the object, $A_i(\theta,\phi)$ defines the angular variation of amplitude in the ith channel $(i = 1,2)$, $D(R)$ defines the range dependence of the amplitude, $\omega(R)$ is the frequency domain range coding and η_i are phase angles.

The present problem is to define the angular variation $A_i(\theta,0)$ so that $\hat{\theta} \equiv \theta$. From the psychophysics of auditory spatial perception it is known that the elevation ϕ cannot be estimated from simple binaural differences[24] with stimuli such as described by Eq. 5.3, therefore ϕ must be considered as an environmental parameter that will influence the estimation of azimuth in the two-dimension display. It is therefore convenient to initially define the system with the object lying upon the equatorial plane and write

$$A_i(\theta) \equiv A_i(\theta,0). \qquad 5.4$$

In accordance with the modified binaural display proposed in Section 3.9, the azimuthal response functions may be written

$$A_i(\theta) = C_i G(\theta) F(\theta+(-1)^i\alpha) \qquad 5.5$$

for $i = 1,2$ where C_i are constants, $G(\theta)$ is an arbitrary illumination function, and α is the angle by which the response $F(\theta)$ has been displaced from the median plane. Substitution of $A_i(\theta)$ for A_i in the localization function

(Eq. 5.2) gives the expression for the estimator

$$\hat{\theta} = K(s) \log_e \left(\frac{CF(\theta-\alpha)}{F(\theta+\alpha)} \right) + \theta_0(s) \qquad 5.6$$

where $C = \frac{C_1}{C_2}$. Since the parameter set s is fixed, it is convenient to write $K = K(s)$ and $\theta_0 = \theta_0(s)$, so that by rearranging Eq. 5.6 the functional relationship to be satisfied by $F(\theta)$ is

$$F(\theta+\alpha) = C \exp\left(\frac{\theta_0}{K}\right) F(\theta-\alpha) \exp\left(-\frac{\theta}{K}\right) \qquad 5.7$$

which if

$$C_1 = C_2 \exp\left(-\frac{\theta_0}{K}\right) \qquad 5.8$$

is satisfied by

$$F(\theta) = \exp\left(-\frac{\theta^2}{4\alpha K}\right) \qquad 5.9$$

which upon substitution into Eq. 5.5 gives the required azimuthal variation of amplitude in the two channels as

$$A_i(\theta) = C_i G(\theta) \exp\left(-\frac{(\theta+(-1)^i\alpha)^2}{4\alpha K}\right). \qquad 5.10$$

The idealized azimuthal response functions of the two receiving apertures are therefore Gaussian in form and are shown in Fig. 5.1 (assuming $G(\theta)$ to be constant). Eqs 5.8 and 5.10 show that the system parameters may be matched to the localization function of an individual by

i Altering the channel gains (C_1 and C_2) according to Eq. 5.8 to compensate for any offset θ_0.

ii Choosing the value of α to match the display to the localization constant K. This is investigated further in Section 5.3.

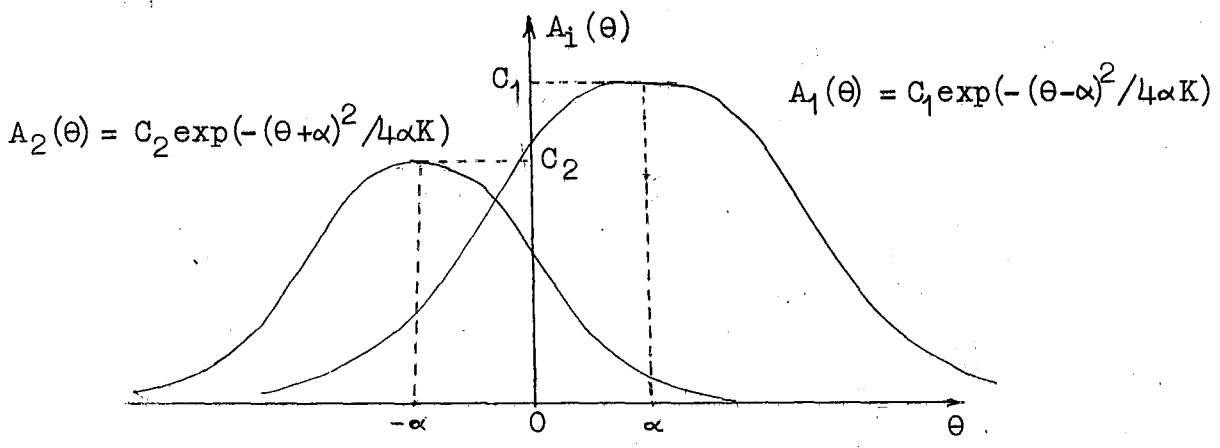


Fig. 5.1 Idealised Variation of Amplitude in the Two Echo-location Channels

(The illumination function $G(\theta)$ is assumed to be constant)

In some of the following analyses it has been convenient to normalize the system parameters with respect to α as follows; write:

$$\hat{\theta}_n = \frac{\hat{\theta}}{\alpha}; \quad \theta_n = \frac{\theta}{\alpha}; \quad K_n = \frac{K}{\alpha}; \quad \theta_{on} = \frac{\theta_0}{\alpha} \qquad 5.11$$

so that normalised functions $A_{ni}(\theta_n)$ and $G_n(\theta_n)$ may be used to define $\hat{\theta}_n$ as

$$\hat{\theta}_n = K_n \log_e \left(\frac{A_{n1}(\theta_n)}{A_{n2}(\theta_n)} \right) + \theta_{on} \qquad 5.12$$

where $A_{ni}(\theta_n) = C_i G_n(\theta_n) \exp\left(- \frac{(\theta_n + (-1)^i)^2}{4K_n} \right)$ 5.13

In presenting results the localization constant is expressed

in units of degrees/decibel, i.e.

$$k = 8.686K; \qquad k_n = 8.686K_n \tag{5.14}$$

where k, k_n are in the modified units.

5.3 The Relationship between the Beamwidth of the Polar Response Function $F(\theta)$ and the Angular Displacement α

Let the polar response $F(\theta)$ from Eq. 5.9 be written

$$F(\theta) = \exp(-\gamma\theta^2) \tag{5.15}$$

where γ is a constant defining the beamwidth, then substitution into Eq. 5.6 gives (assuming Eq. 5.8 is satisfied)

$$\hat{\theta} = 4K\alpha\gamma\theta \tag{5.16}$$

so that the required condition that $\hat{\theta} \equiv \theta$ is given by

$$\alpha = \frac{1}{4K\gamma}. \tag{5.17}$$

Since $\exp(-x) = 0.5$ when $x = 0.693$, the half amplitude (6db) beamwidth of $F(\theta)$, $\psi_{\frac{1}{2}}$, will be

$$\psi_{\frac{1}{2}} = 2\sqrt{\frac{0.693}{\gamma}} \tag{5.18}$$

so that for $\hat{\theta} \equiv \theta$

$$\alpha = \frac{\psi_{\frac{1}{2}}^2}{11.1K} \tag{5.19}$$

or

$$\alpha = \frac{\psi_{\frac{1}{2}}^2}{96k} \tag{5.20}$$

These relationships are important in practice because they enable a system using apertures of known beamwidth to be matched to the auditory characteristics of any individual.

The family of hyperbolae generated by Eq. 5.20 is summarised in Fig. 5.2. The dotted lines in the figure define operating conditions corresponding to values of $k_n = \frac{k}{\alpha}$ used as examples in subsequent analyses.

5.4 The Effect of Object Elevation upon the Azimuth Estimator $\hat{\theta}$

The ratio form of the estimator (Eq. 5.6) implies that for $\hat{\theta}$ to be independent of elevation ϕ, the elevational dependence of the angular response function $A_i(\theta,\phi)$ in Eq.5.3 should be separable, so that Eq. 5.5 becomes

$$A_i(\theta,\phi) = C_i G(\theta) B(\phi) F(\theta+(-1)^i\alpha) \qquad 5.21$$

where $B(\phi)$ completely describes the elevational variation of the response of the receiving apertures.

However, present technology has limited the type of aperture to a plane circular disc (see Chapter 6) in which the amplitude response is a function only of the angle to the normal to the aperture plane. Consider the geometry shown in Fig. 5.3. Let the aperture be in the x-y plane as shown so that the angle χ defining the response is

$$\chi = \cos^{-1}(\cos\theta\cos\phi) \qquad 5.22$$

Then if the angular response is

$$A_i(\theta,\phi) = C_i G(\theta) \exp \frac{-\{\cos^{-1}(\cos(\theta+(-1)^i\alpha)\cos\theta)\}^2}{4K} \qquad 5.23$$

substitution into Eq. 5.6 gives

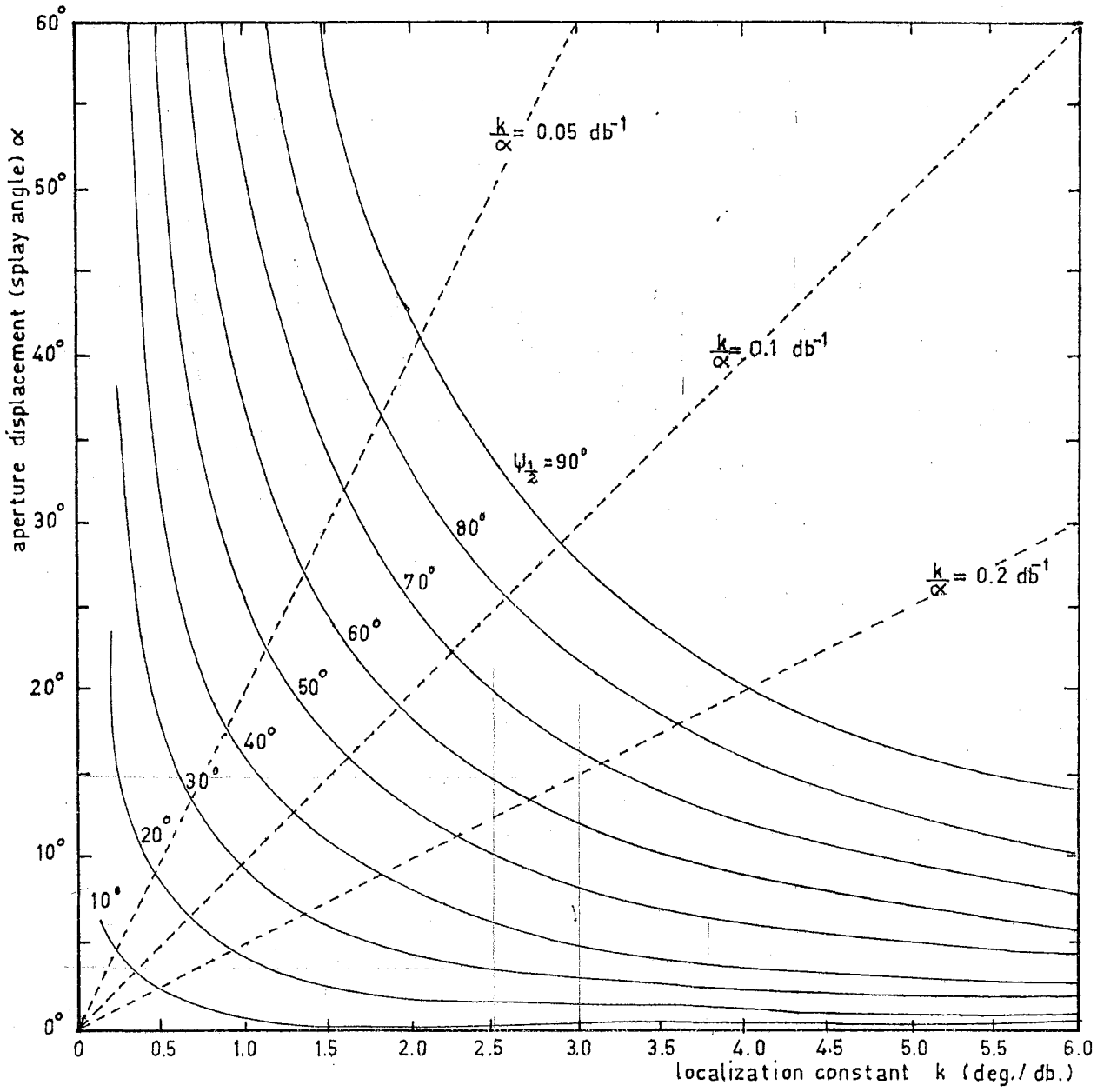


Fig. 5.2 The aperture angular displacement to match the display to a given localization constant.

$$\hat{\theta} = \frac{1}{4\alpha}\left[\{\cos^{-1}(\cos(\theta+\alpha)\cos\phi)\}^2 - \{\cos^{-1}(\cos(\theta-\alpha)\cos\phi)\}^2\right] \qquad 5.24$$

which is of the form

$$\hat{\theta} = g(\alpha) + g(-\alpha)$$

where $g(\alpha)$ is an arbitrary function and therefore $\frac{d\hat{\theta}}{d\alpha} = 0$. For analytic convenience the value $\alpha = \frac{\pi}{2}$ may be substituted in Eq. 5.24 to give the simplified expression

$$\hat{\theta} = \sin^{-1}(\sin\theta\cos\phi) \qquad 5.25$$

which is graphed for increments in ϕ of 20° in Fig. 5.4. The effect of object elevation is to introduce a bias in the azimuth estimator so as to underestimate the true azimuth angle.

5.5 Sensitivity of the Azimuth Estimator $\hat{\theta}$ to the Form of the Angular Response $F(\theta)$

The form of $F(\theta)$ to match the display to the psychophysics of localization was derived in Section 5.2. It is pertinent to examine the effect of perturbations in $F(\theta)$ upon $\hat{\theta}$ in order to determine the sensitivity of the system to the form of $F(\theta)$.

Let the response $F(\theta)$ be perturbed by ε_1 and ε_2 at azimuth angles $\theta-\alpha$ and $\theta+\alpha$ respectively. Then the absolute error $\delta(\hat{\theta}|\theta,\varepsilon_1,\varepsilon_2)$ will be

$$\delta(\hat{\theta}|\theta,\varepsilon_1,\varepsilon_2) = \hat{\theta}-\theta$$

$$= K \log_e\left(\frac{1+\zeta_1}{1+\zeta_2}\right), \qquad 5.26$$

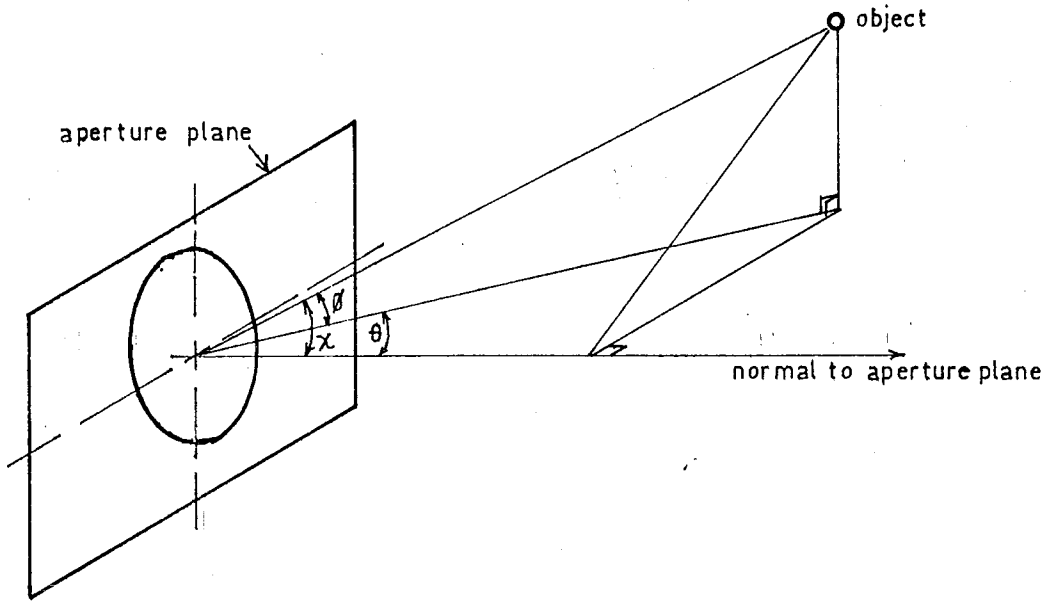


Fig. 5.3 Geometry defining the response angle for an elevated target.

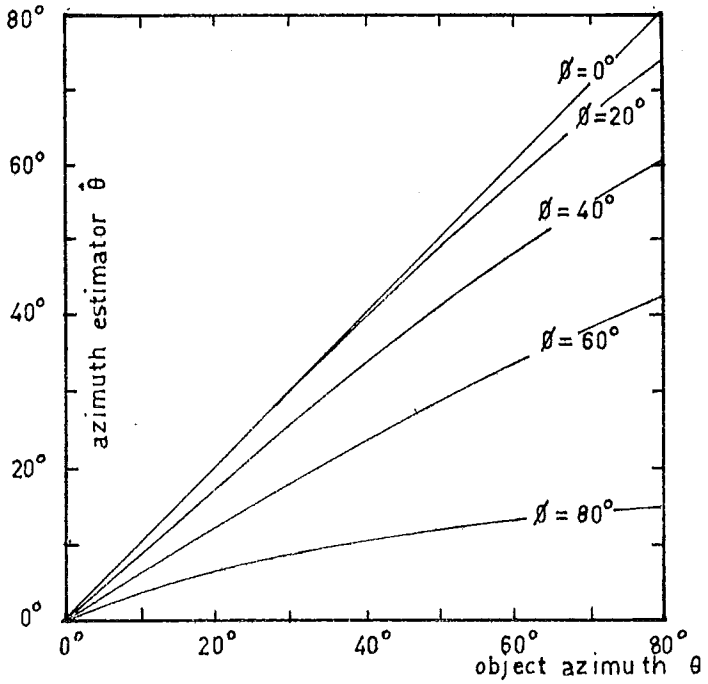


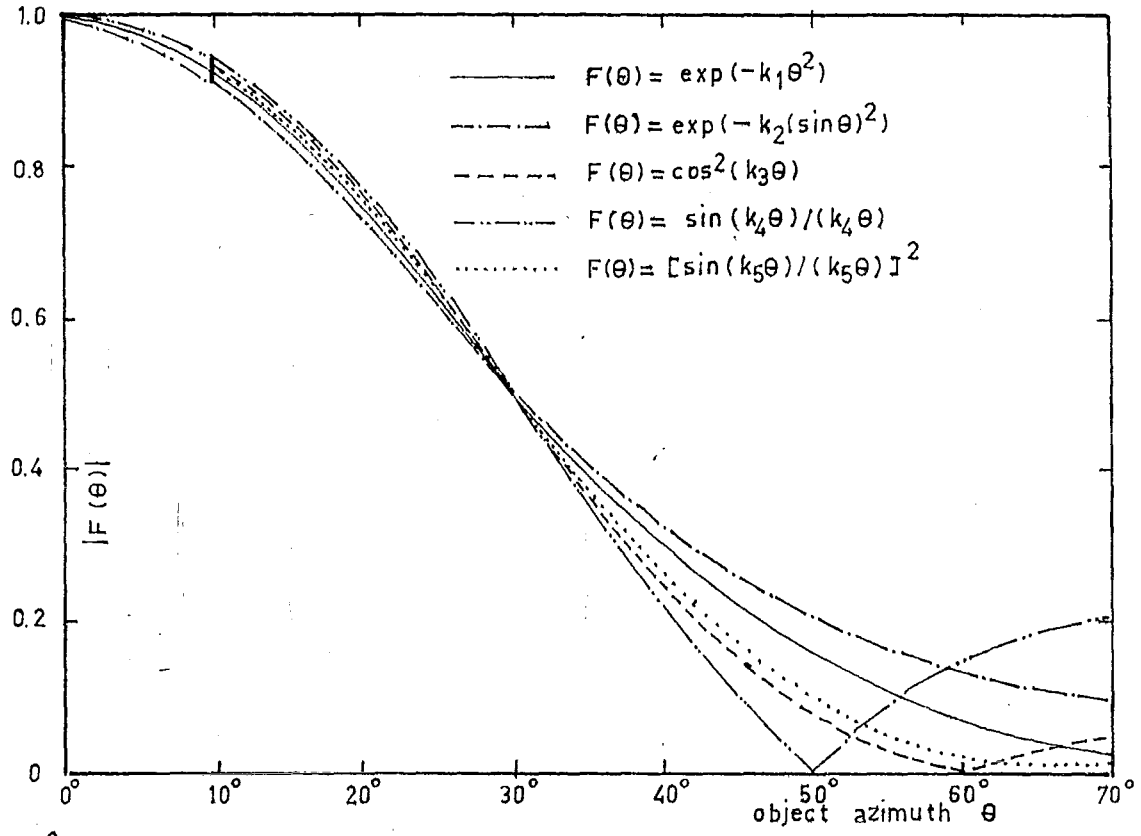
Fig. 5.4 The effect of object elevation upon the azimuth estimator.

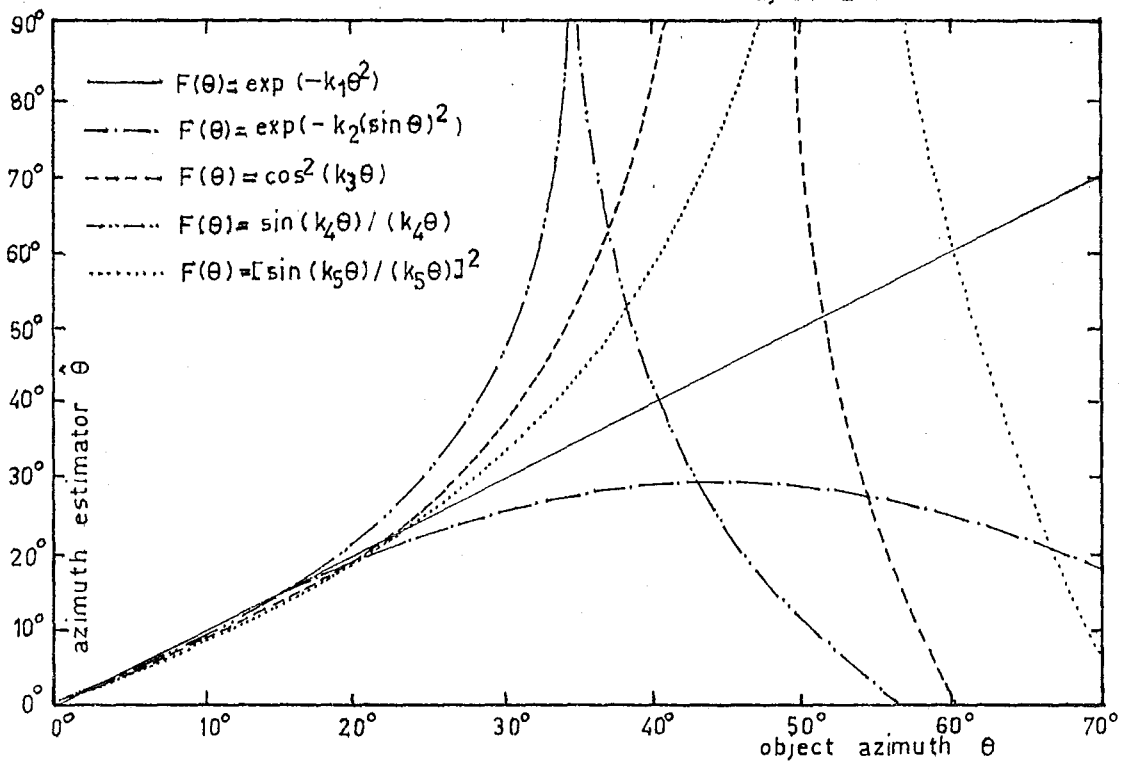
where $\zeta_i = \frac{\varepsilon_i}{F(\theta+(-1)^i\alpha)}$, $(i = 1,2)$. 5.27

This dependence of the estimator upon the fractional error in the response function shows that small absolute values of ε_i in the tails of the responses will cause large errors in the estimator. A more graphic example of this effect may be given by assuming different analytic forms for $F(\theta)$ and using the definition in Eq. 5.6 to define the estimator (assuming θ_0 to be zero). In Figs 5.5a and 5.5b some typical aperture response functions have been adjusted to a common 6 db beamwidth of 60^o and the display parameters adjusted to a localization constant of k = 2.5 degrees/db from Eq. 5.20. The analytic forms of the curves in Fig. 5.5a were chosen because they represent the far-field angular response of common aperture excitation functions or because of their general similarity to the idealised $F(\theta)$ defined in Eq. 5.9. The effects of any sidelobes and zeroes in the responses is manifest in Fig. 5.5b. At large azimuth angles the estimator is extremely sensitive to small absolute errors in $F(\theta)$.

5.6 The Variation of Loudness with Object Azimuth

The effective "window" or system beamwidth of an auditory display must be defined in terms of a psychophysical loudness scale. In the presence of controversy as to the form of the loudness function[116,117], a power law scale[118]





Top: Fig. 5.5a Typical azimuthal polar response functions adjusted to a 60° 6db beamwidth.

Bottom: Fig.5.5b The effect of substituting these functions into the estimation equation (Eq.5.6) with k = 2.5 deg/db.

with binaural loudness summation[119] is assumed here.

Under these assumptions the azimuthal variation in loudness $\Lambda(\theta)$, for stimuli whose amplitude is much greater than the absolute limen, may be written

$$\Lambda(\theta) = q(A_1(\theta) + A_2(\theta))^{0.6} \qquad 5.28$$

where q is a constant defining the loudness unit, and the exponent 0.6 has been empirically determined[118]. Because only the variation in loudness is of interest it is convenient to let q = 1, so that if $G(\theta) = 1$ and $C_1 = C_2 = 1$ substitution for $A_i(\theta)$ in Eq. 5.28 and rearrangement gives the normalised function (Eq. 5.11)

$$\Lambda_n(\theta_n) = 2^{0.6} \exp\left(-\frac{0.15(\theta_n^2+1)}{K_n}\right) \cosh^{0.6}\left(\frac{\theta_n}{2K_n}\right) \qquad 5.29$$

which is plotted in Fig. 5.6 for values of $k_n = 8.686K_n$ over the range of interest for a practical system.

An approximation to the "half loudness" beamwidth may be made as follows. Since $[F(\theta)]^{0.6} = 0.5$ when $\theta = 0.645\psi_{\frac{1}{2}}$, if $F(0.645\psi_{\frac{1}{2}}-\alpha) >> F(0.645\psi_{\frac{1}{2}} + \alpha)$ the display beamwidth will be approximately

$$\psi_{\frac{1}{2}\Lambda} \doteq 2\alpha + 1.29\psi_{\frac{1}{2}}$$

and using Eq. 5.20

$$\psi_{\frac{1}{2}\Lambda} \doteq \psi_{\frac{1}{2}}\left[\frac{\psi_{\frac{1}{2}}}{48k} + 1.29\right] \qquad 5.30$$

which will generally underestimate the true $\psi_{\frac{1}{2}\Lambda}$ because the contribution of the lower amplitude channel has been neglected.

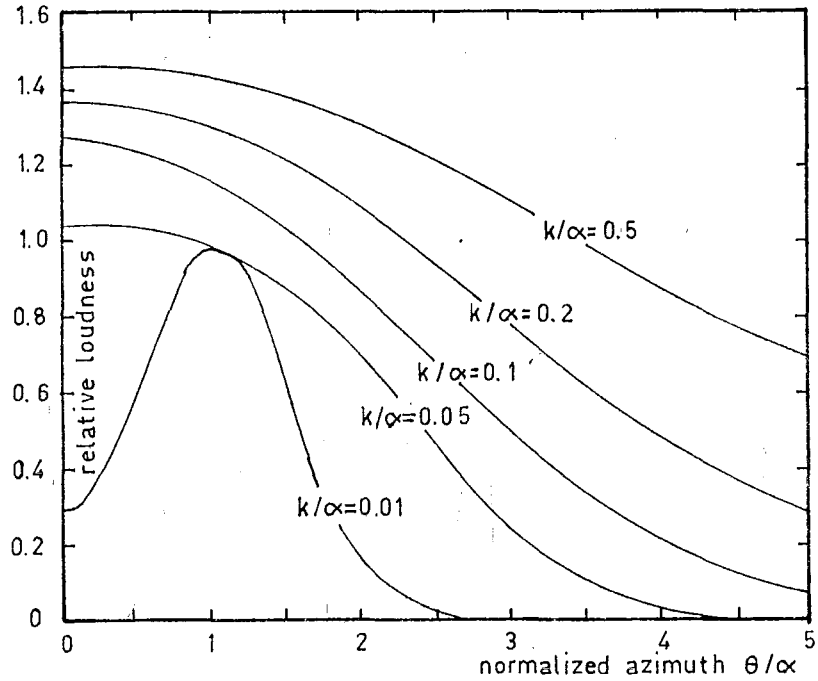


Fig. 5.6 Variation of loudness with azimuth based upon assumed binaural loudness summation and a power law loudness function.

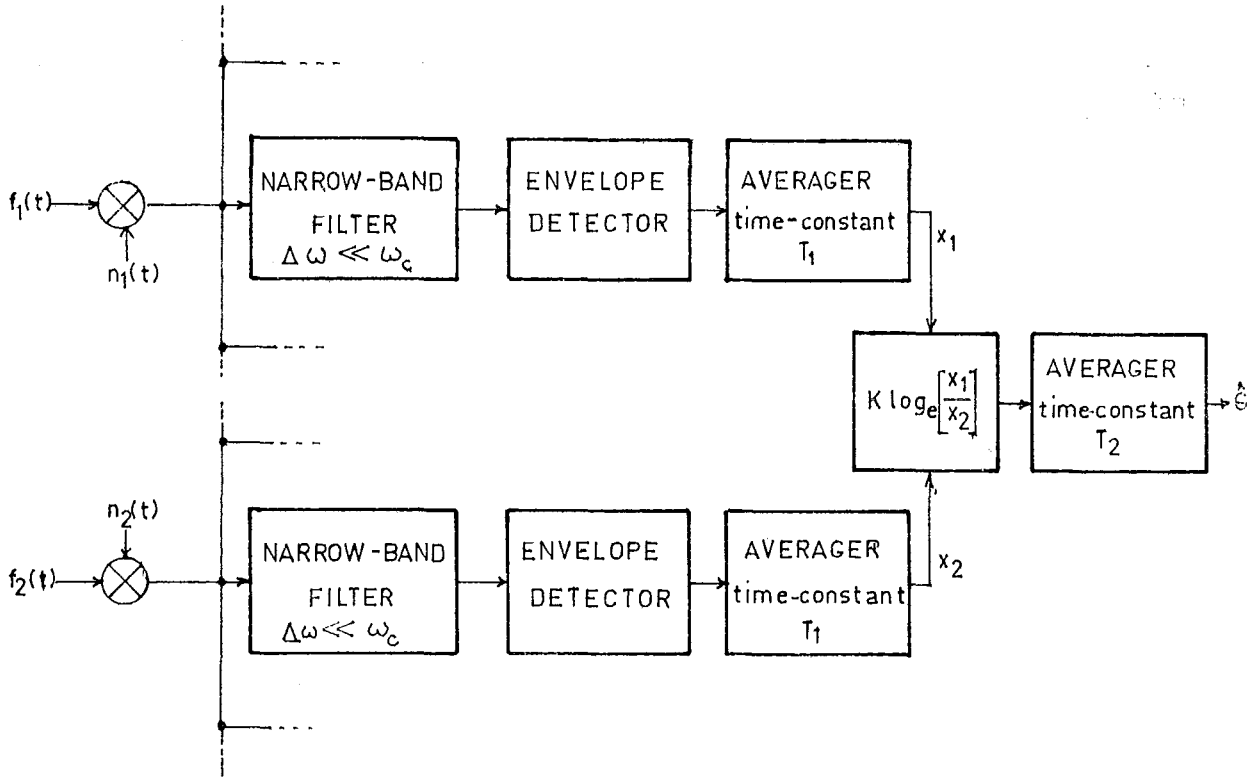


Fig. 5.7 Model of the azimuth estimation process.

5.7 Estimation of Azimuth in the Presence of Additive Gaussian Noise

With the present state of knowledge of the signal processing involved in binaural interaction it is not possible to derive statistics that describe the effects of dichotic white noise upon the localization of tones. The approach taken in this section, therefore, is to examine the sensitivity of the estimator $\hat{\theta}$ to additive white Gaussian noise in the two channels on the basis of the hypothetical estimation process shown in Fig. 5.7. The two channels of this model are based upon a commonly cited model of peripheral auditory signal processing[120], however the extension of this model to include binaural interaction poses additional problems because the site and time-constant of any post-detection averaging will affect the estimator. The analyses here treat two separate cases:

i The post-detection time constant T_1 is small compared to the correlation time of the filter ($T_1 << \frac{1}{2\pi\Delta\omega}$, where $\Delta\omega$ is the effective filter bandwidth), in which case the estimator is based upon comparisons of the instantaneous values of the contaminated envelopes.

ii The limiting form as the time constant T_1 tends to infinity is examined by invoking the ergodic hypothesis so that the estimator is derived from the statistical expectations of the quantities in each

channel.

Schalow[121] reports that signal averaging before amplitude comparison gives a smaller variance on the estimator when S/N > 0 db, whereas averaging the instantaneous estimates gives a smaller variance when S/N < -5 db, when the estimator is based upon the arctan function.

It is relevant to discuss at this point the known features of binaural interaction. Zwislocki[122] points out that any temporal integration must take place at a higher level in the ascending structure than that at which temporal differences in a signal's microstructure are resolved. A similar argument shows that the envelope structure of impulsive stimuli must be retained to the site of binaural interaction. There is evidence that amplitude comparison takes place at a similar level[75] (accessory nucleus in the superior olive) or even more peripheral in the cochlear transduction from mechanical to neural energy[110]. On such arguments the first signal processing method ($T_1 << \frac{1}{2\pi\Delta\omega}$) seems reasonable.

However, there are models of the localization process (including the one presented in Appendix 1) that hypothesize the final site of binaural comparison to be at some high level (possibly cortical)[123,76]. In these models the peripheral binaural interaction is considered only as a signal coding in which the binaural differences are coded into neural activity. Subsequent temporal integration with a time constant much greater than the correlation time of the

filter could conceivably precede the comparison of activity in the two hemi-spheres.

The possibility of non-linear averaging at some unspecified site was introduced by the experiment on the perceptual effect of time-varying IAD in Section 4.5. Until the mechanisms of all such phenomena of auditory localization are elucidated, analyses such as those presented here must be interpreted merely as an attempt to show the sensitivity of the display to noise by examining possible signal processing methods.

The auditory display form derived in this chapter bears a superficial resemblance to the amplitude comparison monopulse (acm) radar system[124]; the noise performance of which has been well documented[125,126,127]. The different form of the angular estimator and the restricting assumptions underlying these analyses (prior matched filtering, small azimuth angle, and even and odd aperture responses) render them unsuitable for the present study. Cooper[128] derived the probability density of an angular estimator based upon cosinusoidal aperture functions. His analysis was a particular case of a general analysis presented below (Section 5.7.2, part i).

In this section the form of the maximum likelihood estimator (MLE) of azimuth in high signal to noise conditions is derived and the performance of the two variants of the

model is investigated when the contaminating noise sources are independent or equal.

5.7.1 The Maximum Likelihood Estimator of Object Azimuth:

Let the object range be constant, and the elevation be zero, so that in the presence of independent noise sources in the two channels the display signals may be written

$$f_i(t) = C_i A_i(\theta)\cos(\omega_c t+\eta_i) + n_i(t) \qquad 5.31$$

where ω_c is the centre frequency of the filter and $n_i(t)$ are independent zero-mean Gaussian noise sources. At the output of the envelope detectors only the random variables $x_i(t)$ are observable. The problem is to define the signal processing on $x_i(t)$ to obtain the most efficient estimate of θ. The method used is that of maximum likelihood[129]. Assume that the quantities $x_i(t)$ are randomly sampled to give the sequences $\{x_{in}\}$, $n = 1,2,\cdots N$, of independent samples from the two channels. (The rationale for assuming a discrete sampled process to describe a continuous process is as given by Marcum[130].)

Let the bandwidth of the filter $\Delta\omega$ be small so that the sequences $\{x_{in}\}$ are described by the Rayleigh-Rice distribution[131] i.e.

$$p(x_{in}|\theta,\sigma_i^2) = \frac{x_{in}}{\sigma_i^2}\exp\left(-\frac{x_{in}^2+A_i^2(\theta)}{2\sigma_i^2}\right) I_0\left(\frac{x_{in}A_i(\theta)}{\sigma_i^2}\right) \qquad x_{in} > 0$$

$$= 0 \qquad x_{in} < 0 \qquad 5.32$$

where σ_i^2 are mean-square noise voltages.

