AURAL DISCRIMINATION OF PITCH IN THE PRESENCE OF NOISE

Richard Eugene Charlton White
1970

A thesis submitted for the degree of
Doctor of Philosophy, in the Faculty of Engineering,
University of London

Department of Electrical Engineering
Imperial College of Science and Technology,
University of London,
1970.

To Sheila, without whom...

ABSTRACT

This thesis describes an investigation of the effects of additive noise on the discrimination, by human listeners, of the relative pitch of two sinusoidal sounds. The experiments were designed to test an Ideal Pitch Discriminator model, an important implication of which was that the ear could match itself to a signal in a noisy situation. However, the results were not consistent with this model, and an explanation in terms of the Critical Band effect was preferred. This effect had not previously been demonstrated in a pitch discrimination task. The Difference Limen for pitch varied with SN ratio according to a curve of characteristic shape, located on the SN axis at a position dependent on the noise bandwidth. The separation between wide and narrow band curves was used to estimate the effective bandwidth of the ear.

A lumped-constant filter model was constructed to simulate the frequency dependent behaviour of the middle ear and basilar membrane. This model could account for the frequency selective (or Critical Band) properties inferred from the subjective experiments. It was also proposed that the perception of pitch was based on the temporal properties of auditory nerve impulses initiated by the zero-crossings of oscillatory basilar membrane movement. It was argued that the generation of such impulses was analogous to the zero-crossing detection performed by an ideal FM discriminator, and it was shown that the effects of noise on such a device could account for the subjective data.

A quite separate aspect of this thesis was the evaluation of a Sequential testing and estimation method. Although a Sequential strategy appeared to cause some instability in a subject's responses, it had a number of advantages over the more commonly used Probit Analysis, particularly for the estimation of Difference Limens.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Prof. E.C. Cherry for his guidance during the course of this work, and even more for his patience and understanding during the writing of this thesis. Thanks are also due to Dr B. McA. Sayers of Imperial College, who was responsible for directing my interest to the field of auditory physiology, and for several helpful discussions. Edith Corliss, of the National Bureau of Standards, Washington, was of great assistance in questions of acoustical measurement techniques.

This work would not have been possible without the stimulation of my colleagues at Imperial College; in particular, I would like to thank Harry Levitt, Tony Marsh, Laurie Moye, Kirit Patel, Don Pearson, and Andrew Sekey. I am also grateful to all those who were patient enough to serve as experimental subjects, including many of the colleagues just mentioned, as well as Peter Bylański, Andrew Murray, Steven Kirk, and Roger Wiley.

I am grateful to Archie White and other members of the technical staff of Imperial College for their help and advice with many problems concerned with the design and construction of apparatus. Thanks are due to the University of London Computer Unit and the Sub-Department of Animal Behaviour at Madingley for the use of their computer facilities. The staffs of the Imperial College Libraries, and of the Philosophical Library, Cambridge, are to be thanked for their efficient services. I am also especially grateful to the Viscount Corvedale for his tracing of a particularly obscure publication.

I received generous financial support during the course of this work, and would like to thank the Department of Scientific and Industrial Research, the Ministry of Aviation, and the Philips Electrical Company.

I am deeply indebted to Professors R.A. Hinde and M.N. Eagle, and particularly to Dr John Morton, who read and criticised the first draft of this thesis. They suggested many additions and improvements which have subsequently been included.

Finally, it is difficult to properly thank my wife, Dr Sheila White who, apart from providing constant encouragement at all stages of the writing of this thesis, has also been responsible for many hours of stimulating discussion of its contents. She has also undertaken the prodigious secretarial task of organising and typing the manuscript.

LIST OF CONTENTS

CHAPTER II. MEASURING PITCH DISCRIMINATION.

CHAPTER III. APPARATUS.

CHAPTER IV. EXPERIMENTAL TECHNIQUE.

CHAPTER VI. MAIN EXPERIMENTS: THE EFFECTS OF NOISE ON PITCH DISCRIMINATION.

CHAPTER VIII. SUMMARY AND SUGGESTIONS FOR FURTHER STUDY.

GLOSSARY OF SYMBOLS AND ABBREVIATIONS

A - A stimulus of reference frequency which is the first of an A-X pair

B - $1/\sqrt{2\pi}$ x the slope of a CN judgement curve at its midpoint

$\hat{B}$ - ML estimate of B

BW - Bandwidth

CB - Critical Band

CBW - Critical Bandwidth

CN - Cumulative Normal

D - A response indicating that X was judged <u>lower</u> in pitch than A in a particular stimulus pair

D' - A composite event formed from U and D judgements, which is used in the construction of s Sequential strategy

d_f - Frequency difference between A and X

DF - Degrees of freedom

DL, Δ - Difference Limen, defined here as the distance between the 71 and 29 pc points on a judgement curve

IPD - Ideal Pitch Discriminator

M - the midpoint of a response curve (point of subjective equality)

$\hat{M}$ - ML estimate of M

ML - Maximum Likelihood

NS - Not significant

P - Probability

PA - Probit Analysis

pc - percent (o/o)

P(F) - Cumulative probability of F

P(U), P(U') - Probability of a U or U' type of response

Q - Ratio of the centre frequency to the half power BW of a bandpass filter

rms - root mean square

Symbols and Abbreviations, cont.

SD, σ	- Standard deviation
SL	- Sensation level, in dB relative to the threshold of audibility
S/N	- Signal to Noise power ratio (equivalent to SN ratio)
SPL	- Sound pressure level (in units of dB relative to .002 dynes/cm^2)
TTS	- Temporary threshold shift
U	- A response indicating that X was judged higher in pitch than A in a particular stimulus pair
U'	- A composite event formed from the U and D judgements which are used in the construction of a Sequential strategy
W-method	- Wetherill Method
X	- A stimulus of variable frequency whose pitch is judged relative to that of A

LIST OF TABLES

LIST OF FIGURES

CHAPTER I. INTRODUCTION

> " The sensation of sound is a species of reaction against external stimulus, peculiar to the ear, and excitable in no other organ of the body, and is completely distinct from the sensation of any other sense.
>
> " As our problem is to study the laws of the sensation of hearing, our first business will be to examine how many kinds of sensation the ear can generate, and what differences in the external means of excitement or sound, correspond to these differences of sensation."
>
> (Helmholtz, 1877, p.7.)

I.1. General Introduction and Statement of Aims

In recent years, engineers have played an increasing part in the study of hearing and speech. The rapid growth of telecommunication systems has made channel bandwidth valuable, and the engineer, who at first concerned himself merely with faithful waveform transmission, now hopes to transmit adequate signals more cheaply by exploiting the limitations of his human customers. This is not possible without a clear understanding of the relevant perceptual aspects of the signals transmitted. The methods used by the communications engineer for the analysis of systems and of their effects on signals have proved valuable in the advancement of the study of hearing. With the extension of Fourier analysis to handle aperiodic and random signals, and the development of Statistical Communication Theory, the engineer now has techniques for describing and analysing signals which are more typical of natural conditions than the sinusoids so frequently used as test stimuli in the past. This thesis describes a study of the perception of pitch from an engineering point of view.

Under natural conditions, the ear is faced with two problems: it must discriminate, and it must select. By discrimination is meant the ability to recognize changes in the physical parameters of a sound stimulus; by selection is meant the ability to separate a wanted sound from an interfering one. The ear is extremely flexible in its use of physical differences between sounds to achieve these ends. For example, the physical variables of spatial position, intensity, frequency, and time characteristics, may be involved in any way. This thesis is concerned with the variable of frequency and how it may be used to achieve both discrimination and separation.

Generally speaking, sensations of pitch can be ordered, from high to low, on a scale which is closely related to the physical variable of frequency, and two sounds of sufficiently different frequency content can be discriminated on the basis of a pitch difference. It is also found that the effect of an interfering sound is decreased if its frequency characteristics are sufficiently different from those of a wanted sound. Previously, these two aspects have been studied separately, and though it has been suggested that there is a close relation between them, this has never been supported by direct experimental evidence. The present work tests the relation between discrimination and selection directly, by studying the perception of pitch in a masking situation.

The problem was considered initially in the framework of a mathematical model which had been developed to apply to the discrimination of pitch in noise, but which had not previously been tested under the appropriate conditions. A closer examination suggested that the perception of pitch might better be accounted for by a model of the peripheral ear which was derived from simple physiological considerations, and viewed the perception of pitch as a dual time-frequency process.

There are five relevant bodies of research which will be reviewed here: first, the work on the perception of pitch which, although closely connected with frequency analysis, has also been shown to be related to the temporal aspects of a sound stimulus; second, those experiments on frequency selectivity which concentrate primarily on signal detection in the presence of noise; third, a brief introduction to Ideal Observer Theory, which forms the basis of the model to be tested; fourth, a review of work on the discrimination of pitch in the presence of noise. Finally, a discussion of the physiology of the ear, which is necessary for the understanding of the model proposed, is reserved for presentation in a more appropriate context, in Chapter VII.

There is a quite separate aspect to the work in this thesis which is concerned with the evaluation of a method of measuring performance in a discrimination task. Some of the theoretical considerations involved are discussed in Chapters II and IV, and an experimental investigation forms the material of Chapter V.

I.2. Theories of Pitch Perception

Although there are some effects which support the classical view of the ear as a frequency analyser, there are others which suggest that the ear may operate in either a temporal, or in a dual time-frequency mode. The following sections contain a selection of those experiments which have been cited as evidence for one view or another. The selection is by no means exhaustive, and the aim has been to present some of the main ideas with a few relevant experiments in each case.

I.2.1. Frequency Analysis

I.2.1.1. Introduction

The view of the ear as a frequency analyser is not new. It was first proposed by Ohm in 1843, with reference to the work of Fourier 20 years earlier, and the theory received strong support from Helmholtz (Helmholtz, 1877; Wever, 1949; Boring, 1950). An essential feature of a frequency analysis theory is some 'place' principle; a place principle states that there is some area in the nervous system where frequencies are distinguished on the basis of differences in spatial characteristics of the activity which they initiate. However, empirical demonstration had to wait until the work of von Bekesy (1947, 1949a, 1949b), who observed the oscillations of the human basilar membrane which were induced by sinusoidal stimuli. He found that the position of the maximally displaced point of the membrane was a function of frequency (see Chap. VII). However, von Bekesy also noted that the position of this maximum was much less sensitive to small frequency changes than would be consistent with the fineness of their discrimination measured psychophysically.

Various attempts have been made to reconcile this difference by looking for ways in which the basic mechanical selectivity of the ear might be refined (Helmholtz, 1877; Huggins & Licklider, 1951; Whitfield, 1957; von Bekesy, 1960). However, because of the difficulties in performing the appropriate physiological experiments, the problem is still unsolved, and suggestions must, for the moment, remain speculative.

I.2.1.2. Direct Evidence for a Frequency Analysis

This section presents some psychophysical results which have been felt to support a frequency-place mode of operation by the ear. On the simplest level, the correlation between the frequency of a pure tone and its pitch (Stevens & Davis, 1938) suggests that the ear is a frequency analyser. More direct evidence comes from experiments using harmonic sounds. If the ear is performing a Fourier analysis, it should be possible for a listener to distinguish between the components of a harmonic sound. Plomp (1964) found that over the fundamental range from 44-2000 Hz, the first 5-8 partials of a complex tone could be distinguished. Higher partials were not separable, implying that frequency analysis may take place, but it is of limited resolution. (See also Plomp & Mimpen, 1968.)

Whether or not harmonics are distinguished, they may not always be useful in discrimination. Flanagan & Saslow (1958) found that the smallest detectable frequency change in the fundamental frequency of a complex sound was slightly less than that for a sinusoid of the same frequency. This implied some use of harmonic information. On the other hand, Henning & Grossberg (1968) did not take this view, and suggested that discrimination of such changes was based on the most discriminable harmonic only. Schodder & David (1960) found evidence of two modes of operation, depending on whether the stimulus bandwidth was large or small (see Section I.2.3.3.). These studies suggest that although the ear performs a frequency analysis, it does not show perfect frequency resolution. The characteristics of the analysing 'filters' appear to affect not only which components of a complex sound can be perceived separately, but also the way in which the energy in the signal can be used effectively.

Other evidence comes from the fact that random sounds may have a pitch which is associated with features of their spectra. For example, filtered noise may have a pitch related to its centre frequency (Michaels, 1957) or to its cut-off frequency (Small & Daniloff, 1967; von Bekesy, 1963b). Both these effects suggest a mechanism which can detect salient features of the frequency spectrum.

I.2.1.3. Indirect Evidence for a Frequency Analysis

Several other effects, although they do not give direct evidence, can be explained readily in terms of a frequency analytic model. According to a place principle, any sound should give rise to a spatial pattern of neural activity whose central location determines the pitch of that sound, and as a corollary, any change in the locus of activity should be interpreted as a change in pitch. Although the most important determinant of the place of activity should be the frequency components of the sound, it is possible that other variables might have a second order effect, and might thus affect its pitch. The examples given here include variations of intensity, auditory fatigue resulting from intense stimulation, and the way in which pitch is affected by the addition of noise.

1) The Effects of Intensity

Stevens & Davis (1938) quote examples of the change in pitch of a tuning fork as it varies in distance from the ear. This effect can be explained as follows: if pitch is identified by the place of origin of those neural channels which respond maximally to a stimulating tone, then at high intensities these channels may saturate, and neural activity may grow more rapidly on either side of the maximum. Unless this growth is symmetrical, this will result in an effective shift of the active area, and would be interpreted as a change in pitch.

Stevens (1935) presented two alternating tones of different frequencies to a listener who was required to adjust their relative intensity until they sounded equal in pitch. He found that while the pitch of pure tones in the 500-2000 Hz range was relatively stable, for higher frequencies it rose, and for lower frequencies it fell, as intensity was increased. However, measurements on larger numbers of subjects have shown significant individual differences in the extent of this effect (Cohen, 1961; Small & Campbell, 1961a). Further, the effects of intensity are not always significant when compared with the constant errors involved in a pitch matching situation. The main conclusion to be drawn at present, seems to be that there may be an effect of intensity on the perceived pitch of a pure tone, but that its magnitude is small compared with other factors.

2) The After-effects of Intense Stimulation

According to a place principle, prolonged intense sounds would be likely to fatigue those neural channels which were most strongly stimulated, causing a temporary rise in their response thresholds. Thus, the effective stimulation pattern of a given sound might well differ pre- and post-fatigue, due to the absence of activity in the fatigued neurons. Several effects may be distinguished which follow prolonged or intense stimulation, and which are often related to the frequency characteristics of the stimulus. These include tinnitus, an after-sensation which may possess pitchlike qualities, TTS (Temporary Threshold Shift), and certain pitch-shift effects.

One example of this is shown by the results of Atherley et al (1968), who found that maximum TTS occurred at a frequency which was a constant amount (1.5 - 2 KHz) above the most intense frequency of a fatiguing noise stimulus; they also described an accompanying tinnitus which was somewhere between the two in frequency. (See also the results of Small & Yelen, 1962, discussed on p.41.) Another effect which has been cited as evidence for a frequency analytic view, is diplacusis which is a pitch-shift observed when the same pure tone is presented alternately to the two ears (Licklider, 1958). Ward (1963) exposed one ear of each of his listeners to loud periodic pulses, and observed changes in diplacusis relative to control measurements at all frequencies above 100 Hz. However, Skinner & Antinoro (1968) employing a tracking method which allowed the onset and offset of fatigue to be studied, failed to find significant diplacusis effects after intense stimulation with a pure tone. As mentioned above, it seems that the choice of a test technique may have a considerable influence on the results of this type of experiment.

3) <u>The Effects of Noise</u>

It has been reported that the pitch of a pure tone may change with the addition of noise. This can occur either with a narrow band noise whose spectrum is offset relative to the signal frequency, or with a wideband noise. In either case, a place principle would argue that the location of the combined stimulus pattern of signal-plus-noise could be effectively different from that of signal alone, and might be responsible for the change in pitch which is observed.

Schubert (1950), using a binaurally presented sinusoid, reported interaural pitch differences of nearly a semitone when wideband masking noise was injected to one earphone. (Control measurements showed that the effect was not explicable in terms of any change in loudness due to the threshold shift caused by the noise.) Egan & Meyer (1950), using a narrow band noise signal, found that the pitch of a sinusoid tended to move in a direction away from the noise spectrum; the pitch of a higher frequency signal moving upwards, while that of a lower frequency signal moved downwards. Similar effects were found by Webster & Schubert (1954), but here, upward shifts were more marked than downward ones. Similarly, Webster et al (1952) studied the change in pitch of a pure tone which was placed in an octave gap in the noise spectrum, and found significant upward shifts for signals in the lower part of the gap, with smaller downward shifts for signals in the upper part.

Webster & Schubert (1954) assumed that noise had a masking effect which suppressed signal activity in those channels which were most sensitive to the noise spectrum. This would cause a movement of the effective signal pattern away from the noise spectrum. They also suggested that the stimulation pattern of a pure tone was asymmetrical in frequency due to the frequency response of the basilar membrane, which is known to have a more gradual slope in the low frequency direction (von Bekesy, 1949b). This would account for the greater effect of noise spectra below the signal frequency.

More recently, the significance of the change in pitch in the presence of noise was questioned by Allanson & Schenkel (1965). They used band-limited noise added to the signal, and found that although the pitch tended to move away from the noise band, the magnitude of the effect was not much greater than the normal constant errors of judgement found even in the absence of noise (cf. Cohen, 1961). Further, they found very large inter-subject differences (23 subjects were used), indicating that comparison of the results of different studies using only a few subjects may be of doubtful value.

To complicate the question further, von Bekesy (1963b) found an exactly opposite effect, where the pitch of a pure tone changed in a direction towards the noise spectrum. (He also demonstrated this effect using a model basilar membrane applied to the skin of the forearm.) He suggested that signal and noise were combined into a single unified pattern of activity. It would follow from this that the addition of a narrow band noise would cause the centre of the pattern to move towards the noise frequency (i.e., the pitch of a pure tone would move in that direction). These questions have not yet been resolved.

I.2.1.4. Evaluation

This section presented some general types of experiment, the results of which have been used as evidence for a classical frequency-place theory of auditory analysis. By more modern standards they are unconvincing, since the results have not always proved easy to replicate using different test techniques and more subjects. In particular, more recent work has shown that experiments involving interaural pitch comparisons may be unreliable. This is not to say that frequency analysis does not exist, but that the strict frequency-place principle can be criticised because it is too restrictive. Other work, which will be presented in the following sections, has suggested that hearing involves more flexible mechanisms which are able to use both time and frequency information.

I.2.2. Temporal Analysis

I.2.2.1. Introduction

Evidence of the inadequacy of a pure frequency-place theory of pitch perception was available at the time of Ohm and Helmholtz, but was ignored for nearly 100 years. Seebeck, in 1841 (see Schouten, 1940a), using a siren, had shown that the pitch of some sounds corresponded to a frequency which was, in fact, missing from the Fourier spectrum. The view of the ear as a Fourier analyser was so compelling however, that even when Fletcher (1924) also found the missing fundamental effect, the most likely explanation was felt to be that the fundamental component was reintroduced by some nonlinear distortion process prior to analysis. (See also Stevens & Davis, 1938, who alluded to a similar explanation.) However, more recent work has shown that this effect cannot be explained so simply.

Schouten (1940a) performed experiments analogous to Seebeck's, using an optical siren which allowed the precise control of the amount of the fundamental component in a sound. Even when this frequency was cancelled completely, the pitch of the complex sound remained unchanged at a value corresponding to the missing component. This work presented the first serious challenge to strict frequencies theories of auditory analysis in that Schouten showed beyond doubt that a pitch sensation could be experienced in the absence of any Fourier component corresponding to that pitch. He noted that the waveforms of his stimuli showed a periodicity corresponding to the perceived pitch, and suggested that the sensation might arise from a temporal, rather than a spectral analysis of the stimulus. Schouten's theory (1940b, 1940c) included a preliminary frequency analysis, and was thus, in principle, a dual time-frequency model; this will be discussed in section I.3. However, some pitch effects can be explained in terms of a temporal analysis only, and these will be described here.

I.2.2.2. The Experimental Evidence

A striking finding was that interrupted random noise could have a pitch corresponding to its interruption rate (Miller & Taylor, 1948). (See also Harris, 1963, who found it possible to perform pitch matches between interrupted noise and a periodic pulse train for frequencies up to at least 750 Hz.) Miller & Taylor's interpretation was that their results proved a temporally based form of analysis, since the power spectrum of the noise was uniform and would be unchanged by periodic interruption. Cramer & Licklider (1957) refined this explanation, and distinguished between two aspects of the pitch of this type of signal, one of which was based on its spectral properties, and the other of which was related to its temporal features.

Time delay may also be a source of pitch sensation. Thurlow (1958) showed that a pitch sensation could be associated with the time delay between two pulse trains (see also Small & McClellan, 1963). A time delay pitch was also observable with random waveforms and with random repetition of pairs of pulses (McClellan & Small, 1966, 1967). A pitch may also be experienced with continuous time-delayed signals. Bassett & Eastmond (1964) showed that wide or narrow band noise reflected from a flat surface, could be heard with a pitch which varied inversely with the distance of the observer; pitch values from 200 to 2000 Hz were reported. They suggested that the pitch sensation was due to an aural difference tone introduced by the nonlinearity in the ear. However, McClellan & Small (1966) pointed out that there was also a temporal relation in Bassett & Eastmond's signals, in that the original and a similar version delayed by the reflection path, were both present in the sound field of the listener. They suggested autocorrelation as a possible means of extracting a temporally based pitch sensation.

It had been suggested earlier that the temporal analysis used by the ear may be of an autocorrelational type (e.g., Licklider, 1951, 1959; Sayers & Cherry, 1957). (This was a natural consequence of work on the extraction of radar signals from noise (e.g., Lee, 1950), and was also related to a proposal by Jeffress (1948) of a neural mechanism responsive to interaural time delay.) A paper by Fourcin (1965) describes the possible sources of a pitch sensation in a random stimulus. Using clipped noise into which correlation was introduced by adding the noise to a delayed version of itself, he was able to manipulate the location and spacing of the regular peaks which were present in the stimulus spectrum. Using a method of pitch matching to a sinusoid, he found that the pitch sensations could not be explained on the basis of a frequency analysis by the ear, and suggested an autocorrelational process as a more likely alternative (see also Bilsen, 1966, 1967).

Cramer & Huggins (1958) provide a classic example of pitch resulting from binaural interaction, which implies the existence of a central process capable of integrating information from the two auditory nerves (see Licklider, 1959). They presented white noise from the same generator through earphones, with the input to one ear channel passed through a narrow band phase-shift network. A pitch was heard which corresponded to the frequency of the phase shift, even though the phase-shifted signal itself had no pitch quality (see also Fourcin, 1959). Other experiments suggest that this integration is not always perfect. For example, if a signal of reference frequency is presented to one ear, with the unknown in the other, the CDL (Contralateral Difference Limen, a measure of the minimum detectable frequency difference between two binaural signals) may be greater than for similar signals presented monaurally (Small & Brandt, 1963; Webster, 1969).

Apart from any consideration of the pitch of sounds, it is known that temporal information must be available in some situations from the extremely high sensitivity of the ear to changes in interaural time delay which are associated with changes in the spatial position of a sound source (see Levitt, 1964, and references cited therein). Further evidence for the transmission of temporal information comes from the finding that two otherwise unrelated signals, when presented binaurally, may give rise to a single fused image if they possess temporal features in common (see Broadbent & Ladefoged, 1957; Leakey _et al_, 1958; David _et al_, 1959; Harris, 1963).

I.2.2.3. _Evaluation_

The findings presented here were not intended to deny the existence of a frequency analysis, but were grouped together because a temporal mechanism provided a sufficient explanation. It is suggested that they be considered together with the frequency analysis results of the previous section as investigations of two aspects of the behaviour of the ear which, jointly, provide support for a dual mechanism which will be discussed in the following section.

In any case, since the temporal and frequency domains are related in an exact manner by the Fourier transformation, any absolute choice between them is difficult. When a choice has been implied, it has usually been with the aim of finding the most parsimonious explanation of a set of experimental facts. It should be remembered that the concept of parsimony may be heavily influenced by the ideas of Fourier analysis, which is by no means the only possible method of signal representation, and which may not be the closest to that used by the ear. (See Gabor, 1946, and Tondorff, 1962, for some alternative possibilities.)

I.2.3. Dual Process in Pitch Perception

I.2.3.1. Introduction

As mentioned earlier (p. 35), Schouten (1940b, 1940c) did not deny the existence of a preliminary frequency analysis, but made a distinction between two different components of a theory of hearing which had not been recognised previously. The first is the process of initial analysis, which may be responsive to the frequency characteristics of a sound; the second is concerned with how the results of the preliminary analysis are transmitted by the auditory nerve to the brain. According to a strict frequency-place theory, the only information transmitted was the position of the activity on the basilar membrane. Schouten argued that the auditory nerve might be capable of transmitting more than mere place information, and that a pitch sensation might be associated with the temporal properties of neural activity. The implication of this is that the perception of pitch may be a dual process, operating on either the place of origin of individual nerve fibres, or on the temporal features of the information which they carry. (See also Wever, 1949.)

In the following two subsections, some experiments will be presented which support a dual mechanism of pitch perception. The first is devoted to experiments which show that the pitch of a stimulus is not always related to its frequency characteristics. The second contains some experiments which use stimuli containing conflicting time and frequency information, and which show that judgements associated with one mode or the other can be obtained. The reader is also referred to the papers of Huggins & Licklider (1951), Licklider (1959), Tonndorff (1962), von Bekesy (1963a), and Plomp (1967) for other discussions of time and frequency concepts.

I.2.3.2. Characteristics of Residue Signals; Masking and Fatigue

Schouten (1940a) noted that the pitch sensation corresponding to the fundamental frequency of a harmonic sound could be resolved into two components: one which was aroused by the fundamental frequency, if present, and the other was a 'residue' which was sharp in quality and was related to the periodicity of a combination of several high harmonics. Further studies of the residue phenomenon have given more insight into the nature of the temporal analysis which is involved. The effect is a low frequency one, and the residue pitch is most marked for periodicities below 800 Hz (Ritsma, 1962). The temporal analysis appears to preserve the features of both the fine structure and the envelope structure of the stimulus (Schouten et al, 1962; Ritsma, 1962; Ritsma & Engel, 1964; see also Thurlow & Small, 1955).[†]

When signals possessing a residue pitch are used in masking or fatigue situations, it is found that although their pitch characteristics are related to their temporal properties, the effects of masking

---

[†] There is an anomalous effect, first found by deBoer (cited in Schouten et al, 1962), which is observed with an amplitude modulated carrier wave. DeBoer noted that when the modulating frequency was raised slightly, raising the envelope frequency, the pitch actually dropped. It has subsequently been suggested (Schroeder, 1966; Fischler, 1967; Fischler & Cern, 1968) that the frequency dependent behaviour of the basilar membrane may account for this secondary pitch-shift which is difficult to explain on the basis of any purely temporal analysis of the unmodified stimulus waveform.

and fatigue depend on their spectral properties. Licklider (1959, p. 118) used noise to selectively mask the frequency regions above or below 1 KHz. Two types of signal were tested: a low frequency sinusoid and a highpass filtered pulse train, both of which gave rise to the <u>same</u> low pitch sensation. While the low frequency noise masked the sinusoid completely, the pulse train continued to be audible with a low pitch. Conversely, high frequency noise masked only the pulse train. A similar effect was shown by Small & Campbell (1961b) using bands of noise to mask a pulsed sinusoidal signal (2.2 KHz) which had a pronounced residue pitch at 150 Hz. They found that only noises with significant energy in the 2.2 KHz region were effective. Small & Yelen (1962) used the same signal as a fatiguing stimulus, and found a rise in the hearing threshold for signals in the region of 2 KHz, while thresholds for 150 Hz signals were not significantly affected (see p. 31).

These results are difficult to explain on the basis of any single process theory. Rather, they imply a two-stage process involving a preliminary filtering followed by temporal analysis of the filtered stimulus. Presumably, the initial frequency analysis would involve a 'place' mechanism which was selective on the basis of the spectral properties of the stimulus. (This stage would account for the filtering of noise frequency components which were markedly different from that of the signal, and would also be prone to desensitization by over-stimulation.) Temporal analysis would be performed on the result of this filtering operation, and would give rise to the residue pitch. If a pitch sensation were derived from a temporal analysis, it would follow (as Schouten had argued) that different neural channels might carry similar pitch information.

I.2.3.3. Perception of Signals with Conflicting Information

The experiments cited above suggested a dual analysis process of which the two stages appeared to have different functions; the frequency analysis being involved in the effects of noise and fatigue, while the temporal analysis was used to derive a pitch sensation. There are other experiments which implicate both types of process in the perception of pitch itself. These involve the use of stimuli containing temporal and frequency information which present conflicting pitch cues; it has been found that judgements may be associated with either mode.

Schodder & David (1960), by changing the frequency of only the lower of a pair of sinusoids, were able to produce a situation in which conflicting information was present, since the frequency of the envelope of the complex changed in the opposite direction to that of the lower component. For closely spaced sinusoids, the subjects' judgements were correlated with the periodicity of the stimulus envelope, suggesting that they were able to integrate the two frequencies into a single percept and base their judgement on its temporal properties. For wider spacings, judgements were based on changes in the lower component only, suggesting that the two components were perceived separately.

Flanagan & Guttman (1960a; see also 1960b, and Guttman & Flanagan, 1964) used periodic pulse trains to study the temporal and spectral modes of pitch perception. By inverting the polarity of some pulses, it was possible to change the relationship between fundamental frequency and pulse rate. At low pulse rates, any two such trains were matched for pitch on the basis of the number of pulses per second, regardless of polarity or fundamental frequency; at frequencies above 100-200 Hz, the trains were matched for fundamental frequency. It was possible to explain the effects by reference to a model of the mechanical behaviour of the basilar membrane. At different points along the model membrane, displacement functions could be found which might be expected to initiate nerve impulses at either an impulse rate or a fundamental frequency rate; the relative magnitudes of these two types of disturbance changed with impulse frequency in a way which paralleled the psychophysical data.

Rosenberg (1965) studied the effect of masking noise on Flanagan & Guttman's stimuli (see p. 42), and found that he could control the type of judgement elicited. With highpass filtered noise, pitch matches were predominantly to the fundamental frequency of the reference stimulus, while lowpass noise caused a switch of attention to the pulse period. The cut-off frequencies of the low- and high-pass filters were both 1 KHz, suggesting that temporal cues arose from auditory channels which were normally sensitive to frequencies above this value, and that frequency information was simultaneously present in the lower frequency channels.

A transitional frequency of 1 KHz is also found in other situations. For example, in the localization of sounds, time differences between the two ears appear to be useful only up to about 1.5 KHz (David <u>et al</u>, 1959; Mills, 1960). Webster (1969) found two distinct types of behaviour in his listeners, and suggested that this might be due to his use of a test frequency (1 KHz) which was in the region of transition between different modes of analysis. He found that his subjects could be divided into two groups; for one of these, the CDL (p. 37) was approximately twice that found monaurally, while for the other, it was somewhat less.

I.2.3.4. Summary

The experiments described in this section suggest that the ear is capable of using both time and frequency features of a stimulus in arriving at a pitch sensation. Certain of the results suggest that this is a quite flexible process, in that for some stimuli the attention of a listener can be directed to one aspect or the other. Thus, under natural conditions, the ear may be adaptive in that it will select one mode or another, depending upon the information available. Results have been presented which are difficult to explain on any other than a multi-stage basis, and suggest that pitch cues of more than one type may be simultaneously present in the perception of a sound.

I.3. Auditory Frequency Selectivity; the Critical Bandwidth

I.3.1. Introduction

Until now, frequency analysis has been discussed only in relation to discrimination. Another area in which the ear shows frequency dependent behaviour is in the separation of wanted from unwanted sounds, and this will be called frequency selectivity. This section is concerned with auditory frequency analysis from the frequency selectivity viewpoint.

Various experiments suggest the existence of a critical frequency interval such that those frequency components of a stimulus which fall within this interval are treated in one way, and those falling outside it are treated in another. This has been demonstrated in areas as diverse as loudness of complex sounds (Zwicker et al, 1957), masking (Fletcher, 1940), monaural phase sensitivity (Zwicker, 1952), auditory frequency analysis (Plomp, 1964), and musical consonance (Plomp & Levelt, 1965). This critical frequency, the CBW (Critical Bandwidth), suggests some basic frequency selective process in hearing.

The classic demonstration of the CB (Critical Band) effect was by Fletcher (1940), who showed that not all the spectral components of of a wideband noise were equally effective in masking a pure tone; he interpreted this as due to the operation of some frequency selective process. As the masking aspect of the CBW is the most relevant to the experiments described in this thesis, other aspects will not be mentioned here.[+] More detailed discussions of different determinations of the CBW are given by Green (1958) and Scharf (1961, 1966).

I.3.2. The CBW in Masking Experiments

Fletcher (1940) presented tonal signals of different frequencies in a background of continuous masking noise, and asked subjects to adjust the signal to a just-detectable level. Noises of bandwidths from 30-8000 Hz were used, centred on the signal frequency in each case. He found that for a given power density, the masking effect of a noise was independent of its bandwidth as long as it was larger than a certain critical value, the CBW. Fletcher suggested that for each frequency there exists a sharply defined region on the basilar membrane, and that only that part of the noise spectrum which stimulated

---

[+] Two terms will be used here: CB (or Critical Band) and CBW (Critical Bandwidth). The first refers to the frequency response characteristic of the mechanism responsible for frequency selective behaviour, and the second describes the effective frequency range which is included in one CB.

this region was effective in masking the tone. Variation of noise bandwidth outside the critical region could have no effect on the detectability of the sinusoid, since this region was always fully stimulated. Once the noise bandwidth fell below the critical value however, the effective masking signal would fall in direct proportion to the reduction of noise bandwidth. In other words, the CB effect could be described as the operation of an 'aural filter' which was tuned to the signal frequency.

Fletcher made two assumptions which allowed him to estimate the CBW: the first was equivalent to stating that the CB was rectangular in form, and the second was that a pure tone would be detectable when its intensity was just equal to that of the noise within a CB. From these assumptions, it follows that if the threshold ratio (R_t) of signal energy to noise power density is plotted as a function of noise bandwidth, the results should lie along one of two straight lines. For narrow band noise, where the SN ratio is unaffected by the CB 'filtering', a line of slope 1 (or equivalently, 3 dB/octave) passing through the origin should be followed (i.e., R_t = noise bandwidth). For wideband noise, points should lie on a horizontal line, since variation of noise bandwidth would have no effect on its masking power (i.e., R_t = CBW). The intersection of these two lines would occur at the point where the noise bandwidth was just equal to the CBW. Using this rationale, Fletcher made estimates of the CBW (Fletcher, 1940; or see Green, 1958, p. 8).

According to Fletcher's assumptions, measurement of the R_t for wideband noise should give the CB directly (the Critical Ratio Method). Measurements were reported by Hawkins & Stevens (1950) and Bilger & Hirsh (1956), and are in close agreement with Fletcher's. However, the validity of the two assumptions (rectangular CB's and perfect integration of noise power within a CB) has been questioned in more recent work, and the results of these studies are presented in the following two sections.

I.3.2.1. <u>The Shape of the CB</u>

The rectangular CB assumption was first questioned by Schafer <u>et al</u> (1950). They used band-limited masking noises with nearly rectangular spectra, since these should show most clearly the discontinuity predicted by Fletcher. Their results were more consistent with a gradual transition of slope in the region of the CBW than with the sharp change expected on Fletcher's hypothesis. Schafer <u>et al</u> suggested that this might be due to a non-rectangular CB, and showed that a curve based on the expected variation in output power of a single tuned filter operating on variable bandwidth noise was a better description of their data.

The non-rectangularity of the CB was also shown by Webster <u>et al</u> (1952), who performed the converse masking experiment using wideband noise with a spectral gap in the region of the signal frequency. Masking was greatest for signals whose frequencies lay outside the range of the gap, and least for signals placed centrally within it. However, it was found that the increase of masking as the signal frequency moved away from the centre of the gap was less rapid than expected from a strict interpretation of a rectangular CB; the data were fitted better by CB's shaped like resonance curves. Using this method, numerical estimates of the CBW were obtained by deBoer & Bos (1962).

Greenwood (1961) also used a masking situation to measure the

CBW, but made no assumptions, either of shape or of the integration of noise power. He used noises of a wide range of bandwidths, and in each case measured a complete masked audiogram (the threshold intensity of a sinusoidal signal for a wide range of noise band centre frequencies). He found that the audiograms fell into two categories, depending on the noise bandwidth. For narrow band noises, the variation of threshold intensity with frequency was triangular in form, with the peak occurring when the signal was approximately centred in the noise band. As the noise band was widened, the audiogram acquired a flat top and became trapezoidal. His explanation was that the flat top would appear at the noise bandwidth which was just equal to the CBW, beyond which the effective masking power of the noise was limited by the CB filtering action. Thus, he defined the CBW as that noise bandwidth at which the masked audiogram just began to change from a triangular to a trapezoidal form. His estimates were wider by a factor of 2-3 than values based on the Critical Ratio Method, and agreed with other determinations based on loudness, threshold measurements, phase sensitivity, and tonal masking (see Scharf, 1966).

Swets <u>et al</u> (1962) made estimates of the CBW based on several different shapes of the CB, and found that the values varied by a factor of about 2, depending on the specific CB characteristic assumed. They pointed out that if the CB did not represent an ideal rectangular filtering operation, then it was likely that the effect would depend on the particular test stimuli selected. They suggested that failure to consider the true shape of the CB might account for some of the discrepancies between individual estimates of the CBW.

I.3.2.2. Integration of Noise Power

Fletcher assumed that noise power at the output of the CB filter is the only determinant of its masking effect, and that masking increases at 3 dB/octave of bandwidth for subcritical bands of noise. It follows from this that for subcritical bands of noise, changes of bandwidth and power density can be traded exactly, provided that total power remains constant. This apparently is not true. It is found that although masking changes with power density in the expected way, variation of noise bandwidth has a smaller effect than predicted by Fletcher's assumption, and data lie closer to a slope of 1.5 dB/octave (Hamilton, 1957; Swets et al, 1962; van den Brink, 1964; see also deBoer, 1962, for a compilation of results).

This departure may be because the intensity fluctuations of narrow band noise make discrimination of the signal more difficult, and tend to offset any gain in detectability due to a lower noise level (Bos & deBoer, 1966). Another possibility is that a pitch quality is associated with band-limited noise, which becomes more distinct as the bandwidth is narrowed. Michaels (1957) showed that the pitch DL (Difference Limen, or minimum detectable frequency change) for narrow band noise decreases as its bandwidth is reduced, indicating an increased distinctiveness in its pitch. Possibly the difficulty of detecting a sinusoidal signal increases as the noise becomes more distinct in pitch, and thus sounds more like the signal. This, too, would tend to offset the increase in detectability due to reduction of the noise power.

Greenwood's results (1961) do not agree with the above. He found that the total power within a CB was an adequate predictor of the masking effect of a noise. As long as noises of subcritical width were used, the peak point of the resulting triangular (see p. 48) normally increased in height in direct proportion to increases in noise power due to variation of either bandwidth or power density.[+] Thus, his results support the ideal integration of noise power suggested by Fletcher. Note however, that Greenwood's test stimuli were different from the narrow band random noises normally used, in that they were generated by a modulation process. This point will be discussed later, in Chapter VII.

I.3.3. Evaluation

This section has described the CB phenomenon and some of the assumptions underlying its estimation in a masking situation. Although the methods of estimation used have often been based on somewhat arbitrary assumptions, there is still general agreement amongst different studies. Numerical estimates of the CBW are remarkably consistent in view of the number of different situations in which it has been studied; at 1 KHz, estimates have ranged from 40 Hz to 200 Hz (deBoer & Bos, 1962). There is also agreement on the way in which the estimates vary with frequency, remaining roughly constant below 1 KHz and increasing with frequency above this value. Although the CB concept has been demonstrated in so many different areas, there is as yet, no unified theoretical framework to explain why the CBW should apply so universally.

I.3.4. The Relation between the CBW and Pitch Discrimination

One theoretical approach particularly relevant to this thesis was

---

[+] This statement was true only for a constant signal frequency, since Greenwood did find that the threshold SN ratio for a tone in noise of about CB width, tended to decrease with frequency from about -3.5 dB at 450 Hz to -8 dB at 3250 Hz.

an attempt to explain the discrimination of frequency changes in terms of the CB concept. Briefly, the view was taken that the frequency selectivity implied by the CB effect in masking, implies the existence of some internal auditory 'filter'. If the tuning of this filter were fixed, then changes in the frequency of a sinusoid at the input could be discriminated on the basis of changes in the amplitude of the filter output (Schafer et al, 1950; Corliss, 1967).

Schafer et al noted that the DL for pitch varies with frequency in a similar way to the CB, though at a level which is perhaps 20 times lower (see Shower & Biddulph, 1931, for DL data, and Green, 1958 p. 24, for a comparison of the DL and CB data). Schafer et al suggested that this could arise if pitch discrimination was based on the transformation of frequency changes into energy changes by the CB filter, as outlined above. (Thus, a just-discriminable frequency change should be that value which leads to a just-discriminable energy change in the filter output.) Using their estimates of the CBW, Schafer et al calculated the energy changes which would be associated with the DL's measured by Shower & Biddulph.[+] They compared these values with independent estimates of the intensity DL for a sinusoidal signal (Riesz, 1928) and found a close agreement. Their conclusions have been confirmed by Corliss (1967) using the same theoretical framework to intrepret a large amount of data assimilated from a number of different sources.

[+] CBW's were estimated from masking data in a detection experiment by fitting curves derived from an analysis of the behaviour of a single tuned filter in the presence of variable bandwidth noise.

Apart from the fact that a frequency-oriented model such as this cannot account for the temporal effects described in sections I.2.2. and I.2.3., there are two major objections to this approach. First, no attempt was made to show that the CB does indeed exist in a discrimination task; CBW estimates were all taken from detection experiments which might call for the use of a quite different strategy by the ear. Second, both Schafer et al and Corliss used CBW estimates which were narrow by modern standards (roughly 20 Hz at 1 KHz); Greenwood's values, for example, are nearly ten times this figure. In sum, the attempts to explain pitch discrimination in terms of a CB filtering process have led to a model which is rather vague, and whose evaluation must be indirect. However, the idea is nonetheless interesting, and one reason for the experimental situation used in this thesis (the discrimination of pitch in the presence of noise), was to test directly whether a CB concept could be applied to a discrimination situation.

There is also a different viewpoint, developed by Sekey (1962, 1963), which specifically predicts the effects of random noise on pitch discrimination. The essential feature of his model was also a filtering process which had the dual functions of separation and discrimination and was, in this respect, similar to the model above. The difference lay in the approach used to derive the model (the Theory of Ideal Observers), and this is described in the next section.

I.4. Ideal Observer Theory; Its Implication in Auditory Theory

I.4.1. Introduction

The theory of Ideal Observers is important in that it separates the problem of processing a signal from that of making a decision about it, and has been found to be a powerful way of looking for the basic invariants in many different types of sensory task. The basic theory can be found in the papers of Peterson et al (1954), Marill (1956), Tanner (1960), Tanner et al (1960), and Green (1960). An original bibliography was complied by Creelman (1960), and more recently, Levitt (in press) has produced a review of research within the Ideal Observer framework.

Given certain assumptions of the properties of a signal corrupted by random noise, it is possible to place mathematical bounds on the precision with which some task involving that signal can be performed. In each situation, application of the theory leads to the specification of an ideal, or optimal scheme for performing the task in question, as well as an indication of the minimum error rate attainable by an 'ideal' observer. The behaviour of the ideal observer can then form the basis for assessing the performance of a human observer given a similar task, and it has been found that in some situations, the discrepancy is not very great (e.g., Marill, 1956). However, the significance of the theory in the context of this thesis does not lie in the comparison of relative efficiencies, but rather in its implications for the mode of operation of the auditory nervous system. It can be shown that the optimum processing operation for a signal corrupted by additive guassian noise involves a filtering process in which the spectral characteristic of the receiver is matched to that of the signal (see e.g., Woodward, 1953). This implies that the ear, too, is able to adjust its parameters to match those of the signal in question.

I.4.2. Sekey's Ideal Pitch Discriminator (IPD)

Essentially similar models of human pitch discrimination, based on Ideal Observer Theory, have been proposed by Cardozo (1962) and

Sekey (1963). Here, only Sekey's model will be discussed in detail, as this is the model tested in this thesis. Sekey's derivation will not be repeated here; it is very clearly presented in his paper or alternatively, reference might be made to Woodward's work (1950, 1953) on the estimation of the time delay of radar signals. (Sekey's and Woodward's results are similar in many ways if the variables of time and frequency are interchanged.)

Sekey considered an observer in a frequency discrimination situation, who was forced to judge whether the second (X) of two brief sounds was higher or lower than the first (A). The test sounds were two gated segments of pure tone for which the frequency of X was different from that of A by a controlled amount. Sekey described the situation from a spectral viewpoint, since the major effect of the frequency difference was to shift the spectrum of the second sound bodily on the frequency axis. He then derived an optimal strategy for the detection of the spectral shift of X relative to A. This model will be referred to here as the IPD (Ideal Pitch Discriminator).

Since the signals are assumed to be corrupted by noise, exact measurement is impossible, and the optimal operation performed by the IPD model is that of forming the posterior probability distribution of the spectral shift. A decision is then made of whether the most probable value of the shift (F_{Max}) is positive or negative.[ǂ] Due to the noise, F_{Max} is itself a random variable, which Sekey showed is approximately normally distributed about a mean value corresponding

---

[ǂ] For the symmetrical forced choice situation used here, the decision as to whether F_{Max} is positive or negative is equivalent to an indication of whether X is higher or lower than A.

to the true frequency shift (F_o). Since F_{Max} is a random variable, it is possible for decision errors to be made since there is a finite probability that F_{Max} will have the opposite sign to F_o.

Sekey showed that the probability of a decision error, P(e), was given by:

$$P(e) = \text{erf}\left[(R/2)^{1/2}.a.F_o\right] \tag{1}$$

where: $R/2$ = the ratio of signal energy to noise power density

a = a measure of the effective duration of the signal

F_o = the frequency shift between the first and second signal

$$\text{erf}(X) = 1/\sqrt{2\pi}\int_{-\infty}^{x}\exp(-y^2/2)dy$$

I.4.2.1. The Predicted Effect of Noise Level

Eq. 1 shows how performance depends on the parameters of the signal and on the SN ratio. For a given error rate, either frequency shift (F_o) or duration (a) can be traded for S/N according to a square root law; a change by a factor of two in frequency shift compensating for a change in noise power density by a factor of four. Put another way, if a minimum detectable frequency shift is defined according to some fixed value of error probability (P(e)), then this value should change by a factor of 2 for each 6 dB change in R.

I.4.2.2. The Predicted Effects of Bandwidth and Relation to the CB

The second important feature of Sekey's model is that it implies that the ear is able to adjust its parameters to match a signal. The essential stage in the formation of the posterior distribution (and hence, finding F_{Max}) is the computation of the cross-correlation between the A and the X signals. To achieve this, it would be necessary for the signals to be stored for sufficient time for the correlation to be formed. Alternatively, a matched filtering technique could be used; this is a method of finding the cross-correlation between two signals by passing one signal through a filter whose impulse response is defined by the waveform of the other. These two procedures are mathematically identical, and one or the other may be selected in practice, purely on the grounds of convenience. The important point is the implication that the receiver is able to adapt itself to a situation, either by the use of a perfect memory for the storage of reference signals, or by adjusting the parameters of some filter to exactly match the signal waveform. Indeed, there is evidence from experiments on signal detection in noise that the ear may adjust itself to match the signal over some range (Hamilton, 1957; Green <u>et al</u>, 1957; Green, 1957, 1958; Creelman, 1961; van den Brink, 1964).

The optimal nature of matched filtering can also be understood from a SN ratio viewpoint: the matching means that the filter bandwidth is equal to that of the signal, and thus those parts of the noise spectrum which do not contain signal energy will be excluded. This is a common sense approach to minimising the effect of the noise without sacrificing a significant amount of the signal energy. Provided that the noise spectrum is uniform, at least over the range occupied by the signal, its bandwidth will then be unimportant -since any additional frequency components are removed by the filtering-. (This explains why the bandwidth of the noise does not appear in the expression for P(e) in Eq. 1.)[+] This bandwidth-independence leads to a prediction

---

[+] This point is discussed more fully by Lawson & Uhlenbeck (1950) and Green <u>et al</u> (1957); see also Harmon (1963) for a simple treatment.

which is similar to the CB effect: the error rate for the discrimination of the pitch of a signal in additive noise should be the same for all noise bandwidths greater than the signal bandwidth, provided that noise power density is kept constant.

I.4.2.3. Experimental Verification

To conclude the discussion of Sekey's model, his experimental verification should be mentioned briefly. He measured the error rate in pitch judgements for tone pulse pairs as a function of the duration of the signals and of the frequency shift between reference (A) and unknown (X). The signals were essentially noise-free, and to account for the imperfect performance of his subjects, internal noise was postulated. (Since it was found that the error rate did not vary significantly with signal level, the power density of the internal noise was presumably proportional to signal energy.) From a comparison of the results with the relation described by Eq. 1, Sekey concluded that his matched filter model was a better predictor of his data than a similar device which was not able to vary its bandwidth. (For this argument, see Sekey, 1962.)

I.4.2.4. Evaluation

Quite apart from the assumption that the ear is capable of performing the precise filtering and integrating operations required, Sekey's own evaluation of his model was incomplete. He used only noise-free signals, and assumed the existence of internal noise of the appropriate type. Thus, his model was not really tested under the conditions for which it was derived: to explain the discrimination of signals corrupted by additive guassian noise. A more satisfactory test situation would be one where noise was added to the signal externally, to allow the proper control of its characteristics.

The main experiments in this thesis (Chapter VI) measured the discrimination of pitch as a function of noise bandwidth and power density. Apart from directly testing Sekey's model, this situation is analogous to Fletcher's original CB experiment, and therefore provides a direct test of whether a CB could indeed be demonstrated for discrimination.

The concluding section of this chapter contains a brief review of previous experiments on the effect of noise on the discrimination of pitch.

I.5. The Discrimination of Pitch in the Presence of Noise

One effect of noise which was discussed earlier (Sec. I.2.1.3.) is the simple pitch-shift effect where a signal with added noise may consistently be judged different in pitch from a similar noise-free version. This section is concerned with experiments in which the resolution of frequency differences between two signals (i.e., discrimination) is influenced by the addition of noise to them both. There have been few studies in this area, and even these have not tested a very wide range of conditions. In view of the small amount of data and the differing experimental situations, no consistent picture has yet emerged of the way in which masking noise affects the discrimination of pitch.

It has generally been found that the discrimination of frequency changes is relatively resistant to the addition of masking noise; the DL increases significantly only at noise levels which are sufficient to almost completely mask the signal. Harris (1947, 1948a, 1966) found that the DL for pitch was relatively unaffected unless quite low

SN ratios were used; discrimination was possible for signal levels as low as 3-5 dB above the detection threshold for the same signal in noise. This suggests a parallel with the discrimination of noise-free signals, where the DL increases as a signal approaches its threshold of audibility (Shower & Biddulph, 1931). Possibly, the effect of noise is simply to raise the threshold of audibility, causing the DL to increase in an analogous way to the noise-free situation. Harris (1947) found that this was not the case, since a given signal increment caused a greater reduction in the DL for the masked signal than for the noise-free signal at a corresponding level above threshold. On the other hand, Brandt & Small (1963) found that the effects were equivalent, and they could explain their data by using the argument presented above (i.e., in terms of the equivalent level of signal above threshold).

A final experimental approach which is worth noting, compared the performance of an Ideal Observer with DL's measured for two human listeners as a function of signal duration and SN ratio (Cardozo, 1962). It was found that the trends were somewhat similar, although the subjects were different by an order of magnitude. The interesting aspect of this approach is that since the Ideal Observer is able to match itself to a received signal (see p. 55), the parallel with perceptual data suggests that humans may also behave adaptively in a discrimination situation.

It appeared from a review of the literature on the discrimination of pitch in noise, that no study had been carried out which was sufficiently comprehensive to allow the proper evaluation of either of the theoretical approaches of interest here (the relation between the CB and discrimination, and the application of an Ideal Observer model, the IPD). It was therefore felt that a detailed investigation in this area might be a useful contribution to a rather sparse body of data.

I.6. Concluding Remarks

This chapter was concerned with two different aspects of auditory frequency analysis: the discrimination of pitch, which was shown to involve both frequency and temporal principles, and frequency selectivity, as implied by the CBW and signal detection experiments. Two theoretical approaches were described, those of Schafer et al and Sekey; both of these represent attempts to treat the two aspects of frequency analysis in a unified way.

Schafer et al assert that the same filter is responsible for all types of auditory frequency analysis, and hence that its operation can be studied in either a masking or a discrimination situation. Sekey made essentially the same claim, but with the addition of a matched filter property which follows from the behaviour of an Ideal Observer. However, neither of these models has been tested satisfactorily. Schafer et al for example, assumed that the CBW was fixed, and applied estimates based on masking data to a discrimination task. Sekey, on the other hand, did not really test his model under the conditions for which it was designed, in that he used noise-free signals and assumed that the 'internal noise' of a listener was of the appropriate type.

Both these models can be tested more directly by observing the effect of variable bandwidth noise on pitch discrimination. Sekey's model made two definite predictions which could immediately be tested in such an experimental situation: 1) the relation between noise bandwidth, signal bandwidth, and the pitch DL and 2) the square root trading relation between frequency shift and SN ratio. Bringing together the two aspects of auditory frequency analysis into a single experimental situation might serve to clarify any relationship between them. For example, most determinations of the CBW in masking situations have used fixed frequency signals; it might perhaps be expected that when signal frequency is uncertain, as in a pitch discrimination task, that wider CBW's would be observed.

The experiments in this thesis are felt to illustrate that the study of hearing should be made from as broad a viewpoint as possible, and that abstract modelling techniques, although extremely valuable, should make use of known physiological data. Such an integrated approach may not only point the way to more parsimonious models of hearing, but should hasten progress towards the time when the communications engineer, the physiologist, and the psychologist can begin to meet on common ground.

I.7. General Organisation of the Thesis

The Preliminary experiments in this thesis, presented in Chapter V, concern methods of measuring pitch discrimination, and are introduced separately in the next chapter (II) and in the last section of the chapter concerned with experimental techniques (IV). Chapter III contains a description of the apparatus used in all experiments.

The Main experiments, in Chapter VI, were designed to investigate whether Sekey's IPD model could predict the effects of variable bandwidth noise on the discrimination of pitch, and to study the application of a CB concept in pitch discrimination.

The results of the Main experiments led to the proposal of a simple model of the peripheral ear, which accounted for the data better than the IPD model. This is presented in Chapter VII together with a review of the relevant work on auditory physiology. A summary and some suggestions for future research are presented in Chapter VIII.

CHAPTER II. MEASURING PITCH DISCRIMINATION

Introduction This chapter concerns the problem of measuring pitch discrimination, and discusses some possible approaches. In particular, the test methods adopted by the writer will be justified and described in detail.

In the context of this thesis, pitch discrimination will be defined operationally: if a subject gives overt responses which are correlated with changes in the frequency of a sound stimulus, then he is said to be discriminating between the corresponding pitch sensations. One common measure of discrimination is the DL (Difference Limen) for pitch, which can be defined as that frequency change which is just detectable. Another measure which is applicable in some cases is the Midpoint (M), which is the frequency difference between two stimuli at which they are judged equal in pitch. (This is sometimes called 'the point of subjective equality'.)

The problem then, is to set up a useful measure of discrimination. This may depend on: 1) the kinds of judgement a subject is asked to make, 2) the way in which the stimuli are presented, and 3) any statistical methods which are used to form estimates from his response data. The methods used will be introduced along these lines.

II.1. Response Data from Forced Choice Judgements

The main experiments in this thesis were planned to test the predictions of Sekey's model mentioned in Chapter I. For consistency with this model, a symmetrical forced choice judgement situation was used. In this, a listener's sensitivity is measured by providing two stimuli for comparison: a reference (A) and an unknown (X), whose frequency differs by an amount d_f, and whose relative pitch is to be judged. For each A-X pair, the frequency of X is equally likely to be higher or lower than A, and the subject is forced to indicate the direction of the pitch difference.

A result which could be obtained from a typical pitch judgement experiment involving a binary forced choice is shown in fig. 1. The horizontal axis represents values of the frequency shift (d_f), which varies from a large negative to a large positive value. The vertical axis is the proportion of trials on which X was judged higher in pitch than A. The subject changes from consistently indicating that X is lower (response 'D') to consistently indicating that X is higher than A (response 'U'). An ogival curve can be used to describe this tendency (see fig. 1). The sharpness of the changeover indicates the sensitivity of the listener to changes in frequency. This sharpness can be specified either by the slope of the curve, or in terms of the DL, which is the distance on the frequency axis between two defined percentage points.

II.2. Presentation of Stimuli

Two methods of presenting stimuli were considered: the method of Constant Stimuli, and the method of Constant Stimulus Differences (Guilford, 1954). In the first, Constant Stimuli, a set of d_f values is chosen in advance and a prearranged number of presentations are made using those values in an approximately random order; methods of this type will be called 'Random'. The advantage of a Random method is that the subject cannot produce misleading results by following any consistent strategy. It has the further advantage that the subject can vary his concentration, as the d_f values usually include both difficult and easy judgements in the U and D directions. However, it may exagerate certain effects due to the subject's inability to judge stimulus pairs in isolation. For example, the probability of a U-response to a small positive change is likely to be reduced if the previous stimulus was a large positive change. The extent of this effect is difficult to predict, as 'large' and 'small' must be interpreted in relation to the listener's DL, which is not known when the actual values of d_f are being chosen.

In the second method, Constant Stimulus Differences, the stimuli are presented in <u>sequences</u> within which the d_f value is changed unidirectionally in small steps; sequences usually follow increasing and decreasing directions alternately. Methods of this type will be called 'Sequential'. The advantages of a Sequential method are: first, that very different stimuli never appear close together and second, by making the change of direction of the sequences dependent on the subject's responses, a test can be made equally difficult for each subject. One drawback is that, since the stimuli are not randomly presented, it becomes necessary to disguise their sequence from the subject (Cornsweet, 1962). Further, it is possible that Sequential methods, by setting up a feedback situation in which each stimulus presentation depends on previous responses, might be more sensitive to a subject's inability to judge each A-X pair independently (Blackwell, 1952).

In addition to the subjective considerations mentioned above, the choice of a method of stimulus presentation would influence the way in which the DL was estimated from the response data. This meant that it was difficult to make a final choice without evaluating each method under working conditions.

II.3. Estimates of Performance

In setting up a definition of a DL, two considerations are involved: the form of the response curve and the parameters which specify that curve. Here, changeover curves of the CN (Cumulative Normal) type were assumed. There were several reasons for this. First, there are probably many random processes whose effects combine to give the variability which leads to an ogival curve and so, by the central limit theorem, it is not surprising that the resultant should have normal properties. Second, data gathered and analysed by the writer which assumed this hypothesis, gave self-consistent results. Third, as only the central part of the changeover region was investigated (and differences between the normal and other ogives would tend to show up more in the tails), the choice of the normal curve was probably not critical. Once a mathematical description for the ogive has been selected, the problem becomes that of estimating the parameters of the response curve from forced choice binary judgements. Some appropriate methods will now be discussed.

II.3.1. Probit Analysis

PA (Probit Analysis) is an optimal method based on the principle of ML (Maximum Likelihood) for estimating the parameters of a CN curve (Finney, 1952). An ML approach is attractive, particularly when large quantities of data are available, because it then yields normally distributed estimates which are efficient (i.e., of minimum variance).

The data required are a series of d_f values, the total number of presentations made at each, and the corresponding number of U-responses. The analysis derives two estimates: $\hat{M}$, which is an estimate of the d_f value at which 50 pc of the U-responses occur and $\hat{B}$, which is $1/\sqrt{2\pi}$ x the slope of the fitted curve at $\hat{M}$. $\hat{M}$ is normally taken as an estimate of the midpoint, or point of subjective equality. $\hat{B}$ is a measure of acuity of pitch discriminations, large slopes implying a high resolution for frequency differences, while small slopes imply poor resolution. Alternatively, a DL can be defined as the SD (Standard Deviation) of the CN curve ($\hat{\sigma}$) which = $1/\hat{B}$.[+]

The variance properties of the estimates are affected by the d_f values selected. This was studied for the large sample case by evaluating the expected SD's of $\hat{M}$ and $\hat{B}$ in the equations given by Finney (1952). The evaluation led to the selection of a design using five values of d_f which were symmetrical about zero, and equally spaced by approximately 1 DL. The chief advantages of this design were that it gave reasonably efficient estimates of $\hat{B}$ and $\hat{M}$, while not being too dependent on prior knowledge. Secondly, the five points allowed a χ^2 test for goodness of fit to be made. The details of the basis for this choice are given in Appendix 1.

