Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

CENTERING IN PARALLEL CHANNEL SYSTEMS

BY

ALLEN KATZ

A DISSERTATION

PRESENTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF ENGINEERING SCIENCE IN ELECTRICAL ENGINEERING

AT

NEWARK COLLEGE OF ENGINEERING

Newark, New Jersey
1971

ABSTRACT

Several types of signal processing systems in which the signal flows along parallel channels in a fashion similar to the auditory system have been investigated. The effect of excitation with signals containing both single and multiple spectral peaks (formants) was considered. In particular, the effect of nonlinear interaction between channels, referred to as centering, in the presence of noise was studied.

These systems were investigated for their value, both as information processing networks and as models of the auditory system.

The analysis indicates that parallel channel systems, in general, exhibit excellent performance in the presence of noise, and that a parallel channel system, with a limited overall bandwidth, can be made to process large amounts of information per unit time if used in conjunction with an appropriate centering network. Furthermore, these systems permit detailed frequency analysis of signals in the presence of noise without impairing their temporal discrimination capability.

Of the centering processes investigated, maximum likelihood centering provides an optimum estimate of formant

frequency in the presence of noise, while lateral inhibitory centering probably represents the most practical process for implementation.

The performances of various centering processes are compared to the known characteristics of the auditory system, and the most promising of these, lateral inhibitory centering, is employed in a model of the peripheral auditory system.

The response of this model, when simulated on the digital computer, correlates closely with many of the characteristics of the peripheral auditory system. The model, however, does not adequately "explain" the spectral resolving ability displayed by the ear. An extension of the model was suggested which should not be subject to this limitation.

APPROVAL OF DISSERTATION

CENTERING IN PARALLEL CHANNEL SYSTEMS

BY

ALLEN KATZ

FOR

DEPARTMENT OF ELECTRICAL ENGINEERING

NEWARK COLLEGE OF ENGINEERING

BY

FACULTY COMMITTEE

APPROVED: _____ Chairman

_____

_____

_____

_____

NEWARK, NEW JERSEY

JUNE, 1971

ACKNOWLEDGMENTS

The author is indebted to Dr. M. Zambuto, Professor of Electrical Engineering, for his advice and constructive commentaries during the course of this investigation. He also wishes to give acknowledgment to the faculty of the Department of Electrical Engineering, Newark College of Engineering, for many helpful suggestions. Particular thanks are extended to Dr. M. Kurland, Dr. A. H. Marsh, and Dr. B. Son.

The financial support of Newark College of Engineering is also gratefully acknowledged.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
APPROVAL PAGE	iv
ACKNOWLEDGMENTS	v
NOTATIONS	ix
1. INTRODUCTION	1
2. PARALLEL CHANNEL INPUT CONSIDERATIONS	8
3. BASIC PARALLEL CHANNEL SYSTEM	16
4. PARALLEL CHANNEL SYSTEMS WITH IDEAL CENTERING	30
5. PRACTICAL CENTERING PROCESSES	50
6. CENTERING BY LATERAL INHIBITION	78
7. A REVIEW OF PERTINENT ANATOMICAL AND PHYSIOLOGICAL PROPERTIES OF THE EAR	106
8. A MODEL OF CENTERING IN THE PERIPHERAL AUDITORY SYSTEM	122
9. CONCLUSION	148
10. RECOMMENDATIONS	160
APPENDIXES	
A. Calculation of the Distribution of Estimated Frequency of P.C.S. with Centering in the Presence of Intrinsic Noise	164
B. Probability of Occurrence of Firing Pattern Y_i in the Case of Extrinsic Noise	172

TABLE OF CONTENTS (continued)

APPENDIXES Page

C. Calculation of the Probability Distribution of the Estimated Frequency of a P.C.S. with Centering in the Presence of Extrinsic Noise 175

D. Computer Simulation of the Peripheral Auditory System 185

E. Computer Program Listings 189

REFERENCES 204

VITA 209

LIST OF FIGURES

Figure		Page
1.	Time-frequency domain representation of a single formant signal	9
2.	Possible generators of multiple formant signals	11
3.	Time-frequency domain representation of a multiple formant signal	13
4.	Basic parallel channels system	17
5.	Probability of correct detection as a function of N for a basic P.C.S.	21
6.	Z_T as a function of N and false alarm rate for a basic P.C.S.	22
7.	MTR as a function of N for a fixed bandwidth basic P.C.S.	24
8.	MTR as a function of N for a fixed information rate basic P.C.S.	26
9.	MTR as a function of normalized system bandwidth for a basic P.C.S.	28
10.	P.C.S. with centering and sample bandpass characteristics	31
11.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with maximum likelihood centering in the presence of extrinsic noise	36
12.	σ_f as a function of Z for a 100 channel P.C.S. with maximum likelihood centering in the presence of intrinsic noise	37
13.	$P(CD/S_1)$ as a function of Z and ΔF_S for a 100 channel P.C.S. with maximum likelihood centering in the presence of intrinsic noise	40