Then define the log-likelihood function $L(\theta)$ as the logarithm of the joint probability density of the two sequences[129], i.e.

$$L(\theta) = -N \log_e(\sigma_1^2\sigma_2^2) + \sum_{i=1}^{2}\sum_{n=1}^{N}\left[\log_e x_{in} - \frac{A_i^2(\theta)}{2\sigma_i^2} - \frac{x_{in}^2}{2\sigma_i^2} + \log_e I_0\left(\frac{x_{in}A_i(\theta)}{\sigma_i^2}\right)\right]. \qquad 5.33$$

The maximum likelihood estimator $\ddot{\theta}$ is a solution to the system of equations

$$\frac{\partial L(\theta)}{\partial\theta} = 0; \quad \frac{\partial L(\theta)}{\partial\sigma_1^2} = 0; \quad \frac{L(\theta)}{\partial\sigma_2^2} = 0 \qquad 5.34$$

under the constraint that $\frac{\partial^2 L(\theta)}{\partial\theta^2} < 0$. Now

$$\frac{\partial L(\theta)}{\partial\theta} = \sum_{i=1}^{2}\left[\left[\sum_{n=1}^{N}\frac{I_1\left(\frac{x_{in}A_i(\theta)}{\sigma_i^2}\right)}{I_0\left(\frac{x_{in}A_i(\theta)}{\sigma_i^2}\right)}\frac{A_i'(\theta)x_{in}}{\sigma_i^2}\right] - \frac{NA_i'(\theta)A_i(\theta)}{\sigma_i^2}\right] \qquad 5.35$$

where $A_i'(\theta) = \frac{\partial A_i(\theta)}{\partial\theta}$ and although the ratio of modified Bessel functions can be expressed as a continued fraction[132] or the ratio of two infinite series, such forms are not amenable to analytic manipulation. It is shown in Appendix 4 that for $z >> 3$

$$\frac{I_1(z)}{I_0(z)} \doteq 1 - \frac{1}{2z}. \qquad 5.36$$

Accordingly if the signal to noise ratio is large so that

the probability that $x_{in} < \frac{3\sigma_i^2}{A_i(\theta)}$ may be neglected, Eq. 5.35 may be written

$$\frac{\partial L(\theta)}{\partial \theta} = N \sum_{i=1}^{2} \left[\frac{A_i'(\theta)}{A_i(\theta)} - \frac{A_i'(\theta)A_i(\theta)}{\sigma_i^2} - \frac{A_i'(\theta)\bar{x}_{i\cdot}}{2\sigma_i^2} \right] \qquad 5.37$$

where $\bar{x}_{i\cdot} = \frac{1}{N} \sum_{n=1}^{N} x_{in}$.

Similarly, using Eq. 5.36 the other equations in Eq. 5.33 may be written

$$\frac{\partial L(\theta)}{\partial \sigma_i^2} = \frac{N}{2(\sigma_i^2)^2} (A_i^2(\theta) - 2A_i(\theta)\bar{x}_{i\cdot} + \overline{x_i^2}_{\cdot} - \sigma_i^2) \qquad 5.38$$

where $\overline{x_i^2}_{\cdot} = \frac{1}{N} \sum_{n=1}^{N} x_{in}^2$. Then substituting for $A_i'(\theta)$ in Eq. 5.37 and solving Eqs 5.37 and 5.38 simultaneously gives the required solution for the MLE as

$$\ddot{\theta} = K \log_e \left(\frac{\bar{x}_{1\cdot} + \sqrt{(\bar{x}_{1\cdot})^2 - \frac{3}{4}\overline{x_1^2}_{\cdot}}}{\bar{x}_{2\cdot} + \sqrt{(\bar{x}_{2\cdot})^2 - \frac{3}{4}\overline{x_2^2}_{\cdot}}} \right) \qquad 5.39$$

and as σ_i^2 tends to zero for both i = 1,2 the limiting form is

$$\lim_{\sigma_1 \to 0,\ \sigma_2 \to 0} \ddot{\theta} = K \log_e \left(\frac{\bar{x}_{1\cdot}}{\bar{x}_{2\cdot}} \right) . \qquad 5.40$$

As the sequence length N tends to infinity the means $\bar{x}_{i\cdot}$ and $\overline{x_i^2}_{\cdot}$ may be replaced by the statistical expectations and the ergodic hypothesis may be invoked to equate the time

and ensemble averages, thus generalising Eq. 5.39 to the continuous processing of the model in Fig. 5.7.

5.7.2 Estimation of Azimuth in Noise by Direct Comparison of Envelope Amplitudes ($T_1 << \frac{1}{2\pi\Delta\omega}$):

i The Probability Density of the Estimator $\hat{\theta}$ in the Presence of Independent Noise Sources: In this section the first variant of the processing model is examined by deriving the probability density function (pdf) of the estimator $p(\hat{\theta}|\theta,\rho)$, where ρ is the signal to noise ratio defined below, when $n_1(t)$ and $n_2(t)$ are statistically independent. This situation would arise when the two noise sources represent thermal noise generated in the receiver channels. Let the statistics of the detected envelopes be described by Eq. 5.32 with the restriction that $\sigma_1^2 = \sigma_2^2 = \sigma^2$, and further let $C_1 = C_2 = C$ (i.e. $\theta_0 = 0$) and $G(\theta) = 1$ in Eq. 5.5 so that

$$A_i(\theta) = CF_i(\theta) \qquad 5.41$$

where $F_i(\theta) = F(\theta+(-1)^i\alpha)$. Since the estimator is in ratio form the sequences $\{x_{in}\}$ may be normalized as

$$\nu_{in} = \frac{x_{in}}{\sqrt{2}\sigma} \qquad 5.42$$

and if the maximum signal to noise ratio in each channel is written

$$\rho = \frac{C}{\sqrt{2}\sigma} \qquad 5.43$$

the pdf of the normalised amplitude in each channel is

(from Eq. 5.32)

$$p_1(\nu_i|\theta,\rho) = 2\nu_i \exp(-[\nu_i^2+\rho^2 F_i^2(\theta)])\ I_0(2\rho F_i(\theta)\nu_i) \quad \nu_i \geq 0 \qquad 5.44$$

$$= 0 \quad \nu_i < 0$$

Define a variable $z = \frac{\nu_1}{\nu_2}$; the pdf of z is[133]

$$p_2(z|\theta,\rho) = \int_{-\infty}^{\infty} |\nu_2| p_1(z\nu_2|\theta,\rho) p_1(\nu_2|\theta,\rho) d\nu_2 \qquad 5.45$$

$$= 4z \exp(-\rho^2[F_1^2(\theta) + F_2^2(\theta)])$$

$$\cdot \int_0^{\infty} \nu_2^3 \exp(-\nu_2^2(1+z^2)) I_0(2\rho z\nu_2 F_1(\theta)) I_0(2\rho\nu_2 F_2(\theta)) d\nu_2 \qquad 5.46$$

which is a form of Weber's 2nd exponential integral[134], which gives a solution:

$$p_2(z|\theta,\rho) = \frac{2z}{\pi d^4} \exp(-\rho^2(F_1^2(\theta) + F_2^2(\theta)))$$

$$\cdot \int_0^{\infty} \exp\left(-\frac{a^2+b^2-2ab\cos\xi}{4d^2}\right)\left[1 - \frac{a^2+b^2-2ab\cos\xi}{4d^2}\right] d\xi \qquad 5.47$$

where $a = j2\ zF_1(\theta)$, $b = j2\ F_2(\theta)$, $d = 1+z^2$, $(j = \sqrt{-1})$.

Then

$$p_2(z|\theta,\rho) = \frac{2z}{d^4} \exp(-\rho^2[F_1^2(\theta) + F_2^2(\theta)]) \exp\left(-\frac{a^2+b^2}{4d^2}\right)$$

$$\cdot \left[\left(1 - \frac{a^2+b^2}{4d^2}\right) I_0\left(\frac{ab}{2d^2}\right) + \frac{ab}{2d^2} I_1\left(\frac{ab}{2d^2}\right)\right] \qquad 5.48$$

$$= \frac{2z}{(1+z^2)^2} \exp\left(\frac{-\rho^2(z^2F_2^2(\theta) + F_1^2(\theta))}{1+z^2}\right)$$

$$\cdot \left[\left(1 + \frac{\rho^2(z^2F_1^2(\theta) + F_2^2(\theta))}{1+z^2}\right) I_0\left(\frac{2\rho^2 zF_1(\theta)F_2(\theta)}{1+z^2}\right)\right.$$

$$+ \frac{2\rho^2 z F_1(\theta)\ F_2(\theta)}{1 + z^2}\ I_1\left(\frac{2\rho^2 z F_1(\theta) F_2(\theta)}{1 + z^2}\right)\Bigg] \qquad 5.49$$

To this point the analysis has been general; the variable is now transformed to the estimator $\hat{\theta}$ using the relationship

$$p_4(\hat{\theta}|\theta,\rho) = \frac{1}{K} \exp\left(\frac{\hat{\theta}}{K}\right) p_3\left(\exp\left(\frac{\hat{\theta}}{K}\right)|\theta,\rho\right) \qquad 5.50$$

which upon substitution into Eq. 5.48 and suitable rearrangement gives

$$p_4(\hat{\theta}|\theta,\rho) = \frac{1}{2w^2K} \exp\left(-\frac{\rho^2}{2}\{(1-u)F_1^2(\theta) + (1+u)F_2^2(\theta)\}\right)$$

$$\cdot \left[\{1 + \frac{\rho^2}{2}((1+u)F_1^2(\theta) + (1-u)F_2^2(\theta))\} I_0\left(\frac{\rho^2 F_1(\theta)F_2(\theta)}{w}\right)\right.$$

$$\left. + \frac{\rho^2 F_1(\theta)F_2(\theta)}{w}\ I_1\left(\frac{\rho^2 F_1(\theta)F_2(\theta)}{w}\right)\right] \qquad 5.51$$

where $u = \tanh\left(\frac{\hat{\theta}}{K}\right)$, $w = \cosh\left(\frac{\hat{\theta}}{K}\right)$.

Computed examples of $p_4(\hat{\theta}|\theta,\rho)$ are shown in Fig. 5.8 using the normalization with respect to α given by Eqs 5.11. The pdf of $\hat{\theta}$ is known as the likelihood function $\ell_1(\theta)$ for a single observation. If m independent observations are available, the likelihood function $\ell_m(\theta)$ is

$$\ell_m(\theta) = \prod_{n=1}^{m} \ell_n(\theta) \qquad 5.52$$

and the inverse probability relation[18] can be used to derive the posterior distribution $p_5(\theta)$

$$p_5(\theta) = p_6(\theta)\ell_m(\theta)$$

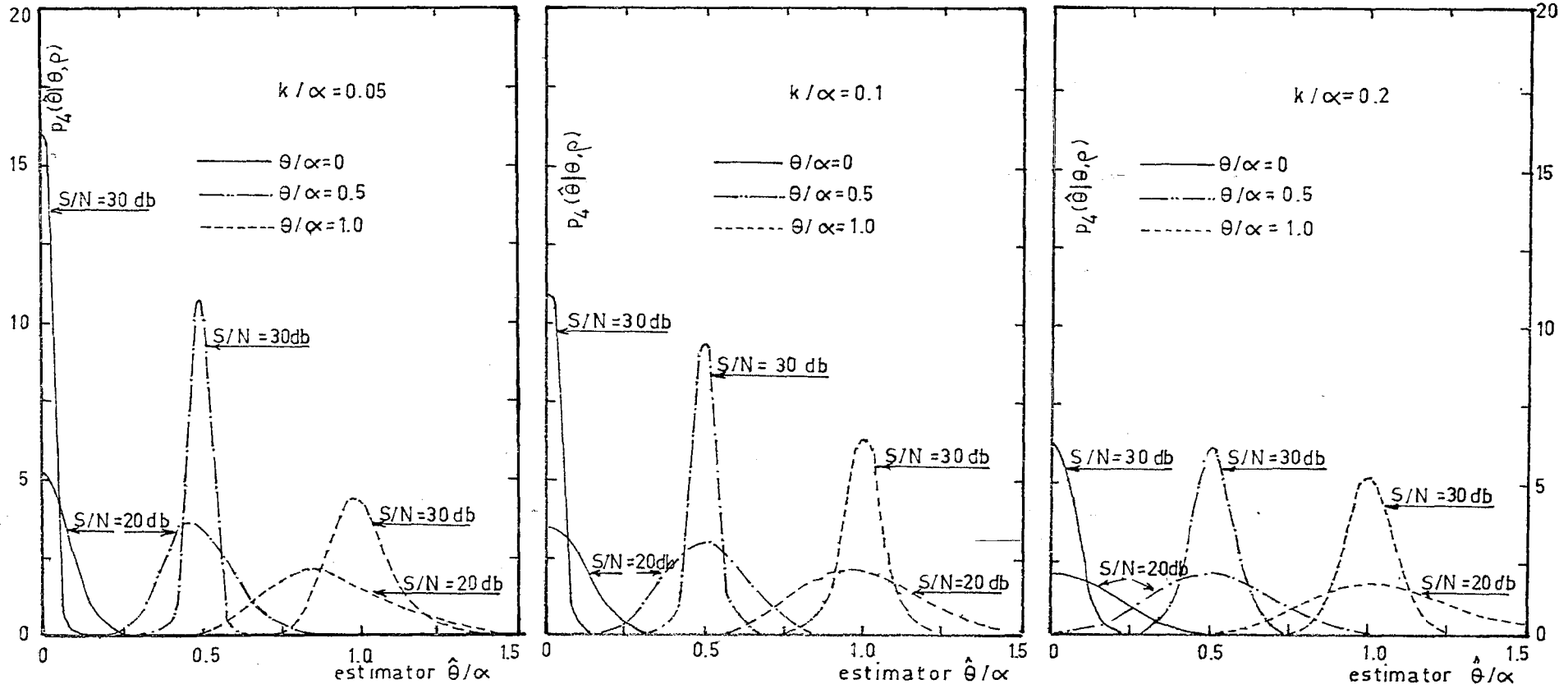


Fig. 5.8 Typical examples of the likelihood function of the azimuth estimator with interference from uncorrelated additive Gaussian noise sources.
($T_1 << \frac{1}{2\pi\Delta\omega}$)

where $p_6(\theta)$ is the a-priori distribution of θ. 5.53

If the model is extended to include post-estimation averaging, the limiting form as the time constant T_2 becomes large may be found from the ergodic principle, i.e.

$$E\{\hat{\theta}\} = \int_{\infty}^{\infty} \hat{\theta} p_4(\hat{\theta}|\theta,\rho) d\hat{\theta} \qquad 5.54$$

$$= \lim_{T_2 \to \infty} \frac{1}{T_2} \int_0^{T_2} \hat{\theta} dt, \qquad 5.55$$

which may be calculated by substituting $p_4(\hat{\theta}|\theta,\rho)$ from Eq. 5.51. It can be seen from Fig. 5.8 that the effect of the noise on such an estimator is tobias the estimate toward $\theta = 0$.

ii <u>The Probability Density of the Estimator in the Presence of Correlated Noise Sources ($n_1(t) = n_2(t)$)</u>: The above analysis requires statistical independence between the samples x_{in}, which results if the noise sources $n_i(t)$ or the phase angles η_i are independent random variables. In this section the particular case of statistical dependence when $n_1(t) = n_2(t) = n(t)$ and $\eta_1 = \eta_2 = \eta$ in Eq. 5.31 is examined. This situation typifies the effect of environmental noise, either from the echolocation object space or from the ambient noise in the listener's environment.

In the absence of the signal let the noise at the filter output be described by the narrow-band approximation

$$n_f(t) = \zeta(t) \cos(\omega_c t + \xi) \qquad 5.56$$

where the frequency ω_c has been arbitrarily equated to that of the signal (Eq. 5.31), and the random variables ζ and ξ (the time dependence being implicit) are described by the Rayleigh distribution

$$p_1(\zeta) = \frac{\zeta}{\sigma^2} \exp \frac{-\zeta^2}{2\sigma^2} \qquad \zeta \geq 0 \qquad 5.57$$

$$= 0 \qquad \zeta < 0$$

$$p_2(\xi) = \frac{1}{2\pi} \qquad 5.58$$

where σ^2 is the mean square amplitude of $n_f(t)$.

When both signal and noise are present the instantaneous normalised envelope amplitudes ν_i may be found by vector addition

$$\nu_i = (\rho^2 F_i^2(\theta) + a^2 + 2\rho F_i(\theta) a \cos(\xi-\eta))^{\frac{1}{2}} \qquad 5.59$$

where $a = \frac{\zeta}{\sqrt{2}\sigma}$ so that

$$p_3(a) = 2a \exp(-a^2) \qquad a \geq 0 \qquad 5.60$$

$$= 0 \qquad a < 0$$

If the time origin is chosen so that $\eta = 0$ the ratio $z = \frac{\nu_1}{\nu_2}$ is the positive root of

$$z = \left[\frac{\rho F_1^2(\theta) + a^2 + 2\rho F_1(\theta)\, a \cos \xi}{\rho F_2^2(\theta) + a^2 + 2\rho F_2(\theta)\, a \cos \xi}\right]^{\frac{1}{2}} \qquad 5.61$$

The probability that z lies in the elementary interval $z_1 < z < z_1+dz$ is the sum over all ξ_1 of the probability that ξ lies in the interval $\xi_1 < \xi < \xi_1+d\xi$ and simultaneously

that a lies in either of the regions $a_1 < a < a_1+da$ or $a_2 < a < a_2+da$, where a_1 and a_2 are real positive roots of the quadratic in a formed by rearranging Eq. 5.61. Then expressing the limiting form of the summation as an integral, the pdf for z is

$$p_4(z|\theta,\rho) = \sum_{m=1}^{2} \int_0^{2\pi} p_3(a_m) \left|\frac{da}{dz}\right|\Bigg|_{a=a_m} p_2(\xi)\, d\xi \qquad 5.62$$

$$= \frac{1}{\pi} \sum_{m=1}^{2} \int_0^{2\pi} a_m \exp(-a_m^2) \left|\frac{da}{dz}\right|\Bigg|_{a=a_m} d\xi \qquad 5.63$$

where evaluation takes place separately for the two terms in those regions of ξ for which $R(a_m) > 0$ and $I(a_m) = 0$ for $m = 1,2$ and where $R(.)$ and $I(.)$ are real and imaginary parts, then

$$\frac{da}{dz} = \frac{-z(a^2 + 2a\, F_2(\theta)\cos\xi + \rho^2 F_2^2(\theta))}{a(z^2-1) + \rho(z^2 F_2(\theta) - F_1(\theta))\cos\xi} \,. \qquad 5.64$$

The variable in Eq.5.64 may be transformed as in Eq. 5.50. i.e.

$$p_5(\hat{\theta}|\theta,\rho) = \frac{1}{K} \exp\left(\frac{\hat{\theta}}{K}\right) p_4\left(\exp\left(\frac{\hat{\theta}}{K}\right)|\theta,\rho\right) \qquad 5.65$$

The examples of $p_5(\hat{\theta}|\theta,\rho)$ given in Fig. 5.9 were computed from Eqs 5.61-5.63 using numerical integration methods. The curves differ from those of the previous analysis in that

$$\lim_{\theta\to 0} p_5(\hat{\theta}|\theta,\rho) = \delta(\hat{\theta}) \qquad 5.66$$

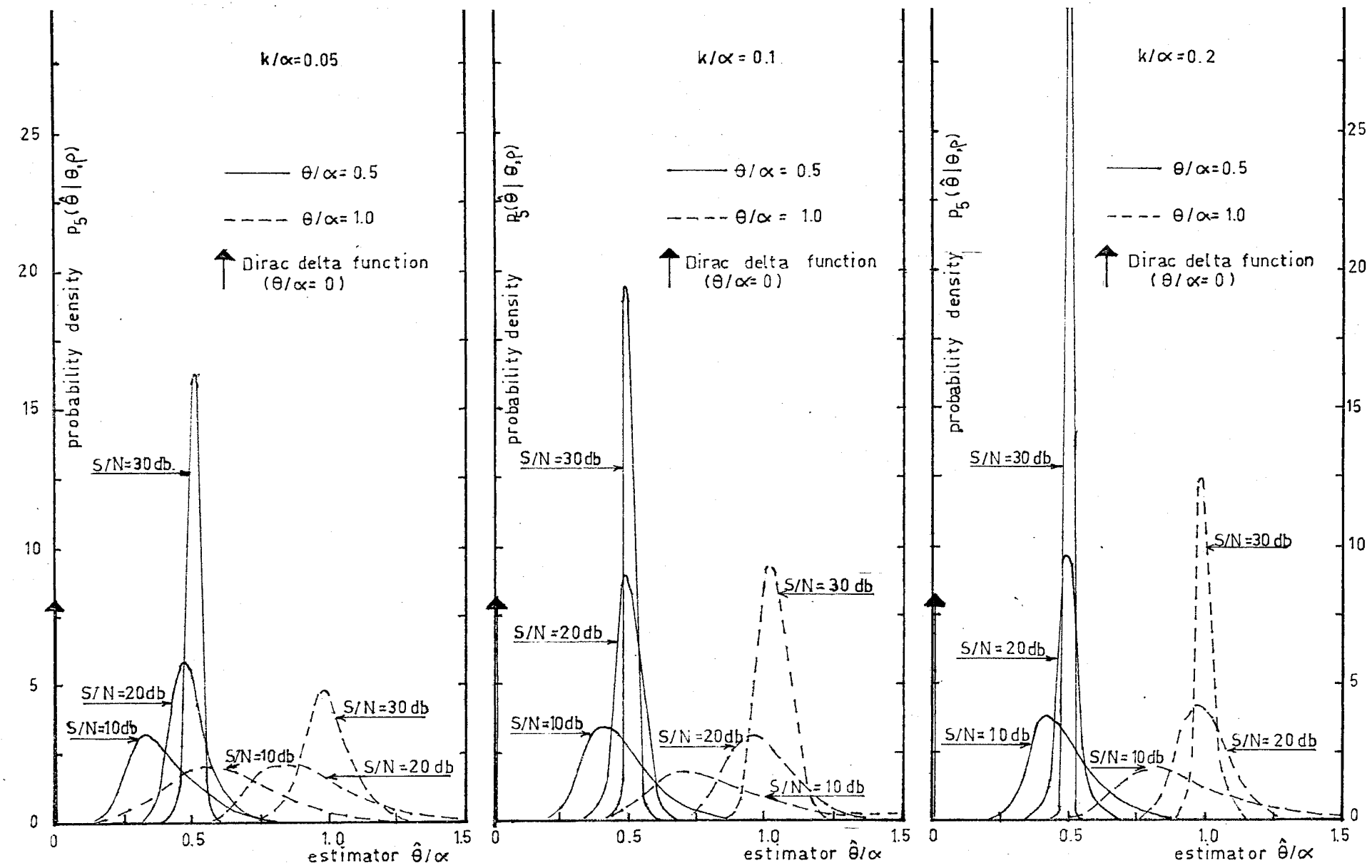


Fig. 5.9 Typical examples of the likelihood function of the azimuth estimator with interference from identical additive Gaussian noise sources. ($T_1 \ll \frac{1}{2\pi\alpha}$)

where $\delta(\hat{\theta})$ is the Dirac delta function, because when $\theta = 0$ both envelopes are identical, i.e. $\nu_1 = \nu_2$.

5.7.3 Estimation of Azimuth in Noise with Post-Detection Averaging ($T_1 >> \frac{1}{2\pi\Delta\omega}$):

In general a knowledge of only the first-order pdf of the input waveform is insufficient to describe the output statistics of a finite memory element. Therefore only the limiting form of the second variant of the model as $T_1 \to \infty$ is examined here. This form of the model does not depend upon any statistical dependence between the random signals in the two channels.

Equating the ensemble and time averages by the ergodic principle gives

$$\lim_{T_1 \to \infty} \frac{1}{T_1} \int_0^{T_1} x_i(t)\,dt = E\{x_{in}\}$$

where x_{in} are independent samples of $x_i(t)$. The estimator in this case may be written

$$\hat{\theta} = K \log_e \left(\frac{E\{x_{1n}\}}{E\{x_{2n}\}} \right).$$

In particular if the samples x_{in} are described by Rayleigh-Rice statistics (Eq. 5.31) the expectations will be[135]

$$E\{x_{in}\} = \Gamma(2) 2^{\frac{1}{2}} \, {}_1F_1\left(-\tfrac{1}{2}; 1; -\frac{A_i^2(\theta)}{2\sigma_i^2}\right)$$

where $\Gamma(\cdot)$ is the gamma function and ${}_1F_1(\cdot;\cdot;\cdot)$ is the

confluent hypergeometric function (defined in Appendix 4).

Then

$$\hat{\theta} = K \log_e \left(\frac{{}_1F_1(-\tfrac{1}{2};1;-\rho^2 {}_1F_1{}^2(\theta))}{{}_1F_1(-\tfrac{1}{2};1;-\rho^2 {}_2F_2{}^2(\theta))} \right)$$

if $\sigma_1 = \sigma_2$ as before. Typical curves of the estimator $\hat{\theta}$ are plotted in Fig. 5.10; the effect of the noise is an introduction of bias toward $\theta = 0$ into the estimate.

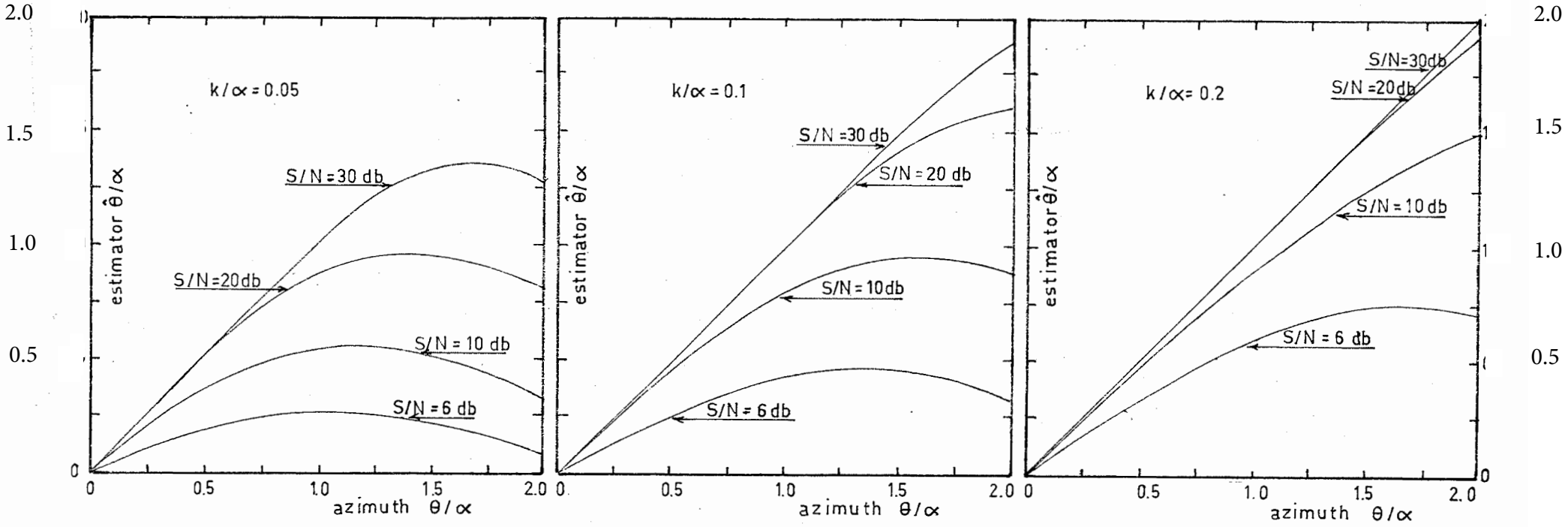


Fig. 5.10 The effect of additive noise sources on the azimuth estimator ($T_1 \gg \frac{1}{2\pi\Delta\omega}$)

CHAPTER 6

THE EFFECT OF WAVEFORM AND SYSTEM PARAMETERS UPON THE AUDITORY DISPLAY.

6.1 Introduction

The analyses of the display presented in Chapter 5 were idealised in that no consideration was given to the effects of the elements of a practical echolocation system, or to the characteristics of the environment. Such analyses are useful in defining the display but do not give an understanding of system performance in a complex and dynamic environment. In this chapter some topics relating the signal processing and system elements to the characteristics of the display are discussed, with an emphasis upon the effects of the wide-bandwidth linear FM waveform.

6.2 System Performance in Noise: Optimization of the Signal to Noise Ratio at the Display.

In the normal temporal analysis of radar returns attention is often directed to an optimization of the instantaneous signal to noise ratio ρ_t defined as

$$\rho_t = \frac{\text{peak instantaneous output signal power}}{\text{expected output noise power}} \qquad 6.1$$

and the linear (matched) filter that will maximise this ratio is optimum for signal detection by other criteria (likelihood ratio or inverse probability).[137] The impulse response h(t) of the matched filter for a received signal r(t)

is $$h(t) = C\, r(T_d - t) \qquad 6.2$$

where the gain C and delay T_d are arbitrary constants.

If however, the delay of the received signal is to be presented in the frequency domain, as in the auditory display under discussion, a different signal to noise criterion must be adopted. Accordingly, define a frequency domain signal to noise ratio ρ_ω as

$$\rho_\omega = \frac{\Phi_{os}(\omega_c)}{E\{\Phi_{on}(\omega_c)\}} \qquad 6.3$$

where $\Phi_{os}(\omega)$ and $\Phi_{on}(\omega)$ are the energy density spectra of finite samples of the output in the presence of signal alone and noise alone respectively and ω_c is the detection angular frequency.

Assume a non-dispersive single object environment with an impulse response $h_e(t)$ of the form

$$h_e(t) = A\, \delta(t - \tau) \qquad 6.4$$

where A and τ are arbitrary and $\delta(t)$ is the Dirac delta function. The received signal r(t) after convolution with $h_e(t)$ will be

$$r(t) = As(t - \tau) \qquad 6.5$$

where s(t) is the transmitted waveform. In accordance with the duality of the operations of multiplication and convolution in the time and frequency domains, the equivalent of the optimising linear filter (with an impulse response given by Eq. 6.2) in the time domain, is specified to be a dimensionless multiplier g(t) that exists for a period T.

The form of g(t) to optimise ρ_ω is now derived.

In the presence of additive white Gaussian noise n(t) and the signal r(t), the multiplier output f(t) will be

$$f(t) = g(t)[r(t) + n(t)] \quad -\frac{T}{2} < t < \frac{T}{2} \qquad 6.6$$

$$= 0 \qquad |t| > \frac{T}{2}$$

where g(t) exists in the period $|t| < \frac{T}{2}$. Let the noise have a one sided power spectral density $\frac{N_o}{2}$, and let the autocorrelation function of the output in the presence of noise alone $\psi_{on}(\tau)$ be

$$\psi_{on}(\tau) = \int_{-\infty}^{\infty} n(t)n(t+\tau)g(t)g(t+\tau)dt \qquad 6.7$$

Because $\psi_{on}(\tau)$ is computed from a finite sample of the stationary random process n(t) it will be a random variable. The statistical expectation may be found by averaging over all such samples with a time origin λ i.e. define

$$\psi_{on}(\tau|\lambda) = \int_{-\infty}^{\infty} n(t)n(t+\tau)g(t-\lambda)g(t-\lambda+\tau)dt \qquad 6.8$$

then invoking the ergodic principle

$$E\{\psi_{on}(\tau)\} = \lim_{T_1\to\infty} \frac{1}{T_1} \int_{\frac{-T_1}{2}}^{\frac{T_1}{2}} \psi_{on}(\tau|\lambda)d\lambda \qquad 6.9$$

and substituting $\zeta = t - \lambda$

$$E\{\lambda_{on}(\tau)\} = \int_{-\infty}^{\infty} g(\zeta)g(\zeta+\tau) \lim_{T_1\to\infty} \frac{1}{T_1} \int_{\frac{-T_1}{2}}^{\frac{T_1}{2}} n(t)n(t+\tau)dt\,d\zeta \qquad 6.10$$

$$= \psi_{gg}(\tau) \circ \psi_{nn}(\tau) \qquad 6.11$$

where $\psi_{gg}(\tau)$ and $\psi_{nn}(\tau)$ are the autocorrelation functions of the multiplier and noise respectively. The expectation of the output energy density spectrum is given by the convolution integral

$$E\{\Phi_{on}(\omega)\} = \int_{-\infty}^{\infty} \Phi_g(\zeta)\Phi_n(\omega - \zeta)\,d\zeta \qquad 6.12$$

$$= \frac{No}{2} \int_{-\infty}^{\infty} \Phi_g(\omega)\,d\omega \qquad 6.13$$

With the signal alone at the input, the output energy spectral density may be found directly

$$\Phi_{os}(\omega) = \left| \int_{-\infty}^{\infty} r(t)g(t)e^{-j\omega t}\,dt \right|^2 \qquad 6.14$$

and substitution of Eqs. 6.13 and 6.14 into Eq. 6.3 gives

$$\rho_\omega = \frac{\left| \int_{-\infty}^{\infty} r(t)g(t)e^{-j\omega t}dt \right|^2}{\frac{No}{2} \int_{-\infty}^{\infty} g^2(t)\,dt} \qquad 6.15$$

from Parseval's theorem. Now one form of the Schwarz inequality[138] states

$$\left| \int_a^b u(x)v(x)\,dx \right|^2 \leq \int_a^b |u(x)|^2 dx \int_a^b |v(x)|^2 dx \qquad 6.16$$

so that Eq. 6.15 reduces to the inequality

$$\rho_\omega \leq \frac{2E_r}{No} \qquad 6.17$$

where $E_r = \int_{-\infty}^{\infty} r^2(t)dt$ is the energy of the received signal. Equality in Eq. 6.17 is given (for real r(t) and g(t)) only if

$$g(t) = Cr(t)\exp(\pm j[\omega_c t + \phi]) \qquad 6.18$$

where the gain C and phase ϕ are arbitrary constants.

The upper bound upon the signal to noise ratio ρ_ω imposed by Eq. 6.17 is identical to that at the output from a matched filter and is dependent only upon the energy in the finite sample of the signal. Thus for maximum detectability the period T of the multiplier g(t) should be chosen to coincide with the full duration of the received signal r(t).
Eq. 6.18 states that the optimum multiplier is simply a frequency translated version of r(t), shifted by the detection angular frequency ω_c.

In particular, for an echolocation system using the linear FM waveform

$$s(t) = \exp(j[\omega_o t + \tfrac{1}{2}mt^2] \qquad 0 < t < T \qquad 6.19$$

in complex notation, the received signal after convolution with the environment impulse response (Eq. 6.4) is

$$r(t) = As(t)\exp(j[m\tau t + \omega_o \tau + \tfrac{1}{2}m\tau^2]) \quad \tau < t < T < \tau \qquad 6.20$$

so that if $T >> \tau$, s(t) approximates the optimum multiplier for frequency domain detection at an angular frequency $\omega_c = m\tau$.

Therefore the multiplicative pre-detection processing used in the linear FM system produces an auditory display with near optimum signal to noise properties. This does not mean that the auditory analyzer is optimum for detection of the received signals. In practical situations the effective bandwidth of the auditory system (critical bandwidth)[120] will be much greater than that required for detecting $\Phi_{os}(\omega_c)$ in the presence of noise.

6.3 The Display Statistics of an Extended Object.

An assumption that has been implicit to this point, and one that is normally implicit in narrow-band radar signal analyses[18], is that all reflecting objects in the environment are characterized by an impulse response that is a Dirac delta function (see Section 2.2). In many practical situations a better description of the scattering characteristics of an object is to assume it to be an aggregate of discrete point reflectors, evenly distributed over some finite extent in azimuth and delay.[139] Then the object impulse response as viewed through the two receiver channels may be written

$$h_i(t) = \sum_{n=1}^{N} \frac{B_n A_i(\theta_n)}{\tau_n^2} \delta(t - \tau_n) \qquad 6.21$$

where the amplitudes B_n are normally distributed with zero mean, the azimuthal responses $A_i(\theta)$ are as defined in Eq. 5.10 and the delays τ_n are such that the delay extent $\Delta\tau$ is much

less than the mean delay τ. Then let

$$b_n = \frac{B_n}{\tau^2} \qquad 6.22$$

so that Eq. 6.21 may be written as

$$h_i(t) = \sum_{n=1}^{N} b_n A_i(\theta_n)\,\delta(t - \tau_n) \qquad 6.23$$

The discrete point scatterer model of the environment is shown in Fig. 6.1.

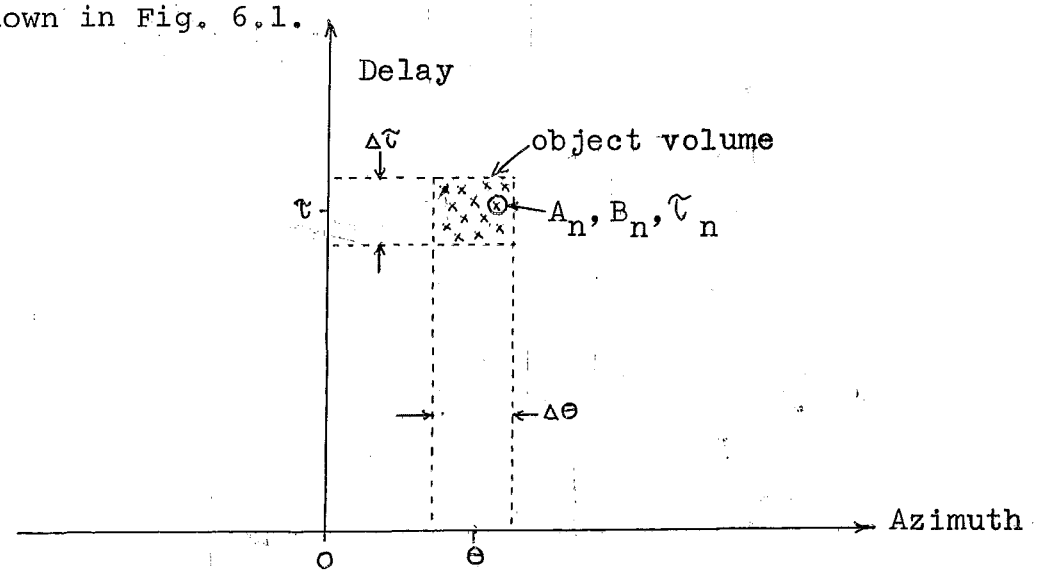


Fig. 6.1. Discrete Point Scatterer Model of an Extended Object.