Theoretically, PA had a number of drawbacks as a technique for use with subjective data. First, for efficient estimation, the values of d_f would have to be chosen with some prior knowledge of $\hat{M}$ and $\hat{B}$. In many experimental situations, there is inadequate prior knowledge of parameter values and indeed, these values might vary during the course of an experiment (Levitt, 1964). (PA estimates average values and is not, strictly speaking, appropriate in these cases.) Second,

---

[+] The invariant property of ML estimation means that $\hat{\sigma}$ and $\hat{B}$ need not be estimated separately. The use of $\hat{\sigma}$ is equivalent to defining the DL as the distance between the 69 and 31 pc points on the CN curve.

the theory of PA assumes large samples; as the number of judgements available from a subjective test might typically be 50 to 100, it was not clear whether this was large enough. The properties of ML estimation when used with small samples of data are not so well known, but it has been found that estimates may be biased and considerably more variable than suggested by the asymptotic formulae (Wetherill, 1963). A final point is that the amount of computation involved in fitting a probit curve is considerable; there might be a delay of some days before the results of an experiment were available.

II.3.2. Sequential Methods

One drawback of PA is its dependence on the proper choice of d_f values. The use of a Sequential strategy offered a possible solution to this problem. The essential feature of all Sequential strategies is that each presentation depends in some way on the previous responses given by a subject, thus allowing the course of a test (i.e., the choice of d_f values) to be matched to the subject's performance. As will be seen however, the value of such methods does not merely lie in the efficient placing of observations for a subsequent probit analysis. Sequential rules can be designed to concentrate observations in some chosen region of a subject's response curve, and it is possible to estimate the parameters of the curve directly from the clustering properties of the data. The methods to be described here require that d_f be changed in small steps, rather than at random, and was therefore classed as an example of the method of Constant Stimulus Differences (see p. 64).

II.3.2.1. Dixon and Mood's Up and Down Method

Dixon and Mood (1948) proposed a method for estimating, from quantal responses, the mean and SD of a normal population. This technique can be regarded as the prototype of the more general method which was eventually adopted.

Initially, a set of equally spaced values of d_f is defined, and a single presentation is then made at one of the chosen values.[†] If the response is 'U', the next presentation is made at the next lower value of d_f; if the response is 'D', the subsequent presentation is made at the next higher value. The value of d_f for the third presentation is chosen according to the second response, and so on. Eventually, a pattern of responses appears such as that shown in fig. 2. If the current response in a U-D strategy is made at a level of d_f above the 50 pc point (M), it is more probable that the next presentation will be at a lower level. Similarly, for levels below M, upward movement is more likely. Thus, a set of observations will tend to be grouped in the region of M.

Dixon and Mood proposed two simple ML estimators which make use of the clustering property of the data. The first is a measure of central tendency (which is an estimate of M), and the second is a measure of the spread of the data about this centre (an estimate of σ, or 1/B). These estimators have the advantage of requiring much less computation than the fitting of a probit curve.[††]

[†] Dixon and Mood showed that subsequent analysis is simpler if the spacing interval is less than 2σ (or $2/\hat{B}$).

[††] Since a large proportion of the observations will be made at d_f's close to M, the data are well placed for a subsequent estimate of M by PA, if this is desired. However, at the same time, the data are not so well placed for estimating B. (see Appendix 1.)

II.3.2.2. Wetherill's Sequential Method

The properties of the Up and Down Method were derived from large sample theory, and it has been found that when small samples are used, their behaviour is not always as expected. Wetherill (1963) tested a number of sequential rules, and found that while estimations of M were generally satisfactory and agreed with large sample theory, the estimates of B could be biased and much more variable than expected. He concluded that satisfactory DL estimates could not be obtained from sequential data by existing techniques and suggested a new approach. (See also Wetherill & Levitt, 1965.)

The basis of Wetherill's revision is that DL's are estimated indirectly from the difference on the frequency axis between two percentage points. The points themselves are obtained by transforming the original response curve in such a way that a U-D rule can be used to estimate percentage points other than the 50 pc point, or M. The technique operates in an analogous way to Dixon and Mood's method, but with U replaced by U' and D replaced by D'. The events U' and D' are formed from the real responses (U and D) in a way which depends on the percentage points to be estimated.

Some examples of U' and D' events are given in table 1. These have been generated by deciding on a maximum number of consecutive presentations to be made at each level of d_f. U' is then defined as a sequence of this number of U-responses. The occurrence of a D-response before this maximum number, defines an event of the D' type. The effect of using these rules is to transform the original response curve into one whose midpoint is where $P_{U'} = .50$. (N.B., independent responses are implicitly assumed.) This is shown in fig. 3, where the vertical projection from the point where each curve crosses 50 pc (i.e., $P_{U'} = .50$) will show, on the original parent curve (i.e., $U = U'$), the percentage point estimated. Rules for complementary percentage points (e.g., 71 pc and 29 pc) are found by substituting D' for U', and D for U.

By analogy with the U-D rule, a measure of central tendency, such as Dixon and Mood's, will estimate the value of d_f for which $P_{U'} = .50$. However, an alternative estimator was proposed by Wetherill, which is simpler to compute and has better small sample properties. As a U-D sequence is almost completely specified by its turning points (marked 1, 2, ...etc. in fig. 2), Wetherill used the arithmetic mean of the d_f values of an even number of turning points to estimate the required percentage point. In other words, if the series of d_f values between and including two consecutive turning points is called a <u>run</u>, the Wetherill Estimator is an average of the midpoints of alternate runs; each midpoint can be considered as a separate estimate of the percentage point in question. If estimates of two complementary percentage points are obtained, then M can be estimated by the average of these (assuming an anti-symmetrical response curve), and the DL can be defined in terms of their difference.

For several reasons, entry 2 of table 1 was selected. <u>First</u>, the transformations cause non-linearity in the resultant curves (see fig. 3), which becomes more marked for higher percentage points; entry 2 shows the least amount of bias. <u>Second</u>, fewer responses are involved in generating the patterns of entry 2, which means that the strategy is quicker to recover from occasional 'wild' responses, and also is quicker to track any drift. <u>Third</u>, the distance between the 71 and 29 pc points corresponds to about 1.1 σ; this is convenient, as σ is a useful working definition of the DL. Fourth, if two complementary strategies are used to place data at the 71 and 29 pc points, both M and B can subsequently be estimated fairly efficiently by a probit analysis, if desired.

The recording and estimation procedures used are described in Chapter IV (section IV.3.; see also fig. 14).

II.4. Conclusions

In this chapter, two possible methods of selecting stimulus pairs for presentation to a subject have been described: the Random method of Constant Stimuli and Wetherill's Sequential method, which is an example of a method of Constant Stimulus Differences. Also, two methods have been discussed which estimate the mean and SD of a CN curve: Probit Analysis, which is normally used with data obtained by the Random method, and the Wetherill Estimator, which is used with data obtained by the Sequential Method.

As an estimation method, PA was an obvious choice since it is the appropriate optimal method, given the assumptions made about the response curve (p. 66). One of its drawbacks was that prior knowledge was desirable for the selection of d_f values, although the use of a Sequential technique might overcome this. Secondly, the amount of data available might not always be large enough for the optimal properties to be achieved. Finally, the amount of computation involved was considerable, and prevented immediate assessment of an experiment. Wetherill estimation appeared to have a number of advantages. First, it could give an extremely rapid estimation of the DL and M. Second, it could never give estimates which fall outside the range of d_f values tested. (This is not true of PA, where it is possible to get extreme estimates, particularly with small amounts of data.)

From theoretical considerations alone, it was not possible to select one method in preference to the other. The properties of the Wetherill Method have been studied less extensively than those of Probit Analysis, and it was not known how they compared in an experimental situation. Some experiments are presented in Chapter V, in which the methods described above were tested under working conditions so that a suitable choice could be made for use in the Main experiments.

CHAPTER III. APPARATUS

Introduction In most of the experiments to be described, a listener was required to make judgements about the relative pitch of a pair of signals which he heard in earphones. The stimuli were generated by adding a controlled amount of filtered white noise to a sinusoidal signal. It is the purpose of this chapter to indicate the main features of the experimental apparatus.

The apparatus will be described in four sections. The first section deals with most of the electronic components; many of these are quite standard and need only a brief mention. The second section is concerned with the methods used to generate narrow bandwidth noise. Thirdly, a source of random numbers was found useful in many experimental situations. Since methods such as card shuffling were too slow, an electronic random number generator was built, and is described in section III.3. Finally, in many experiments it was necessary to estimate the effective power ratio of signal to noise (S/N) which was influenced by the properties of the earphones worn by the subject. The earphones and the methods used for their testing and calibration are described in section III.4.

III.1. Electronic Equipment

Fig. 4 shows a block diagram of the main components of the experimental apparatus. The apparatus was designed to produce A-X signal pairs of adjustable duration, time separation, and frequency difference. The generation of these was controlled by a timer, which allowed adjustments in 1 mS steps.

III.1.1. The Timer

A 1 KHz clock pulse generator (see fig. 4) was connected to an eleven stage binary divider (÷ 2048). An automatic reset circuit would stop the divider whenever 2048 pulses had been counted; counting would start again when the 'start' key was pressed. Thus, by repeatedly pressing the key, the divider could be made to cycle once through all of its different states. Five 11-input AND gates were connected to the divider, and by varying the method of connection, each gate could be made to produce a pulse at any one of the 2048 instants (each instant corresponding to a particular state of the divider). Since the clock frequency was 1 KHz, the divider would change its state every 1 mS, allowing the time of occurrence of each gate pulse to be varied in steps of 1 mS. Four of the output pulses were used to control the two signal gates, one 'on' and one 'off' pulse being required for each.

As shown in fig. 4, the input to the divider could be derived either from the signal oscillator, or from an independent time source (the 1 KHz clock). Thus, the AND gate outputs would either be synchronized with a sinusoidal input to the signal gates, or not. The common input connection was used for 1 KHz test signals; for signals of other frequencies, signal and timing were separately derived in order to maintain a constant 1 mS time unit. This meant that at 1 KHz, the segments of tone which formed the A and the X signals were always switched in the same phase (at a positive going zero-crossing). For the other frequencies, switching was in random phase, because of the independence of signal and time source. (The use of a separate time source also meant that the duration of the signals remained constant, independent of frequency shift.) In practice it was found that either method gave equivalent results (see discussion in Chapter VI, section VI.5.3.1.).

The fifth pulse was used to set a bistable circuit ('frequency control') which was reset at the end of the divider cycle. DC restorers were connected to the bistable outputs so that each restored output was normally zero. When the bistable was set by a pulse from gate 5, one restored output took up a positive and the other an equal negative value. One of the two outputs was selected by a manual switch, and was connected through an attenuator to the frequency modulator of the oscillator. Thus, by correctly setting the switch and the attenuator, the oscillator frequency could be changed upwards or downwards by a selected amount.

A diagram of the trigger pulse sequence which was normally used appears in fig. 5. Since the output of gate 5 appeared just after the end of the first signal pulse (A), the second signal pulse (X) had a frequency different from A by an amount which depended on the settings of the selector switch and attenuator. As fig. 5 shows, each stimulus pair consisted of a 256 mS tone pulse of reference frequency (A) followed, after a half second delay, by a second pulse of similar length, but of different frequency (X). The DL results potentially depend on both the duration and the time separation of these pulses, but their effect was not investigated here. However, the reasons for the particular choice will be outlined.

Time separation The time interval between the A-X pulses was chosen to be similar to that used by Cardozo (1962) and Sekey (1963). Konig (1957) found some variation in the DL over the range 0.3 - 2.4 sec., but this was less than 15 pc. Chistovich (1960) found little difference between separations of about 170 mS and greater. Thus, the choice of a half second delay was probably not critical.

Signal Duration A number of studies have shown that the pitch DL varies inversely with signal duration at short durations, becoming constant at long durations. Estimates of the 'breakpoint' have varied between 50 mS (Cardozo, 1962) and about 200 mS (Liang & Chistovich, 1960). (These differences may be due to the use of different test frequencies, since the effect of changing duration may be a function of frequency. Another consideration was that since it is likely that any BW (Bandwidth) adjustment mechanism operates over a limited range only, signal duration should not be extremely long. Sekey (1963) studied signal durations of 32 mS and below, and may not have examined the full range of BW's to which adjustment was possible. Cardozo (1962) found that 64 mS was probably the limit of adjustment in pitch discrimination. However, there is evidence from detection experiments that adjustment may range from about 20 mS to about 300 mS, depending on the subject (e.g., Green et al, 1957). In any case, it is generally agreed that the pitch DL does not change much with duration above 250 mS, and 256 mS was chosen so that signal duration should not have too strong an effect on the DL. The possible implications of this choice will be discussed again in Chapter VI.

III.1.2. The Signal Gates

Since stimuli were presented in A-X pulse pairs, it was important that each pulse be switched on or off in a transient-free manner, and that there was no extraneous sound between the pair which might give additional cues to the subject. To achieve this, signal switching was performed by transistor operated balanced gates.

Each gate (see fig. 4 and circuit diagram in fig. 6) consisted of a pair of linear amplifiers driving a common output resistor. The amplifiers were controlled by a bistable circuit such that only one was active at any time. Thus, a signal applied to one amplifier was either transmitted to the output or not, depending on the state of the bistable circuit. The balanced design enabled the mean level of the output to be kept constant whether or not the signal was present; this minimized any transients. Audible transients due to imperfections in the gate were about 80 dB below the normal signal level and the on-to-off transmission ratio was 88 dB for frequencies below 20 KHz. The frequency response of each linear amplifier was uniform over the audio-frequency range. (N.B. The signal waveforms in fig. 6 are not test stimuli, but were chosen at a sufficiently high frequency (1.3 MHz) to show the limitations of the circuit.)

III.1.3. The Amplifiers

The same basic circuit was used for all amplifiers in fig. 4, and are detailed in fig. 7. Most were of unity gain (as shown), but other values were obtained simply by variation of the feedback resistor. The 600 Ω input impedance allowed direct matching to standard attenuators and filters; any number of independent signals could be added by connecting sufficient input resistors. The output impedance of about 0.5 - 1 Ω meant that normal loading effects were negligible, and was also useful in providing electrical damping for the earphones. The frequency response was flat over the audio-frequency range, and the maximum output level into a 10 Ω earphone corresponded to 105 dB SPL.

III.1.4. Other Equipment

<u>Oscillator</u> A Bruel and Kjaer BFO (type 1012) was used, and was

modified to allow external control of its frequency by the method described above.

Attenuators Solartron 600 ohm attenuators (type AT201) were used. The smallest steps were 1 dB, with an accuracy of ± 0.1 dB.

Meter A Solartron Model JM 1067, true rms voltmeter was used for most signal and noise level measurements. Its accuracy over the audio-frequency range was about ± 0.2 dB.

Lowpass Filter A Mullard filter (type GFF 001/01) was set to 4.5 KHz. Its attenuation was more than 50 dB at all frequencies greater than about 5.3 KHz.[†]

III.2. Generation of Variable Bandwidth Noise

Two different methods were used to generate narrow band noise. In the first, wideband noise was applied to a bandpass filter; in the second, a high frequency carrier was amplitude-modulated by lowpass filtered noise, producing a band of noise centred upon the carrier frequency. These two methods are described below.

III.2.1. Filter Method

Wideband (20 KHz) gaussian noise was filtered by an LCR circuit with adjustable positive feedback. By varying the gain of the feedback loop, bandwidths (3 dB) from about 5 to 1500 Hz could be obtained at a

---

[†] The frequency response of the whole apparatus is shown in fig. 8. Two curves are plotted: one with, and one without, the lowpass filter.

centre frequency of 1 KHz. Tape recordings (EMI RE301) of the filter output were made for a range of BW's from 10 to 1280 Hz.

Tape recordings were used because the parameters of the filter were neither very easy to change nor very stable at narrow BW's; thus, the circuit was too cumbersome to use in a subjective test. An advantage of the method was the simplicity of the apparatus required, but there were also several drawbacks. Fluctuations of the speed of the tape recorder caused the noise spectrum to move on the frequency axis; the peak-to-peak variation was 0.6 pc. Secondly, the non-uniformity of the tape meant that there were continuous amplitude variations of about ± 1 dB and occasional 'drop outs' were found where the level fell by 6 - 12 dB. Finally, the time taken to make and check a set of good recordings was considerable. The need for a tape recorder could have been obviated by building a more complex filter which was stable and quick to adjust, but this was not pursued.

III.2.2. Noise Generation by Modulation

A block diagram of the noise modulator is shown in fig. 9. Wide-band noise was passed through a lowpass filter, yielding a random signal containing frequencies up to F_L. This signal was applied to the input of a signal gate which was switched repetitively by a square wave signal of frequency F_C from the carrier oscillator. Thus, the gate operated as a balanced modulator to generate a series of random amplitude sinusoids of frequencies F_C, $3F_C$, $5F_C$,......etc. The second lowpass filter was set at about half an octave above the carrier frequency, and thus a single modulated sinusoid of frequency F_C appeared at the filter output. This signal had an amplitude spectrum of width $2F_L$ centred upon F_C, and was used as a masking stimulus in some experiments which required narrow band noise. The performance of the complete modulator was checked by replacing the noise generator with a sinusoidal oscillator. The output spectrum was non-ideal for two reasons: first, due to the non-linearities in the gate and in the second lowpass filter, components also appeared at harmonics of F_C. The strongest of these was -60 dB relative to the basic noise band. Second, due to imperfect balance in the gate, components also appeared in the frequency range below F_L. These were greater than 40 dB below the main noise band.

The generation of narrow band noise by modulation did have several advantages: BW and centre frequency were easily changed by varying F_L and F_C, and the first lowpass filter could be of quite a simple type since its rate of attenuation with frequency was effectively multiplied many times by the modulation process. However, in order to avoid a gap in the spectrum of the modulated noise, it was important that both the gate and the first lowpass filter should have an extended low frequency response. (These were uniform from the cut-off frequency, F_L, down to 1.5 Hz.) Further, the spectrum of the noise generator should be level down to as low a frequency as possible. For this, a laboratory noise generator was modified to improve its low frequency output. Its power density spectrum is shown in fig. 10 and is based on measurements made with the Radiometer Wave Analyser.

III.2.3. Comparison of the Two Methods of Noise Generation

Examples of the noise spectra generated by the two methods are shown in figs. 11 and 12. Both were measured using a Radiometer Wave Analyser, FRA2, together with the Solartron voltmeter; the analysis

BW was 4 Hz. A comparison of figs. 11 and 12 shows that even though the nominal BW of the modulated noise was slightly wider, its spectrum was much more nearly rectangular due to the effective increase in the filter slope as the result of the modulation process.

The modulation system was found to be more reliable than the tape recorder, and a comparison experiment (using four subjects) showed no difference between the effects of filtered and modulated noises of similar powers and BW's. It was therefore decided to use the modulator for most of the experiments in which narrow band noise was required.

III.3. The Random Number Generator

The basis of the random number generator was a three-stage binary divider with adjustable feedback which allowed the number of stable divider states to be set to any number from 1 to 8. If the number of allowed states was 'n', then when the counter was driven by a periodic pulse train it would spend 1/n of its time in each state. Thus, when the input pulses were suddenly disconnected by a manually controlled switch, the circuit should stop with probability 1/n in one of its n states. This resting state was decoded by auxiliary circuits and presented as a random number on an illuminated display.

The random properties of the output depended on the assumption that the driving pulse train was independent of the starting and stopping signals. The pulse rate used (25 KHz) precluded any voluntary control of the output by the operator, as the counter could go through 300 or more cycles within his reaction time. However, there remained the possibility of some hidden synchronization (via the power supply rail, for example) and the only satisfactory check on this was whether or not the generated numbers appeared random. This was checked by recording sequences of its indications. It was found that firstly, successive indications were statistically independent (i.e., $P(x_i x_j) = P(x_i)P(x_j)$). Secondly, the distribution of run lengths (sequences of identical indications) were consistent with those expected from independent indications.

III.4. The Earphones and their Calibration

It was important to be able to specify the acoustic stimulus delivered to the listener in normal use, as well as to have some simple objective method of comparing earphones to enable the detection of changes in performance due, for example, to accidental damage. The testing method used by the writer was chosen primarily with the second requirement in mind.

III.4.1. The Artificial Ear

A standard method of objective earphone calibration involves the measurement of the pressure produced in an enclosed cavity (coupler) when the earphone is excited by a known sinusoidal voltage. The combination of coupler and pressure transducer constitutes an artificial ear. Several artificial ears have been proposed (British Standard 2042, 1953; Delany & Whittle, 1965), all of which are intended for use with supra-aural earphones which bear directly on the surface of the pinna. They were designed to simulate as closely as possible the acoustic behaviour of the eardrum and canal. However, a considerable disadvantage of supra-aural earphones is that they attenuate ambient noise relatively little, and may also increase inter-subject variability (Shaw, 1966). This may be due to the formation of air cavities between the convolutions of the pinna and the earphone cap (Delany, 1964). For these reasons, circumaural earphones were chosen which make only light contact with the pinna; the headband pressure is applied through an annular cushion to the surrounding skin. Although circumaural earphones have the advantages of providing high attenuation of ambient noise, of possibly reducing subject variability, and of being comfortable to wear, there is the drawback that no standard coupler has yet been defined for their calibration.

The coupler which was used followed a proposal by Shaw & Thiessen (1962). It consisted of a brass disc, six inches in diameter, with a central hole. A Bruel and Kjaer 36 mm condenser microphone cartridge (type MK0001) fitted the hole, with its diaphragm in the surface plane of the disc. The circumaural earphone was held onto the coupler by a weight of about 500 gm, in addition to its own; the joint between the earphone cushion and coupler was sealed with vaseline. Electrical signals were supplied to the test earphones by an amplifier of low output impedance (0.5 - 1Ω) and its frequency response could be recorded by a Bruel and Kjaer automatic level recorder.

III.4.2. The Earphones

Using the artifical ear described above, Dr. H. Levitt and the writer tested most types of commercially available earphone, concluding that none was very satisfactory. Special units were therefore constructed, which gave good results at reasonable cost (about £12 per pair). Standard moving coil audiometer driver units (Telephonics TDH39) were mounted in a plastic circumaural shell (Anticoustic Ltd); empty space inside the shell was filled with a mixture of foam plastic and cotton wool damping material. The low frequency response of the earphone could be controlled by the density of the foam plastic used. Fig. 13

shows frequency curves obtained with two different types of foam. Pair 48 was used for all subjective tests.

III.4.3. The Relation between Real and Artificial Ear Performance

The interpretation of several experiments depended on knowing the exact ratio of S/N, and so it was necessary to specify the acoustic stimulus delivered to the listener. As it was not possible to measure the sound pressure in the ear canal, all stimuli were defined in terms of corresponding artificial ear measurements. Shaw & Thiessen (1962), using a probe microphone, found a maximum difference of about 10 dB between real ear and flat plate coupler pressures. The same conclusion was reached by the writer using a psychophysical method involving measurement of the hearing threshold.

Hearing thresholds for 10 ears (five subjects) were measured at seven frequencies in the range 100 Hz to 5 KHz. The subjects all had pure tone audiograms within normal limits; these were measured independently. It was assumed that the average of the threshold measurements was an estimate of the British normal hearing threshold (Whittle and Robinson, 1961). A comparison was made between the 'real ear' pressures (as derived above) and the measured sound pressure generated by the same earphones on the flat plate coupler. For frequencies below 4 KHz, agreement was fairly close; the biggest discrepancy was 8.8 dB at 2 KHz. It was therefore assumed that measurements of signals in the artificial ear were a more suitable indication of conditions at the subject's eardrum than measurements made electrically. All figures for SN ratios in this thesis represent acoustic values measured in the artificial ear.

CHAPTER IV. EXPERIMENTAL TECHNIQUE

Most experiments were performed in a fairly similar way. In this chapter, those details of procedure and analysis which were common to all experiments are described. Any differences from these will be described in the Methods sections of the individual experiments.

IV.1. Subjects

The subjects used were all male postgraduate students, less than 25 years of age, who were paid at the rate of five shillings per hour. Two selection criteria were used: first, that the pure tone audiogram for each subject was within normal limits,[+] and second, that the pitch DL for 1 KHz sinusoids in the absence of noise was less than 6 - 8 Hz. (It was found by experience that a subject who could not perform at this level after three trials, was likely to be unstable in later performance.)

IV.2. Test Conditions

For all pitch judgement experiments, the initial instructions to the subject were the same: "At each presentation you will hear two sounds separated by a short interval. The pitch of the first is to be regarded as a reference. You must answer each time whether the second sound is higher or lower than the first. When you have responded, the next pair of sounds will follow after a short delay."

[+] Audiograms were made by Miss Quinn of the Royal National Throat, Nose, and Ear Hospital, London.

A signalling box was held on the subject's knees and contained two silent switches which operated lights on the experimental apparatus. When making a pitch judgement, he was instructed to press one switch to indicate 'higher', and the other to indicate 'lower'. In addition, the signalling box contained a microphone which allowed two-way communication between the experimenter and the subject.

The test apparatus and the experimenter were in the main laboratory to prevent the transmission of unwanted cues to the subject. The subject was seated, by himself, in a small adjoining room (8' by 4') wearing circumaural earphones. The combined attenuation of outside noise introduced by the room and the earphones was about 70 - 80 dB. The laboratory was kept as quiet as possible during test sessions.

Stimuli were presented binaurally in the same phase and, except where otherwise indicated, at the same SL (Sensation Level) for each ear (50± 3 dB). The average length of a test session was 20 minutes (30 minutes maximum), and the subject was allowed a short rest halfway through. Test sessions on the same day were separated by at least three hours.

IV.3. Recording of Data

A subject's performance could be described by making two estimates from his responses. These were the DL, defined as the distance on the frequency axis between the 71 and 29 pc points of the response curve, and the point of subjective equality, or M. Initially, a Random method, based on the method of Constant Stimuli was compared with a Sequential technique based on the method of Constant Stimulus

Differences. (These were described in Chapter II.) The former was the normal method of gathering data for Probit Analysis (PA), and involved making the same number of stimulus presentations (eight) at each of five values of frequency shift, d_f (0, $\pm f_d$, $\pm 2f_d$). The spacing f_d was about 0.8 - 1 DL, and the order of presentations was random. Before recording of data commenced, about 10 - 15 stimuli were presented to accustom the subject to the situation. A CN curve was fitted to the subsequently recorded response data to obtain the DL and M.

The second method was a Sequential technique developed by Wetherill. Here, before starting a test, a spacing interval was selected which defined a large range of d_f values. (As above, this interval was about 0.8 - 1 DL.) The actual d_f values used however, were dependent on the behaviour of the subject. After 10 - 15 practice judgements, subsequent stimuli were presented according to one of two sequential rules, designed to estimate the 71 and 29 pc points of the response curve (see table 1). The rules were interleaved at random to disguise them from the subject; the random number generator (see p. 82) indicated which rule was to be followed for each presentation. The procedure commenced by making presentations at equal positive and negative values of d_f. Thereafter, the rules were used to modify the starting values to match the subject's performance (see below).

Each response given by a subject was recorded on a test sheet as either U (higher) or D (lower). The rule used for the 71 pc strategy was that any D-response would cause the next presentation of that strategy to be made at the next higher value of d_f. The d_f value would be reduced only after two consecutive U-responses were given at a particular level. The 29 pc strategy was an exact mirrorimage of the above. An example of an experimental record is shown in fig. 14.

The Wetherill estimator was a simple measure of central tendency for the data recorded by each strategy, and depended on the oscillating behaviour of the response pattern. This estimator is described on p. 71, and involved an arithmetic average of the midpoints of runs 1, 3, 5, 7, etc. From each test, two estimates were immediately obtained, which were the d_f values corresponding to the 71 and 29 pc points. The DL was then defined as the difference between these, and the midpoint as their average. The use of this estimation method meant that each test involved a different number of presentations. Also, these were not always equally divided between the 71 and 29 pc strategies, since it was necessary to continue each strategy until the required number of turning points, or runs, were recorded. Most estimates were based on four runs per strategy.

IV.4. Experimental Design and Method of Analysis

For most experiments, factorial designs were used. This had the advantage of allowing both the individual and interactive effects of several experimental variables to be studied simultaneously. An Analysis of Variance was used to test for the significance of simple main effects and interactions using a fixed effects model (see e.g., Brownlee, 1960). This was felt to be appropriate, since the experimental variables were closely controlled and the same group of subjects was used for most of the experiments. It was generally found that differences between subjects were large and consistent. Thus, the use of a fixed effects model meant that any conclusions drawn were applicable only to the subjects tested. It was also found however, that all subjects tended to react in the same qualitative way to changes in the experimental variables, so that conclusions reached should at least predict qualitative effects for a wider population.

IV.4.1. Analysis of Variance; Assumptions and their Validity

One powerful advantage of the Analysis of Variance is the small number of assumptions on which it depends. The two most likely to be important in practice are: 1) that errors are normally distributed and 2) that errors are of uniform variance. The importance of these assumptions has been discussed by Cochran (1947), and his conclusions will be summarized here. Both theoretical studies and sampling trials show that although strictly they apply only to normal populations, the F and t (2-tailed) distributions are not much affected by reasonable departures from normality. (Cochran suggests that the true significance level corresponding to an estimated probability of 1 pc, may lie between 0.5 and 2 pc.)

The more critical assumption seems to be that of homoscedasticity[†]. Analysis of Variance estimates an average error variance; if errors are non-uniform, the variances of some measurements will be underestimated, while those of others will be overestimated. The effect of this is difficult to predict. Error variance may be heterogeneous in two different ways. First, the variance of a quantity may be functionally related to its mean value. Variability of this kind can be dealt with by computing the error and other mean square estimates from some non-linear transformations of the observations, rather than from the observations themselves (Bartlett, 1947). It can be shown that if a set of random variables, y_i, have mean values, $\overline{Y}_i$, and variance s^2_i, where:

$$s_i = f(\overline{Y}_i)$$

[†] i.e. uniformity of variance

then the function

$$g(y) = \int 1/f(y)dy \qquad (2)$$

has a variance which is roughly independent of its mean value.[‡] For example, for all the DL figures quoted in this thesis, it was found that variance was proportional to $(\text{mean})^2$. Thus, from Eq. 2, the appropriate transformation to stabilise the variance was:

$$g(y) = k\int 1/y(dy) = k(\log y) \qquad (3)$$

It was found that this transformation not only stabilised error variance, but also tended to make the error distribution more nearly normal. The log transformation had the additional advantage of reducing the interaction effects between subjects when they occurred. This might be expected in view of Weber's law, in that a logarithmic scale should be more appropriate than a linear one.

The second situation is where error variance is heterogeneous without any underlying functional relationships. This was true for all midpoint data, where the variance was a function of the DL, rather than of M itself. (The DL and M tended to be independent.) Errors of this type should strictly be handled by using a weighted Analysis of Variance, but suitable values for the weights are not always available. In practice, it was found that a transformation had an adequate stabilising effect, although not theoretically appropriate.

[‡] This result depends on the probability density function for 'y' being reasonably symmetrical, and on the function 'f' being reasonably smooth. This seems to be fairly well satisfied in practice. However, see Curtiss (1943) for a more detailed treatment.

IV.4.2. Averaging Transformed Variables

In general, the inverse transform of an arithmetic average of transformed quantities is different from the corresponding average of the original data. More concisely, if the inverse transform is defined by g^{-1}, then

$$1/N\Sigma y_i \neq g^{-1}(1/N)\Sigma g(y_i).$$

It was felt that averages in the transform domain were better, since the non-linearity of the transformation prevented an undue weight being given to the more variable observations. (In practice, the difference between averaging the transformed and the untransformed data was not very great; a comparison for the DL data of experiment P-1 (see Chapter V) showed that the inverse transforms of averages were about 20 pc smaller than their directly estimated counterparts.)

All analyses and significance tests in later parts of this thesis were performed on transformed data. Where mean squares are quoted, they are in logarithmic units. Mean values have been obtained by applying the inverse transform to appropriate averages of the transformed data.

CHAPTER V. PRELIMINARY EXPERIMENTS: MEASUREMENT OF PITCH DISCRIMINATION

Introduction The main purpose of this thesis was to describe and explain the effects of additive noise on the discrimination of pitch. However, the choice of a suitable measurement technique was important, and in Chapter II some preliminary considerations were discussed. The two methods which were described in some detail were the Random method of Constant Stimuli with an appropriate curve fitting technique (PA), and the method of Constant Stimulus Differences which involved a Sequential strategy. In particular, the Wetherill method (W-method), which was associated with a very simple estimation technique, appeared to have certain advantages. However, it had been developed for the testing of plastic pipes, and its value in psychophysics might depend, for example, on how far its assumption of the independence of successive responses could be satisfied. Levitt (1964), in experiments involving judgement of the apparent intra-cranial position of a sound image, had favoured the W-method, since it both yielded an efficient estimate of the DL and was able to track the changes in M which sometimes occurred.

An independent evaluation was undertaken partly to gain experience with the method, and partly because the experimental situation was somewhat different from Levitt's. The first experiment (P-1) was designed to compare the W-method with the better known PA under working conditions. The results of this suggested further experimental tests, and these form the rest of the material in this chapter.

V.1. Comparison of Probit Analysis and the Wetherill Method under Working Conditions: The Effects of Noise Bandwidth on the Pitch DL (Experiment P-1).

V.1.1. Introduction

Experiment P-1 was first conceived as a general preliminary study of the effects of additive variable BW noise on pitch discrimination. It also provided an opportunity to evaluate the W-method, by comparing it under working conditions with PA. The discussion of the results will be confined here to this comparison; their relevance to pitch discrimination will be dealt with in Chapter VI.

V.1.2. Methods

Stimuli were generated by adding tape-recorded noise of constant power density to a 1 KHz sinusoid. Eight noise BW's in a logarithmic range from 10 to 1280 Hz were used, and each noise band was centred on the signal frequency. In the absence of noise, the signal level was set at 50 ± 2 dB above threshold for each subject, and was then kept constant as S/N was changed.† (The average values measured in the artificial ear are shown in Fig. 26.)

The principles of the two methods of stimulus presentation to be tested were described in Chapter II. In the Random method, five equally spaced (by about 1 DL) values of d_f were chosen in advance, and a total of eight presentations were made at each in random order. For the Sequential method, complementary 71 and 29 pc strategies (entry 2 of table 1) were chosen, again with a level spacing of about 1 DL. The two strategies were interleaved at random, and the process was continued until 40 judgements were recorded.

---

† Signal level was kept constant, and S/N varied by changing noise level only.

Each subject was tested with eight different masking noise BW's. Both Random and Sequential methods were used at each noise condition, and the experiment was replicated twice to allow a direct estimate of error. In each test session with a given subject, two groups of 40 judgements were recorded; thus, each subject participated in 16 sessions. Test conditions for each session were allocated at random.[+] The experimental design is summarized below:

Subjects (S)	4	
Bandwidth (B)	8	10, 20, ...1280 Hz
Test Method	2	Random, Sequential
Replications	2	

V.1.3. Results and Discussion

V.1.3.1. Analysis of the Data

Three sets of analysis were possible from the data obtained. These were: 1) PA of the randomly presented data, 2) PA of the sequentially presented data, and 3) W-estimates of the sequential data. This allowed the W-method to be assessed separately as a method of presenting data, and as an estimation method in its own right.

PA finds the slope of a fitted curve ($\hat{B} = 1/\hat{\sigma}$), while the W-technique estimates a DL directly. As noted on p. 71, the expected value of a W-estimate is 1.1σ. Therefore, for the purposes of comparison, all probit estimates were converted to DL's by the relation:

[+] Due to the random order, the 'error' term also included any real differences between occasions.

DL = 1.1/B. Midpoint estimates could, of course, be compared directly. The results are shown in tables 2, 3, and 4 and will be discussed in the subsequent sections. Table 5 presents a summary of subject averages.

V.1.3.2. Properties of the Error Variance

Since any test of significance would depend on the properties of the experimental error, this was examined first. A disappointing result was that the replication error was neither normally distributed nor stable in variance. The non-normality was shown by plotting cumulative distributions ~~histograms~~ of the average deviation (from their mean value) of the two data in each cell of tables 2 and 3 $(x_1 - x_2)/2$.[†] The distributions ~~histograms~~ are shown in fig. 15. Because only two data per cell were available, it was not possible to test for any asymmetry in the distribution, and therefore only half of each distribution ~~histogram~~ has been plotted. Both DL and M data clearly show departures from the linearity which would be found with normally distributed data. Further, the variance of the M data (shown by the slope of the curves) appears to differ for the Random and the Sequential presentations.

The lack of homoscedasticity was shown in two ways: first, the estimate of SD for each cell ($|x_1-x_2|/2$) was plotted against the corresponding mean value, $(x_1+x_2)/2$. It was found that SD(DL) was correlated with DL for both the Random and Sequential data (Pearson R_p = .61). SD(M) was more highly correlated with DL (R_p = .64) than with M (R_p = .19). Second, the data from tables 2, 3, and 4 were divided into two equal parts (on the basis of high and low means):

[†] The Wetherill estimate data of table 4 could have been used instead of that of table 3; the differences were not large.

80 Hz and below, and 160 Hz and above. The replication mean square was estimated for each part, and the results are shown in table 6. (The figures in brackets should be ignored for the present.) If the error variance were uniform, the ratio of the two estimates for each table should have been distributed as $F_{16,16}$; this null hypothesis was clearly untenable.[†] These properties of the data made a strong case for using a transformation before proceding further. The function used was: $\text{sign}\ (x)\log_{10}(1 + |x|)$, a simple extension of the basic log transformation given on p. 91. This was chosen because the data could be either positive or negative, and might contain small numbers and zeros.

For several reasons a logarithmic transformation was appropriate to the DL data. First, it is commonly found that the error of a subjective measure is proportional to its mean; indeed, the relatively high correlation between SD and mean mentioned on p. 96 tends to confirm such a linear relationship for the present data. Second, the error distribution was made more nearly normal, as shown in fig. 16. Finally, variance figures were calculated for the transformed data; these are the bracketed figures in table 6. In all cases, the F ratios (for the DL data) were reduced to a non-significant level.

As the lower correlation figures showed, there was less evidence of a functional relation between SD and mean for M, making the

---

[†] Although an inter-subject difference in variance was possible, this was ignored at this stage. Examination of the data suggested that, in any case, subject differences in this respect were small when compared with the effect of noise bandwidth.

logarithmic transform less appropriate on theoretical grounds. However, in practice, it did have a beneficial effect in reducing the departures from normality, as shown in fig. 16. Also, the SD's were made more similar for the two sets of data; the transformed F ratios in table 6 are considerably lower than for the raw data, although two are significant at the 5 pc level. It might have been possible to find a different transform which would lead to better M properties, but this was not pursued.

In conclusion, the residual heterogeneity of the transformed data seemed acceptable, at least insofar as Analysis of Variance was concerned, since the F test was known to be tolerant of reasonable departures from homoscedasticity.[+] (See also sections IV.4.1. and IV.4.2.).

V.1.3.3. Comparison of Test Methods

In this section, the data of tables 2, 3, and 4 will be examined with a view to assessing the relative merits of Random *versus* Sequential stimulus presentation, and Probit *versus* Wetherill estimation. There were however, two subsidiary questions which could be answered at this stage: the validity of the CN curve assumption, and whether

---

[+] A Study by Norton (see Lindquist, 1956, p. 78) of the effects of non-uniform variance, examined more extreme departures than found here (variance ratios of 4:1 and 8:1), and found small increases in the probability levels of significance tests. For example, a test of an apparent probability level of 2.5 pc might have a true probability of approximately 5 pc. On this basis, it was felt that the log transform could be used without serious misgivings.

subject's performance was more variable between than within occasions. These two points will be discussed first.

1) The CN Assumption

The PA computation generated a goodness of fit measure, χ^2, for each fitted curve. Using both the Random and Sequential data, a total of 128 χ^2 values were available. These are plotted as a histogram in fig. 17. For comparison, a true χ_3^2 is shown. The fit is good, although there is a suggestion that too many large values were found. This reflects one weakness of the Probit method: if the d_f values are poorly chosen, so that several groups contain all, or no U-responses, the fitted curve can become highly unstable, resulting in wild estimates and extreme values of χ^2. This situation sometimes arose with the rather small samples used here, even with prior knowledge of the subject's performance.

2) Stability of the Subject's Performance

PA finds both $\hat{M}$ and $\hat{B}$ and their expected SD's. An independent estimate of error was also available from the replication of the experiment. If the two sets of estimates were in agreement, this would suggest that occasion effects were small. (It would further suggest that the asymptotic formulae for SD($\hat{M}$) and SD($\hat{B}$) were valid for samples of size 40.) Accordingly, the observed and predicted SD's for both $\hat{M}$ and $\hat{B}$ probit estimates were calculated, and are plotted in figures 18 - 21. The $\hat{M}$ calculations were straightforward. For each pair of replications, PA gave four estimates: $\hat{M}_1$, SD($\hat{M}_1$), $\hat{M}_2$, SD($\hat{M}_2$).

The observed SD was: $|\hat{M}_1 - \hat{M}_2|/\sqrt{2}$

The predicted SD was: $|SD^2(\hat{M}_1) + SD^2(\hat{M}_2)|^{1/2}/\sqrt{2}$

For slope estimates, PA also gave four corresponding estimates: $\hat{B}_1$, $SD(\hat{B}_1)$, $\hat{B}_2$, and $SD(\hat{B}_2)$. For consistency, these were transformed to the corresponding figures for DL's.

The observed SD was: $|1/\hat{B}_1 - 1/\hat{B}_2|/\sqrt{2}$

The predicted SD was taken as: $|SD^2(\hat{B}_1)/\hat{B}_1{}^4 + SD^2(\hat{B}_2)/\hat{B}_2{}^4|^{1/2}/\sqrt{2}$

The expected result was that the plotted points should be scattered about a line of slope 1 (i.e., observed SD = predicted SD). Although there was considerable scatter, there was a general tendency for the points to lie about the line (see figs. 18 - 21). The number of points lying above and below each line are shown in table 7, below:

Table 7. Distribution of Points in Figures 18 - 21.

	DL Random Data	DL Sequential Data	Midpt Random Data	Midpt Sequential Data
Above the line:	8	14	15	22
Below the line:	23	17	17	10
Significance (Sign test)	.05	NS	NS	NS

A sign test suggested that significantly more DL points lay below the the line for the random data (see also fig. 18). Even in cases where points were equally divided above and below the line, the size of the departures was not symmetrical (on the log scale), being greater in the downward direction. This may have been due to the χ^2 distribution of the replication error estimate. Assuming that the line of slope 1 did predict the expected value of the observed SD (i.e., that the predicted SD's were <u>fixed</u> values and there were no occasion effects), confidence limits (P = 0.1) were calculated for the observed SD's, and are also shown in figs. 18 - 21. These account for the departures in all cases except for the Sequential midpoint data in fig. 21. Here, it seemed that a real occasion effect may have been present; the observed SD's being about two times too large.

In view of the limited data, further examination of this result was not attempted. The essential point was the finding of corroborative evidence for the absence of occasion effects in all except the Sequential midpoint data.

3) Methods of Data Allocation

Fig. 22 shows the relation between Probit estimates made on the Random and Sequential data of tables 2 and 3. Table 8 shows the Analysis of Variance for DL's and M's. This analysis compares the properties of the Random and Sequential methods of data allocation when PA was used. For DL's, only two effects were significant: the noise BW and subject differences (see Chapter VI). There was no suggestion that method of data allocation was important. (The relatively high correlation implied by Fig. 22a confirms this.) For M, three effects were found: noise BW, subjects, and interaction. This was because two subjects (AHM and PB) seemed to have larger midpoints when the Sequential method was used. This is also shown in Fig. 22b, where Sequential data midpoints tend to be larger than those of Random data. (This point will be discussed more fully at the end of this chapter.)

4) Methods of Estimation

Table 9 compares the properties of the two estimation methods for the Sequential data. Since each comparison was made using the same data, the situation was different from that shown in table 8. Although, in principle, 64 DF were available for the estimation of replication variance, all significance tests were made on the basis of 32 DF. This ensured that any correlations between the estimation methods did not influence the true significance level.

The data allowed 32 cross-comparisons between estimation methods using the cell means of tables 3 and 4 (four subjects x eight BW's). The mean squares for these have been shown in two components: the first was the overall effect of estimation method (E) averaged over all conditions; the second was all residual interactions of estimation with other factors. For DL's, the residual term is significantly less than the replication mean squares, while the main effect (E) is significantly greater. For M, the interaction term was again significantly small, and the main effect was not significant. The small residual mean squares indicate that differences due to the estimation methods were small when compared with differences between occasions. Therefore, the choice of estimation method was not critical.

The reason for the large main effect for DL's can be seen in fig. 23, where Wetherill <u>versus</u> Probit estimates are plotted. Here, although highly correlated, Probit DL estimates are about 24 pc smaller than W-estimates (see also table 5). This was surprising since the expected 10 pc difference between the two estimates had already been allowed for. The difference might indicate a small sample bias in the ML estimates, or that the simple theory of the W-estimator outlined in Chapter II, was inaccurate. (Quite a small shift in the true percentage point estimated on the response curve - from 71 to 75 pc - could have caused the observed 24 pc difference.) In any case, this was not a serious drawback in view of the consistency of the effect.

V.1.4. Summary and Conclusions

All methods tested gave fairly similar results for DL estimation, as tables 2 - 5 show. M estimation using the Wetherill method, was less satisfactory. It appeared that the properties of the errors were dependent on the method of data collection (see fig. 15). Although a logarithmic transformation did have a stabilising effect, it was still clear that sequentially obtained M results were more variable than the random data (cf. table 6a with 6c). Whether this was because the technique was more sensitive to a real instability in the midpoint, or whether instability resulted from the use of the technique, was not certain. The latter situation might have arisen for a number of reasons: possibly the random selection of one of two strategies was not sufficient to disguise the rule from the subject. Alternatively, some inherent relationships between responses might be involved to which the rule was reacting in an unstable way. (For example, if a subject had a tendency to give markedly more or less than 50 pc U-responses on average, the Sequential rule could never stabilise itself.)

Experiment P-1 was performed because the Wetherill Sequential method was relatively new, and its properties as a subjective testing technique were not well known. The conclusion reached was that for DL estimation, there appeared to be no difference between a Random and a Sequential method of data allocation (table 8), and that the two estimation methods tested (Probit and Wetherill) gave highly correlated, though consistently different results (fig. 23). However, the differences were not large, and the Wetherill method was preferred for its greater simplicity. (The comments of the subjects showed that there was no preference for either method of data presentation; indeed, some subjects could never tell which was in use on a given occasion. This was reassuring, since it suggested that the constant level of difficulty imposed by the Sequential strategy was not a disadvantage in practice.

It is also worth mentioning that in comparing the relative efficiencies of the methods (i.e., the relative variances of the estimates obtained from a given number of judgements), this experiment probably tended to favour the Probit method, since the results of preliminary trials were available when deciding the range of d_f values to be used.

The Sequential rule was possibly inferior for estimation of M. This was not serious, since the main interest in this work was in the DL for pitch, and the W-Sequential method with its associated estimators was therefore adopted for all further experiments. However, before proceeding further with the main experiments, it was decided to test the properties of the Sequential rule in more detail. This might show whether the more variable midpoints were real effects or artifacts of the method. The remainder of this chapter is devoted to several experiments having some bearing on this.

V.2. On the Statistical Properties of the Wetherill Method.

(Experiment P-2)

V.2.1. Introduction

Probit Analysis indicates how a choice of spacing for levels of a stimulus parameter affects the precision of the estimates obtained (see Appendix 1). It had thus far been assumed that the most suitable value of level spacing to use with the Wetherill method was about 0.8 DL, which was the best value for PA. Since it had been decided to use the W-method extensively, some further study of the effects of level spacing seemed necessary.

Further, one drawback of the W-method was that, unlike PA, there was no standard way of obtaining an estimate of error from a single experiment. Each estimate of DL and M had so far been obtained from the means of the midpoints of a number of 'runs', and it was possible that their variance might be used as a measure of error. Thus, the second object of this experiment was to compare this measure with the error estimate obtained by replication which had been used so far.

V.2.2. Methods

Wideband noise (4.5 KHz) was used to mask a 1 KHz sinusoid; the SN ratio was about -4 dB, making all DLs about 3 Hz. Five values of level spacing were chosen: 0.5, 1, 2.5, 5, and 10 Hz (covering the range from about 0.2 to 4 DL units). Two independent Wetherill estimates were made for each of four subjects at each spacing. For each subject, the test conditions were allocated at random. The design did not distinguish variance within a single session from variance between sessions; it merely allowed a comparison of variance between tests with that within a single test, regardless of when individual tests were administered. This seemed adequate in view of the result of Experiment P-1, that occasion differences were small.

The Wetherill 71/29 pc rules were used, and six run pairs were recorded under each condition. Separate estimates of DL and M were derived as the difference and the mean of the midpoints of each pair of corresponding runs; this gave six estimates of the DL and M for each test. The 'within test' variances were derived from these figures, and allowed a total of 40 DF for estimation at each level spacing. The mean squares for replication ('between tests') and subject differences were found in the usual way. In the tables, separate analyses are shown for each spacing to indicate the effect of spacing on variance.

V.2.3. Results and Discussion

A reassuring feature of the results, shown in tables 10a and 10b, was that the expected values of both DL's and M's were not normally affected by the choice of level spacing. Therefore, results could not be seriously biased by an unfortunate choice of spacing.

The variance properties were more complex. First, mean squares between tests were always greater (in eight cases, significantly so) than within-test variance. Thus, the latter could not be considered a satisfactory estimate of experimental error. Second, the two error variances reacted differently to changes in level spacing: while within-test variance increased steadily with level spacing, replication variance showed no consistent pattern for the smallest spacings, and increased markedly at wider level spacings. The optimal choice of spacing appeared to be about 1 DL, since this gave the lowest between-test variance with the most economical number of presentations.

V.2.4. Summary and Conclusions

This experiment showed that while a choice of level spacing for the 71/29 pc Wetherill strategy was not crucial, a value of around 1 DL was optimal, allowing good efficiency with a minimal number of judgements. Thus, it seemed that the same level spacing was most suitable for both PA and the W-method.

The finding that between-test variance was sometimes significantly greater than within-test variance was surprising. The interpretation of Experiment P-1, based on large sample Probit theory, had been that subjects were not more variable between, than within, occasions. The effect might have been found here either because the Probit theory was wrong (i.e., subjects were really more variable) or because the runs in the Wetherill strategy were not independent. For M, correlation between runs would explain the finding of P-1, that midpoints determined from Sequential data seemed to be more variable. A possible next step was to see whether this effect was subjective or statistical in origin by testing the strategy with the effects of subjects removed.

V.3. The Statistical Properties of the Wetherill Method from Synthetic Response Data (Experiment P-3)

V.3.1. Introduction

Experiment P-2 had shown that DL and M estimates were more variable between tests than within a single test. This suggested that either the subjects were more stable during a single test, or that the Sequential rule itself caused correlated responses. The aim of this experiment was to examine the properties of the Sequential method itself, in isolation from any subjective effects; this required a method of objectively synthesizing 'pitch judgements'. The method used made it very simple to record 'judgements' quickly, and it was possible to collect enough data for a fairly reliable assessment of the contribution of the rule itself to the effects found previously.

V.3.2. Methods

Apparatus A pulse-counting frequency discriminator was used to produce the responses; the arrangement of the system is shown in fig. 24. Stimuli were generated in exactly the same way as for a normal subjective experiment; wideband noise and a 1 KHz signal were used. The threshold device generated an output whenever its input voltage passed through zero. For each signal pulse (i.e., A or X) from the test apparatus, the counter displayed the time required for the first 100 zero-crossings. Thus, for a pair of signal pulses, two counts were displayed successively, and the sign of their difference was used as a 'judgement' of whether the frequency had changed upwards or downwards. The variance of each count increased as noise was added and hence, an increasing number of errors was made. Since error probability was also a function of the magnitude of the frequency change, it was possible to estimate a 'DL' in the same way as for a human listener.

Design A single fixed value of S/N was chosen which made the expected value of DL about 15 Hz. Two level spacings were used (0.8 and 0.4 DL). The 71/29 pc Wetherill strategies were used, and one or the other was chosen at random for each presentation. When eight complete runs had been recorded for each strategy, W-estimates of the DL and M were made. A new trial was started by returning to standard values of d_f (approximately ± 1 DL). The experiment was replicated 50 times for each spacing, and about 5500 presentations were made in all.

V.3.3. Results and Discussion

V.3.3.1. Synthetic Results

A preliminary examination of the data was made by plotting cumulative histograms of the DL and M values in the same way as for Experiment P-1. These are shown in figs. 25a to 25d. As expected, the error distribution was very close to normal on the original scale (cf. figs 25b and 25d with fig. 15 for the subjects). This meant that transformation of the data was unnecessary. However, for comparison with the subjective data, the log(1+X) transform was used for subsequent analysis. (The transformation had no great effect on the error distribution.)

Table 11 shows the mean values and replication variance for the DL and M estimates respectively; separate figures are shown for each value of spacing. The choice of level spacing had no significant effect on the replication variance or on the expected value of M. However, it was found that DL's were significantly smaller by about 15 pc for the narrower spacing. An effect of this magnitude was found as significant here only because the model enabled the collection of a large amount of data under very stable conditions. In practice, a bias of this order would not be important, in view of the typical SD of one DL estimate which is of the order of 30 pc.[+] However, it was felt that this point might be worth further study, although it was not pursued in this thesis.

V.3.3.2. Comparison between Model and Subjects

Table 12 includes the corresponding values from Experiment P-1, and compares the DL and M for the model and subjective data. (The average DL's have not been shown, as no attempt was made to relate the noise levels in the two experiments.) The table shows that error variance was similar for both subjects and model if the log transform

---

[+] A typical variance figure was .015, leading to an SD of about .12 (logarithmic); i.e., about 30 pc of mean value.

was used. This is equivalent to saying that the SD was about the same proportion of the mean value for both subjects and model.

Midpoint Data Average midpoints for model and subjects were both significantly different from zero. However, the average for the model was small enough to be explained in the following way: first, the experimental apparatus was set to produce frequency shifts, while the model discriminator measured period changes. The change in average period produced by an upward frequency shift of about 10 Hz would be slightly less than an equal downward shift. This would cause a small positive bias of the order of 1 pc of the DL. Second, the accuracy of calibration of the experimental apparatus was only about 1 - 2 pc. These two considerations could jointly account for the results obtained. However, it does not seem possible to explain the larger subjective midpoints in this way, if only because individual subjects were significantly different in this respect.

Serial Correlation Effect The relationship between successive runs on a given occasion was assessed by finding an estimate of DL and M for each of four run pairs recorded, and calculating correlation coefficients between: a) the 1st and 2nd, b) the 1st and 3rd, and c) the 1st and 4th, estimates. These figures are shown in tables 13a and 13b for DL's and M's. For the model, the DL data show correlations between runs, but this falls with increasing separation. In any case, the order of the correlations is not high enough to seriously affect estimates of error derived from four or more runs. On the other hand, the M data show more persistent correlation effects, especially at the narrower level spacing. These were high enough to preclude the use of run variance as an error estimate.

Corresponding correlation estimates are shown for the data of

Experiment P-2. Although the amount of data available was too small to obtain stable correlation coefficients, it is still clear that serial correlation effects were much more marked for M than for DL. The 'miscellaneous' column is derived from a larger amount of data, amassed from several experiments in which a level spacing of about 1 DL was used. This shows the effect more clearly; correlations tending to fall with increasing separation between runs, but remaining higher for M's than for DL's.

The serial correlations found for the subjective and the model DL's were not very different. This indicated that the effects were probably a property of the Sequential rule itself, and that there was no serious interaction between subjects and strategy. For M however, the subjective correlations were higher and more persistent, suggesting that the behaviour of the subjects themselves might be important. The effect of serial correlations would be to reduce the effective number of independent judgements available from a given test. This could explain the results of Experiment P-1, namely that M's estimated from a Sequential strategy tended to be more variable and their variance tended to be under-estimated by the asymptotic Probit formulae.

Wetherill & Levitt (1965) note that high inter-run correlations may exist, but do not discuss this point in detail. One way of deriving a more realistic error value might be to multiply the variance estimated from a number of successive runs by a constant factor to allow for their non-zero correlation (this would be a factor of 2 in the case of DL's). Alternatively, runs could be combined in groups. Neither of these possibilities was examined further.

V.3.4. Summary and Conclusions

The findings of Experiment P-3 were: 1) that correlation effects were present in all the data, both for models and subjects, and for DL's and M's, although they varied in magnitude. Therefore, error estimation from run variance was never used. 2) The model and subjective DL's were similar. Therefore, whatever drawbacks a Sequential technique might have, it was at least a reliable way of measuring DL's. 3) For M, there was a real difference between model and subjects; the model results showed less bias and a significantly lower serial correlation.

A comment by one of the more biased listeners (AHM) was of interest. He found that once an experiment was under way, he lost any sense of absoluteness of a pitch change, and tended to base his judgements on whether the current stimulus pair sounded similar to, or different from, the last pair. (I.e., he felt he could not judge each stimulus pair independently.) The expected behaviour of the Sequential rule is based on the assumption that each judgement is independent, and it seemed likely that by following any drift, the Sequential rule would make such a subject unstable. The situation would be worsened by the serial correlation effect introduced by the rule, so that a few 'wild' responses at the beginning could influence the course of a whole test.

V.4. On the Subjective Effects of the Wetherill Method (Experiment P-4)

V.4.1. Introduction

The experiments thus far had suggested that the W-Sequential rule tended to introduce statistical dependencies into sequences of otherwise independent responses. The effect was complex, and appeared to have a more marked influence on M than on DL estimation. It also seemed that subjects could introduce additional correlation effects -again, especially in the case of midpoints-. This might have occurred because the selection of one of two strategies at random was not sufficient to disguise the rule from the subjects.

This experiment was designed to test the contribution of the subjects, by interrupting the sequences with randomly timed 'irrelevant' stimuli. (These were normally presented stimuli which would be disregarded in the structuring of the Sequential strategy.) While this could have no effect on the statistical properties of the rule itself, it might reduce its influence on the subjects. The chief difference between the randomly and the sequentially presented data in Experiment P-1 was that the latter were more variable. It was expected that the inclusion of the irrelevant stimuli should make the sequential data more stable and hence, more like the random data of Experiment P-1.

V.4.2. Methods

The stimuli and SN ratios were similar to those of Experiment P-1 except that only two noise BW's (40 and 160 Hz) were used. Two types of Sequential rule were used: the first was the same as in P-1, and will be denoted '<u>S</u>'. The second, denoted '<u>S+R</u>', was such that the value of d_f for each stimulus could, with equal probability, be chosen a) from the upper strategy (71 pc), b) from the lower strategy (29 pc), or c) at random from a set of five values equally spaced by about 0.8 to 1 DL.

Presentations were continued until four runs (i.e., eight turning points) had occurred in both upper and lower strategies. The W-method was used to derive a single DL and M estimate. This meant that the estimates were based on different numbers of responses from Experiment P-1 (mean 45, range 32-57; cf. 40, for Experiment P-1). Four subjects were used, and the experiment was replicated twice. A factorial design was used, which is summarised below:

Subjects (S)	:	4	
Bandwidths (B)	:	2	40 and 160; S/N as in P-1 (see fig. 26)
Methods	:	2	
Replications	:	2	

V.4.3. Results and Discussion

The DL results and an Analysis of Variance are shown in table 14a. The DL was apparently unaffected by the inclusion of the random stimuli. The means and variance of the S and S+R data were very similar, and the figures were comparable with corresponding ones from Experiment P-1 (see tables 5 and 6). Thus, for DL estimation at least, the selection of one of two strategies at random, seemed to adequately disguise the Sequential rule (see Smith, 1961; Cornsweet, 1962).

The results for M appear in table 14b. It was found that the variance of the S+R data was about 0.6 times that of the S data. Although the experiment was too small to make this difference significant, it paralleled the finding of Experiment P-1, where the variances of the Sequential and Random data were in a similar ratio. This might indicate that the inclusion of random stimuli tended to have a stabilising effect, by introducing an occasional reorienting judgement.