LIST OF FIGURES (continued)

Figure		Page
14.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with maximum likelihood centering in the presence of extrinsic noise	44
15.	σ_f as a function of Z for a 100 channel P.C.S. with maximum likelihood centering in the presence of extrinsic noise	45
16.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with maximum likelihood centering in the presence of extrinsic noise	46
17.	Diagrammatic display of ΔF_{ss} required by maximum likelihood centering process	48
18.	Median channel centering process	51
19.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with M.C. centering in the presence of intrinsic noise	56
20.	σ_f as a function of Z for a 100 channel P.C.S. with M.C. centering in the presence of intrinsic noise	57
21.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with M.C. centering in the presence of intrinsic noise	58
22.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with M.C. centering in the presence of extrinsic noise	60
23.	σ_f as a function of Z for a 100 channel P.C.S. with M.C. centering in the presence of extrinsic noise	61
24.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with M.C. centering in presence of extrinsic noise	63
25.	Mean frequency centering processes	64
26.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with M.F. centering in the presence of intrinsic noise	67

LIST OF FIGURES (continued)

Figure		Page
27.	σ_f as a function of Z for a 100 channel P.C.S. with M.F. centering in the presence of intrinsic noise	68
28.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with M.F. centering in the presence of intrinsic noise	69
29.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with M.F. centering in the presence of extrinsic noise	72
30.	σ_f as a function of Z for a 100 channel P.C.S. with M.F. centering in the presence of extrinsic noise	73
31.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with M.F. centering in the presence of extrinsic noise	74
32.	Diagram displaying ΔF_{ss} required by a M.F. centering process	77
33.	Two channel lateral inhibitory processes (a. nonrecurrent, b. recurrent)	80
34.	Nonrecurrent L.H. network and possible response characteristics	82
35.	Recurrent L.H. network and possible response characteristics	84
36.	Response characteristic of L.I.C. L.H. centering process	89
37.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of intrinsic noise	93
38.	σ_f as a function of Z for a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of intrinsic noise	94
39.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of intrinsic noise	96

LIST OF FIGURES (continued)

Figure		Page
40.	$P(f_E/f_s)$ distribution of a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of extrinsic noise	98
41.	σ_f as a function of Z for a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of extrinsic noise	99
42.	$P(CD/S_1)$ as a function of Z and ΔF_s for a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of extrinsic noise	100
43.	$P(f_m/f_s)$ distribution of a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of intrinsic noise	104
44.	σ_m as a function of Z for a 100 channel P.C.S. with (L.I.C.) L.H. centering in the presence of intrinsic noise	105
45.	Schematic and functional block representations of the peripheral auditory systems	107
46.	Transmission characteristics of the middle ear	109
47.	Cochlear portion of the inner ear	110
48.	Relative amplitude of the envelope of vibration along the cochlear partition as a function of exciting frequency	112
49.	(a) Maximum value of displacement of the cochlear partition as a function of frequency. (b) Position of maximum displacement of the cochlear partition as a function of frequency	113
50.	Amplitude of displacement of the cochlear partition as a function of position and frequency	115
51.	Transducer characteristics of individual auditory nerve fibers	119
52.	Tuning curves of primary auditory nerve fibers	121

LIST OF FIGURES (continued)

Figure		Page
53.	A comparison of the fiber firing thresholds expected on the basis of (a) the spectral response of the basilar membrane, (b) experimental observation, (c) the model	124
54.	Logarithmic display of the impulse response of the cochlear partition	129
55.	Idealized inhibition and response areas of a typical auditory nerve fiber	131
56.	Lateral inhibitory centering process used in the model	135
57.	Centering action produced by the model as a function of summer gain	137
58.	Centering action produced by the model as a function of signal level	139
59.	Temporal characteristics of the centering action of the model	141
60.	Comparison of tuning curve produced by the model to that observed by Kiang	142
61.	Centering action produced by the model in response to an approximately 2 msec. tone burst	144
62.	Centering action produced by the model in response to two continuous tones	146
63.	Approximate system bandwidth reduction achievable by a maximum likelihood centering process as a function of overlap factor	152
A-1.	Machine computation of $P(f_E/f_S)$ in intrinsic noise	165
A-2.	Decision processes for intrinsic noise case	167
A-3.	Lateral inhibitory decision process	170

LIST OF FIGURES (continued)

Figure		Page
C-1.	Machine computation of $P(f_E/f_S)$ in extrinsic noise by means of Monte Carlo simulation	176
C-2.	Relationship of extrinsic noise energies of two adjacent channels	181
D-1.	Machine simulation P.C.S. model of peripheral auditory system	186

NOTATIONS

A_c = Amplitude of segment of noise spectrum common to two adjacent channels.