Upon excitation with the transmitted waveform s(t), the environment will give rise to a received signal in each channel

$$r_i(t) = \sum_{n=1}^{N} b_n A_i(\theta_n)\,s(t - \tau_n). \qquad 6.24$$

Let s(t) be the linear FM waveform

$$s(t) = \exp(j[\omega_o t + \tfrac{1}{2}mt^2]) \qquad 0 < t < T \qquad 6.25$$
$$= 0. \qquad \text{otherwise.}$$

After multiplicative processing and filtering, the signal at the auditory display will be

$$f_i(t) = \sum_{n=1}^{N} b_n A_i(\theta_n) \exp(j[m\tau_n t + \phi_n]) \qquad 6.26$$

where ϕ_n is uniformly distributed between 0 and 2π.

Let $\omega_c = m\tau$ so that

$$f_i(t) = \exp(j\omega_c t) \sum_{n=1}^{N} b_n A_i(\theta_n) \exp(j[m\Delta\tau_n t + \phi_n]) \qquad 6.27$$

where $\Delta\tau_n = \tau_n - \tau$. For large N, it follows from the central limit theorem that the quantity in the summation is normally distributed with zero mean and since the covariance of the real and imaginary parts

$$E\{ \sum_{n=1}^{N} b_n A_i(\theta_n) \cos(m\Delta\tau_n t + \phi_n) . \sum_{n=1}^{N} b_n A_i(\theta_n) \sin(m\Delta\tau_n t + \phi_n) \} = 0 \qquad 6.28$$

they are therefore statistically independent. Because $\Delta\tau << \tau$ the narrow-band approximation may be used to define the quantity in the summation of Eq. 6.27 as the envelope and phase modulation ($y_i(t)$ and $\xi_i(t)$) which will be described by the Rayleigh distribution i.e.

$$f_i(t) = y_i(t) \exp(j[\omega_c t + \xi_i(t)]) \qquad 6.29$$

$$\text{where } p_1(y_i) = \frac{y_i}{2\sigma_i^2} \exp(-\frac{y_i}{2\sigma_i^2}) \qquad 6.30$$

$$p_2(\xi_i) = \frac{1}{2\pi},$$

the time dependence of y_i and ξ_i being implicit.

The variance σ_i^2 may be found as follows. Since θ_n is uniformly distributed over the interval $\theta - \frac{\Delta\theta}{2} < \theta_n < \theta + \frac{\Delta\theta}{2}$, for small $\Delta\theta$ the azimuthal responses $A_i(\theta_n)$ will be approximately uniformly distributed over the range $A_i(\theta - \frac{\Delta\theta}{2}) < A_i(\theta_n) < A_i(\theta + \frac{\Delta\theta}{2})$ with density $\Delta\theta A_i'(\theta)$ where $A_i'(\theta) = \frac{\partial}{\partial\theta} A_i(\theta)$. The variance σ_i^2 will be equal to the variance of the real and imaginary parts in Eq. 6.28 i.e.

$$\sigma_i^2 = \frac{N\sigma_b^2}{2}\left[A_i^2(\theta) + \frac{1}{3}\left(\frac{\Delta\theta}{2} A_i'(\theta)\right)^2\right] \qquad 6.31$$

$$= \frac{N\sigma_b^2 A_i^2(\theta)}{96(\alpha K)^2}\left[48(\alpha K)^2 + (\Delta\theta(\theta + (-1)^i\alpha))^2\right] \qquad 6.32$$

where σ_b^2 is the variance of the amplitudes b_n and α, K are the display parameters defined in Eqs. 5.5 and 5.6.

Consider now the covariance between the envelope generating Gaussian distributions in the two channels. Let

$$x_i = \sum_{n=1}^{N} b_n A_i(\theta_n)\exp(j[m\Delta\tau_n t + \phi_n]) \qquad 6.33$$

so that

$$E\{x_1 x_2\} = E\{\sum_{n=1}^{N} b_n A_1(\theta_n)\exp(j[m\Delta\tau_n t + \phi_n])$$

$$\cdot \sum b_n A_2(\theta_n)\exp(j[m\Delta\tau_n t + \phi_n])\} \qquad 6.34$$

$$= \sum_{n=1}^{N} E\{b_n^2 A_1(\theta_n) A_2(\theta_n)\exp(j2[m\Delta\tau_n t + \theta_n])\} \qquad 6.35$$

or $E\{x_1 x_2\} \neq 0$.

The two modulation functions $y_1(t)$ and $y_2(t)$ will not therefore be statistically independent. The effect of the object extent upon the azimuth estimator $\hat{\theta}$ derived from instantaneous values of the envelope amplitudes (Section 5.7.2) cannot be stated without knowledge of the joint probability density $p(y_1, y_2)$. However, for the second variant of the estimation model in Section 5.7 with a large post-detection integration time constant, the limiting form as T_2 tends to infinity may be found because it is not affected by statistical dependence i.e.

$$\lim_{T_2 \to \infty} \hat{\theta} = K \log_e \frac{E\{y_1\}}{E\{y_2\}} \qquad 6.36$$

where y_i are defined by Eq. 6.29 and

$$E\{y_i\} = \sqrt{\frac{\pi}{2}}\, \sigma_i$$

$$\lim_{T_2 \to \infty} \hat{\theta} = K \log_e \frac{A_1(\theta)\left[48(\alpha K)^2 + (\Delta\theta(\theta-\alpha))^2\right]^{\frac{1}{2}}}{A_2(\theta)\left[48(\alpha K)^2 + (\Delta\theta(\theta+\alpha))^2\right]^{\frac{1}{2}}} \qquad 6.37$$

$$= \theta + \tfrac{1}{2} K \log_e \frac{48(\alpha K)^2 + (\Delta\theta(\theta-\alpha))^2}{48(\alpha K)^2 + (\Delta\theta(\theta+\alpha))^2} \qquad 6.38$$

The second term represents a bias in the estimator toward the median plane. Fig. 6.2 shows the magnitude of the bias as a function of the angular extent $\Delta\theta$ for normalised

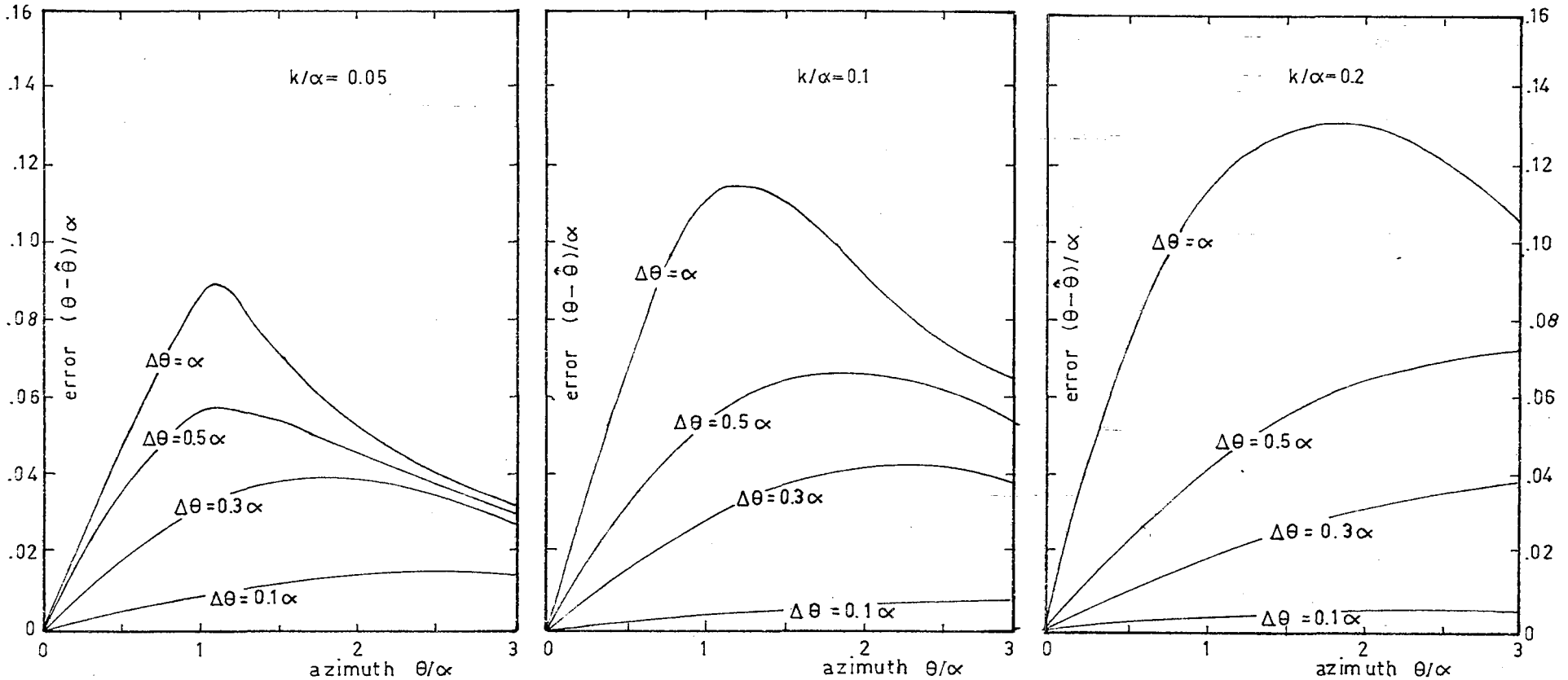


Fig. 6.2 Error in the azimuth estimator caused by object extent.

values of the display parameters (Eq. 5.11).

It is interesting to note in passing that the envelope modulation functions y_i contain information on the delay extent $\Delta\tau$ of the object. In his classic work on the properties of noise, Rice[131] showed that the expectation of the number of envelope maxima per second, M, for Rayleigh noise is a function only of the bandwidth

$$E\{M\} = 0.6411\ (2\pi\Delta\omega)$$
$$= 4.02\ (m\Delta\tau) \qquad 6.39$$

from Eq. 6.27.

It has been shown in this section that for objects with a finite extent in both range and azimuth, the signals at the auditory display are not pure tones as assumed in the derivation of the display in Chapter 5, but have envelopes described by the Rayleigh distribution. There is a need therefore for a study of the psychophysics of the localization of Rayleigh noise, as a function of bandwidth and dichotic statistical dependence, to define the display in these situations.

6.4 The Effect of Object Movement Upon the Range Dimension of the Display.

The Doppler effects, caused by radial motion of a scattering object, must be considered as a time compression (or dilation) for the wide-band signals used in the FM mobility

aids. The normal radar approximation of a frequency shift is not valid when the bandwidth is an appreciable fraction of the carrier frequency.

Consider an environment with a time-varying impulse response $h_e(t)$ of the form

$$h_e(t) = \frac{B}{\tau^2(t)} s(t - \tau(t)) \qquad 6.40$$

where the time dependence of the delay $\tau(t)$ is due to the movement of the object. The received signal $r_i(t)$ will be

$$r_i(t) = \frac{BA_i(\theta)}{\tau^2(t)} s(t - \tau(t)) \qquad 6.41$$

and in particular, if $s(t)$ is the linear FM waveform (Eq. 6.25) $r_i(t)$ will be

$$r_i(t) = \frac{BA_i(\theta)}{\tau^2(t)} \exp(j[\omega_o(t - \tau(t)) + \tfrac{1}{2}m(t - \tau^2(t))]) \qquad \tau(0) < t < T + \tau(T) \qquad 6.42$$

Let the mean delay $\tau(\frac{1}{2}T)$ be much greater than the increment in delay $(\tau(T) - \tau(0))$ and let

$$b = \frac{B}{\tau(\frac{1}{2}T)} \qquad 6.43$$

so that after the multiplicative processing and filtering the display signals $f_i(t)$ may be approximated as

$$f_i(t) = bA_i(\theta)\exp(j[\omega_o\tau + m\tau t - \tfrac{1}{2}m\tau^2)] \qquad 6.44$$

where the time dependence of τ is implicit. The instantan-

eous angular frequency $\omega_c(t)$ is

$$\omega_c(t) = m\tau + (\omega_o + m(t - \tau)) \frac{d\tau}{dt} \qquad 6.45$$

The first term in Eq. 6.45 is clearly the quasi-static range coding, the second is an impressed time-varying error which may be written as an error in the range estimator

$$\varepsilon = R(t) - \frac{c\omega_c(t)}{2m} \qquad 6.46$$

$$= -\frac{c}{2m}(\omega_o + m(t - \tau))\frac{d\tau}{dt} \qquad 6.47$$

where c is the velocity of propagation.

For an object moving with uniform radial velocity v toward the transmitter the delay will be

$$\tau = \frac{2(R_o - vt)}{c-v} \qquad 6.48$$

where R_o is the range at t = 0. Substitution into Eq. 6.45 gives

$$\omega_c(t) = (1 - \xi)\omega_o + \frac{\xi}{\zeta} m\tau_o + (1 - \xi^2)mt \qquad 6.49$$

where $\tau_o = \frac{2R_o}{c}$; $\xi = \frac{c+v}{c-v}$; $\zeta = 1 - \frac{v}{c}$. If $c >> v$ the instantaneous angular frequency will be approximately

$$\omega_c(t) \doteq m\tau_o - \frac{2v}{c}\omega_o + \frac{4v}{c}t \qquad 6.50$$

and the equivalent range error is

$$\varepsilon \doteq v\left(\frac{T}{\Delta\omega} + 2t\right) \qquad 6.51$$

where $\Delta\omega = \frac{mT}{\omega_o}$ is the fractional bandwidth of s(t).

For radial motion with constant acceleration, i.e. for

an equation of motion of the form

$$R = R_o - (ut + \tfrac{1}{2}ft^2) \qquad 6.52$$

where u is the radial velocity when t = 0 and f is the radial acceleration; the delay τ is

$$\tau = \frac{1}{4}f[((c - v)^2 + 2fR)^{\frac{1}{2}} - (c - v)] \qquad 6.53$$

and

$$\frac{d\tau}{dt} = \frac{2f\tau - 4v}{f\tau + 2(c-v)} . \qquad 6.54$$

Substitution into Eq. 6.45 allows numerical computation of $\omega_c(t)$.

The effects of the Doppler time compression/dilation upon the range dimension of the auditory display are shown for a typical trajectory in Fig. 6.3. The echolocation system parameters are shown in the figure. The object trajectory consists of a period of constant acceleration from a stationary position, a period of constant radial velocity, and uniform deceleration to zero velocity. The range was chosen to be within the operating range of the binaural mobility aids. The sensitivity of the signal processing to Doppler effects is clearly demonstrated in the figure, for although the maximum radial velocity of the trajectory is less than 0.005 of the acoustic velocity of propagation in air the maximum range error indicated by $\omega_c(t)$ is approximately 0.3 of the true range.

The effect of the sweep direction (the sign of m

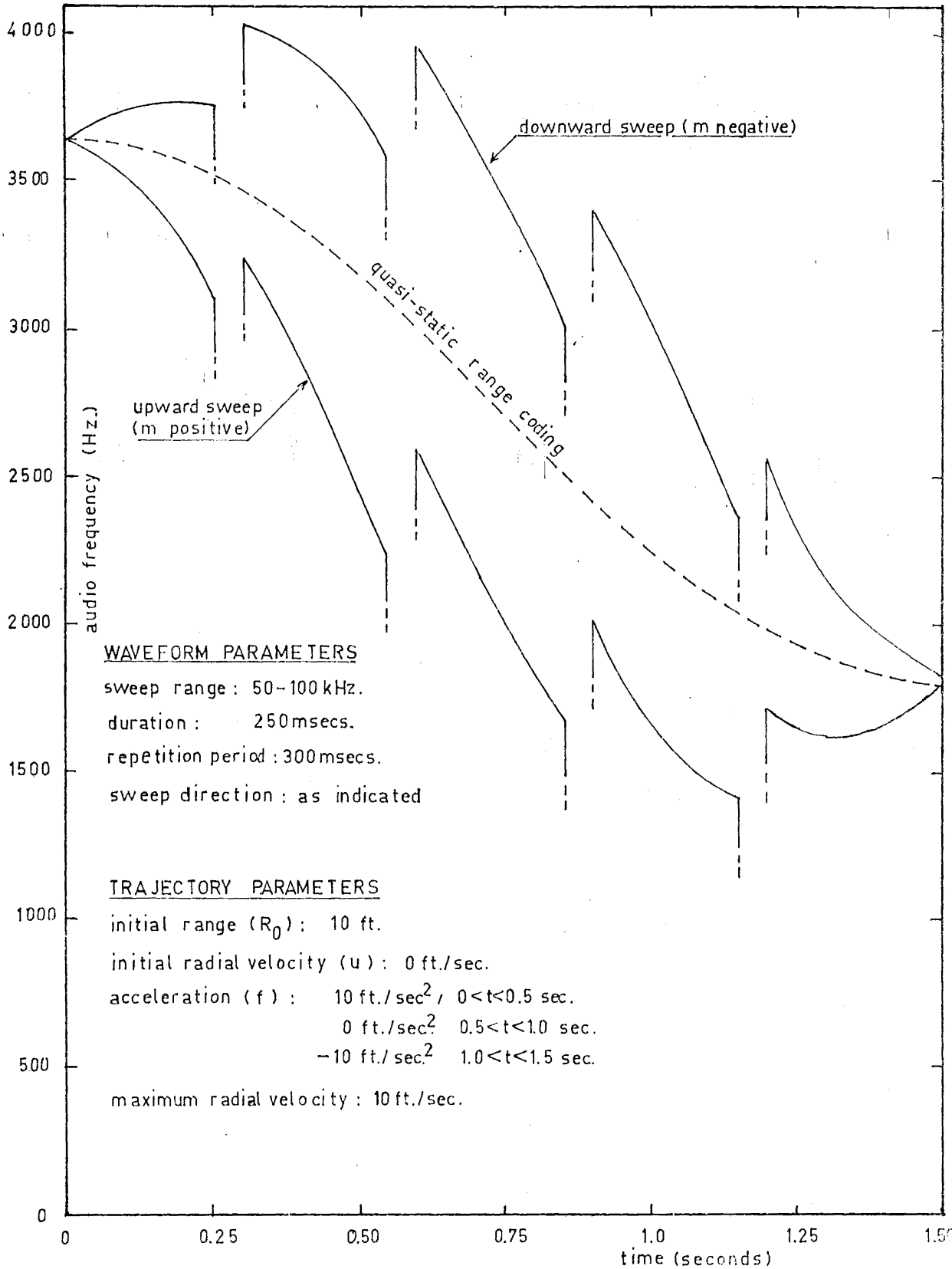


Fig. 6.3 An example of the distortion in the range coding caused by Doppler effects. The maximum velocity in the trajectory shown is 5 ft/sec.

in Eq. 6.25), is important in the context of the kinaesthetic sensory feedback from a mobility aid (Section 1.3). In the example given in Fig. 6.3 the initial acceleration toward the system appears to make the object recede when m is negative, and the radial acceleration is accentuated for positive m. The range of an object approaching with constant velocity will be overestimated for m negative, and underestimated for m positive. The effect of these distortions in the range coding upon mobility performance was investigated in the following experiment.

A binaural mobility aid, as designed by Martin[14], was modified so that the sweep direction could be reversed with a negligible change in range coding. A single female subject was instructed to walk toward a 2 inch diameter vertical steel pole and stop at a "safe" distance. The subject was always started at a distance of 15 feet from the pole but was randomly oriented relative to it so that the task included the detection of the object and motion toward it. The particular subject was chosen because she had shown the greatest adaptation to the auditory display out of a group of four sighted (but blindfolded) subjects who had participated in the initial mobility training programme as designed by R.W. Pugh. She was, at the time, the most experienced user of the mobility aid, and had received extensive training with the negative m range coding.

The sweep direction was randomly changed between trials. A total of 20 observations of the stopping distance for each sweep direction were recorded. No training in the task was given and the mobility aid was switched off after each trial.

The results (measured to the tip of the shoe nearest the pole) were

m positive	mean stopping distance 22.7"
m negative	mean stopping distance 13.0"

with a standard deviation in each case of approximately 8". The significance of the difference between the mean values of the stopping distances was computed from the t statistic ($t = 3.84$, $df = 38$: $p < 0.01$). It was concluded that for this subject the sweep direction, and the associated Doppler distortion in the range coding, was significantly affecting the motor actions in a mobility task. In the presence of an impending obstacle, the display with m positive caused the subject to take earlier avoiding action.

6.5 The Effect of the Receiving Aperture Characteristics upon the Auditory Display

The previous analyses have assumed non-dispersive receiving apertures with an azimuthal response function $A_i(\theta)$ defined in Eq 5.10. The present state of the technology of wide-band air-borne ultrasonic transducers does not allow non-dispersive apertures to be constructed, so that the received signals $r_i(t)$ will be modified in structure

by the aperture response. Let the waveform $s(t)$ be transmitted into an environment consisting of a single object at azimuth θ.

Then

$$r_i(t) = (s(t) \otimes h_e(t)) \otimes h_i(t|\theta) \qquad 6.55$$

where $h_e(t)$ is the environmental impulse response, $h_i(t|\theta)$ is the impulse response of the receiving aperture of the ith channel ($i = 1,2$). Note the similarity of this equation to Eq. 2.6 where it was stated that the recognition of $h_i(t|\theta)$ could serve to characterize θ. It was subsequently shown that because the signal processing used to generate the range dimension of the display produced narrow bandwidth auditory stimuli, the use of auditory directional coding of this form was inhibited.[22]

In the frequency domain, Eq. 6.55 may be written

$$R_i(\omega) = S(\omega)H_e(\omega)H_i(\omega|\theta) \qquad 6.56$$

where $R_i(\omega), S(\omega), H_e(\omega), H_i(\omega|\theta)$ are the Fourier transforms of $r_i(t), s(t), h_e(t), h_i(t|\theta)$ respectively. The idealised aperture would have a response

$$H_i(\omega|\theta) = CA_i(\theta) \qquad 6.57$$

where C is an arbitrary constant. Smith[140] has developed underwater sonar arrays with a non-dispersive azimuthal response over a wide bandwidth by using a frequency dependent amplitude taper, but these arrays exhibit side-lobes which render them unsuitable for the auditory azimuth estimator $\hat{\theta}$ (see Section 5.5).

This section examines three aspects of the aperture response. First the ideal linear aperture distribution to generate $F(\theta)$ satisfying Eq. 5.9 is derived, then the response of the electrostatic transducers used in the mobility aid is described. Finally the definition and measurement of the variation of the audio voltage with azimuth is described.

6.5.1 The Linear Aperture Function to Generate a Far-Field Azimuthal Response of the form $\exp(-\gamma\theta^2)$

Let the idealised variation of auditory stimulus level be given by Eq. 5.15 i.e.

$$F(\theta) = \exp(-\gamma\theta^2) \qquad 6.58$$

so that the value of α must be chosen according to Eq. 5.17 to match the display to the individual.

The Fourier transform relationship between the far-field pressure response $P(u)$ where $u = \sin\theta$ and the aperture distribution $f(x)$ for a linear aperture is

$$P(u) = \int_{-\infty}^{\infty} f(x) e^{jkxu}\, dx \qquad 6.59$$

where k is the wavenumber ($k = \frac{2\pi}{\lambda}$). The corresponding inverse transform is

$$f(x) = \int_{-\infty}^{\infty} P(u) e^{-jkxu}\, du \qquad 6.60$$

where the contour of integration involves θ in the imaginary region ($\sin\theta > 1$) as well as real θ. Define $P(u)$ to be

zero for $|u| > 1 - \xi$ where $0 < \xi < 1$. If $P(u)$ is even Eq. 6.60 may be rewritten

$$f(x) = \int_{-1}^{1} P(u)\cos(kxu)\,du \qquad 6.61$$

which upon a change of variable and substitution for $F(\theta)$ gives

$$f(x) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp(-\gamma\theta^2)\cos(kx\sin\theta)\cos\theta\,d\theta. \qquad 6.62$$

Now because $P(u)$ was defined to be zero in the imaginary region, the limits of the integral may be extended, so that for γ greater than unity a good approximation is

$$f(x) = \int_{-\infty}^{\infty} \exp(-\gamma\theta^2)\cos(kx\sin\theta)\cos\theta\,d\theta \qquad 6.63$$

$$= \int_{-\infty}^{\infty} \exp(-\gamma\theta^2)\left[J_0(kx) + 2\sum_{n=1}^{N} J_{2n}(kx)\cos(2n\theta)\right]\cos\theta\,d\theta \qquad 6.64$$

$$= J_0(kx)\int_{-\infty}^{\infty} \exp(-\gamma\theta^2)\cos\theta\,d\theta + 2\sum_{n=1}^{\infty} J_{2n}(kx)\int_{-\infty}^{\infty} \exp(-\gamma\theta^2)\cos(2n\theta)\cos\theta\,d\theta \qquad 6.65$$

$$= \sqrt{\frac{\pi}{\gamma}}\exp(-4\gamma)\left[J_0(kx) + 2\sum_{n=1}^{\infty} \exp\left(-\frac{n}{\gamma^2}\right)\cosh\left(\frac{n}{\gamma}\right)J_{2n}(kx)\right] \qquad 6.66$$

where $J_\nu(x)$ is the Bessel function of the first kind. Computed curves for integer values of γ, normalised so that $f(0) = 1$, are shown in Fig. 6.4. The inverse transform was computed and showed that for $\gamma > 2$, truncation of the aperture at $\pm 2\lambda$ introduced negligible errors into $F(\theta)$.

The frequency dependent amplitude taper function required to produce a non-dispersive $F(\theta)$ may be calculated directly from Eq. 6.66 by fixing x and varying k. This is not investigated further here because, as discussed below, no method of controlling the surface vibrational characteristics of the electrostatic transducers used in the binaural mobility aid has been found.

6.3.2 Measured Vibrational Characteristics of Electrostatic Transducers

The solid-dielectric electrostatic transducers used in the mobility aid have been extensively investigated and described by Martin.[14] The construction of these transducers is shown diagrammatically in cross-section in Fig. 6.5. A diaphragm of 0.00015" thickness Mylar*, metallized on one side, is held under slight tension against a circular 1cm diameter back-plate which has concentric rectangular grooves 0.02" deep and 0.01" wide at radial increments of 0.02".

The surface vibrational amplitude distribution of these units was investigated with a probe microphone (O.D. = 0.035", I.D. = 0.020"). The microphone, developed by the author,

*Trade name: Du Pont

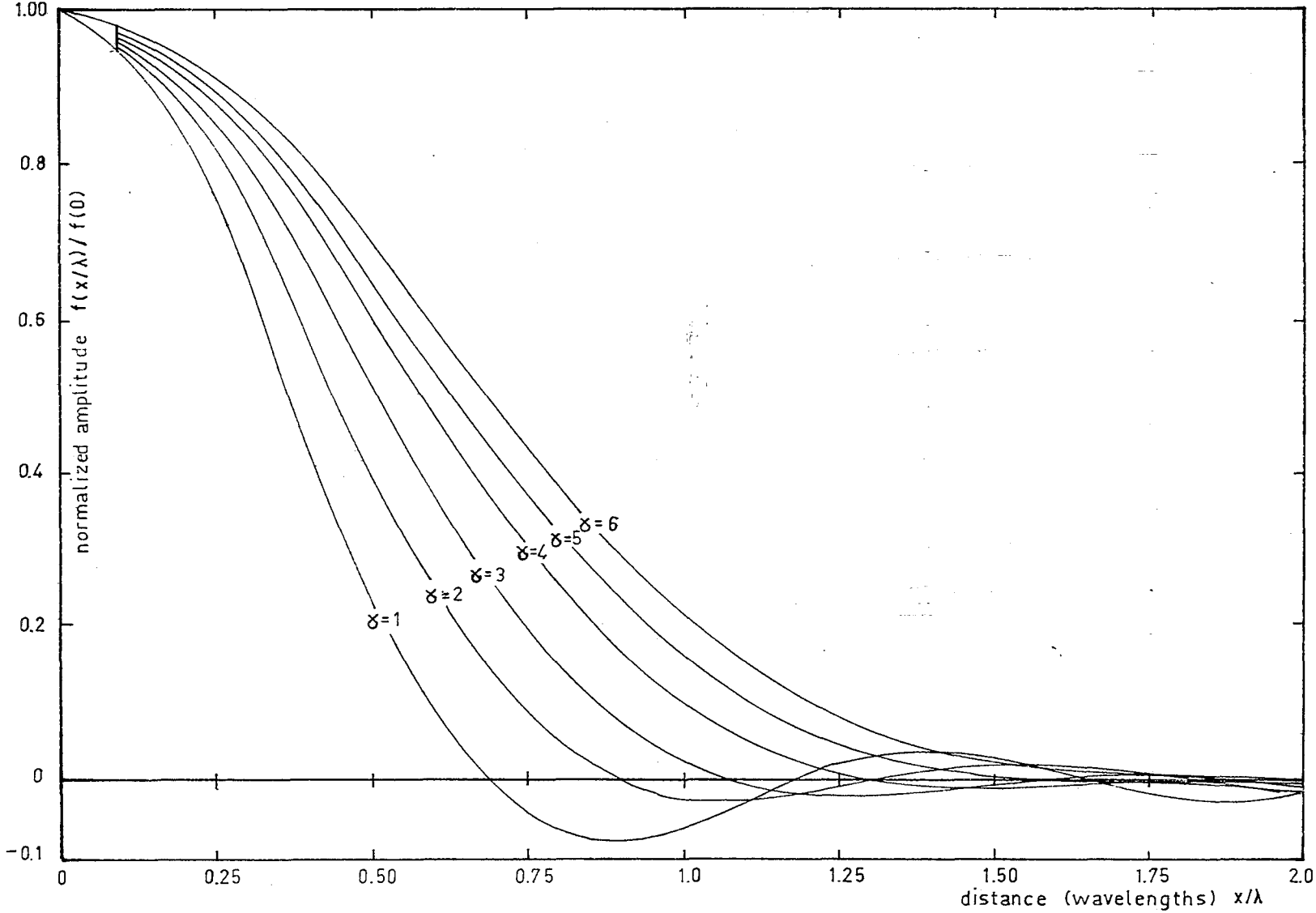


Fig. 6.4 The linear aperture excitation/response function to generate a far field angular response function of the form $\exp(-\gamma\theta^2)$.

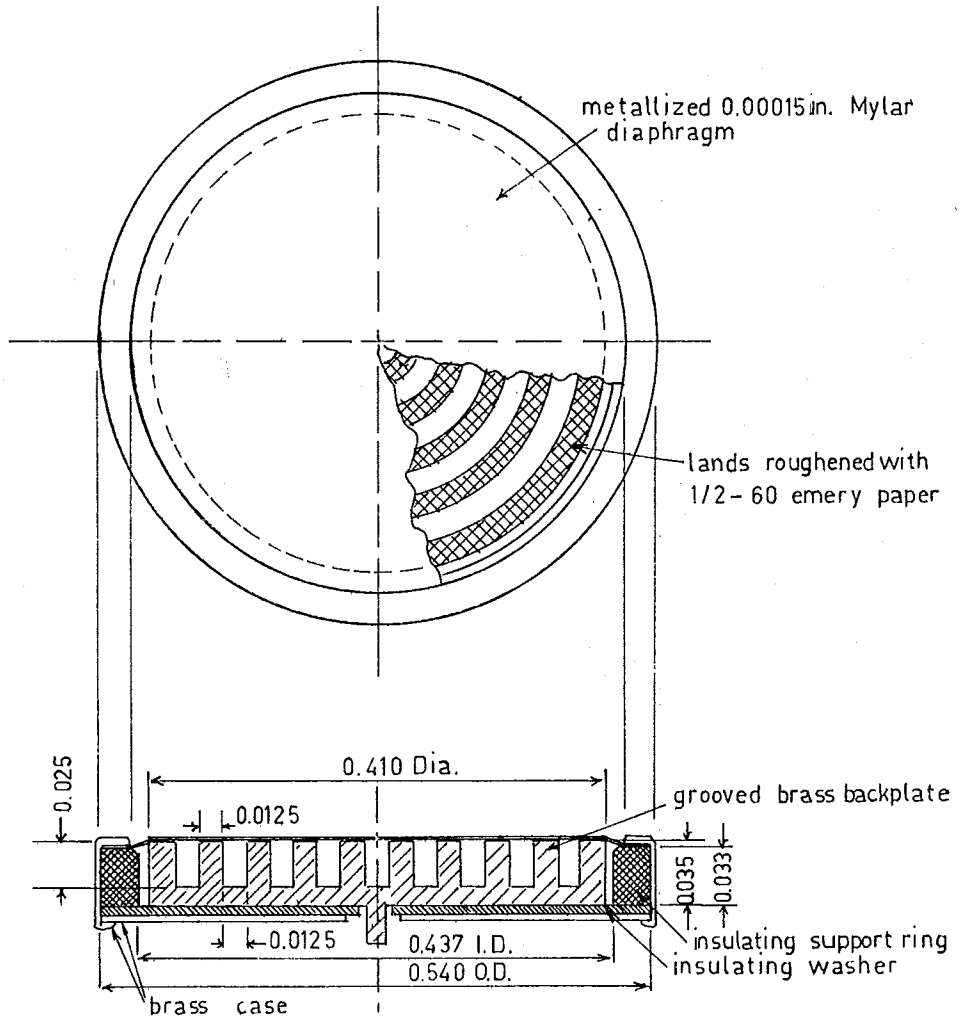


Fig. 6.5 Solid-dielectric electrostatic ultrasonic transducer construction (after Clark).

is described by Martin.[14] Care was taken in the construction to minimise interference with the pressure distribution, and measurements were made at several heights above the membrane (0.2 mm to 2 mm) to determine the effects of the probe. It was concluded that the probe was not significantly distorting the pressure distribution.

Some typical measurements are shown in Fig. 6.6 as

i a contour representation of the amplitude distribution across a transducer diaphragm at 40kHz

ii amplitude variations across two diameters at right angles at four different frequencies.

Other measurements made using the probe microphone are described by Martin.[14] These results showed that at frequencies below 40kHz the amplitude distribution was unimodal, but at higher frequencies higher order vibrational modes introduced further maxima. Nominally identical transducers were found to bear little resemblance to one another in their response patterns, particularly at frequencies greater than 60kHz. It would appear that the vibrational amplitude is subject to local influences such as back-plate roughness, residual diaphragm polarization, and local diaphragm stress.[14] Such factors have meant that it has not yet been possible to construct transducers with predictable vibrational characteristics. Until the mode of operation of these transducers is known it will not be possible to predict or control the azimuthal response.

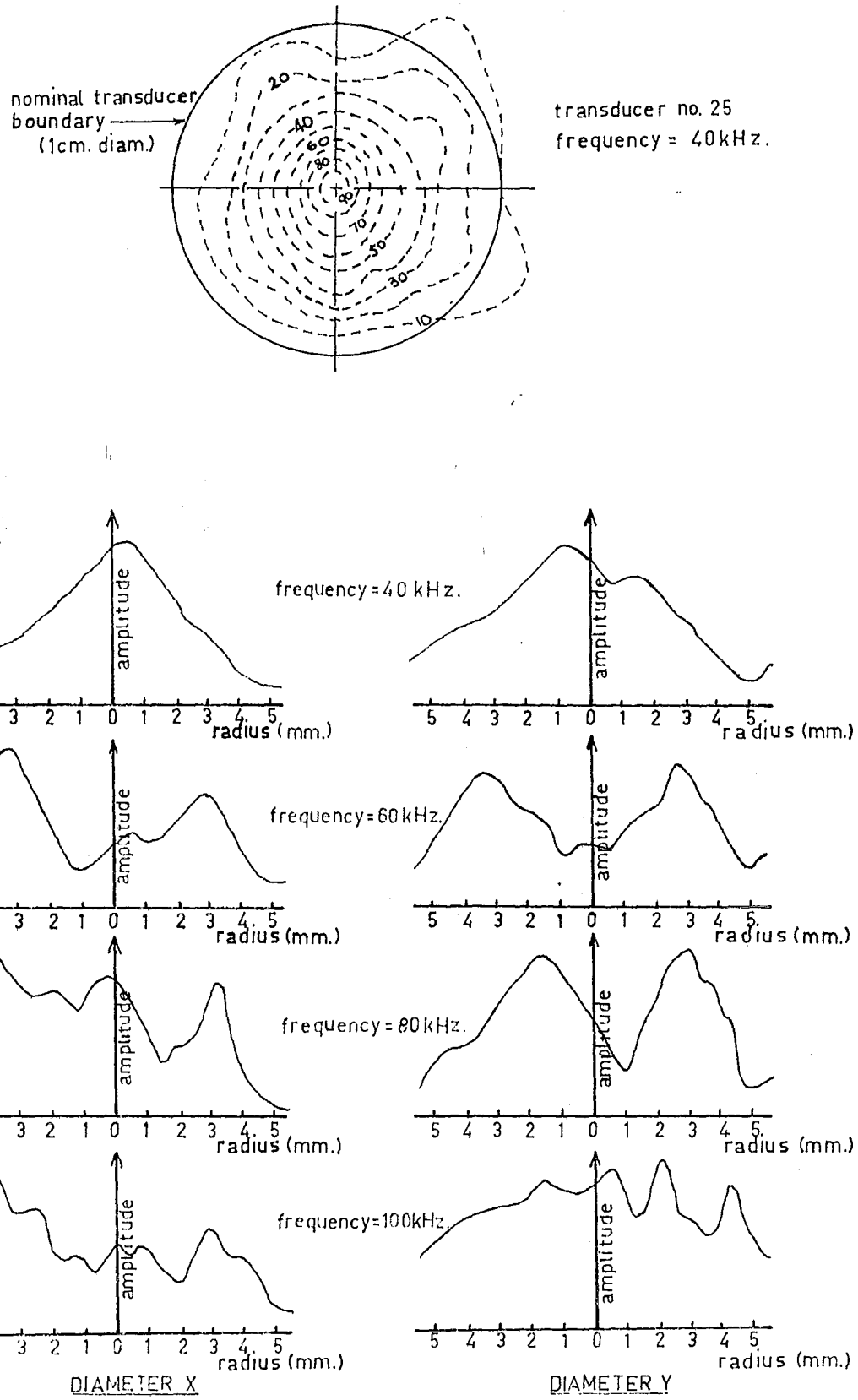


Fig. 6.6 Examples of the variation of surface response of a transducer.