A type of behaviour which predictably occurred with some subjects when a Sequential rule was used, was that although they would start with the upper and lower strategies being symmetrical about zero, they steadily drifted as the experiment progressed, usually moving in a positive direction. Thus, the lower variability of the S+R data was probably due to the subject being brought back to a fixed reference point at intervals.

V.4.4. Summary and Conclusions

A reassuring result of Experiment P-4 was that the inclusion of random stimuli had no significant effect on the properties of the DL estimates. This meant that the simple expedient of running two Wetherill strategies simultaneously, and selecting one or the other at random, was sufficient to disguise the sequence from a subject. (According to their comments, subjects were unable to distinguish whether the S or S+R strategy was being used.) In the case of the midpoints, the lower variabilities of the S+R data suggested that results might depend on the method of stimulus presentation.

The conclusion reached therefore, was that while DL estimation using the Wetherill method was straightforward and efficient, M estimation was probably rather unreliable. The extreme simplicity of the W-estimation procedure was still greatly in its favour, and since the DL was the chief interest in the present study, it was decided to use the technique for the main experiments.

V.5. General Summary and Discussion of Midpoints

The experiments described in this chapter were all designed to assess the Wetherill Sequential method and its associated estimators under working conditions. It was found that for the estimation of subjective DLs, the technique performed as well as the better known Probit Analysis while being much simpler to use (Experiment P-1). (As with PA, it was found that a range of d_f values spaced by about 1 DL was satisfactory for the application of the W-technique (Experiment P-2).) The non-randomness of the presentation sequence was apparently sufficiently disguised by selecting one of two strategies at random. This was shown in two ways: first, by the fact that a mechanical 'pitch' discriminator, whose responses were independent, gave results similar to those of the subjects (Experiment P-3), and second, by the finding that the inclusion of unrelated random stimuli had no effect on the DL (Experiment P-4). A disappointing result was that adjacent runs in a Wetherill strategy were correlated; a typical figure would be about 0.6 for DL's. This meant that the variance of the runs within a single occasion would tend to be biased, and its use as an error estimate was not considered further.

V.5.1. Midpoint Estimation

While DL estimation was straightforward and not markedly dependent on the test method used, this was not true for midpoints. As the main experiments in this thesis were concerned with the study of the pitch DL, midpoints were an interesting side issue only, and the only detailed discussion of midpoint results appears here. The main findings can be summarized as follows:

1. Estimates of M were significantly more variable when a Sequential rule was used (see table 6), ($P(F) \leq .05$). The inclusion of random stimuli (Expt P-4) appeared to reduce this effect.

2. Estimates of M obtained from adjacent runs recorded in a Wetherill strategy were correlated; this correlation was much higher for

subjective data than for comparable data from a model discriminator (Expt P-3).

3. There were significant differences between the M's of individual subjects and, on average, M was significantly different from zero whichever method was used (see tables 2b, 3b, and 4b).

4. M's were affected by changes in the noise level, but the effect was small when compared with differences between subjects, and not all subjects were affected in the same way (see tables 8b, 10b, and 14b). This was in marked contrast with the DL data (cf. tables 8a, 10a, and 14a).

V.5.2. Midpoint Variance Properties and Correlation Effects

It seemed likely that points 1) and 2) above, were simply two different aspects of a basic property of the sequentially obtained data; that correlation between successive responses reduced the effective amount of data. Experiment P-3 showed that the Sequential rule itself could introduce correlations into a sequence of independent responses. Serial correlation would reduce the effective number of independent data and thus, in any case, the Sequential midpoints would tend to be more variable. However, the subjective data showed that even further correlations, in addition to any due to the rule, could be introduced by the subjects.

The subjective correlation effects might have been due to an inability to judge stimuli in isolation, and the Sequential rule, from its mode of operation, could be very sensitive to this. For example, during a test each stimulus pair could be chosen from the same strategy (in which case the change in d_f would tend to be more than one step). Thus, the approach described by one of the subjects (AHM) of changing his response whenever the current stimulus pair sounded different from the last pair (see p. 112), could yield stable DL's but would lead to unstable M's. (Indeed, this approach would probably give measurable DL's even in the absence of a standard stimulus.) The finding of Experiment P-4 that error variance was reduced by the inclusion of random stimuli, is of interest here. By ensuring that, at intervals, the subjects heard symmetrical values of d_f, the random stimuli might be expected to stabilise their midpoints. This is discussed further in the next section.

V.5.3. Subjective Bias

The finding that individual subjects had significantly different M's, and that the average M of three (sometimes four) subjects was significantly different from zero, was unexpected. Although sequential M's tended to be greater, even the average for the random data was significant: the average M in table 2b was 1.39 Hz.[+] This result implied a preference for D-responses when d_f was, in fact, zero. Of 576 judgements recorded for $d_f = 0$, only 224 (39 pc) were U-responses.

[+] Using the replication mean square from table 2b to test a null hypothesis of zero, gave: $t = 9$, 32 DF, $P < .0005$. Even allowing a large safety margin for the known heterogeneity, this value was still highly significant. However, subject differences accounted for most of the variance of the data. Considering the 1.39 Hz as the mean of four midpoint values, each characteristic of one listener, gave: $t = 2.3$ 3 DF, $P > 0.10$. This showed that, although the individuals tested gave repeatable non-zero midpoints, the data were not inconsistent with a zero population average.

This could arise from the fact that whenever two complementary sequential strategies were used, the rules would attempt to force a subject to give about 50 pc U-responses on average. Possibly not all subjects operated this way. Although each experiment started with values of the unknown frequency (X) symmetrical about the reference (A), in practice, this situation would not be maintained for very long. Any preference by the subject for giving U or D responses would cause the Sequential rule to move the group of X frequencies in such a direction as to reduce this tendency. (This would mean that A was no longer symmetrically placed among the X's.) Since subjects would expect to maintain a fixed percentage of U-responses, this asymmetry might perhaps cause a further drift in M as the subject tried to compensate. This situation would be further worsened by the serial correlation effects mentioned above. (See also Campbell, 1969, who found biased judgements when using a sequential strategy to measure detection thresholds for signals in noise.) Harris (1948b) condemned the use of a fixed reference frequency, as a listener need not judge stimuli on a pair-by-pair basis, but could use some composite reference derived from a number of previous presentations. Possibly some listeners were prone to forming a biased reference of this kind.[†]

[†] To check this, AHM was retested using a 320 Hz masking noise at the level used in Experiment P-1. In the first half of a single experimental session, the presentations were made in a conventional way using a fixed reference frequency. In the second half, the reference for each presentation was chosen at random from the range 980, 985, ... 1020 Hz. The Wetherill method was used; six run pairs were

-continued on p. 120-

The reports of other workers on pitch discrimination have varied. The majority have found that bias was normally non-significant when judgements were made by successive comparison (e.g., Koester, 1945; Postman, 1946; Harris, 1948b, 1952). However, Tresselt (1948) and Flanagan & Saslow (1958) did find significant biases, though in each case this could have been due to an unusual experimental situation. Tresselt presented A-X pairs with additional background sinusoids of different frequencies; Flanagan & Saslow's work was concerned with the discrimination of synthetic vowel sounds, where frequency changes were accompanied by amplitude changes.

There are a number of problems in evaluating previous work on this topic. First, some inconsistencies may be due to the use of different measures of bias, some workers using error scores, and some using the parameters of fitted curves. Second, and more seriously, some studies where bias was <u>not</u> reported, are of no value in this context, since it was not stated whether any tests for bias were, in fact, made (e.g., Stevens, 1952; Henning, 1967). Finally, even where

---

(Continued from p. 119.) recorded for each condition. The results were: DL's of 8.1 and 12.5 Hz, and M's of 15.3 and 12.9 Hz for the fixed and random standards, respectively. None of these figures was significantly different from corresponding figures in table 4. However, both M's were significantly different from zero ($P< .01$, single tailed t-test). Although the result was based on a small sample (approximately 100 judgements), it suggested that the use of a fixed standard was not a primary cause of bias.

significant bias effects were found, the method of error estimation was not always described (e.g., Harris, 1952), and as already noted, the results of Experiment P-1 showed a significant bias only if subjects were treated as fixed effects.

V.5.4. Effects of Other Variables

A noticeable feature of most of the preliminary experiments was that M estimates were relatively little affected by any of the variables studied (see p. 117). In all cases, differences between subjects accounted for the major part of M variance. Thus, the differences between the Random and Sequential test methods discussed here should be viewed in perspective. Although real effects were undoubtedly present, they were small compared with the consistent inter-subject differences found. The analysis of the midpoint data of later experiments only tended to confirm the findings of this chapter, that although midpoints were quite large in some cases, they were relatively little affected by any experimental variables, other than subjects.

V.5.5. Conclusions

In this discussion, two main points have emerged: first, that individual subjects had consistently different midpoints, and that differences between subjects accounted for most of the variations in the M data. This had not been widely reported before. Second, the use of a Sequential strategy gave less stable midpoints than did a fully randomized presentation of stimuli. (It should be noted however, that this difference was not very large, amounting to a variance ratio of less than 3:1.) Its use for DL estimation had sufficient advantages to justify its adoption for the main experiments.

It was concluded than an experiment-by-experiment discussion of midpoint data would be repetitive, and would obscure the presentation of more interesting results. For this reason, midpoint data will not be presented for any of the main experiments, and the only discussion of midpoint effects appears here.

CHAPTER VI. MAIN EXPERIMENTS: THE EFFECTS OF NOISE ON PITCH DISCRIMINATION

Introduction This chapter includes a number of experiments designed to elucidate the effects of masking noise on the pitch DL, and to evaluate the IPD (Ideal Pitch Discriminator) discussed in section I.4.2. The IPD model made two explicit predictions: 1) that the DL should be independent of BW for noise BW's wider than that of the signal (about 10 Hz), and 2) that for constant BW noise, the DL should be inversely proportional to the square root of the SN ratio. A third, implicit prediction, was that for a given signal BW, these effects should be independent of signal frequency. As these predictions were fairly straightforward to test, their study seemed a natural starting point.

A second reason for studying the effects of noise was to repeat the classical CB (Critical Band) experiments in a pitch discrimination situation. The frequency selectivity implied by the CB effect, and the fact that the pitch DL and the CBW (Critical Bandwidth) are similarly frequency dependent, has led to the idea that pitch discrimination involves a CB 'filter' which transforms frequency changes at its input into amplitude variations at its output; these are then interpreted as changes in pitch (see section I.3.4.). Schafer et al (1950) and Corliss (1967) have proposed quantitative models using CBW values estimated from signal detection experiments. However, the assumption that results of a detection experiment can be used to explain pitch discrimination may be unjustified. The experiments in this chapter were performed to see if the CB effect could be demonstrated in a pitch discrimination situation.

VI.1. The Effects of Noise Bandwidth on Pitch Discrimination (Experiments P-1 and I)

VI.1.1. Introduction

Experiments P-1 and I both tested the effects of the same noise BW's on the DL for pitch. However, Experiment P-1 had the additional purpose of evaluating two test methods (PA and the Wetherill Sequential method), and the experiment was discussed from this point of view only in Chapter V. It was decided to repeat the experiment (as seven months had elapsed during the remaining preliminary experiments) to ensure that the subjects' performance had remained stable. Here, those results of Experiment P-1 obtained by the Wetherill method will be discussed together with the data from Experiment I.

In both experiments, masking noise of constant power density and variable BW were used, making the SN ratio inversely proportional to the noise BW. This was done as a counterpart of the classical CB experiment to see whether a CBW could be demonstrated in a discrimination, as well as in a detection, task. On this basis, the expected result was that the DL for pitch should increase with noise BW only up to a certain value, increase of BW beyond this point having no further effect. The value of the 'CBW' if found, could then be compared with the predicted IPD value which would be about 10 Hz if the ear matched its BW to that of the signal used. Thus, this experiment also tested the first prediction of the IPD model.

VI.1.2. Methods for Experiment I

The stimuli were generated and presented in exactly the same way as described for P-1 (see p. 94). However, there were some differences: first, only one replication was made, as the preliminary experiments had shown no interaction between subjects and BW's. (This meant that an interaction mean square could be used in lieu of an error estimate.) Second, a different stopping rule for the W-method was used: whereas in Experiment P-1, 40 judgements were allowed for each DL estimate, here, a test was continued until 12 runs had been recorded in each strategy (71 and 29 pc). On average, this required 95 judgements. Thus, that half of Experiment P-1 for which the Sequential method was used, involved about the same number of judgements as did Experiment I (2600 and 3000 respectively). However, in the former these were divided between two replications.

VI.1.3. Results and Analysis

VI.1.3.1. Separate Results

The results of Experiments P-1 and I are shown in tables 15 and 16. The DL's show a similar dependence on noise BW for all subjects; the small interactions in the Analysis of Variance tables were also evidence of this. A BW increase from 40 to 160 Hz caused a rapid rise in the DL; further increase in BW had relatively less effect, the DL rising, if at all, very much more slowly.[ǂ] This was similar to the

---

[ǂ] A feature of table 15 was that in all cases, the DL at 320 Hz was less than that for either 160 or 640 Hz. This may have been due to the use of tape-recorded noise. During Experiment P-1, it was observed that the calibration signal used to set the reference noise level at 320 Hz was not very stable. More care was used in Experiment I, and table 16 shows the 320 Hz dip in only one case. Since, as later experiments will show, noise level was a very critical variable, the use of tape-recorded noise was subsequently abandoned.

classical CB effect demonstrated by Fletcher (1940).

A gross indication of the overall similarity of the two experiments was given by their grand means, which were 10.1 and 10.8 Hz. On a more detailed level, corresponding cells of tables 15 and 16 were generally quite highly correlated. When comparing the results of the two experiments, it must be remembered that although each involved approximately the same number of judgements, Experiment I was singly, while P-1 was doubly, replicated. Thus, assuming independence between successive runs, the error variance of Experiment I should have been approximately 0.5 times that of P-1. Comparing the replication figure from P-1 with the interaction term of I (which should be an estimate of error), gave an actual ratio of 0.88. This could have been due partly to a real interaction effect, and partly to the serial correlation which could be introduced by a Sequential strategy (see p. 110). Comparing the two interaction terms directly, gave a ratio of 0.55, which was closer to the expected value.

VI.1.3.2. Combined Results

Since the two sets of data appeared similar, and since there was no reason to believe that error variance had changed markedly in the interim, the differences between occasions were tested for significance by analysing all the data on the basis of three factors: subjects, BW's, and occasions.[ł] The results are shown in table 17b. The only

---

[ł] This word is misleading in the sense that more than two occasions were involved. The occasion factor here distinguishes between two groups of data, each of which was concentrated in time and which were gathered seven months apart.

effects were those of noise BW's and subjects. Since the mean square for occasion was not significant, it was concluded that the subjects had remained stable over the seven month period, supplementing the finding of Experiment P-1 that there was probably no significant short-term variation either. The results of Experiments P-1 and I were averaged, and appear in table 17a.

Subject Averages There was no evidence of interaction between subjects and noise BW's, although the effect of each factor was individually significant. This indicated that all subjects were affected in qualitatively the same way by the change of noise BW, while showing repeatable quantitative differences at the same time. Thus, averages over subjects were an appropriate indication of trends in the data.

The Effect of Noise BW Table 17a shows that for increasing BW from 10 to 40 Hz, the average DL rose from 3.4 to 5.3 Hz; between 40 and 160 Hz, a rise of 5.3 to 18.7 Hz occurred. Wider BW's had less effect, and no subject gave a significantly larger DL for 1280 than for 160 Hz. For all subjects, the difference between 40 and 160 Hz DLs was significantly greater than that between 160 and 1280 Hz (t test, 2-tailed, $P< .01$), although the change in S/N was actually greater in the latter case.

The averaged results of the two experiments are shown in fig. 26. The plotted points are means for each subject; the solid line shows the result of averaging over subjects. The vertical bars indicate the estimated ± 1 SD.[†] The generally consistent ordering of the

---

[†] These estimates of error were relevant to a repetition of the same experiment with the same subjects; i.e., the subjects were

-continued on p. 128-

points for each noise BW would be expected from the small interaction effects mentioned earlier. Fig. 26 suggested approximating the BW-DL relation by three linear segments, indicated by the dashed lines. A variance ratio was used to test the hypothesis that in the ranges of 10 to 40, and 160 to 1280 Hz, noise BW had no effect (i.e., that the horizontal segments were appropriate). The results are shown in table 18 below:

Table 18. Partitions of the Variance due to Noise BW.

	10-40	40-160	160-1280	10-1280
F ratio	4.4	20.0	3.2	47.0
Significance	.05	.005	.05	.005

Although all the F ratios were significant, those for 10-40, and 160-1280 Hz were relatively small. A more detailed examination of the data of tables 15, 16, and 17 showed the reasons for this: first, for the 10 - 40 Hz, a genuine rise in DL was probably occurring, as all subjects showed a greater DL for 40 than for 10 Hz. Second, for 160-1280 Hz, there was no such consistent trend, and the main reason for the significant F ratio was the 320 Hz dip mentioned in the footnote on p. 125. Thus, the three segment approximation shown in fig. 26 did account for a large part of the variation due to noise BW. (Later measurements on the same subjects under noise-free conditions are

---

(continued from p. 127.) regarded as <u>fixed effects</u>. In both experiments the mean square for subjects was highly significant, indicating large inter-individual differences. Confidence intervals for an experiment treating subjects as a random sample would have been at least twice as large.

included in fig. 26 for comparison. They confirm that the effect of the masking noise was small at BW's below 40 Hz.) It was concluded that the pitch DL was probably not a linear function of S/N when S/N was varied by means of noise BW.

VI.1.4. Discussion

The relative constancy of the average DL at BW's below 40 Hz merely showed that the limit of discrimination had been reached, and that, as expected, decrease of noise below some minimal level could reduce the DL no further. However, the small rate of increase of the DL in the 160 - 1280 Hz region was more interesting. This effect could be explained by a cascade connection of a linear bandpass filter tuned to the signal frequency and a pitch discriminator whose performance depended on the S/N at the filter output. The filter would effectively exclude noise components at frequencies far removed from that of the signal and thus, the noise power at its output would tend to remain constant at wide BW's. Depending on the assumed shape of the filter characteristic and the properties of the discriminator, the filter BW would probably lie between about 40 - 160 Hz. Such a model is identical with that proposed by Fletcher (1940) to explain the CB effect in signal detection and, in fact, most estimates of the CBW at 1 KHz have also fallen in the same range (see e.g., Scharf, 1966).

The IPD model could also be regarded as a tuned filter and discriminator; however, its filter would be 'matched' to the signal (i.e., its impulse response would be determined by the signal waveform) and thus, its BW would be the same as that of the signal. The 256 mS signals used in this experiment had a BW of the order of 8 - 10 Hz, and it seemed unlikely than an internal filtering process of this order could account for the data of fig. 26; a value of the order of

80-160 Hz would seem to fit the data better. What the data could not show however, was whether the suggested aural BW of around 100 Hz was fixed, or whether it represented the lower limit of adjustment of some approximation to an IPD.

There were other explanations of the results which were different in principle. <u>First</u>, it seemed possible that as noise BW increased, its greater masking effect (due to the decrease in SN ratio) might be offset by its becoming <u>less tonal</u> in quality and thus, more distinct from the signal. As a result of this, the DL might remain constant with increasing BW. <u>Second</u>, at the wide noise BW's, the stimuli were loud (approximately 70 dB SL) because of the high noise level. Possibly, some non-linear process operated to reduce the extreme excursions at high noise levels, thereby reducing the changes in S/N. If the DL varied in some smooth inverse way with S/N at the output of the non-linear processor, the effect observed in fig. 26 could occur. The location of such non-linearity could be in the middle ear, which is known to have a noise-limiting function, or it might occur at the neural level (see Chapter VII).

VI.1.5. Summary and Conclusions

The main purpose of Experiments P-1 and I had been to see whether the masking effects of variable BW noise were consistent with the IPD model and, related to this, whether an effect analogous to the CBW phenomenon could be shown. For each of four listeners, the pitch DL appeared to be a discontinuous function of masking noise BW, and although individual results differed in magnitude, they did not differ in direction.

The results could be explained by postulating an 'aural bandpass filter' tuned to the signal frequency, and were in qualitative agreement with both the IPD and CB models. It did not seem worthwhile to make any numerical BW estimates from the data, since this would require arbitrary assumptions of the shape of the 'filter' characteristic as well as of the relation between the pitch DL and the SN ratio at the filter output. However, fig. 26 suggested that any estimate would probably lie between 40 and 160 Hz, and even the lower of these was several times larger than would be expected on an IPD basis. This range is that encompassed by most published estimates of the CBW at 1 KHz. The CB effect had not previously been shown in a discrimination task; detection situations having received more attention. The appearance of this effect in discrimination prompted the suggestion that a filtering process might be a common feature of several types of auditory analysis.

However, at the level of explanation used here, other quite different processes seemed plausible. These involved non-linear operations on loud noises, or a confusion between the pitch of a signal and that of narrow band noise. The results of Experiments P-1 and I did not enable any distinction between these explanations, and further experiments were designed to test them.

VI.2. The Effects of Spectral Distribution on Pitch Discrimination (Experiments II-A, II-B, and II-C)

Introduction The filtering explanation for the results of Experiments P-1 and I was not the only one possible. An alternative was that noise of different BW sounds different; i.e., noises of 160 Hz and below sound progressively more tonal, while at BW's greater than this, the tonal quality begins to disappear. Possibly, the increased difficulty in pitch judgements due to the lower SN ratio with wider BW noise (see fig. 26), was offset by the fact that it sounded less tonal, and therefore less like the stimulus. This section is devoted to three experiments which tested this by varying the relative spectral and tonal properties of signal and noise.

VI.2.1. The Effects of a Gap in the Noise Spectrum (Expt II-A)

VI.2.1.1. Introduction

This experiment compared two wide band noises of equal power, where the spectrum of one contained a gap centred on the signal frequency. (The two noises were not readily distinguishable by ear.) On the 'filter hypothesis', the gapped noise should have little masking effect, since it contained no components in the region of the signal. If noise power and tonality were the only relevant variables, then the masking effect of the two noises should be similar.

VI.2.1.2. Methods

A 1 KHz sine wave signal was used with two types of masking: a) wide band noise whose spectrum was uniform to 4.5 KHz (which was wide compared with any expected CBW), and b) noise whose spectrum contained a gap created by a bandstop filter; this gap extended from about 0.7 to 1.5 KHz. The maximum noise attenuation which occurred at the centre of the gap, 1 KHz, was about 18 dB.

The SN ratio was adjusted separately for each noise to -5.8 ±0.3 dB. Thus, the only difference between the two conditions was in the spectral distribution of the noises. Two subjects were used, and DL's were measured, using the W-method, under conditions of no noise and masking by each type. The experiment was replicated twice.

VI.2.1.3. Results and Discussion

The results in table 19 show that mean DL's for gapped noise were similar to those measured without noise, while DL's for uniform noise were about twice as large. The t-value for this difference was significant at the .05 level. The difference between the two masking noises was that the spectrum of one contained components in the region of the signal frequency, while the other contained a gap, about 800 Hz wide, in which the signal was placed. The negligible masking effect of the gapped noise implied that pitch discrimination involved some process which was able to separate wanted from unwanted stimuli on the basis of their spectral properties. This was additional support for filter model mentioned earlier.

VI.2.2. The Effects of a Gap in the Signal Spectrum (Expt II-B)

VI.2.2.1. Introduction

A converse experiment to II-A used a non-sinusoidal signal (a 1 KHz carrier-suppressed AM signal signal), whose spectrum consisted of two sinusoids separated by twice the modulation frequency, and was symmetrically placed about 1 KHz. Thus, by using a suitable noise BW and varying modulation frequency, the overlap of signal and noise spectra could be controlled. In other words, whereas in Experiment II-A the signal was placed in a gap in the noise spectrum, here, the noise was placed in a gap in the signal spectrum. It was expected that if there was no overlap in the spectra of signal and noise, masking effect would be small.

VI.2.2.2. Methods

Noise of 80 Hz BW was used, with modulating frequencies of 25 and 250 Hz, giving signal BW's of 50 and 500 Hz. Thus, for the 50 Hz BW, the signal spectrum was included within the noise band; for the 500 Hz BW, the spectrum lay symmetrically about the noise. (Subjectively, the narrow band signal in the presence of noise was very indistinct in pitch, while the wide band signal was clearly heard as separate.) Each signal was tested with and without masking noise. Data was gathered using the Wetherill method from only one subject (RLW), as the initial results were so marked.

VI.2.2.3. Results and Conclusions

The results, in table 20, show that whereas pitch judgements of the narrow band signal were considerably more difficult in the presence of noise, the wide band signal was little affected. Thus, the important criterion for determining whether masking would take place, again appeared to be whether or not the noise spectrum overlapped that of the signal.

VI.2.3. The Effect of Relative Spectral Shifts (Expt II-C)

VI.2.3.1. Introduction

This experiment tested the effect of spectral overlap of signal and noise in a different way, by shifting the signal relative to the noise on the frequency axis. According to a filter model, maximum masking should occour when signal and noise occupied the same frequency region.

VI.2.3.2. Methods

A masking noise of 80 Hz BW centred on 1 KHz was used. Signal frequencies of 750, 1000, 1250, and 1500 Hz were tested, both with added noise (at an SN ratio of 0.8 dB) and in a noise-free condition. Three subjects were used, and each was tested twice under each condition using the Wetherill method.

VI.2.3.3. Results and Discussion

The results are shown in table 21, and plotted in fig. 27. The solid lines represent an average of the three subjects at each point; the upper and lower curves represent the noisy and noise-free conditions respectively. The curves show that there was a significant masking effect at the 1000 and 1250 Hz frequencies only. This is consistent with the existence of a noise reducing filter centred on the signal frequency. There is a suggestion that the 1500 Hz signal was more affected than the 750 Hz signal. This might perhaps indicate an asymmetrical filtering process which is more responsive to lower frequencies. A phenomenon analogous to this has been shown with the threshold shifts produced by pure tone masking, in that masking has been found to be more effective when the masking tone was below the signal frequency (Wegel & Lane, 1924).

A further point is that the noise in this experiment was deliberately chosen to be narrow enough in BW to have a somewhat tonal quality. The pitch difference between the noise and the 1250 Hz signal was quite marked; even so, the noise caused a significant increase in the DL at this frequency. This was additional evidence that the masking effect of narrow band noise was not due to its similarity in pitch to that of the signal. (The inverse experiment, shifting noise relative to signal, gave the same results.)

VI.2.4. Joint Summary and Conclusions

In each of these experiments there was a different relation between the pitch of the signal and that of the noise. In Experiment II-A, the noise had no pitch quality; in II-B, although signal and noise tended to occupy a restricted spectral region, their pitches were different because of the low frequency envelope of the signal. In II-C, signal and noise sounded more or less similar, depending on the signal frequency; even here, the finding that the 1250 Hz signal was masked by a 1000 Hz noise (which sounded different) suggested that pitch similarity was not an important variable. The results of all three experiments were consistent with the hypothesis that noise would have a significant masking effect only when it occupied the same spectral region as the signal, regardless of the relative pitch of signal and noise.

Some experiments leading to a similar conclusion have been discussed in Chapter I (section I.2.3.2.). They showed that the masking and fatigue properties of signals possessing a residue pitch are determined by their frequency spectra, and are not related to their pitch quality.

VI.3. Comparison of the Effects of Noise Bandwidth and Power Density (Experiment III)

VI.3.1. Introduction

Experiment I had shown that the effect of noise on the pitch DL was not always linearly related to the SN ratio. (A 1280 Hz noise did not have a significantly greater masking effect than a 160 Hz noise, even though it was 7.6 dB more powerful.) Two alternatives to a linear filtering mechanism were suggested to account for the constancy of the DL at wide noise BW's: 1) the relative pitch of signal and noise, or 2) an amplitude dependent, non-linear process. The first point was examined in experiments II-A, II-B, and II-C; this experiment was designed to test the second possibility. This was done by varying the SN ratio (i.e., noise amplitude) in two ways: a) by noise BW, as in Experiments P-1 and I, and b) by power density. If a linear filter were involved, the DL should be more sensitive to variations in power density than to changes in BW. If amplitude were the relevant factor, they should have equal effects. As a further check for non-linear effects, the noise levels were slightly lower than those of Experiments P-1 and I, thus reducing the loudness of the stimuli.

VI.3.2. Methods

A 1 KHz signal was used with two noises of 160 and 1280 Hz BW; each was assessed at the same two values of power density. The two SN ratios for the 160 Hz noise were +0.9 dB and +2.9 dB; those for the 1280 Hz noise were -6.7 dB and -4.7 dB. (I.e., the SN ratios were 3 and 5 dB higher than those used in Experiments P-1 and I for the same conditions.) Three subjects were tested once each under each of the four noise conditions using the W-method.

VI.3.3. Results and Conclusions

The DL results and Analysis of Variance are shown in table 22. Only two effects were significant: subject differences and power density. A change in power density of 2 dB had a greater effect on the DL than a 7.6 dB change due to varying BW (cf. columns 1 and 3;

2 and 4, with 1 and 2; 3 and 4). The very marked effect of even a 2 dB change in power density made the flattening of the curve in fig. 26 more striking, and was strong confirmation that the relative constancy of the DL found in Experiments P-1 and I for noises between 160 and 1280 Hz was a real effect. That this result had been obtained at noise levels lower than those used in Experiments P-1 and I, made any non-linearities less likely.

VI.4. The Effects of Noise Level on Pitch Discrimination (Expt IV)

VI.4.1. Introduction

The results so far could have been generated by a system composed of a linear filter followed by a pitch discriminator, which was consistent with the IPD as well as other types of model. However, the IPD made three predictions: 1) the BW of the proposed filter was of the order of the signal BW (here, 8 - 10 Hz), 2) that the DL should be proportional to the square root of noise power density, and 3) that the effects described by 1) and 2) should be independent of frequency. The results of Experiments P-1 and I (see fig. 26) suggested that the first prediction was an underestimate; a value of 100 Hz would seem to fit the data better. This experiment was designed to test the second and third predictions by using sinusoidal test stimuli of different frequencies with added wideband noise of variable power density.

The filter model mentioned above would predict certain relationships between the results of this experiment (i.e., varying power density) and those of Experiments P-1 and I (i.e., varying BW). For variations in the power density of constant bandwidth noise, the SN ratio at the filter output should be proportional to that at its input.

On the assumption that the filtered S/N was the only determinant of the resolution of the discriminator, it would be expected that the pitch DL should increase continuously as noise power density was raised. (The IPD model would predict a linear relation with a slope of a factor of 2 in DL per 6 dB of power density.) However, the effects of varying noise BW (as in P-1 and I) would be of two types: for noise of BW's less than that of the 'filter', the external and internal SN ratios will be similar, and thus the results should be comparable with those obtained by varying power density. For wider noise BW's, the internal S/N (and hence the DL) would not vary. Therefore, it was expected that the data in fig. 26, if plotted in terms of SN ratio, should show a similar trend to those of the present experiment at low noise BW's, and an increased divergence at wide noise BW's. The point at which divergence began, might be used to give an estimate of the internal filter BW. This is analogous to Greenwood's method (1961) of estimating the CBW in a detection situation (see p. 48).

VI.4.2. Methods

Tonal frequencies of 0.5, 1, 2, and 3 KHz were used, with a masking noise derived from a noise generator via a 4.5 KHz lowpass filter. (Thus, the problems associated with tape-recorded noise, mentioned on p. 125, were avoided.) It was estimated that the SN ratio was stable to ±0.3 dB. Initially, a set of S/N values separated by 5 dB was chosen, attempting to cover the range between the DL becoming on the one hand unmanageably large, and on the other, too close to its noise-free value. Each of three subjects was tested twice at each S/N value, and twice with noise-free stimuli; the order of testing was randomized, and the W-method was used.

Testing was started at 1 KHz and all results were recorded at this frequency before passing to the others. This was done to avoid wasting data: if it appeared that the range of S/N values was inappropriate, this could be rectified. The initial results showed that there was a marked 'threshold' S/N value below which the DL increased rapidly. In the region of the 'threshold', a 5 dB interval was too great, and for the remaining three frequencies, a value of 3 dB was used.

There was another difference between the test conditions at 1 KHz and those at other frequencies. In all previous experiments, a signal frequency of 1 KHz was used, and in order to comply with the derivation of Sekey's model, the A and the X signals of each pair were always switched in the same phase. Here, for the first time, other frequencies were used, and it was convenient to use a method which meant that the signals were switched in random phase (see p. 75). The effects of this will be discussed together with the results of Experiment V, which allowed a comparison of fixed vs. random phase switching at 1 KHz (see section VI.5.3.1.).

VI.4.3. Results and Conclusions

The results for each frequency are shown in tables 23a to 23d.[+]

[+] The S/N values in tables 23a to 23d were different at each frequency, even though a constant 3 dB spacing was used (except at 1 KHz, as mentioned above). This was because the SN ratios were acoustical values measured in the Artificial Ear (i.e., they were affected by the frequency responses of the earphones).

Analyses of Variance of the data are shown in table 24. For the first time there was evidence of significant interaction effects involving subjects. Their importance was not great however, as the mean squares for subject and S/N differences were at least three times larger than the largest interaction effect. In other words, insofar as trends in the data are concerned, averaging over subjects had a small effect. Figures 28a to 28d show these averages (solid lines); individual data from tables 23a to 23d are also shown.

VI.4.3.1. Frequency and Noise Level; the IPD Predictions

Several features are of interest. First, as might be expected, the noise-free DL's tended to rise with frequency. Second, the effect of SN ratio was not linear; at each frequency, the DL was roughly constant until some 'threshold' value of S/N was reached, and then rose rapidly. In all cases, once the S/N was below 'threshold', the average DL rose at a rate very close to a factor of 2 per 3 dB change in S/N (as indicated by the dotted lines of this slope in the figures). This was twice the rate predicted by the IPD model. Third, the 'threshold' value of S/N rose with frequency, increasing from about -10 dB at 500 Hz to about +2 dB at 3 KHz. The effect of frequency was therefore twofold: the curves moved upwards and to the right at higher frequencies. Since the signal duration was constant, it was expected, according to the IPD model, that the curve relating the DL and S/N would be independent of the signal frequency used. This was clearly not so; although the average DL curves remained similar in shape, they were steadily displaced to the right as frequency rose.

VI.4.3.2. Comparison with Experiments P-1 and I

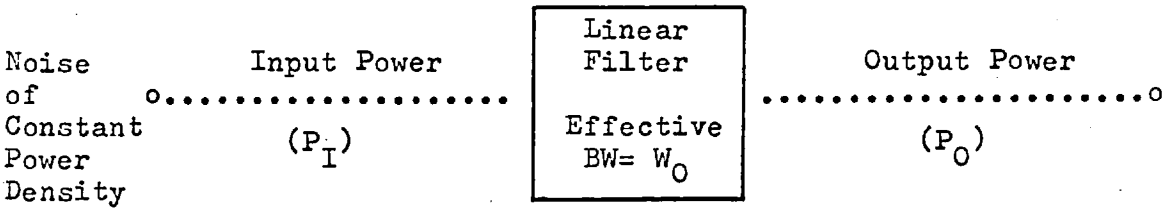
As mentioned on p. 139, on the basis of a linear filter followed by a pitch discriminator, certain parallels could be drawn between the results of Experiments P-1, I, and IV. For narrow noise BW's, variation of BW (as in P-1 and I) should have an equivalent effect to variation in power density (IV), while for wide noise BW's, variation in power density should have a greater effect than variation of BW. According to a filter model, a narrow band noise would have a given masking effect at a lower total noise power, due to its spectral concentration in the frequency region surrounding the signal. This is shown in fig. 28b (the 1 KHz data) where the averages from P-1 and I have been replotted on an S/N scale for direct comparison. (The data are separated horizontally because of the different noise BW's used in the two experiments.)

The main point to be made is that the two curves are approximately parallel for noise BW's less than 160 Hz (i.e., a line of slope 2 per 3 dB fitted fairly well through the points of the steeply rising portion of each curve). It would be possible to estimate a filter BW from fig. 28b, arguing that the BW at which the two curves cease to be parallel corresponds to the effective BW of the ear. It is interesting that the figure obtained (approximately 160 Hz) is of the same order as published CBW's (which range from 40 to 200 Hz at 1 KHz). However, there is an objection to this approach, which is that it attempts to impose an arbitrary threshold on what is, in fact, a continuous function. The finding that the narrow band data of Experiments P-1 and I were separated from the wideband data of Experiment IV, when plotted in terms of S/N, suggested a less arbitrary method of BW estimation, which is the basis of the next experiment.

VI.5. A New Method of Estimating the Filter Bandwidth (Expt V)

VI.5.1. Introduction

As mentioned above, the more usual methods of BW estimation were rejected because of the assumptions necessary. An alternative was suggested by the finding of Experiment IV, that varying the BW of narrow band noise yielded a similar curve to variation of the power density of wide band noise -but displaced on the S/N axis-.[†] The separation of these two curves might be used as a measure of the noise reducing effect of the aural filter. Consider the system shown below:



Where Noise 1 is narrow: P_1, $W_1<W_0$

and Noise 2 is wide: P_2, $W_2>W_0$

If W_1 is narrow enough for the filter output power (P_0) to equal the input power (P_I) when noise 1 is applied, and W_2 is wide compared with the filter BW (W_0)

then, for Noise 1: $P_0(1) = P_I(1)$

and for Noise 2: $P_0(2) = (W_0/W_2)P_I(2)$

[†] Similar results would presumably be obtained by varying the noise power density, rather than the BW of the narrow band noise. This point will be shown later.

If $P_0(1) = P_0(2)$

Then $W_0 = (P_1 / P_2)W_2$ (4)

Hence, W_0 could be estimated using only measurements made at the filter input, given some indication of when $P_0(1) = P_0(2)$. This method could be applied to subjective results by choosing two noises of wide (W_2) and narrow (W_1) BW's relative to the CBW. The pitch DL could then be measured for a few values of S/N at each noise BW. Plotting DL against S/N, the expected result would be two curves of similar shape, but separated horizontally. On the assumption that the DL was determined only by the SN ratio at the filter output, the displacement would be a measure of the reduction of the effective power of the wideband noise by the filtering process, and could be used as an estimate of P_1/P_2, enabling W_0 to be found in Eq. 4. (Experiments P-1, I, and IV provided some basis for this assumption.)

The success of this method would, of course, depend on sufficient prior knowledge of the true CBW to allow a correct choice of BW's; this should not be a problem in practice. (A correct choice would be verified by the results, since if W_i were too large, then Eq. 4 would give an estimate of $W_0 = W_1$. In other words, any estimate greater than W_1 would indicate a narrow enough choice of BW.) Experiment V was planned to investigate the use of this method under a wide range of conditions.

VI.5.2. Methods

A factorial design was used in which five factors were included.

1) Subjects: three subjects were tested.

2) Frequency: signal frequencies of 0.5, 1, 2, and 3 KHz were tested.

The signals were all of random phase. (Although it was felt from Experiment IV that gating phase was not an important variable, this design would make it possible to compare random vs. fixed phase switching for the 1 KHz data (see pp. 75 and 140).

3) <u>Bandwidth</u>: Wideband noise of 4.5 KHz BW was used; this figure being large compared with any expected CBW value. The selection of the narrow BW was a compromise between two considerations: first, the use of very narrow BW's meant that large values of d_f could cause the signal frequency to move outside the spectral region masked by the noise. Second, the CBW itself was expected to be of the order of 10-20 pc of signal frequency, and the narrow noise BW should be less than this.[†] A figure of 8 pc of signal frequency was chosen. This seemed reasonable in view of the fact that noise was generated by the modulation system described in section III.2.2., which produced narrow noise spectra with a steep cut-off slope.[††] Since the rate of attenuation of the noise bands was greater than that expected from any physiological filter, it would mean that for practical purposes, the narrow band noise power would be unaffected by any CBW filtering, provided that the CBW itself was about 10 pc or greater. This will be discussed later.

[†] The results usually quoted are of this order, and Experiments P-1, I, and IV suggested a figure of 160 Hz for a signal of 1 KHz.

[††] The cut-off slope was 24 dB/octave. But 'octaves' here are expressed in terms of the cut-off frequency of the lowpass filter used in the noise modulator (see p. 81), rather than absolute frequency. For example, for a BW of 900-1100 Hz, the octave points would lie at 800 and 1200 Hz. (See also figs. 11 & 12.)

4) Signal level: In all previous experiments, signal level had been set at a standard 50 dB above threshold. Here, the two levels (15 and 50 dB) were compared. If the wide and narrow band noises had different effects, even at the 15 dB level, it would provide further evidence against any non-linear effects due to the high noise power. (It would also be interesting to see whether any estimate of CBW was dependent on signal level.)

5) S/N value: Four values of S/N were tested. Those for wideband noise were near the 'threshold' region, based on the results of Experiment IV (see p. 141). The narrow band values were selected, after preliminary trials, to have a similar masking effect. This meant that the actual values of S/N depended both on the signal frequency and on whether the wide or narrow band noise was used. Apart from this however, the same S/N values were used for all subjects and at each signal level.

The above description is summarized below:

Factorial Design of Experiment V.

Subjects	:	4	
Frequencies	:	4	(0.5, 1, 2, and 3 KHz, switched in random phase)
Signal Level	:	2	(15 and 50 dB above threshold)
Noise BW	:	2	(wide, 4.5 KHz and narrow, 8 pc of signal frequency)
S/N ratio	:	4	(in 3 dB steps)
Replications	:	1	(5-factor interaction used as error estimate)

VI.5.3. Results and Discussion

The DL results are shown in tables 25a to 25d, and are plotted in in figures 29a to 29d. Separate Analyses of Variance were performed for each frequency and noise BW. The results of these are shown in tables 26a to 26d, and will be discussed by factors after a discussion of the effects of switching phase.

VI.5.3.1. The Effects of Switching Phase

As mentioned earlier, one difference between the 1 KHz and other frequencies in Experiment IV was the phase in which the signal was switched (fixed vs. random; see p. 140). It might be assumed that fixed phase switching would be more consistent with the exact knowledge of the signals assumed by the IPD. In the present experiment, the 1 KHz signals were switched in random phase to give some basis for comparison. Comparing figs. 28b and 29b, shows that the data overlap, suggesting that listeners were unable to make use of consistent signal phase.

There are three points to be noted: 1) the situation might well have been different if noise-free signals had been tested. In that case, the different switching phases are clearly perceptible. With additive noise, switched signals tend, in any case, to contain an initial transient of random amplitude and polarity, and so it is not surprising that signal phase was relatively unimportant. 2) Some confirmation of this comes from the work of Green (1961) and Green & Sewell (1962), who found evidence that observers were not always certain of the frequency, duration, and time of occurrence of a signal, even when these parameters were kept fixed. 3) This cannot be regarded as a firm conclusion, and should be tested using a wider range of conditions. If confirmed, it would be further evidence against the IPD model, as originally proposed.

VI.5.3.2. Subjects

In all cases subjects were significantly different, whereas the mean squares for interactions involving subjects were not (with one exception, at 2 KHz, narrow band). Thus, averages over subjects, shown by the solid lines in figs. 29a to 29d, were adequate descriptions of trends in the results.

VI.5.3.3. The Effects of Frequency

There was a tendency for the data to move upwards as frequency increased. For the narrow band noise, this was the most marked effect observed, and the data for all four frequencies are similar if simply shifted vertically. (This vertical movement was due to the known increase of the pitch DL with frequency.) This would be consistent with a filtering hypothesis which would predict that for noise BW's narrower than the aural BW, S/N should be the only determinant of the DL; for noises of wider BW's, this would not be so. The wideband data showed a similar vertical shift, but also moved to the right (towards higher values of S/N) as frequency increased. The decreasing separation between the two curves was interpreted as the result of an increase in the effective aural BW.

VI.5.3.4. The Effects of SN Ratio

As expected, the DL rose as S/N was reduced, but the function did not appear to be linear. More information on this effect was provided by Experiment IV, and it was found that the wideband data corresponding to the high signal level exactly overlap the Experiment IV results obtained under comparable conditions (cf. figs. 28a-d with figs. 29a-d). As in Experiment IV, the DL appeared to increase by a factor of 2 per 3 dB change in S/N. The lines used in estimating the filter BW were of this slope, and fitted quite well. The results suggested that the effect of S/N for narrow band noise was similar, but data were not available on the effect over such a wide range of S/N values.

VI.5.3.5. The Choice of Narrow Noise BW

The finding that although four different narrow noise BW's were used, effects appeared at about the same values of S/N, suggested that the chosen values were narrow enough (cf. 'narrow'data in figs 29a-d). An additional check was possible for the 1 KHz data; the data from Experiments P-1 and I overlap the narrow band data from this experiment for all BW's narrower than 160 Hz (cf. fig. 28b, "Expt I" with fig. 29b, "narrow"). Bearing in mind that the narrow band noises in this experiment were more nearly rectangular than those used in Experiments P-1 and I (cf. figs. 11 and 12), the choice of an 80 Hz BW at 1 KHz was narrow enough. Finally, using this method, CBW estimates (see table 27) were wider than the narrow noise BW's chosen. (Refer also to Eq. 4 and subsequent discussion.)

VI.5.3.6. Signal Level

Where present, the effect of signal level could be adequately described by a vertical shift of the data on the DL axis, but some anomalies were found. In three cases (500 Hz wide and narrow band, and 2 KHz wide band) the mean squares for signal level were large, and in two cases, significantly so (tables 26a and 26c). In the remaining five, the mean square for level is very small or zero, indicating that the average results at high and low levels at these frequencies were *more* similar than expected by chance. This finding, which might imply some correlation between test conditions, could not be explained.

There were several reasons for expecting performance to vary with signal level. First, measurements of the pitch DL in the absence of noise have shown that it would be higher at an SL of 15 dB than at 50 dB (Shower & Biddulph, 1931). Thus, it might be predicted that for noisy signals, the DL would also be greater at the low signal levels. In fact, in 21/32 cases, the average DL in figs. 29a-d was larger at the low level. If only the highest S/N values were considered at each frequency (i.e., the closest to the noise-free condition), this proportion was 7/8.

Another effect might work in the opposite direction (reducing the DL at low signal levels), particularly at low S/N values. The masking effect of wideband noise might depend on which of its components were above the hearing threshold of the subject. Since threshold tends to rise at high and low frequencies, the effective power of the wideband noise might be reduced at the low signal level. The non-uniform frequency response of the earphones (fig. 13) would also influence this effect. Although factors such as these might account for the detailed differences between the data, the main point was that any effects of signal level were fairly small compared with those of other factors. To a good approximation, the results at each frequency could be described as two curves of similar shape separated on the S/N axis; the separation being independent of signal level. This was the form of the result predicted by the filter hypothesis described in section VI.5.1., and implied that the BW of the filter did not change with signal level.

VI.5.3.7. Estimation of the Aural Bandwidth

As was predicted by the filter hypothesis, the curves for wide and narrow band noises were similar in shape - especially at the high signal level - but displaced on the S/N axis. The displacement was a function of frequency, decreasing as frequency rose. This was consistent with an aural filter whose BW increased with frequency.

As it appeared that the results were generally not very different for high and low signal levels, the method of BW estimation outlined earlier was used with the high signal level only. (It can be seen from inspecting the curves, that similar estimates would be found for the low signal level at all frequencies except 500 Hz, where the wide band curve was poorly defined.) Averages over subjects were used, as the data were insufficient for individual subject estimates to be made, and since there were no interaction effects. The separation between the wide and narrow band data was measured by fitting lines of slope 2/3 dB by eye to the average DL's for the lowest S/N values. The distance between the two lines (in dB), together with the corresponding estimates of the aural BW are shown in the last two lines of table 27.

For the three lower frequencies, there is general agreement with published estimates of the CBW in other situations (see table 27), although the figures for Experiment V tend to be somewhat larger. This might indicate that the CBW varies with the task in hand. It has been suggested by several studies, mainly in signal detection, that the CB may vary (Hamilton, 1957; Creelman, 1961; van den Brink, 1964). However, at 3 KHz, a very large value was found which was larger even than the largest other published value (de Boer & Bos, 1962; line 4 of table 27). Unfortunately, there is no direct way of checking either of these figures, since deBoer & Bos's work represent the only determination of the CBW using gapped rather than band-limited noise, and Experiment V is the only determination of the CBW in a discrimination task.

There are several possible reasons for the discrepancies, and the present results do not justify distinction between them:

1) The estimates quoted here are equivalent rectangular BW's. Since any CB filter would probably not be rectangular in shape (see Section I.3.2.1.), these figures are higher than the 3 dB values which are a more conventional measure of BW.

2) Although this method does not require the same assumptions as more conventional methods, there are two factors which tend to limit its accuracy: first, the CBW estimate is derived as a fraction of the BW of the wide band noise. It is difficult to measure this BW because of the non-uniform response of the earphones and the uncertainties of their interaction with the subjects. A comparison of the powers of the wide and narrow band noises using the Artificial Ear indicated that the effective wide noise BW might be about 4 KHz, rather than 4.5 KHz. Thus, the BW estimates in table 27 might be 10 pc too large. Second, the CBW depends on the distance in dB between the wide and narrow curves; a comparatively small error on the dB scale can have quite a large effect on the results.

3) The results depend on the assumption that the power of a noise is the only relevant factor in determining its masking effect. It may well be that at high frequencies particularly, noise envelope is equally important. It is fair to say though, that the same criticism can be made of most other determinations of the CBW.

VI.5.4. Summary and Conclusions

Experiment V showed that over a fairly wide range of conditions, the effect of noise on pitch discriminability, as measured by the pitch DL, was consistent with the existence of a filtering process. Further, it was possible to make estimates of the filter BW by a new method which did not depend on arbitrary breakpoints. The estimates were of the same order of magnitude, though larger, than conventional estimates of the CBW obtained from other types of experiment (see table 27), and varied in the same way with frequency. Thus, it was concluded that the ear shows frequency selectivity in making pitch judgements of noisy signals. From the wide applicability of the CB concept, which the present work extended to the discrimination situation, the filtering process might be viewed as a common preliminary to various kinds of auditory analysis.

VI.6. Evaluation

The experiments described so far showed that pitch discrimination under noisy conditions could not easily be explained by the IPD model. The main points of disagreement were: although a filtering process appeared to be involved, its BW was at least an order of magnitude too large, and was also frequency dependent (Experiments P-1, I, and V). Further, the DL was twice as sensitive as predicted to changes in S/N (Experiment IV).

An important feature of the IPD model was its ability to match itself to the signal. In one sense, the results clearly showed that some kind of adaptation was taking place, since the ear was apparently able to restrict its sensitivity to the spectral region occupied by the signal (Experiments II-A, II-B, and II-C). However, adaptation in the IPD sense had a more precise meaning, in that the ear was presumed to be capable of an exact matching which took into account the detailed structure of the signal. The discrepancies between signal and estimated aural BW's showed that this was not the case (Experiment V).

(Further suggestive evidence came from the finding that the phase of a signal was not an important variable; see p. 147.) Some type of adaptation may indeed be possible, but it may have been that the signals used here were of the wrong type to show the effect. This could be tested by repeating experiment V using a number of different signal BW's to see if these were reflected in the aural BW estimates obtained.

The IPD model would also show a threshold effect at high noise levels (Woodward, 1953). At this point, its behaviour would become more strongly dependent on noise level, and might indeed reflect the relation between DL and power density found here. It is difficult to test this, since the behaviour of an IPD-type of model below threshold depends on the assumptions about the prior distribution of signal parameters, and it is not easy to put bounds on this. In this situation, its behaviour then becomes analogous to that of a detector, rather than a discriminator of signals. Thus, an IPD model, to fit the data found here, would presumably be one designed to match signals of wide BW and to be (already) limited by some internal noise to working at, or below, its threshold value. This type of explanation was felt to be too complex, and rather than attempt to account for the above discrepancies by modifying the existing model, a simpler explanation was sought based on aural physiology. This forms the material of the next chapter.

CHAPTER VII. THE MODEL EAR

Introduction

As a result of the experiments in Chapter VI, it was suggested that the ear was able, in a pitch discrimination task, to separate wanted from unwanted stimuli on the basis of their spectral characteristics. The effective BW implied by this frequency selectivity was dependent on frequency, and was close to previously published estimates of the CBW. Rather than suggest alterations to the IPD model which was being tested, it was felt that a simpler model might exist which could be tied to what is known about aural physiology. Before discussing the proposed model in more detail, it may be useful to briefly introduce some salient features of the operation of the ear. (More comprehensive reviews of anatomy and physiology may be found in Wever, 1949; von Bekesy, 1960; Whitfield, 1957; Field et al, 1959; more detailed discussion may be found in Rosenblith, 1961; de Reuck & Knight, 1968, and other references cited therein.) The present review is presented solely as a means of justifying the major features of the Model. More detailed considerations will be discussed in context.

VII.1. The Operation of the Ear

VII.1.1. General Introduction

The word 'ear' will be used here to mean the entire sensory system concerned with the perception of sounds. This system can be divided into four main parts: first, the eardrum (or tympanic membrane) which is responsive to pressure variations in the external air; second, the middle ear mechanism which has the function of transmitting vibrations in a controlled way to the fluid filled inner ear cavity; third, the inner ear, or <u>cochlea</u>, where fluid pressure variations initiate neural activity; fourth, the <u>auditory nervous system</u>, both peripheral and central.

A diagrammatic section of the peripheral human ear is shown in fig. 30. Sound waves entering the external auditory meatus impinge on the tympanic membrane, causing it to vibrate. The vibrations of this membrane are transmitted by the three ossicles (malleus, incus, and stapes) across the middle ear cavity to the fluid filled bony labyrinth which contains both the vestibular system (which controls body equilibrium) and the inner ear mechanism. One purpose of the middle ear is to achieve a closer match between the impedances of the external air and the cochlear fluid; it also contains a reflex mechanism which prevents damage to the inner ear by loud sounds (the middle ear reflex).

A diagram of a part section through the cochlea is shown in fig. 31. In the centre of the cochlear duct is a closed sac (scala media) formed by the basilar and Reissner's membranes. (The latter is simple in structure, and probably serves merely to contain the endolymph fluid which fills the scala media.) The basilar membrane and the highly complex organ of Corti which rests on it, are essential to the process by which pressure variations in the fluid initiate impulses in the auditory nerve. Resting on the basilar membrane are the hair cells, usually one inner and three to five outer rows. These are constrained to follow, as closely as possible, the movements of the basilar membrane. Protruding from the upper ends of the cells are hairs, whose free ends are fixed in the tectorial membrane. This is a separate structure which is relatively resistant to movement, and is connected to the organ of Corti only via the hairs of the hair cells. Thus, any movement of the basilar membrane results in the application of a mechanical force to the hairs.

In general, each hair cell is connected with many dendrites, and each dendrite with many hair cells, although the interconnections are less complex for the inner hair cell system than for the outer. The axons of the primary neurons combine to form the cochlear section of the auditory nerve, which terminates in the cochlear nucleus in the brain stem. From here, neural connections run via several intermediate nuclei to at least two areas in the cerebral cortex. At each level, certain regularities have been shown which represent, in a more or less orderly way, points on the basilar membrane.

Finally, it has been shown that there are some different nerve fibres (running from the brain stem) to each cochlea (see e.g., Field *et al*, 1959). Stimulation of these has an inhibitory effect on the response to sounds, and they may be concerned with the control of the middle ear reflex, among other things. Their presence may indicate central control of the initial sound analysis performed by the cochlea (see e.g., Wiederhold & Peake, 1966; Dewson, 1967, 1969).

VII.1.2. Frequency Dependent Mechanical Behaviour

VII.1.2.1. The Middle Ear

As mentioned earlier, the middle ear has the function of efficiently coupling the tympanic membrane and the stapes. At frequencies below 2 KHz, the middle ear ossicles function as a rigid lever, and vibrations are transmitted with a pressure increase of about 18 times across the middle ear cavity (from tympanic membrane to stapes). However, transmission becomes less ideal at frequencies above about 2 KHz.

Fig. 32 shows a description of this effect, based on impedance measurements made by Zwislocki (1957).[+] This behaviour is analogous to that of a lowpass filter with a cut-off frequency of about 1.5 KHz.

VII.1.2.2. <u>The Basilar Membrane</u>

The movements of the basilar membrane have the dual purpose of performing a preliminary analysis of incoming sounds and of stimulating the primary auditory neurons. Pressure waves in the perilymph caused by stapes movements initiate travelling waves which run from base to apex. In their course along the membrane, these waves increase slowly in amplitude, reaching a maximum at some point, before decaying rapidly to zero. The locus of maximum amplitude is close to the stapes at high frequencies, moving apically as frequency is reduced. Thus, the basilar membrane functions roughly as a mechanical analyser, <u>mapping frequency into position</u>. The detailed behaviour was analysed by von Bekesy (1947, 1949a, 1949b), who observed the travelling waves elicited by constant amplitude sinusoidal oscillations of the stapes, and plotted the amplitudes of vibration of a number of points along the membrane as a function of frequency. Some of his results are shown in fig. 33. What the curves show is the asymmetry which follows from the unequal growth and decay rates of the travelling waves. The result is that the frequency analysis performed by the membrane is between band-pass and lowpass filtering. Teas & Henry (1968) also demonstrated this

---

[+] The effect of the external meatus, which has a resonance at around 4 KHz, is to compensate to some extent for middle ear loss at higher frequencies. This point will not be considered further, as frequencies below 3 KHz are the main concern in this study.

by recordings of the electrical activity from restricted regions of the basilar membrane in response to wideband noise. In sum, the relation between air pressure variations and the displacement of a given point on the basilar membrane can be described fairly accurately by the cascade connection of a lowpass filter (middle ear transmission) and a bandpass filter (travelling wave behaviour).

VII.1.2.3. Hair Cell Stimulation

Although movements of the basilar membrane are regarded as basic to the hearing process, the precise relationship between these movements and hair cell activity is also important, and not yet fully understood. As mentioned earlier, the hair cell bodies are fixed firmly to the basilar membrane, while the hairs themselves are embedded in the tectorial membrane, which is relatively resistant to movement. Basilar membrane movement is greatest in the centre and least at the edges, where it is rigidly attached to the inner and outer walls of the cochlear duct. The tectorial membrane, apart from being attached to the hairs, has only one point of attachment: at the inner wall of the cochlear duct. Therefore, when the basilar membrane moves in response to pressure waves, it causes the tectorial membrane to rock about its point of attachment, applying a stress to the hairs which is probably partly bending and partly tensile/compressive. Application of mechanical stress to the hair cells causes initiation of activity in the auditory nerve.

Various experiments using intense sounds to destroy parts of the Organ of Corti, have suggested that the inner and outer hair cells show differential sensitivity (see Wever, 1949). Von Bekesy (1953a, 1953b), using a vibrating needle to stimulate the tectorial membrane, also suggested that the inner and outer hair cells may be sensitive to different directions of stress. He argued that the structure of the organ of Corti could make the magnitude of stress applied to some hair cells dependent on the gradient of the basilar membrane, rather than simply on its displacement. A gradient sensitive process would have the effect of enhancing the boundaries of the displaced areas of the membrane, and would tend to transform the frequency response curves described earlier, into ones which were narrower and more symmetrical.

VII.1.3. The Electrical Activity of the Ear

Recordings of activity in the auditory nerve show that direct analogue representation of stimulus waveform is lost, and that whatever attributes of a stimulus are transmitted must be coded in terms of time, frequency, or spatial patterns of all-or-none nerve impulses. Stimulation with pure tones has shown that some auditory nerve fibres respond to a restricted range of frequencies, particularly at low intensities. As intensity rises, this range increases markedly in the low frequency direction. That is, the response curve of the fibre tends from a bandpass to a lowpass form (Tasaki, 1954; Katsuki, 1961). The refractory period of an individual nerve fibre limits its maximum discharge rate to a few hundred impulses per second, and thus it can directly transmit only quite low frequencies. However, although a fibre may not respond to every cycle of the stimulus, it does tend to produce spikes at intervals of an integral number of periods (Tasaki, 1954; Rose *et al*, 1968). It has been suggested that stimulus frequencies up to at least 2-3 KHz could be represented by a group of such synchronized fibres firing in rotation (the Volley Principle; Wever, 1949).

Neural units of the cochlear nucleus behave in a similar way, except that their frequency selectivity is less dependent on stimulus level. The higher auditory centres also contain pure-tone-sensitive elements but there are many which, although excited by complex sounds, are unaffected by sinusoids (see Ades, 1959). One would expect this since the pure tone is not a commonly occurring signal. Those elements which are sensitive to sinusoids however, show some frequency selectivity which, although poorer, is less dependent on intensity than that of more peripheral units.