Above: Contour representation
Below: Response across two diameters at right angles to one another.

6.5.3 The Audio Azimuthal Response with a Wide-band Echolocation Waveform

The results of the dichotic localization experiment using the prototype system described in Section 3.6, indicated that the differences in the envelope modulation functions were not affecting the image placement. The correlation between the centring IAD and the mean IAD between the two channels ($r = 0.97$) indicated that a suitable empirical form of the estimator was

$$\hat{\theta} = K \log_e \frac{\overline{y_1(\theta)}}{\overline{y_2(\theta)}} \qquad 6.67$$

where $\overline{y_i(\theta)}$ $(i = 1,2)$ are the mean envelope levels in the two channels. This emprical estimator is similar to that proposed in Section 5.7.3. The experiment described in Section 4.4 also indicated that for small envelope modulation differences with repetitive stimuli Eq. 6.68 is a suitable description of the auditory localization function.

Let $\theta_i = (\theta + (-1)^i \alpha)$ and write

$$\overline{y_i(\theta)} = \overline{y(\theta_i)} \quad . \qquad 6.68$$

If the transmitted waveform $s(t)$ has finite energy, then

from Parseval's theorem the received energy $P(\theta_i)$ will be

$$P(\theta_i) = \int_{-\infty}^{\infty} |s(\omega)|^2 |H(\omega|\theta_i)|^2 |H_e(\omega)|^2 d\omega \qquad 6.69$$

from Eq. 6.57, where $H(\omega|\theta_i) = H_i(\omega|\theta)$.

If s(t) is the linear FM waveform (Eq. 6.25) with a large time-bandwidth product, $s(\omega)$ may be approximated by[141]

$$\begin{aligned} s(\omega) &= \sqrt{\frac{1}{2m}} && \omega_o < |\omega| < \omega_o + mT \\ &= 0. && \text{otherwise} \end{aligned} \qquad 6.70$$

For simplicity assume that $|H_e(\omega)| = 1$ then the mean envelope amplitudes may be related to the rms amplitude by a form factor $f(\theta_i)$

$$\overline{y(\theta_i)} = \frac{f(\theta_i)}{T\sqrt{m}} \left[\int_{\omega_o}^{\omega_o + mT} |H(\omega|\theta_i)|^2 d\omega \right]^{\frac{1}{2}} \qquad 6.71$$

The azimuthal response $\overline{y(\theta)}$ was measured using the set-up shown in Fig. 6.7. The MkI version of the mobility aid[14] used to generate the linear FM waveform was modified so that the sweep limits could be altered. It was found that nominally identical transducers differed significantly in the variation of average envelope amplitude, Martin found similar differences in single frequency polar responses. The responses of two typical transducers are shown in Fig. 6.8. For most transducers it was

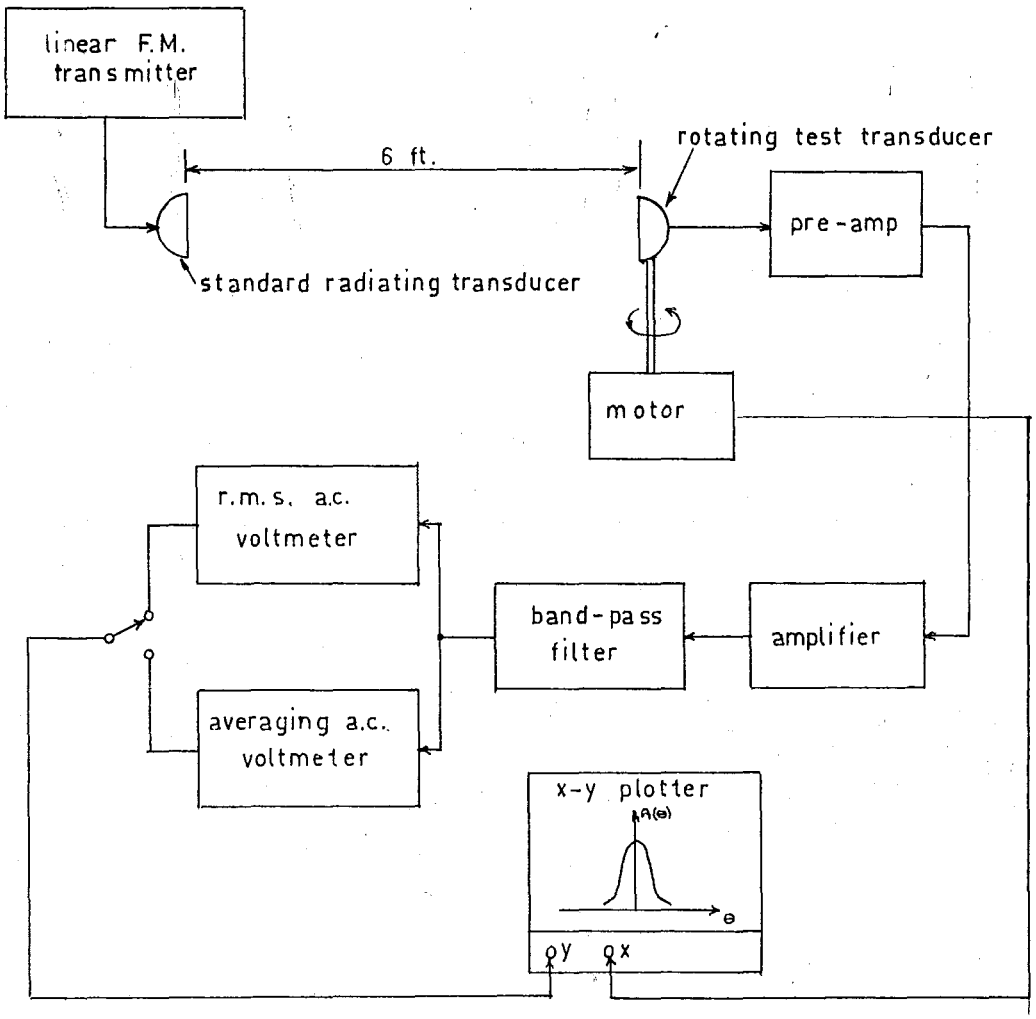


Fig. 6.7 Experimental set up to measure the azimuthal variation of audio voltage.

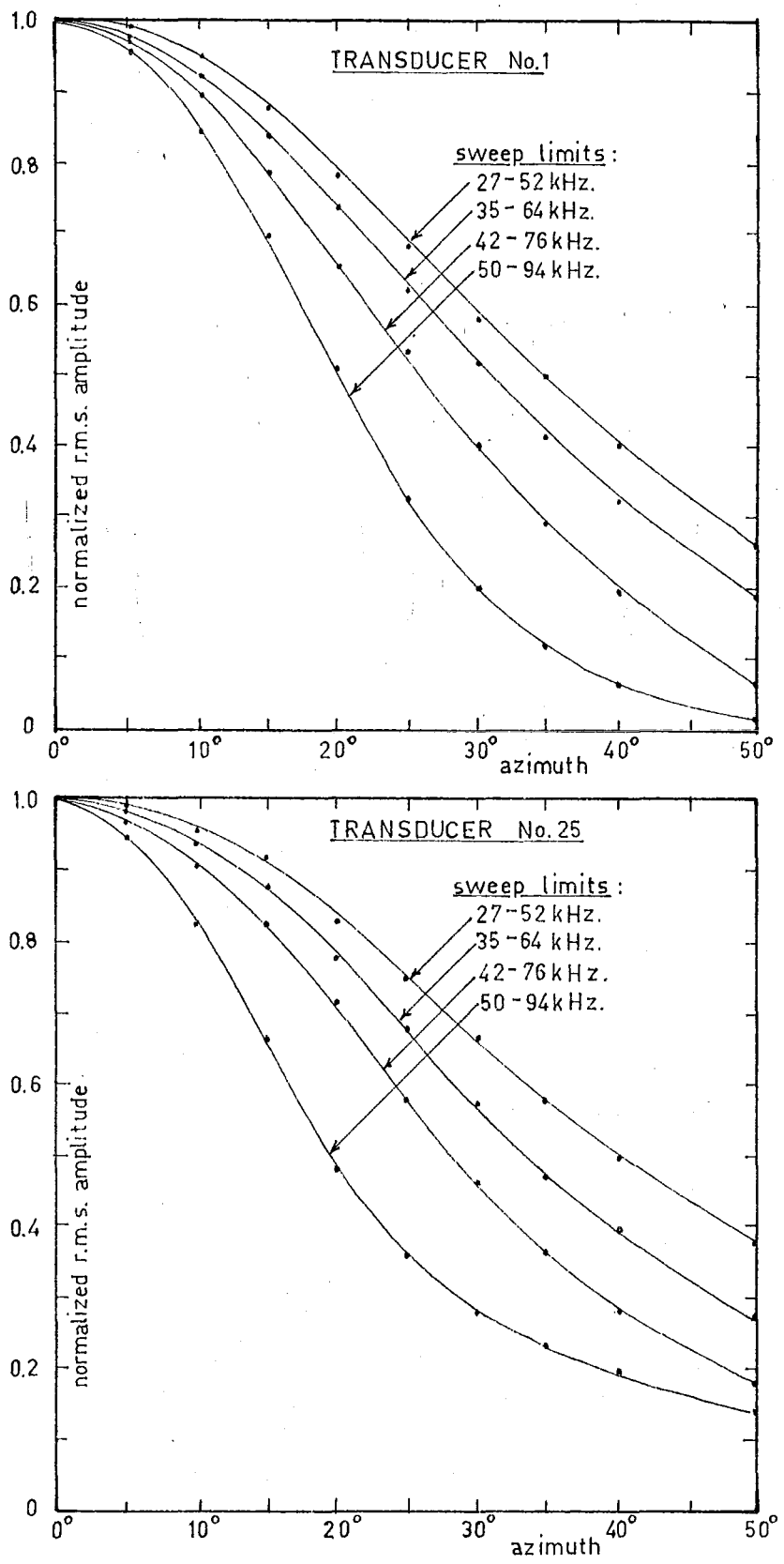


Fig. 6.8 Azimuthal variation of audio voltage for two transducers.

possible to fit a curve of the form $F(\theta) = \exp(-\gamma\theta^2)$ to the responses with a negligible error to an azimuth of approximately 30°. The procedure for the manufacture of mobility aids has been to select, from a batch of transducers, pairs having a response closely approximating $F(\theta)$ given by Eq. 6.58 with an equal 6db beamwidth.

Fig. 6.9 shows the normalised variation of rms amplitude for circular piston mounted in an infinite baffle, with a bandwidth $mT = \omega_o$. The curves were computed by substitution of the directivity function for a piston into Eq. 6.72 i.e.

$$H(\omega|\theta) \propto \omega \frac{2J_1\left(\frac{\omega a \sin\theta}{c}\right)}{\frac{\omega a \sin\theta}{c}} , \qquad 6.72$$

where a is the piston radius.

In Fig. 6.10 the average 6db beamwidth $\psi_{\frac{1}{2}}$ of ten transducers is plotted against the median frequency $\omega_o + \frac{1}{2}mT$. For comparison the computed 6db beamwidth for a piston source with an octave bandwidth (from Eqs 6.71 and 6.72), and the beamwidth of a piston source at the median frequency are also plotted. The empirical least squares line of best fit for the measured values was

$$\psi_{\frac{1}{2}} = 51 - 0.207\ (f_1 + f_2) \text{ degrees}, \qquad 6.73$$

where f_1 and f_2 are the sweep limits in kHz.

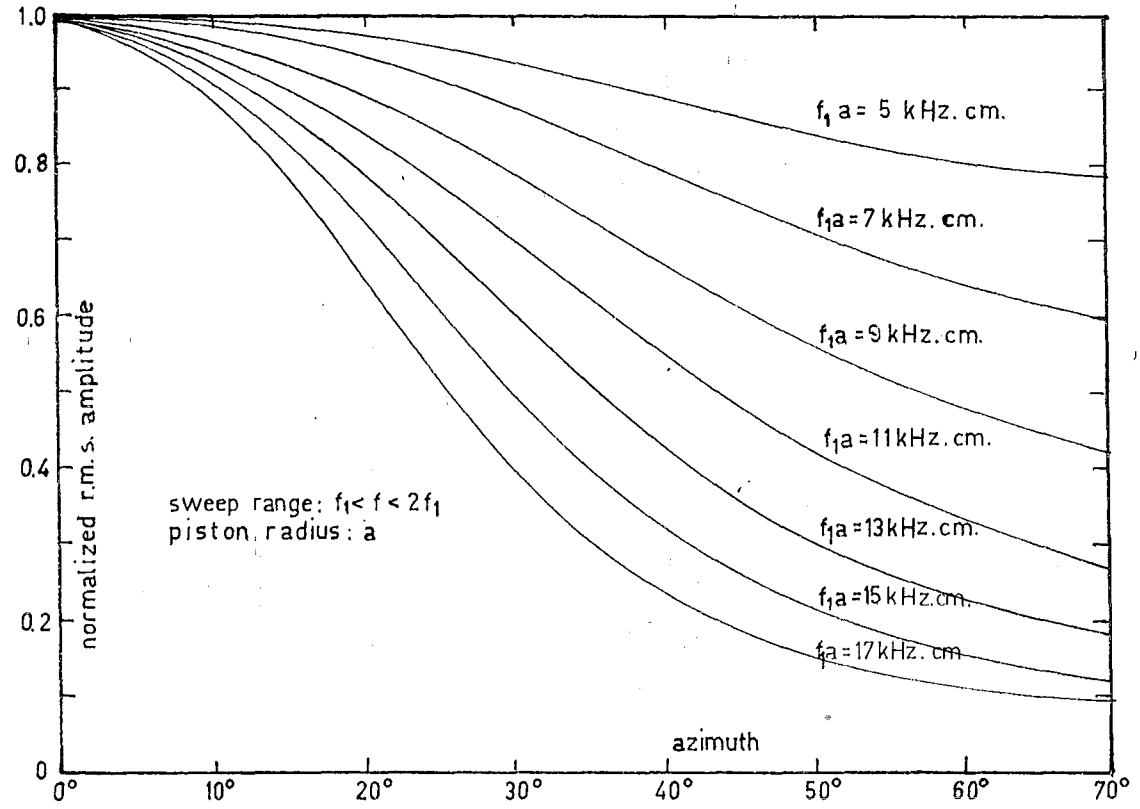


Fig. 6.9 Computed variation of rms audio voltage for a piston aperture in an infinite baffle.

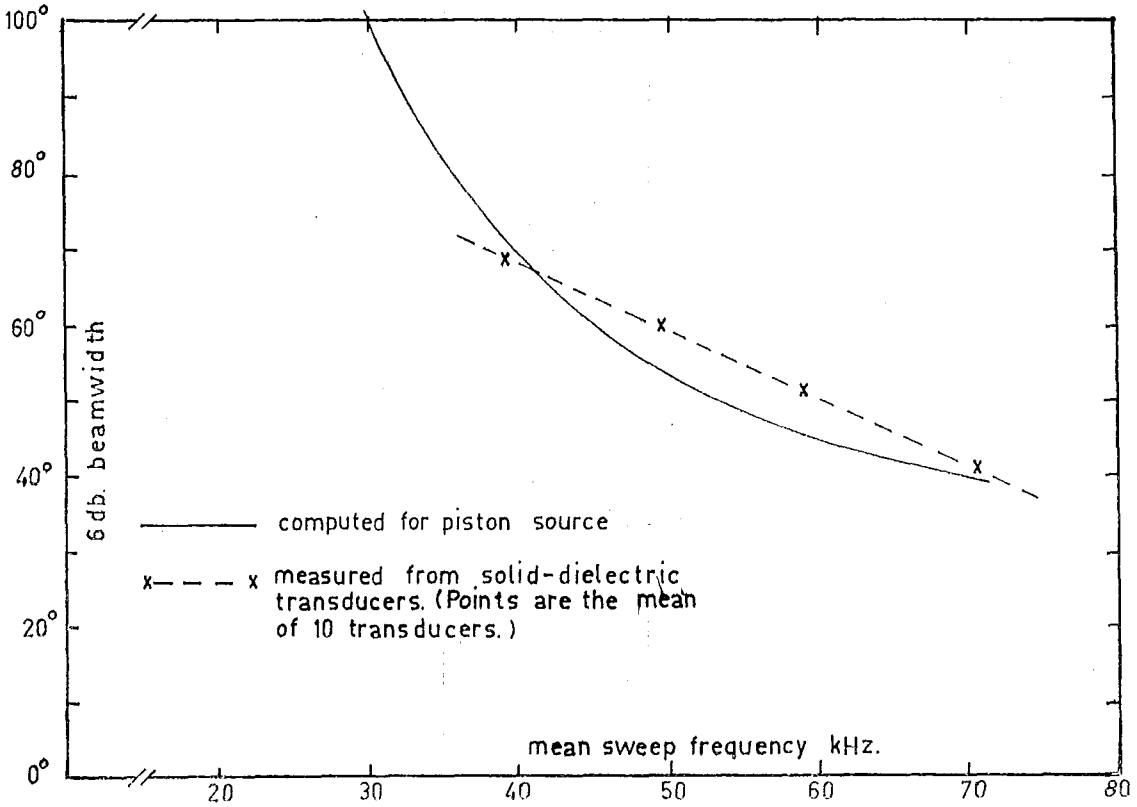


Fig. 6.10 Comparison between measured audio beamwidths of electrostatic transducers and the computed beamwidth for a piston source.

CHAPTER 7

MEASURES OF AUDITORY RESOLUTION WITH A FREQUENCY DOMAIN DISPLAY

7.1 Introduction

It was shown in Chapter 2 that to overcome the inhibitory nature of the precedence effect, in the localization of the components of a multiple object display, it was necessary to code propagation delay into the frequency domain. There is, however, little available evidence of the perceptual improvement in resolution afforded by a frequency domain display. It could be expected that an effective resolution interval might be defined in terms of the critical bandwidth, as proposed in an earlier work[143]. It is shown in this chapter that a simple definition of perceptual resolution of this type does not form an adequate description of the auditory resolution capabilities in a two-dimensional frequency x IAD display.

Two definitions of auditory resolution were investigated by the author. The first is simply the increment in the stimulus magnitude (frequency) necessary for the recognition of the presence of two simultaneous components; the second definition requires that the resolvable increment in stimulus magnitude is large enough for both recognition and lateral identification of each component. Two experiments

are described in this chapter. The first measured the minimum frequency difference at which subjects were able to hear two tones simultaneously, the second experiment required subjects to identify one tone and determine its lateralization in the presence of an interfering tone.

The second experiment used an on-line hybrid computing system with complete control of the experimental strategy by the digital computer. The techniques of interfacing the subject to the computer are believed to be original.

7.2 Simultaneous Two-Tone Frequency Resolution

Thurlow and Bernstein[142], and Plomp and Steeneken[143] have investigated the monaural difference in frequency necessary for the perception of both components of a two-tone stimulus. The first group used the method of limits with two subjects to obtain the threshold for the perception of the component pitches at a sensation level of 30db. Plomp and Steeneken used a larger sample of 20 subjects and the method of adjustment (average error) to obtain estimates of the frequency difference between two tones adjusted to the criteria of

i maximal dissonance (roughness)

ii minimum interference between the tones.

The second of these criteria may be used as a definition of auditory resolution in frequency. The authors reported that judgments made against this criterion agreed well with the critical bandwidth measurements by Zwicker, Flottorp and

Stevens[144] between 700 Hz and 3000 Hz. At lower and higher frequencies the experimental frequency differences were smaller than the commonly accepted critical bandwidths.

The results of these two monaural studies are summarised in Fig. 7.1. Both references reported wide individual differences between subjects.

The experiment reported here used the method of adjustment to measure the frequency difference between the two tones to the criterion that subjects heard the two components "clearly and distinctly". This study differed from the two cited above in that binaural stimuli were used in a factorial design with IAD and pulse duration as additional factors.

Experimental Method: The subject was seated in an anechoic chamber and listened to the two-component stimuli through insert headphones (as described in Section 4.3). At each presentation of the stimulus, the subject's task was to adjust the frequency of the oscillator generating the secondary tone in relation to a fixed frequency primary tone according to the criterion of the minimum frequency difference in which the two components could be heard "clearly and distinctly". All judgments were made from an initial condition of zero frequency difference, and the secondary oscillator could only be adjusted to be the lower frequency of the two tones.

The experiment is shown in block diagram form in Fig. 7.2. An amplitude stabilised Wein bridge oscillator was used to

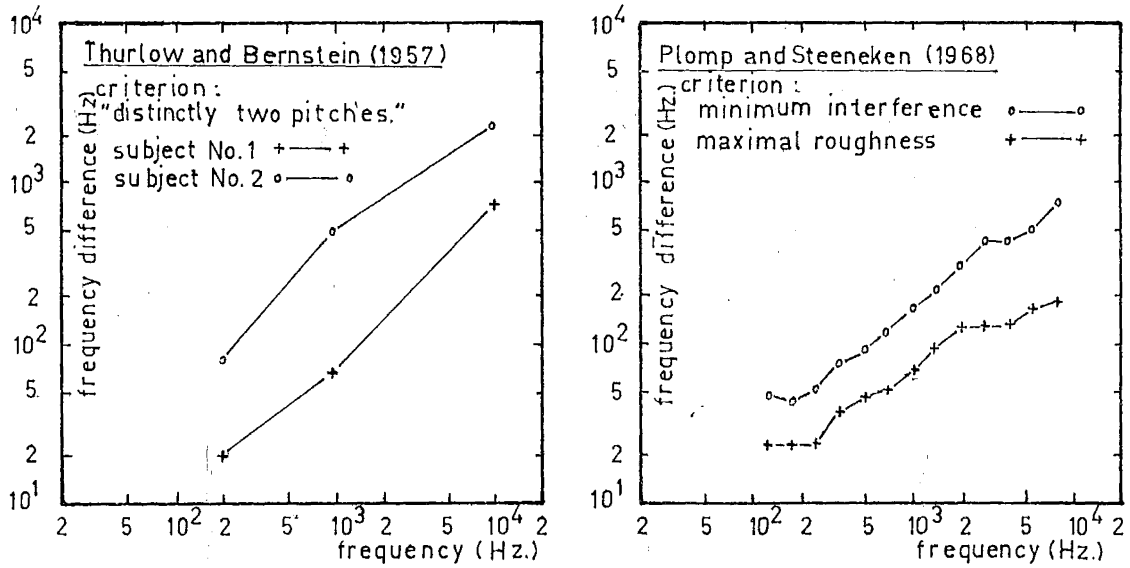


Fig. 7.1 The results of other workers estimating the minimum frequency difference between two tones according to the indicated criteria.

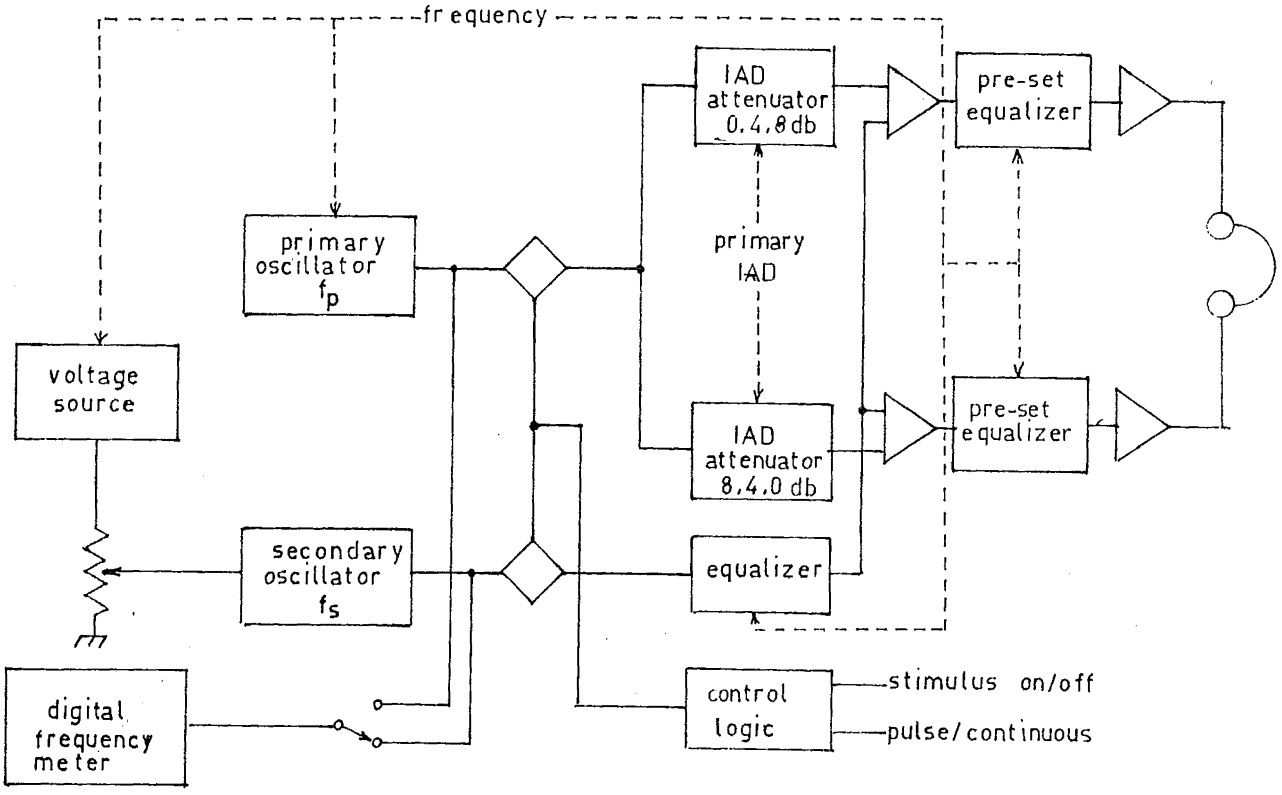


Fig. 7.2 Experimental apparatus used in estimating the frequency resolution of the auditory system.

generate the primary tone. Push-button selection of the frequencies was incorporated for repeatability. The secondary oscillator was a Hewlett-Packard 3300A function generator used in the voltage controlled mode. The subject's control was an uncalibrated potentiometer that gave a linear variation of the secondary frequency with rotation from a minimum of 50 Hz to the primary frequency. The stimulus envelopes were controlled by FET gates with a rise/fall time of 5 msecs to minimise transient clicks. Pre-set resistive equalizers were included to maintain a constant sensation level at all of the primary frequencies. After each response the frequency difference was measured with a digital frequency meter.

The stimulus parameters used were as follows:

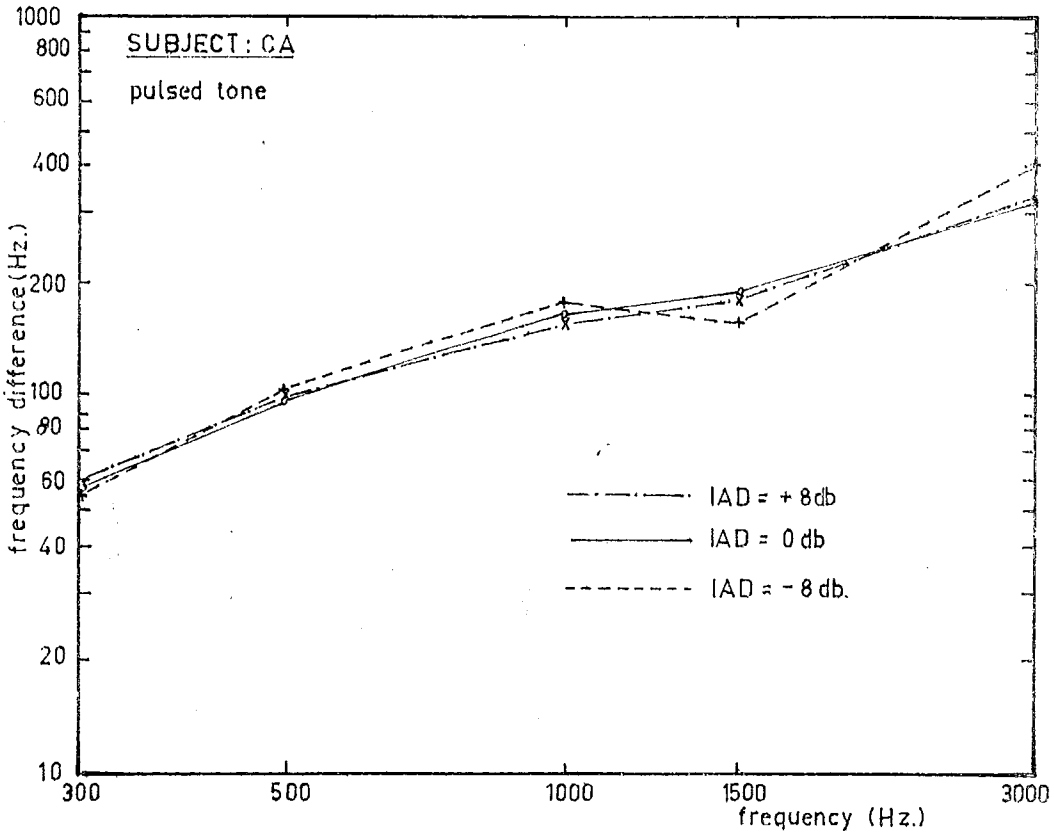
Primary frequency:	300, 500, 1000, 1500, 3000 Hz
Primary tone IAD:	+8, 0, -8 db (re the left ear)
Envelope:	Continuous tone, and pulsed tone (repetition period 330 msecs, duration 300 msecs).
Sensation level:	60 db for each component tone at the primary frequencies, as measured in a 2 cc coupler (ASA) with the insert tubes attached.

The envelope parameters of the pulsed tone condition were chosen to simulate the envelope structure of the binaural mobility aid.

The experiment was designed as a three way factorial (Primary frequency x IAD x Envelope structure), with the stimulus presentation order randomised between and within repetitions of the complete block. The mean value of the difference frequency was computed from ten observations in each cell.

Six male undergraduate students were used as subjects. None of them had previously participated in psychophysical experiments, and all were given audiometric threshold tests to ensure the absence of threshold abnormalities. The subjects reported daily through the week for a period of three weeks; each daily session lasted for ½ hour. The first four sessions for each subject were spent in familiarisation with the task and during this period results were not recorded. No corrective feedback of any kind was given to subjects throughout the experiment, but the criterion for adjusting the frequency difference was repeated verbally to the subject before the start of each session.

Results: A lack of homoscedasticity (equality of variance) over the complete block precluded rigorous statistical analysis using the analysis of variance technique for the factorial design. However, the significance of the three main effects was obvious. The results of a single subject are shown in Fig. 7.3, all subjects gave results similar in form. It was obvious that the IAD factor was not affecting



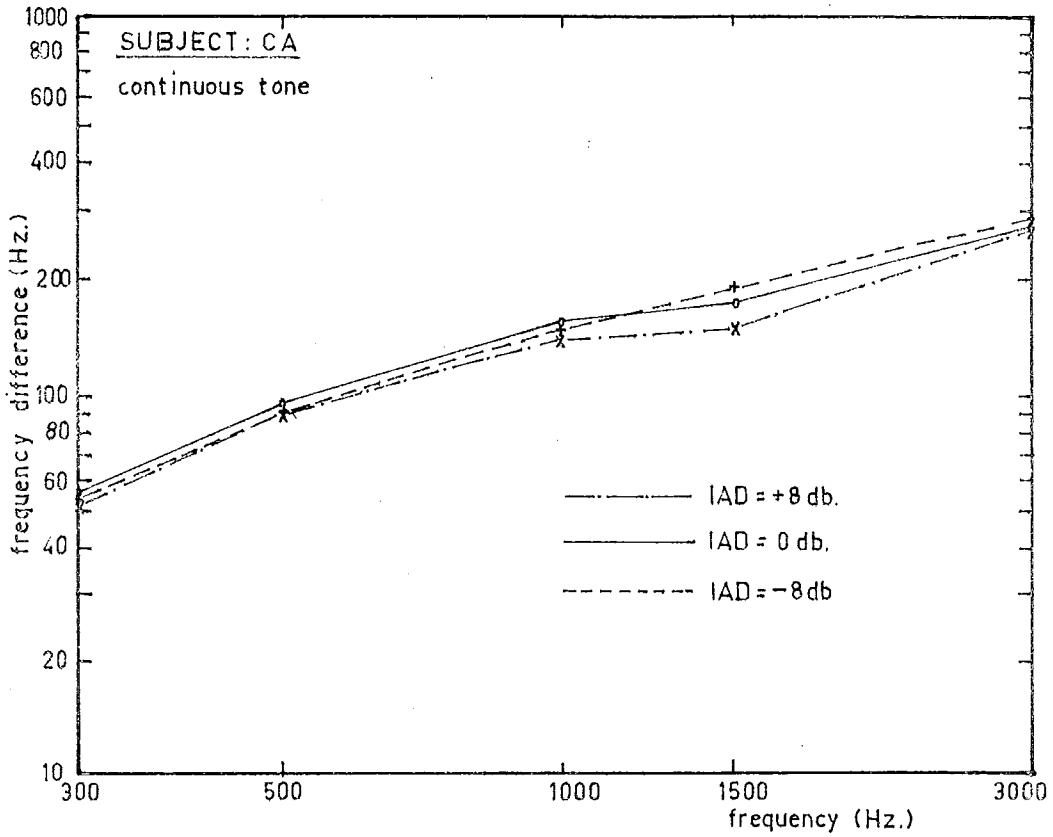


Fig. 7.3 Results of frequency resolution experiment for subject CA.

the results of any subject; the data from the three IAD levels were therefore pooled to obtain an estimate of the difference frequency based upon 30 observations. The frequency dependence of the results of all subjects is shown in Fig. 7.4 for both the pulsed and continuous tone conditions.

Discussion: The large differences in the results of individual subjects reported in the earlier studies were verified in this experiment. A subject's performance in a task of this nature will depend upon the interpretation placed upon the judgment criterion. The vague criterion of "clear and distinct" component images was purposely used for this experiment so that subjects would define their own resolution interval. The magnitude of the results of subjects DE, CG, LG and JR is clearly much smaller than the accepted values of the critical bandwidth and when compared with the results of the experiment by Plomp and Steeneken (Fig. 7.1) indicate the effect of the criterion given to the subjects upon the response in an experiment of this type.

The frequency dependence of the results of all subjects (except JR for continuous tones) showed a monotonic increase in the resolution interval Δf with the median frequency. The reason for the anomolous results of JR is not known. For all other subjects the results may be expressed as a psychophysical power law, i.e.

$$\Delta f = q \left(\frac{f_p + f_s}{2} \right)^n \qquad 7.1$$

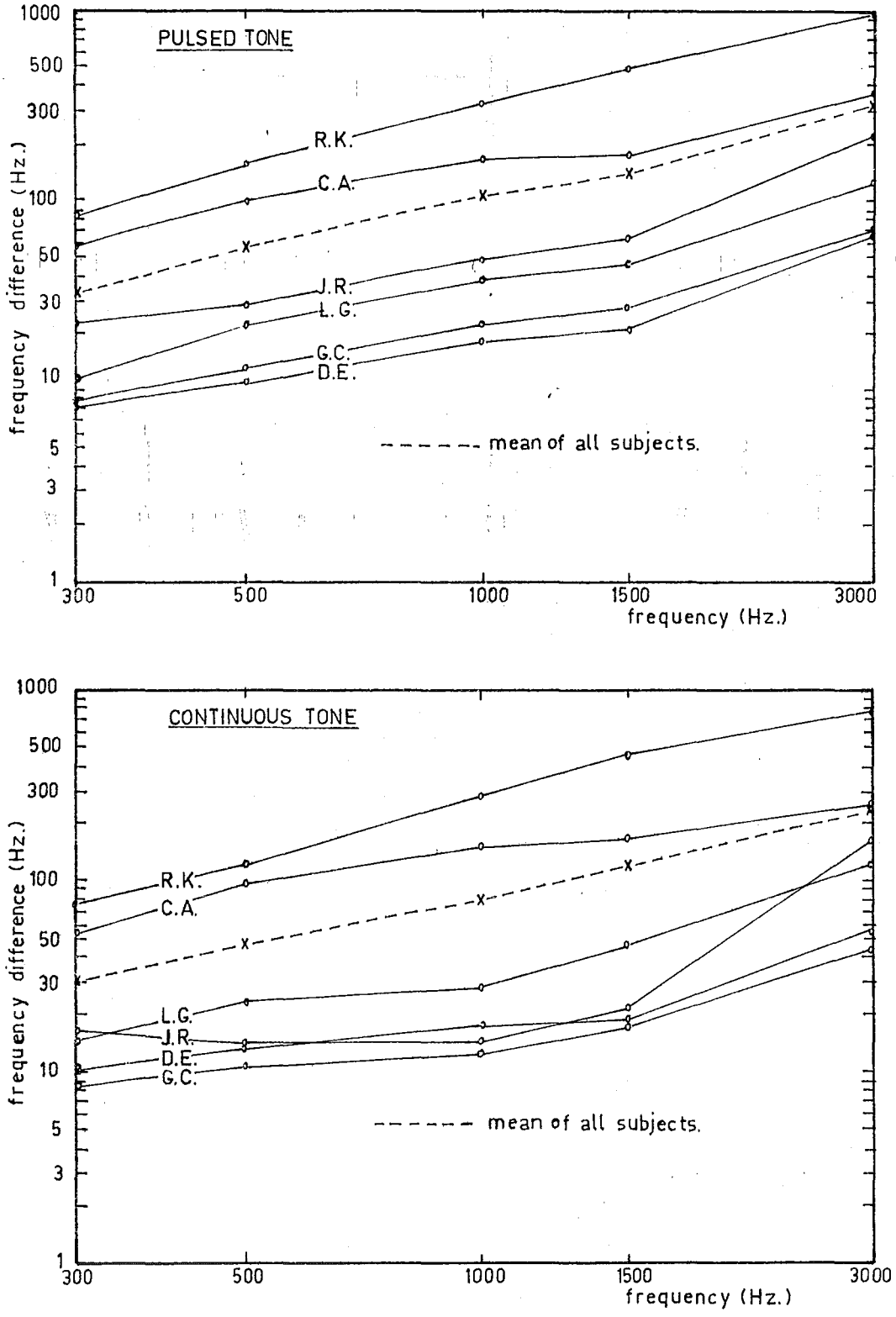


Fig. 7.4 Frequency dependence of the minimum frequency difference between two tones for perception of both, for all subjects.

where Δf is the frequency difference, f_p and f_s are the primary and secondary frequencies, q is a constant of proportionality (varying between 0.04 and 0.4 in the present experiment) and the exponent n is near to unity.

The fact that IAD in the primary tone did not appear to affect the resolution interval supports a hypothesis of cochlear interference between the components. This result means that the range resolution in a two-dimensional display (frequency x IAD) will not be affected by angular separation of two objects in the environment.

All six subjects reported, at times, the presence of interaction tones as they adjusted the frequency of the secondary oscillator. The most common report was of the $(2f_s - f_p)$ interaction tone when the primary frequency was 3000 Hz. This particular combination tone has been extensively studied with the conclusion that the site of the generating non-linearities lies within the cochlea[145,146]. The spurious components may be reduced by decreasing the stimulus amplitudes. Thurlow and Bernstein reported that interaction between the two component tones was minimised at a sensation level of 10 db. However, an auditory display, particularly for a mobility aid, will normally be required to operate against an ambient background noise level that would mask tonal signals at this low intensity. In these situations the ramifications of interaction tones in a multi-component frequency domain range display must be accepted.

7.3 The Ability of the Auditory System to Resolve the Lateralization of a Tone in the Presence of an Interfering Tone

The experiment described in this section was designed to estimate the resolution capabilities of the auditory system in a task requiring both recognition and identification of a primary tone in the presence of a secondary tone. The ability of subjects to determine the lateralization of the primary tone, offset from the median plane by IAD, was measured as a function of the frequency of the interfering secondary tone.

The effect of the secondary tone was estimated by measuring the degradation in the difference limen (DL) for IAD as a function of the frequency difference between the two components. The definition of the DL used for this experiment was similar to that used by Mills[48] in his studies of dichotic localization phenomena. If, in a two-alternative forced choice experiment, a plot is made of the fraction of responses of one type against the stimulus magnitude, the resulting psychometric function[147] may be used to define the DL as follows. If the stimulus level at which the probability of a response of the given type is x is denoted as L_x, the DL may be defined as

$$DL = \tfrac{1}{2}(L_{.75} - L_{.25}) \qquad 7.2$$

as in Fig. 7.5. For experimental expediency it has been

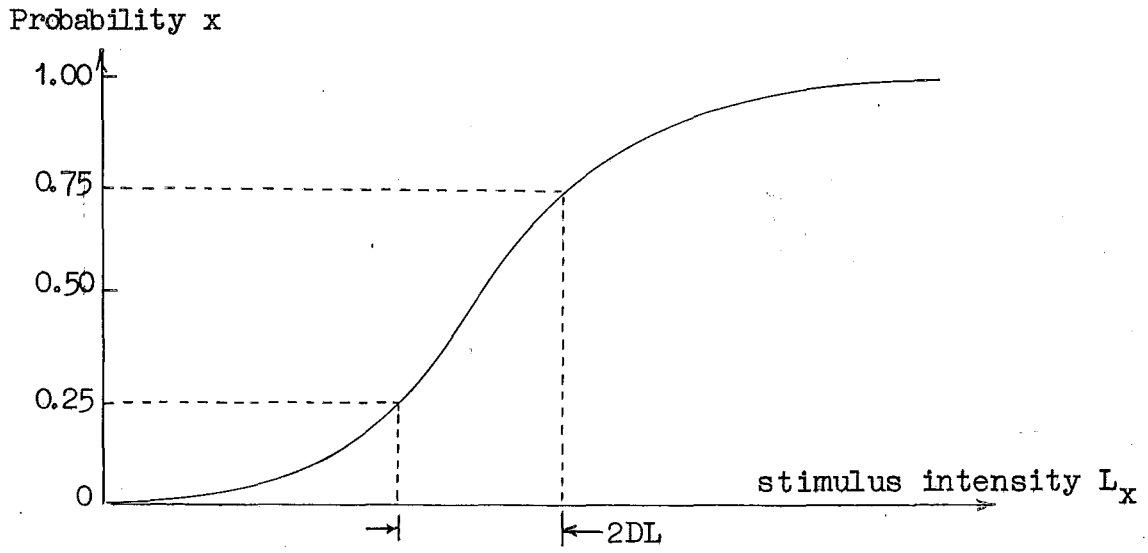


Fig. 7.5 Definition of the Difference Limen from the Psychometric Function

convenient to define a modified difference limen DL' as

$$DL' = \tfrac{1}{2}(L_{.707} - L_{.293}) \qquad 7.3$$

for reasons given later. The relationship to the more usual definition of DL (Eq. 7.2) may be easily established by assuming some parametric form for the psychometric function; it is commonly assumed to be either cumulative normal, logistic, or linear. For example, for a cumulative normal distribution DL = 1.09DL'.

In defining a DL for a dichotic difference the fraction x may be defined from responses of lateralization "to the right". In estimating the effect of an interfering tone upon the DL a degradation factor DF may be defined as

$$DF = \frac{DL_{ps}}{DL_{p}} \qquad 7.4$$

where the subscripts p and s refer to the presence of the

the primary and secondary components respectively. If the psychometric functions defining DL_p and DL_{ps} have the same parametric form then

$$DF = \frac{DL'_{ps}}{DL'_{p}} \qquad 7.5$$

The Psychophysical Method of Estimating the Difference Limen: The classical psychophysical techniques (method of constant stimuli, method of limits etc.) have been widely used in the past but suffer from many disadvantages, the most important of which is a lack of experimental efficiency. Modern statistical techniques have been developed to obtain greater efficiency by concentrating the stimulus level about the point of interest on the psychometric function.

If the parametric form of the psychometric function is known, which will not generally be the case, maximum likelihood methods[148] may be used to obtain efficient estimates of any L_x. However, non-parametric estimators are to be preferred, especially if there is little a-priori knowledge on the form of the psychometric function. Sequential and adaptive strategies that are not critically dependent upon the parametric form of the psychometric function have been adapted for use in psychoacoustics to give efficient estimates of L_x with small asymptotic bias. PEST (Parameter Estimation by Sequential Testing)[149,150] which uses Wald's sequential stopping rule[151] to terminate the testing sequence, and the

up-down method[152] together with its variants the UDTR (Up-Down Transformed Response)[153,154] and BUDTIF (Block Up-Down, Two Interval Forced Choice)[155,156] have been used in various psychoacoustic contexts.

The experiment described here was designed around an on-line hybrid computer implementation of the UDTR technique, which combines experimental efficiency with simplicity of the strategy and small asymptotic bias. The technique is described in detail by Wetherill[157] and is summarised here.

The observations in the two-alternative experiment are taken at a set of equally spaced stimulus levels S_k. Then for certain values of x the UDTR method tracks L_x as follows. Take a finite series of observations at the stimulus level that is the best initial estimate of L_x, and concurrently calculate the fraction x' of positive responses. As soon as $x' < x$ terminate the series and increment the stimulus level. If after the predetermined length of the series, n, $x' \geq x$ then decrement the stimulus level before starting the next series. For example, if the response pattern required to decrement the stimulus level is n positive responses (i.e. any negative response will make $x' < x$ and terminate the series) the level L_x tracked by the strategy will be given by

$$x^n = 0.5. \qquad 7.6$$

The original up-down rule[152] set n = 1 and therefore tracked the level $L_{.5}$. The effect of the UDTR strategy is to transform the response pattern, and therefore to track the

transformed response at x = 0.5. Some typical UDTR strategies based upon the response pattern of n consecutive positive responses are shown in Table 7.1. Note that L_x for $x < 0.5$ may be tracked by decrementing the stimulus level after n negative responses in which case the level will be given by

$$(1-x)^n = 0.5 \qquad 7.7$$

TABLE 7.1 SOME POSSIBLE UDTR STRATEGIES

n	Response Pattern to Decrement s_k	Response Pattern to Increment s_k	Level tracked by Strategy
1	1	0	0.5000
2	11	0,10	0.7071
3	111	0,10,110	0.7940
4	1111	0,10,110,1110	0.8409
5	11111	0,10,110,1110,11110	0.8705

Two additions to the basic UDTR strategy were used in the experiment. Consider the hypothetical experimental run with $n = \frac{2}{3}$ as shown in Fig. 7.6. Wetherell proposed an estimate $\overline{W}$ of L_x based only on those series in which there is a change of stimulus direction (series ~~1, 3, 5, 6, 8~~ 2, 4, 6, 7, 8, 10 in Fig. 7.6). Then $\overline{W}$ is defined as

$$\overline{W} = \frac{1}{m} \sum_{k=1}^{m} W_k \qquad 7.8$$

where the experiment is terminated after m changes of response

type and where W_k are the stimulus levels at which a change of response type occurred. Although the $\overline{W}$ estimate has a small asymptotic bias caused by the lack of symmetry in the

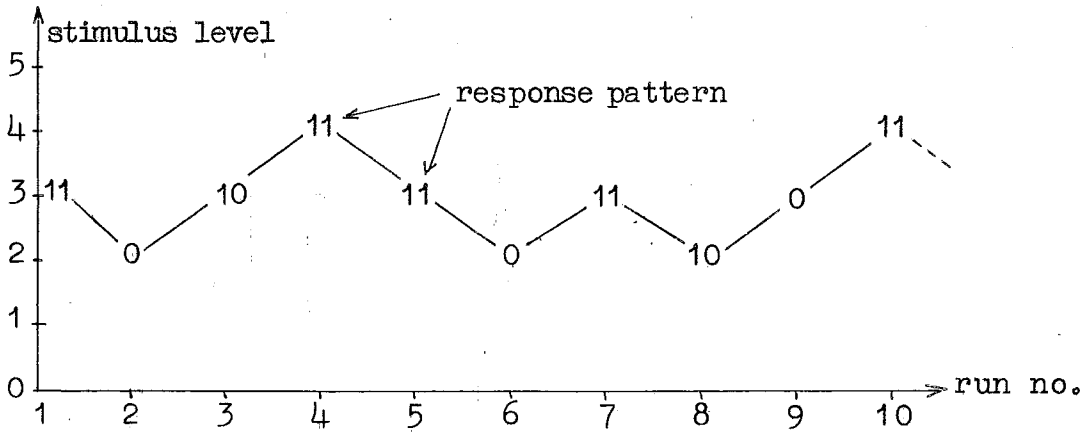


Fig. 7.6 A Hypothetical UDTR Experiment

transformed response curve, this may be reduced by an appropriate choice of stimulus increment magnitude. Further possible bias in the estimate caused by an incorrect initial estimate of L_x was largely eliminated by following Wetherell's Routine 17 which states that the first m_1 changes of response type should be used to calculate a $\overline{W}$ statistic, which is then used as an initial estimate to restart the sequence with one half of the original stimulus increment magnitude.

The dichotic DL in IAD was estimated by using the strategy with n = 2 in Table 7.1 to estimate $L_{.707}$ and its inverse (response pattern 00 to decrement the stimulus magnitude) to estimate $L_{.293}$. These two strategies were interleaved with equal probability of either at each trial to reduce sequential dependencies[158]. The initial estimate

of the DL was set at 2 db with the stimulus increment set at 0.4 db. Wetherill's Routine 17 was used with a reinitialization after six changes of response type, followed by the experimental run based upon 16 changes of response type with the increment in IAD set at 0.2 db.

On-line Computer Implementation of UDTR: The repetitive nature and logical (even if randomised) stimulus sequence structure of the psychophysical experiment make it ideally suited to control by digital computer. Only the simplest of the modern sequential or adaptive statistical techniques can be used in a manually controlled experiment, so that if full advantage is to be taken of the efficiency of these methods some programmed control of the experiment is necessary.

Previous reports on the computer control of psychophysical procedures have shown that the major problem is in the interfacing of the subject to the computer and the control of the stimulus. Three basic techniques have been used:

i. To have the computer generate the stimuli in real time and to apply these to the subject's headphones through digital to analogue (D - A) converters[159].

ii. To have the computer control a magnetic tape transport mechanism with previously recorded stimuli[160].

iii. To use the computer's internal digital pulses applied directly to the headphones as "click" type stimuli[161].

All of these techniques present difficulties in a general psychoacoustic experiment because they either require a

specialised computer interface or else lack flexibility.

This experiment used a hybrid computing system to take advantage of the stimulus-response processing capabilities of both analogue and digital computers. The experimental strategy and data analysis was controlled by the digital computer with the stimulus control and timing executed through the elements of the analogue computer. The only external equipment needed was an oscillator, a pair of headphones and three push-button switches.

The experimental set-up was designed around the hybrid computing system in the Electrical Engineering Dept., University of Canterbury. At the time of the experiment (1969) this comprised an EAI 640 digital computer with 8192 words of memory, an EAI 580 analogue computer with 60 amplifiers and 50 servo-set potentiometers and a basic EAI 693 hybrid linkage unit containing a single D-A convertor (the RDAC) and A-D data transfer through the digital voltmeter (DVM).

The analogue computer patching is shown in Fig. 7.7. (The analogue computing symbols used are defined in Appendix 5). A separate phase of the programme, not shown in the figure, was used to set the levels of the oscillators generating the primary and secondary tones and to equalise the headphone responses. Because of the limited data transfer capability of the system it was necessary to set the amplitudes of the dichotic stimuli serially on to the RDAC and to multiplex the output on to four analogue storage units

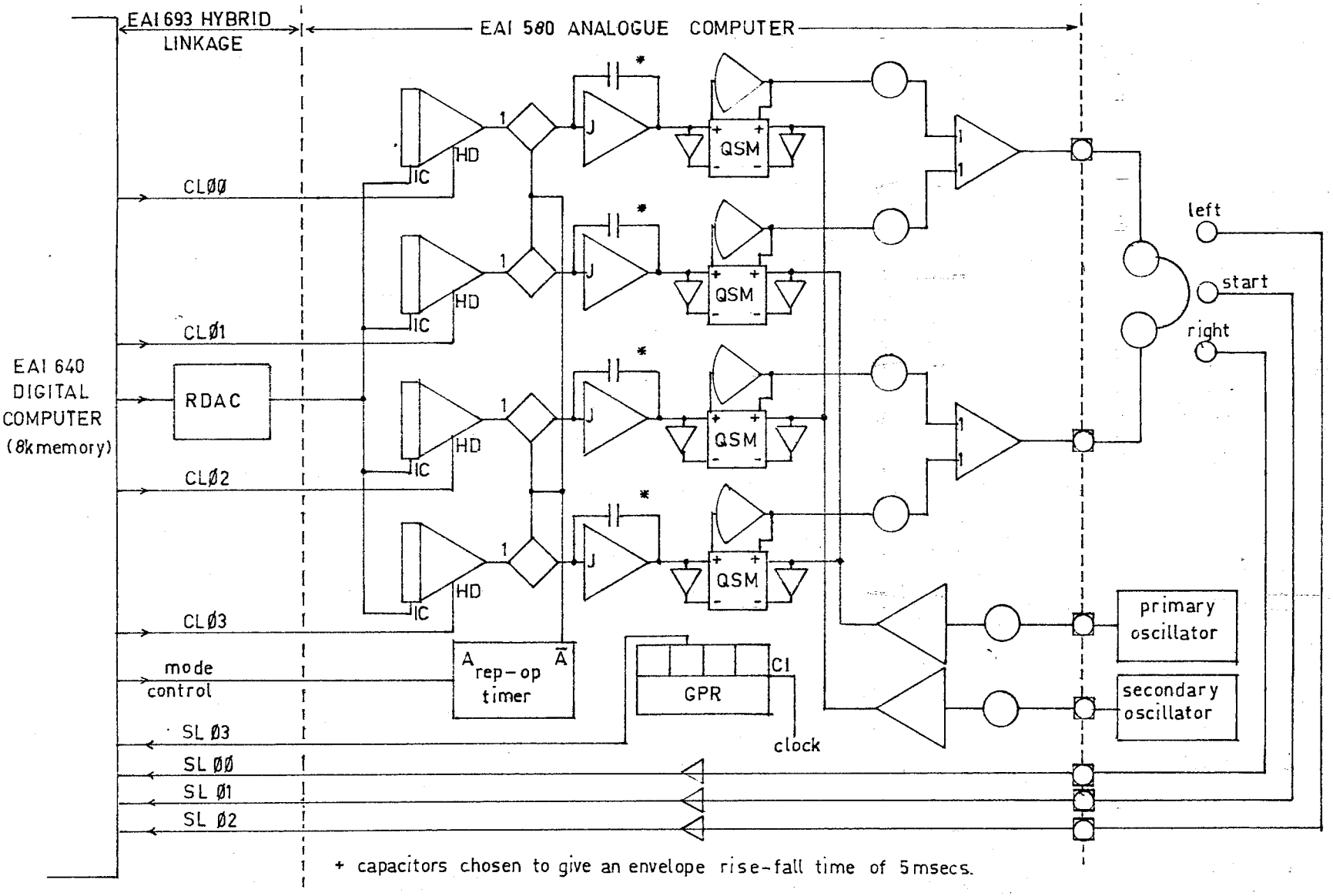


Fig. 7.7 Analogue computer patching for on-line hybrid estimation of the DL in IAD. (see Appendix 5 for analogue computer symbols).

by using the control lines (CLØØ-CLØ3). The envelope parameters (duration and repetition period) were controlled by the repetitive operation (rep-op) timer. Digitally controlled analogue (D/A) switches with a rise/fall time of 5 msecs were used to gate the stimulus magnitude on to high accuracy (0.01%) quarter-square multipliers (QSMs) to control the amplitude of the externally generated tones. Normal inverting amplifiers were used to sum the primary and secondary components and to drive the headphones.

The subject communicated with the digitally controlled strategy through the sense lines; one was used to initiate the session, two more were used as response lines. The fourth sense line was connected to an astable multivibrator; the state of this sense line was interrogated after each response to determine whether the next stimulus presentation should be for the $L_{.707}$ or $L_{.293}$ strategy. Because the response time is random the two strategies were interleaved with equal probabilities at each trial.

The digital computer programme was written in a version of the EAI HOI (Hytran Operations Interpreter) language that was modified by the author to improve the flexibility of communication with the EAI 580 analogue computer. This language is a high level interpreter type of language designed specifically for EAI hybrid computing systems. The flow diagram for the experiment is shown in Fig. 7.8. Only one strategy ($L_{.707}$) is shown, the other was identical in form.

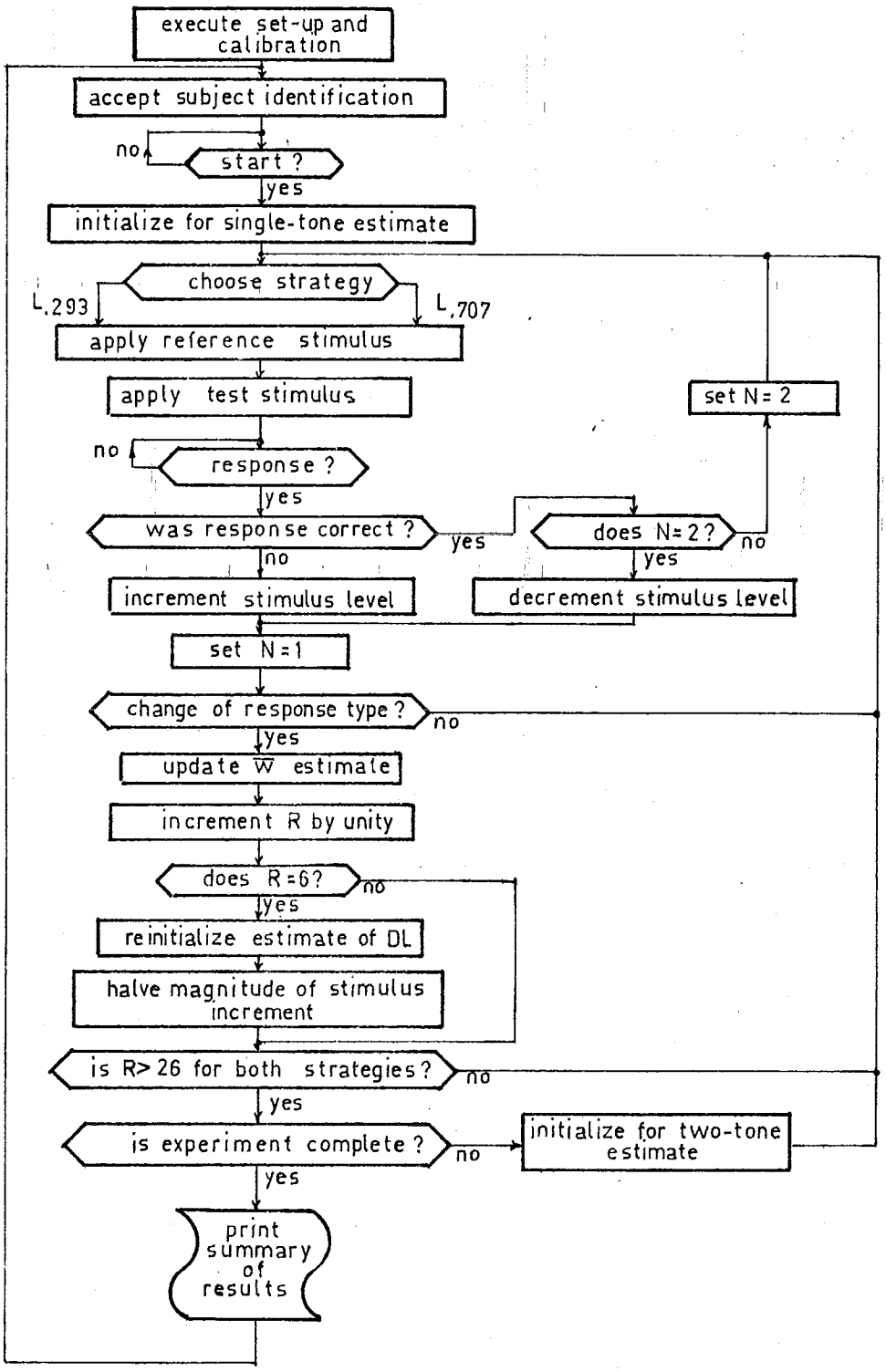
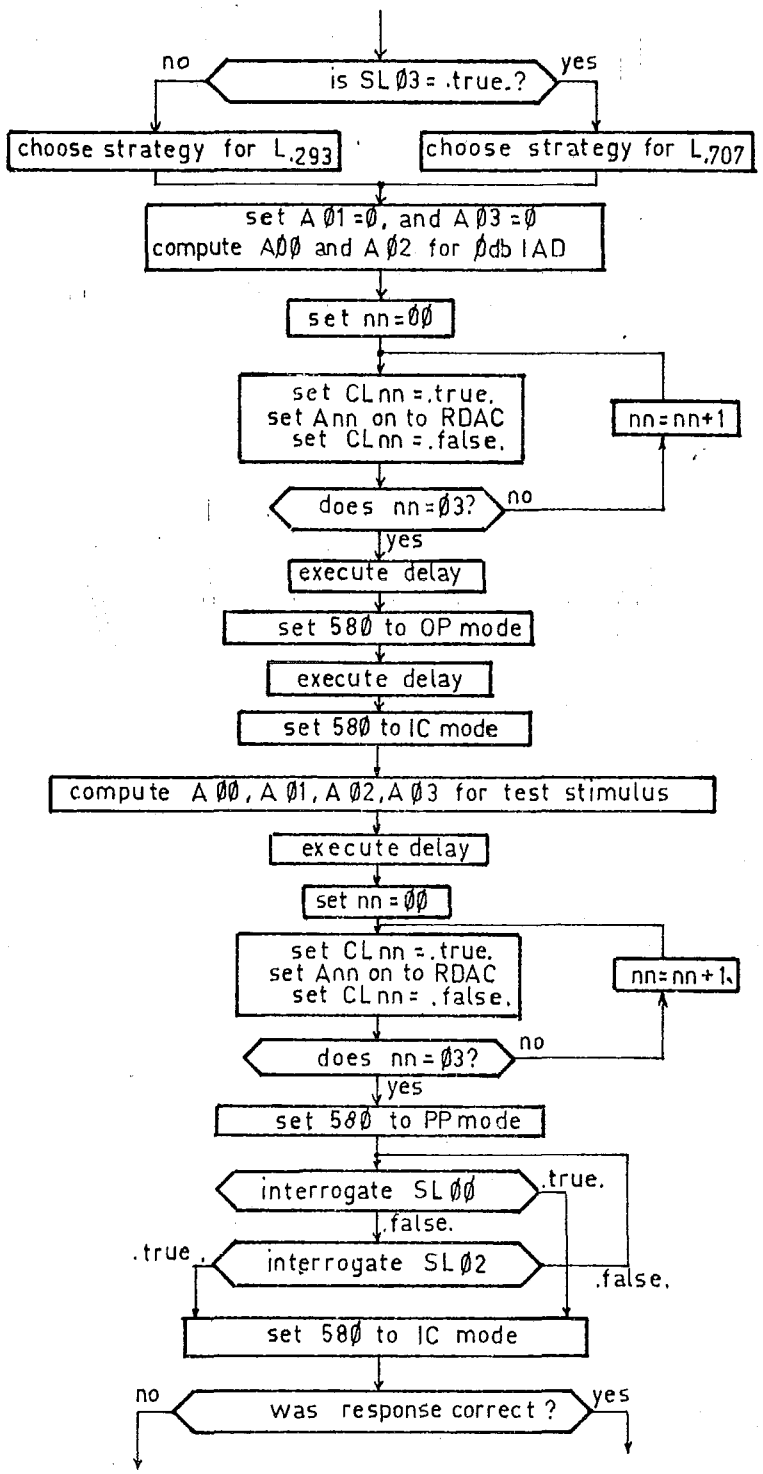


Fig. 7.8 Flow diagram for on-line UDTR estimation of the degradation in the DL. The hybrid operation is shown on the next page.



examine the state of a register to choose between strategies

set stimulus component amplitudes on to T/S amplifiers.

apply reference stimulus

set stimulus component amplitudes on to T/S amplifiers for test stimulus.

apply test stimulus

await response

NOTE : AØØ and AØ2 are the amplitudes of the dichotic primary components.
AØ1 and AØ3 are the amplitudes of the dichotic secondary components.

Fig. 7.8 contd. Hybrid computer control of the UDTR experiment. This page shows the stimulus application and response testing stages shown on the previous page.

Appendix 6 describes a method of controlling a simple UDTR experiment in IAD with an analogue computer alone. Although this method lacks the flexibility of the full hybrid control described above, it could be used to advantage in those laboratories where only an analogue computer is available.

Experimental Method: In a pilot experiment, using the method of constant stimuli, it was found to be necessary to balance the IAD between the primary and secondary components, when estimating DL_{ps}, to prevent lateralization responses based upon the position of a nett image without the resolution and identification of the primary tone. The secondary component of the stimulus was therefore presented with IAD equal in magnitude but opposing that of the primary tone.

Each experimental session was divided into two parts; estimation of DL_p then estimation of DL_{ps}, separated by a break of approximately three minutes during which time a summary of the strategies for estimating DL_p were printed out. For all sessions the primary frequency was 1500 Hz. The secondary frequency was randomised between sessions between 300, 400, 500, 600, 800, 1000, 1200, 2400 and 4000 Hz. At least two sessions were spent in estimating DL_{ps} for each secondary frequency.

The stimulus timing is shown diagrammatically in Fig. 7.9. A one second duration reference stimulus consisting of

the primary component alone with 0 db IAD was applied before each test stimulus. This was done for two reasons:

i To establish the median plane as the point of reference for lateralization judgments and therefore to reduce any sequential dependence of responses.

ii To aid in the identification of the primary component in the estimation of DL'_{ps}.

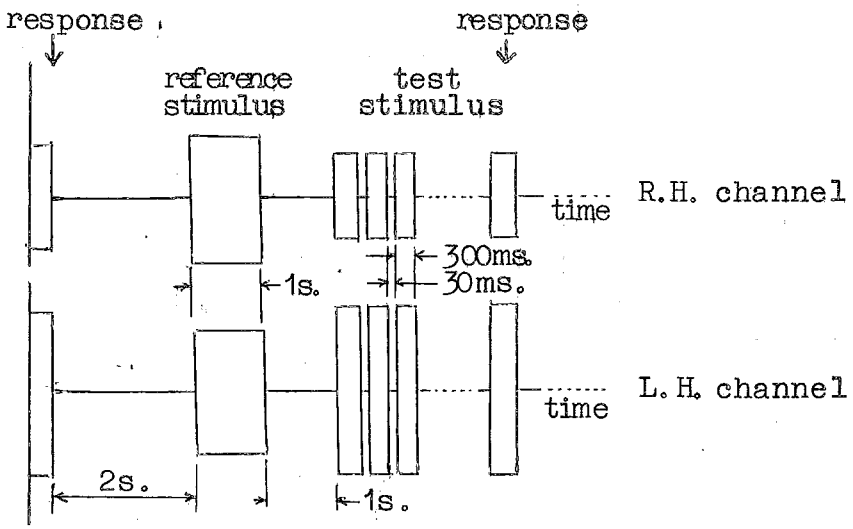


Fig. 7.9 Stimulus Timing Diagram

The subjects were instructed to listen to the reference stimulus and to decide whether the primary component of the test stimulus was to the left or right of the reference. The subject's console consisted of three push buttons, one to initiate the experiment, the other two to indicate the lateralization of the test stimuli. The experiment was conducted in a small room adjacent to the computer laboratory.

Six subjects reported daily during the week for ½ hour sessions, over a period of five weeks. All were male students in engineering and all had previously participated in experiments in auditory localization conducted by the author. Audiometric tests were given to all subjects to ensure the absence of threshold abnormalities. The first four sessions were spent in familiarisation with the task; the results from these sessions were discarded.

Results: All subjects were able to lateralize the primary tone and gave consistent estimates of DL_p throughout the experiment as summarised in Table 7.2. Each entry is based upon 20 estimates of DL_p from the experimental sessions. The estimate of DL_p was derived from DL'_p as measured by the UDTR strategy by assuming a cumulative normal psychometric function ($DL_p = 1.09\ DL'_p$). The overall between-subjects mean (0.78 db) compares favourably with the DL estimated by Mills[48] using the method of constant stimuli (0.70 db) at a sensation level of 50 db. It was not possible to estimate the significance of the between-subjects differences in a two way classification (subjects x days) because there is no known way of determining the variance of an estimate of L_x from the UDTR method. The between-subject differences were tested in a one-way analysis of variance ($F = 1.49$, $df = 5, 19$; $p > 0.1$); there was no evidence rejecting a hypothesis of equality of the means.

TABLE 7.2 ESTIMATES OF SINGLE TONE DL IN IAD

f_p = 1500 Hz

Subject	Mean DL_p	Standard deviation σ_{DL_p}
CA	0.84	0.465
PMC	0.63	0.270
LG	0.65	0.314
JDH	1.04	0.314
ARJ	0.97	0.320
DR	0.60	0.180
Between Subjects	0.78	0.186

The presence of the secondary tone in the estimation of DL'_{ps} generally gave elevated thresholds. However, for three subjects the estimated DL'_{ps} was depressed relative to DL'_p when f_p and f_s were widely separated (f_s = 300, 400 Hz); which indicated that these subjects were able to use the extra information in the balanced nature of the stimuli, when both tones were present, to increase the slope of the psychometric function. Subjects were not informed that the primary and secondary components contained opposing IAD.

At small frequency separations (f_s between 1200 Hz and 4000 Hz) subjects failed to give convergent estimates of DL'_{ps}. Under these stimulus conditions it was not uncommon

for the UDTR strategy to increase the IAD level to the dynamic limit of the analogue computer (> 20 db). One subject in particular (PMC) was unable to lateralize the primary tone in the presence of the secondary tone except under the conditions of widest separation (f_s = 300, 400 Hz). All other estimates of DL'_{ps} were non-convergent. No improvement was noticed in this subject throughout the course of the experiment.

The results of the experiment are summarised in the degradation factors (Eq. 7.5) plotted in Figs 7.10 and 7.11. The first figure shows the overall mean degradation factor in IAD calculated from the results of all subjects except PMC. Fig. 7.11 summarises the results of three individual subjects. In this case DF was calculated from the mean DL'_p for the subject (Table 7.2) and the mean DL'_{ps} from the two experimental sessions at each value of f_s.

Discussion: The results plotted in Figs 7.10 and 7.11 show that the interference between the lateralization (and hence localization) of two tones extends over a much wider frequency difference than would be predicted from the critical bandwidth of the auditory system (Fig. 7.1). An attempt was made to correlate the degradation factor with the cochlear interference between the two tones, using Flannagan's model of the response of the basilar membrane[73], but it was found that the degradation in the DL predicted from either energy

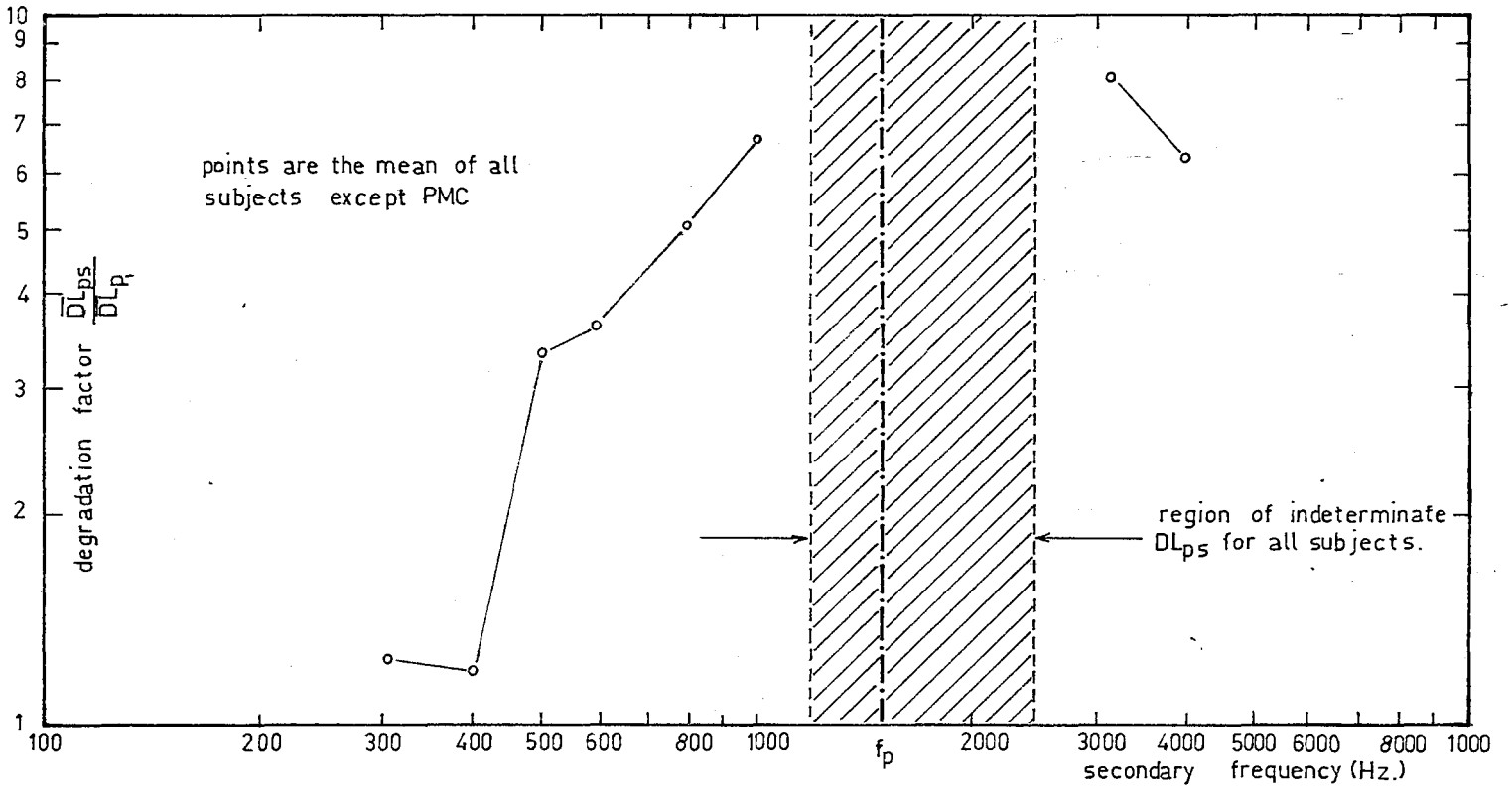


Fig. 7.10 Mean degradation in the DL caused by the presence of the secondary tones for all subjects except PMC.