VII.2. The Proposed Model: Components

VII.2.1. Introduction

The model proposed here consists of two main parts: the first (which was actually constructed) is concerned with the filtering operation of the middle ear and cochlea, and with a simple representation of the first stage of neural transduction. This part of the model is assumed to be essentially deterministic. The second part (which is theoretical) suggests how the results of the analysis performed by the first part might be used by the brain in arriving at a pitch judgement. The second part also accounts for the finding that subjective results are imperfect, even in the absence of external noise. As will be shown, it was not necessary to consider the second part of the model in great detail in order to compare the results with subjective data.

The experiments described in Chapter VI could be interpreted using a model involving three components: 1) a linear filter tuned to the signal frequency, 2) a 'pitch' discriminator operating on the signal at the filter output, and 3) a decision process which would compare the discriminator output under the A and X conditions and generate an appropriate response. The next three sections show that the physiological concepts presented previously can be useful in constructing a model of this type, and a definite proposal is made. The remainder of this chapter is devoted to a comparison between the subjective results and those obtained from the model.

VII.2.2. The Filter

In Experiment V it was found that the aural BW was of the same order as the classical CBW estimates. This led to the suggestion that the same peripheral filtering process might be involved in many types of auditory analysis. Since the middle ear and cochlea are known to have filtering properties of the type outlined in the previous section, it was interesting to see how far these alone could account for the results obtained with subjects.

The basilar membrane could be regarded as a set of mechanical filters which map input frequency into position. The analogy of a finite set of filters is not strictly accurate because of the continuous nature of the membrane, but it is possible to represent the frequency selectivity of a given point on the membrane by a lumped constant filter. Flanagan (1962) proposed a simple design method for representing the frequency dependent effects measured by von Bekesy (1947, 1949a, 1949b) and Zwislocki(1957). A model filter was constructed to this design, which consisted of: a) a lowpass section designed to simulate middle ear losses, followed by b) one of four bandpass filters to represent the displacement behaviour of four discrete points on the basilar membrane maximally sensitive to the four frequencies tested (0.5, 1, 2, and 3 KHz). A facility for representing membrane gradient was also included. This involved a second bandpass filter for each frequency, which represented a point 0.5 mm distant from the corresponding point represented by the 'displacement filter'. The difference between the outputs of these two filters was taken as a measure of gradient. A detailed description of the design is given in App. 2.

The frequency response curves of the 'displacement filter' are shown in fig. 34 (cf. fig 33, von Bekesy's observations). The filter BW corresponds to a Q value of the order of 2. It is noticeable that because of the middle ear characteristic, the 2 and 3 KHz points are not maximally sensitive at their resonance frequencies.[†] (The implications of this for the present model will be discussed later.) Similar curves for the 'gradient filter' are shown in fig. 35. These differ from fig. 34 in that the curves are narrower in BW and more symmetrical.[††]

A simple method of comparing the two types of filter was suggested by the results of Experiment V. This involved applying a signal and noise at a known SN ratio to the input of the filters, and measuring the reduction in noise power effected by each type for comparison with the subjective data. This is presented as Experiment VI-A below.

[†] Although these response curves are maximally sensitive at some frequency lower than their resonance points, no other points on the membrane are any more sensitive at those resonance frequencies.

[††] Another factor, the time delay introduced by the propagation velocity of waves along the basilar membrane was not included. Estimates of this time delay were available from von Bekesy's data for frequencies less than 1 KHz (1949b). They suggested that a reasonable figure would be of the order of 100 µS per mm. The effect of including a 60 µS delay was tested at 500 Hz; the effect was negligible.

VII.2.2.1. The Effects of Basilar Membrane Filtering: Displacement or Gradient Sensitivity? (<u>Experiment VI-A</u>)

<u>Introduction</u> According to the model proposed here, the mechanical behaviour of the basilar membrane, by causing the hair cells to be excited in a frequency-selective manner, should account for the filtering effects found in the main experiments. As suggested earlier, this might happen in two ways. The simplest way is stimulation by displacement of the basilar membrane; alternatively, the gradient of membrane displacement might be involved. This experiment was designed to test the two different kinds of model filter by applying a signal and wideband noise to the model input, and measuring the improvement in SN ratio due to the 'filtering' for comparison with the subjective figures of table 27.

<u>Methods</u> A continuous signal of 0.5, 1, 2, or 3 KHz was added to 4.5 KHz BW noise at a constant SN ratio. This formed the input to an earphone which was placed on the Artificial Ear. A suitably amplified version of the condenser microphone output was applied to the 'eardrum' terminal of the model filter (see App. 2). The output SN ratio was measured at the appropriate 'membrane point' to determine the effective reduction of noise power at each signal frequency.

<u>Results</u> The first row of table 28 shows the S/N enhancement of the Displacement Filter as a function of frequency. As might be expected, the figures tend to fall with frequency, due to the constant percentage BW characteristic of the filters. However, the asymmetry introduced by the middle ear characteristic had the effect of increasing this frequency dependence so far above the expected 3 dB per octave of signal frequency, that at 3 KHz, the SN ratio was worse at the output than at the input.

As the second row shows, the inclusion of gradient sensitivity compensated for the middle ear effect, leading to a positive S/N enhancement at all frequencies. The effect of this can be assessed by comparing row 2 with row 3 of the table; the latter shows the expected S/N enhancement if the middle ear effect were removed entirely. (These figures were calculated on a -3 dB/octave basis from the 500 Hz measurement.) The effects of gradient sensitive filtering are most marked at the two higher frequencies, as might be predicted from figs. 34 and 35. The fourth row of table 28 shows corresponding figures derived from the subjective data of Experiment V (from table 27). These are consistently larger by 4-5 dB than the Gradient Filter results, but differ increasingly from the Displacement Filter at the higher frequencies.

Discussion and Conclusions If the model were in fact a realistic analogue of the joint behaviour of the middle ear and basilar membrane, then this experiment implied that some type of sharpening process must be involved. This followed from the parallel tendencies of the results of gradient filtering and the subjective data of Experiment V. (The Displacement Filter was actually a handicap at the higher frequencies.) Therefore, on the basis of this experiment, the Gradient Filter was incorporated into the final model.

This approach can however, be criticized on two grounds. First, although the Gradient Filter results were parallel with those of the subjects, there was a consistent 4-5 dB difference, indicating that the effective subjective BW's were 2-3 times narrower. This could be because of more sophisticated aural sharpening mechanisms (neural or mechanical), or might be due to differences between von Bekesy's measurements on dead specimens and the behaviour of the subjects tested here. Johnstone and Boyle (1967), using guinea pigs, have found basilar membrane BW's about half those quoted by von Bekesy. Thus, a smaller BW is at least possible in some mammalian cochleae, whether or not it is ever to be found in humans. Second, this model implies that the selection of a region of the membrane is fairly precise, and that the BW effect is a consequence of the rather broad selectivity of the basilar membrane as well as the patterns of neural connections to it. However, to consider only a single point in a continuous system is an oversimplification. It might be approximately realistic if the ear is able to select for periodic local activity such as would be initiated by a sine wave signal in noise. This selection might occur at the cochlear level, either by local inhibition, or as a result of feedback from a more central analysis.

VII.2.3. The Pitch Discriminator

Two types of discrimination process were considered. First, whenever a filtering mechanism is involved, its frequency selective characteristics may be used to transform changes in frequency into corresponding changes in amplitude. (Pitch discrimination would then become a process of detecting changes in the amplitude of the filter output. One such model was discussed in the introductory chapter, section I.3.4.) An alternative would be to analyse the temporal properties of the filtered signal to obtain the pitch directly. While these two mechanisms would yield similar results in many situations, some phenomena, e.g., the missing fundamental, indicate that more weight is given to temporal than to spectral properties of signals where the two types of information are in conflict. These two possibilities will be discussed in more detail in the next two sections.

VII.2.3.1. Energy Discrimination

If the discrimination process involved a spectral analysis performed with a filter, then the DL for pitch should depend on the effective slope of the filter characteristic and on the minimum energy change detectable. Presumably, the filter would be the same one which had the effect of reducing noise power, and the minimum detectable energy change might be expected to be the same as the DL for intensity (typically, 0.5 dB (Riesz, 1928)). This type of model is attractive in view of what is known of the properties of nerve transmission: that individual neurons, because of their refractory periods, cannot directly transmit frequencies above a few 100 Hz. Thus, an explanation in terms of preliminary filtering performed largely or entirely in the cochlea, followed by transmission of relatively slowly changing energy information along the auditory nerve, is immediately plausible. (Note that this is essentially a 'place' theory of frequency discrimination, in that pitch is supposed to be signalled by the location of the region on the basilar membrane where maximum activity is occurring.)

A 'place' theory of pitch discrimination requires some process analogous to filtering, together with a method of discriminating changes of energy. There were reasons for rejecting this type of explanation, however. For example, the range of Q-values suggested for energy-detecting models has been from 40 to 150, while those inferred from Experiment were about 5 or less.[+] (This is in agreement with

---

[+] Most proposed energy-detecting discriminators assume a critical band filter of 'bell shaped' selectivity characteristics (LCR and gaussian shapes are two common examples). The important point is not

-continued on p. 168-

electrophysiological data in that no auditory neurons have ever been shown to have very narrow frequency response characteristics.) Thus, to explain pitch discrimination in this way would appear to require an order of magnitude increase in filter slope, or an equivalent reduction in the intensity DL.

Another objection to this type of approach is that the energy at a filter output might change for two reasons: either because the frequency of the input signal has changed, or because of a change in its intensity. It is reasonable to infer that the nervous system is able to distinguish these two cases, as if it could not, pitch discrimination would become almost impossible in any situation where the frequency response of the transmission medium is not smooth. (A telephone is an obvious example of a highly non-smooth response.) However, if the distinction is indeed made, it means that the use of the intensity DL to explain pitch change phenomena is of doubtful value.

Finally, the pitch of some sounds is not always determined by

---

[+] (cont. from p. 167) so much the particular shape of the filter, which has a second order effect on the estimated BW's, but that all filters of this type have a BW which is related to cut-off slope. It is quite possible that the effective filter characteristic of the ear is not of this type at all, but is much more rectangular, with steep cut-off slopes and a comparatively wide BW. Such a suggestion has been made by Whitfield (1957). This type of filter would not only be consistent with the relatively wide CBW's measured here, but its sharp edges could make a very high differential sensitivity to frequency change possible.

their frequency characteristics (see sections I.2.2 and I.2.3. in Chapter I). An energy oriented model would not account for these temporal effects. An alternative explanation, which did not suffer the above drawbacks, and which seemed to fit the data better, is presented in the next section.

VII.2.3.2. Temporal Analysis

Although it is not known exactly how the pattern of basilar membrane movement is transformed into nerve impulses, it has been found that frequency information for signals up to several KHz is present in the auditory nerves of the cat (Kiang & Goldstein, 1962), guinea pig (Tasaki, 1954), and monkey (Rose *et al*, 1968). This information is contained in synchronized nerve impulses which tend to occur at the same point in each cycle of the input signal. On a basic level, information on the frequency of a periodic stimulus would be available from those hair cells whose thresholds were at about the mean value of the signal. (If sufficient hair cells with different thresholds were involved, this would allow complete information on the stimulus wave form to be encoded, but this has never been shown.)

The proposed model involves a combination of frequency and time principles in that there is a preliminary filtering followed by a temporal analysis. The preliminary filtering might be mediated chiefly by the properties of the cochlea which, as shown, could more or less account for the effective BW's measured by noise masking. A subsequent temporal analysis of the filtered stimulus might account for the fine discrimination. This model has a number of advantages. First, it is consistent both with the results which suggest that pitch is a 'place' phenomenon (see section I.2.1.) and with the demonstrations that temporal information is sometimes involved (sections I.2.2, I.2.3).

Second, it avoids the problems encountered by a simple energy discriminating model of distinguishing two possible causes of energy change. It raises the problem however, of how temporal information might be transmitted at high frequencies, but this might be catered for by the Volley principle mentioned on p. 160, which suggests that groups of nerve fibres may fire in synchronized rotation at least up to several KHz.

If the auditory nerve transmits temporal information by synchronized nerve impulses, a conceivable method of analysis of the paired stimuli used in these experiments would be the comparison of a pulse count for the first signal with a corresponding count for the second. A 'judgement' of the direction of the frequency (pitch) change is made according to the sign of the count difference. If only those impulses which occur at the zero-crossings of the input signal are considered, the process becomes analogous to the operation of an FM receiver. This has a complex behaviour in the presence of noise: there is a sharp threshold at an SN ratio of about 0 dB, above which the input noise has a small effect on the output variance. Below the threshold, the noise has the effect of reducing the mean discriminator output for a given frequency shift; that is, the discriminator is made less sensitive by noise. Eq. 5 describes the effect of SN ratio (see Rice, 1948).

$$\bar{\phi} = d_f(1-\exp(-S)) \qquad (5)$$

where:

$\bar{\phi}$ = the expected change in the output level of the discriminator

d_f = the change in frequency of an input sinusoid (steady value)

S = the input SN power ratio (signal plus additive guassian noise)

Eq. 5 shows how the discriminator output falls with increase of noise level at its input.[+] It also shows how the SN ratio can be traded for input frequency change to maintain a constant mean output level. The LHS of Eq. 5 remains roughly constant for S>1, but falls rapidly as the ratio decreases below 1. If the neural transduction process can be likened to a zero-crossing discriminator, then it is possible that the increase of the DL with noise level might arise because of the reduction in sensitivity described by the equation. In other words, if the pitch judgement mechanism could detect only a certain minimal change in the discriminator output, then with added noise, a larger input frequency change would be necessary to maintain a just detectable output. These points will be discussed more fully in subsequent sections.

[+] In fact, the effect of noise on the discriminator output would be twofold: the reduction in mean value (according to Eq. 5), and an increase in variance. It is possible that the increased variance of the discriminator output due to the noise might add to the basic internal noise of the discrimination process itself, necessitating an even larger input d_f. This second factor was ignored, because the output variance of an FM discriminator changes with input noise level only at the 'threshold' (i.e., at 0 dB and above). Below the threshold, the output noise level becomes constant. Hence, this would cause only a shift of the DL curve on the S/N axis. In any case, ignoring this effect meant only that the results would perhaps indicate a lower bound for subjective performance, rather than some true value.

VII.2.4. Decision Making

The discussion in the previous sections suggested preliminary signal transformations of filtering and zero-crossing detection, but left completely open the mechanism by which decisions were made. Decision making, according to the proposed model, would have to involve the comparison of zero-crossing counts for the A and X pulses. A 'decision' of U or D would then be made according to the sign of the difference. Further, it was necessary to include some internal random element to limit accuracy in order to account for the imperfect responses given by subjects even in the absence of external noise.

The non-linearity involved in the neural transduction process meant that any internal noise in the model must be assumed to lie at the neural level. If the constant random element were in the linear (receptor) stages of the system, this would mean that the SN ratio (and hence, the DL) would vary with signal level in the absence of noise. Subjective data show that this is not so (Shower & Biddulph, 1931). It was assumed that the noise would result in a given fixed change in the discriminator output being indicated correctly only with a given probability. As will be shown later, this was adequate to allow comparison with subjective data.

VII.3. The Complete Model

A block diagram of the complete model is shown in fig. 36, and consists of the Gradient Filter, a zero-crossing threshold detector, a pulse generator (which represents the initiation of impulses by the hair cells), and a frequency counter which indicates the number of zero-crossings in a unit time. The remainder of the model was not constructed, but was assumed to operate by comparing the zero-crossing counts of the A and X pulses. Some indication of the behaviour of the early stages of the model is shown in fig. 37, which shows the response of the 'middle ear', 'basilar membrane', and zero-crossing detector to a short 'sound' impulse of electrical origin.

The model was assumed to operate in the following way: in the absence of external noise, the DL would be determined by the 'noise' inherent in the decision maker, with the result that a given input frequency change (d_{fo}) would be correctly indicated only with a given probability. (Presumably, the relation between the d_f and probability would be of the CN type, mentioned on p. 66.) The effect of external noise would be to reduce the sensitivity of the discriminator, according to Eq. 5. In order to maintain the same (just detectable) mean change in its output, it would be necessary to increase the input d_f at least to d_{fN}, where:

$$d_{fN} \geq d_{fo} / 1-\exp(-S) \qquad (6)$$

In other words, Eq. 6 would describe the relation between the DL and the noise level.

The complete model was tested by measuring the mean value of the discriminator output under conditions exactly comparable with the subjective experiments.

VII.3.1. A Test of the Complete Model (Experiment VI-B)

VII.3.1.1. Introduction

It has been suggested in previous sections that pitch discrimination might be based on the temporal analysis of the properties of a broadly filtered version of a sound stimulus. Although it is not known what temporal properties are preserved for transmission along the auditory nerve, it is probable that some information on zero-crossings is present. Considering only this feature allowed a parallel to be drawn between the operation of the hair cells and zero-crossing detection as performed by an FM discriminator. It was suggested that the increase of DL with noise power density found in Experiments IV and V, might be due to the known reduction in sensitivity of such a discriminator as noise level rose. Experiment VI-B was designed to test this possibility by analysing the same stimuli used in the subjective experiments with the Model Ear system.

VII.3.1.2. Methods

Signal and noise were generated in exactly the same way as in the subjective experiments (see Experiment V, p. 144, except that a single high signal level was used). After transmission through the Artificial Ear to allow for the frequency response of the earphones, the sounds were applied to the Model Ear. Measurements were made at all signal frequencies (0.5, 1, 2, and 3 KHz) using both wide and narrow band noises (as in Experiment V). The threshold detector (see fig. 36) generated a short impulse whenever a positive going zero-crossing occurred at the Model output (see fig. 37). The signals here were continuous, and the counter indicated the mean number of zero-crossings per second.

If an input frequency shift of d_{fo} were used, the change in the mean zero-crossing rate in the absence of noise would, of course, be d_{fo}. With added noise, the difference would be less than this because of the reduced discriminator sensitivity. If d_{fN} was the difference in the presence of noise, then the ratio: $\underline{d_{fo} / d_{fN}}$ could be used as a measure of the increase in d_{fo} necessary to maintain a constant output change (i.e., the 'DL' of the model increased by this ratio). This method specified the mean value of the discriminator output as a function of noise level, and implied that the increase in variance of the discriminator output at higher noise levels was not significant when compared with the variability inherent in the decision process (see footnote on p. 171).

<u>Method of Analysis</u> The values of d_{fo} / d_{fN} obtained for the model varied between 1 at low noise levels, and some value >1 at high noise levels. These measured ratios were normalized, using the noise-free DL's measured for the subjects at each frequency. This was done in the following way: the subjective data were first plotted, and curves derived from Eq. 6 were drawn to give the best fit by eye. (The ordinate of the horizontal portion of the curves was always close to the average value of the noise-free DL, though drawing curves through all the points avoided absolute reliance on the single noise-free measurement.) The value of the 'noise-free' ordinate was used as a multiplier for the d_{fo} / d_{fN} ratios measured for the model; these were then plotted on the same scale. (Since both axes of the curves were logarithmic, normalization involved merely a constant shift of the plotted points.) The effect of this procedure was to equate the model results with the average subjective DL's, under noise-free conditions.

VII.3.1.3. <u>Results and Discussion</u>

The averaged values of d_{fo} / d_{fN} are plotted, separately for each frequency, in figs. 38a to 38d as 'DL' against SN ratio. Subjective data from Experiments IV (for wide band) and V (for narrow band) are included for comparison. (Experiment IV data were given here, as there was more information on the effects of wideband noise. As already shown, on p. 148, the Experiment V results would overlap these. Only the data for the high signal level (50 dB) are shown.)

Model Results

For the narrow band noise, where filtering had little effect on the SN ratio, it was expected that Eq. 6 should predict the results after allowing for the normalizing shift. This was found to be the case. The lines drawn for the narrow band model data are simply vertically shifted versions of Eq. 6 with negligible horizontal adjustment. For the wideband data, the SN ratio at the discriminator input was effectively increased by the filtering process, and therefore the data fell at a lower value of S/N. A curve derived from Eq. 6 was adjusted horizontally to give the best fit by eye.

Table 29 shows that, as with the subjective data, the separation between the wide and narrow band curves decreased with frequency; this would be consistent with an increasing filter BW.[+] The model and subjective figures were roughly parallel; BW's were estimated in the same way for both, and are generally wider for the model by a factor of two or three.

[+] A comparison can be made between the separations measured in this experiment and the S/N reductions measured for the Gradient Filter in VI-A (see table 28); the agreement is fairly close. This was a useful check, since Eq. 6, from which all the curves were derived, depends on the assumption of a noise spectrum which is fairly symmetrical about the signal frequency. The agreement between the two sets of figures suggested that the approximation was good enough.

Comparison between Subjects and Model

1. The Shapes of the Curves The Eq. 6 curve fitted the wideband subjective data quite closely at all frequencies. However, at all frequencies except 500 Hz, the narrow band data might be better described by a curve with a point of inflection about 2-3 dB wide at an SN ratio of around +4 dB. When this effect was noticed, re-examination of the data of Experiments P-1 and I (see figs. 26 and 28b) also showed a discontinuity between 20 and 40 Hz BW, where the SN ratio was also between +6 and +3 dB.

Greenwood (1961) has described what may be a related effect. In measuring masked audiograms for 1 and 3 KHz tones, he found that for narrow band noises (only) the threshold level of the tone was a discontinuous function of the noise power. The threshold signal level increased in proportion to noise level until a noise SL of 40-60 dB was reached.[ǂ] At this value, the threshold abruptly decreased by about 3 dB before continuing to increase proportionately as the noise level was raised further. The same effect was not found for BW's wider than the CBW (see also Campbell, 1964). There is also related evidence from Ward (1960) who found a critical noise level below which recovery from noise-induced TTS proceeded as if the noise had been removed entirely, but above which recovery time was increased.

[ǂ] 50 dB would correspond to 0 dB S/N at the high signal level in Experiment V. It is interesting that while this effect can be seen in the 1, 2, and 3 KHz data at this level, it is present only at the low level at 2 KHz (see fig. 29c). This would be consistent with Greenwood's finding that absolute noise level, rather than SN ratio was the relevant variable.

Greenwood suggested that this discontinuity might reflect the sound level at which basilar membrane movement becomes large enough to stimulate the inner hair cells. Although this is possible, it does not by itself account for the difference between the effects of wide and narrow band noises.[+] Possibly the CBW reflects neural connections to a contiguous group of inner hair cells lying in a restricted region of the membrane. The activity of this group might have different effects depending on how the noise stimulated similar groups nearby; i.e., on whether the noise was above or below the CBW. (This type of effect might be more likely to occur with the inner hair cell system which is known to be more regularly innervated than the outer system.)

This threshold effect with narrow band noise has not been widely reported, and it could be more than coincidence that both Greenwood and the writer used what was essentially an identical method of noise generation (see p. 24 and section III.2.). As mentioned in Chapter III, the noise generation system produced a random signal which was different from bandpass filtered white noise, although a comparison experiment (see p. 82) suggested that the masking effects of the two types were similar. However, this point might warrant further investigation.

One reason for the different published values of subjective CBW's, as well as for the difference between the model and subjective results quoted here, might follow from the possible threshold effect found for narrow band noise. Using the technique outlined in

---

[+] There is a suggestion however, that for the wideband data for 2 and 3 KHz, a similar threshold effect might also be present, but the effect is less marked.

Experiment V, BW estimates might be expected to vary by the width of the threshold step (3 dB, i.e., a factor of 2), depending on whether the separation between the wide and narrow band curves was measured above or below the threshold point. There is obviously insufficient data to make this measurement worthwhile here, but it is possible that an effect like this, if unrecognized, could account for the discrepancies between published CBW estimates, which always assume a smooth relationship between noise power and masking.

2. The Location of the Narrow Band Curves According to this model, the location on the S/N axis of the narrow band curves should be predicted by Eq. 6, since the filter would have no effect on the SN ratio. Thus, the narrow band model data which, as already shown, were closely fitted by Eq. 6, provided a standard of performance against which the subjective data could be compared. Any differences between the curves could be ascribed to differences between the real and model discrimination processes. (The wideband curves could not be compared in this way since their location would depend on two factors: the behaviour of the discriminator, and the effective BW of the filtering process. Note that the wideband subjective curve changes its position more rapidly with frequency than the narrow band curve, from lying about 2.5 dB to the left of the model at 500 Hz (fig. 38a), to lying about 4 dB to the right at 3 KHz (fig. 38d).)

The subjective narrow band curves were always to the right of the model, and the separation between them increased with frequency (from about 2.5 dB at 500 Hz to about 6 dB at 3 KHz). The size of the separation depends partly on the normalizing factor used for the subjective data (see p. 175), but even so, its variation with frequency seems fairly regular. There might be two reasons for this: first, the results of Experiment VI-B were obtained as mean DL values, and the assumption was made that the noise level itself did not contribute a significant variance. This assumption was probably not strictly correct, and could well account for some separation between the curves. Second, it is tempting to speculate that the increased separation with frequency might be a result of the Volley Principle (see pp. 160, 170). If the temporal information required by this model were to be carried by the auditory nerve, groups of fibres firing in rotation would have to be involved. Since the refractory period of an individual fibre would prevent it from transmitting frequencies above about 500 Hz, the size of the group would increase with frequency. This would increase the number of transmission paths involved and hence, the variability of the transmitted time information. At the same time, high frequency tones stimulate a more restricted region of the basilar membrane, making fewer paths available. That is, with increasing frequencies the real ear would tend to become less and less like the model ear, due both to the increase in variability of transmitted timing information and to the reduced possibility of averaging this out over a number of different paths.

3. Subjective and Model Bandwidths Table 29 shows BW estimates for model and subjects using the method outlined in Experiment V (p. 144), but measuring separation between the curves, rather than between straight lines. (This allowed better use to be made of all the data.) These estimates were very close to those shown earlier in Experiment V (cf. table 27). The subjective and model data show some interesting parallels. The model BW's are about 2-3 times those of the subjects for each frequency. (A factor of about 2-3 in BW seemed small in view of the simplicity of the model.) An agreement of this order

suggests that the filtering effects observed with noise masking were cochlear in origin and further, that it is unlikely that pitch discrimination operates via the detection of changes in the output of an extremely narrow band filter. Analysis of the temporal properties of a much less selective filter seems a more likely mechanism.

Although BW's roughly double with increase of frequency (table 29), subjective BW's increase more rapidly with frequency than model BW's. It is to be remembered that the model represents an ideal system, and one factor which has been omitted is that timing of nerve impulses will become less precise at high frequencies. Possibly, the real ear combats the increased variability by attempting to average over more neural channels as frequency increases. This might effectively widen its BW relative to that of the model.

VII.3.2. Evaluation, Summary and Conclusions

The object of this chapter was to see whether the effects of noise on the DL could be explained by using known auditory physiology. Using a simple electronic analogue of the frequency-dependent behaviour of the middle ear and cochlea, it was possible to account to a large extent, for the stimulus filtering properties inferred from the noise masking experiments done with subjects.

The model contained a filter with an effective Q-value of the order of 2, which represented a simple relation between the mechanical movement of the basilar membrane and the excitation of the hair cells. It was argued that the process of transduction of membrane movement into nerve impulses was analogous to the operation performed by an ideal FM discriminator, and it was suggested that the essential information on the pitch of a sinusoidal signal was contained in the temporal properties of the nerve impulses initiated by the zero-crossings of basilar membrane movement. It was further suggested that judgement of the relative pitch of a pair of signals might involve a comparison of corresponding zero-crossing counts. It was shown that a model, operating on these principles, could account for the effects of noise on the DL for pitch measured psychophysically.

One important implication of this model is that all stimuli go through a fairly similar type of preliminary filtering which is due to the mechanical properties of the ear itself. This would certainly be consistent with what appears to be the very fundamental nature of the Critical Band concept, which has been shown to be applicable in many different perceptual situations. It also means that more attention might be devoted to the real frequency characteristics of the filtering process, since any selectivity based on the behaviour of the basilar membrane is probably neither very narrow nor very symmetrical.

The chief feature of the subjective data which required explanation by any acceptable model were: the apparent existence of a linear filtering process which was tunable to the signal frequency, and a DL which was proportional to the S/N (power) ratio. The IPD model as originally proposed, did not fit the results very well. It is very difficult to design any finite set of experiments to either prove or disprove all features of a model. However, the reasons for preferring the present model were its consistency with known physiological phenomena, and its simplicity. On these grounds, as well as its better fit with the subjective data, the model discriminator proposed here was preferred to the IPD.

In sum, it is felt that the work presented here illustrates a profitable approach to auditory modelling: a compromise between mathematical elegance and reality.

CHAPTER VIII. SUMMARY AND SUGGESTIONS FOR FURTHER STUDY

Introduction The work in this thesis can be divided into three separate parts: first, the preliminary experiments, which tested the use of a Sequential method for measuring pitch discrimination; second, the main experiments, which tested the effects of additive noise on pitch discrimination as well as the application of a Critical Band concept to discrimination; third, the suggestion of a physiologically oriented model, which accounted for the findings of the main experiments. This chapter is intended to provide a concise summary of the thesis and at the same time, to indicate a few lines along which further development might occur. A full discussion is not included, as the implications of a number of points have already been discussed in context.

VIII.1. The Preliminary Experiments

VIII.1.1. Summary

Discrimination of pitch was specified in terms of a judgement curve of the cumulative normal type, and it was found that the Wetherill Sequential method gave efficient estimates of the parameters of this curve with very little effort. The technique had the further property of making a test of equal difficulty for all subjects.

The method was assessed by comparing it with Probit Analysis, which is the optimal estimation procedure for CN curves. It was found that DL estimates obtained by the two methods were essentially identical, although estimates of the midpoint (M) tended to be more variable when a Sequential rule was used. The higher variability appeared to be due to a very large serial correlation effect which was found for the midpoint data. The correlation seemed to arise partly from the statistical properties of the rule itself, and partly from the behaviour of the subjects.

An unexpected finding of the preliminary experiments was that some subjects could have a non-zero midpoint. This was not due to a particular test method, since it was found with both Probit and Wetherill techniques. Some possible reasons for this were discussed, but the matter was not pursued further.

VIII.1.2. Suggestions for Further Study

VIII.1.2.1. The Wetherill Sequential Technique

The implication of the preliminary experiments was that the behaviour of the Sequential rule was not quite as predicted by simple theory, and that there was some interaction between the rule and the subjects. M data showed this effect strongly, but even in the case of the DL, consistent differences were found between estimates from randomly and from sequentially presented stimuli. There were three particular points which would be worth further study, and which might clarify any interaction between subject and strategy.

1. The Use of a Reference Sound This problem has not often occurred in the past, since the most extensive application of Sequential methods has been in signal detection, where reference sounds are not normally used. The only previous use of the Sequential method in a discrimination situation was by Levitt (1964), but even here reference sounds were not always present. (The subject's task was to judge the intercranial position of a binaural sound image relative to the centre of his head.) With an A-X discrimination task, there is the problem of

how the reference sound (A) is to be presented. Here, the frequency of A was kept fixed, and therefore did not always lie centrally in the range of X frequencies used. Possibly, bias might be reduced if the frequency of A was varied so that it was always in the centre of the range of X frequencies used. This could easily be done with the Wetherill method, by making running estimates of a subject's midpoint, and using these to correct the starting value of A. Alternatively, A could be varied at random from presentation to presentation, either about a fixed mean value, or about the trend mentioned above. Possibly, one or the other of these techniques would prevent a subject from building up a bias based on a number of previous presentations (see detailed discussion in section V.5.3.). (Incidentally, the Sequential rule could be tested in a reference-free discrimination situation with a subject who had absolute pitch. One could present single X sounds and ask, for example: "Is this higher or lower than middle C?". It might be interesting to see how the Sequential rule behaved in this situation.)

<u>2</u>. <u>Level Spacing</u> Wetherill's work showed that increased estimation efficiencies could be obtained by changing the level spacing used by a Sequential rule during the course of a test. This problem was not considered here, but was left at the stage of finding a single level spacing for general use (Experiment P-2). It was shown that a rather narrow level spacing (approximately 1 DL) was most suitable from a statistical point of view, but this meant that the changes from one presentation to the next tended to be small, which may have been difficult for the subjects. It may well be that the use of variable level spacing could have significant subjective effects, as it would give a means of varying the difficulty of a test, while keeping the advantages of the Sequential rule.

3. Inclusion of Irrelevant Stimuli The experimental technique used throughout this thesis was a very simple one, involving two randomly interleaved Sequential strategies; thus, every response by the subject had some influence on subsequent presentations. One consequence of this has already been mentioned: the operation of the strategies would attempt to force a subject to give about equal numbers of U and D responses. A simple extension of this method might be to present occasional A-X pairs with a d_f value corresponding to the estimated midpoint. Responses to these stimuli would not be used in the structuring of the Sequential rule, and would reduce the rather rigid control of response frequency imposed by the two complementary interleaved rules. Some evidence that this type of procedure might be effective was found in Experiment P-4, but the point seems worth further study.

VIII.1.2.2. Subject Bias

As indicated, some subjects showed a strong positive bias. One immediate question is, how far was this a consequence of the use of a Sequential strategy? This might be answered by some of the experiments suggested above, or an alternative might be to take a simulation approach. For example, the output of the zero-crossing detector of the Model Ear could be fed into a computer programmed to make 'pitch decisions' by comparing counts. Constraints on judgements could be introduced in ways which might correspond with how a subject appears to operate. For example, what would happen if the computer attempted to maintain a roughly equal number of U and D judgements in the short term? Or, if X were not judged relative to A on each occasion, but relative to the previous X?

This judgement simulator could also be used to test whether the use of a forced choice method (as used here) might have disruptive results. This might have some bearing on the occasions in a test where a subject is uncertain, and will guess in a characteristic way. It would be possible, for example, to introduce biased guessing when the pulse counts of a particular A-X pair differed by less than a critical amount, and to see the effect of such a strategy on the operation of a Sequential rule. The behaviour of such a simulation system might suggest more suitable test methods for final evaluation with human subjects.

The finding that biases could also appear when stimuli were presented in random order (Experiments P-1 and P-4), and yet were not found with automatically generated responses (P-3), suggested that factors <u>other than the test methods</u> were involved. Where subjects were biased, these biases sometimes became larger as noise level increased. Figs. 34 and 35, which describe the combined response curves of the basilar membrane and the middle ear, suggest a possible reason for this. Since the filter sensitivity falls more quickly in the high frequency direction, this would mean that an upward frequency change would lead to a slightly greater reduction in signal intensity than an equal downward frequency change. Thus, SN ratio might change with frequency in a non-symmetrical way, falling more rapidly as the signal frequency was increased than as it was decreased. (This is equivalent to saying that the frequency changes in the upward direction would have to be larger than those in the downward direction in order to give the same probability of detection, and would lead to a positive bias.) Of course, this argument is based on many assumptions of the relation between the lumped constant filters of the Model, and the behaviour of the real basilar membrane and hair cells. However, it could be tested by introducing non-uniformities into the noise spectrum to see whether these bias effects could be manipulated.

VIII.2. The Main Experiments

VIII.2.1. Summary

The main experiments measured the effects of additive noise on pitch discrimination with the object of evaluating an optimum model (the Ideal Pitch Discriminator, or IPD). The IPD model made various predictions which did not fit the results of these experiments. However, this did not necessarily mean that a Decision Theory approach was inappropriate. A point that has often been disregarded is that Ideal Observer models usually involve two stages: first, the processing of an input signal to give the quantities required for a decision; and second, the decision process itself, which is quite separate. The main experiments were designed to test only the first part of the IPD model, which was concerned with the processing of the original signals. The lack of correspondence between the predictions of the IPD and the subjective data was not surprising, as signal processing in the IPD was defined in a mathematically optimal way, requiring operations which there is no reason to believe that the ear can perform.

It was found that the masking effect of noise appeared to be limited by a preliminary filtering process, in that an increase of masking noise above a certain BW had no effect. This was similar to results found in the classical Critical Bandwidth (CBW) experiments, but had not previously been shown in a pitch discrimination task. This interpretation was confirmed by the results of introducing a gap into an otherwise continuous noise spectrum, as well as by moving the signal relative to the noise on the frequency axis. Finally, noise

BW and power density were varied in a complementary way, and it was found that for BW's wider than about 100 Hz, power density had a much more significant effect than BW. Below 100 Hz, power could be varied with equal effect by changing either BW or power density.

Later experiments measured the relation between the DL and SN ratio for different noise BW's. It was found that the relation could be described by a curve of standard shape whose location on the S/N axis depended on noise BW. This was again interpreted as the result of a filtering process. A filter would respond to the total energy of signals whose spectra were narrower than its own BW, so that for any narrow band noise, S/N alone should be sufficient to describe its masking effect. (I.e., the effective S/N at the filter output would be similar to that measured externally.) For wide band noises, the output noise level would be less than the input, and therefore curves plotted in terms of external SN ratio would be displaced, due to the filtering action. It was proposed that the distance between the wide and narrow band curves be used as a measure of the effective aural BW. Although this method estimated effective rectangular BW, it depended neither on assumptions of the shape of the aural filter, nor on the definition of an arbitrary breakpoint in a continuous curve, and therefore seemed to be an improvement on previous methods of estimating the CBW.

The aural BW's estimated using this method tended to be larger than values obtained from detection experiments, although they were of the same general order. CBW's have not before been estimated in a discrimination situation, and the discrepancies might indicate that aural BW varied according to the task at hand. In a sense, the CB concept itself implies that the ear is able to adjust one of its parameters to a signal (i.e., the centre frequency), and it might be that where signal frequency is uncertain, as in pitch discrimination, other parameters would be adjusted, leading to a wider aural BW.

VIII.2.2. Suggestions for Further Study

VIII.2.2.1. The Variability of the CBW

The experiments in this thesis imply that the ear does not possess a precisely adjustable filter, as suggested by the IPD for example, and it seems more likely that the finest frequency resolving unit is the CB. Thus, if the parameters of the ear are indeed adjustable, then the CBW presumably represents the lower limit of such adjustment. Any adaptability might occur in two ways: either by variation of the BW of a single filter (upwards from this lower limit), or by the combination of a number of separate CB's. Some experiments (see e.g., Marill, 1956; Creelman, 1959) have suggested that the ear may operate on a scanning principle, concentrating on only one CB at a time. Other work (Green, 1957) has suggested that CB's may be combined simultaneously.

An immediate extension of the experiments in this thesis would be to use signals of shorter duration (and thus, broader BW's). It might be expected that a repetition of experiment V under these conditions would lead to wider BW estimates. Another possibility would be to measure the discrimination of pitch for harmonic signals in the presence of noise. Selective masking could show which components were used, and also whether attention could be switched from some components to others to combat masking in the most effective way. (This would extend the work of Flanagan & Saslow, 1958, and Schodder & David, 1960, who also studied harmonic signals, but without masking.)

A different view of the CBW might be obtained in a situation involving frequency uncertainty. Greenberg & Larkin (1968) have shown that where the frequency of a signal is varied at random, its detectability is constant as long as the range of variation remains within a CB; for wider ranges, detectability is reduced. It would be interesting to see whether similar findings would apply in a discrimination situation, where the frequency of the A signal was varied at random from presentation to presentation. The informal experiment described on p. 119 suggested that such random variation would have no effect, but it might be worth repeating under a wider range of conditions. If the BW of the ear were adjustable, then it might be wider in this situation than when the reference frequency was constant.

VIII.2.2.2. The CB in Discrimination

1. A Method of Estimating the CBW The CBW derived from Experiments P-1 and I was obtained by the 'classical method' of finding the breakpoint in a curve relating a signal parameter to the noise BW, and was within normal limits for other tasks. However, the CB estimates obtained by the method proposed in Experiment V tended to be slightly larger, particularly at high frequencies. It might be of interest to apply this technique directly to a detection task (by plotting the threshold signal against noise level for wide and narrow band noises) to see how far the larger BW's found in discrimination are due to the estimation method itself, and how far to the difference in tasks. It is felt that this method could be used with a wide range of tasks involving masking, and might give a more universal and less arbitrary way of measuring CBW's.

2. The Width of the CB in Discrimination Of the four estimates of the CBW made in Experiment V (see table 27), those at the three lowest

frequencies (0.5, 1 and 2 KHz) agreed fairly closely with previously published values obtained in different situations. The most obvious discrepancy between the estimates in table 27 was the value at 3 KHz. This might be accounted for physiologically, in that at high frequencies there is greater variability within a single channel, and at the same time, there are fewer channels over which this variability can be averaged out (see p. 180). The interesting feature of a noisy situation is that it introduces a conflict: in effect, there are two 'noises' to be considered, an internal one due to the timing variability, and the imposed external one. The methods of dealing with these might be directly contradictory. (Maximum reduction of internal noise would be achieved by averaging the activity of all neurons responding to the signal. This would mean however, the inclusion of activity from a fairly wide region on the basilar membrane, which would increase the effects of the external noise.) Thus, the effective BW at high frequencies might be a compromise between a wide value (to accomplish the maximal amount of averaging) and a narrow value (to exclude the maximal amount of noise).

An interesting experiment would be to measure the CB using both band-limited noise and wideband noise with a gap, centred on the signal frequency in each case. If the above argument were true, it would be expected that at low frequencies, CBW estimates with each type of noise should be similar; while at high frequencies, wider values should be obtained with gapped- than with bandpass-noise. Further, if it were possible to measure a CB effect in a noise-free situation (by introducing frequency uncertainty, for example), then one would expect that at low frequencies, the BW estimates should be similar to those obtained using noise. At high frequencies, this would not be so, and the BW estimate using noise should be less than the noise-free one.

VIII.3. The Proposed Model Ear

VIII.3.1. Summary

A physiologically oriented model of pitch discrimination was proposed, and an electronic model was constructed to see whether the effects observed in the main experiments could be accounted for by what is known of the behaviour of the peripheral ear. Included in the electronic model were circuits which simulated the frequency dependent behaviour of the middle ear transmission, and of selected points on the basilar membrane. It was suggested that nerve impulses might be initiated either by the displacement of the basilar membrane or by its gradient. It was found that the 'gradient' model came closer to accounting for the CBW effect found with subjects.

A parallel was drawn between the initiation of nerve impulses and the zero-crossing detection performed by an FM discriminator. Curves describing the reduction in sensitivity of such a device by additive noise, were of the same shape as the subjective data. There was no evidence that the effective BW of the ear was very narrow in a pitch discrimination task. Thus, explanations of pitch discrimination in terms of the detection of changes of energy at the output of a selective filter were not satisfactory, and an alternative process was suggested involving temporal analysis of the aural filter output.

The decision process in the proposed model was assumed to involve a probabilistic comparison of zero-crossing counts, as it was necessary to account for the imperfect decisions given by subjects to noise-free signals. A significant feature then, was the placing of the internal noise at the decision stage; this avoided a difficulty sometimes encountered, that internal noise must vary with signal level. It was found that the predictions of the complete model paralleled subjective data fairly closely.

VIII.3.2. Concerning Models of the Mechanical Behaviour of the Ear

VIII.3.2.1. Middle Ear and Basilar Membrane Behaviour

The first property of the Model was the filtering due to the middle ear and basilar membrane (see figs. 35 and 36). The Model suggests that the rather broad stimulus filtering necessary to explain the results of these experiments could take place in the cochlea. Indeed, this level of the system has been shown to predict other effects involving the perceptual aspects of transient stimuli, notably some of their pitch qualities (see p. 43).

Some investigations have suggested that the simple view of the basilar membrane presented here is incomplete. For example, measurements of the frequency sensitivity of single neurons in the auditory nerve has shown a much more critical frequency dependence than would be predicted by this Model (e.g., Katsuki, 1961; Weiss, 1964). More recently, de Boer (1967) used a correlation technique to infer the effective BW of a point on the basilar membrane from the relation between a random noise stimulus and recordings of the auditory nerve. Although his BW estimates were of a similar order to those of the Model proposed here (see table 27), he found evidence of much steeper cut-off slopes. It should be noted that this Model makes a gross approximation to the real ear by considering only discrete points in what is actually a continuous system. Thus, it makes no allowance for the existence of travelling waves in the real ear, and of the way in which these may influence adjacent points to the ones considered.

The true properties of the ear can be found only from detailed physiological measurements. This would require however, a very precise experimental technique, and it is not yet possible to study the auditory nervous system in a stage-by-stage way which would be required for detailed evaluation. It is to be noted that most models of the mechanical behaviour of the cochlea, including the present one, are based on the experiments of von Bekesy, some of which were done 35 years ago, and which have not been verified in detail. In the last few years, workers have begun to use new methods of observing small movements (e.g., Johnstone & Boyle, 1967), and answers will have to wait until these techniques are perfected.

VIII.3.2.2. The Initiation of Nerve Impulses

The second property of the Model was the initiation of nerve impulses by the zero-crossings of the basilar membrane movement (the threshold detector in fig. 36). This is the means by which temporal information is supposedly transmitted. The synchronization between nerve impulses and periodic stimuli has already been mentioned (see section VII.2.3.2.), and formed the basis for the proposal of this Model. The work of de Boer, mentioned above can be taken as indirect evidence of such synchronization at frequencies perhaps as high as 5 KHz. (Correlation between neural activity and the originating sound would not exist unless time synchrony were present.)

There is other evidence that this view of the initiation of nerve impulses as always occurring at the same fixed level of basilar membrane displacement is too simple. One extension of the Model is suggested by the work of Weiss (1964), who compared the statistics of impulses initiated in a model neuron with physiological data.

He found that by including some type of non-linear transducer (to represent the hair cell behaviour) together with a variable threshold (to represent the refractory behaviour of nerves), he could obtain fairly close parallels with the behaviour of real neurons. If the present model were extended in a similar way, by using a non-linear function of basilar membrane gradient to initiate nerve impulses, one consequence would be a change in the effective BW. This might make for closer parallels between model and subjects.

The present Model suggests some further neurophysiological experiments which might test the proposals made here. Briefly, these would all be based on the hypothesis that the statistical properties of single unit recordings, in response to a sine wave in the presence of random noise, should be describable in terms of FM receiver theory. Namely: 1) there would be some decrease in variance with increased signal level, 2) with an increase of noise level, there would be an increase in time variability up to a certain point only; beyond this, variability would be independent of the signal, 3) for a given input frequency, the change in mean impulse rate should reduce according to Eq. 6 (on p. 173).

It seems very likely that information in the auditory nerve is carried in the form of threshold crossing signals of one sort or another. This type of signal however, is difficult to treat theoretically. The important point here is not whether the Model is accurate or not, but whether it makes definite predictions which can be tested.

VIII. 3.2.3. Discrimination

The third, most physiologically speculative stage of the Model, suggests that pitch perception involves only the analysis of zero-crossing information (the counter in fig. 36). Although it fits the subjective data fairly well, the Model does not throw much light on how discrimination might actually occur. The problem is not alleviated by the comparative lack of physiological data on the stimulus transformations which occur at higher levels of the auditory system, and it is here that the analogy between the Model ear and the real ear probably breaks down completely. The counter used in these experiments involved a reference clock which was used to time the counting of impulses. There is no evidence that the nervous system has a clock, or counts impulses with any precision.

A more plausible mechanism for the analysis of zero-crossing information might involve the mapping of temporal into spatial dimension, following the original proposal made by Jeffress (1948) (see also Licklider, 1959). Essentially, this would involve chains of neurons with each neuron introducing a characteristic small delay. It would also involve systems for comparing the delayed and undelayed signals for coincidence. Coincidences would tend to occur at values of delay corresponding to the period of a stimulating signal. In fact, this mechanism is analogous to the operation of autocorrelation (Cherry, 1961; see also p. 36). There is some evidence that systems like this do indeed exist in some parts of the brain (e.g., Freeman & Nicholson, 1970). The structure of the nervous system makes it perhaps more suited to perform complex discriminations in the spatial rather than the temporal domain, as it is remarkable for its complex connections, rather than for its precise timing. Obviously what is needed is some way of tracing how this type of information is used by more central parts of the auditory nervous system, and whether there are any mechanisms which might perform the kinds of analyses suggested.

Some experiments by Pollack (1968) appear to contradict this Model. He found that 'jitter' (i.e., random displacement of various kinds) imposed on a periodic pulse train tends to be undetectable at low frequencies. This appears to contradict a temporal interpretation of pitch discrimination, because it is at low frequencies that timing information should be most precisely transmitted. There are two points to be made: 1) How would the present Model process Pollack's signals? 2) How do the auditory nerve recordings change when jittered, rather than exactly periodic, stimuli are used?

Obviously, this model as it stands, would not predict all pitch effects. There are some signals which appear to be associated with the envelope properties, rather than with fine waveform structure. This would not be difficult to allow for, by postulating some type of envelope detection at the cochlear level, followed by encoding of the zero-crossing properties of the envelope. It may be that the viscous properties of the tectorial membrane might cause some hair cells to be stimulated more by the envelope structure than by the fine structure of basilar movements (Davis, 1959). Again, a more realistic model of nerve impulse excitation, such as that used by Weiss (p. 195) would account for some of these effects.

VIII.3.2.4. Decision

The fourth stage of the Model involves a decision based on those features of the stimulus which are encoded by the preliminary operations of filtering and zero-crossing detection. It is felt that at this stage mathematical modelling should start. The conclusion drawn from the present work is that there is little point in mathematical elegance for its own sake, and that, where possible, models should be based on any relevant knowledge of how auditory stimuli are processed by the ear. Starting with this limitation, it then makes more sense to develop optimal models, since there is no doubt that the brain as a whole might be able to operate in a very adaptive way. For example, Siebert (1965) derived an optimal model for detection and discrimination of signals coded in terms of auditory nerve impulses. He showed that the counting of zero-crossings was in some sense optimal, and showed parallels between the operation of his model and certain psychophysical phenomena.

A difficulty with constructing models of brain function is that they have to be unspecific in their location, and independent of the reliability of individual components (since, for example, neurons are continually dying). A more fruitful approach may ultimately be with methods which attempt to describe the operation of a whole system probabilistically, in terms of the operation of unreliable single components.

VIII.4. Concluding remarks

In the past there have been three separate approaches to the problems of audition: that of the physiologist, that of the engineer, and that of the psychologist. Physiologists, although working with the real system, have been hampered by technical difficulties and the use of very simple signals. The communications engineer, on the other hand, has been able to analyse complex signals, but has done so in an unrealistic way, tending to view the ear as a rigid idealized system. The psychologist has effectively ignored the two above approaches, and has tended to over-concentrate on stimulus-response experiments, describing the intervening relation in an idealised 'black box' way, if at all. In the words of the immortal bard:

"And so these men of Hindostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right
And all were in the wrong."

(from "The Blind Men and the Elephant"

by John Godfrey Saxe)

There have recently been fewer signs of the previously drawn barriers between these three fields, and auditory theory has reached a very exciting stage because of this. It is felt that the most promising future of the development of the study of hearing is an interdisciplinary one, combining the unique skills that each approach has to offer. It is hoped that the work in this thesis has made some contribution to this.

TABLES

Table 1. Wetherill Sequential Rules for U' and D' Events.

Entry No.	Response Type		Expected Central Tendency[+]
	U'	D'	
1	U	D	d_f: 0.50
2	UU	D, UD	d_f: 0.71
3	UUU	D, UD, UUD	d_f: 0.79
4	UUUU	D, UD, UUD, UUUD	d_f: 0.84
5	UUUUU	D, UD, UUD, UUUD, UUUUD	d_f: 0.87

[+] Value of d_f for which $P_{U'} = 0.5$

Table 2. Probit Analysis of Random Data (Experiment P-1).

a. DLs

Subj. / BW	LM		PB		AHM		KP		MEAN
10	1.7	1.9	4.4	7.1	3.2	3.6	2.3	2.8	3.13
20	1.9	1.7	5.1	4.6	3.8	2.1	4.2	5.8	3.42
40	2.5	3.1	6.6	5.5	4.0	6.4	4.4	6.7	4.68
80	5.1	4.6	12.4	5.7	5.9	3.9	7.9	4.3	5.86
160	9.0	6.7	26.6	21.0	8.4	7.2	33.0	8.5	12.58
320	7.0	10.8	22.3	14.7	6.1	5.0	14.3	9.7	10.14
640	9.9	10.3	17.6	16.0	32.1	18.5	19.1	18.5	16.74
1280	8.9	11.9	20.6	21.8	11.7	12.5	8.1	18.2	13.42

Grand Mean: 7.51

b. Midpoints

Subj. / BW	LM		PB		AHM		KP		MEAN
10	0.5	1.5	0	4.6	1.7	2.7	0	-1.9	0.71
20	0.7	-0.1	3.9	1.9	2.8	-0.4	-2.5	-0.6	0.32
40	0.6	0.9	4.1	1.8	3.6	3.6	0	0.9	1.53
80	0.5	2.7	8.2	6.2	1.2	2.5	-0.7	-1.1	1.31
160	-1.3	-1.1	11.1	14.5	0.7	5.5	6.2	-2.2	1.35
320	-1.2	-1.6	0.4	8.4	2.9	2.9	-2.9	0.8	0.42
640	1.3	2.3	4.1	9.8	11.6	2.4	6.0	6.0	4.48
1280	-1.4	1.7	9.6	18.7	2.4	7.1	2.7	5.0	3.43

Grand Mean: 1.39

Table 3. Probit Analysis of Sequential Data (Experiment P-1).

a. DLs

Subj. / BW	LM		PB		AHM		KP		MEAN
10	1.4	2.2	7.7	3.5	5.0	3.7	3.2	2.4	3.33
20	2.2	1.4	4.6	4.1	5.0	3.3	2.9	7.3	3.52
40	3.7	2.6	5.1	5.8	5.9	5.4	3.3	3.3	4.25
80	3.3	5.0	6.7	11.6	6.7	6.2	4.5	11.5	6.44
160	11.2	11.1	23.2	25.4	10.6	8.1	15.8	22.7	14.85
320	5.4	5.6	12.8	26.6	6.9	11.0	10.0	16.5	10.46
640	8.3	11.1	45.4	37.6	9.7	19.1	29.5	15.8	18.68
1280	12.1	19.8	19.5	17.9	14.7	10.9	16.8	29.5	16.95

Grand Mean: 8.14

b. Midpoints

Subj. / BW	LM		PB		AHM		KP		MEAN
10	1.3	3.2	3.3	1.6	2.8	3.4	2.3	2.2	2.43
20	4.1	1.8	1.2	0.5	3.8	-1.3	-3.2	0.8	0.58
40	4.3	1.6	2.1	4.3	5.1	5.5	-5.1	-2.2	1.14
80	3.1	2.3	8.8	6.8	4.4	2.4	1.3	-10.8	1.79
160	5.3	3.4	34.4	12.0	0.7	12.6	-5.7	9.9	4.16
320	-8.3	-5.0	9.2	19.5	7.5	10.4	-0.7	1.2	1.17
640	0.1	9.8	21.0	5.6	8.5	14.5	-4.8	-3.6	2.15
1280	8.1	1.7	21.4	15.7	-0.04	9.6	8.9	2.6	5.55

Grand Mean: 2.05

Table 4. Wetherill Estimates of Sequential Data (Experiment P-1).

a. DLs

Subj. BW	LM		PB		AHM		KP		MEAN
10	0.9	2.7	9.6	1.9	4.3	6.3	4.6	4.1	3.70
20	2.5	1.3	5.0	6.1	5.7	4.8	4.6	9.1	4.43
40	3.4	3.7	6.1	3.4	8.7	4.6	3.3	4.3	4.46
80	5.6	7.0	8.8	18.3	6.9	9.4	8.3	5.6	8.18
160	15.8	15.0	24.6	30.0	14.4	13.7	24.5	35.0	20.43
320	7.5	7.5	16.7	32.5	11.6	16.2	14.2	20.0	14.28
640	13.8	12.5	40.0	46.1	13.1	27.5	38.7	23.7	24.00
1280	18.8	22.8	23.6	19.5	16.3	12.5	25.5	40.0	21.28

Grand Mean: 10.12

b. Midpoints

Subj. BW	LM		PB		AHM		KP		MEAN
10	1.3	2.7	2.4	1.0	3.2	3.2	1.9	-2.1	1.35
20	3.8	1.5	0	1.2	3.5	-0.7	-3.6	0.5	0.48
40	3.6	1.5	1.4	2.5	4.4	3.6	-5.0	-2.2	0.82
80	2.8	1.8	5.9	8.7	1.6	2.2	1.7	0.3	2.48
160	2.9	2.5	30.7	10.0	0.3	10.7	-4.0	12.5	3.57
320	-7.6	-5.5	6.7	16.3	7.5	9.4	0.4	-2.5	0.75
640	0.6	8.8	15.0	-0.5	6.6	8.8	-2.4	-3.2	1.33
1280	6.8	0.1	19.0	13.6	-0.7	8.8	11.1	5.0	4.64

Grand Mean: 1.63

Table 5. Subject Means of Data from Tables 2, 3, and 4.

(Experiment P-1)

Estimation BW	Probit of Random (DL)	Probit of Sequential (DL)	Wetherill of Sequential (DL)	Probit of Random (M)	Probit of Sequential (M)	Wetherill of Sequential (M)
10	3.13	3.33	3.70	0.71	2.43	1.35
20	3.42	3.52	4.43	0.32	0.58	0.48
40	4.68	4.25	4.46	1.53	1.14	0.82
80	5.86	6.44	8.18	1.31	1.79	2.48
160	12.58	14.85	20.43	1.35	4.16	3.57
320	10.14	10.46	14.28	0.42	1.17	0.75
640	16.74	18.68	24.00	4.48	2.15	1.33
1280	13.42	16.95	21.28	3.43	5.55	4.64

Table 6. Replication Mean Squares for Subdivisions of Tables 2, 3, and 4 (Experiment P-1).

a. Probit Analysis of Random Data.

	DL	M
	Raw (Transform)	Raw (Transform)
≤80 Hz	2.8 (.010)	1.8 (.058)
≥160 Hz	32.3 (.019)	13.1 (.155)
F ratio[†]	8.5 (1.9)	7.2 (2.7)

b. Probit Analysis of Sequential Data.

	DL	M
	Raw (Transform)	Raw (Transform)
≤80 Hz	3.8 (.015)	7.3 (.138)
≥160 Hz	27.8 (.015)	49.7 (.244)
F ratio[†]	7.3 (1.0)	6.8 (1.8)

c. Wetherill Estimation of Sequential Data.

	DL	M
	Raw (Transform)	Raw (Transform)
≤80 Hz	6.9 (.024)	2.7 (.094)
≥160 Hz	36.7 (.011)	44.8 (.310)
F ratio[†]	5.3 (0.5)	16.6 (3.3)

[†] $F_{16,16}(.05) = 2.34$

Table 8. Comparison of Probit Estimates of Random and Sequential Data Presentation (Experiment P-1).

a. Analysis of Variance for DLs.

Source	Mean Squares	DF	F	P(F)
Subjects (S)	.559	3	39	≤.005
Noise BW (B)	1.087	7	76	≤.005
Method (M)	.031	1	2.1	-
(SB)	.024	21	1.6	-
(SM)	.002	3	0.14	-
(BM)	.006	7	0.42	-
(SBM)	.018	21	1.2	-
Replication	.0145	64	-	-

b. Analysis of Variance for Midpoints.

Source	Mean Squares	DF	F	P(F)
Subjects (S)	4.240	3	28	≤.005
Noise BW (B)	.580	7	3.9	≤.005
Method (M)	.355	1	2.4	-
(SB)	.348	21	2.3	≤.01
(SM)	.305	3	2.0	-
(BM)	.147	7	1.0	-
(SBM)	.221	21	1.5	-
Replication	.149[+]	64	-	-

[+] This is an average of two significantly different figures:
a) .107 - random data
b) .191 - sequential data
P(F)≤ .05

Table 9. Comparison of Probit and Wetherill Estimation for Sequential Data (Experiment P-1).

a. Analysis of Variance for DLs.

Source	Mean Squares	DF	F	P(F)
Estimation (E)	.230	1	14	≤.005
Residual	.0044	31	.27	≤.005
Replication	.016[†]	32	-	-

[†] This is the mean of .015 for Probit and .017 for Wetherill

b. Analysis of Variance for Midpoints.

Source	Mean Squares	DF	F	P(F)
Estimation (E)	.128	1	.62	-
Residual	.039	31	.20	≤.005
Replication	.197[††]	32	-	-

[††] This is the mean of .19 for Probit and .20 for Wetherill

Table 10. Effects of Level Spacing (Experiment P-2).

10a. DLs

	Spacing in Hz (approx. DL units)									
SUBJ.	0.5 (0.2)		1.0 (0.4)		2.5 (1.0)		5.0 (2.0)		10.0 (4.0)	
AHM	1.3	2.2	1.3	3.4	3.3	1.9	5.0	4.2	5.8	0
KP	1.4	3.7	2.9	4.2	2.9	3.1	2.5	1.9	1.7	2.5
RLW	2.0	1.8	2.3	2.4	4.1	2.7	0.8	5.8	0	10.0
SK	2.2	3.2	3.7	3.4	1.7	3.8	3.3	5.4	0.8	9.2
MEAN	1.95		2.72		2.48		2.61		1.55	
Total Presentations	661		582		489		444		408	

(No two means are significantly different.)