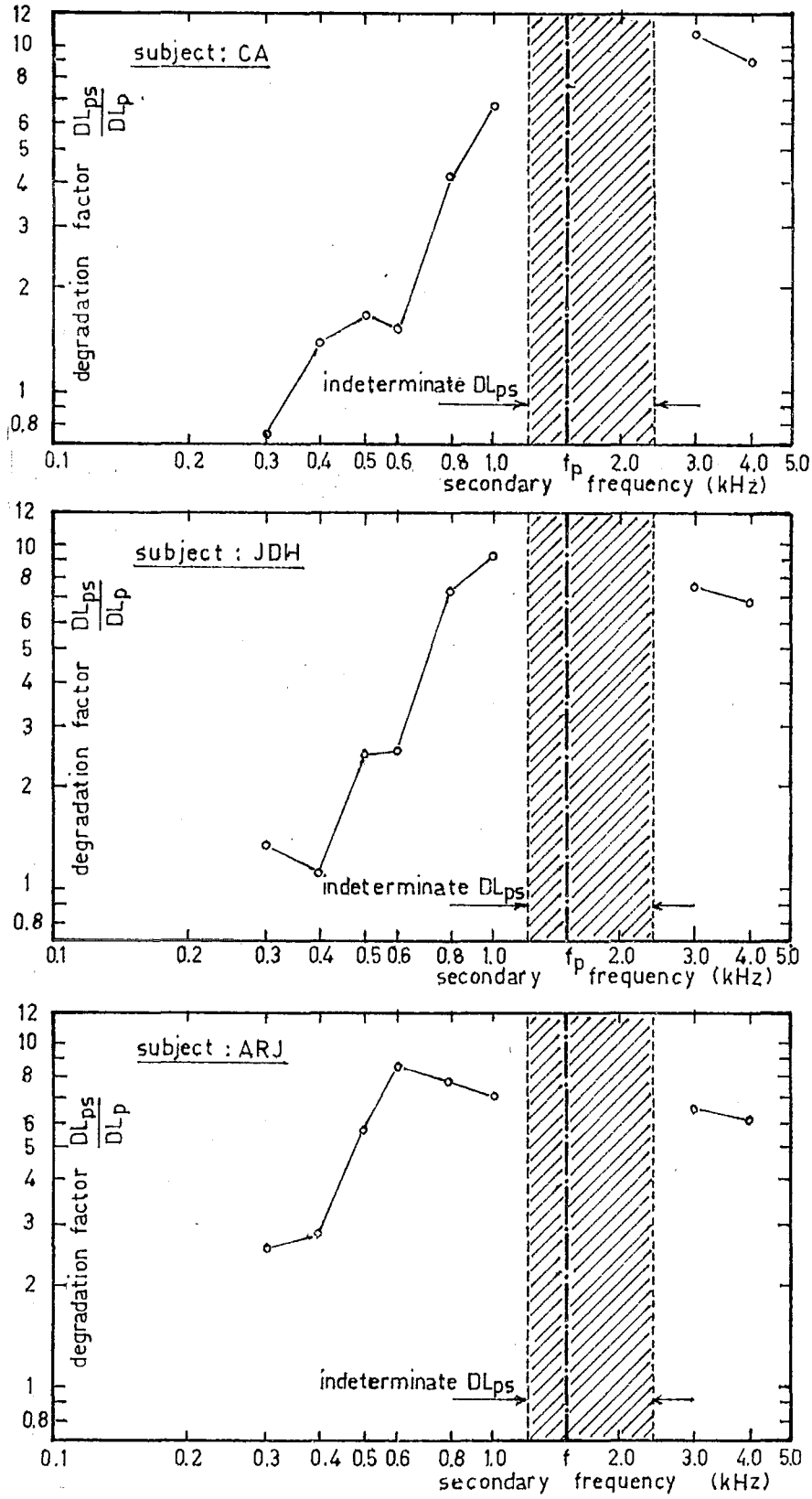


Fig. 7.11 Degradation factors for three subjects.

or peak amplitude comparison was very much less than that observed in the results of the experiment. There is therefore evidence of central interaction between the neural patterns generated by simultaneous tones of different frequency. The region of indeterminate DL'_{ps} (1200 - 2400 Hz) was greater than that indicated by the critical bandwidth or values of Δf measured in the previous experiment.

The conclusion drawn was that perception of both components of a two-tone stimulus is not a sufficient condition for lateral (or azimuthal) resolution. This was confirmed by the subjects who reported that although they were able to hear both components they had difficulty in determining lateralization. It is therefore not possible to define a simple Δf x IAD resolution cell for the dichotic display.

Both of the experiments reported in this chapter have been concerned only with a static dichotic display. The effect of the secondary cues of head-movement and kinaesthesis upon resolution in a dynamic environment have yet to be investigated. The display simulation techniques described in the next chapter provide a foundation for such a study.

C H A P T E R 8

STUDIES IN DYNAMIC AUDITORY LOCALIZATION

8.1 Introduction

The azimuthal variation in amplitude in the dichotic channels of the display developed in Chapter 5, was based upon the parametric form of the localization function empirically derived from a static experiment using constant dichotic stimuli (Chapter 4). The adequacy of any such binaural display in a real environment cannot be predicted directly from this classical psychoacoustic situation because secondary cues afforded by kinaesthesis, memory, head movement and intermodal association are eliminated by the experimental conditions. In this chapter localization experiments with free head movement are described, using an analogue computer simulation of the idealised display in a single object environment. The experiments were designed to investigate the use of adaptive strategies to obtain rapid estimates of the localization function. The object of the experiments was to define a possible 'clinical' testing procedure to be used in the matching of the display of the mobility aid to each user.

The simulator was developed with the broader aim of providing a research tool for the study of requirements of a sensory display for human mobility. For this reason a

discussion of the role of a simulator in the psychophysics of mobility is presented below.

8.2 The Role of a Mobility Aid Simulation

A complete specification of any non-visual display to aid the blind must be related to normal human interaction with the environment and the enhancement of this interaction by the display. This area of research has been largely neglected.

The need for computer simulation of mobility aids as a research tool has been recognized[162] and investigated by Baecker.[163] This feasibility study investigated the possibility of the simulation of a pencil-beam mobility aid. It was concluded that the simulation would be extremely slow and would require specialised computer hardware. No consideration was given to the display or the sensory requirements of the man-machine interface.

The author proposes that a more profitable approach would be to study the mobility performance of subjects in a simple simulated environment. A rigorous study could then be made of the effects of the display parameters upon performance and the sensitivity of the mobility function to perturbations in the display parameters. The simulator would provide a well defined environment with strict control of the experimental stimuli and a provision for monitoring the subject's actions. The major problem associated with the use of an actual

mobility aid for this type of study is that the system and environment cannot be controlled or accurately specified (Chapter 6).

Consider, for example, two features of the display developed in this thesis that can only be resolved in a dynamic environment. The localization experiments described in Chapter 4 showed that individual subjects have a pronounced frequency dependence in the regression coefficient β of the localization curve. Although no way has been found of compensating for the resulting range dependence of the azimuthal dimension of the display, no subject (blind or sighted) using the mobility aid has reported a distorted auditory space. The sensitivity of mobility performance to induced distortions in IAD-azimuth angle relationship can only be studied in a simulated environment. Similarly, although it was shown in Chapter 7 that the ability to resolve the azimuthal direction of a dichotic tone was severely degraded by the presence of a secondary tone, it has been the experience of users of the mobility aid that resolution of the azimuth of two objects is possible. The factors affecting resolution with head movement could only be elucidated in a well controlled experimental environment - such as provided by a display simulator.

8.3 Analogue Computer Simulation of the Idealised Dichotic Display

Consider the geometry of a subject in a single object environment as in Fig. 8.1.

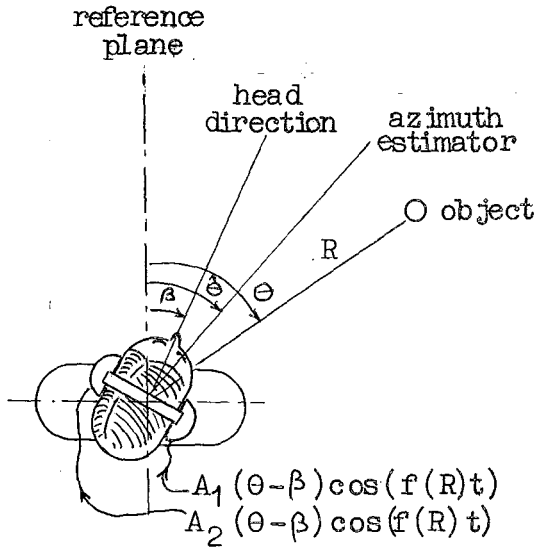


FIG 8.1 GEOMETRY OF THE SIMULATED ENVIRONMENT

Define all angles relative to the median plane through the body so that if the head is directed at an angle β the effective azimuth, relative to the auditory egocentre, of an object with azimuth θ is $\theta-\beta$. Then, as in Chapter 5, let θ' be the response to a stimulus with dichotic IAD: a suitable model is

$$\theta' = \beta + K \log_e \frac{A_1(\theta-\beta)}{A_2(\theta-\beta)} + \theta_0 + \xi \qquad 8.1$$

where K, $A_i(\theta)$ $(i = 1,2)$, θ_0 are defined in Chapter 5 and where ξ is a zero mean normal random process. Then define an estimator $\hat{\theta} = E\{\theta'\}$. If $A_i(\theta)$ is given by

$$A_i(\theta) = \exp\left(-\gamma[\theta + (-1)^i\alpha]^2\right) \qquad 8.2$$

where γ is a constant (Eq. 5.15) then

$$\hat{\theta} = \beta + 4K\gamma(\theta-\beta)\alpha + \theta_0 \qquad 8.3$$

Eq. 8.2 forms the basis of the display simulation shown

in Fig. 8.2, which was designed around the computing elements of an EAI 580 analogue computer. The quantities $(\theta-\beta\pm\alpha)$ were formed in summing amplifiers (see Appendix 5 for the notation of analogue computing components). A single time-shared diode function generator was used to generate the two amplitude weighting functions $A_i(\theta-\beta)$. To minimise set-up time a ten segment approximation to $F(\theta) = \exp(-\gamma\theta^2)$ for $\theta > 0$ was used, with an absolute value circuit to retain the modulus of θ when $\theta < 0$. The time sharing was accomplished through synchronised pairs of A/D switches and T/S amplifiers driven by a four bit ring-shift register. The T/S amplifiers were up-dated every four msecs. Alternative methods of generating the Gaussian weighting functions $A_i(\theta-\beta)$ using squaring circuits and preset exponential generators were tried but were found to have significant inherent errors.

The signal source was a Hewlett-Packard 3300A function generator used in the voltage controlled mode. No provision was made for simulating Doppler effects in the range coding (Chapter 6); this addition should be made for mobility studies. The range dependence of signal amplitude was included by using a non-linear network (R^{-2}) and multiplier. The azimuthal amplitude dependence of the dichotic stimuli was controlled by quarter-square multipliers (0.01% accuracy). The parallel logic was used to control the stimulus timing and envelope parameters. The pulse duration and repetition period were controlled from the rep-op timer through A/D switches which

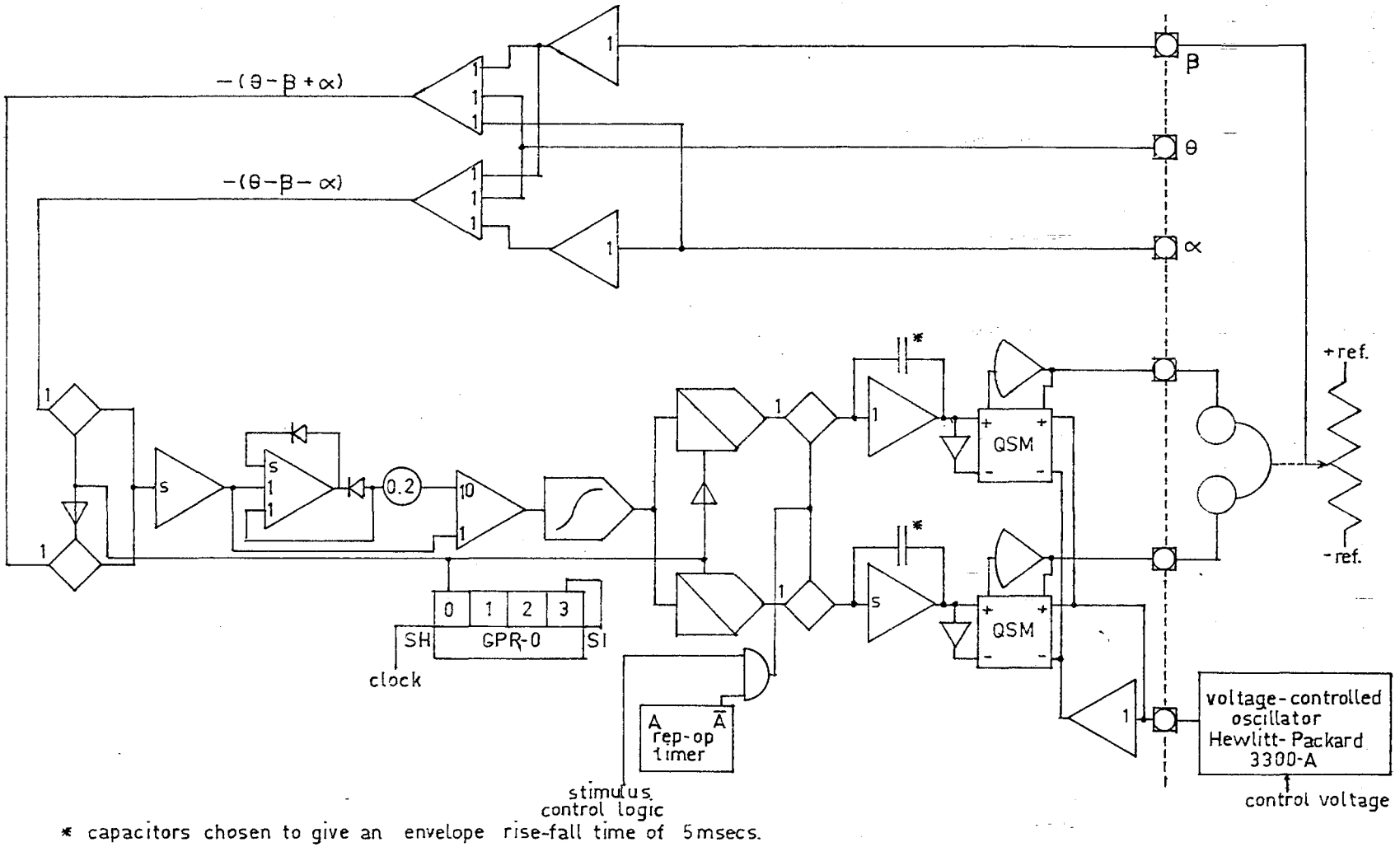


Fig. 8.2 Binaural display simulator, based upon the EAI 580 analogue computer, using a single time-shared diode function generator. (See Appendix 5 for analogue computer symbols).

gated the voltages $A_i(\theta-\beta)$ on to the multipliers through amplifiers with a first-order rise time of 5 msecs.

The expansion of the simulator to multiple object environments is limited by the number of available analogue multipliers. Each simulated object requires two multipliers as amplitude modulators for $A_i(\theta-\beta)$ and a further multiplier for the range dependence of the amplitude. The time shared diode function generator may be extended to at least a four object environment.

8.4 Interactive Methods of Estimating the Localization Constant K

The basic simulator described above was extended to include the subject's azimuthal response in a feedback loop containing an adaptive strategy in order to estimate the value of the localization constant K in conditions of free head movement.

Consider the man-machine system shown in Fig. 8.3.

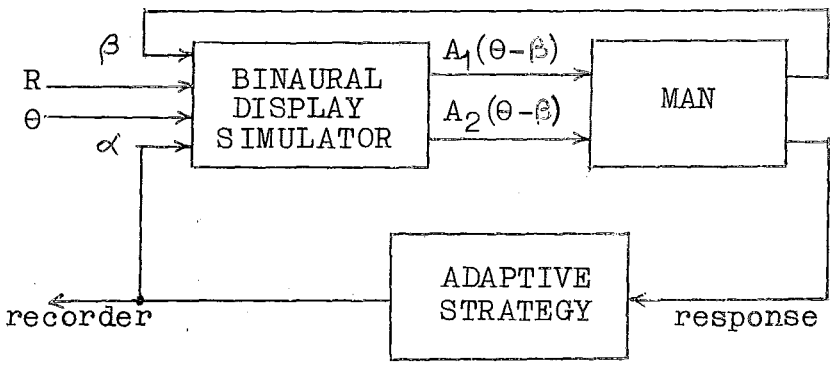


FIG. 8.3 THE USE OF A DISPLAY SIMULATION TO ESTIMATE K

If the value of α is controlled by some strategy that forces $\hat{\theta} \equiv \theta+\theta_0$, the value of K may be estimated directly from Eq. 5.17, i.e.

$$K = \frac{1}{4\alpha_c\gamma} \qquad 8.4$$

where α_c is the value of α when $\hat{\theta} \equiv \theta+\theta_0$. In this condition the expectation of the subject's response will remain constant and will be independent of the head direction β. For all other values of α the image direction will depend upon β (Eq. 8.3).

Three methods of controlling α were evaluated, in separate experiments, against three criteria:

i The ability of naive subjects to perform the task required by the method.

ii The rate and variability of the convergence of the estimate of α_c.

iii Repeatability of estimates of α_c.

For each method estimates of K were made at six frequencies (300, 500, 1000, 1500, 2000, 3000 Hz). The sensation level in each channel was set to 60 db, as measured in a 2 cc coupler, when $A_i(\theta-\beta) = 1$. All experiments were carried out with $\theta = 0$, and with the value of γ chosen to make the 6 db beamwidth $\psi_{\frac{1}{2}}$ equal to 90°.

The subject was seated in a room adjacent to the computer laboratory. The head direction β was monitored from a potentiometer, affixed to the headset upon which was mounted

a freely sliding vane that was restrained in azimuth only. In this way pitch and roll movements of the head were not significantly restricted. Fig. 8.4 shows a subject seated in the experimental apparatus with the EAI 590 hybrid computing system in the background. The image direction indicator held by the subject was used in the second and third experiments described below.

8.4.1 Direct Control of α by the Subject

In the first experiment the subjects were given direct control of α, through a potentiometer, and were instructed to find the position of the control at which the azimuthal position of the auditory image remained fixed in space. The method that they were told to use was to note any shift in image position between the extremes of the head movement and to null this shift with the control. It was found that a joy-stick or a thumb-wheel was the most convenient method of adjusting α.

Although the author was able to give convergent and repeatable estimates of α_c, the 16 other subjects who took part in the experiment were unable to perform the required task at all. Even when given detailed instructions on the effect of the control all subjects were unable to perform the task with any reliability. Estimates of α_c were scattered over a range of approximately 5:1 for each subject.

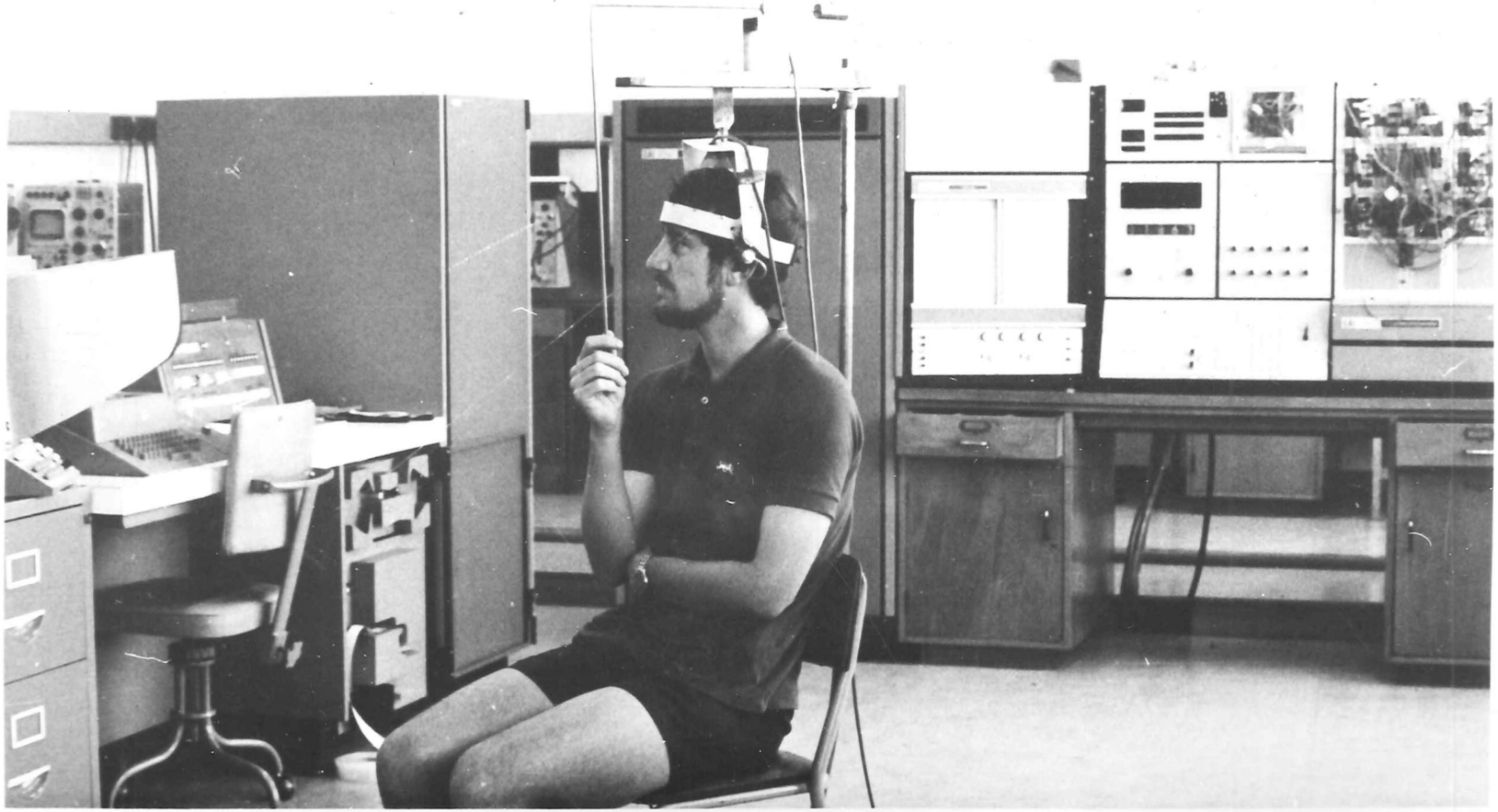


Plate 8.1 Equipment and computing facilities used in the interactive localization experiments. The subject's head position is monitored by the sliding vane attached to his headset and the image direction indicator is affixed to the boom above the head. (The EAI 640 digital computer is to the left, the EAI 693 hybrid linkage unit is behind the subject, and the EAI 580 analogue computer is at right).

It was therefore concluded the method was not a satisfactory procedure for estimating α_c because of the complexity of the interaction between the subject's control, the head direction and the image direction. This method of controlling α was therefore abandoned.

8.4.2 Integral of Response Error

The subject's task was modified to that of simply tracking the image direction as the head was slowly turned from side to side. The response θ' was measured with a potentiometer mounted above the subject's head (Plate 8.1) and the difference between the desired response θ and the actual response was integrated in the direction required to force the difference $\theta-\theta'$ to zero. The time dependence of α was given by

$$\alpha(t) = C \int_0^t (\theta-\theta')\,\mathrm{sgn}(\beta-\theta)\,dt + \alpha(0) \qquad 8.5$$

where C is a constant defining the system open-loop gain. The signum of the effective azimuth $-(\beta-\theta)$ is necessary to force the strategy to converge upon α_c. The complete feedback system and the analogue computer implementation of the strategy are shown in Fig. 8.5. Substitution for θ' in Eq. 8.5 shows that the strategy is the linear sum of a deterministic and a random process. Assume the head to be fixed and let $\theta_0 = 0$, so that for the ideal (error free) observer differentiation of Eq. 8.5 gives

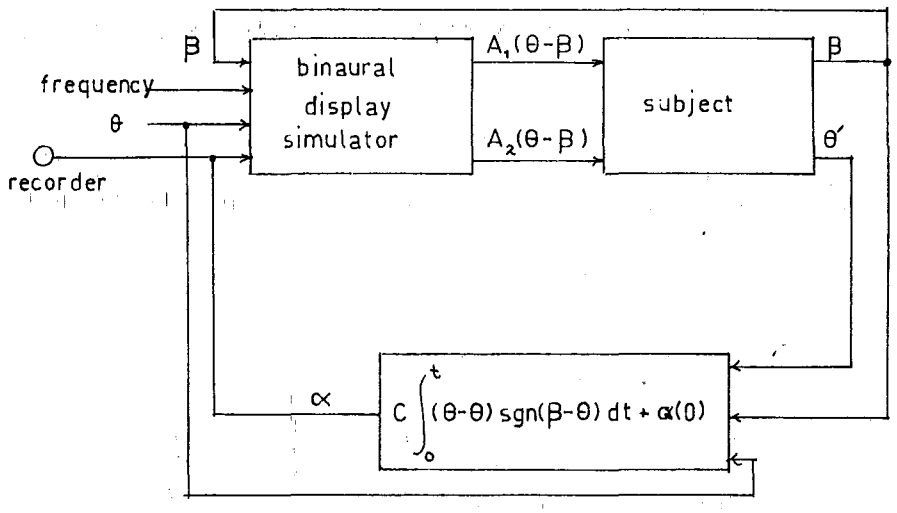


Fig. 8.5a The complete feedback loop for the integral of signed error method of controlling .

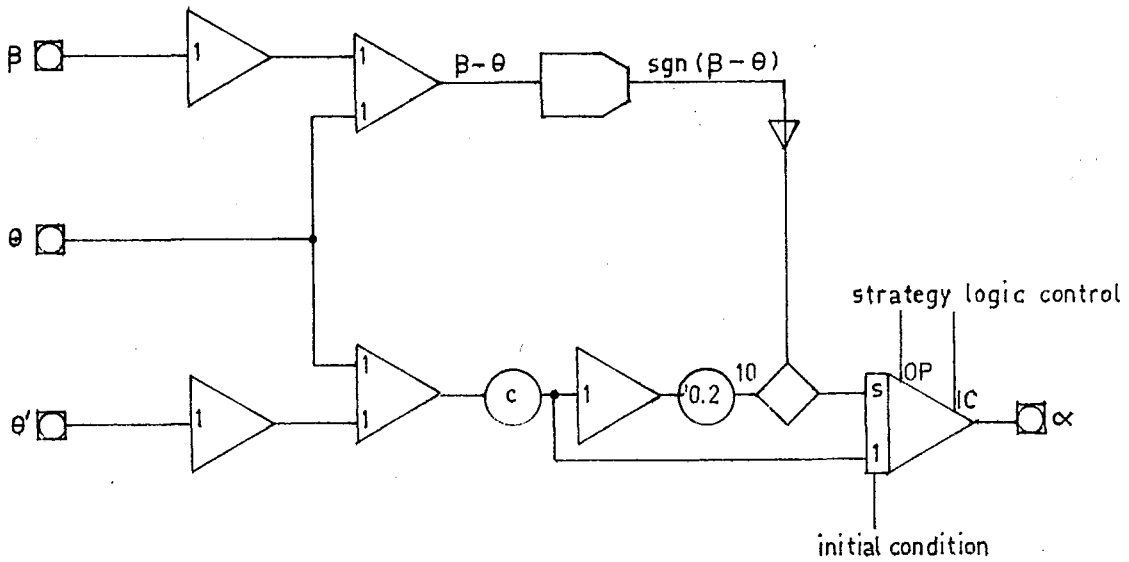


Fig. 8.5b Analogue computer patching to control .

$$\frac{d\alpha(t)}{dt} + 4CK\gamma(\theta-\beta)\alpha(t) + C\theta = 0 \qquad 8.6$$

and if the initial condition $\alpha(0) = 0$ is substituted the solution becomes

$$\alpha(t) = \frac{1}{4K\gamma}\left[1 - \exp(-4KC\gamma(\beta-\theta)t)\right] \qquad 8.7$$

so that

$$\lim_{t\to\infty} \alpha(t) = \frac{1}{4K\gamma}. \qquad 8.8$$

Because

$$\lim_{t\to\infty} \int_0^t \xi dt = 0 \qquad 8.9$$

the convergence will be unaffected by a noisy observer.

This strategy forces the subject to respond to some predetermined angle θ. Similar techniques could be used in other psychophysical situations in which it is desired to measure the stimulus magnitude required to evoke a response of a given magnitude on some response continuum. The subject need not be aware of the response magnitude being tracked. The method is, in fact, similar to that used in the Békésy audiometer[164] where a constant rate of change of stimulus magnitude in the direction indicated by a binary response is used to measure auditory thresholds.

A group of ten subjects was used to evaluate the method both with the head fixed, and with free head movement. All subjects gave convergent results with the head fixed but most failed to do so when free head movement was allowed. In these cases large amplitude fluctuations occurred in the recorded

value of $\alpha(t)$. Two examples of the variation of α during an experimental run are shown in Fig. 8.6. The first is from one of the few subjects who consistently gave convergent estimates of α_c with head movement, the second is from a subject who failed to give any convergent estimates of α_c.

There were several factors that contributed to the failure of this method. The first was a lack of knowledge of the transfer function of each observer in this tracking task so that the gain C to produce stable operation had to be found by trial and error. The task required the constant attention of the subject, for any lapse in attentiveness allowed the value of α to integrate away from α_c. Furthermore the task was sensitive to the magnitude of β. (Note that if $\beta = \theta$ the subject will be unaware of changes in α and divergent estimates will occur.) The major factor was the presence of θ_0 in the localization function. Unless $\theta_0 = 0$ the value of α to make $\hat{\theta} = \theta$ will be different for each value of β. Therefore, for successful operation of the method, any offset θ_0 must be compensated as in Eq. 5.8. Although the compensating IAD could be measured by an extension of this technique in which $\hat{\theta}$ was forced to zero, this was not investigated because of the other difficulties mentioned above.

8.4.3 Integral of the Signed Derivative of the Response Error

The previous method was modified so that the time differential of the response error (and hence the response itself)

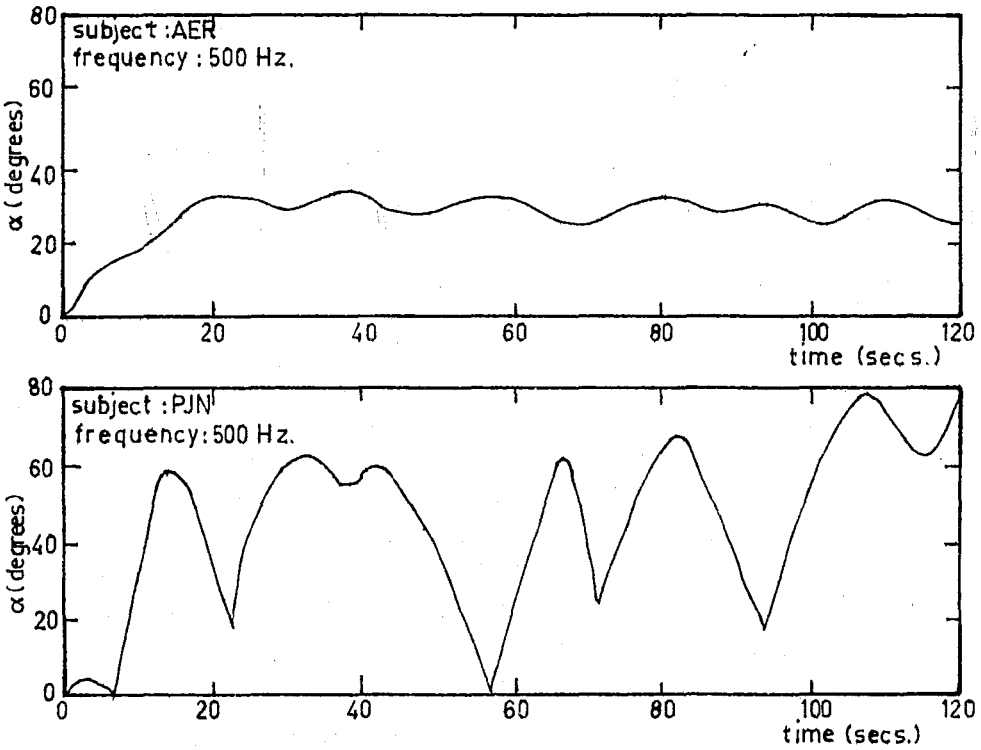


Fig. 8.6 Examples of stable and unstable experimental runs estimating the value of α to match the display to the localization function.

was integrated with a correct sign to form α. The subject's task was modified slightly to include the additional requirement that definite judgments of the image direction should be made with the head direction alternating about θ. Then at each judgment the value of α was incremented by an amount proportional to the total response change.

Let successive judgments be defined by the zero-crossings of the angle $\beta-\theta$, and let the nth such judgment be defined by the period $t_{n-1} < t < t_n$ where t_n is the time of the nth zero-crossing. Then the strategy increments α by $\delta\alpha_n$ according to

$$\delta\alpha_n = C \int_{t_{n-1}}^{t_n} \frac{d\theta'}{dt} \operatorname{sgn}(\beta-\theta)\, dt \qquad 8.10$$

$$= C(\theta'_n - \theta'_{n-1}) \operatorname{sgn}(\beta-\theta) \qquad 8.11$$

so that $\delta\alpha_n$ is defined only by the responses θ'_n at the successive zero-crossings of $\beta-\theta$. The subject's response strategy between these times does not affect the convergence of α_n. Let

$$\theta'_n = \beta_n + 4K\gamma\alpha_n(\theta - \beta_n) + \theta_0 + \xi_n \qquad 8.12$$

where the random variable ξ_n is distributed with zero mean, so that

$$E\{\delta\alpha_n\} = \frac{C(\beta_n-\beta_{n-1})(1-4K\gamma\alpha_{n-1})\operatorname{sgn}(\beta_n-\theta)}{1 + 4CK\gamma|\beta_n-\theta|} \qquad 8.13$$

which is independent of θ_0. Therefore this strategy does not depend upon prior compensation for any offset in the localization function. The convergence of α_n may be demonstrated by letting the head direction by symmetrically displaced about $\beta = \theta$, i.e.

$$\beta_n - \theta = (-1)^n(\beta-\theta) \qquad 8.14$$

This simplification, which is not a necessary condition for convergence, gives

$$E\{\delta\alpha_n\} = \frac{2C|\beta|(1-4K\gamma\alpha_{n-1})}{1+4KC\gamma|\beta-\theta|} \qquad 8.15$$

and

$$E\{\alpha_n\} = \alpha_{n-1} + D(1-4K\gamma\alpha_{n-1})$$

where

$$D = \frac{2C|\beta|}{1+4KC\gamma|\beta-\theta|}$$

and therefore

$$E\{\alpha_n\} = D\left[1 + \sum_{p=1}^{n-1}(1-4KD\gamma)^P\right] + (1-4KD\gamma)^n\alpha_0, \qquad 8.16$$

which is similar to the expression for an exponentially smoothed average[165]. For the case of a noisy observer, the estimation procedure may be viewed as the above deterministic successive approximation with a superimposed random component. Then the procedure may be written in the form of a generalised stochastic approximation[166]

$$\alpha_n = T_{n-1}(\alpha_0, \cdots, \alpha_{n-1}) + Z_{n-1} \qquad 8.17$$

where $T_{n-1}(\alpha_0, \cdots, \alpha_{n-1})$ is the error free transformation, and Z_{n-1} is the random noise component. Because

$$|T_n(\alpha_0,\cdots,\alpha_{n-1})-\alpha_c| = (1-4KD\gamma)|\alpha_{n-1}-\alpha_c| \quad 8.18$$

$$= (1-4KD\gamma)^{n-1}|\alpha_0-\alpha_c| \quad 8.19$$

if $0 < 1-4KD\gamma < 1$ then the procedure is convergent upon α_c with probability 1 in the mean square sense if

$$\sum_{n=1}^{\infty} E\{Z_n^2\} < \infty \quad 8.20$$

and

$$E\{Z_n\} = 0, \quad 8.21$$

by Dvoretsky's theorem[167] (Special Case 1). The first of these latter two conditions depends upon the subject's strategy in that for absolute convergence ($\lim_{n\to\infty} E\{(\alpha_n-\alpha_c)^2\} = 0$) the response error ξ_n must be a monotonic function of the magnitude of the response change. In the present experiment this was generally observed to be the case. Eqs 8.16 and 8.17 both indicate that

$$\lim_{n\to\infty} \alpha_n = \alpha_c.$$

The rate of convergence will depend upon the factor $(1-4DK\gamma)$ and will therefore be variable between subjects. It can be seen from Eq. 8.19 that rapid convergence will result if $4KD\gamma$ is made near to unity.

The strategy and its analogue computer implementation is shown in Fig. 8.7. The only difference to the previous strategy is the inclusion of an analogue differentiator. (The usual problem of high frequency noise in a differentiator is eliminated by the following integrator.) The constant C (Eq. 8.10) which defines the open-loop gain of the system was chosen empirically as a compromise between the rate of convergence and the disquieting effect of having the image position react violently to any action taken with the image indicator.

The same group of ten subjects who took part in the evaluation of the previous method (Section 8.4.2) were used to evaluate this experimental procedure. All subjects gave convergent estimates of α_c although the rate of convergence varied widely between subjects. During this part of the experiment the effect of different initial conditions α_0 upon the response was investigated. It was concluded that convergence was most rapid for naive subjects if α_0 was made large, so that identification of the task was obvious from the rapid lateralization that occurred with small head movements.

The convergence of α_n from different values of α_0 is shown for two subjects in Fig. 8.8. Several subjects reported spontaneously that in the convergent condition ($\alpha_n = \alpha_c$) the image appeared to originate from the indicator no matter which way they turned their heads.