Analysis of Variance

Spacing/DL	Mean Squares Within Tests (DF= 40)	Mean Squares Between Tests (DF= 4)	Mean Squares for Subjects (DF= 3)
0.2	.018	.107^{++}	.061^{+}
0.4	.017	.082^{++}	.078^{++}
1.0	.055	.116	.038
2.0	.085	.413^{++}	.165
4.0	.100	1.660^{++}	.285

$^{+}$P(F)$\leq$ .05 $^{++}$P(F)$\leq$ .01

Table 10. Effects of Level Spacing (Experiment P-2).

Table 10b. Midpoints

	Spacing in Hz (approx. DL units)									
SUBJ.	0.5 (0.2)		1.0 (0.4)		2.5 (1.0)		5.0 (2.0)		10.0 (4.0)	
AHM	3.4	3.3	6.3	2.9	3.6	3.9	5.8	5.4	4.6	5.0
KP	1.3	-0.6	-0.8	0.3	0.4	0.01	1.3	-0.01	5.0	3.8
RLW	0.4	-0.1	0.2	1.4	0.2	0.9	4.6	1.8	0	0
SK	-1.1	-1.6	-0.1	-2.0	-1.9	-1.0	-0.8	0.2	-4.6	-1.3
MEAN	0.250		0.495		0.349		1.19		1.07	
Total Presentations	661		582		489		444		408	

(Only the largest and smallest means are significantly different, $P \leq .05$, 2-tailed.)

Analysis of Variance

Spacing/DL	Mean Squares Within Tests (DF= 40)	Mean Squares Between Tests (DF= 4)	Mean Squares for Subjects (DF= 3)
0.2	.013	$.225^{++}$	1.98^{++}
0.4	.028	$.272^{++}$	3.42^{++}
1.0	.035	.056	2.05^{++}
2.0	.060	$.332^{++}$	1.69^{++}
4.0	.095	$.666^{++}$	3.83^{++}

$^{++}P(F) \leq .01$

Table 11. Analysis of the Transformed DL and Midpoint Data (Experiment P-3).

Level Spacing	Mean Values (DL)	Variance (DL)	Mean Values (M)	Variance (M)
0.4 DL	12.9[†]	.018	0.39[††]	.26
0.8 DL	15.5[†]	.022	0.18	.23

[†] The difference between these two means is significant: $P < .01$.

[††] This is significantly different from zero: $P \leq .05$.

……

Table 12. Comparison of Subjective and Model Data (Experiment P-3).

	DL	Midpoint	
	Variance	Mean	Variance
Model	.020	0.28[†]	.248
Subjects	.017	1.63[††]	.203

[†] $P \leq .05$

[††] $P \leq .01$

Table 13. Summary of Results of Correlation Analysis of Model and Subjective Data (Experiment P-3).

Table 13a. DL Results

	Approx.[+] Level Spacing	Correlation: Run 1 and Run 2	Correlation: Run 1 and Run 3	Correlation: Run 1 and Run 4
I. Expt. P-3 (Model Data) N= 50	0.4	.43††	.16	.04
	0.8	.35††	.11	-.02
II. Expt. P-2 (Subjective Data) N= 8	0.2	.70†	.62	.67
	0.4	.62	.24	-.13
	1.0	.38	-.38	.28
	2.0	.25	.83†	.71†
	4.0	.23	.48	.78†
III. Miscellaneous Data (Subjects) N= 100	1.0	.49††	.42††	.17

[+] Approximate level spacing per unit DL

$^{\dagger}P \leq .05$

$^{\dagger\dagger}P \leq .01$

Table 13 cont. Summary of Results of Correlation Analysis of Model and Subjective Data (Experiment P-3).

Table 13b. Midpoint Results

	Approx^{+} Level Spacing	Correlation: Run 1 and Run 2	Correlation: Run 1 and Run 3	Correlation: Run 1 and Run 4
I. Expt. P-3 (Model Data) N= 50	0.4	$.66^{\dagger\dagger}$	$.53^{\dagger\dagger}$	$.31^{\dagger}$
	0.8	$.29^{\dagger}$	.20	.23
II. Expt. P-2 (Subjective Data) N= 8	0.2	$.96^{\dagger\dagger}$	$.95^{\dagger\dagger}$	$.95^{\dagger\dagger}$
	0.4	$.97^{\dagger\dagger}$	$.93^{\dagger\dagger}$	$.90^{\dagger\dagger}$
	1.0	$.87^{\dagger\dagger}$	$.85^{\dagger\dagger}$	$.83^{\dagger\dagger}$
	2.0	.51	$.77^{\dagger\dagger}$	$.82^{\dagger\dagger}$
	4.0	.41	.40	.21
III. Miscellaneous Data (Subjects) N= 100	1.0	$.77^{\dagger\dagger}$	$.67^{\dagger\dagger}$	$.57^{\dagger\dagger}$

$^{+}$Approximate level spacing per unit DL

$^{\dagger}P \leq .05$

$^{\dagger\dagger}P \leq .01$

Table 14. Comparison of S and S+R Data Presentation (Experiment P-4).

Table 14a. DL Results

BW / Subj.	S				S+R			
	40		160		40		160	
PB	10.0	8.8	11.0	14.0	6.3	5.0	29.0	23.0
AHM	3.1	5.9	17.0	17.0	3.7	4.1	16.0	8.1
KP	5.0	4.7	29.0	16.0	5.6	1.6	19.0	20.0
RLW	3.4	1.9	9.4	15.0	2.2	2.2	10.0	11.0
MEAN	4.83		15.29		3.55		15.63	

Analysis of Variance

Source	Mean Squares	DF	F	P(F)
Subjects (S)	.122	3	8.8	≤.005
Method (M)	.020	1	1.4	-
Bandwidth (B)	2.039	1	145	≤.005
Interactions	.025	10	1.8	-
Replication	.0139[†]	16	-	-

[†] This is the average of .011 for S and .016 for S+R.

Table 14 cont. Comparison of S and S+R Data Presentation (Experiment P-4).

Table 14b. Midpoint Results

BW Subj.	S				S+R			
	40		160		40		160	
PB	3.9	6.9	9.4	21.0	5.0	5.0	24.0	11.0
AHM	2.2	0.5	-1.6	5.3	0.9	1.7	0.6	7.2
KP	-1.3	-1.1	5.6	4.4	3.9	2.7	0.6	8.8
RLW	1.3	1.4	0.9	1.3	1.1	3.3	-1.3	-0.3
MEAN	0.95		2.86		2.63		2.48	

Analysis of Variance

Source	Mean Squares	DF	F	P(F)
Subjects (S)	.947	3	8.1	≤.005
Method (M)	.100	1	.83	-
Bandwidth (B)	.155	1	1.3	-
Interactions	.253	10	2.1	-
Replication	.118†	16	-	-

---

†This is the average of .147 for S and .089 for S+R.

Table 15. Effect of Noise Bandwidth on the DL (Experiment P-1).

a. Average DLs

Subj. / BW	LM	PB	AHM	KP	MEAN
10	1.7	4.5	5.2	4.3	3.7
20	1.8	5.5	5.2	6.5	4.4
40	3.6	4.6	6.4	3.8	4.5
80	6.3	12.7	8.1	6.9	8.2
160	15.4	27.2	14.1	29.3	20.4
320	7.5	23.4	13.7	16.9	14.3
640	13.1	43.0	19.0	30.2	24.0
1280	20.7	21.5	14.3	32.0	21.3
				Grand Mean:	10.1

b. Analysis of Variance

Source	Mean Squares	DF	F	P(F)
Subjects	.246	3	20	$\leq$.005
Bandwidths	.747	7	44	$\leq$.005
Interactions	.027	21	1.6	-
Replications	.017	32	-	-

Table 16. Effect of Noise Bandwidth on the DL (Experiment I).

a. Average DLs

Subj. / BW.	LM	PB	AHM	KP	MEAN
10	0.8	6.6	3.5	3.5	3.1
20	3.7	9.0	4.8	5.6	5.5
40	4.2	11.7	5.4	5.8	6.3
80	5.8	29.2	5.4	21.4	12.1
160	8.8	27.5	15.8	22.1	17.2
320	10.8	16.7	12.5	30.0	16.2
640	11.2	34.0	9.9	35.0	19.2
1280	13.3	39.2	20.8	24.5	22.8
				Grand Mean:	10.8

b. Analysis of Variance

Source	Mean Squares	DF	F	P(F)
Subjects	.321	3	22	≤.005
Bandwidths	.304	7	20	≤.005
Interactions	.015	21	-	-

Table 17. Comparison of Averaged DL Results of Experiments P-1 and I.

a. Averaged Results of P-1 and I

BW \ Subj.	LM	PB	AM	KP	MEAN
10	1.2	5.5	4.3	3.9	3.4
20	2.7	7.1	5.0	6.0	4.9
40	3.9	7.4	5.9	4.7	5.3
80	6.0	19.4	6.6	12.3	10.0
160	11.7	27.3	14.9	25.6	18.7
320	9.0	19.8	13.1	22.5	15.2
640	12.1	38.3	13.8	32.6	21.5
1280	16.6	29.0	17.3	28.0	22.0
				Grand Mean:	10.5

b. Joint Analysis of Variance

Source	Mean Squares	DF	F	P(F)
Subjects (S)	.413	3	29	$\leq .005$
Bandwidths (B)	.660	7	47	$\leq .005$
Occasions (P1/I)	.012	1	98	-
S x B	.014	21	1	-
(P1/I) x B	.017	7	1.2	-
(P1/I) x S	.031	3	2.2	-
Remainder	.014	21	-	-

Table 19. The Effect on the DL of a Gap in the Noise Spectrum (Experiment II-A).

Noise / Subject	NO NOISE	GAPPED NOISE	UNIFORM NOISE
RLW	1.3 2.7	4.6 0.7	4.7 4.7
KP	3.3 2.9	4.4 1.9	7.2 5.3
Average	2.5	2.5	5.4[†]

[†] $P(t)_{\text{Gapped - Uniform}} \leq .05$

……

Table 20. The Effect on the DL of a Gap in the Signal Spectrum (Experiment II-B).

BW / Noise	50 Hz	500 Hz
No Noise	1.6	2.2
80 Hz Noise	12.1	3.3

Table 21. The Effect on the DL of Shifting the Signal Relative to the Noise Spectrum (Experiment II-C).

a. No Noise

	Signal Frequency (Hz)							
Subject	750		1000		1250		1500	
AHM	5.6	2.5	1.9	1.9	2.5	5.6	5.0	2.5
KP	1.3	3.8	3.1	1.9	3.1	5.0	7.5	5.6
RLW	3.8	2.5	1.9	1.3	2.5	5.0	5.7	6.3
MEAN	3.0		2.0		3.8		5.2	

b. With 80 Hz Noise (S/N= +0.8 dB)

	Signal Frequency (Hz)							
Subject	750		1000		1250		1500	
AHM	1.9	2.5	8.1	8.8	9.1	7.8	5.6	7.5
KP	5.0	2.5	10.0	7.5	5.0	6.3	6.9	3.8
RLW	5.0	2.5	3.8	5.6	5.6	5.6	5.6	6.2
MEAN	3.1		7.0†		6.4†		5.8	

†Significant masking effect: P≤ .01

Table 22. Comparison of the Effects on the DL of Varying Noise Bandwidth and Power Density (Experiment III).

	Low Power Density		High Power Density	
	Col. 1 S/N= 2.9	Col. 2 S/N= -4.7	Col. 3 S/N= 0.9	Col. 4 S/N= -6.7
Subj. \ BW	160 Hz	1280 Hz	160 Hz	1280 Hz
AHM	3.8	6.1	12.1	8.3
KP	5.2	8.3	6.3	14.1
RLW	3.3	1.9	4.6	6.2
MEAN	4.0	4.8	7.1	9.0

Analysis of Variance

Source	Mean Squares	DF	F	P(F)
Subjects (S)	.087	2	5.8	$\leq .05$
Bandwidths (B)	.017	1	1.1	-
Power Density (P)	.151	1	10.0	$\leq .05$
Residual	.015	7	-	-

Table 23. Effects of Noise Level on the DL (Experiment IV).

Table 23a. 500 Hz

Subj. S/N	AHM		KP		RLW		MEAN
-24	36.3	65.0	15.0	11.9	8.2	8.8	17.9
-21	12.5	11.3	5.0	11.6	6.9	4.1	7.9
-18	5.6	4.4	5.6	3.8	3.4	2.9	4.2
-15	2.8	3.4	5.9	2.2	3.0	1.5	2.9
-12	2.5	2.1	2.1	1.8	2.1	2.3	2.1
-9	2.4	1.6	1.3	1.4	2.1	1.8	1.7
No Noise	1.3	1.4	1.8	2.4	1.3	1.3	1.6

Table 23b. 1 KHz

Subj. S/N	AHM		KP		RLW		MEAN
-16	45.0	36.2	31.2	15.0	11.9	7.2	20.3
-14	13.1	27.5	11.9	13.8	4.3	5.9	10.9
-11	5.0	5.0	6.3	8.8	4.7	4.4	5.5
-6	2.8	5.6	4.7	6.3	2.2	1.6	3.5
-1	1.6	4.4	2.2	4.4	2.4	1.7	2.6
+4	2.5	3.9	2.0	3.6	3.0	2.2	2.8
No Noise	2.0	3.1	3.3	2.4	2.7	1.1	2.3

Table 23 cont. Effects of Noise Level on the DL (Experiment IV).

Table 23c. 2 KHz

Subj. S/N	AHM		KP		RLW		MEAN
-12.3	51.8	42.5	73.7	22.5	16.9	21.9	33.4
-9.3	30.0	33.8	10.0	11.3	11.9	10.0	15.6
-6.3	6.9	12.5	14.4	5.0	8.8	6.6	8.5
-3.3	6.3	13.7	5.6	11.6	6.3	4.7	7.5
-0.3	8.8	5.3	4.4	6.9	5.0	3.8	5.2
+2.7	4.1	8.8	3.8	5.9	4.7	1.3	4.3
No Noise	4.7	4.4	4.4	3.8	3.4	3.1	3.9

Table 23d. 3 KHz

Subj. S/N	AHM		KP		RLW		MEAN
-6.5	61.2	72.0	57.5	73.8	35.0	73.8	60.4
-3.5	68.8	46.2	21.2	40.0	8.1	15.6	26.7
-0.5	15.8	18.2	26.2	28.8	8.8	7.5	15.7
+2.5	10.0	11.9	10.6	9.4	6.3	8.1	9.0
+5.5	12.5	8.1	10.6	14.4	4.7	4.1	8.3
+8.5	11.9	6.9	10.0	10.6	5.6	5.6	8.1
NO Noise	5.6	5.0	8.8	11.9	5.0	6.3	6.8

Table 24. Analysis of Variance of the Results of Experiment IV (From data of Tables 23a to 23d).

	Source	Mean Squares	DF	F	P(F)
a) 500 Hz	S/N	.608	6	61	≤.005
	Subjects	.125	2	12.5	≤.005
	Interaction	.041	12	4.1	≤.005
	Replication	.010	21	-	-

	Source	Mean Squares	DF	F	P(F)
b) 1 KHz	S/N	.558	6	36	≤.005
	Subjects	.242	2	15.5	≤.005
	Interaction	.031	12	2	-
	Replication	.016	21	-	-

	Source	Mean Squares	DF	F	P(F)
c) 2 KHz	S/N	.585	6	22	≤.005
	Subjects	.167	2	6.3	≤.01
	Interaction	.022	12	0.1	-
	Replication	.027	21	-	-

	Source	Mean Squares	DF	F	P(F)
d) 3 KHz	S/N	.594	6	41	≤.005
	Subjects	.368	2	25	≤.005
	Interaction	.038	12	2.6	≤.025
	Replication	.014	21	-	-

Table 25. Effects of Noise Level and Bandwidth on the DL (Experiment V).

Table 25a. 500 Hz

1) High Signal Level

S/N	Wide Bandwidth				Narrow Bandwidth			
Subj.	-18	-15	-12	-9	-3	0	+3	+6
AHM	6.3	2.8	2.8	2.8	8.8	6.6	2.0	1.3
KP	8.2	5.0	4.4	2.2	14.7	5.0	2.8	3.1
RLW	4.4	2.1	1.6	1.1	3.0	2.2	2.5	1.4
MEAN	6.1	3.1	2.7	1.9	7.5	4.3	2.4	1.8

2) Low Signal Level

S/N	Wide Bandwidth				Narrow Bandwidth			
Subj.	-18	-15	-12	-9	-3	0	+3	+6
AHM	7.2	4.4	2.5	2.3	15.3	5.3	3.3	4.8
KP	6.9	4.4	5.6	5.0	20.9	7.5	3.4	2.1
RLW	3.1	3.6	4.4	2.3	7.2	4.4	5.6	2.0
MEAN	5.4	4.1	4.0	3.0	13.3	5.6	4.0	2.8

Table 25 cont. Effects of Noise Level and Bandwidth on the DL (Experiment V).

Table 25b. 1 KHz

1) High Signal Level

S/N Subj.	Wide Bandwidth				Narrow Bandwidth			
	-16	-13	-10	-7	-1.75	+1.25	+4.25	+7.25
AHM	18.8	5.6	5.6	2.5	19.4	7.5	5.6	4.7
KP	17.5	11.3	7.2	4.7	22.5	11.9	12.2	4.1
RLW	13.8	7.2	1.3	3.4	5.4	4.4	3.6	2.2
MEAN	16.6	7.7	4.0	3.4	13.5	7.4	6.4	3.5

2) Low Signal Level

S/N Subj.	Wide Bandwidth				Narrow Bandwidth			
	-16	-13	-10	-7	-1.75	+1.25	+4.25	+7.25
AHM	9.4	10.6	6.3	3.9	17.5	6.9	4.4	2.6
KP	12.2	11.3	5.9	3.0	18.1	15.6	5.0	6.1
RLW	9.4	5.9	4.7	4.1	8.5	6.9	3.5	4.3
MEAN	10.3	9.0	5.6	3.6	13.0	9.1	4.3	4.1

Table 25 cont. Effects of Noise Level and Bandwidth on the DL (Experiment V).

Table 25c. 2 KHz

1) High Signal Level

S/N Subj.	Wide Bandwidth				Narrow Bandwidth			
	-12.3	-9.3	-6.3	-3.3	-3	0	+3	+6
AHM	45.0	18.8	9.7	15.0	32.5	17.5	5.6	8.4
KP	31.9	18.1	9.4	8.7	58.8	37.5	10.6	10.0
RLW	16.3	5.0	7.5	4.1	20.6	14.4	6.9	8.4
MEAN	28.7	12.2	8.8	8.4	34.2	21.2	7.5	8.9

2) Low Signal Level

S/N Subj.	Wide Bandwidth				Narrow Bandwidth			
	-12.3	-9.3	-6.3	-3.3	-3	0	+3	+6
AHM	65.0	18.8	21.3	25.0	33.1	18.1	13.8	13.5
KP	82.5	17.5	12.5	15.6	46.3	23.8	5.0	6.9
RLW	27.5	8.1	8.1	11.9	31.3	10.0	5.3	9.8
MEAN	53.0	13.9	13.0	16.7	36.3	16.3	7.2	9.7

Table 25 cont. Effects of Noise Level and Bandwidth on the DL (Experiment V).

Table 25d. 3 KHz

1) High Signal Level

S/N Subj.	Wide Bandwidth				Narrow Bandwidth			
	-6.5	-3.5	-0.5	+2.5	-2.75	+0.25	+3.25	+6.25
AHM	45.0	30.6	30.0	12.5	71.3	45.0	31.3	28.8
KP	50.0	25.0	28.1	15.0	58.8	60.0	35.0	18.8
RLW	43.8	17.5	11.9	8.8	36.2	15.0	10.0	14.1
MEAN	46.2	23.8	21.9	11.8	53.5	34.6	22.4	19.8

2) Low Signal Level

S/N Subj.	Wide Bandwidth				Narrow Bandwidth			
	-6.5	-3.5	-0.5	+2.5	-2.75	+0.25	+3.25	+6.25
AHM	65.0	25.0	21.2	22.5	86.2	34.4	26.3	19.4
KP	86.3	23.1	21.2	16.3	91.3	45.0	13.8	21.3
RLW	33.8	15.6	8.4	11.6	38.8	21.9	15.6	6.9
MEAN	57.6	20.8	15.7	16.2	67.4	32.4	17.8	14.3

Table 26. Analysis of Variance for DL Data of Experiment V.

Table 26a. 500 Hz

1) Wide Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.089	3	5.3	≤.005
Subjects (S)	.093	2	5.4	≤.01
Signal Level (L)	.035	1	2.1	-
NL x S	.005	6	.31	-
NL x L	.010	3	.59	-
S x L	.007	2	.42	-
NL x S x L	.012	6	.72	-

2) Narrow Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.317	3	19	≤.005
Subjects (S)	.085	2	5.1	≤.01
Signal Level (L)	.142	1	8.4	≤.005
NL x S	.035	6	2.1	-
NL x L	.005	3	.30	-
S x L	.015	2	.89	-
NL x S x L	.014	6	.83	-

Table 26 cont. Analysis of Variance of DL Data of Experiment V.

Table 26b. 1 KHz

1) Wide Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.208	3	12	$\leq .005$
Subjects (S)	.051	2	3.0	$\leq .05$
Signal Level (L)	0	1	-	?
NL x S	.014	6	.83	-
NL x L	.028	3	1.7	-
S x L	.014	2	.83	-
NL x S x L	.015	6	.89	-

2) Narrow Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.269	3	16	$\leq .005$
Subjects (S)	.193	2	11	$\leq .005$
Signal Level (L)	0	1	-	?
NL x S	.015	6	.89	-
NL x L	.016	3	.95	-
S x L	.029	2	1.7	-
NL x S x L	.009	6	.54	-

Table 26 cont. Analysis of Variance of DL Data of Experiment V.

Table 26c. 2 KHz

1) Wide Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.372	3	22	$\leq .005$
Subjects (S)	.293	2	17	$\leq .005$
Signal Level (L)	.209	1	12	$\leq .005$
NL x S	.011	6	.65	-
NL x L	.016	3	.95	-
S x L	.007	2	.42	-
NL x S x L	.010	6	.59	-

2) Narrow Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.500	3	30	$\leq .005$
Subjects (S)	.067	2	4.0	$\leq .05$
Signal Level (L)	.001	1	.08	-
NL x S	.022	6	1.3	-
NL x L	.006	3	.38	-
S x L	.052	2	3.1	$\leq .05$
NL x S x L	.010	6	.59	-

Table 26 cont. Analysis of Variance of DL Data of Experiment V.

Table 26d. 3 KHz

1) Wide Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.335	3	20	≤.005
Subjects (S)	.543	2	8.5	≤.005
Signal Level (L)	0	1	-	?
NL x S	.007	6	.43	-
NL x L	.023	3	1.4	-
S x L	.005	2	.29	-
NL x S x L	.006	6	.35	-

2) Narrow Band

Source	Mean Squares	DF	F	P(F)
Noise Level (NL)	.348	3	21	≤.005
Subjects (S)	.280	2	17	≤.005
Signal Level (L)	.008	1	.51	-
NL x S	.006	6	.36	-
NL x L	.015	3	.89	-
S x L	.005	2	.31	-
NL x S x L	.022	6	1.3	-

Table 27. Comparison of Some Published Estimates of the CBW with the Results of Experiment V.[†]

	Signal Frequency (Hz)			
	500	1000	2000	3000
Fletcher (1940) (Masking) CBW (Hz)	40	70	110	160
Zwicker et al (1957) (Loudness) CBW (Hz)	110	160	295	450
Greenwood (1961) (Masking) CBW (Hz)	100	186	300	450
de Boer & Bos (1962) (Masking) CBW (Hz)	140	220	420	800
Experiment V (Discrimination) CBW (Hz)	120	200	375	1270
(Wide-Narrow Separation, in dB) (cf. figs. 29a - 29d)	(15.7)	(13.5)	(10.8)	(5.5)

[†] Some of these figures are approximate, since they were read from published graphs.

Table 28. Comparison of the Noise Reducing Properties of the Basilar Membrane Filters with Subjective Results (Experiment VI-A).[+]

	Signal Frequency (Hz)			
	500	1000	2000	3000
Displacement Filter	10.4	7.5	1.8	-6.3
Gradient Filter	10.9	8.2	5.1	1.5
(Without the 'Middle Ear')	(10.4)	(7.4)	(4.4)	(2.6)
Subjective Data (Experiment V)	15.7	13.5	10.5	5.5

[+] The figures indicate the SN ratio at the filter output relative to 0 dB S/N at the filter input for noise of 4.5 KHz bandwidth.

Table 29. Comparison of the Separation (in dB) between the Wide and Narrow Band Curves and Estimated CBW's (in Hz) for Model and Subjective Data (Experiment VI-B).

Signal Frequency (Hz)	Subjective Data[+] (Experiment V)		Model Data (Experiment VI-B)	
	Separation (dB)	Estimated CBW (Hz)	Separation (dB)	Estimated CBW (Hz)
500	17.0	90	11.8	298
1000	13.5	201	9.5	505
2000	9.3	530	6.5	1010
3000	5.0	1425	2.5	2530

[+] Note that the subjective figures given here are slightly different from the separations shown in table 27. The small differences are due to the different methods of fitting curves through the plotted points; the figures in this table are probably more reliable.

FIGURES

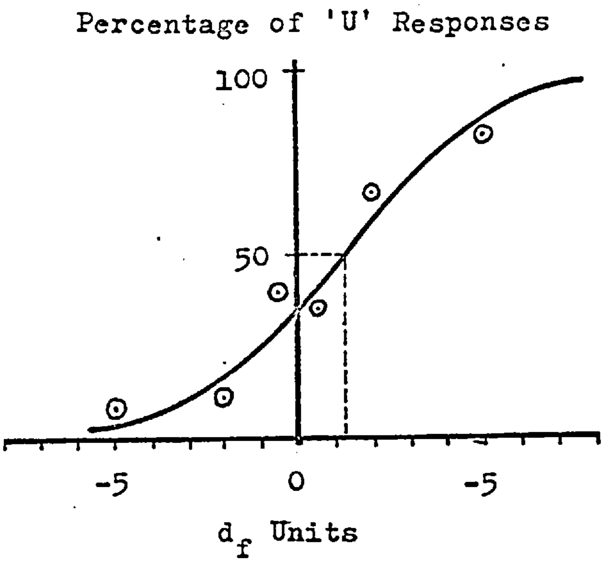


Fig. 1 Typical Result Obtained from a Pitch Judgement Experiment.

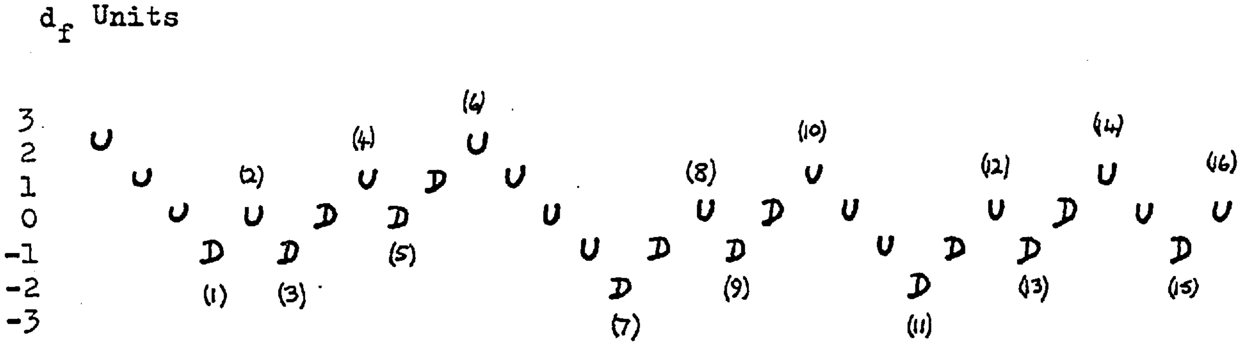
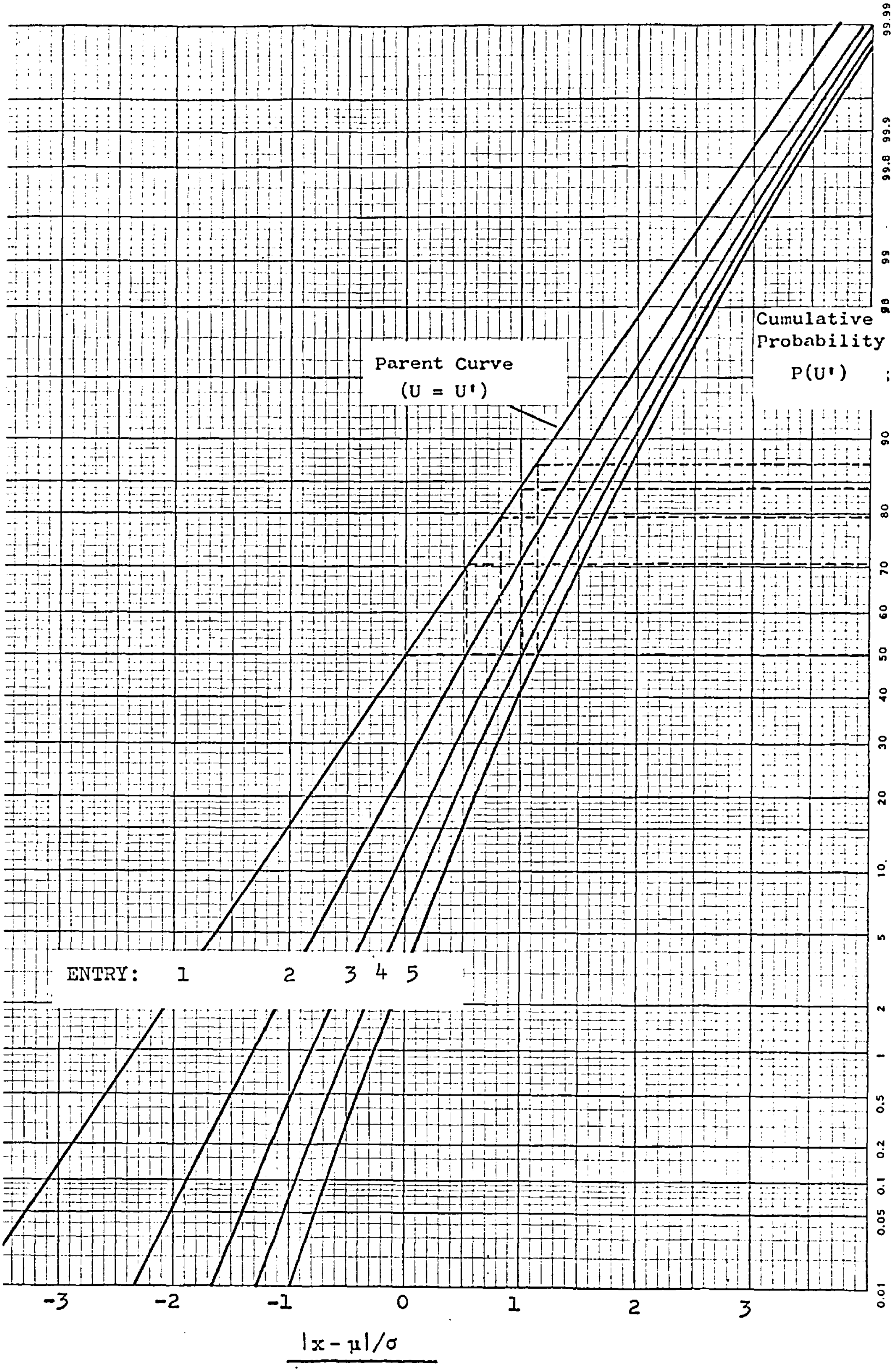


Fig. 2 Pattern of Responses Generated by the Use of Dixon and Mood's Up and Down Method.

Fig. 3 Response Curves Generated by Wetherill Sequential Rules.



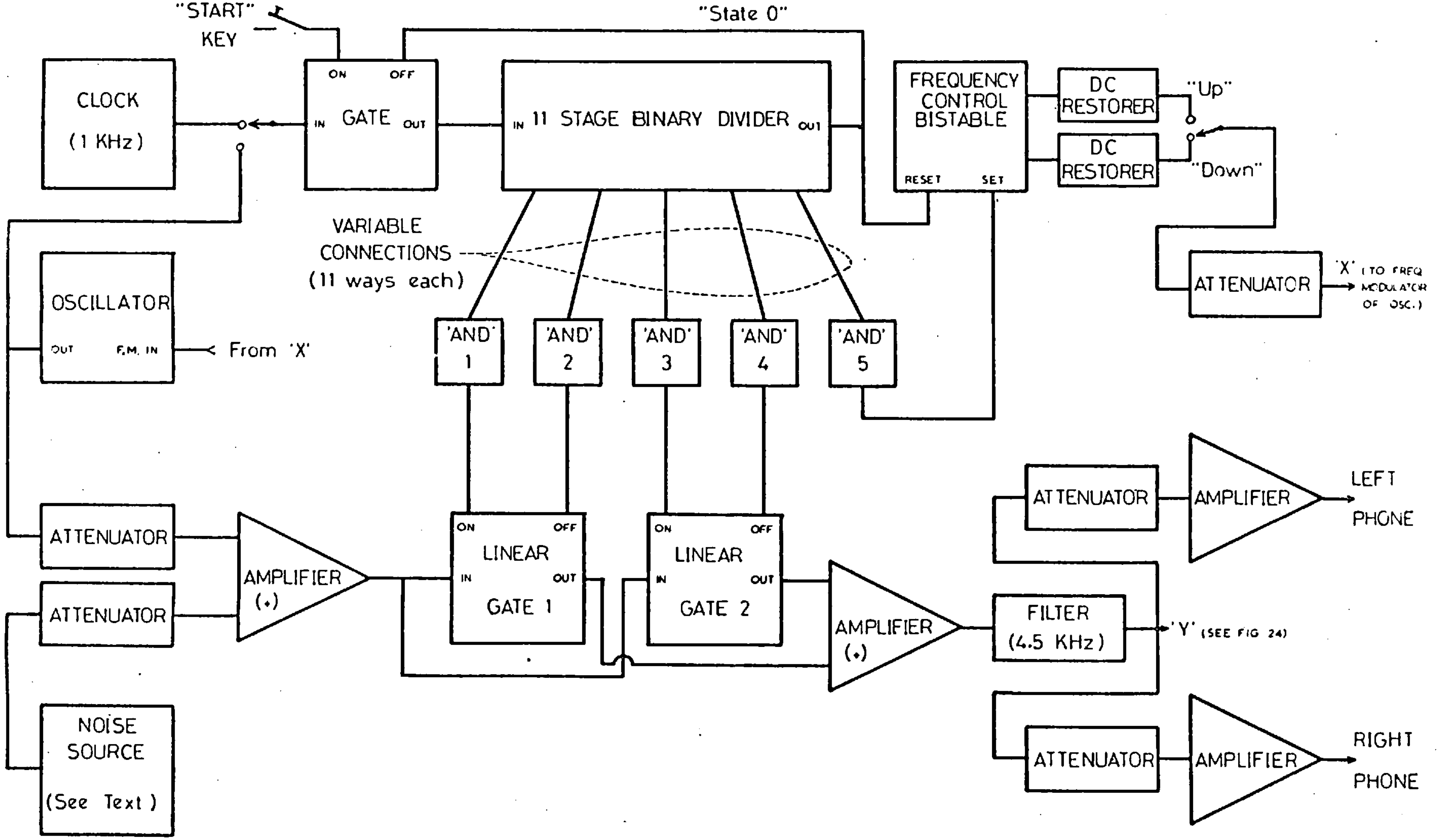
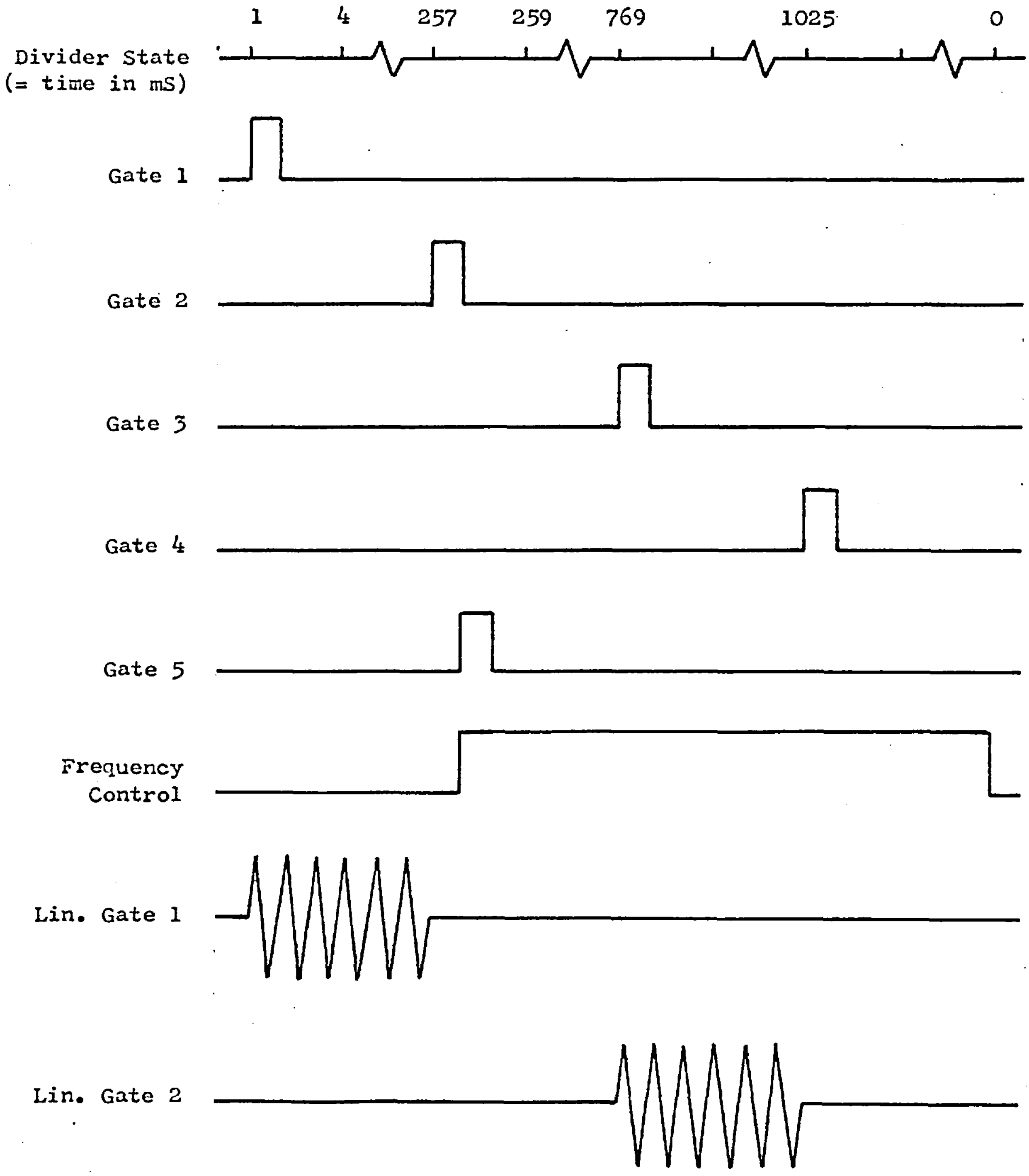
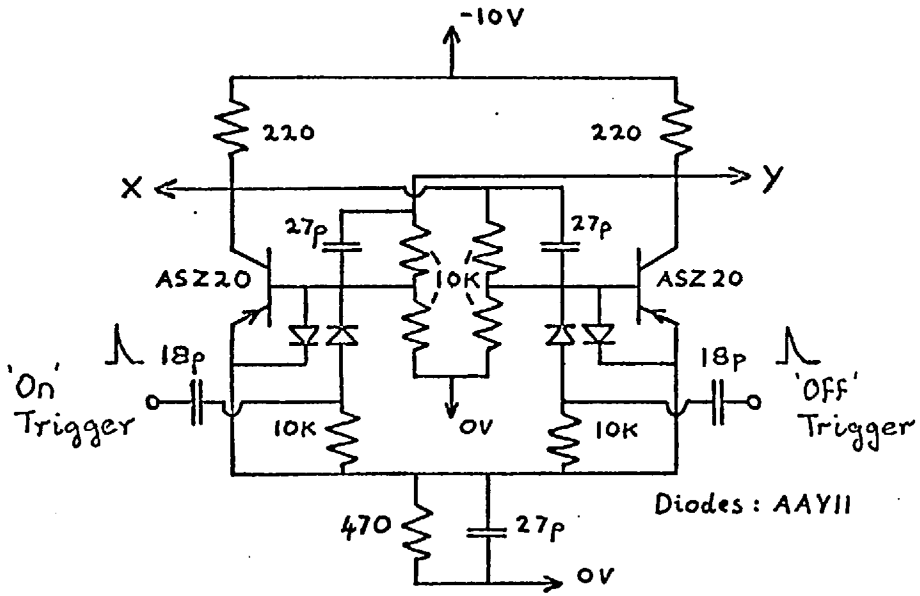


Fig. 4. Block Diagram of the Experimental Apparatus.

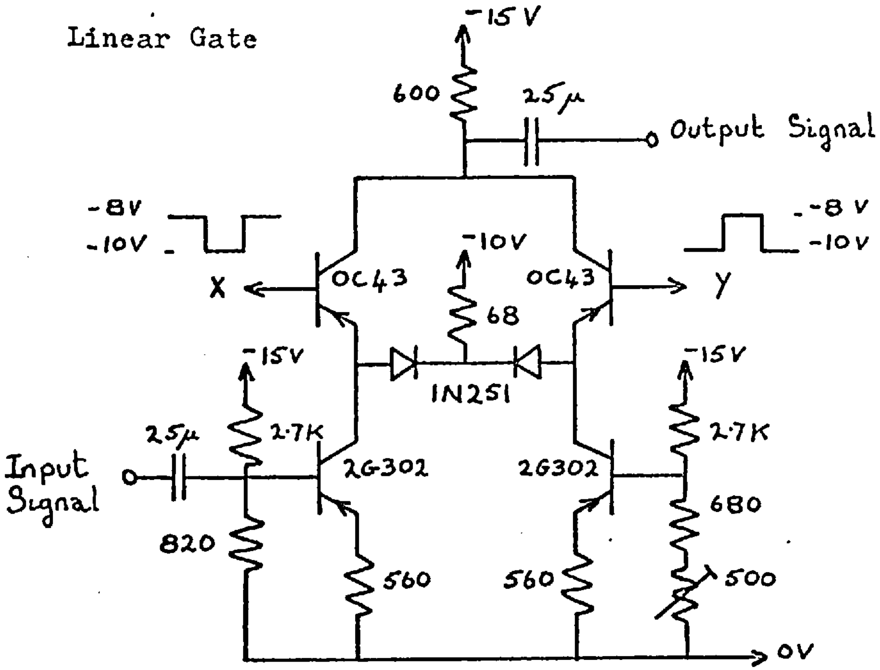
Fig. 5 Trigger Pulse Sequence Used to Generate an A-X Sequence.



Control Bistable



Linear Gate



Gated Segments of 1.3 MHz Sinewave, showing transient aberrations.

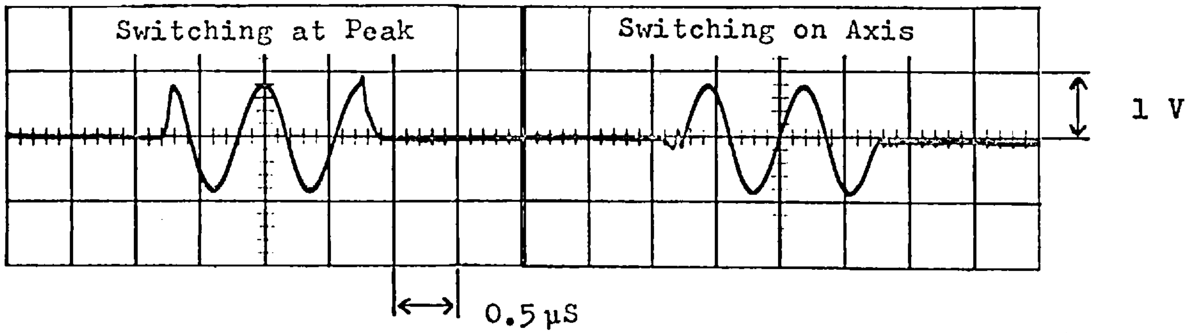


Fig. 6 Circuit and Output Waveforms for Transistor Operated Signal Switch.

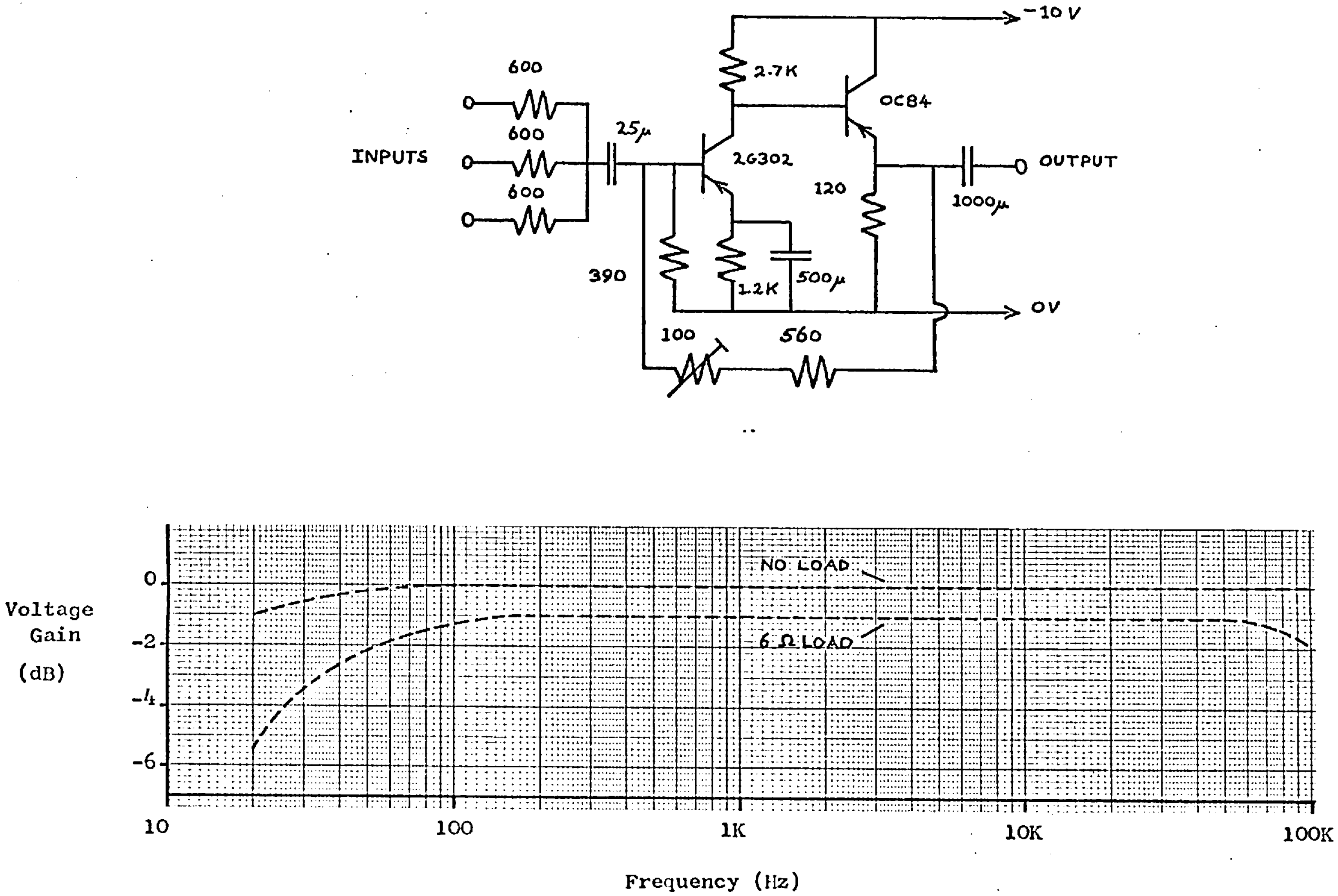
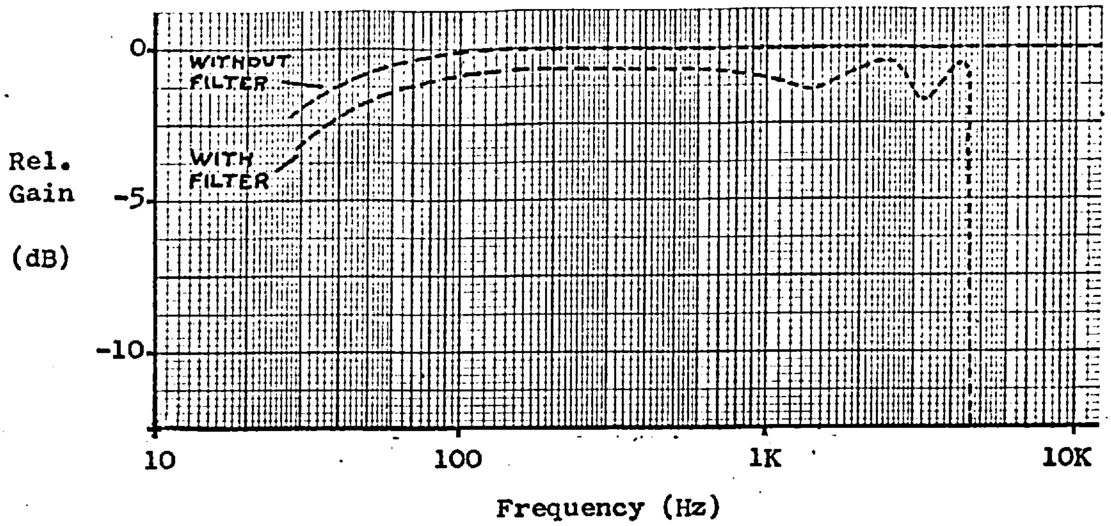


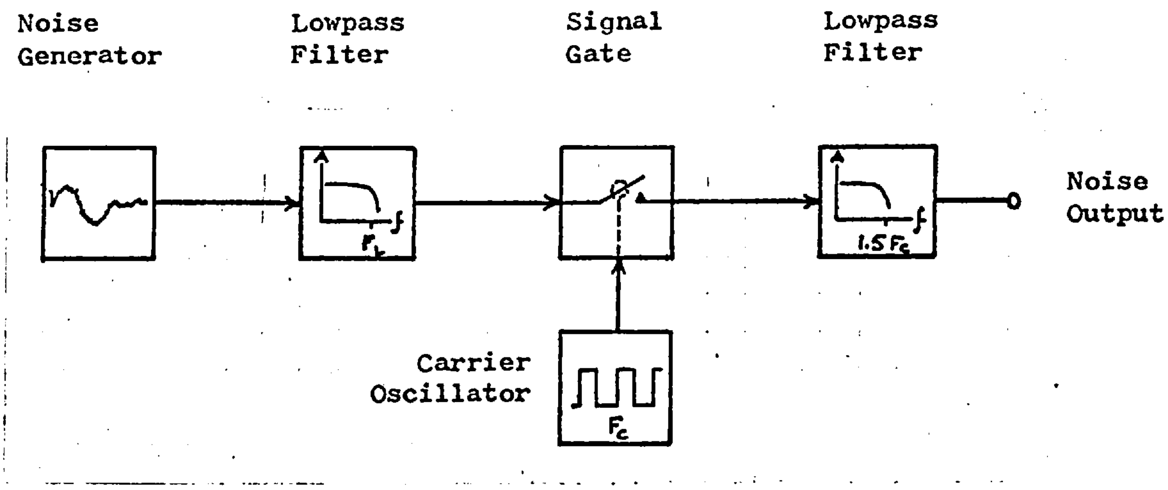
Fig. 7 Circuit and Frequency Response of Transistor Operated Feedback Amplifier.

Fig. 8 Frequency Response of the Complete Apparatus.



••

Fig. 9 Block Diagram of the Noise Modulator.



••

Fig. 10 Power Density Spectrum of the Noise Generator.

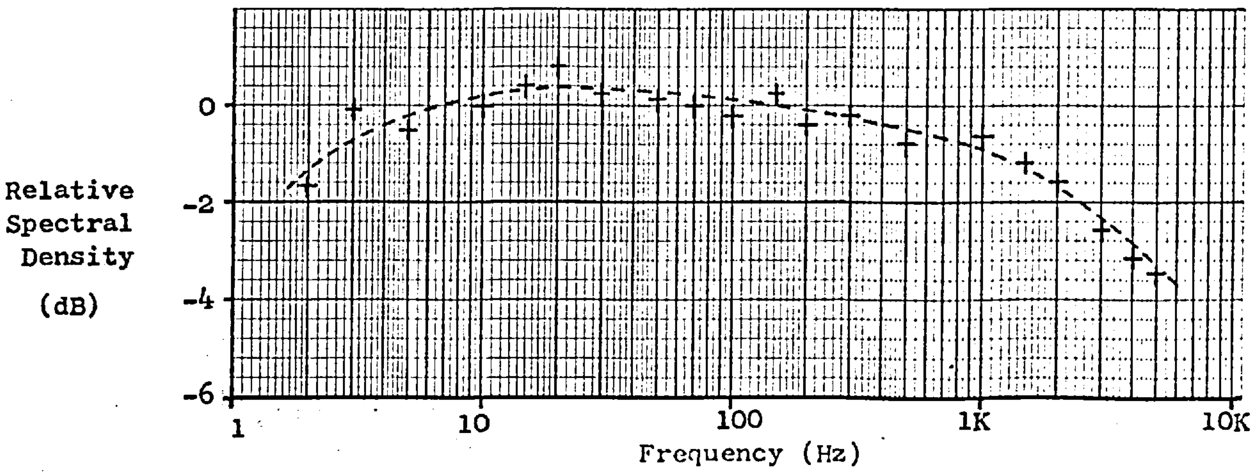
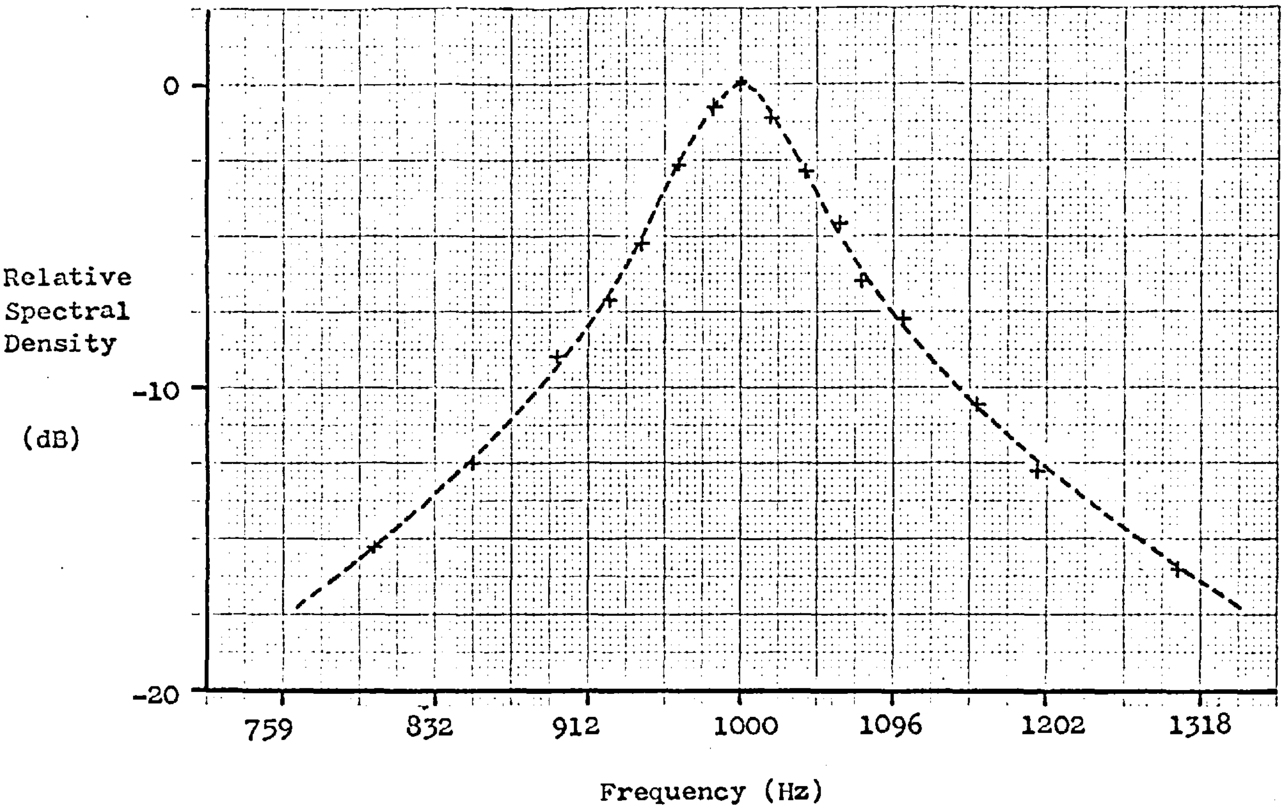


Fig. 11 Power Density Spectrum of Filtered Noise of 80 Hz Bandwidth.



……

Fig. 12 Power Density Spectrum of Modulated Noise of 100 Hz Bandwidth.

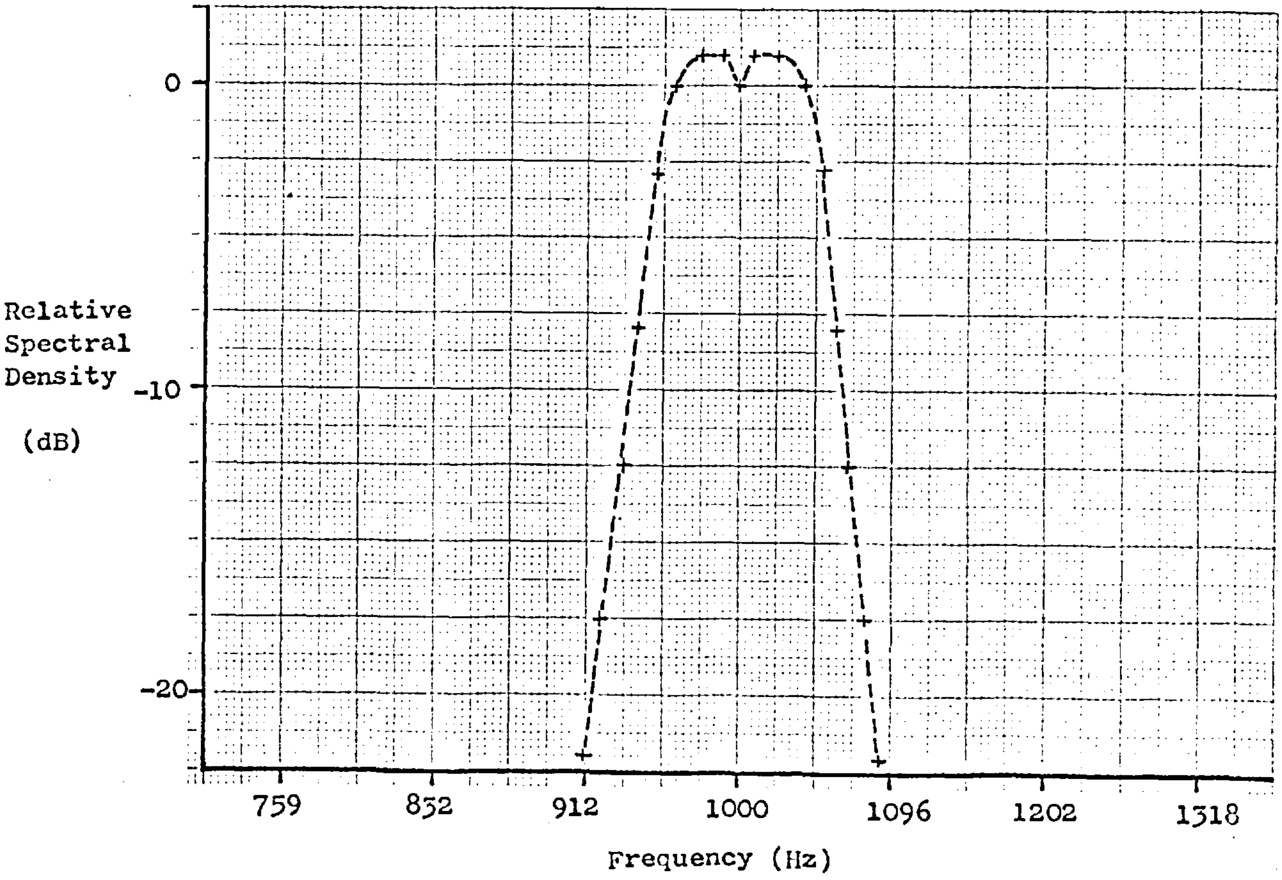


Fig. 13 Frequency Response Curves of the Earphones Using Two Different Types of Foam.

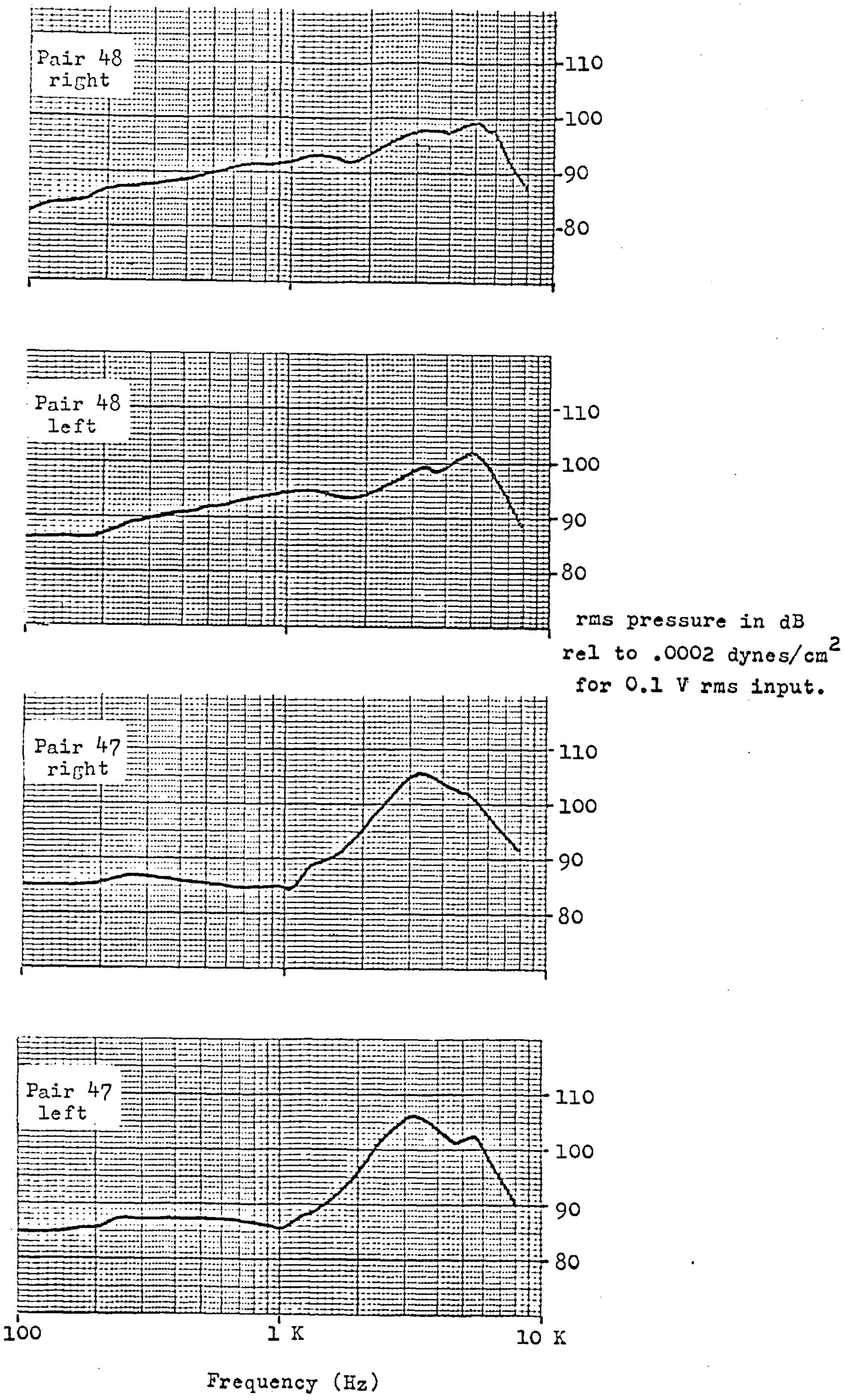
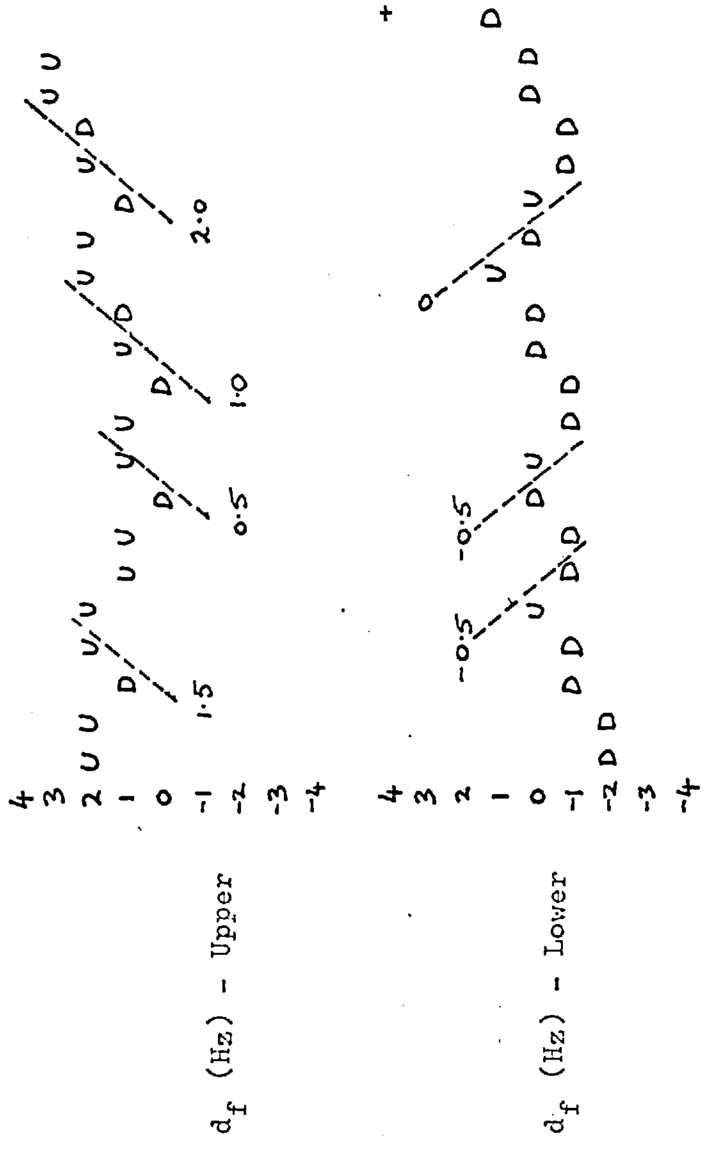


Fig. 14 Application of the Wetherill 71/29 Percent Strategy.

(This data is for Subject AHM for the noise-free condition at 500 Hz in Experiment V; see table 25a.)



$$\overline{d}_f\,(\text{upper strategy}) = (1{\cdot}5 + 0{\cdot}5 + 1{\cdot}0 + 2{\cdot}0)/4 = 1{\cdot}25\ \text{Hz}$$

$$\overline{d}_f\,(\text{lower strategy}) = (-0{\cdot}5 - 0{\cdot}5 + 0 + 1{\cdot}0)/4 = 0$$

$$DL = \overline{d}_f\,(\text{upper strategy}) - \overline{d}_f\,(\text{lower strategy}) = 1{\cdot}3\ \text{Hz}$$

$$M = \left[\overline{d}_f\,(\text{upper strategy}) + \overline{d}_f\,(\text{lower strategy})\right]/2 = 0{\cdot}6\ \text{Hz}$$

FIG. 15. Distribution of DL and Midpoint Estimates (Experiment P-1).

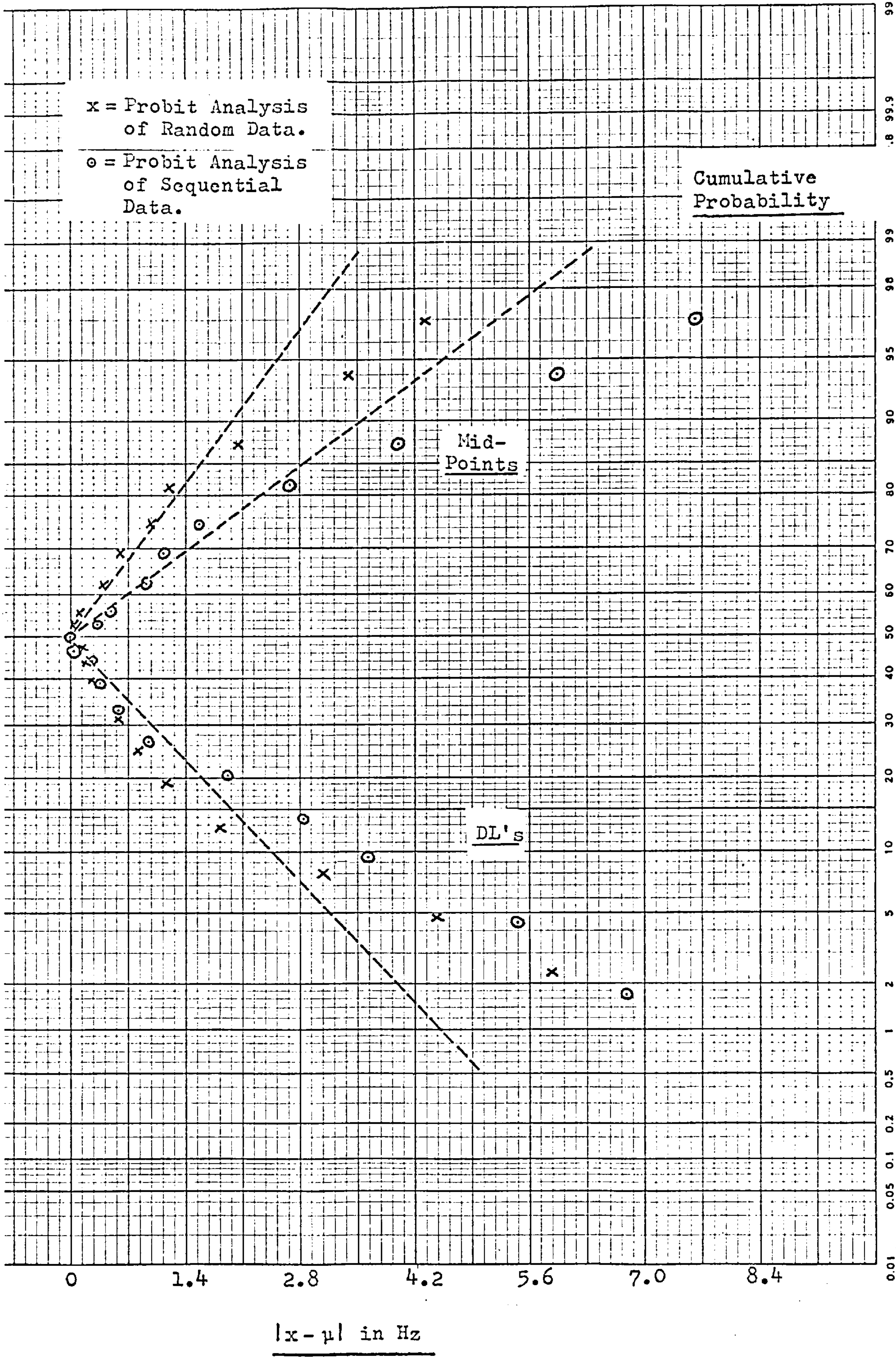


Fig. 16 Distribution of Transformed DL and Midpoint Estimates. (Experiment P-1)

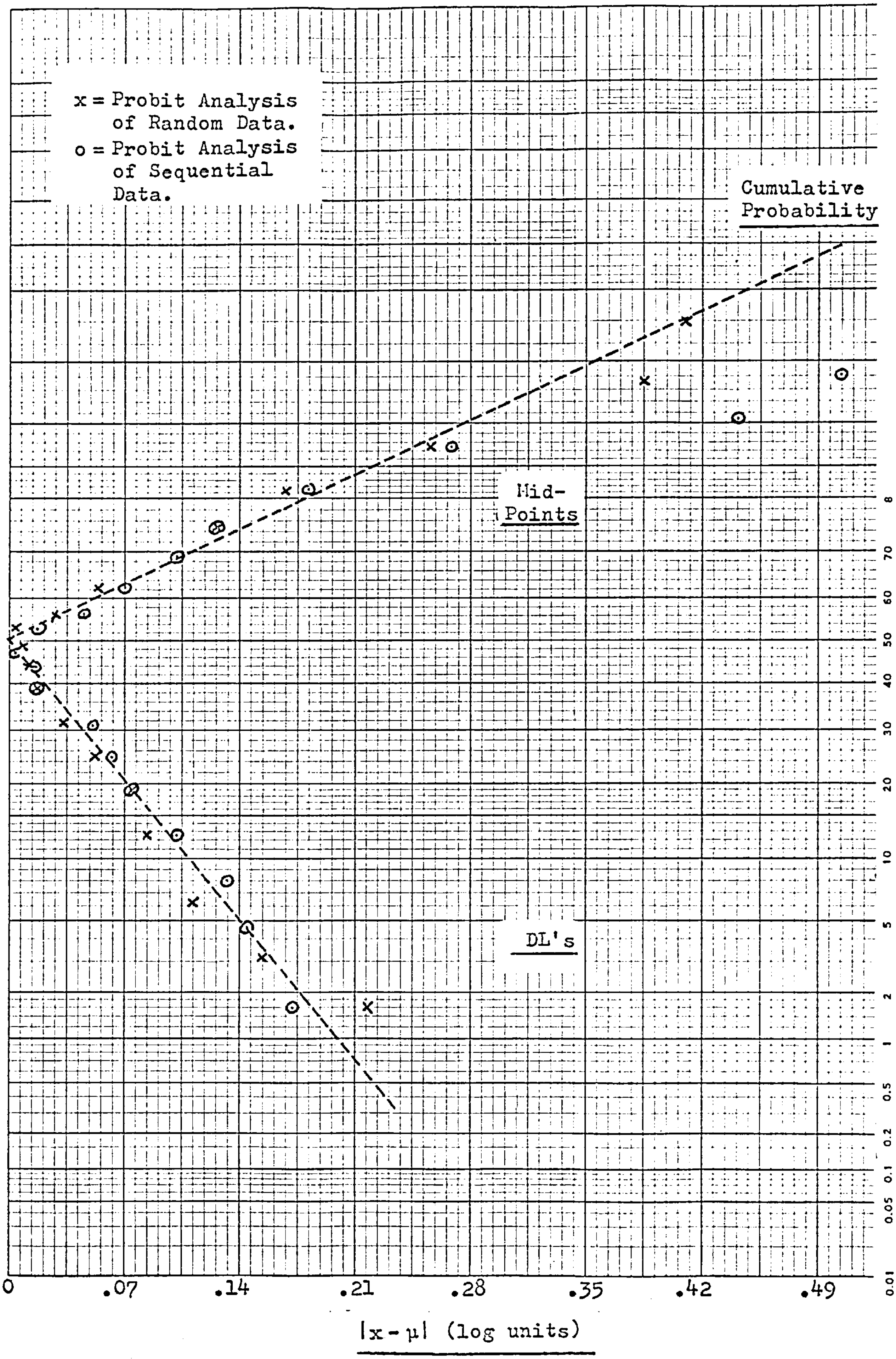


Fig. 17 χ^2 for Probit Curves Fitted to Data from Experiment P-1.

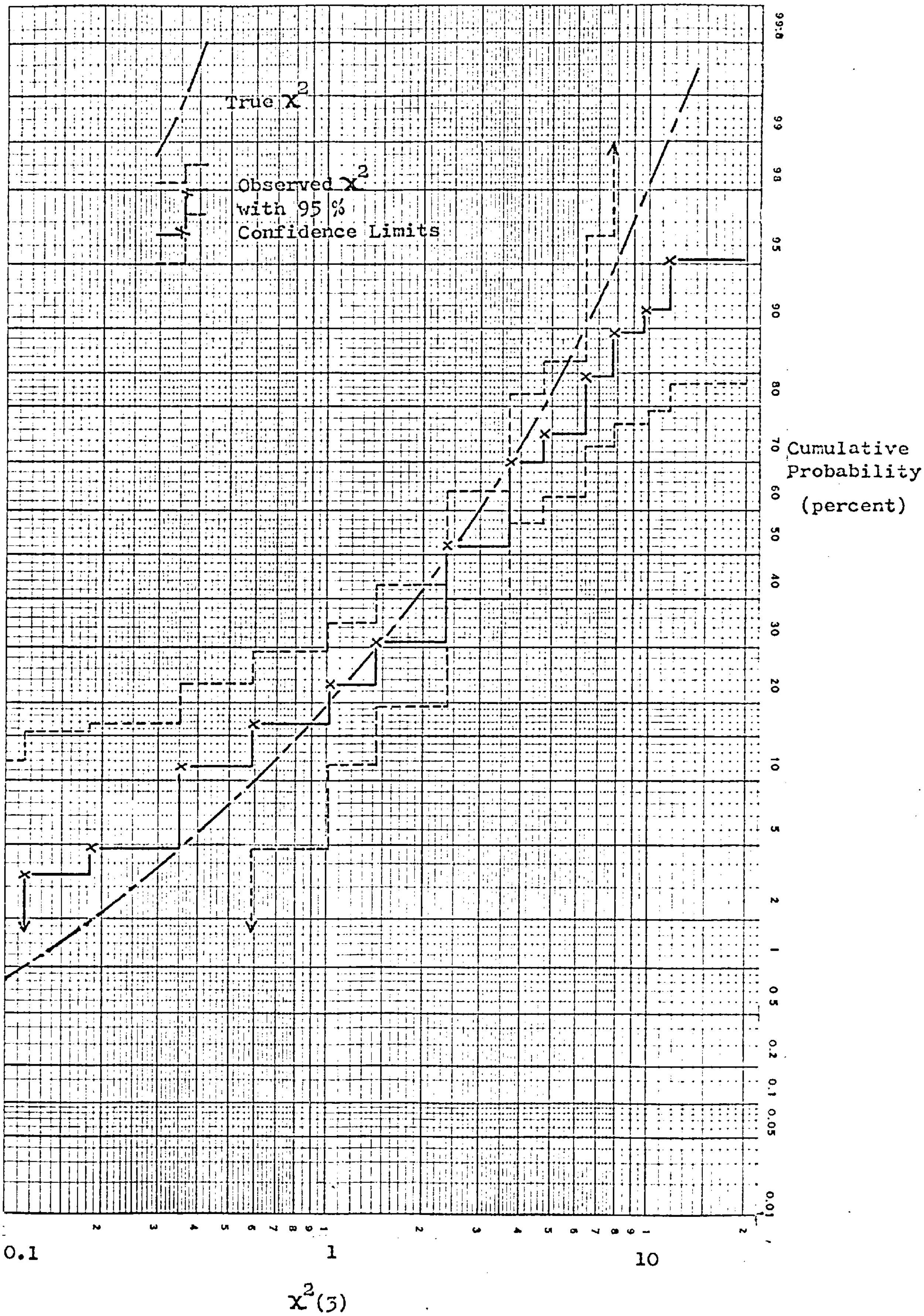
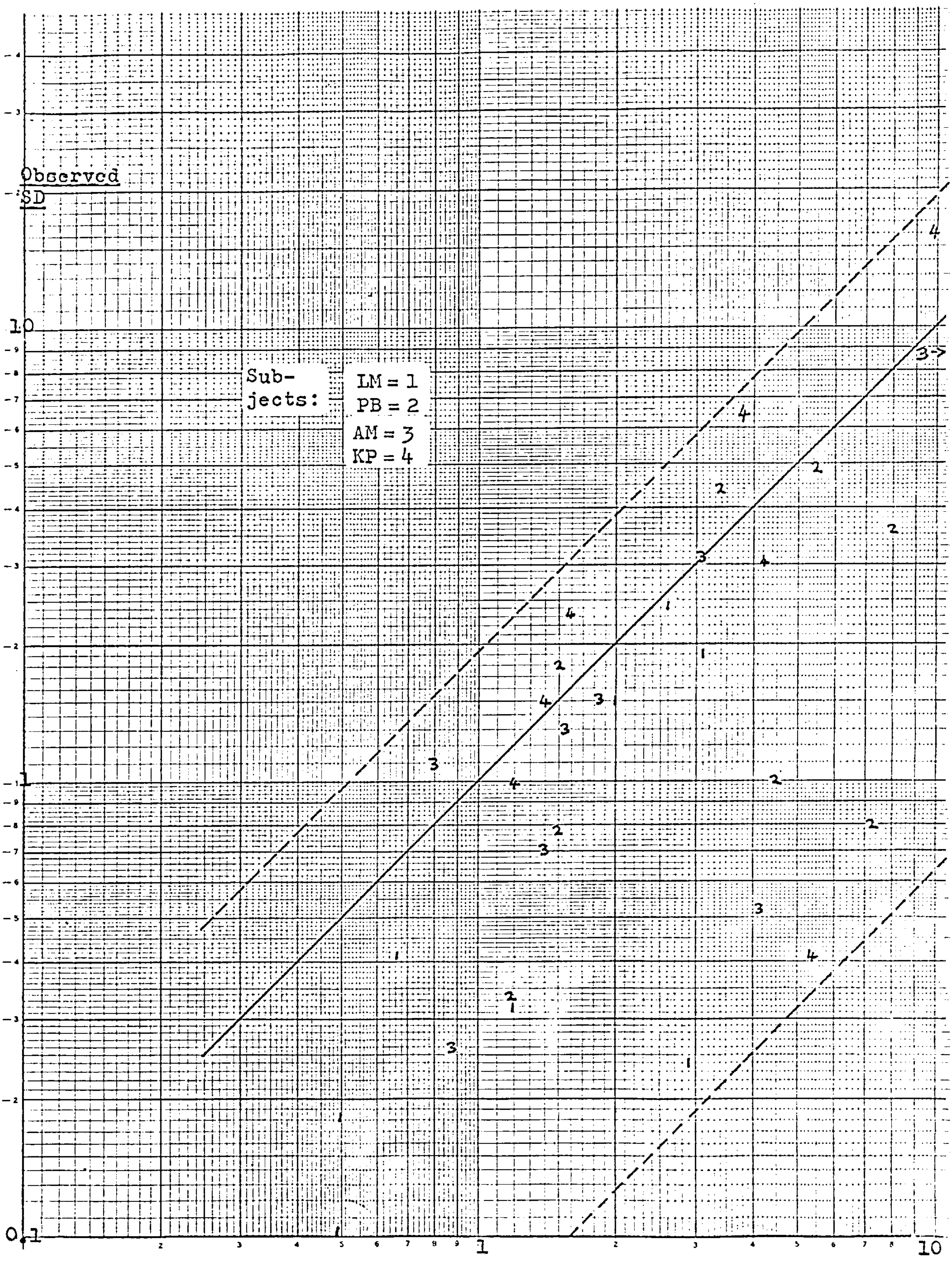


Fig. 18 Observed vs. Predicted SD's for DL's (Random Data).

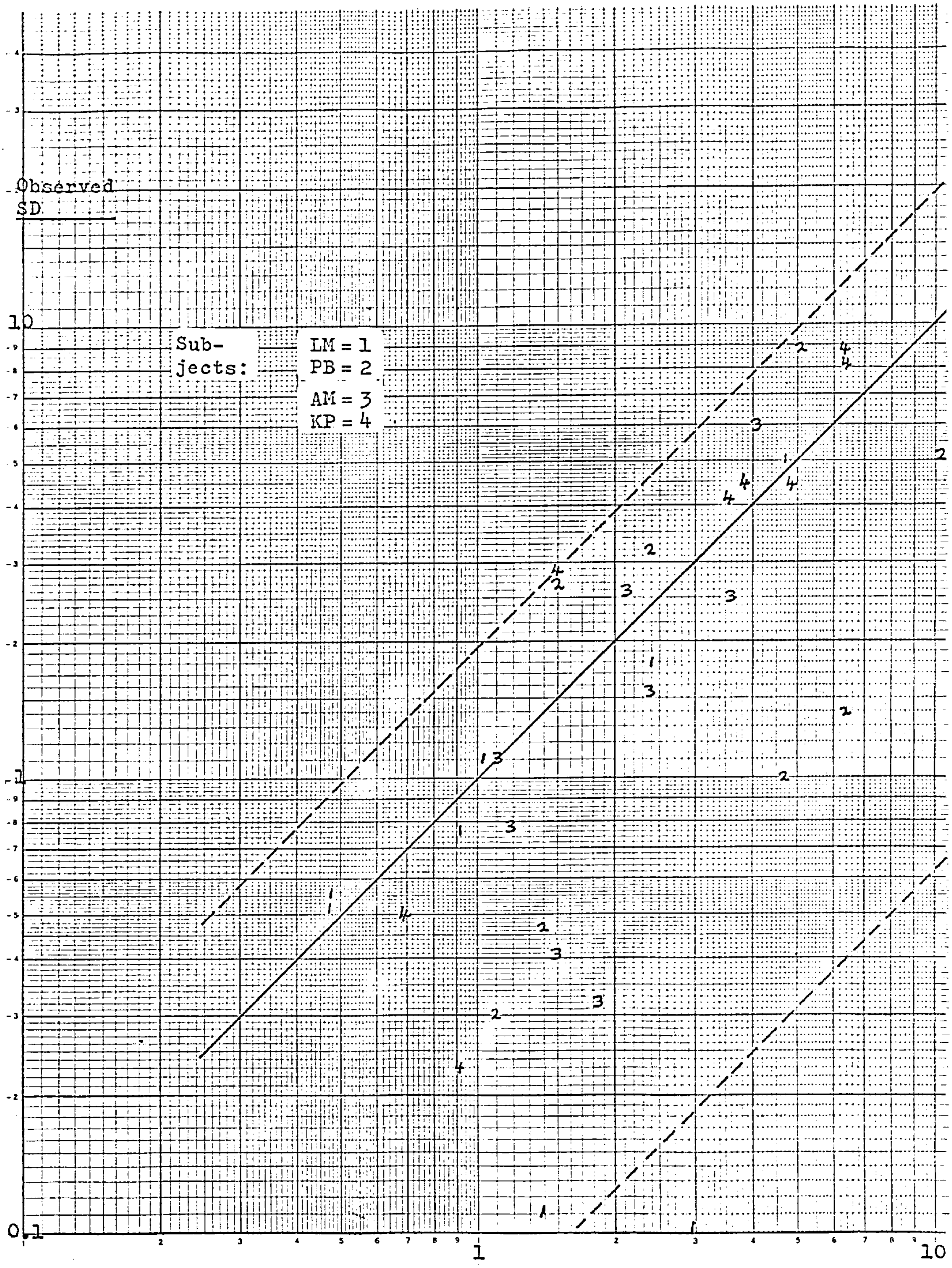
(Experiment P-1)



(------ 10 pc Confidence Interval)

Fig. 19 Observed vs. Predicted SD's for DL's (Sequential Data).

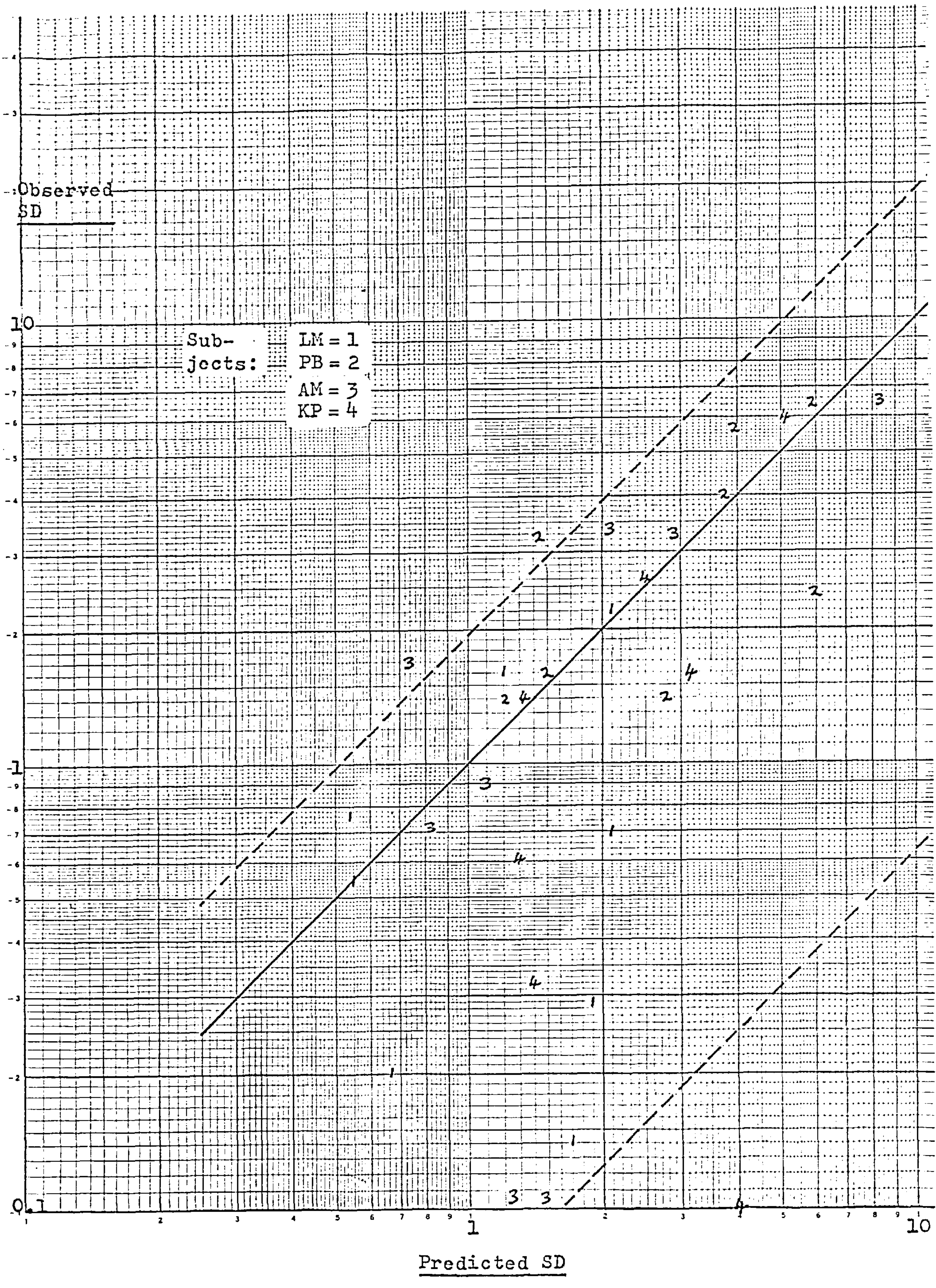
(Experiment P-1)



(----- 10 pc Confidence Interval)

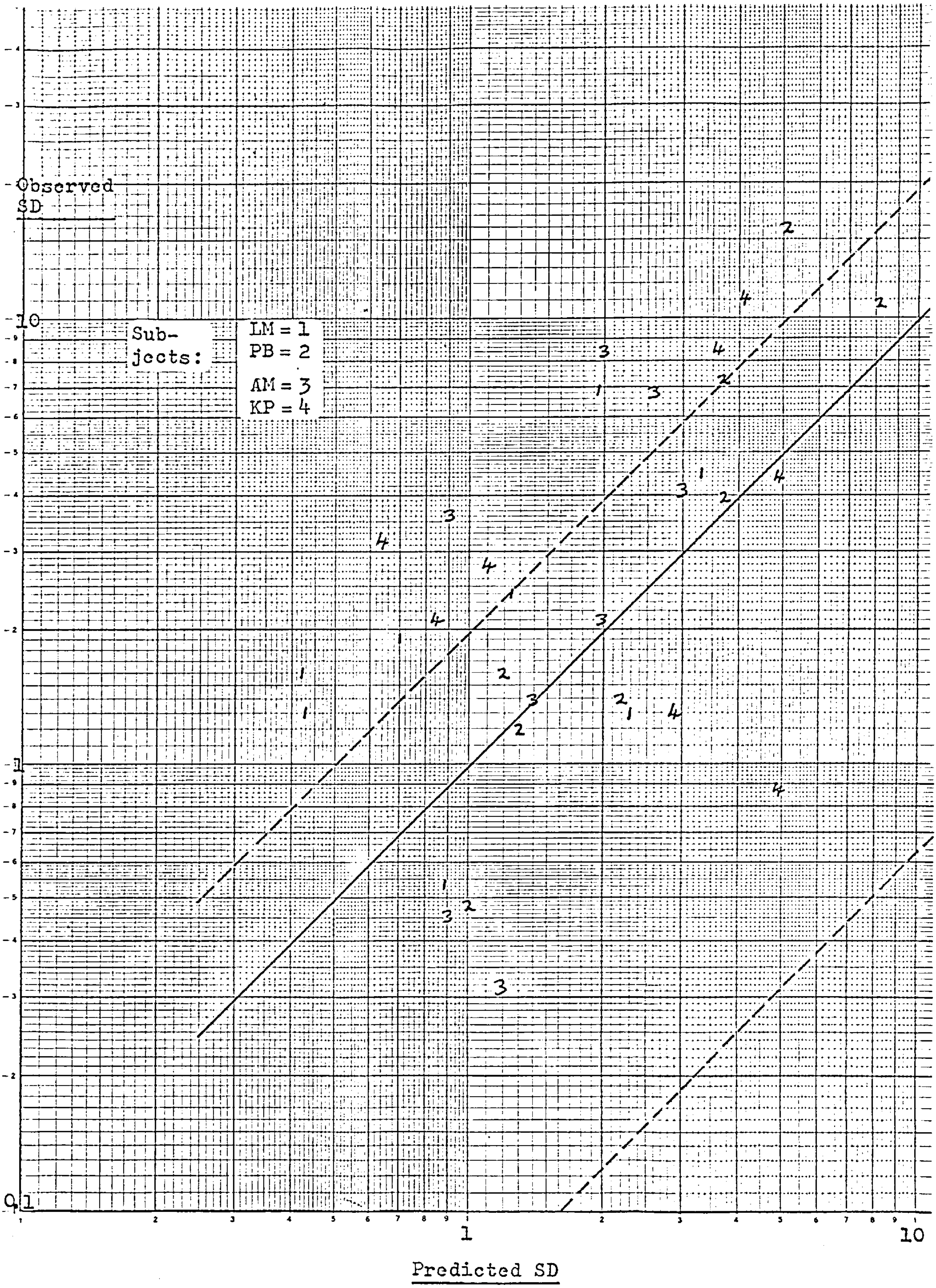
Fig. 20 Observed vs. Predicted SD's for M's (Random Data).

(Experiment P-1)



(------ 10 pc Confidence Interval)

Fig. 21 Observed vs. Predicted SD's for M's (Sequential Data).
(Experiment P-1)



(------ 10 pc Confidence Interval)

Fig. 22 Comparison of Probit Analyses of R and S Data (Expt P-1).

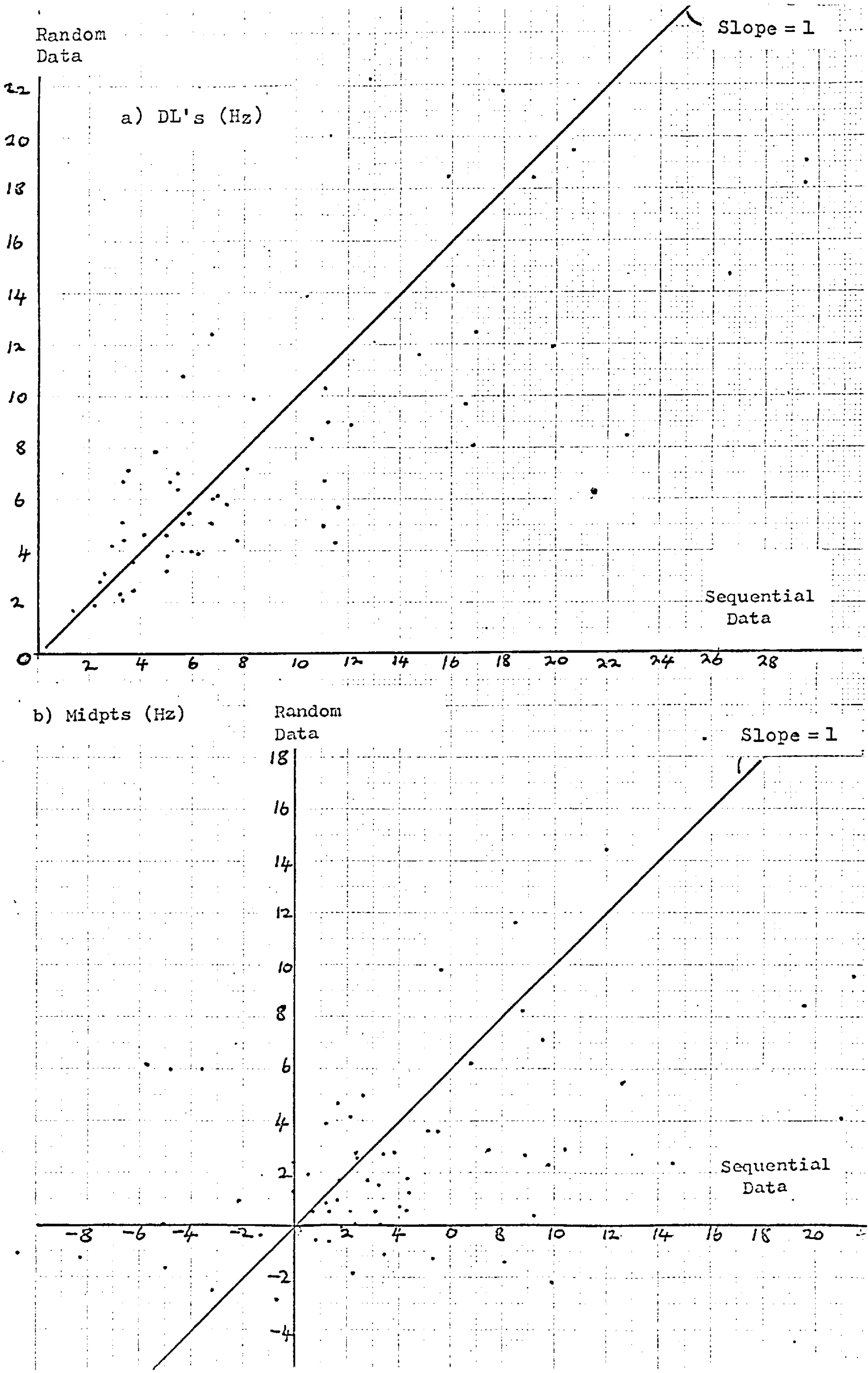
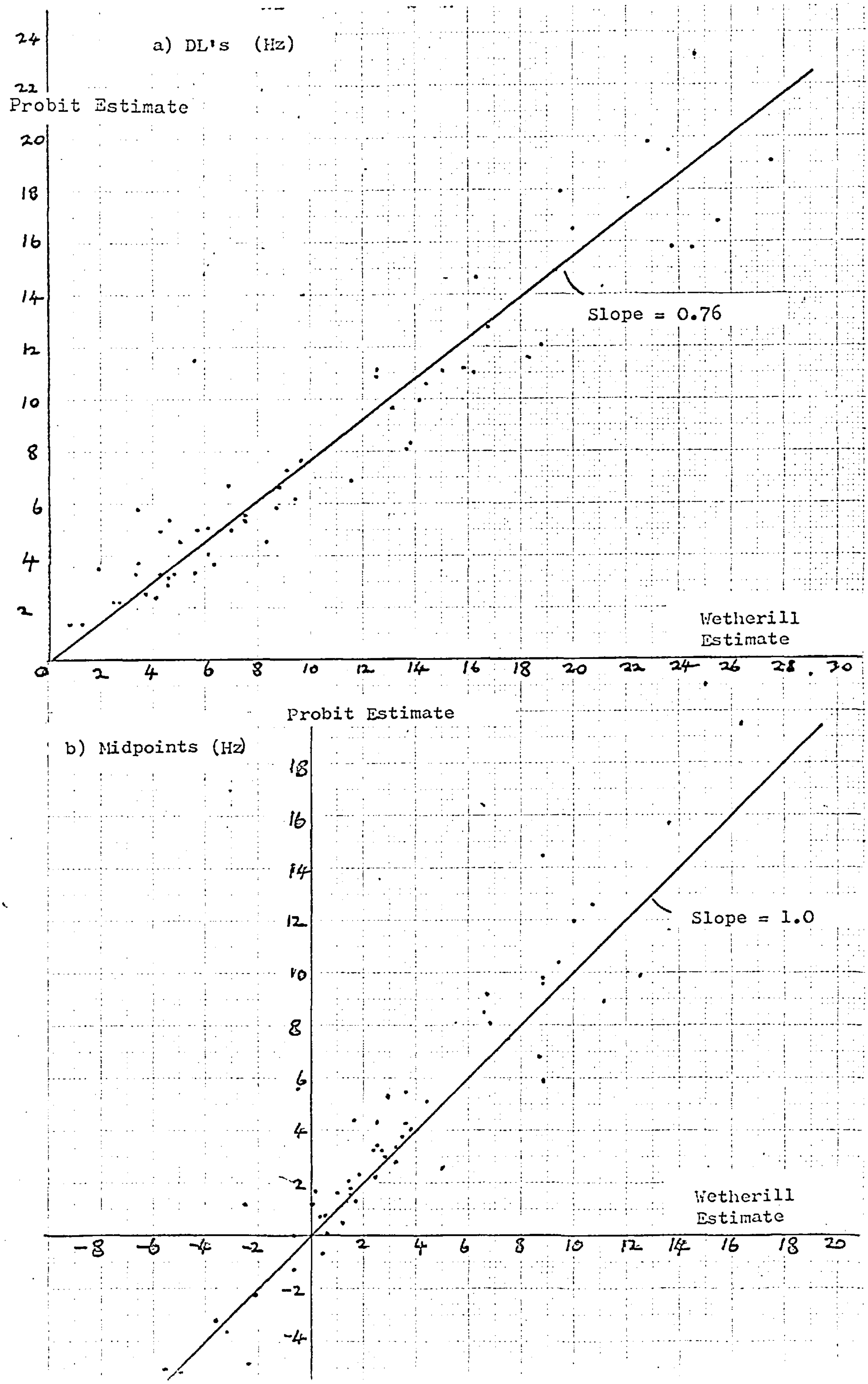


Fig. 23 Probit vs. Wetherill Estimates for Sequential Data.
(Experiment P-1)



Input:

A-X signal plus noise pulses from point Y in Fig. 4.

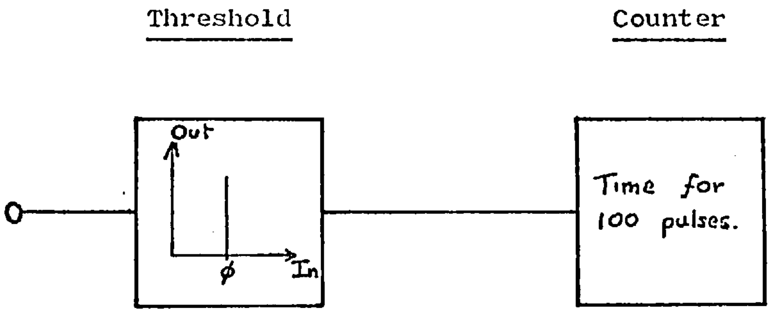


Fig. 24 Arrangement of Pulse Counting Frequency Discriminator (Experiment P-3).

Fig. 25a Error Distribution for 'Synthetic' DL's (Expt P-3).

(Level Spacing $\sim$ 0.4Δ).

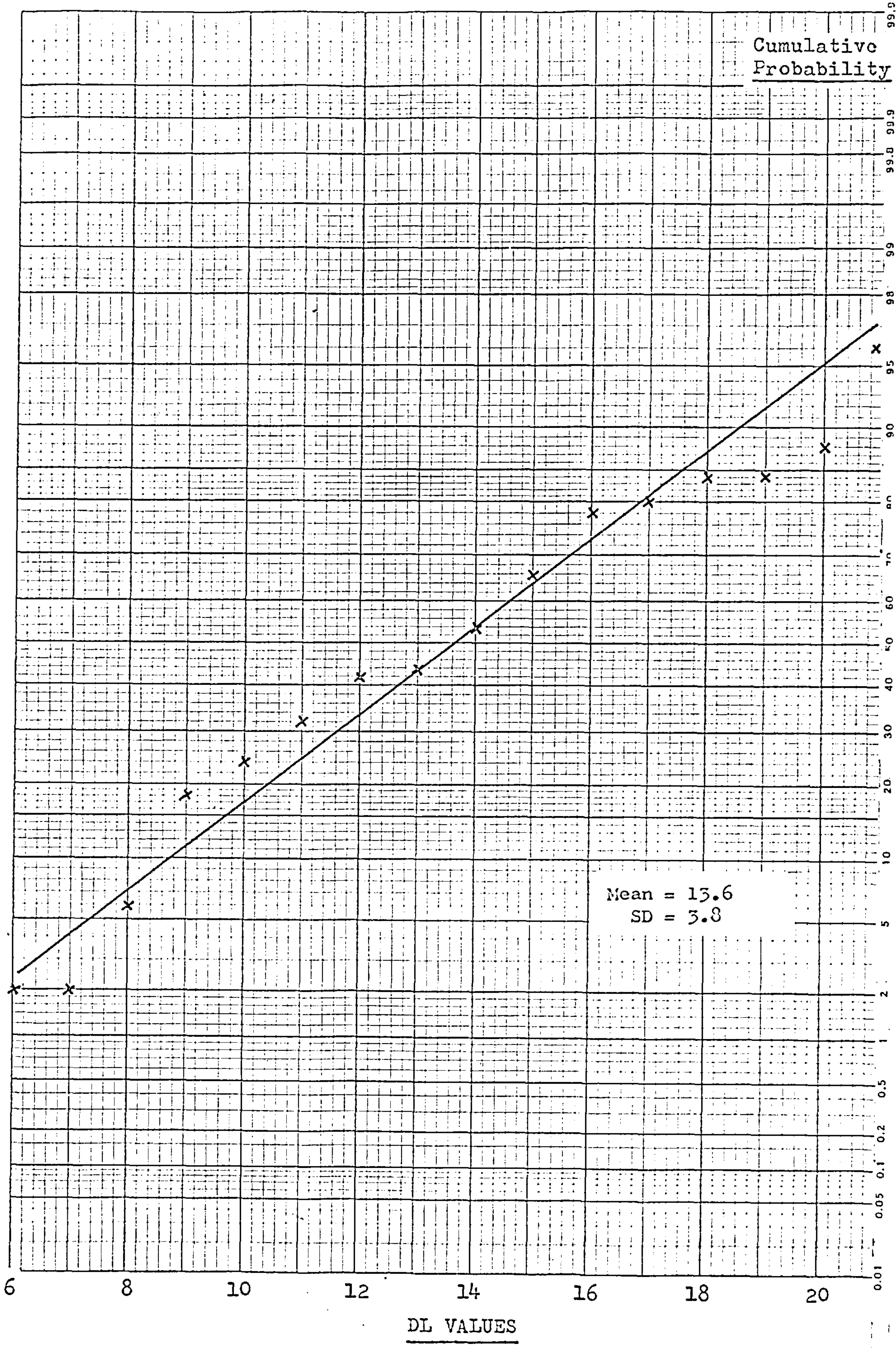


Fig. 25b Error Distribution for 'Synthetic' DL's (Expt P-3).

(Level Spacing ~ 0.8 Δ).

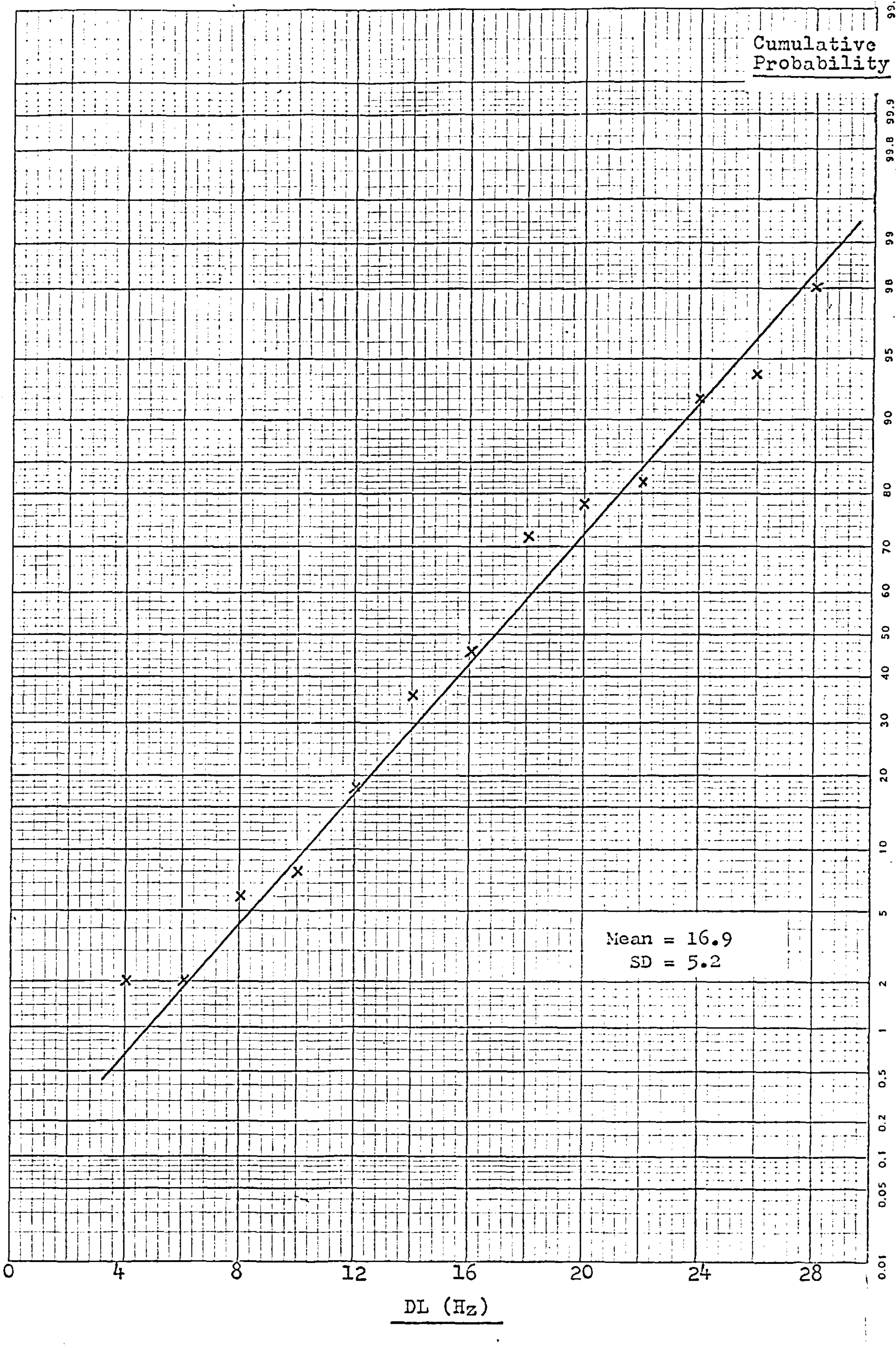


Fig. 25c Error Distribution for 'Synthetic' Midpoints (Expt P-3).

(Level Spacing ~ 0.4Δ).

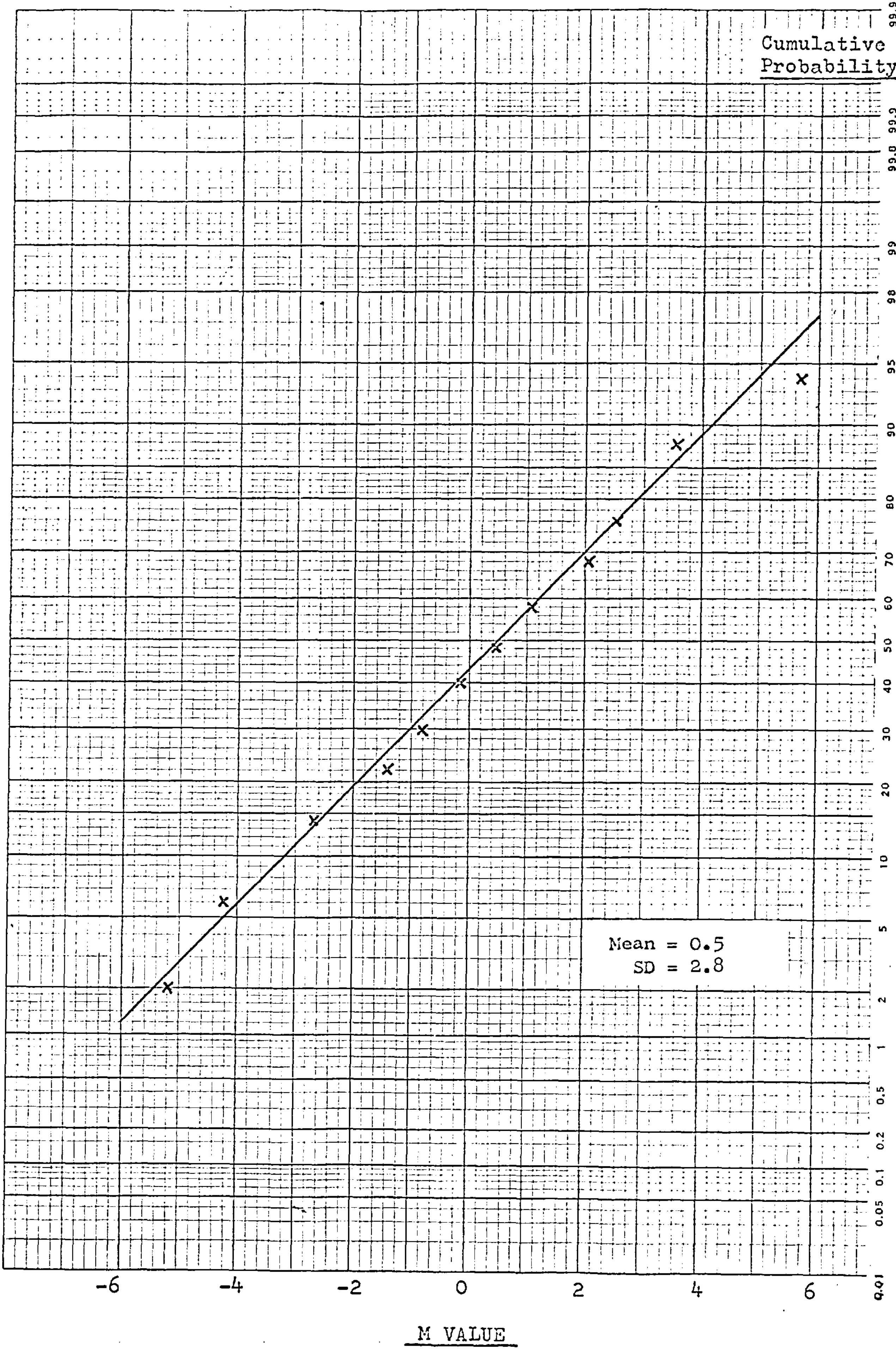


Fig. 25d Error Distribution for 'Synthetic' Midpoints (Expt P-3)

(Level Spacing ~ 0.8 Δ).

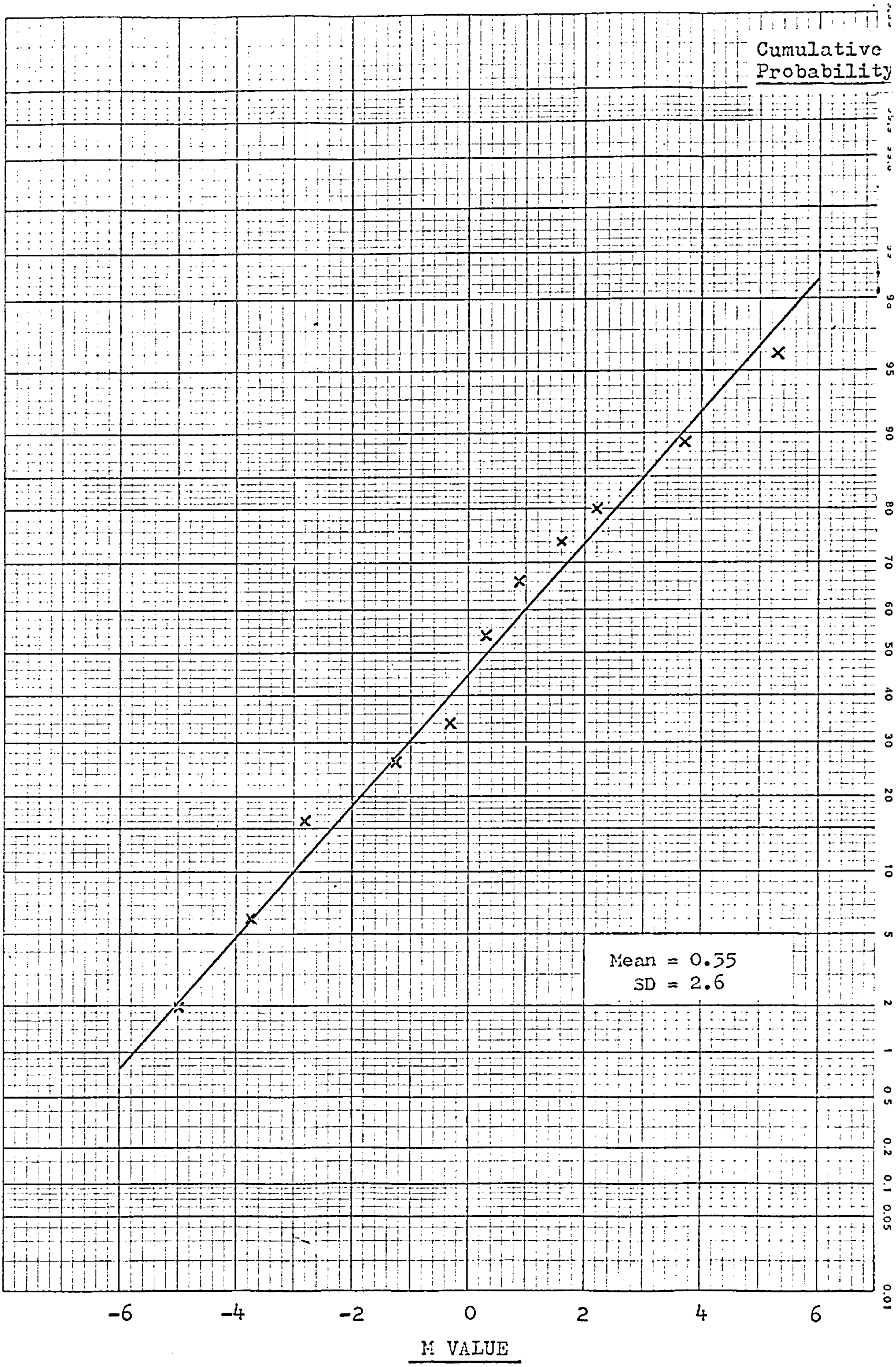


Fig. 26 Combined Results of Experiments P-1 and I.