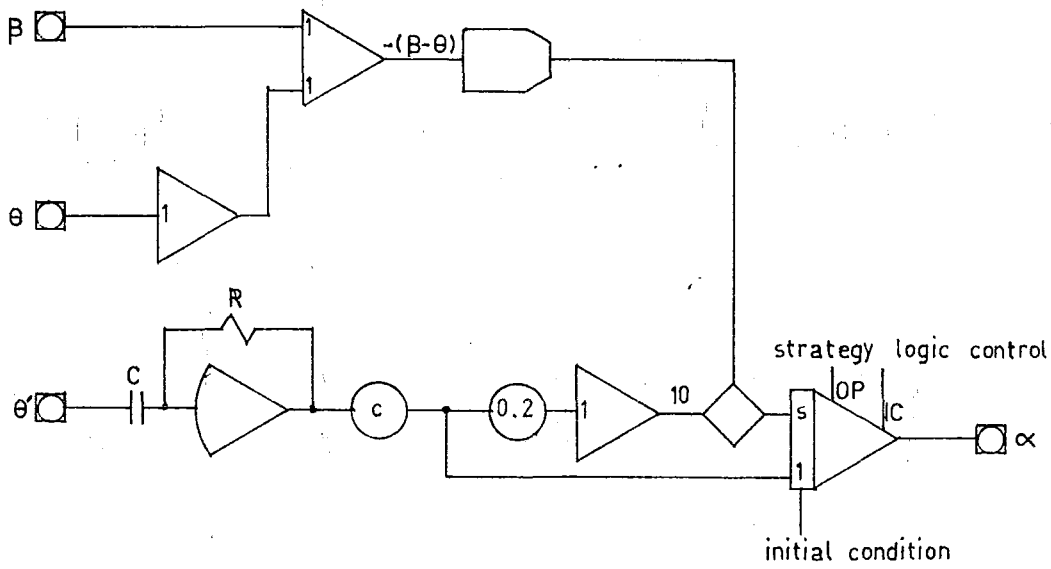


Fig. 8.7 Analogue computer patching for the integral of the signed derivative of response method for controlling α.

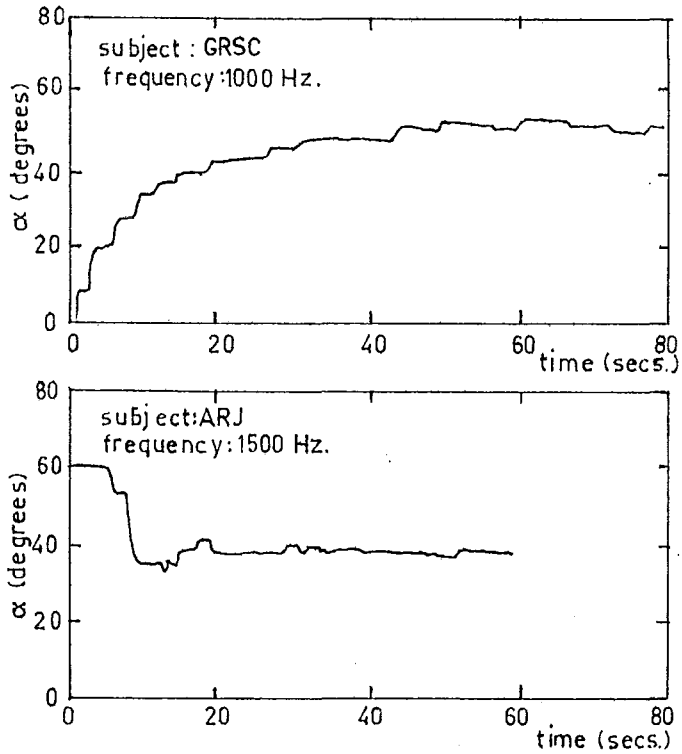


Fig. 8.8 Examples of the variation of during an experimental run for two subjects. Note the different initial conditions, and the step-like change in α at each response.

Four subjects were retained for an extended series of tests to estimate the within-subject variability of the estimates of α_c as measured by this procedure. Each of these sessions consisted of 18 trials of one minute duration, with three trials at each of the six frequencies in randomised order. An estimate of $k = 8.686K$ (Eq. 5.14) was made from each trial. All four subjects showed a pronounced daily variation in k at most frequencies. An F test was used to compare the between-days variance with the estimated within-cell variance for a days x frequency analysis. The between days difference was generally significant at the 0.05 significance level for most frequencies. Fig. 8.9 compares the daily estimates of k with the overall mean of four separate sessions for each of two subjects.

Two subjects (PSB and ARJ) had participated in the static dichotic localization experiment described in Section 4.3 approximately 12 months earlier. The frequency dependence of the localization function as measured by the two experiments is compared in Fig. 8.10; there are clearly pronounced differences between the results obtained in the two experiments.

It is not possible, without further corroborative evidence, to state whether the static or dynamic localization methods is preferable in the estimation of α_c to match the display to an individual. It was found that the variance between the estimates of k obtained in this experiment was greater than that of the regression coefficient β in the

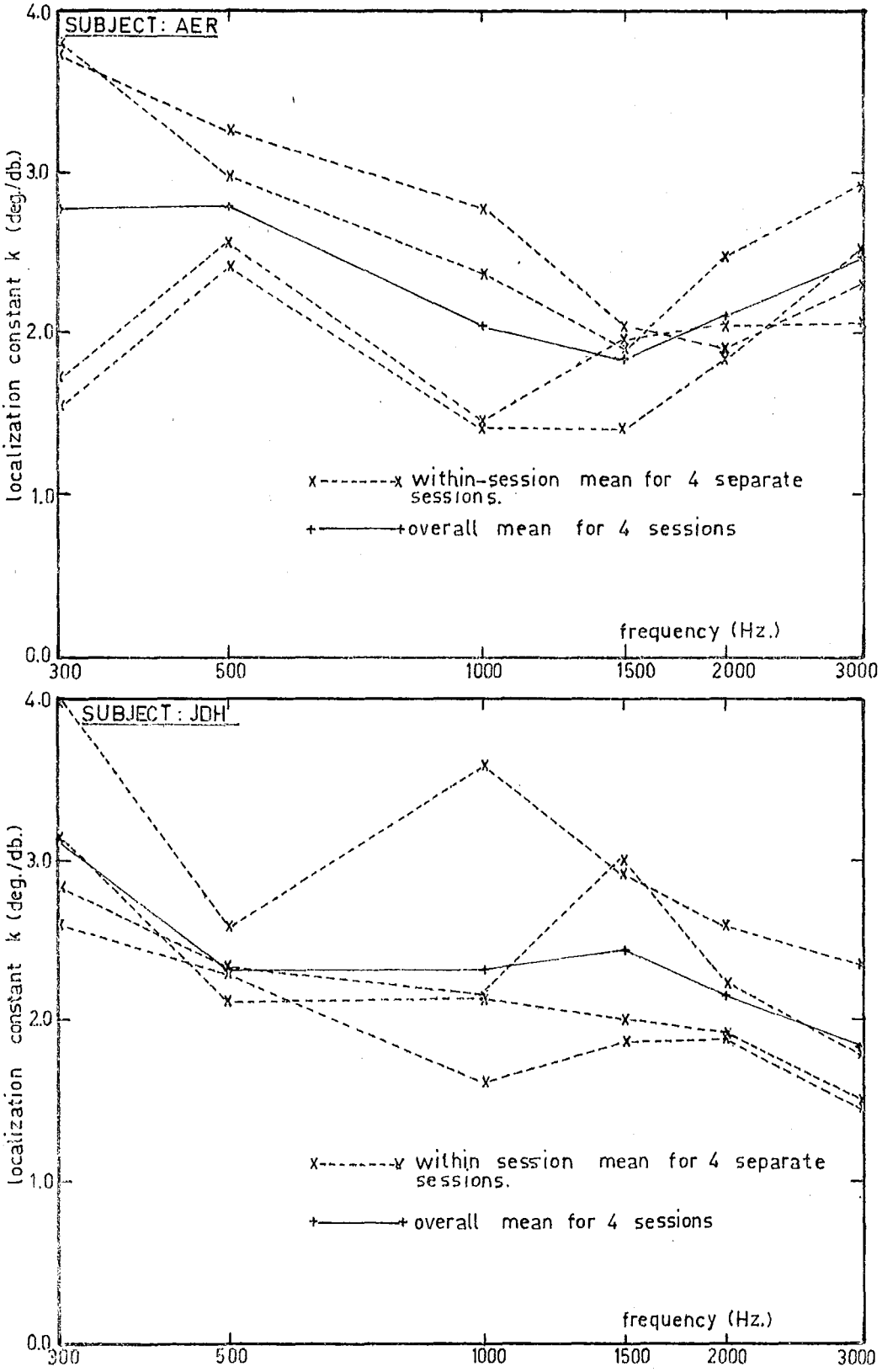


Fig. 8.9 Comparison of daily mean responses for two subjects over four sessions.

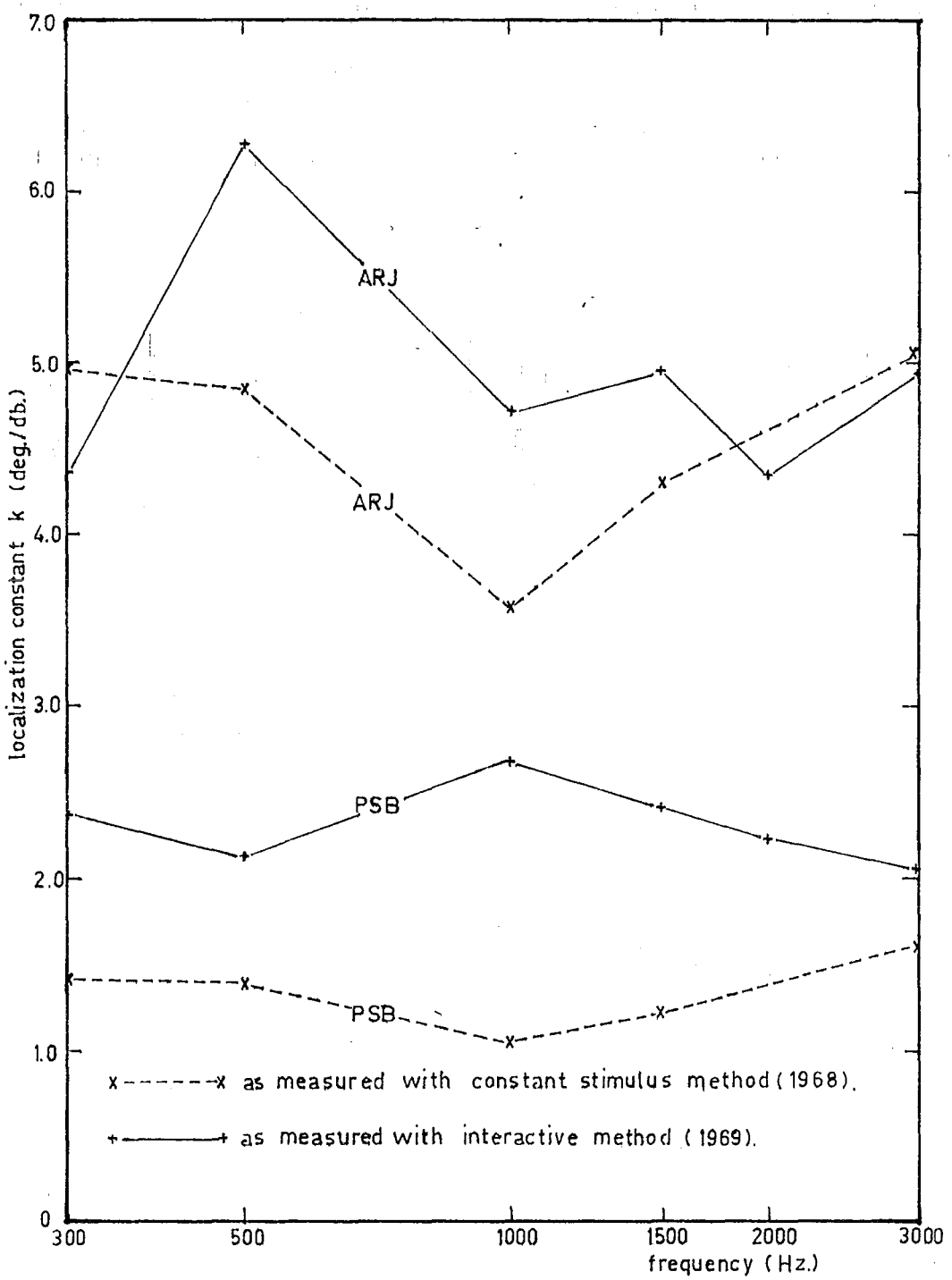


Fig. 8.10 Comparison of mean results of the interactive method of estimating the localization function with those obtained in the dichotic experiment 18 months previously, for two subjects. (see Fig. 4.3a)

static localization experiments with constant stimuli (Chapter 4). Pollack and Rose[27] reported a similar reduction in localization accuracy in a free-field experiment when free head-movement was allowed. There is a need for further investigation of the effects of head movement upon auditory localization in order to determine which experimental method is preferable.

The day to day variation in the estimates of k cannot be simply explained. The possibility of a time-varying localization function is ignored in the classical psychoacoustic methods where the data is obtained over a period of many days. It is conceivable, however, that the variability between days could be explained in terms of subject motivation. Experienced subjects realise that the estimation procedure will eventually reduce the need for changes in the indicator position between successive responses to zero. Therefore lapses of attention to the task could lead to an apparently convergent estimate of α_c.

These factors should be further investigated in a set of experiments designed to compare the classical and dynamic estimates of k within each experimental session.

8.5 Discussion

The author is not aware of any previous experiments using dichotic stimuli in which head movements provided secondary cues. There does not appear to be any literature

on the perceptual effects of distortions in the variation of IAD as a function of source direction. The available literature on binaural distortions only describes the effects of a uni-lateral attenuation[168]. Experiments such as those described above therefore provide a new method of studying the effects of secondary cues and the adaptation of the auditory system to distortions in the binaural function.

The use of an analogue computer simulation of the azimuthal display developed in this thesis demonstrated that subjects were able to perceive a dichotic image that remained fixed in space irrespective of the head direction. The use of the parametric form of the localization function, determined in static experiments, to specify a dynamic display was therefore justified.

CHAPTER 9

SUMMARY AND CONCLUSIONS

In this final chapter the principal results of the work presented in this thesis are summarised and possible areas for future research are indicated.

It was shown in Chapter 2 that the behavioural characteristics of human audition place a severe restriction upon the form of signal coding that can be used in a single dimensional auditory display. The use of the normal echo-location range display of direct propagation delay is precluded by the precedence effect, which is essential for passive listening in reverberant surroundings; which inhibits the perception and localization of all but the nearest of a field of correlated sound sources. For this reason it was concluded that a frequency domain range coding is optimal for the auditory display of multiple objects. The linear F.M. aids for the blind as proposed by Kay have an auditory display in which the stimulus frequency is directly proportional to object range.

In Chapter 3 a set of experiments on the localization phenomena associated with the prototype binaural mobility aid is described. The first of these experiments gave the first quantitative evidence that the display from this system contained binaural differences that could be recognized by the localization mechanism. It was then shown that of the three

binaural differences in the system specification only IAD was effective in displacing the image from the median plane. From the results of these experiments it was concluded that the frequency domain range coding had destroyed coherence in the microstructure of the binaural stimuli for an object not on the median plane. The IFD itself was ineffective in shifting the image and the nature of the stimuli was such that envelope delays were not recognized by the localization mechanism. The IFD did not destroy binaural fusion for all subjects, demonstrating that coherence in the microstructure is not a necessary condition for fusion of pulsed tonal stimuli, however fusion was intermittent for some subjects.

These experiments provided a basis for the display specification. It was decided that those binaural differences that were dependent upon propagation delays (ITD and IFD) should be eliminated, and the requirement was set that the display should be matched to the natural localization function for IAD. Because the literature contained few quantitative data on localization with IAD it was necessary to design and complete a series of experiments to verify the linear relationship between IAD (expressed in logarithmic units) and the image azimuth. The "within subject" and "between subjects" variation in the coefficients of the localization function were estimated over a wide range of stimulus parameters. It was found that the slope of the localization curve was dependent upon the stimulus frequency for all subjects, with a pronounced dip in the region of 1000Hz. Furthermore these

experiments showed that wide variations existed in the magnitude of the slope of individual localization functions, indicating that the display should be adjusted to each individual.

The linear relationships between IAD and image azimuth was used to define the form of the transducer angular response functions to preserve equality between object and image azimuth angles in an echolocation display. It was shown that under the constraint that the angular response should be an even function about the normal to the aperture surface, the required form was Gaussian, with the two receivers displaced by a fixed angle in opposite directions about the median plane. Differences in individual localization functions can be accommodated by varying the splay angle to match the slope of the localization function and the relative channel gains to match the intercept. Further analyses demonstrated the sensitivity of the azimuthal display to perturbations in the angular response of the aperture; in particular the monotonic nature of the object-image azimuth was shown to be destroyed by sidelobes or zeroes in the response.

The effect of additive Gaussian noise in the two channels, upon the azimuthal display, was examined by deriving the probability density functions for the azimuth estimator for an idealized model of the localization process. It was shown that the noise biases estimates toward the median plane. No attempt has yet been made to verify the applicability of the model to human localization of tones contaminated by noise.

In Chapter 6 a set of analyses relates the signals of the linear F.M. echolocation system to the display. A frequency domain signal to noise criterion was defined and used to show that the multiplicative signal processing maximises the signal to noise ratio. The signal to noise ratio attainable with this form of processing is identical to that at the output of a linear matched filter.

A discrete point reflector model of an extended reflecting object was used to show that a reflecting object of finite extent in range and azimuth gives rise to display signals described by Rayleigh statistics. The target's azimuthal extent was then shown to bias estimates of its "centre of gravity" toward the median plane.

The response of a linear aperture to generate a far-field angular response of the required Gaussian form was then derived. Present technology does not allow accurate control of the response of air-borne ultrasonic transducers (solid-dielectric electrostatic) of the type used in the mobility aids, therefore no attempt has been made to control the angular response. However, measurements upon the transducers has shown that they can be selected, with a high yield, to closely approximate the required angular response. Future attention should be directed toward the manufacture of repeatable and controlled transducers.

The final section of this thesis is devoted to psychophysical measures of the display performance. In Chapter 7

the attainable resolution in a static two component environment was investigated in two separate experiments. In the first experiment subjects demonstrated that the minimum frequency separation between two tones required to allow each to be heard separately could be approximated by a power law with an exponent near unity. The second experiment showed that a much greater frequency separation was necessary before the lateralization of one tone could be determined. Even at separations greater than one octave the lateralization ability of some subjects was impaired. Neither of these experiments allowed the subjects the kinaesthetic factors associated with headmovement. A set of experiments is required that will estimate the improvement in azimuthal resolution when free headmovement is allowed.

A computer simulation of the idealised auditory display was described in Chapter 8. The purpose of this simulator was to study localization phenomena associated with the display when headmovement was allowed. This type of study has the advantage over direct use of the mobility aid that the display parameters can be controlled and changed at will, and the subjects responses can be monitored and processed in real time. The simulation was extended to include the subject in a feedback loop with an adaptive strategy controlling the angular displacement of the two simulated aperture response functions. It was found that by incrementing the angular displacement by an amount proportional to the change in

the indicated image direction between successive judgements a rapid and convergent estimate could be made of the coefficients of the localization function. However, these estimates showed a pronounced day to day variation for all subjects which indicated that the localization function might not be static. Further research is necessary to establish the stability of the localization function.

The author recommends that the following areas of research be investigated to further the work presented in this thesis:

i) Localization of tones with IAD: There is a need for further study of the IAD localization in order to determine the stability of the coefficients of the localization function and to define a "clinical" testing procedure to enable the system to be matched to each individual. These studies should compare and correlate the localization function as measured by classical and adaptive methods within each session to eliminate biases caused by the experimental methods.

ii) The localization function for Rayleigh noise: It was shown in Chapter 6 that for an extended target the signals at the display are described by Rayleigh statistics. The relationship between the localization function for tones and narrow band Rayleigh noise should be investigated. Although the author can see no reason why the two cases should differ this should be verified.

iii) The effect of ambient noise upon localization: The effect of both wide band and narrow band noise upon the

localization of tones should be studied to determine the sensitivity of this display form to ambient environmental conditions. Such a study would provide a possible method of testing the validity of the localization model presented in Chapter 5.

iv) Azimuthal resolution with headmovement: It is suggested that the computer simulation described in Chapter 8 be extended to a multiple object environment so that the resolution capabilities of the display may be studied in conditions of free headmovement.

v) The relevance of this display to the sensory requirements for human mobility: It is suggested that experiments involving motor actions in a simulated environment be conducted to determine the sensitivity of the sensory-motor system to the display parameters. The results of this study should be correlated with the results of field evaluations of mobility aids designed around the display presented here.

vi) The technology of the mobility aid: Because the azimuthal information encoding takes place at the receiver transducers these elements must be well controlled and repeatable in their operation. There is a need for greater understanding of the operation of the solid-dielectric electrostatic transducers used in the mobility aid, and effort should be placed upon the manufacture of transducers with repeatable characteristics and methods of controlling the surface amplitude taper to produce the angular response of the required form.

APPENDIX 1

A Model of Binaural Spatial Localization

The model presented here was not conceived as a quantitative model of the psychological process of localization; it is intended to show that the action of the known binaural phenomena retains the spatial information impressed upon wideband signals by the head and pinnae.

The Role of the Outer Ear: Consider a remote sound source with a power density spectrum $\Phi(\omega)$. Directional information is impressed upon wideband signals by the head and pinnae as a linear operation on the input signal.

$$\Phi'_i(\omega) = \Phi(\omega)\left|H_i(\omega;\theta)\right|^2 \qquad \text{A1.1}$$

where $\Phi'_i(\omega)$, $i = 1,2$, are the excitation at the two eardrums, and $H_i(\omega,\theta)$ is the transfer function of the outer ear. It has been shown that this function is extremely dependent upon the spatial coordinates θ of the source. The model states that the coordinates are inferred from some monotonic function of the ratio of the two power spectra, for example

$$G(\omega;\theta) = \frac{\Phi_1(\omega)}{\Phi_2(\omega)} = \left|\frac{H_1(\omega;\theta)}{H_2(\omega;\theta)}\right|^2 \qquad \text{A1.2}$$

The Function of the Cochlea: The transducer mechanism from vibrational to neural energy is located in the cochlea, shown schematically in Fig. A1.1. Although shown stretched

out the cochlea is in fact a spiral cavity in the temporal bone. The primary innervation is spread along the basilar

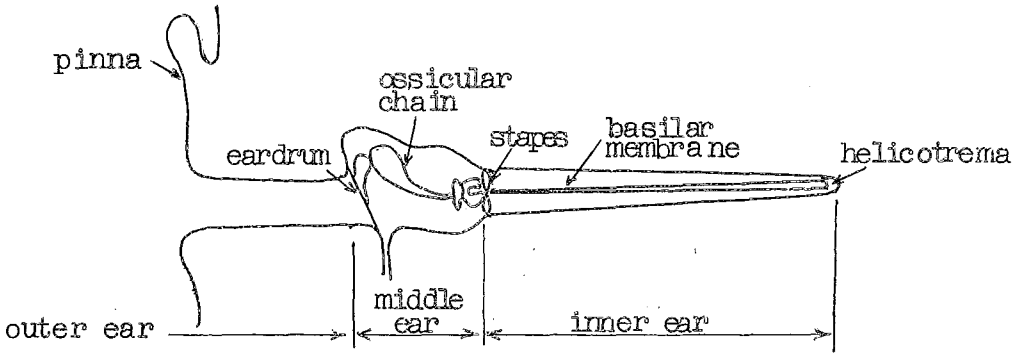


FIG A1.1: SCHEMATIC REPRESENTATION OF THE EAR

membrane in the cochlear partition. This membrane has interesting mechanical properties in that it is light and stiff in the basal (stapes) region and tapers to become more massive and compliant at the apical (helicotrema) end[70]. This feature causes maximal vibrational amplitude at the basal end for high frequency tones, and a movement of the maximum toward the helicotrema as the frequency is lowered[71]. The membrane acts as a form of mechanical frequency analyzer with the peripheral neurons responding to shear stress between the tectorial membrane and the organ of Corti[72].

The motion of a point on the membrane distance ℓ from the stapes (Fig. A1.1) may be written as a convolution

$$y_\ell(t) = f(t) \otimes h_\ell(t) \qquad \text{A1.3}$$

where $y_\ell(t)$ is the shearing motion

$f(t)$ is the excitation at the stapes

$h(t)$ is the response of the point ℓ to an impulse at the stapes. Flanagan[73] has derived the form of the impulse

takes place only between fibres originating from corresponding points on the two cochlae, i.e. place is preserved in the ascending system.

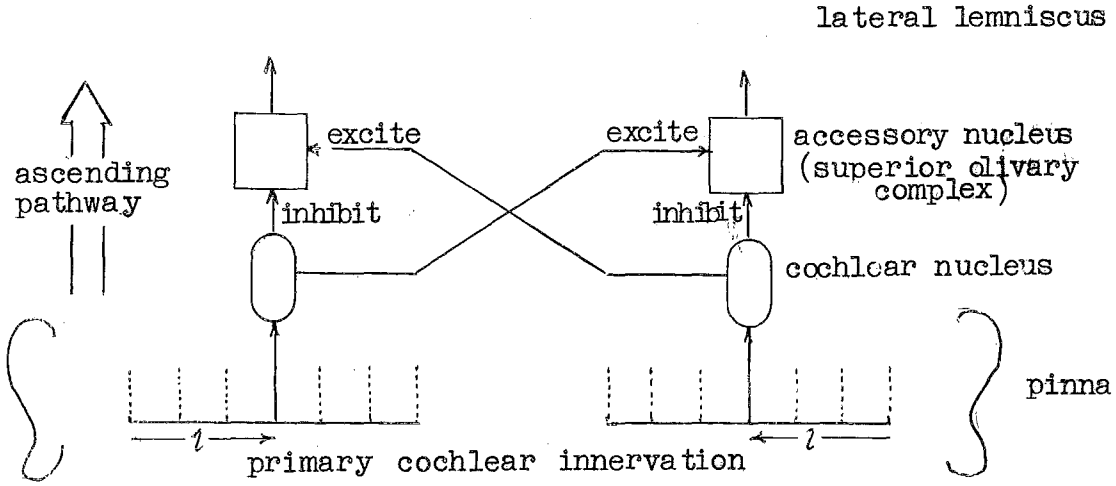


FIG A1.2: SIMPLIFIED SCHEMATIC OF PERIPHERAL BINAURAL INTERACTION

It is generally accepted that binaural differences are resolved at the lowest level of interaction, namely the accessory nuclei in the superior olivary complex. The interaction at this level has been investigated by several groups[75,76]. Both nuclei receive afferents from the ipse- and contralateral cochlear nuclei. It has been found that contralateral stimulation will excite the efferents from the accessory nuclei while ipsilateral stimuli will inhibit these efferents. The results of Hall[76] indicate that it is reasonable to write the outputs of the two separate nuclei in the form

$$N_i(l) = a \log[b\Phi_C^{\ell}(\omega_l)] - d \log \frac{\Phi_I(\omega_l)}{\Phi_C(\omega_l)} - f_l(\tau) \qquad \text{A1.10}$$

response $h_\ell(t)$ and has shown that the frequency domain transfer function of the point ℓ may be expressed in normalised form as

$$H_\ell(\omega) = H(\frac{\omega}{\omega_\ell})$$

where ω_ℓ is the angular frequency of maximum vibrational amplitude of the point ℓ. Then if the spectrum $\Phi_i'(\omega)$ is slowly varying with ω the power density spectrum of the vibration at the point ℓ will be

$$Y_\ell(\omega) \doteq \Phi_i(\omega) \left|H(\frac{\omega}{\omega_\ell})\right|^2$$

and the power of the vibration will be by Parsevals theorem

$$\int_{-\infty}^{\infty} y_\ell^2(t)\,dt \doteq \frac{E(\ell)}{2\pi} \Phi_i(\omega_\ell)$$

where $E(\ell) = \int_{-\infty}^{\infty} \left|H(\frac{\omega}{\omega_\ell})\right|^2 d\omega$

As a further simplification Fechner's Law[74] is assumed to describe the neural transduction process, i.e. we write the neural output as being proportional to the logarithm of the excitation, which is taken as either the energy of vibration or the power density spectrum at $\omega = \omega_\ell$.

Then the activity $N_i(\ell)$, $i = 1,2$, may be written

$$N_i(\ell) = a \log[b\Phi_i'(\omega)] \qquad \text{A1.9}$$

where a and b are constants of proportionality.

The Ascending Auditory System: A simplified schematic representation of the peripheral centres of binaural interaction is given in Fig. A1.2. It is assumed that interaction

where the subscripts C and I refer to the contralateral and ipsilateral stimuli, and where $f_\ell(\tau)$ is an odd function of the interaural delay τ.

It suffices to show, for this model, that the directional information has not been destroyed and can be derived from a direct comparison of activity at some higher centre. For example, consider the modulus of the difference of activities $N_i(\ell)$ and define

$$G(\ell;\theta) = |N_1(\ell) - N_2(\ell)| \qquad \text{Al.11}$$

$$= \left|(a+2d)\ \log\left|\frac{H_1(\omega_\ell,\theta)}{H_2(\omega_\ell,\theta)}\right|^2 - 2f_\ell(\tau)\right| \qquad \text{Al.12}$$

from substitution of Eqs Al.1 in Eq. Al.10.

Consider the case of headphone listening. The modification of the input spectra by the pinnae will contain no spatial information and only the normal dichotic binaural cues of time and amplitude difference may be resolved. True spatial localization, in terms of this model, is then impossible and only the 'sidedness' may be reported.

APPENDIX 2

A High Resolution Spectrum Analyzer

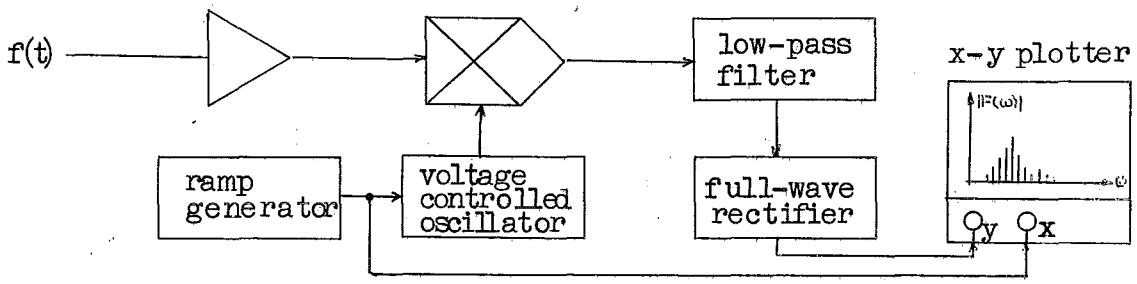


FIG. A2.1 THE ELEMENTS OF THE BASIC SPECTRUM ANALYZER

The method of spectral analysis described here was developed by the author to resolve the fine line structure of the spectra of the auditory display from the linear FM mobility aids. Although simple in concept, it is capable of greater resolution than any known commercial instrument.

Let the signal to be analysed be written

$$V_s(t) = \sum_{n=0}^{\infty} A_n \cos(n\omega_s t + \theta_n) \qquad \text{A2.1}$$

where A_n and θ_n are the amplitude and phase of the nth spectral component and ω_s is the angular repetition frequency. Let this signal be multiplied by a reference signal

$$V_R(t) = A_R \cos(\omega_0 t + \tfrac{1}{2}mt^2 + \phi_0) \qquad \text{A2.2}$$

where A_R is the constant amplitude, ω_0 is the instantaneous angular frequency when $t = 0$, m is the rate of change of angular frequency and ϕ_0 is the phase angle. The difference component of the product is

$$V_m(t) = \sum_{n=0}^{\infty} \frac{A_n A_R}{2} \cos((n\omega_s - \omega_0)t - \tfrac{1}{2}mt^2 + \theta_n - \phi_0) \qquad A2.3$$

and each component of this may be examined in turn by passing $V_m(t)$ through a low-pass filter with a cut-off frequency ω_c such that $\omega_c < \frac{\omega_s}{2}$. Then a quasi-static analysis of the filter output $V_0(t)$ shows that

$$V_0(t) = \frac{A_R A_m}{2} \cos((n\omega_s - \omega_0)t - \tfrac{1}{2}mt^2 + \theta_n - \phi_0)$$

for any n satisfying $|n\omega_s - \omega_0 - mt| \leq \omega_c$ and

$$V_0(t) = 0$$

if $|n\omega_s - \omega_0 - mt| > \omega_c$ for all n.

The amplitude of the oscillating voltage $V_0(t)$ is directly proportional to A_n and may be displayed to complete the spectrum analyzer. The system shown in Fig. A2.1 used a full wave rectifier to give the oscillating voltage the appearance of a spectral 'line'.

In practice the response of the low-pass filter to the transient frequency excitation sets a limit on the usable sweep rate m.

An analogue computer patching diagram for the spectrum analyzer is given in Fig. A2.2. The filter used in the analysis of the audio spectra was an active filter with a cut-off frequency of 1 Hz[78], thus giving a resolution of 2 Hz. The sweep generator was a Hewlett-Packard 3300A function generator used in the voltage controlled mode. The display used in the initial measurements was an X-Y plotter, but it is recommended that a storage oscilloscope be used for future applications because of the faster response.

APPENDIX 3

Analysis of Variance to Test the Significance of Differences in the Linear Regression Slopes of Data derived from a Three Factor Variation of Parameters

Given a total of N groups of data, in which observations are made of a linearly dependent variable y at fixed levels of an independent variable x, from a three way factorial experiment where the number of levels in each factor are n_i, n_j, n_k (i.e. $N = n_i n_j n_k$), for the pth observation a suitable model is

$$y_{ijkp} = \alpha_{ijk} + \beta_{ijk} x_p + \xi \qquad \text{A3.1}$$

where the random variable ξ is normally distributed with zero mean and variance σ^2, and α_{ijk}, β_{ijk} are linear regression coefficients. The problem is to test null hypotheses as to the significance of differences between the estimates of β_{ijk} for the main effects and interactions of the factorial design.

If there is a total of t observations of the dependent variable in each of the N groups, with the same number of observations at any level of the independent variable in all groups, the standard algebraic identity splitting the total sum of squares into between and within groups components may be written

$$\sum_{ijk}\sum_{p}(y_{ijkp}-\bar{y}_{....})^2$$

$$=\sum_{ijk}t(\bar{y}_{ijk\cdot}-\bar{y}_{....})^2+\sum_{ijk}\sum_{p}(y_{ijkp}-\bar{y}_{ijk\cdot})^2 \qquad \text{A3.2}$$

where the dot notation is used to denote means. The within groups sum of squares may be further subdivided, for the analysis of variance for each line shows that

$$\sum_{p}(y_{ijkp}-\bar{y}_{ijk\cdot})^2$$

$$=\beta_{ijk}^2\sum_{p}(x_p-\bar{x}_{\cdot})^2+\sum_{p}[y_{ijkp}-\bar{y}_{ijk\cdot}-\beta_{ijk}(x_p-\bar{x}_{\cdot})]^2 \qquad \text{A3.3}$$

and summing over the N groups gives

$$\sum_{ijk}\sum_{p}(y_{ijkp}-\bar{y}_{ijk\cdot})^2=\sum_{ijk}\beta_{ijk}\sum_{p}(x_p-\bar{x}_{\cdot})^2$$

$$+\sum_{ijk}\sum_{p}[y_{ijkp}-y_{ijk\cdot}-\beta_{ijk}(x_p-\bar{x}_{\cdot})]^2 \qquad \text{A3.4}$$

The last term is a <u>residual sum of squares</u> with N(t-2) degrees of freedom.

Let $W=\sum_{p}(x_p-\bar{x}_{\cdot})^2$

then since

$$\sum_{ijk}(\beta_{ijk}-\bar{\beta}_{...})^2=\sum_{ijk}\beta_{ijk}^2-N\bar{\beta}_{...}$$

the term $\sum_{ijk}\beta_{ijk}^2\sum_{p}(x_p-\bar{x}_{\cdot})^2$ in Eq. A3.4 may be written

$$W\sum_{ijk}\beta_{ijk}^2=W\sum_{ijk}(\beta_{ijk}-\bar{\beta}_{...})^2-NW\bar{\beta}_{...}^2 \qquad \text{A3.5}$$

so that the total sum of squares, Eq. A3.2, may be written

$$\sum_{ijk}\sum_{p}(y_{ijkp}-\bar{y}_{....})^2 = NW\bar{\beta}^2_{...}+\sum_{ijk}t(\bar{y}_{ijk.}-\bar{y}_{....})^2$$

$$+W\sum_{ijk}(\beta_{ijk}-\bar{\beta}_{...})^2+\sum_{ijkp}\sum_{p}[y_{ijkp}-\bar{y}_{ijk.}-\beta_{ijk}(x_p-\bar{x}_{.})]^2$$

A3.6

The 'between group slopes' sum of squares $W\sum_{ijk}(\beta_{ijk}-\bar{\beta}_{...})^2$ and its corresponding N-1 degrees of freedom may be partitioned in components representing the main effects and the interactions, as in a normal analysis of variance for a factorial design. The complete analysis of variance is summarised in Table A3.1.