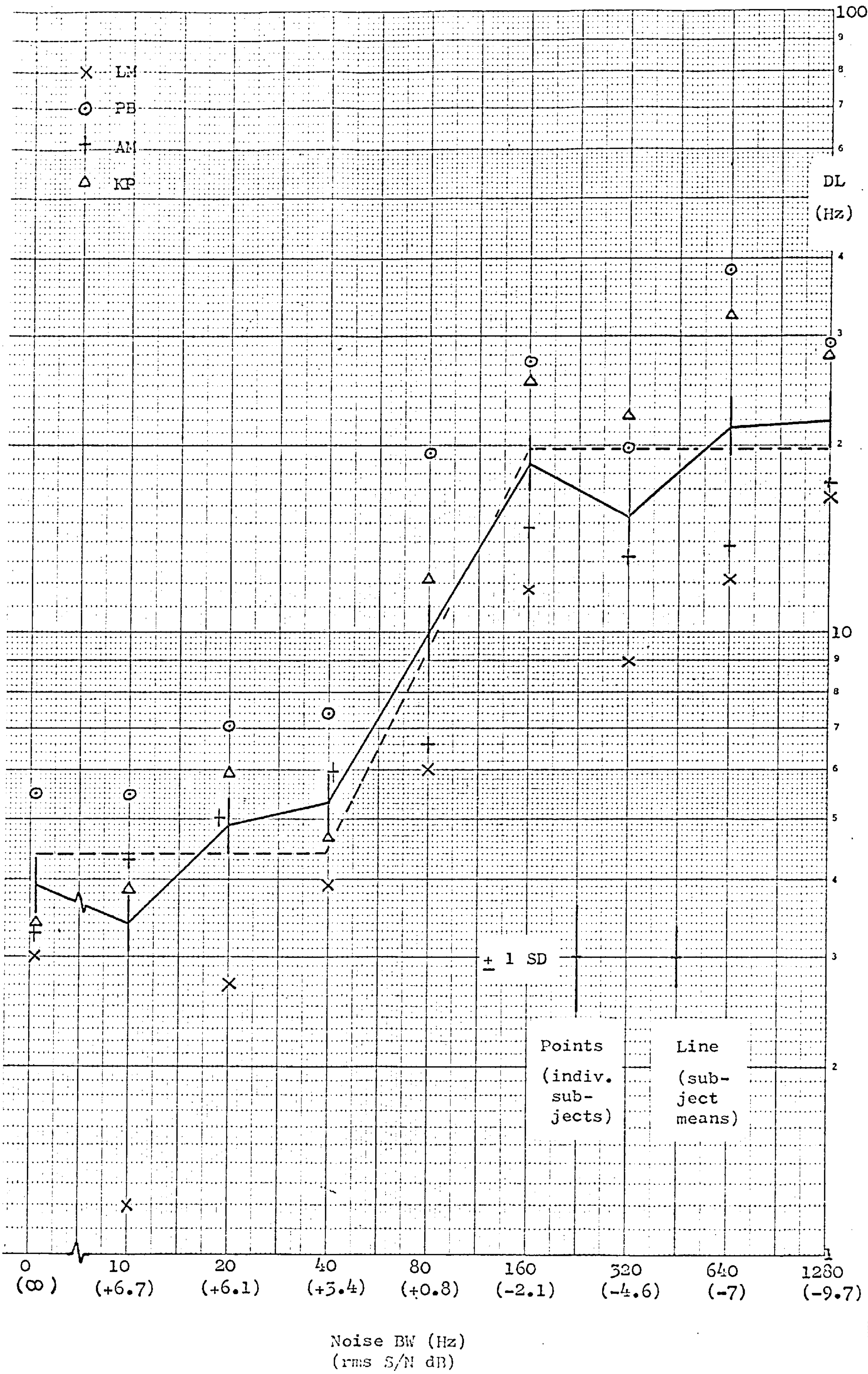


Fig. 27 Effect of Relative Spectral Overlap of Signal and Noise. (Experiment II-C)

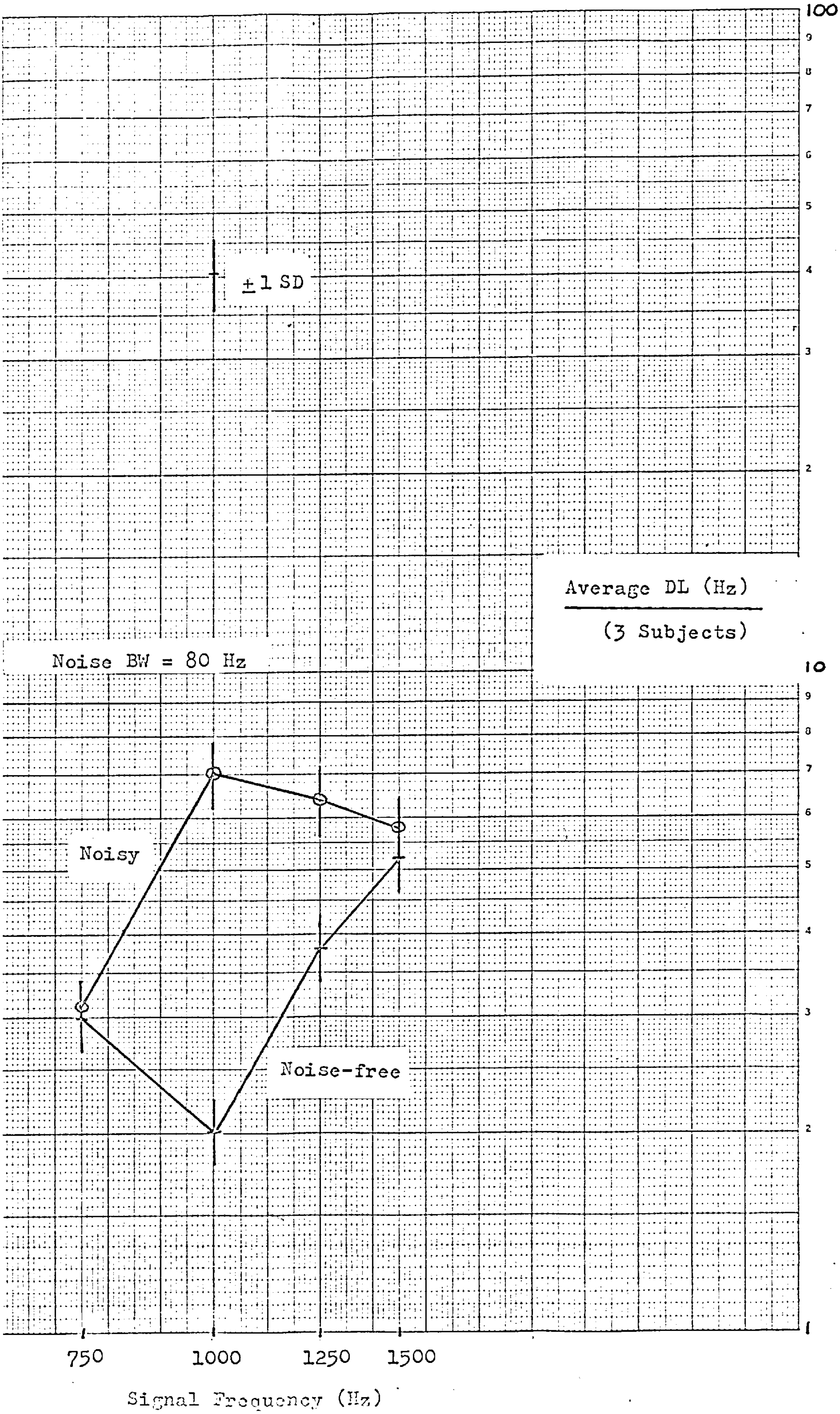


Fig. 28a Effects of Wideband Noise at 500 Hz (Table 23a).

(Experiment P-4)

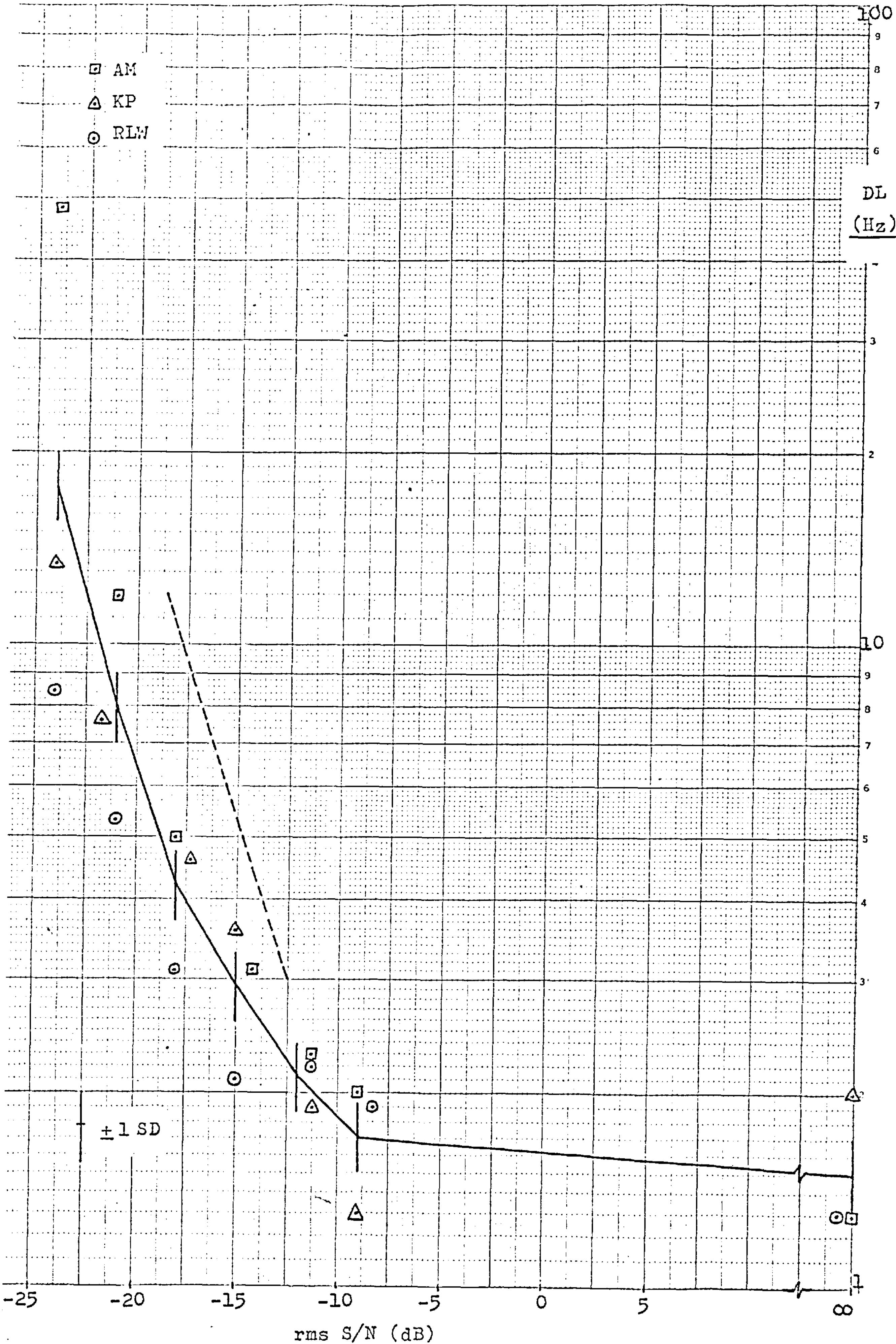


Fig. 28b Effects of Wideband Noise at 1 KHz (Table 23b).

(Experiment P-4)

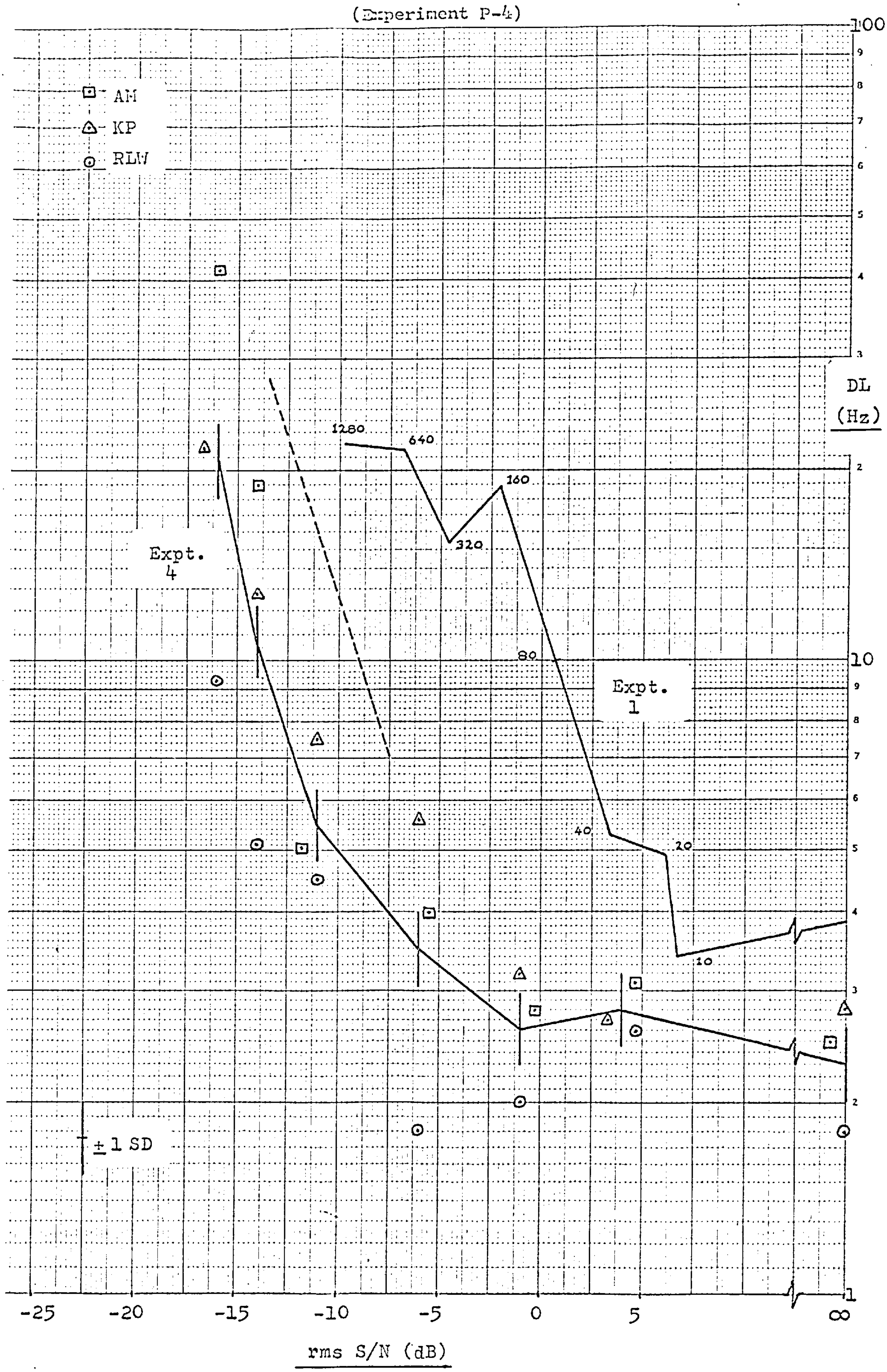


Fig. 28c Effects of Wideband Noise at 2 KHz (Table 25c).

(Experiment P-4)

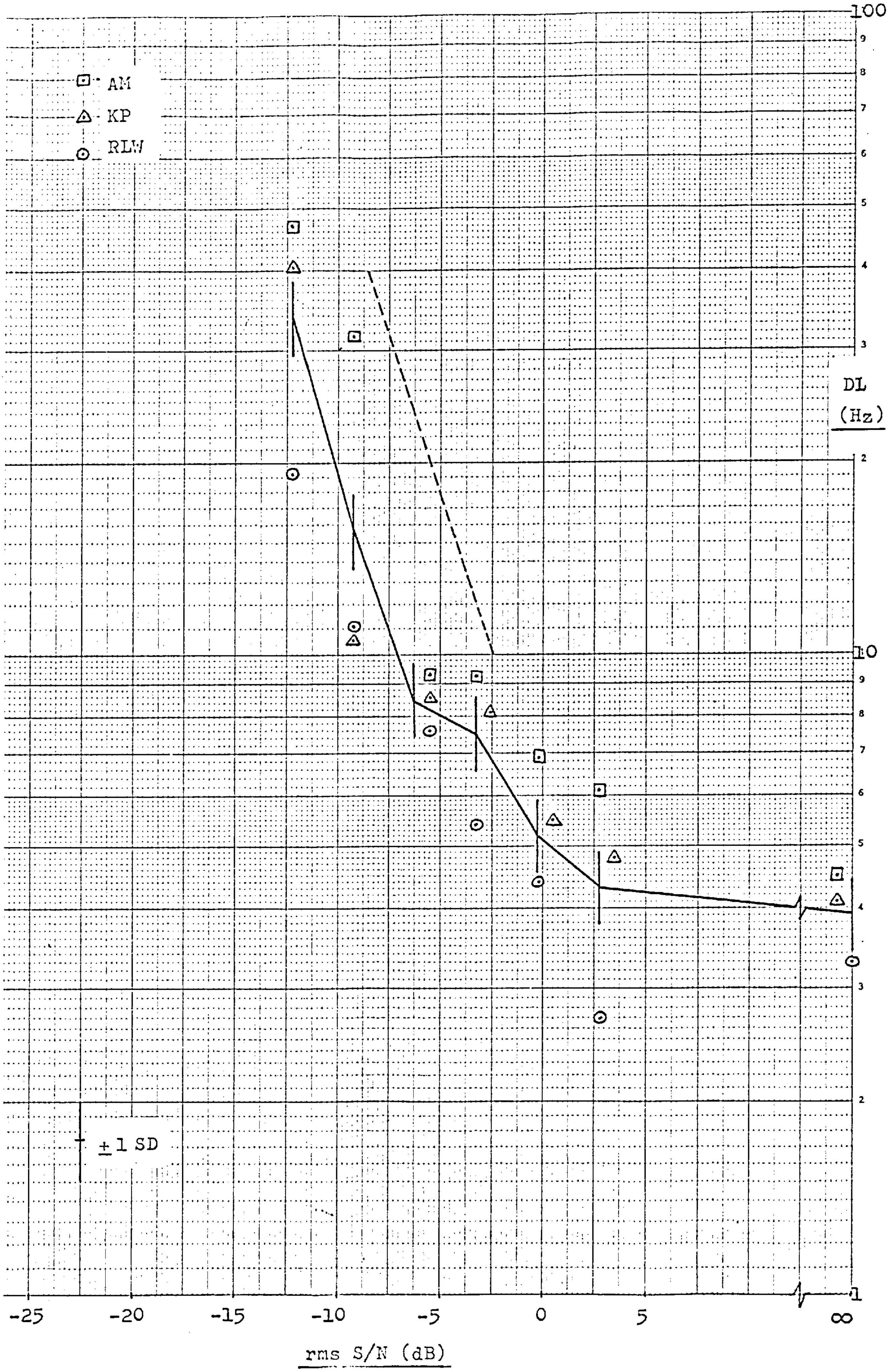


Fig. 28d Effects of Wideband Noise at 3 KHz (Table 23d).

(Experiment P-4)

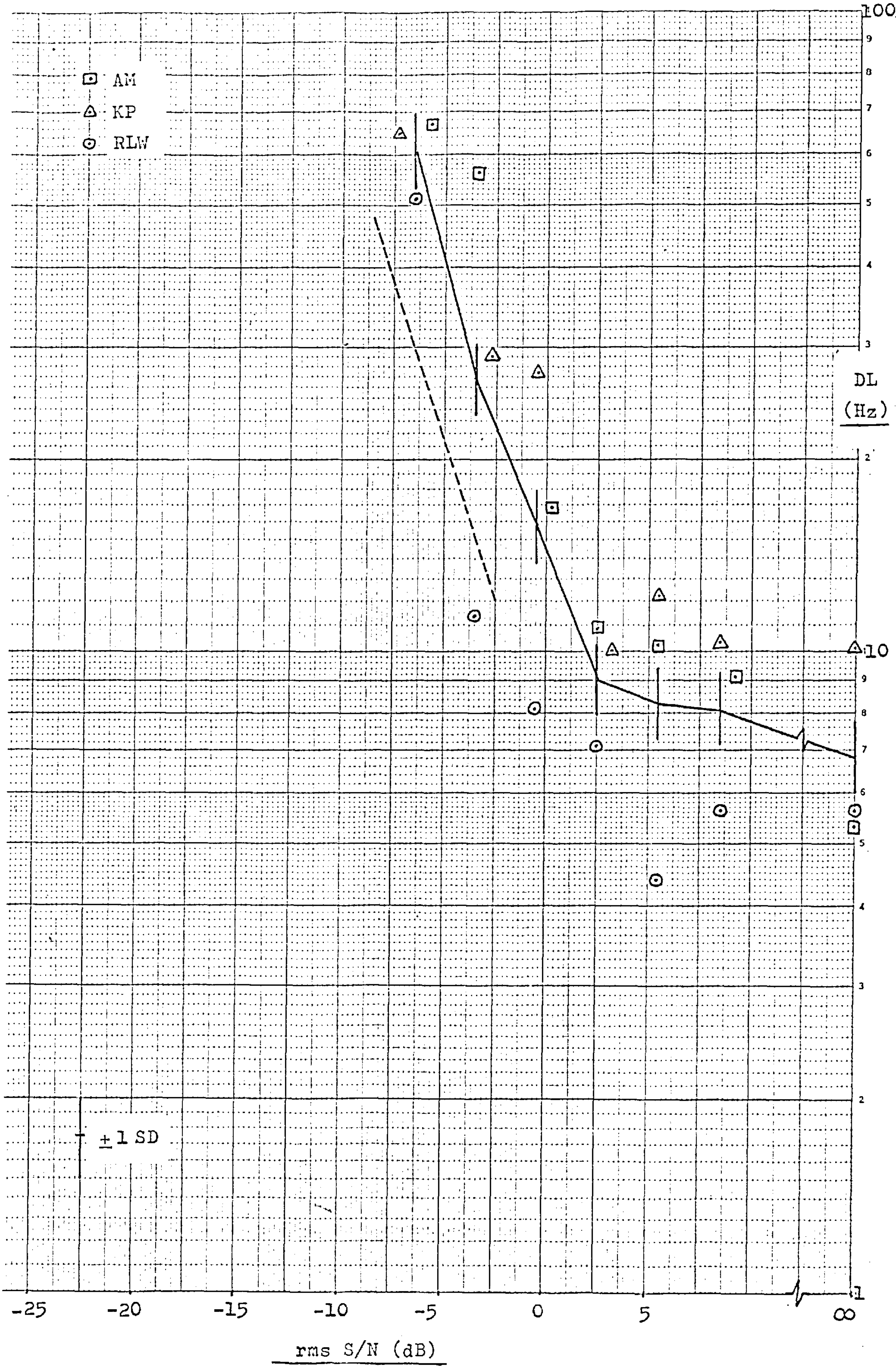


Fig. 29a The Effects of Wide and Narrow Band Noise at Low and High Signal Levels (Experiment V). 500 Hz (from table 25a)

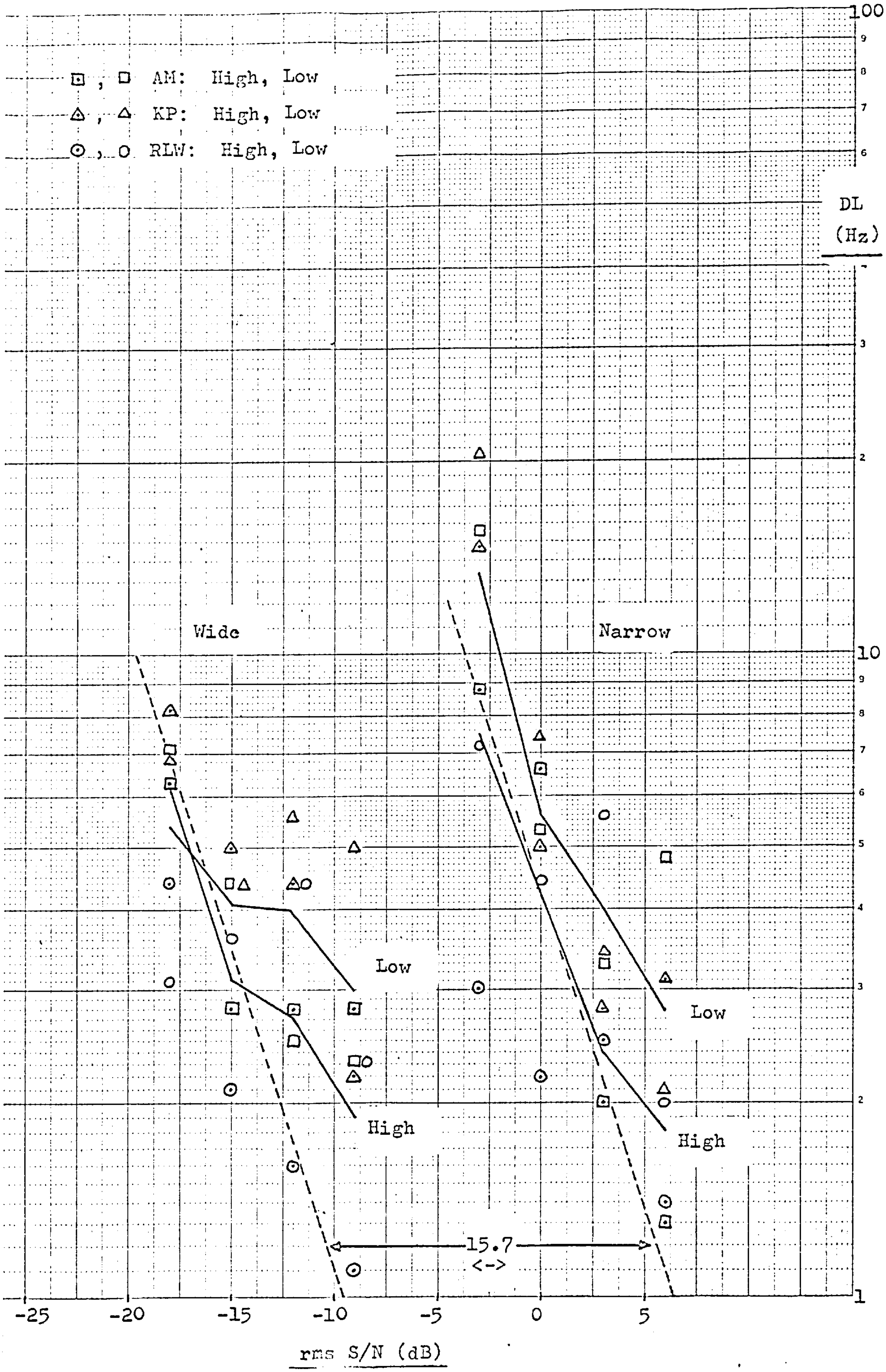


Fig. 29b The Effects of Wide and Narrow Band Noise at Low and High Signal Levels (Experiment V). 1 KHz (from table 25b)

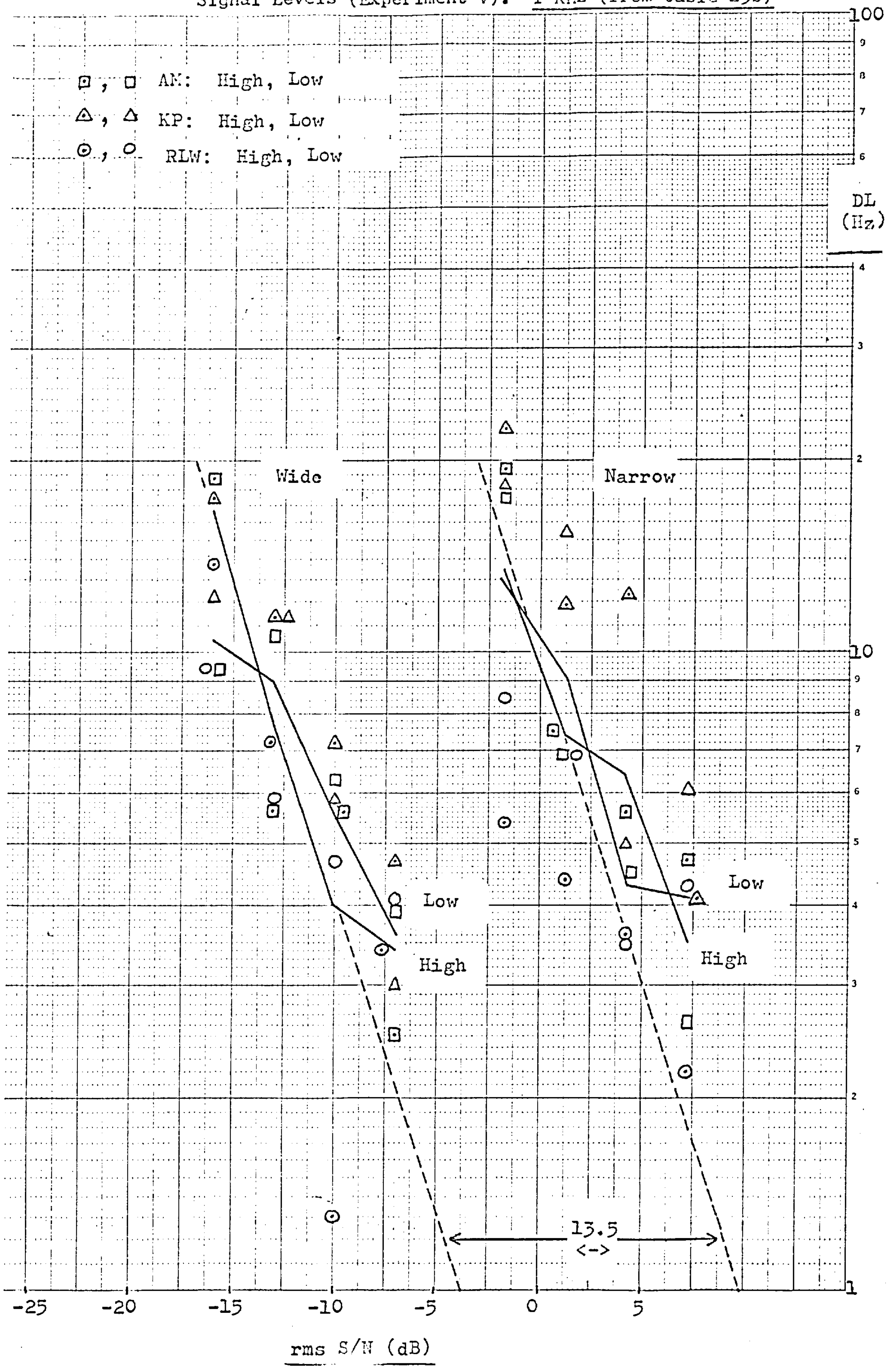


Fig. 29c. The Effects of Wide and Narrow Band Noise at Low and High Signal Levels (Experiment V). 2 KHz (from table 25c)

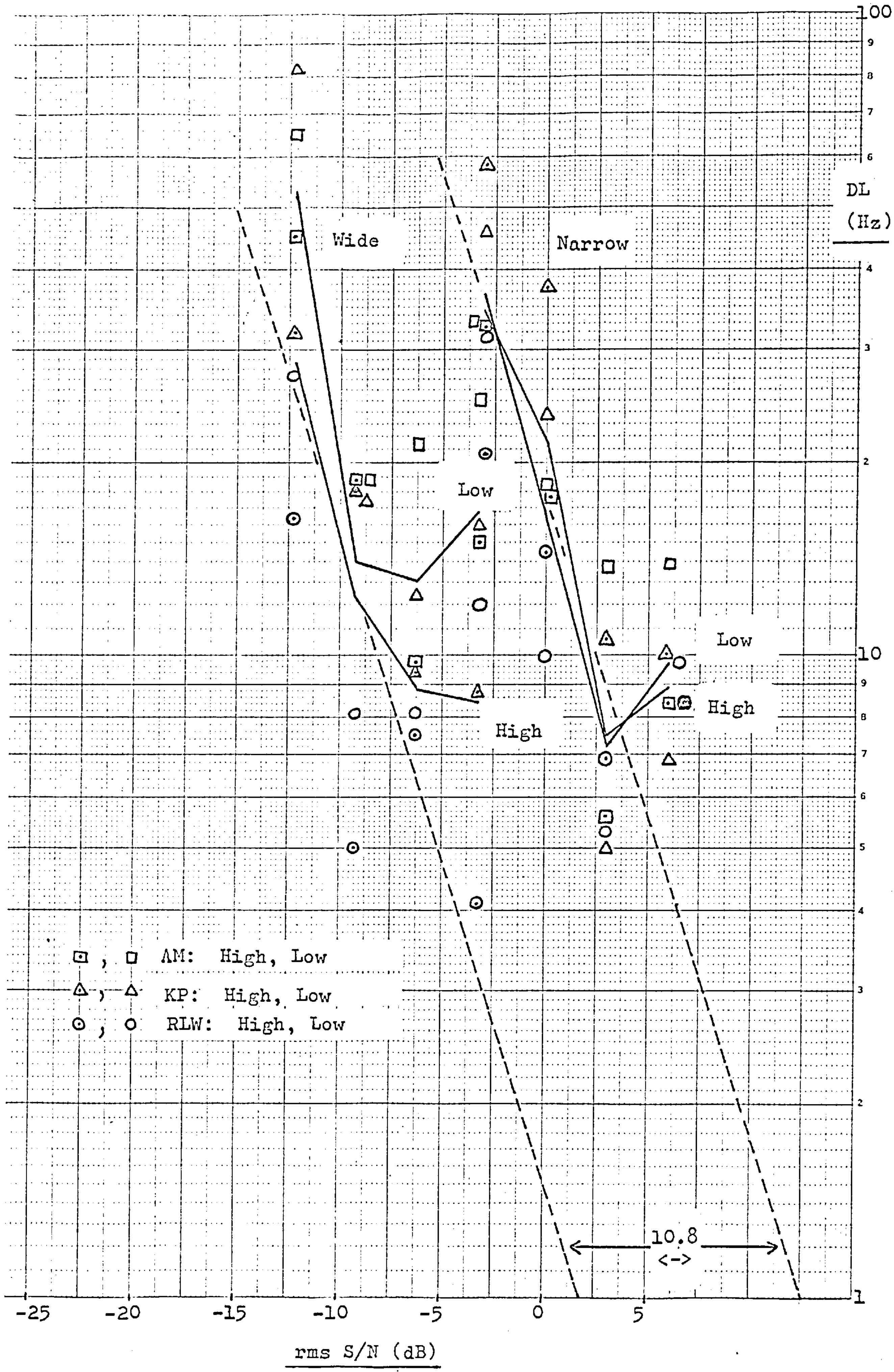
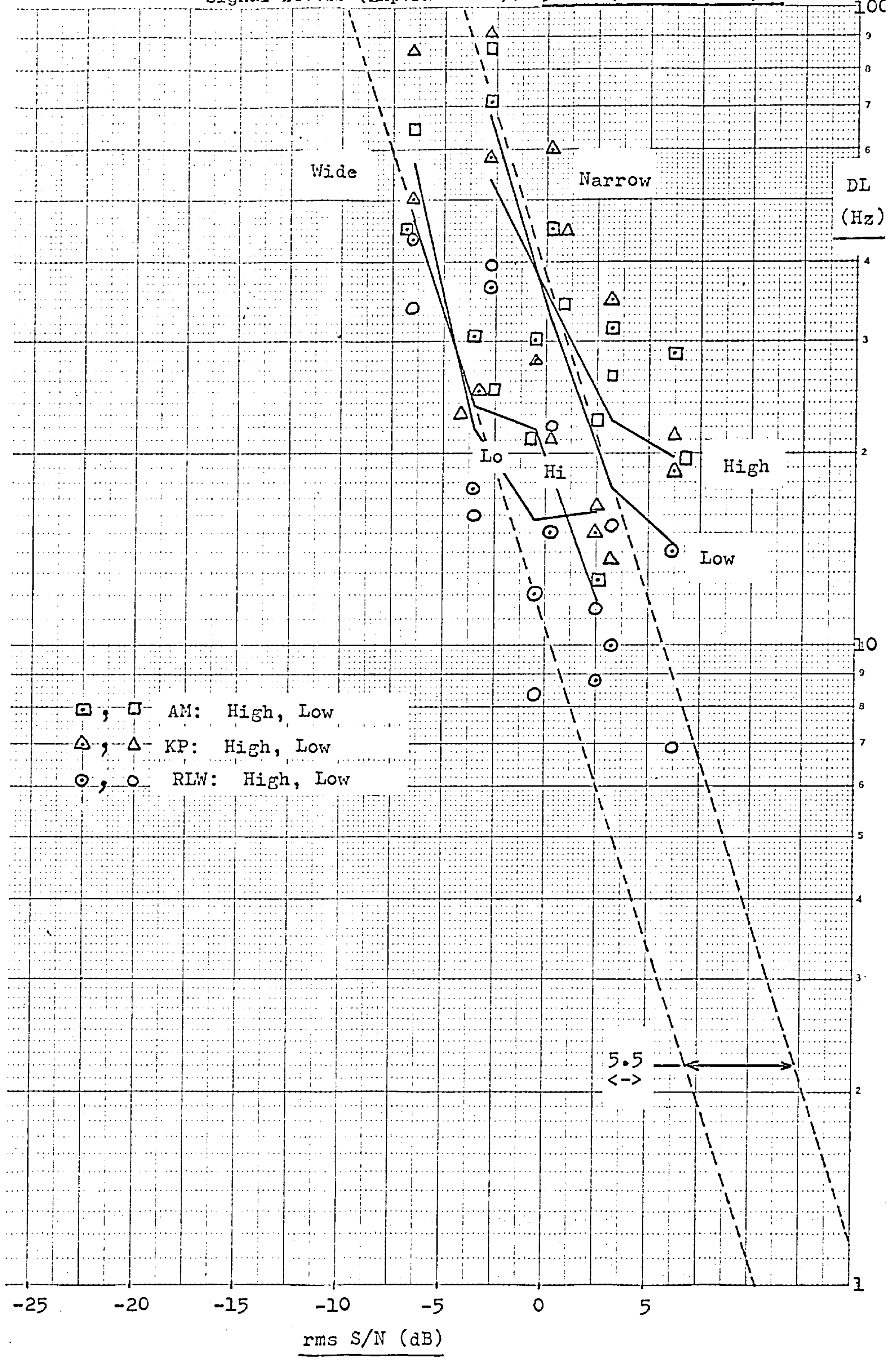


Fig. 29d The Effects of Wide and Narrow Band Noise at Low and High Signal Levels (Experiment V). 3 KHz (from table 25d)



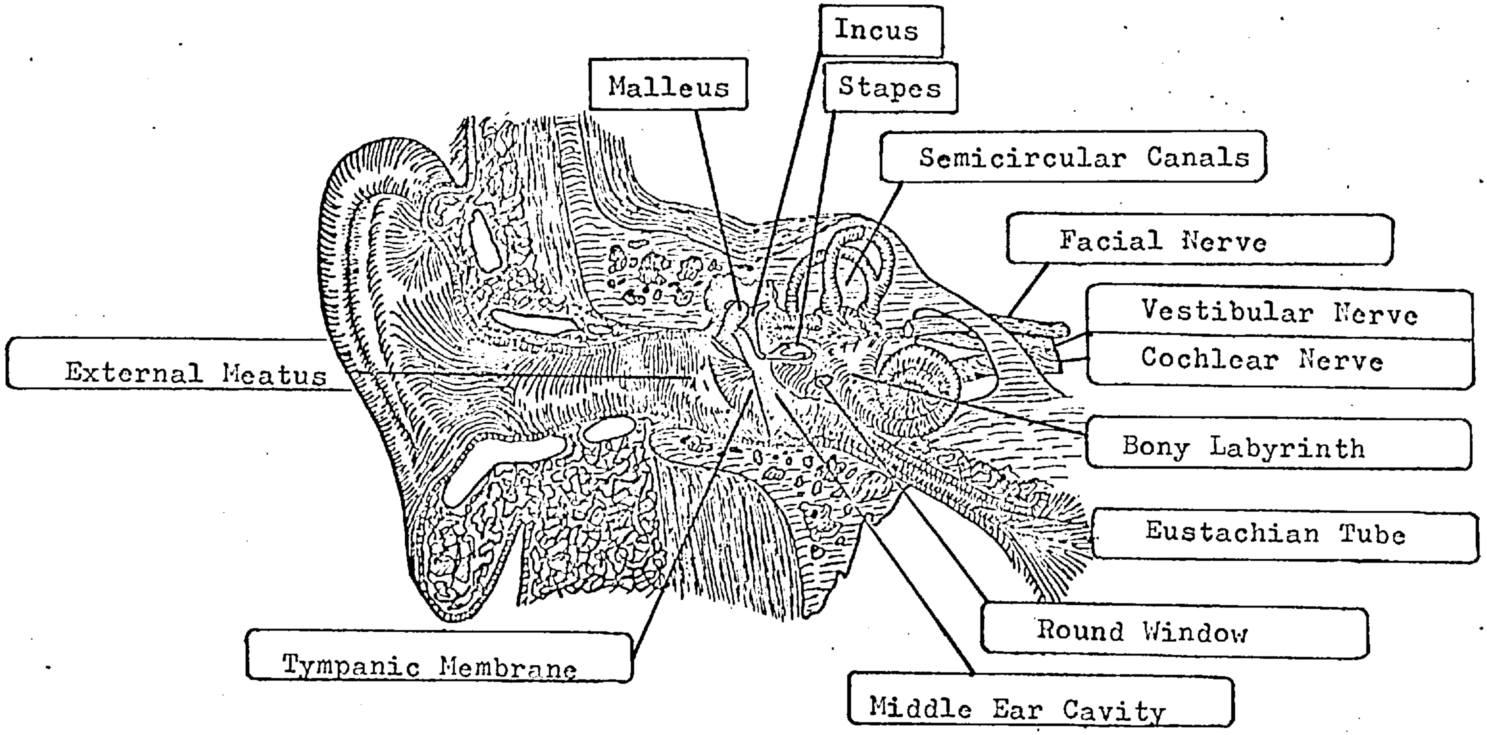
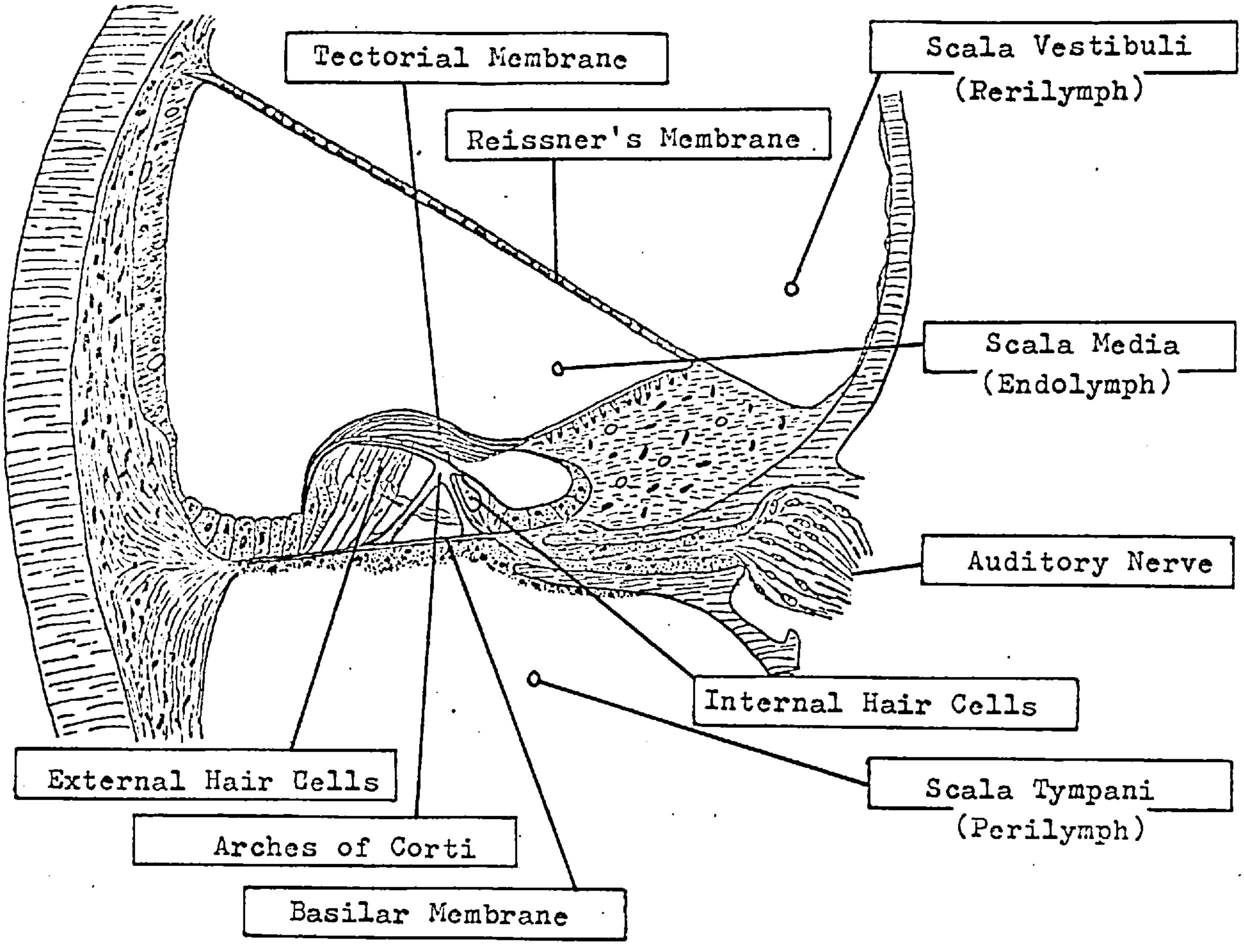


Fig. 30 Diagrammatic Section of the Peripheral Human Ear.

Fig. 31 Part Section of the Human Cochlea.



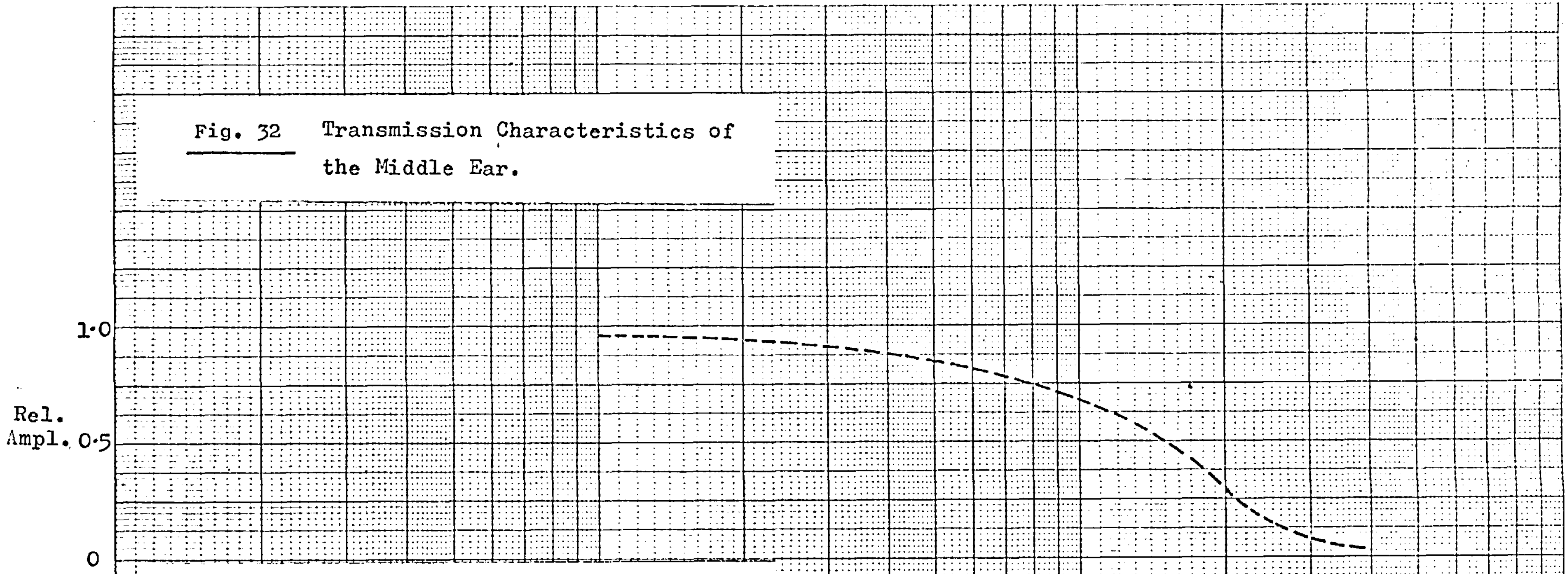


Fig. 32 Transmission Characteristics of the Middle Ear.

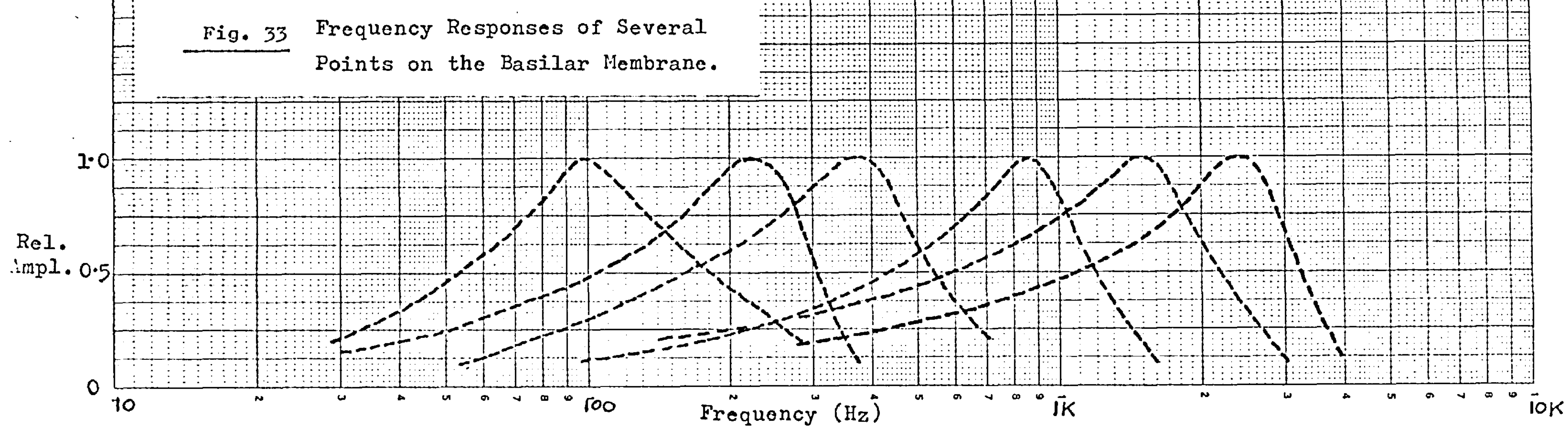


Fig. 33 Frequency Responses of Several Points on the Basilar Membrane.

Fig. 34 Frequency Responses of the Four 'Displacement' Filters.

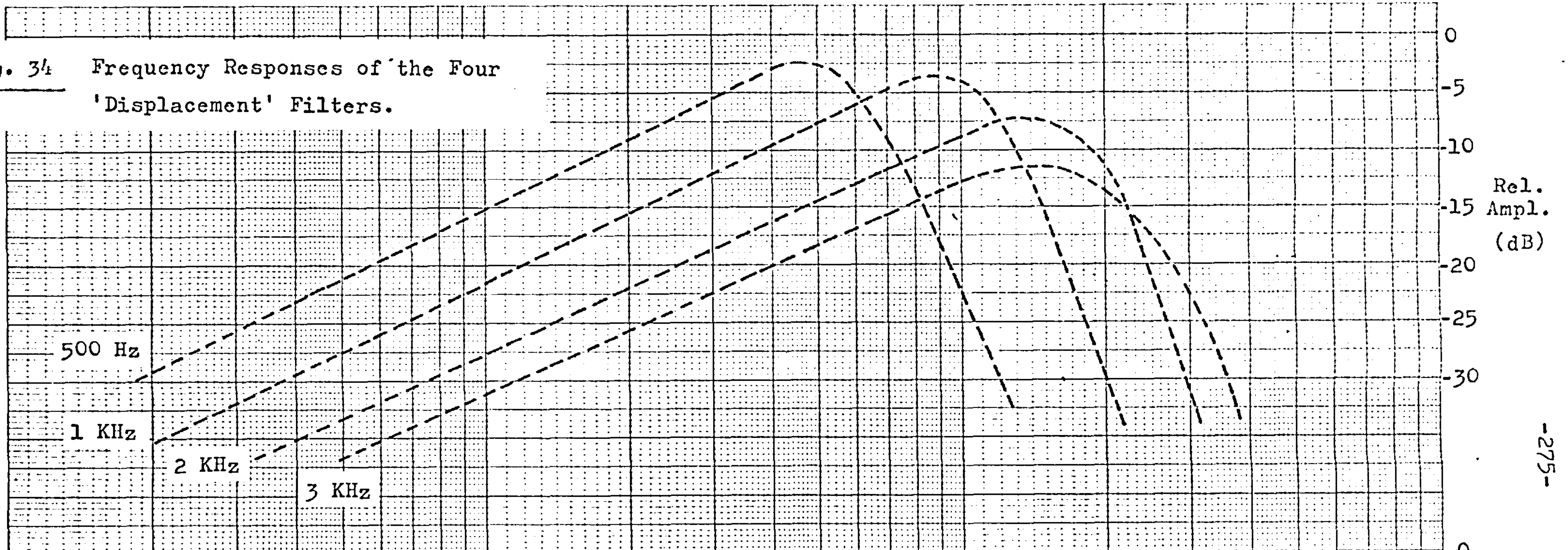
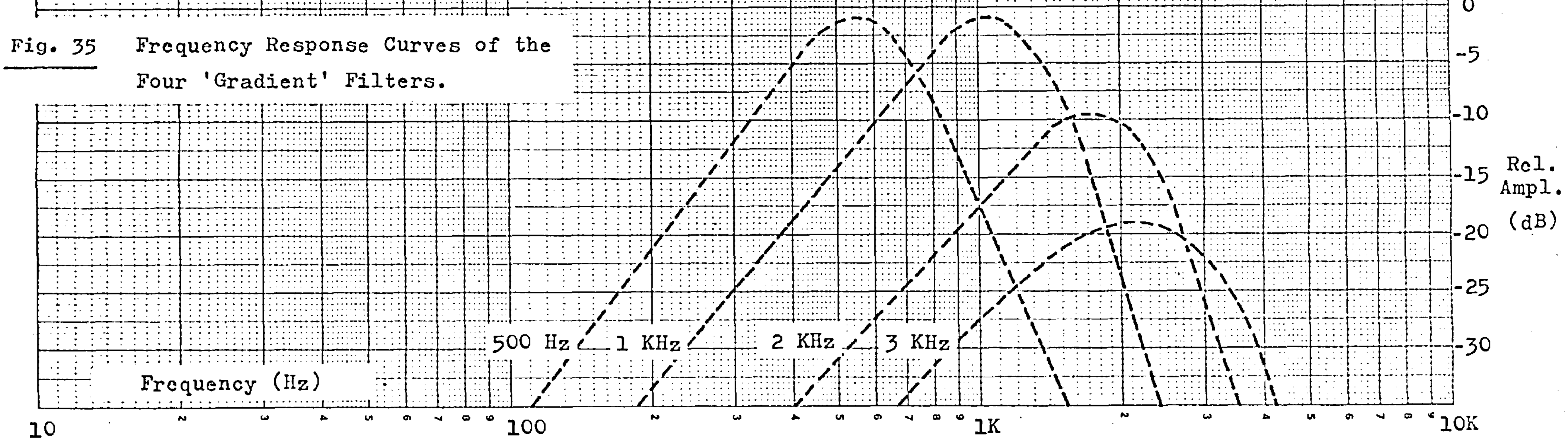


Fig. 35 Frequency Response Curves of the Four 'Gradient' Filters.



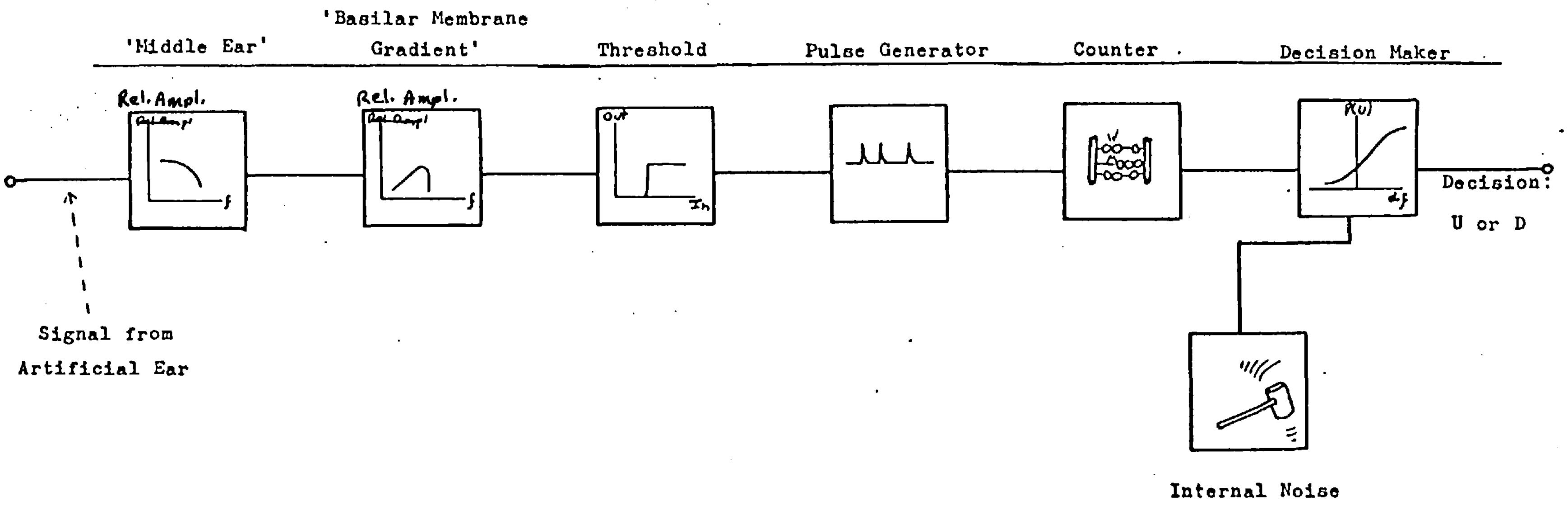
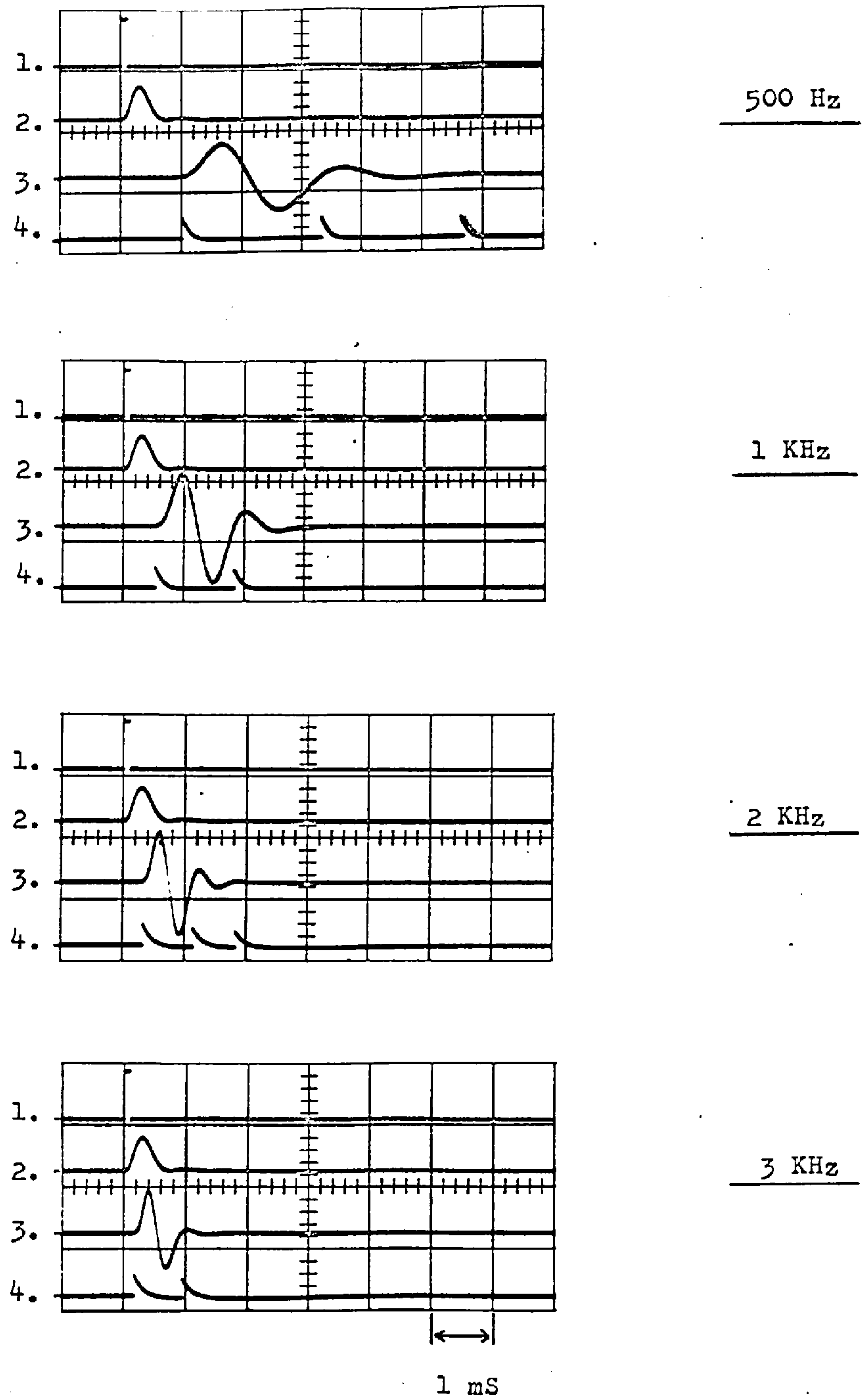


Fig. 36. Block Diagram of the Complete Model.

Fig. 37 Impulse Responses of Different Stages of the Ear Model.



1. Sound Impulse
2. Stapes Displacement
3. Membrane Displacement
4. Nerve Impulses

Fig. 38a Comparison of Model with Subjects at 500 Hz.
(Experiment VI-B)

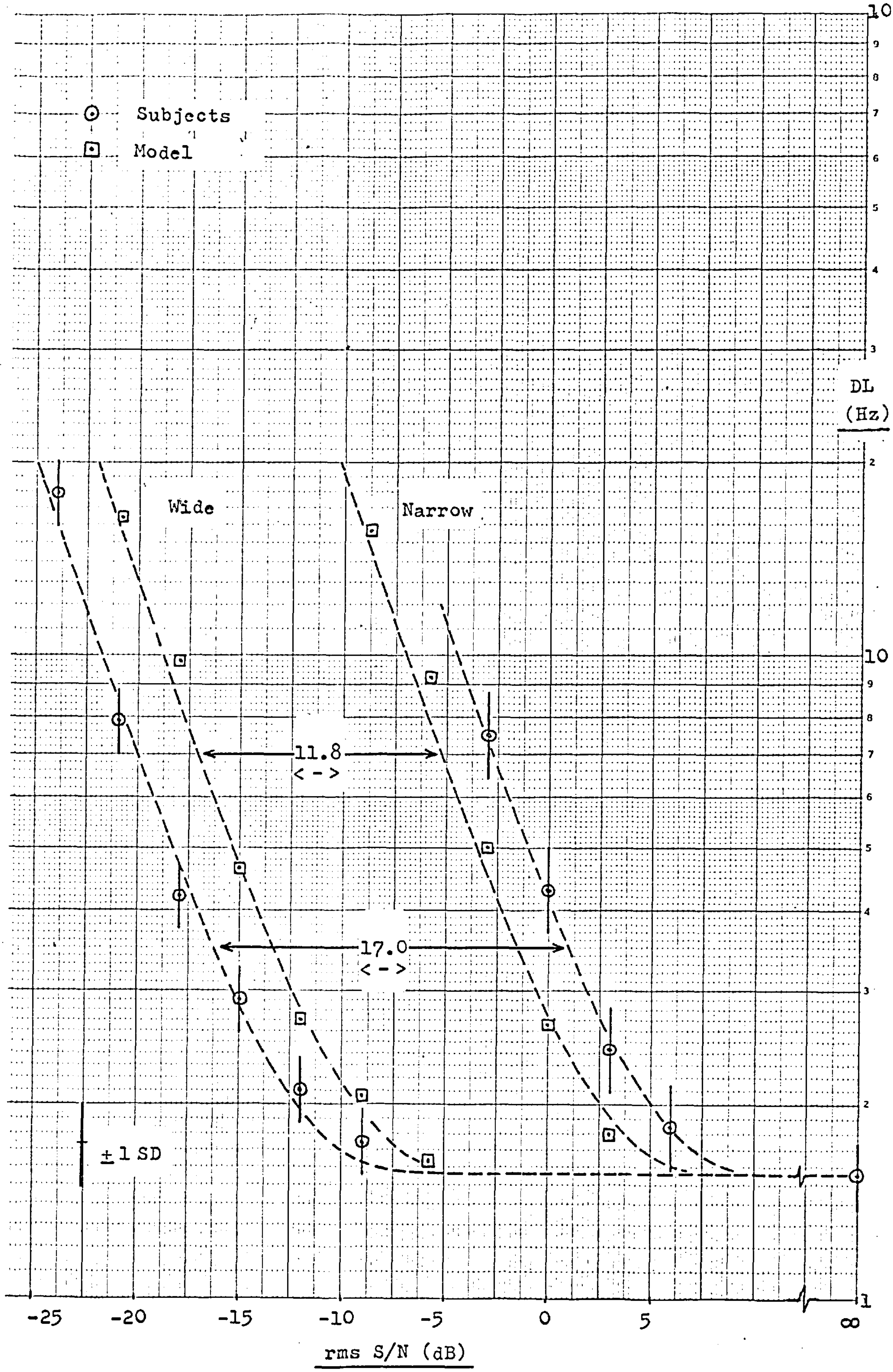


Fig. 38b Comparison of Model with Subjects at 1 KHz.
(Experiment VI-B)

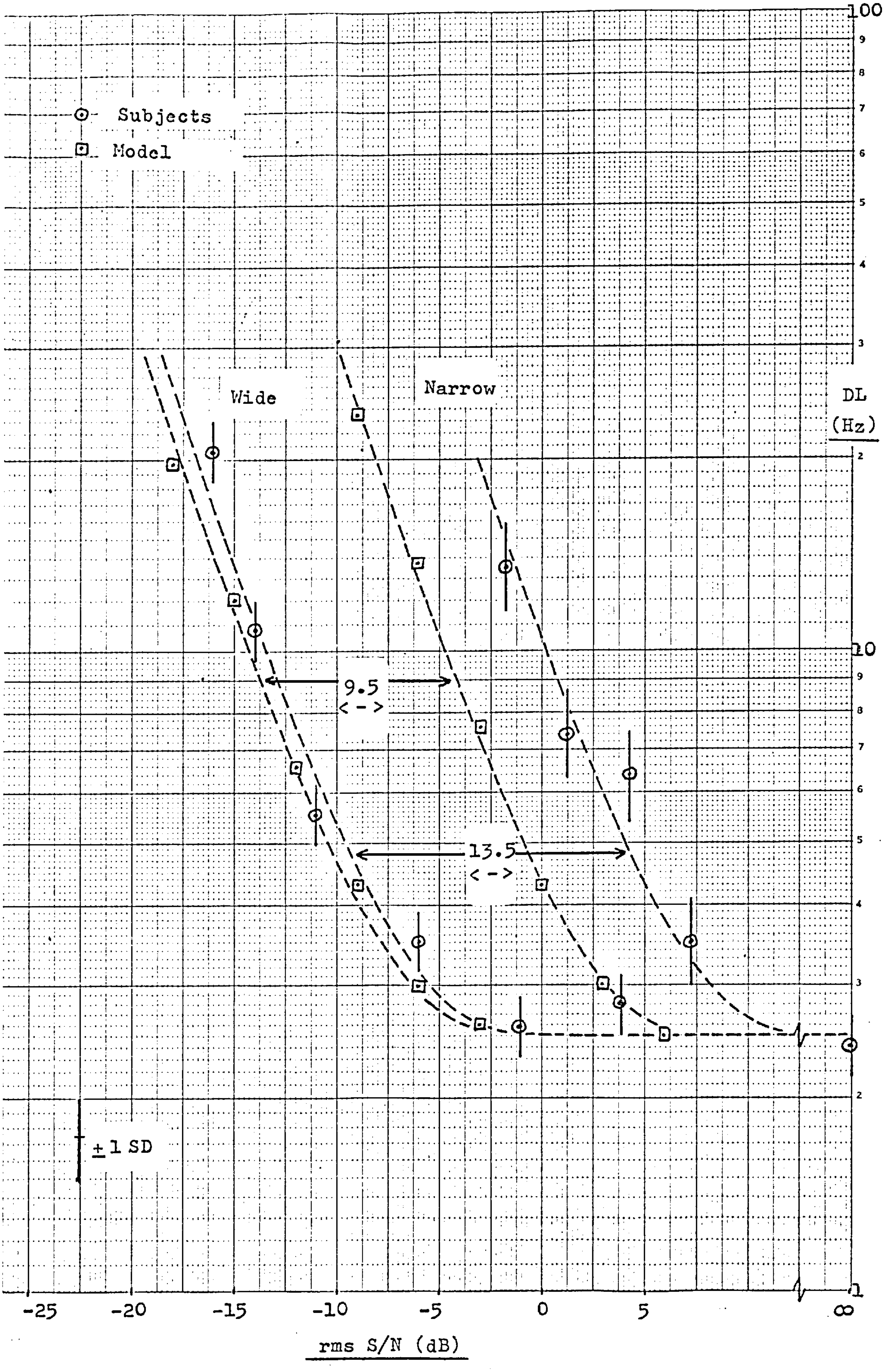


Fig. 38c Comparison of Model with Subjects at 2 KHz.

(Experiment VI-B)

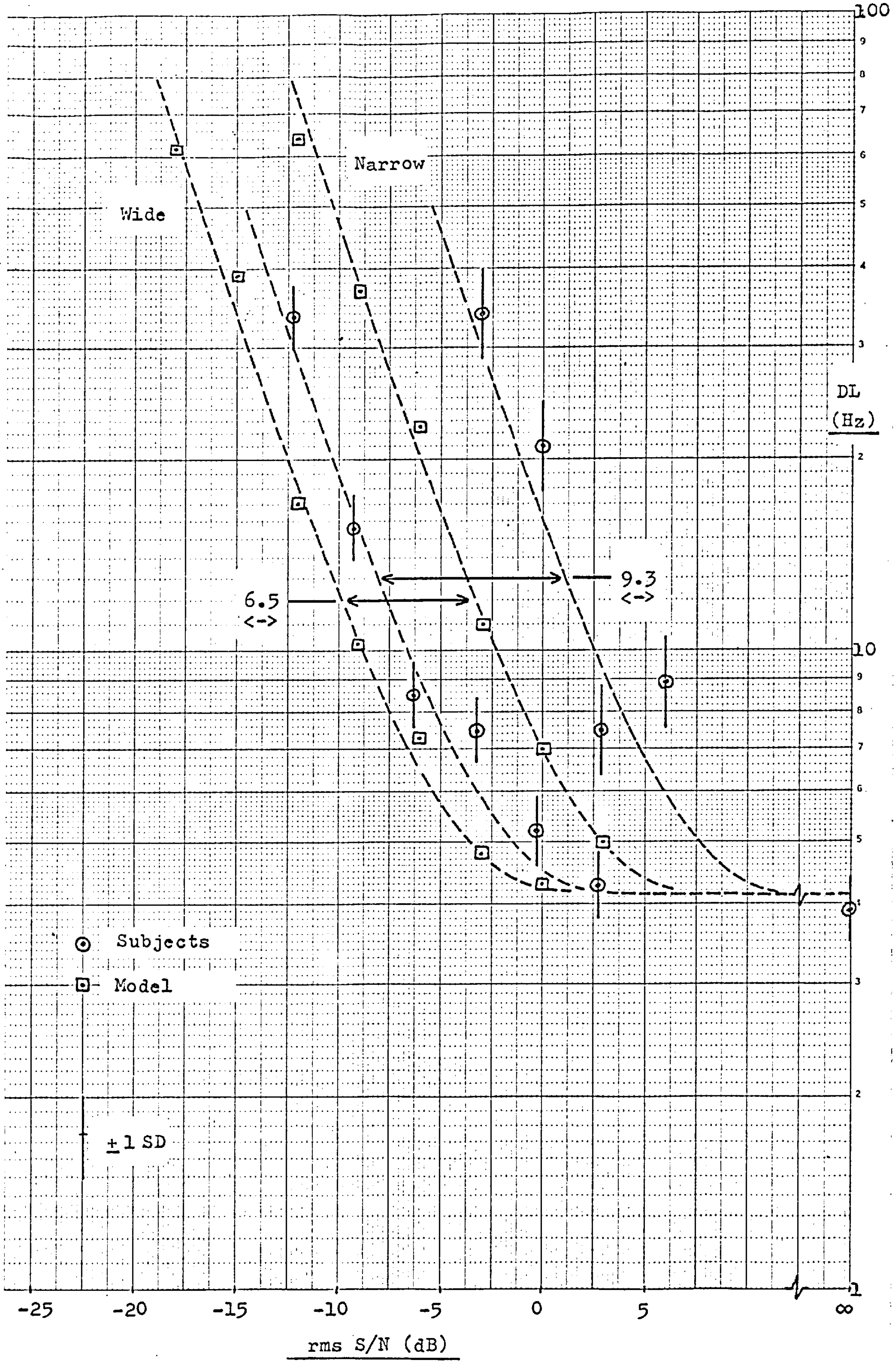
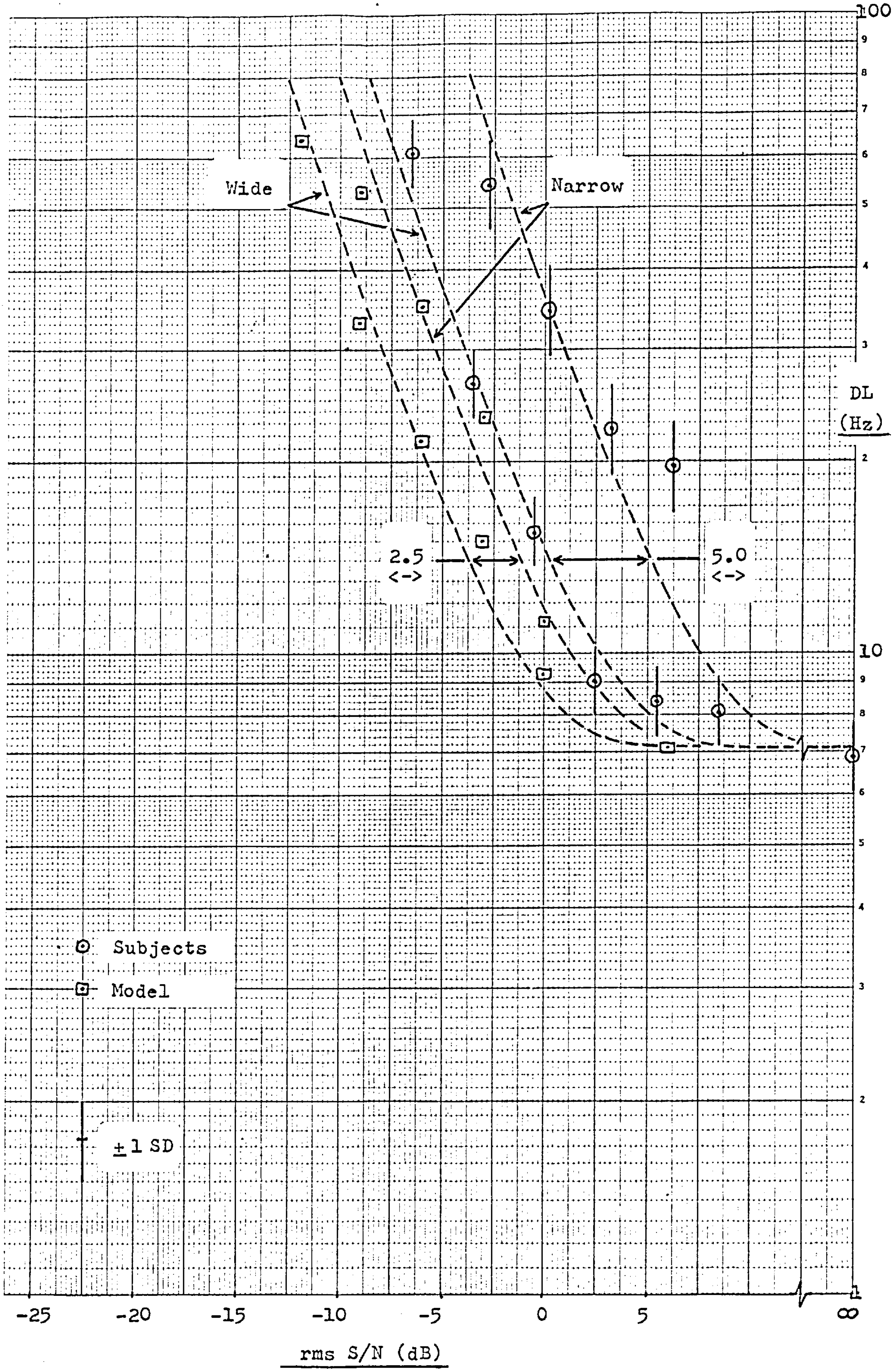


Fig. 38d Comparison of Model with Subjects at 3 KHz.

(Experiment VI-B)



LIST OF REFERENCES[+]

Ades, H.W. (1959) Central auditory mechanisms. (see Field, et al., 1959), p. 585-613.

Allanson, J.T. and Schenkel, K.D. (1965) The effect of band-limited noise on the pitch of pure tones. J. Sound & Vibration, 2: 402-408.

Atherley, G.R.C., Hempstock, T.I., and Noble, W.G. (1968) Study of tinnitus introduced temporarily by noise. JASA, 44: 1503-1506.

Bartlett, M.S. (1947) The use of transformations. Biometrics, 3: 39-53.

Bassett, I.G. and Eastmond, E.J. (1964) Echolocation measurement of pitch versus distance for sounds reflected from a flat surface. JASA, 36: 911-916.

von Bekesy, G. (1947) The Variation of phase along the basilar membrane with sinusoidal vibrations. JASA, 19: 452-460.

-(1949a) The vibration of the cochlear partition in anatomical preparations and in models of the inner ear. JASA, 21: 233-245.

-(1949b) On the resonance curve and the decay period at various points on the cochlear partition. JASA, 21: 245-254.

-(1953a) Description of some mechanical properties of the organ of Corti. JASA, 25: 770-785.

-(1953b) Shearing microphonics produced by vibrations near the inner and outer hair cells. JASA, 25: 786-791.

-(1960) Experiments in Hearing, trans. and edited by E.G. Wever, McGraw-Hill Co., New York.

-(1963a) Hearing theories and complex sounds. JASA, 35: 588-601.

-(1963b) Three experiments concerned with pitch perception. JASA, 35: 602-606.

[+] Note that 'JASA' refers to the Journal of the Acoustical Society of America.

Bilger, R.C. and Hirsh, I.J. (1956) Masking of tones by bands of noise. JASA, 28: 623-630.

Bilsen, F.A. (1966) Repetition pitch: monaural interaction of a sound with the repetition of the same, but phase shifted sound. Acustica, 17: 295-300.

-(1967) Phase sensitivity and (or?) short time analysis of the hearing organ. Acustica, 18: 182-186.

Blackwell, H.R. (1952) The influence of data collection procedures upon psychophysical measurement of two sensory functions. J. Exp. Psychol., 44: 306-315.

de Boer, E. (1962) Note on the critical bandwidth. JASA, 34: 985-986.

-(1967) Correlation studies applied to the frequency resolution of the cochlea. J. Auditory Res., 7: 209-217.

-and Bos, C.E. (1962) On the concept of the critical band. Paper H13, Proc. 4th Int. Congr. Acoust. Copenhagen; publ. Harland & Toksvig, Copenhagen.

Boring, E.G. (1950) A History of Experimental Psychology, Appleton-Century-Crofts, Inc., New York.

Bos, C.E. and de Boer, E. (1966) Masking and Discrimination, JASA, 39: 708-715.

Brandt, J.F. and Small, A.M. (1963) Difference Limen for frequency in the presence of masking. JASA, 35: 1881 (A).

van den Brink, G. (1964) Detection of tone pulses of various durations in noise of various bandwidths. JASA, 36: 1206-1211.

British Standard 2042 (1953) An artificial ear for the calibration of earphones of the external type. Brit. Standards Instit., London.

Broadbent, D.E. and Ladefoged, P. (1957) On the fusion of sounds reaching different sense organs. JASA, 29: 708-710.

Brownlee, K.A. (1960) Statistical Theory and Methodology in Science and Engineering, John Wiley & Sons, Inc., New York.

Brugge, J.F., Anderson, D.J., Hind, J.E., and Rose, J.E. (1969) Time structure of discharges in single auditory nerve fibres of the Squirrel Monkey in response to complex periodic sounds. J. Neurophysiol., 32: 386-401.

Campbell, R.A. (1964) Masker level and noise signal detection. JASA, 36: 570-575.

-(1969) Context and sequence effects with an adaptive threshold procedure. JASA, 46: 350-355.

Cardozo, B. Lopes- (1962) Frequency discrimination of the human ear. Paper H16, Proc. 4th Int. Congr. Acoustics, Copenhagen; pbul. Harland & Toksvig.

Cherry, E.C. (1961) Two ears - but one world. (see Rosenblith, 1961), p. 99-118.

Chistovich, L.A. (1960) Perception of sound sequence. Biofizika, 5: 671-676.

Cochran, W.G. (1947) Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics, 3: 22-38.

Cohen, A. (1961) Further investigations of the effects of intensity on the pitch of pure tones. JASA, 33: 1363-73.

Corliss, E.L.R. (1967) Mechanistic aspects of hearing. JASA, 41: 1500-1516.

Cornsweet, T.N. (1962) The staircase method in psychophysics. Am. J. Psychol., 75: 485-491.

Cramer, E.M. and Huggins, W.H. (1958) Creation of pitch through binaural interactions. JASA, 30: 413-417.

-and Licklider, J.C.R. (1957) Pitch of a chain of pulses of random polarity. JASA, 29: 780 (A).

Creelman, C.D. (1959) Detection of signals of uncertain frequency. Tech. Memo No. 71; Univ. of Michigan Res. Instit., (Dept. of Elec. Eng.), Ann Arbor.

-(1960) Applications of signal detectability theory to psychophysical research: a bibliography. Tech. Memo No. 79; Univ. of Michigan Res. Instit., (Dept. of Elec. Eng.), Ann Arbor.

-(1961) Detection of complex signals as a function of signal bandwidth and duration. JASA, 33: 89-94.

Curtiss, J.H. (1943) On transformations used in analysis of variance. Ann. Math. Stat., 14: 107-122.

David, E.E., Guttman, N., and van Bergeijk, W.A. (1959) Binaural interaction of high-frequency complex stimuli. JASA, 31: 774-782.

Davis, H. (1959) Excitation of auditory receptors. (see Field, et al., 1959), p. 565-584.

Delany, M.E. (1964) The acoustical impedance of human ears. J. Sound & Vibration, 1: 455-467.

-and Whittle, L.S. (1965) Design and validation of a new artificial ear. Paper B14, Proc. 5th Int. Congr. Acoustics, Liege (Ed. by Commins, D.E.).

Dewson, J.H. (1967) Efferent olivocochlear bundle: some relationships to noise masking and to stimulus attenuation. J. Neurophysiol., 30: 817-832.

-(1969) Efferent olivocochlear bundle: some relationships to stimulus discrimination in noise. J. Neurophysiol., 31: 122-130.

Dixon, W.J. and Mood, A.M. (1948) A method for obtaining and analyzing sensitivity data. J. Am. Stat. Assoc., 43: 109-126.

Egan, J.P. and Meyer, D.R. (1950) Changes in pitch of tones of low frequency as a function of the pattern of excitation produced by a band of noise. JASA, 22: 827-833.

Field, J., Magoun, H.W., and Hall, V.E. (Eds.) (1959) Handbook of Physiology - Section I - Neurophysiology, Vol. I, American Physiological Soc., Washington.

Finney, D.J. (1952) Probit Analysis, (2nd edition), Cambridge Univ. Press, Cambridge.

Fischler, H. (1967) Model of the 'secondary' residue effect in the perception of complex tones. JASA, 42: 759-64.

-and Cern, L. (1968) Simulation of the secondary residue effect by digital computer. JASA, 44: 1379-1385.

Flanagan, J.L. (1962) Computational model for basilar membrane displacement. JASA, 34: 1370-1376.

-and Guttman, N. (1960a) On the pitch of periodic pulses. JASA, 32: 1308-1319.

- -(1960b) Pitch of periodic pulses without fundamental component. JASA, 32: 1319-1328.

-and Saslow, M.G. (1958) Pitch discrimination for synthetic vowels. JASA, 30: 435-442.

Fletcher, H. (1924) The physical criterion for determining the pitch of a musical tone. Phys. Rev., 23: 427-437.

-(1940) Auditory patterns. Revs. Mod. Phys., 12: 47-65.

Fourcin, A.J. (1959) Speech perception and bandwidth compression. Report No. 1126, Signals Res. & Dev. Establishment, Ministry of Aviation.

-(1965) Pitch of noise with periodic spectral peaks. Paper B42, Proc. 5th Int. Congr. Acoustics, Liege, (Ed. by Commins, D.E.).

Freeman, J.A. and Nicholson, C.N. (1970) Space-time transformation in the frog cerebellum through an intrinsic tapped delay line. Nature, 226: 640-642.

Gabor, D. (1946) Theory of communication. J. Inst. Elec. Eng., 93(III): 429-457.

Green, D.M. (1957) Detection of multi-component auditory signals in noise. JASA, 29: 1257 (A).

-(1958) Detection of signals in noise and the critical band concept. Tech. Rep. No. 82, Univ. of Michigan (Electronic Defense Group), Ann Arbor.

-(1960) Psychoacoustics and detection theory. JASA, 32: 1189-1203.

-(1961) Detection of auditory sinusoids of uncertain frequency. JASA, 33: 897-903.

-, Birdsall, T.G., and Tanner, W.P. (1957) Signal detection as a function of signal intensity and duration. JASA, 29: 523-531.

-and Sewall, S.T. (1962) Effects of background noise on auditory detection of noise bursts. JASA 34: 1207-1216.

Greenberg, G.Z. and Larkin, W.D. (1968) Frequency response characteristics of auditory observers detecting signals of a single frequency in noise: the probe signal method. JASA, 44: 1513-1523.

Greenwood, D.D. (1961) Auditory masking and the critical band. JASA, 33: 484-502.

Guilford, J.P. (1954) Psychometric Methods, 2nd edition, McGraw-Hill (New York)-Kogakusha (Tokyo).

Guttman, N. and Flanagan, J.L. (1964) Pitch of highpass filtered pulse trains. JASA, 36: 757-765.

Hamilton, P.M. (1957) Noise masked thresholds as a function of tonal duration and masking noise bandwidth. JASA, 29: 506-511.

Harmon, W.W. (1963) Principles of the Statistical Theory of Communication. McGraw-Hill, New York.

Harris, G.G. (1963) Periodicity perception using gated noise. JASA, 35: 1229-1239.

Harris, J.D. (1947) Studies on pitch discrimination in masking II: the effect of signal-noise differential. JASA, 19: 816-819.

-(1948a) Pitch discrimination under masking. Am. J. Psychol., 61: 194-204.

-(1948b) Discrimination of pitch: suggestions towards method and procedure. Am. J. Psychol., 61: 309-322.

-(1952) Pitch discrimination. JASA, 24: 750-755.

-(1966) Masked DL for pitch memory. JASA, 40: 43-46.

Hawkins, J.E. and Stevens, S.S. (1950) The masking of pure tones and of speech by white noise. JASA, 22: 6-13.

Helmholtz, H. (1877) On the Sensations of Tone. (Trans. from the 4th German edition by A. Ellis, 1885) Dover Publications, Inc., New York, 1954.

Henning, G.B. (1967) Frequency discrimination in noise. JASA, 41: 774-777.

-and Grosberg, S.L. (1968) Effect of harmonic components on frequency discrimination. JASA, 44: 1386-1389.

Huggins, W.H. (1952) A phase principle for complex frequency analysis and its implications in auditory theory. JASA, 24: 582-589.

-and Licklider, J.C.R. (1951) Place mechanisms of auditory frequency analysis. JASA, 23: 290-99.

Jeffress, L.A. (1948) A place theory of sound localisation. J. Comp. Physiol. Psychol., 41: 35-39.

Johnstone, B.M. and Boyle, A.J.F. (1967) Basilar membrane vibration with Mössbauer technique. Science, 158: 389-390.

Katsuki, Y. (1961) Neural mechanism of auditory sensation in cats. (see Rosenblith, 1961), p. 561-583.

Kiang, N.Y.S., Goldstein, M.H., and Peake, W.T. (1962) Temporal coding of neural responses to acoustic stimuli. I.R.E. Trans. Info. Theory, February, p. 113-119.

Koester, T. (1945) The time error and sensitivity in pitch and loudness discrimination as a function of time interval and stimulus level. Arch. Psychol., 41: No. 297.

Konig, E. (1957) Effect of time on pitch discrimination thresholds. JASA, 29: 606-612.

Lawson, J.L. and Uhlenbeck, G.E. (Eds.) (1950) Threshold Signals. McGraw-Hill, New York; Chap. 7.

Leakey, D.M., Sayers, B. McA., and Cherry, E.C. (1958) Binarual fusion of low and high frequency sounds. JASA, 30: 222 (L).

Lee, Y.W. (1950) Application of statistical methods to communication problems. Tech. Rep. No. 181, Research Lab of Electronics, M.I.T.

Levitt, H. (1964) Discrimination of Sounds in Hearing. Ph.D. thesis, University of London (Faculty of Engineering).

-(in press) Decision theory, signal detection theory and psychophysics. Chapter in: Human Communication: A Unified View, (Eds. David, E.E. and Denes, P.B.).

Liang, Chic-an and Chistovich, L.A. (1960) Frequency difference limens as a function of tonal duration. Soviet Physics: Acoustics, 6: 75-80.

Licklider, J.C.R. (1951) A duplex theory of pitch perception. Experientia, 7: 128-134.

-(1958) Basic correlates of the auditory stimulus. In: Handbook of Experimental Psychology, (Stevens, S.S., ed.), Wiley, New York;

(NB: published in 1951, 2nd printing: 1958).

-(1959) Three auditory theories. In: Psychology: A Study of a Science, (Koch, S., ed.) McGraw-Hill, New York; p. 42-144.

Lindquist, E.F. (1956) Design and Analysis of Experiments in Psychology and Education. Houghton Mifflin, Co., Boston.

Marill, T. (1956) Detection theory and psychophysics. Tech. Rep. No. 319, Res. Lab of Electronics, M.I.T.

Michaels, R.M. (1957) Frequency difference limens for narrow bands of noise. JASA, 29: 520-522.

Miller, G.A. and Taylor, W.G. (1948) The perception of repeated bursts of noise. JASA, 20: 171-182.

Mills, A.W. (1960) Lateralisation of high frequency tones. JASA, 32: 132-134.

McClellan, M.E. and Small, A.M. (1966) Time separation pitch associated with noise pulses. JASA, 40: 570-582.

McClellan, M.E. and Small, A.M. (1967) Pitch perception of pulse pairs with random repetition rate. JASA, 41: 690-699.

Peterson, W.W., Birdsall, T.G., and Fox, W.C. (1954) The theory of signal detectability. I.R.E. Professional Gp. on Info. Theory; PGIT-4, p. 171-212.

Plomp, R. (1964) The ear as a frequency analyser. JASA, 36: 1628-36.

-(1967) Pitch of complex tones. JASA, 41: 1526-1533.

-and Levelt, W.J.M. (1965) Tonal consonance and critical bandwidth. JASA, 38: 548-560.

-and Mimpen, A.M. (1968) The ear as a frequency analyser II. JASA, 43: 764-767.

Pollack, I. (1968) Detection and relative discrimination of auditory jitter. JASA, 43: 308-315.

Postman, L. (1946) Time error in auditory perception. Am. J. Psychol., 59: 193-219.

de Reuck, A.V.S. and Knight, J. (Eds.) (1968) Hearing Mechanisms in Vertebrates. Ciba Symp.; J.&A. Churchill, London.

Rice, S.O. (1948) Properties of a sine wave and random noise. Bell Sys. Tech. J., XXVII: 109-130.

Riesz, R.R. (1928) Differential Intensity sensitivity of the ear for pure tones. Phys. Rev., 31: 867-875.

Ritsma, R.J. (1962) Existence region of the tonal residue I. JASA, 34: 1224-1229.

-and Engel, F.L. (1964) Pitch of frequency modulated signals. JASA, 36: 1637-1644.

Rose, J.E., Brugge, J.F., Anderson, D.J., and Hind, J.E. (1968) Patterns of activity in single auditory nerve fibres of the Squirrel Monkey. (see de Reuck and Knight, 1968), p. 144-168.

Rosenberg, A.E. (1965) Effects of masking on the pitch of periodic pulses. JASA, 38: 747-758.

Rosenblith, W.A. (Ed.) (1961) Sensory Communication. M.I.T. Press, Massachusetts.

Sayers, B. McA. and Cherry, E.C. (1957) Mechanism of binaural fusion in the hearing of speech. JASA, 29: 973-987.

Schafer, T.H., Gales, R.S., Shewmacher, C.A., and Thompson, P.O. (1950) The frequency selectivity of the ear as determined by masking experiments. JASA, 22: 490-96.

Scharf, B. (1961) Complex sounds and critical bands. Psychol. Bull., 58: 205-217.

-(1966) Critical bands. Special Rept. LSC-S-3, Lab. of Sensory Commun.; Syracuse Univ., New York.

Schodder, G.R. and David, E.E. (1960) Pitch discrimination of two-frequency complexes. JASA, 32: 1426-1435.

Schouten, J.F. (1940a) The residue, a new component in subjective sound analysis. Proc. Kon. Ned. Akad. Wetenschappen, 43: 356-365.

-(1940b) The residue and the mechanism of hearing. ibid., p. 991-999.

-(1940c) The perception of pitch. (Philips Tech. Rev., 5: 286-294.

-, Ritsma, R.J., and Cardozo, B. Lopes- (1962) On the pitch of the residue. JASA, 34: 1418-1424.

Schroeder, M.R. (1966) Residue pitch: a remaining paradox and a possible explanation. JASA, 40: 79-81.

Schubert, E.D. (1950) The effect of a thermal masking noise on the pitch of a pure tone. JASA, 22: 497-499.

Sekey, A.R. (1962) A study of Auditory Perception in the Time-Frequency Domain. Ph.D. thesis, Univ. of London, (Faculty of Engineering).

-(1963) Short-term auditory frequency discrimination. JASA, 35: 682-690.

Shaw, E.A.G. (1966) Ear canal pressure generated by circumaural and supra-aural earphones. JASA, 39: 471-479.

-and Thiessen, G.J. (1962) Acoustics of circumaural earphones. JASA, 34: 1233-1246.

Shower, E.G. and Biddulph, R. (1931) Differential pitch sensitivity of the ear. JASA, 3: 275-287.

Siebert, W.M. (1965) Implications of the stochastic behaviour of primary auditory neurons. Kybernetik, 2: 206-215.

Skinner, P.H. and Antinoro, F. (1968) Study of per- and post-stimulatory fatigue in pitch perception. JASA, 44: 1423-1427.

Small, A.M. and Brandt, J.F. (1963) Differential thresholds for frequency. JASA, 35: 787 (A).

-and Campbell, R.A. (1961a) Pitch shifts of periodic stimuli with changes in sound level. JASA, 33: 1022-1027.

Small, A.M. and Campbell, R.A. (1961b) Masking of pulsed tones by bands of noise. JASA, 33: 1570-1576.

-and Daniloff, R.G. (1967) Pitch of noise bands. JASA, 41: 506-512.

-and McClellan, M.E. (1963) Pitch associated with time delay between two pulse trains. JASA, 35: 1246-1255.

-and Yelen, R.D. (1962) Fatigue as an indicator of pitch channles. JASA, 34: 1987 (A).

Smith, J.E.K. (1961) Stimulus programming in psychophysics. Psychometrika, 26: 27-33.

Stevens, K.N. (1952) Frequency discrimination of damped waves. JASA, 24: 76-79.

Stevens, S.S. (1935) The relation of pitch to intensity. JASA, 6: 150-154.

-and Davis, H. (1938) Hearing; its Psychology and Physiology, (2nd edition), Wiley, New York.

Swets, J.A., Green, D.M., and Tanner, W.P. (1962) On the width of critical bands. JASA, 34: 108-113.

Tanner, W.P. (1960) Theory of signal detectability as an interpretive tool for psychophysical data. JASA, 32: 1140-47.

-, Birdsall, T.G., and Clarke, F.R. (1960) The concept of the ideal observer in psychophysics. Tech. Rep. No. 98, Univ. of Michigan, Dept. of Elec. Eng., Ann Arbor.

Tasaki, I. (1954) Nerve Impulses in individual auditory nerve fibers of Guinea Pig. J. Neurophysiol., 17: 97-122.

Teas, D.C. and Henry, G.B. (1968) Amplitude distributions of cochlear microphonic response to an acoustic stimulus. J. Speech & Hearing Res., 11: 63-76.

Thurlow, W.R. (1958) Some theoretical implications of the pitch of double pulse trains. Am. J. Psychol., 71: 448-450.

-and Small, A.M. (1955) Pitch perception for certain periodic auditory stimuli. JASA, 27: 132-137.

Tonndorf, J. (1962) Time/frequency analysis along the partition of cochlear models: a modified place concept. JASA, 34: 1337-1350.

Tresselt, M.E. (1948) Time errors in successive comparison of tonal pitch. Am. J. Psychol., 61: 335-342.

Ward, W.D. (1960) Latent and residual effects in temporary threshold shift. JASA, 32: 135-137.

-(1963) Diplacusis and auditory theory. JASA, 35: 1746-1747.

Webster, J.C., Miller, P.H., Thompson, P.O., and Davenport, E.W. (1952) Masking and pitch shifts of pure tones near abrupt changes in a thermal noise spectrum. JASA, 24: 147-152.

-and Schubert, E.D. (1954) Pitch shifts accompanying certain auditory threshold shifts. JASA, 26: 754-758.

Webster, R.B. (1969) The contralateral difference limen for pitch. J. Sound & Vibration, 9: 97-100.

Wegel, R.L. and Lane, C.E. (1924) The auditory masking of one pure tone by another. Phys. Rev., 23: 266-285.

Weiss, T.F. (1964) A model for firing patterns of auditory nerve fibers. Tech Rep. No. 418, M.I.T. Research Lab of Electronics.

Wetherill, G.B. (1963) Sequential estimation of quantal response curves. J. Roy. Statist. Soc. (Series B, Methodological), 25: 1-48.

-and Levitt, H. (1965) Sequential estimation of points on a psychometric function. Brit. J. Math. & Stat. Psychol., 18: 1-10.

Wever, E.G. (1949) Theory of Hearing. Wiley, New York.

Whitfield, I.C. (1957) The physiology of hearing. (In: Progress in Biophysics and Biophysical Chemistry, (Eds., Butler, J.A.V. and Katz, B.), vol. 8; Pergamon Press, London, p. 1-47.

Whittle, L.S. and Robinson, D.W. (1961) British normal threshold of hearing. Nature, 189: 617-618.

Wiederhold, M.L. and Peake, W.T. (1966) Efferent inhibition of auditory nerve responses: dependence on acoustic stimulus parameters. JASA, 40: 1427-1430.

Woodward, P.M. (1953) Probability and Information Theory with Applications to Radar. Pergamon Press, London.

-and Davies, I.L. (1950) A theory of radar information. Phil. Mag., 41: 1001-1017.

Zwicker, E. von (1952) Die grenzen der hörbarkeit der amplitudenmodulation und der frequenzmodulation eines tones. Acustica, 2: 125-133.

Zwicker, E. von, Flottorp, G., and Stevens, S.S. (1957) Critical bandwidth in loudness summation. JASA, 29: 548-557.

Zwislocki, J. (1957) Some impedance measurements on normal and pathological ears. JASA, 29: 1312-1317.

APPENDIX I. THE VARIANCE PROPERTIES OF PROBIT ESTIMATES.

The method of Probit Analysis (PA) requires that each observation in an experiment is made at a particular value of d_f, and leads to a response which is either U or D. The subsequent curve fitting is based on the proportions of U-responses in the groups of presentations made at the d_f values chosen. (Here, it has been assumed that the same number of presentations is made in each group.) Ideally, the d_f values should be selected with fairly accurate knowledge of the true parameters of the response curve being estimated; the choice may depend on whether the slope (B) or the midpoint (M) is of greater interest. This is because each proportion is weighted in a way which depends on its location on the response curve. This location will, in turn, depend on both B and M of the curve, relative to the d_f value of the group. In the estimation of M, the greatest weight is given to those groups which are placed close to the true midpoint. In the estimation of B, the greatest weight is given to groups placed about 1.5/B on either side of M.

In practice, not only is some compromise necessary to allow reasonably efficient estimation of both B and M, but allowance must also be made for uncertainties in the prior knowledge of these values. Further, in the case of a subjective experiment, it is desirable that the d_f values should be placed symmetrically about M, so that the total numbers of U and D responses given by a subject are roughly equal. Three designs were chosen for discussion here; all involve an odd number of groups (3, 5, and 7) at equally spaced values of d_f, with the middle group having the value: $d_f = 0$ (i.e., no frequency shift). Thus, assuming unbiased judgements, this middle group should be close to the midpoint. The properties of these designs were assessed by evaluating the formulae for the expected large sample variances of $\hat{B}$ and $\hat{M}$ given by Finney (1952). These are:

$$V(\hat{B}) = 1/\Sigma_i n_i w_i (d_{fi} - \bar{d}_f)^2$$

$$V(\hat{M}) = 1/B^2 \left[1/\Sigma_i n_i w_i + (M - \bar{d}_f)^2 \; V(\hat{B}) \right]$$

where: n_i = the number of presentations in the i^{th} group

w_i = the weighting coefficient for the i^{th} group

Σ_i = the sum over all groups

$\bar{d}_f = \Sigma n_i w_i d_{fi} / \Sigma n_i w_i$

It was assumed that the spacing of the d_f values was 1 unit, and the formulae were evaluated for the range of B between 0.5 - 2 (or DL's between 2 and 0.5 units); this being a range expected in practice. Also, to study the effect of bias, evaluations were made both for $M = 0$ and for a displacement of M by 1 DL unit from 0. The results are shown in fig. A-1.

The curves in fig. A-1 have been plotted on the basis of the same total number of observations in each design. (They are normalized on a per-observation basis, and the SD figures apply to 1/3 of an observation per group in the 3-groups design; 1/5 in the 5-groups design, etc.) The solid curves represent the case where there is no bias (labelled by the number of groups involved); the dotted curves illustrate the effect of a bias of 1 DL unit (labelled by the number of groups, plus S, for 'shifted'). Differences between the groups are also represented by graded line thickness.

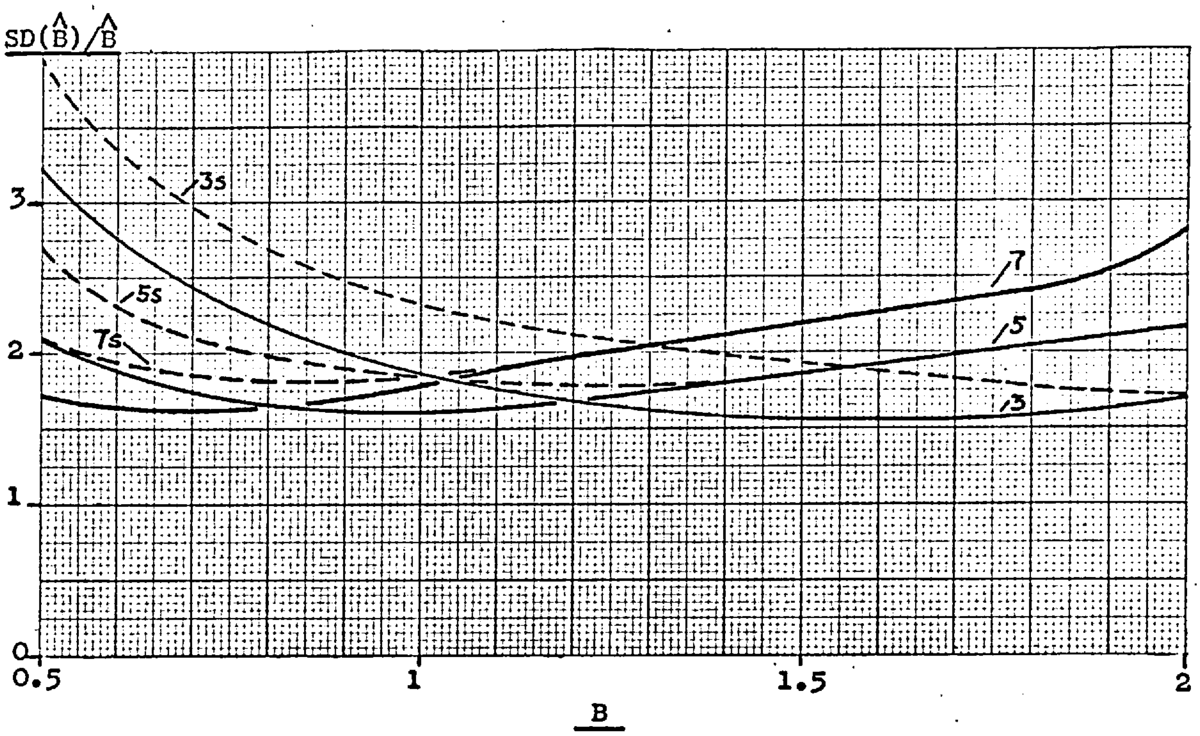
In the case of M (fig. A-1b), the presentation is straightforward, with the expected SD($\hat{M}$) plotted vs. B. The slope data (fig. A-1a) have been presented differently: as $SD(\hat{B})/\hat{B}$ plotted vs. B (i.e., as the standard error of $\hat{B}$). This was for two reasons: 1) this is a more meaningful measure of precision than SD alone, and 2) this form of presentation illustrates, in one curve, the behaviour of both $\hat{B}$ (the estimates derived directly by PA), and the DL (i.e., $1/\hat{B}$), which was used in Experiment P-1. (Since $DL = 1/\hat{B}$, $SD(DL) \cong SD(\hat{B})/\hat{B}^2$. Thus, $SD(DL)/DL$ = Standard Error(DL) = $SD(\hat{B})/\hat{B}$ = Standard Error($\hat{B}$).)

The curves illustrate that, generally speaking, the larger the number of groups in a design, and the larger its spacing (relative to the DL), the more resistant it is to variations in the true parameters of the response curve. However, in most cases, more precise estimates can be obtained with a small number of groups, provided that the parameters of the curve are well known in advance. There is also a practical point here; for subjective testing (where the total number of responses is fairly small), it is preferable to use a small number of groups because this reduces the chances of obtaining groups with all, or no, U responses, which are difficult to handle with PA.

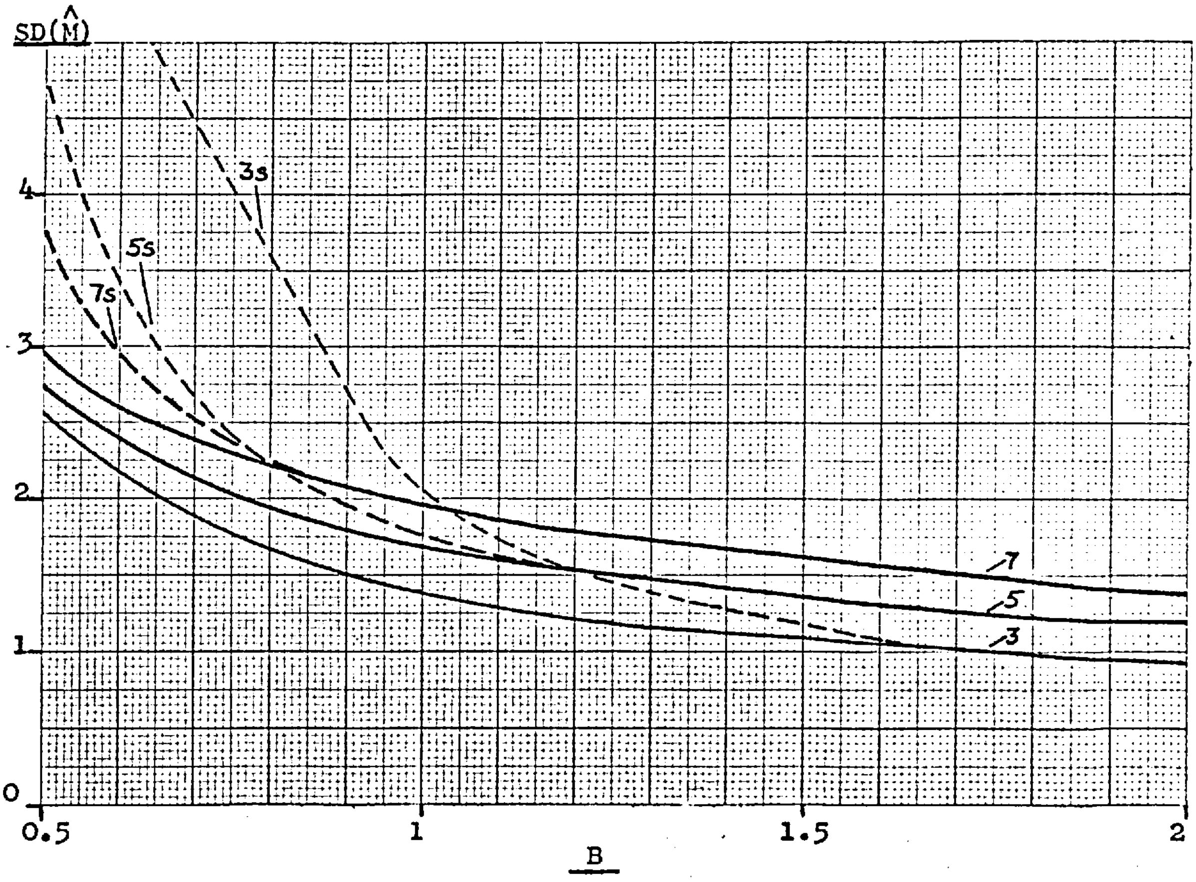
A 5-groups design was finally chosen, with a spacing of approximately 1 DL (1/B). This appeared to be (from fig. A-1a) the best and most stable of the three designs for B estimation, but was somewhat sensitive to the true value of M for M estimation at small values of slope (see fig. A-1b). Some properties of the data obtained with this design are discussed in Chapter V (p. 95, et seq.)

Fig. A-1. Expected Variance of Probit Estimates of Slope and Midpoint for Three Different Experimental Designs, each with a Level Spacing of 1 Unit.

A-1a. Slope Estimates



A-1b. Midpoint Estimates



APPENDIX II. DETAILS OF THE MODEL OF THE MECHANICAL BEHAVIOUR OF THE EAR USED IN EXPERIMENTS VI-A AND VI-B.

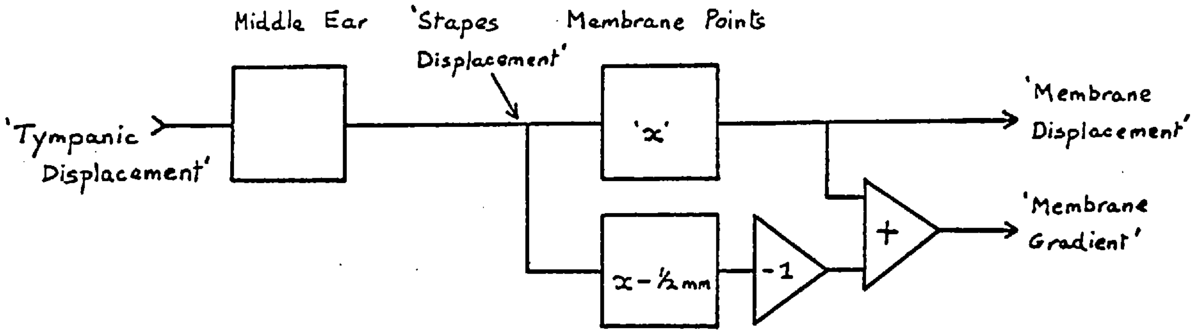
The design of the Model Ear exactly followed a proposal by Flanagan (1962) for lumped constant circuits to simulate the frequency response curves of the middle ear and basilar membrane. As shown in fig. A-2, the middle ear section consists of an RC lowpass- followed by an LCR lowpass-section; while the behaviour of the basilar membrane is simulated by an RC highpass-followed by two LCR lowpass-sections. Changing of frequency was accomplished by switching the components shown in the table at the bottom of fig. A-2.

The performance of these circuits is shown in fig. A-3. Here, the lines are derived from original measurements by Zwislocki (1957) for the middle ear, and by von Bekesy (1947) for the basilar membrane (see Flanagan, 1962). The points represent measurements of the performance of the circuits, and agreement is close except for the phase response of the basilar membrane analogue. This is partly because von Bekesy's measurements include the time delay between the stapes and the membrane point, which would account for the linear phase shift observed. (No time delay was included in the present Model; see footnote on p. 163.)

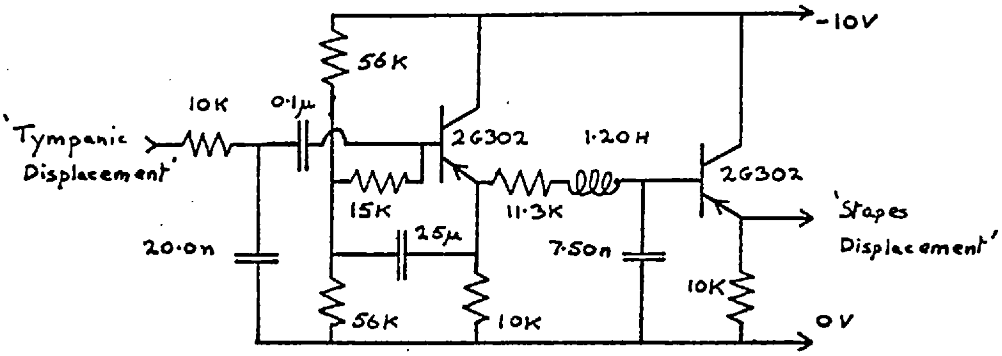
The complete Model (fig. A-2, top) consisted of a middle ear analogue connected to two membrane point analogues, corresponding to 0.5 mm separation. The output of either of the membrane analogues corresponded directly to basilar membrane displacement; taking their difference (as shown in fig. A-2), gave a function corresponding to the gradient.

Fig. A-2. Details of the Model Ear Used in Experiments VI.

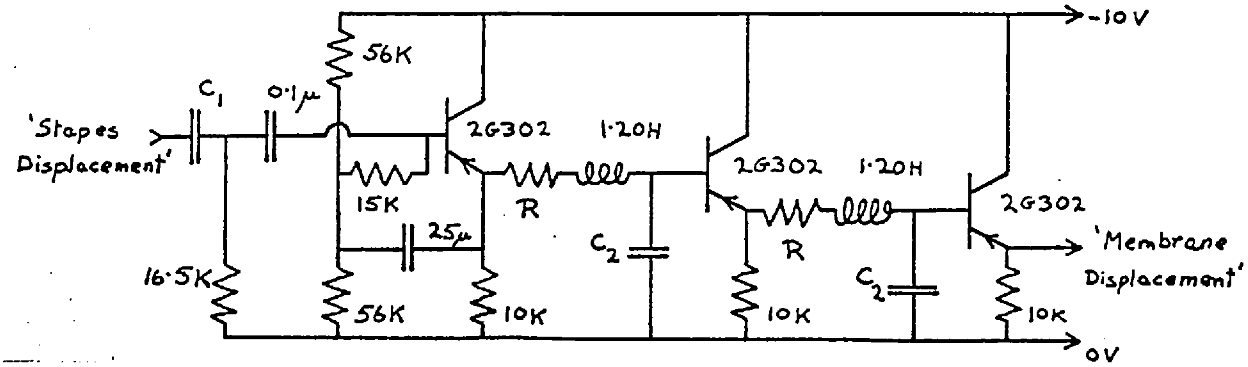
Block Diagram
of Complete
Model



Middle Ear
Analogue



Membrane Point
Analogue



Freq.(KHz)	0.5 ; $-\frac{1}{2}$mm	1.0 ; $-\frac{1}{2}$mm	2.0 ; $-\frac{1}{2}$mm	3.0 ; $-\frac{1}{2}$mm
R (KΩ)	3.78;4.11	7.55;8.22	15.1;16.4	22.6;24.7
C_1 (nF)	19.3;17.7	9.65;8.85	4.82;4.42	3.22;2.95
C_2 (nF)	67.5;56.9	16.9;14.2	4.23;3.55	1.88;1.58

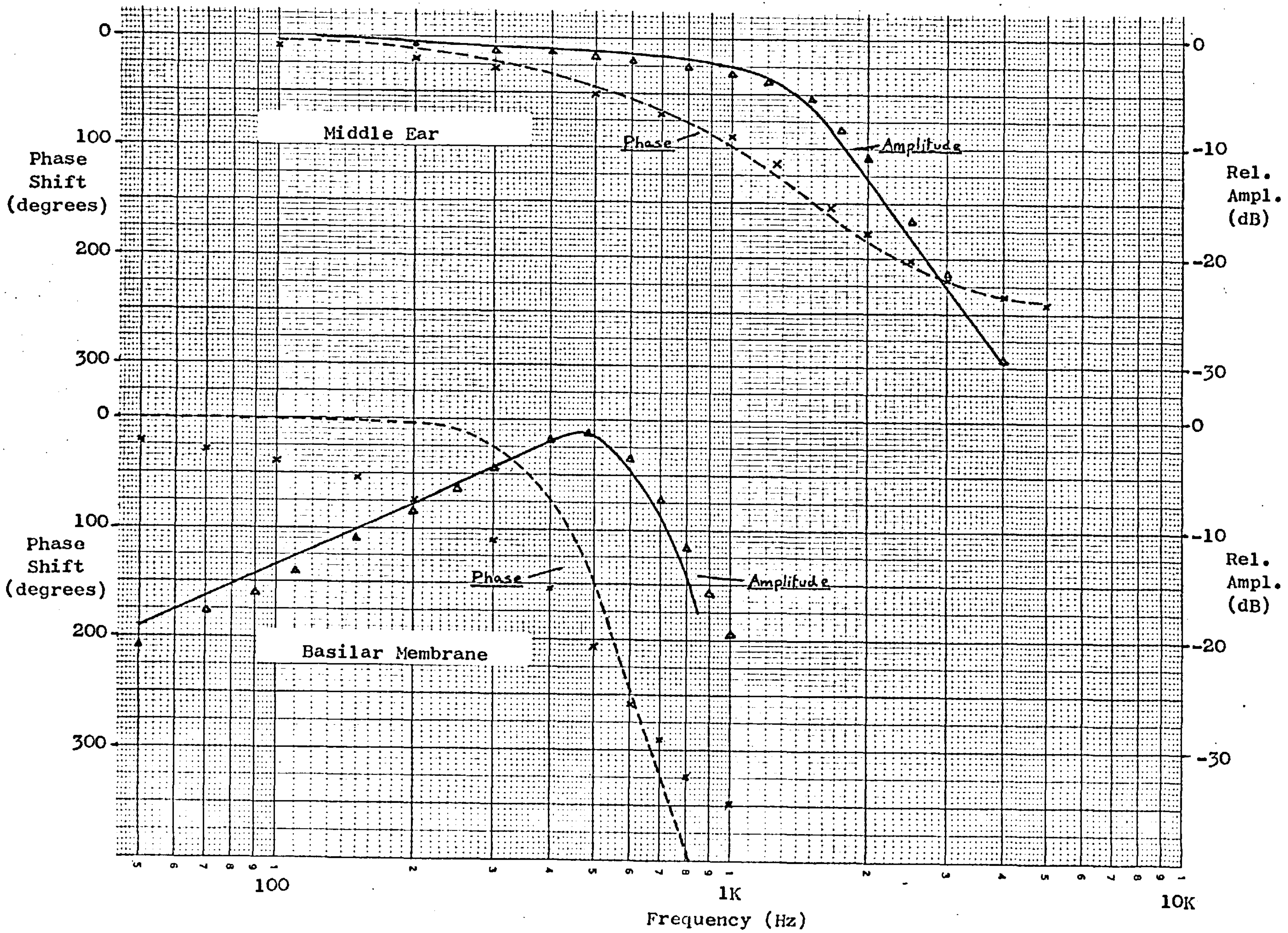


Fig. A-3. Response Characteristics of the Middle Ear and Basilar Membrane Analogues.