TABLE A3.1 ANALYSIS OF VARIANCE FOR LINEAR REGRESSION COEFFICIENTS IN A 3 WAY FACTORIAL EXPERIMENTAL DESIGN

SOURCE OF VARIANCE	SUM OF SQUARES	d.f.
Common Slope:	$NW\bar{\beta}^2_{...}$	1
Between Group Means:	$\sum_{ijk} t(\bar{y}_{ijk.}-\bar{y}_{....})^2$	N-1
Between Group Slopes:		
Factor i	$W\sum_{ijk}(\bar{\beta}_{i..}-\bar{\beta}_{...})^2$	n_i-1
Factor j	$W\sum_{ijk}(\bar{\beta}_{.j.}-\bar{\beta}_{...})^2$	n_j-1
Factor k	$W\sum_{ijk}(\bar{\beta}_{..k}-\bar{\beta}_{...})^2$	n_k-1
Interaction (ixj)	$W\sum_{ijk}(\bar{\beta}_{ij.}-\bar{\beta}_{i..}-\bar{\beta}_{.j.}+\bar{\beta}_{...})^2$	$(n_i-1)(n_j-1)$
Interaction (ixk)	$W\sum_{ijk}(\bar{\beta}_{i.k}-\bar{\beta}_{i..}-\bar{\beta}_{..k}+\bar{\beta}_{...})^2$	$(n_i-1)(n_k-1)$
Interaction (jxk)	$W\sum_{ijk}(\bar{\beta}_{.jk}-\bar{\beta}_{.j.}-\bar{\beta}_{..k}+\bar{\beta}_{...})^2$	$(n_j-1)(n_k-1)$
Interaction (ixjxk)	$W\sum_{ijk}(\beta_{ijk}-\bar{\beta}_{ij.}-\bar{\beta}_{.jk}-\bar{\beta}_{i.k}+\bar{\beta}_{i..}+\bar{\beta}_{.j.}+\bar{\beta}_{..k}-\bar{\beta}_{...})^2$	$(n_i-1)(n_j-1)(n_k-1)$
Residual	$\sum_{ijk}\sum_{p}[(y_{ijkp}-\bar{y}_{ijk.}) - \beta_{ijk}(x_p-\bar{x}_{.})]^2$	N(t-2)
TOTAL	$\sum_{ijk}\sum_{p}(y_{ijkp}-\bar{y}_{....})^2$	Nt-1

APPENDIX 4

An Approximation to $\frac{I_1(x)}{I_0(x)}$ for Large Arguments ($x > 3$)

The confluent hypergeometric function ${}_1F_1(\alpha;\gamma;\beta)$ defined as[136]

$$ {}_1F_1(\alpha;\gamma;\beta) = 1 + \frac{\alpha}{\gamma}\frac{\beta}{1!} + \frac{\alpha(\alpha+1)}{\beta(\beta+1)}\frac{\beta^2}{2!} + \cdots\cdots \qquad \text{A4.1} $$

may be expressed in terms of the modified Bessel function of the first kind, in particular

$$ {}_1F_1(-\tfrac{1}{2};1;-2x) = e^{-x}[(1+2x)I_0(x) + 2xI_1(x)] \qquad \text{A4.2} $$

and

$$ {}_1F_1(\tfrac{1}{2};1;-2x) = e^{-x} I_0(x) \qquad \text{A4.3} $$

Therefore

$$ \frac{I_1(x)}{I_0(x)} = \frac{1}{2x}\,\frac{{}_1F_1(-\tfrac{1}{2};1;-2x)}{{}_1F_1(\tfrac{1}{2};1;-2x)} - (2x+1) \qquad \text{A4.4} $$

Now if $|\beta| >> |\alpha| + 1$ and $|\beta| >> |\gamma - \alpha|$

$$ {}_1F_1(\alpha;\gamma;-\beta) \doteq \frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}\,\beta^{-\alpha} \qquad \text{A4.5} $$

where $\Gamma(x)$ is the Gamma function, so that if $x >> 3$

$$ \frac{I_1(x)}{I_0(x)} \doteq \frac{\Gamma(\tfrac{1}{2})}{\Gamma(\tfrac{3}{2})} - \left(1 + \frac{1}{2x}\right) \qquad \text{A4.6} $$

or

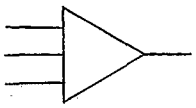
$$ \frac{I_1(x)}{I_0(x)} \doteq 1 - \frac{1}{2x} \qquad \text{A4.7} $$

since $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Eq. A4.7 is the desired approximation.

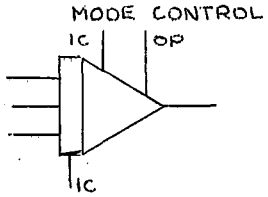
APPENDIX 5

Glossary of Analogue Computer Notation

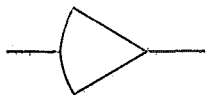
ANALOGUE COMPONENTS:



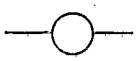
SUMMING AMPLIFIER



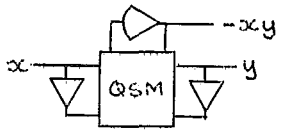
INTEGRATOR



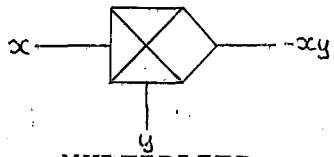
HIGH GAIN AMPLIFIER



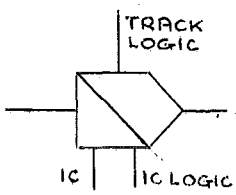
POTENTIOMETER



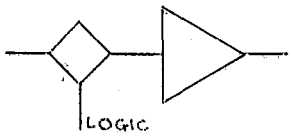
QUARTER SQUARE MULTIPLIER



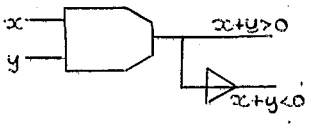
MULTIPLIER



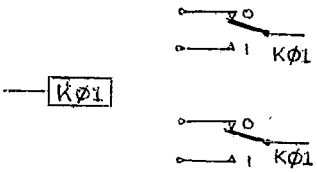
TRACK/STORE (T/S) AMPLIFIER



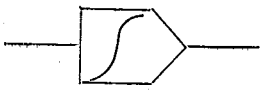
DIGITALLY CONTROLLED ANALOGUE (D/A) SWITCH



ELECTRONIC COMPARATOR

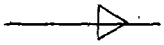


FUNCTION RELAY

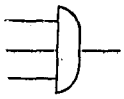


MULTIPLE DIODE FUNCTION GENERATOR

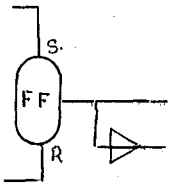
LOGIC COMPONENTS:



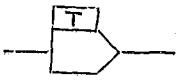
LOGIC INVERTER



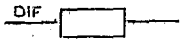
"AND" GATE



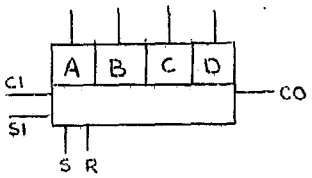
FLIP-FLOP



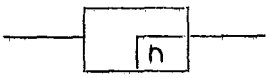
MONOSTABLE TIMER



LOGIC DIFFERENTIATOR



GENERAL PURPOSE REGISTER



DECIMAL DOWN COUNTER

APPENDIX 6

An Analogue Computer Implementation of UDTR

Fig. A6.1 shows a possible analogue computer configuration that could be used to estimate a single DL in IAD using the UDTR technique. Modified analogue accumulators are used to store and increment the stimulus level. The left hand T/S amplifier of each pair is normally in the store mode and the right hand one is in the track mode. At each change of stimulus level this order is reversed for a period T_2 as set by a monostable timer. The output from each of the left hand T/S amplifiers is then multiplied by a fixed fraction of the previous level which is then transferred to the right-hand T/S amplifier when the monostable timer reverts to its resting state. In this way the analogue stimulus level for each channel is stored in the right hand T/S amplifiers.

The binary responses are processed by the logic shown at the top of the figure. A decimal down-counter is used to count the n positive responses necessary to decrement the stimulus level; this counter is reset to its original state by any negative response.

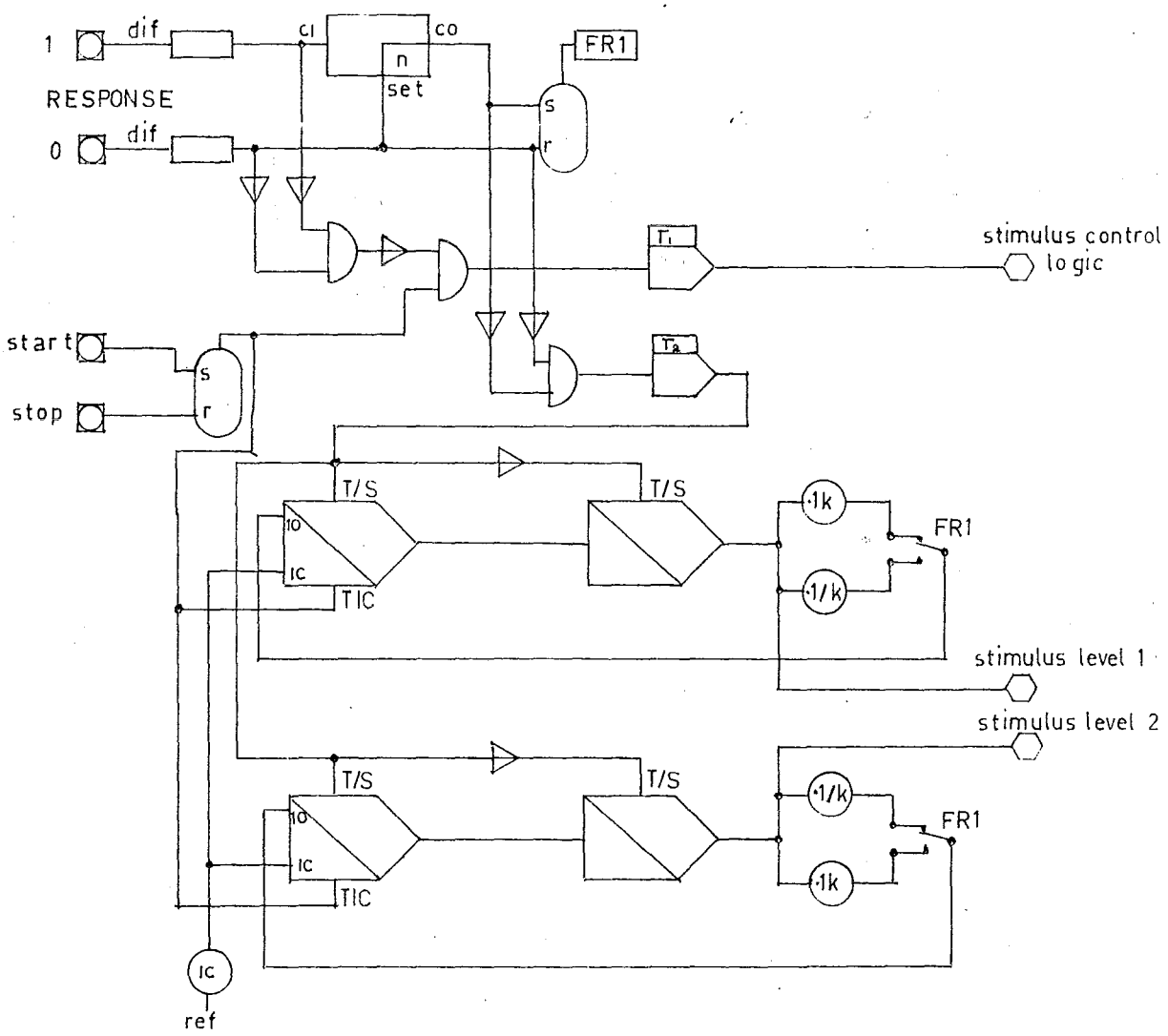


Fig. A6.1 Analogue computer implementation of the UDTR technique to estimate the difference limen in IAD.

REFERENCES

1. Clark, L.L. (Ed) Proc. Int. Cong. on Technology and Blindness. American Foundation for the Blind (AFB); New York (1963) Vols 1-3.

2. Clark, L.L. (Ed) Proc. of the Rotterdam Mobility Research Conference, AFB: New York (1965),

3. Dufton, R. (Ed) Proc. Int. Conf. on Sensory Devices for the Blind. St. Dunstans: London (1966).

4. Frank, W.E. "Instrumentation Requirements in Sensory Aids", Annals N.Y. Academy of Science, 60, 869-875, (1955).

5. Kolers, P.A. "Some General Remarks on Mobility Instrumentation Research", Proc. Rotterdam Mobility Res. Conf. AFB: New York, (1965) pp 235-241.

6. Jacobson, B. "Ambient-Light Obstacle Detector with Tactile Output", Proc. Int. Cong. on Techn. and Blindness, AFB: New York, (1963), Vol 1, pp 187-192.

7. Nelkin, A. "Ultrasonic Aid for the Blind" Proc. Rotterdam Mobility Res. Conf. AFB: New York, (1965) pp 17-26.

8. Kay, L. "Orientation of Bats and Men by Ultrasonic Echolocation" Brit. Comm. and Electronics 8, 582-586, (1961).

9. Kay, L. "A Plausible Explanation of the Bat's Echolocation Acuity" Animal Behaviour 10, 34-41, (1962).

10. Kay, L. "Ultrasonic Guidance for the Blind" Proc. Int. Cong. on Tech. and Blindness, AFB: New York, (1963) Vol 1, pp 137-156.

11. Kay, L. "Ultrasonic Mobility Aids for the Blind" Proc. Rotterdam Mobility Res. Conf. AFB: New York, (1965), pp 9-16.

12. Kay, L. "Ultrasonic Spectacles for the Blind" Proc. Int. Conf. on Sensory Devices for the Blind, St. Dunstans: London (1966) pp 275-292.

13. Martin, G. "Electronics and Transducers for an Ultrasonic Blind Mobility Aid" M.E. Thesis, University of Canterbury, 1969.

14. Carroll, Rev. T.J. "Blindness, What it is, What it does and how to live with it", Little, Brown: Boston (1961).

15. Starkiewicz, W. and Kuliszewski, T. "Progress Report on the Elektroftalm Mobility Aid", Proc. Rotterdam Mobility Conf. AFB: New York (1965), pp 27-38.

16. Gardiner, K.W. and Bliss, J.C. "An Optical to Tactile Image Converter" Proc. Int. Conf. on Sensory Devices for the Blind, St. Dunstans: London (1966) pp 299-308.

17. Tanner, W.P. "The Simulation of Living Systems: On the Question of Substituting one Sense for Another" Proc. Int. Cong. on Tech. and Blindness, AFB: New York (1963) Vol 2, pp 209-220.

18. Woodward, P.M. "Probability and Information Theory with Application to Radar" Pergamon Press: Oxford (1953).

19. Cahlander, D.A. "Echolocation with Wide-band Waveforms: Bat Sonar Signals" Tech. Report 271, Lincoln Laboratory MIT (1964).

20. Grinnell, A.D. and Grinnell, V.S. "Neural Correlates of Vertical Localization by Echolocating Bats" J. Physiol. 181, 830-851, (1965).

21. Gardner, M.B. "Proximity Effect in Sound Localization", J.A.S.A. 43, 163, (1968).

22. Roffler, S.K. and Butler, R.A. "Localization of Tonal Stimuli in the Vertical Plane" J.A.S.A., 43, 6, 1260-1266, (1968).

23. Rubin, E.O., Ware, M.E. and Helson, H. "Anchor Effects in Pitch Localization" Am. J. Psychol., 79, 458-464, (1966).

24. Roffler, S.K. and Butler, R.A. "Factors that Influence the Localization of Sound in the Vertical Plane" J.A.S.A., 43, 1255-1259, (1968).

25. Batteau, W. "Localization of Sound", Pts 1-4, United Research Inc. NOTS TP3109.

26. Wallach, H. "The Role of Head Movements and Vestibular and Visual Cues in Sound Localization", J. Exp. Psychol. 27, 339-368, (1940).

27. Pollack, I. and Rose, M. "Effect of Head Movement on the Localization of Sounds in the Equatorial Plane" Perception and Psychophysics 2, 591-596, (1967).

28. Thurlow, W.R. and Runge, P.S. "Effect of Induced Head-movements on the Localization of Direction of Sounds" J.A.S.A. 42, 480-488, (1967).

29. Perrott, D.R., Elfner, L.F. and Homick, J.L. "Auditory Spatial Orientation" Perception and Psychophysics, 5, 189-192, (1969).

30. Bauer, R.W., Matuza, J.L., Blackmer, R.F. and Glucksberg, S. "Noise Localization after Unilateral Attenuation", J.A.S.A., 40, 441-444, (1966).

31. Freedman, S.J. and Zacks, J.L. "Effects of Active and Passive Functions During Prolonged Atypical Stimulation", Perceptual and Motor Skills, 18, 361-366, (1964).

32. Rice, C.E. "Quantitative Measures of Echodetection in the Blind", Progress Report, NIH Grant NB04738, Stanford Research Institute.

33. Kellog, W.N. "Sonar System of the Blind", Science, 137, 399-404 (1962).

34. Schubert, E.O. and Schultz, M.C. "Some Aspects of Binaural Signal Selection, J.A.S.A., 34, 844-849, (1962).

35. Cherry, E.C. and Bowles, J.A. "Contribution to the Study of the Cocktail Party Effect", J.A.S.A., 32, 884, (1960).

36. Gardner, M.B. "Historical Background of the Haas or Precedence Effect", J.A.S.A., 43, 1243, (1968).

37. Wallach, H., Newman, E.B. and Rosenweig, M.R. "The Precedence Effect in Sound Localization", Am. J. Psychol., 52, 315-336, (1942).

38. Fletcher, H. "Speech and Hearing in Communication" Van Nostrand: New York (1953) Ch. 10.

39. Wenner, C.H. "Frequency Dependence of Aural Difference Tones, J.A.S.A., 43, 780-784, (1968).

40. Jeffress, L.A. and Taylor, R.W. "Lateralization vs. Localization J.A.S.A., 33, 482-483, (1960).

41. Ebata, M., Sone, T and Nimura, T. "Binaural Fusion of Tone Bursts Different in Frequency", Proc. 6th Int. Cong. on Acoustics: Tokyo, (1968), Vol 1, pp A-33 - A-37.

42. Deatherage, B.H. "Binaural Interaction of Clicks of Different Frequency Content", J.A.S.A., 33, 139-145, (1966).

43. Jeffress, L.A., Blodgett, H.C. and Deatherage, B.H. "Effect of Interaural Correlation on the Precision of Centring a Noise", J.A.S.A., 34, 1122-1126, (1962).

44. Cherry, E.C. and Taylor, W.K. "Some Further Experiments upon the Recognition of Speech with one and two ears", J.A.S.A., 26, 554-559, (1954).

45. Woodworth, H.S., "Experimental Psychology", Holt: New York, (1938), p.523.

46. Feddersen, W.E., Sandel, T.T., Teas, D.C. and Jeffress, L.A. "Localization of High-Frequency Tones", J.A.S.A., 29, 988-991, (1957).

47. Firestone, F.A., "The Phase Difference and Amplitude Ratio at the Ears due to a Source of Pure Tone", J.A.S.A., 2, 260-270, (1930).

48. Mills, A.W., "Lateralization of High Frequency Tones", J.A.S.A., 32, 132-134, (1960).

49. Sandel, T.T., Teas, D.C., Feddersen, W.E. and Jeffress, L.A. "Localization of Sound from Single and Paired Sources", J.A.S.A., 27, 842-852, (1955).

50. Teas, D.C. "Localization of Acoustic Transients", J.A.S.A., 34, 1460-1465, (1962).

51. Tobias, J.V. and Zerlin, S., "Lateralization Thresholds as a Function of Stimulus Duration", J.A.S.A., 31, 1591-1594, (1954).

52. Klump, R.G. and Eady, H.R.,"Some Measurements of Interaural Time Difference Thresholds", J.A.S.A., 28, 859-860, (1956).

53. Shaxby, J.H. and Gage, F.H. "Studies in the Localization of Sound", Med. Research Council Spec. Rpt. No.166, 1-32, (1932).

54. Harris, G.G., "Binaural Interactions of Impulsive Stimuli and Pure Tones", J.A.S.A., 32, 685-692, (1960).

55. David, E.E., Guttman, N. and van Bergeijk, W.A. "On the Mechanism of Binaural Fusion", J.A.S.A., 30, 801-802, (1958).

56. Mowbray, G.H. and Gebhard, J.W., "Man's Senses as Information Channels" in "Selected Papers on Human Factors in the Design and Use of Control Systems", H.W. Sinaiko (Ed.), Dover: New York, (1961).

57. Mudd, S.A., "Experimental Evaluation of Binary Pure Tone Auditory Display", J. App. Psychol., 49, 112-121 (1965).

58. Black, W.L., "An Acoustic Pattern Presentation", Res. Bull. AFB, No. 16, 93-132, (1968).

59. Pollack, I., "The Information of Elementary Auditory Displays", J.A.S.A., 24, 745-749, (1952).

60. Garner, W.R., "An Informational Analysis of Absolute Judgement of Loudness", J. Exp. Psychol. 46, 373-380, (1953).

61. Pollack, I. and Fick, L. "Information of Elementary Multi-Dimensional Auditory Displays", J.A.S.A., 26, 155-158, (1954).

62. Miller, G.A., "The Magical Number Seven, Plus or Minus Two: Some Limits in Our Capacity for Processing Information", Psychol. Rev., 63, 81-97, (1956).

63. Nye, P.W., "Psychological Factors Limiting the Rate of Acceptance of Audio Stimuli", Proc. Int. Cong. on Technology and Blindness, AFB: New York, (1963), Vol 2, 99-109.

64. Rupf, J.A.,Jnr. "Time Expansion of Ultrasonic Echoes as a Display Method in Echolocation", Res. Bull. No. 15, AFB (1968), pp 1-34.

65. Fletcher, H., "Speech and Hearing in Communication", Van Nostrand: New York, (1953) Ch.9.

66. Hymans, A.J. and Lait, J., "Analysis of a Frequency-Modulated Continuous Wave Ranging System", Proc. I.E.E., 107 Pt.B, 365-372, (1960).

67. Kay, L. "A Comparison between Pulse and Frequency-Modulation Echo-Ranging Systems", J. Brit. I.R.E., 19, 105-113, (1959).

68. Lane, C.E. "Binaural Beats", Physical Review, 26, 401-412, (1925).

69. Kay, L. "Auditory Perception and its Relation to Blind Guidance", J. Brit. I.R.E., 24, 309-317, (1962).

70. Wever, E.G., "Development of Travelling-Wave Theories", J.A.S.A., 34, 1319-1324, (1962).

71. Von Bekesy, G. "Experiments in Hearing", McGraw-Hill: New York, (1960), Ch.11.

72. Tonndorf, J., "Time/Frequency Analysis along the Partition of Cochlear Models: A Modified Place Concept", J.A.S.A., 34, 1337-1350, (1962).

73. Flanagan, J.L., "Models for Approximating Basilar Membrane Response", Bell Syst. Tech. J., 39, 1163-1191, (1960). "Models for Approximating Basilar Membrane Response Pt.II", Bell Syst. Tech. J., 41, 959-1009, (1962).

74. Granit, R., "Receptors and Sensory Perception", Yale Univ. Press: New Haven, (1955), Ch.1.

75. Galambos, R., Schwartzkopff, J., and Rupert, A., "Micro-Electrode Study of Superior Olivary Nuclei", Am. J. Physiol., 197, 527-536, (1959).

76. Hall, J.L., "Binaural Interaction in the Superior Olivary Complex of the Cat", J.A.S.A., 37, 814-823, (1965).

77. Olson, H.F. "Elements of Acoustical Engineering", Chapman and Hall: London, (1940), Ch.2.

78. Rowell, D., "A High Resolution Spectrum Analyzer", Dept. Memo. No. 20, Dept. of Elect. Eng., University of Canterbury (1968).

79. Hunter, J.D., "The Effect of Distortions in 'Linearly' Frequency Modulated Sonar", Dept. Memo. No.42, Dept. of Elect. Eng., University of Canterbury, (1969).

80. Rao, B.S., "Resolution of Multiple Target Ambiguities in a Twin-Channel Frequency Modulated Radar", I.E.E.E. Trans. on A.E.S., Vol AES-2, 518-528, (1966).

81. Howard, I.P. and Templeton, W.B., "Human Spatial Orientation", Wiley: New York, (1966), p.148.

82. Harris, G.G., "Binarual Interaction of Impulsive Stimuli and Pure Tones", J.A.S.A., 32, 685-692, (1960).

83. Kikuchi, Y., "Objective Allocation of the Sound Image from Binaural Stimulation", J.A.S.A., 29, 124-128, (1957).

84. Von Bekesy, G., "Experiments in Hearing, McGraw-Hill: New York, (1960), p.275.

85. Sayers, B.McA., "Acoustic Image Lateralization Judgements with Binaural Tones", J.A.S.A., 36, 923-926, (1964).

86. Thurlow, W.R., and Elfner, L.F., "Pure-Tone Cross-Ear Localization Effects", J.A.S.A., 31, 1606-1608, (1959).

87. Von Bekesy, G., "Experiments in Hearing", McGraw-Hill: New York, (1960), p.273.

88. Sayers, B.McA. and Cherry, E.C., "Mechanism of Binaural Fusion in the Hearing of Speech", J.A.S.A., 29, 973-987, (1957).

89. Dimmick, K. (Ed.) "Psychoacoustics - A Selected Bibliography", AFB: New York, (1966).

90. Research Index, Vols 1-3, AFB: New York, (1967).

91. Stewart, G.W., and Hovda, O., "The Intensity Factor in Sound Localization; an Extension of Weber's Law, Psychol. Rev., 25, 242-251, (1918).

92. Stewart, G.W., "The Function of Intensity and Phase in the Localization of Pure Tones. I. Intensity", Phys. Rev. 15, 425-432, (1920).

93. Halverson, H.M., "Binaural Localization of Tones as Dependent Upon Phase and Intensity", Am. J. Psychol., 33, 178-212, (1922).

94. Trimble, O.C., "The Relative Roles of the Temporal and the Intensity Factor in Sound Localization", Am. J. Psychol., 41, 564-576, (1929).

95. Guilford, J.P., "Psychometric Methods", McGraw-Hill: New York, 2nd Ed. (1954), Ch.4.

96. Deatherage, B.H. and Hirsh, I.J., "Auditory Localization of Clicks", J.A.S.A., 31, 486-492, (1959).

97. Hanson, R.L., "Sound Localization", J.A.S.A., 31, 830(A), (1959).

98. Moushegain, G. and Jeffress, L.A., "The Role of Interaural Time and Intensity Differences in the Lateralization of Low Frequency Tones", J.A.S.A., 31, 1441-1445, (1959).

99. Trimble, O.C., "Intensity Difference and Phase Difference as Conditions of Stimulation in Binaural Sound Localization", Am. J. Psychol. 47, 264-274, (1935).

100. Whitworth, R.H. and Jeffress, L.A., "Time vs Intensity in the Localization of Tones", J.A.S.A., 33, 925-929, (1961).

101. Banister, H., "Three Experiments on the Localization of Tones", Brit. J. Psychol., 16, 265-292, (1926).

102. Hafter, E.R. and Jeffress, L.A., "Two Image Lateralization of Tones anc Clicks", J.A.S.A., 44, 563-567, (1968).

103. Von Bekesy, G. "Experiments in Hearing", McGraw-Hill: New York, (1960), p.297.

104. Halverson, H.M. "Binaural Localization of Tones as Dependent upon Phase and Intensity" Am. J. Psychol. 33, 178-212, (1922).

105. Sayers, B.McA., "Acoustic Image Lateralization Judgements with Binaural Tones", J.A.S.A., 36, 923-926, (1964).

106. Deatherage, B.H., "Examination of Binaural Interaction", J.A.S.A., 39, 232-249, (1966).

107. Johnson, N.L. and Leone, F.C., "Statistics and Experimental Design, Vol II", Wiley: New York, (1962), p.53.

108. Wine, R.L., "Statistics for Scientists and Engineers", Prentice Hall: Englewood Cliffs, (1964), p.503.

109. Rowell, D., "The Effect of Interaural Amplitude Difference on the Lateral Position of the Binaural Image for Tonal Signals", Dept. Memo No.36, Dept. of Elect. Eng., University of Canterbury, (1968).

110. Sayers, B.McA. and Lynn, P.A. "Interaural Amplitude Effects in Binaural Hearing", J.A.S.A., 44, 973-978, (1968).

111. Sivian, L.J. and White, S.D., "On Minimum Audible Sound Fields", J.A.S.A., 4, 288-321, (1933).

112. Fisz, M., "Probability Theory and Mathematical Statistics", Wiley: New York, (1963),p.394.

113. Massey, F.J.Jnr, "Distribution Table for the Deviation between two Sample Cumulative Functions", Ann. Math. Stat., 23, 435-441, (1952).

114. Whitfield, I.C., "The Auditory Pathway", Edward Arnold: London, (1967), p.1.

115. Perrott, D.R., "Role of Signal Onset in Sound Localization", J.A.S.A., 45, 436-445, (1968).

116. Stevens, S.S., "Psychophysics and the Measurement of Loudness", Proc. 6th Int. Cong. on Acoustics, Tokyo, (1968), Vol 1, pp GP63-GP70.

117. Fletcher, H., "Speech and Hearing in Communication", Van Nostrand: New York, (1953), Ch.11.

118. Stevens, S.S., "The Measurement of Loudness", J.A.S.A., 27, 815-829, (1955).

119. Scharf, B., "Binaural Loudness Summation as a Function of Bandwidth", Proc. 6th Int. Cong. on Acoustics, Tokyo, (1968), vol 1, pp A25-A29.

120. Jeffress, L.A. "Mathematical and Electrical Models of Auditory Detection", J.A.S.A., 44, 187-203,(1968).

121. Schalow, R.D., "Averaging to Reduce Variance in Bearing Estimates", J.A.S.A., 44, 379(A), (1968).

122. Zwislocki, J.J., "Temporal Summation of Loudness: An Analysis", J.A.S.A., 46, 431-441, (1969).

123. Van Bergeijk, W.A., "Variations on a Theme of Bekesy; A Model of Binaural Interaction", J.A.S.A., 34, 1431-1437, (1962).

124. Skolnik, M.I., "Introduction to Radar Systems", McGraw-Hill: New York, (1962), pp 176-178.

125. McGinn, J.W.Jnr, "Thermal Noise in Amplitude Comparison Monopulse Systems", I.E.E.E. Trans. on AES, AES-2, 550-556, (1966).

126. Manasse, R., "Maximum Angular Accuracy of Tracking a Radio Star by Lobe Comparison", I.R.E. Trans. on Ant. and Prop., AP-8, 50-56, (1960).

127. Sharenson, S., "Angle Estimation Accuracy with a Monopulse Radar in the Search Mode", I.R.E. Trans. on Aer. and Nav. Electronics, ANE-9, 175-179, (1962).

128. Cooper D.C., "Errors in Directional Measurements Using the Relative Amplitude of Signals Received by Two Aerials", Proc. I.E.E., 114, 1834-1836, (1967).

129. Hald, A., "Statistical Theory with Engineering Applications", Wiley: New York, (1952), pp 204-207.

130. Marcum, J.I., "A Statistical Theory of Target Detection by Pulsed Radar", I.R.E. Trans. on Information Theory, IT-6, 59-267, (spec. monograph issue), (1960).

131. Rice, S.O., "Mathematical Analysis of Random Noise", Bell Syst. Tech. Journal, 23 and 24. Reprinted in "Selected Papers on Noise and Stochastic Processes", N. Wax (Ed.), Dover: New York, (1954), pp 133-294.

132. Watson, G.N., "Treatise on the Theory of Bessel Functions", Cambridge University Press: Cambridge, (1922), p.153.

133. Burdic, W.S., "Radar Signal Analysis", Prentice-Hall; Englewood-Cliffs, (1968), p.268.

134. Watson, G.N., "Treatise on the Theory of Bessel Functions", Cambridge Univ. Press: Cambridge, (1922), p.395.

135. Blachman, N.M., "Noise and its Effect Upon Communication", McGraw Hill: New York, (1966), Ch.1.

136. Abramowitz, M. and Stegum, I.A., "Handbook of Mathematical Functions", Dover: New York, (1965), pp 504-515.

137. Cook, C.E. and Bernfield, M.E., "Radar signals: An Introduction to Theory and Application", Academic Press: New York, (1967), pp 18-33.

138. Wainstein, L.A. and Zubakov, V.D., "Extraction of Signals from Noise", Prentice-Hall: Englewood Cliffs, (1962), p.83.

139. Delano, R.H., "A Theory of Target Glint or Angular Scintillation in Radar Tracking", Proc. I.R.E., 41, 1778-1784, (1953).

140. Smith, R.P., "Constant Beamwidth Receiving Arrays for Broad Band Sonar Systems", Acustica (to be published).

141. Cook, C.E. and Bernfield, M.E., "Radar Signals: An Introduction to Theory and Application", Academic Press: New York, (1967), pp 137-139.

142. Thurlow, W.R. and Bernstein, S., "Simultaneous Two Tone Pitch Discrimination", J.A.S.A., 29, 515-519, (1957).

143. Plomp, R. and Steeneken, H.J.M., "Interference Between Two Simple Tones", J.A.S.A., 43, 883-884, (1968).

144. Zwicker, E., Flottorp, G., and Stevens, S.S., "Critical Band Width in Loudness Summation", J.A.S.A., 29, 548-557, (1957).

145. Goldstein, J.L. and Kiang,N.Y.S.", Neural Correlates of the Aural Combination Tone". Proc. I.E.E.E., 56, 981-992, (1968).

146. Goldstein, J.L., "Auditory Nonlinearity", J.A.S.A., 41, 676-689, (1967).

147. Guilford, J.P., "Psychometric Methods", McGraw-Hill: New York, 2nd Ed. (1954), pp 139-142.

148. Bush, R.R., "Estimation and Evaluation" in Luce R.D., Bush, R.R., Galanter, E. (Eds), "Handbook of Mathematical Psychology", Vol 1, Wiley: New York, (1963), pp 429-469.

149. Taylor, M.M., and Creelman, C.D., "PEST: Efficient Estimates on Probability Functions", J.A.S.A., 41, 782-787, (1967).

150. Pollack, I., "Interval Patterning and Auditory Gap Detection", Perception and Psychophysics, 4, 147-151, (1968).

151. Wald, A., "Sequential Analysis", Wiley: New York, (1947), pp 88-105

152. Dixon, W.J. and Mood, A.M., "A Method for Obtaining and Analyzing Sensitivity Data", J. Am. Stat. Ass., 43, 109-126, (1948).

153. Wetherill, G.B., and Levitt, H., "Sequential Estimation of Points on a Psychometric Function", Brit. J. Math. and Stat. Psychol., 18, 1-10, (1965).

154. Levitt, H., and Bock,D.E., "Sequential Programmer for Psychophysical Testing", J.A.S.A., 42, 911-913, (1967).

155. Campbell, R.A. and Lasky, E.Z., "Adaptive Threshold Procedures: BUDTIF", J.A.S.A., 44, 537-541, (1968).

156. Campbell, R.A., "Detection of a Noise Signal of a Varying Duration", J.A.S.A., 35, 1732-1737, (1963).

157. Wetherill, G.B., "Sequential Methods in Statistics", Methuen: London, (1966), pp 162-179.

158. Levitt, H., "Testing for Sequential Dependencies", J.A.S.A., 43, 65-69, (1968).

159. Uttal, W.R., "Real-Time Computers: Technique and Applications in the Psychological Sciences", Harper and Row: New York, (1967), Section II.

160. Coker, C.H., Hall, J.L., and Rosenburg, A.E., "Control of Psychophysical Experiments with an On-Line Digital Computer", Proc. 6th Int. Cong. on Acoustics: Tokyo (1968), Vol 1, pp A59-A62.

161. Pollack, I., "Effect of Smooth vs Step Interval Gradients Upon Interval Gap Detection", Perception and Psychophysics, 5, 357-358, (1969).

162. Mann, R.W., "The Evaluation and Simulation of Mobility Aids for the Blind", Res. Bull. No.11, AFB, (1965), pp 93-98.

163. Baecker, R.M., "Computer Simulation of Mobility Aids: A Feasibility Study", Res. Bull. No.16, AFB, (1968), pp 141-206.

164. Hirsh, I.J., "Bekesy's Audiometer", J.A.S.A., 34, 1333-1336, (1962).

165. Brown, R.G., "Smoothing, Forecasting and Prediction of Discrete Time Series", Prentice-Hall: Englewood Cliffs, (1963), p.101.

166. Fu, K.S., "On Learning Techniques in Engineering Cybernetic Systems", Cybernetica, 10, 194-213, (1967).

167. Fu, K.S., "Sequential Methods in Pattern Recognition and Machine Learning", Academic Press, (1968), pp 204-206, (see also ref.166).

168. Perrott, D.R., Elfner, L.F., and Homick, J.L., "Auditory Spatial Orientation", Perception and Psychophysics, 5, 189-192, (1969).

The following papers were presented or published during the course of the project:

Rowell, D. "Blind Mobility - A Problem in Cybernetics" Paper H(E)14, 40th Annual Congress of ANZAAS, University of Canterbury, January 1968.

Kay,L. and Rowell,D. "Auditory Perception and its Relation to Blind Guidance". N.Z. National Electronics Convention, Univ. of Auckland, August 1968.

Rowell,D. and Kay,L. "Auditory Perception and Blind Guidance" Proc. 6th Int. Cong. on Acoustics, Tokyo, August 1968.

The following research notes in the form of Departmental Memoranda have been lodged with the Electrical Engineering Dept, University of Canterbury.

Rowell, D. "Preliminary Report on Research completed in 1966 on the FM Blind Aid". Dept Memo No. 2, April 1967.

Rowell, D. "A High Resolution Spectrum Analyzer". Dept Memo No. 20, February 1968.

Rowell, D. "Blind Mobility - A Systems Study". Dept Memo No. 21, February 1968.

Rowell, D. "The FM Binaural Aid for the Blind". Dept Memo No. 23, March 1968.

Kay,L. and Rowell,D. "Auditory Perception and its Relation to Blind Guidance". Dept Memo No. 24, March 1968.

Rowell, D. "The Effect of Interaural Amplitude Differences upon the Lateral Position of a Binaural Image for Tonal Signals". Dept Memo No. 36, December 1968.

Rowell, D. "Notes on an Interactive Method of Measuring the Auditory Localization Constant of an Individual using Interaural Amplitude Difference". Dept Memo No. 47, June 1969.