A_n = Amplitude of independent segment of noise spectrum.

A_x = Amplitude function of X in expression for $H_x(j\omega)$.

b = Normalized detector threshold = V_{th}/σ.

BW = Overall bandwidth of P.C.S.

C = Constant relating length of basilar membrane and its range of spectral response.

CF = Characteristic frequency (of primary auditory fibers).

CF_{max} = CF corresponding to most basal point of the basilar membrane.

CF_{min} = CF corresponding to the most apical point of the basilar membrane.

CN_L = Complex noise intensity of independent segments of extrinsic noise spectrum ($\Delta f/M$ in bandwidth).

CS_n = Complex signal intensity in channel n.

CV_n = Sample signal plus noise intensity in channel n.

D = Factor establishing the effect of f_E in the expression for the $P(f_E/f_S)$ of a M.C. centering process.

DI = Length of basilar membrane.

e = Indexing variable of summation.

f = Frequency of possible spectral peak.

f_{CF} = Normalized characteristic frequency--CF/f_S.

f_E = Estimated center frequency of an exciting formant.

f_n = Center frequency of the channels of which a P.C.S. is composed; also used to designate (center) frequencies of formants.

f_m = Frequency location of the minimum level of threshold inhibition in an L.H. centering process.

f_s = Frequency at which a (given) exciting formant is centered.

FSK = Frequency shift keying.

FT = Firing threshold.

$G(j\omega)$ = Transfer characteristic of the outer and middle ear.

$g(n,f)$ = Amplitude function indicating the amount of energy which enters the nth channel due to a formant of frequency f.

$g_n(\ )$ = Function relating the output of the nth channel of a L.H. network to its inputs.

$h(x,t)$ = Impulse response of basilar membrane.

$H_n(j\omega)$ = Transfer characteristic of the basilar membrane at a point corresponding to the nth channel of the model.

$H_x(j\omega)$ = Transfer characteristic of the basilar membrane at distance X from the basal end.

I = Maximum information processing rate of a P.C.S.

j_{nm} = Coefficient establishing the inhibitory effect of the input to the nth channel on the mth channel of an L.H. centering process.

k = Indexing variable of summation.

K = Maximum number of simultaneous formants in K formant signal.

k_{nm} = Coefficient establishing the excitatory effect of the input to nth channel on the mth channel of an L.H. centering process.

KI = Constant relating CF to X of the basilar membrane.

K_n = Constant establishing the magnitude of the threshold inhibition in the nth channel of the model.

L = Indexing variable of summation.

L.H. = Lateral Inhibitory.

L.I.C. = Class of L.H. centering process which tends to produce the least inhibition at the center of excitation (least inhibition at center).

ℓ_{nm} = Coefficient establishing the effect of the output of the nth channel on the mth channel of an L.H. centering process.

ℓ_L = Gain of summers (proceeding from high to low) in the model.

ℓ_r = Gain of summers (proceeding from low to high) in the model.

m = Indexing variable of summation.

M = Channel overlap factor ($\Delta f/\Delta f_x$).

M'ary = Multiple channel.

M.C. = Median channel.

M.F. = Mean frequency.

M.I.C. = Class of L.H. centering process which tends to produce maximum inhibition at the center of excitation (maximum inhibition at center).

MTR = Measure of threshold reduction provided by an N channel P.C.S. in comparison with a single channel system.

n = Channel number.

N = Total number of channels of P.C.S.

n' = Indexing variable used in place of n.

N' = Constant related to the minimum number of channels the first moment of an M.F. process must be calculated over in order to avoid the detection of spurious formants.

NC = Number of channels which share common spectra (2M - 1).

p = Probability of a channel being activated when a (single) formant is present in the channel.

P(CD) = Probability correctly interpreting (the presence and spectral composition of) a signal.

$P(CD/S_K)$ = Probability of correctly interpreting a signal given the presence of a signal.

P.C.S. = Parallel channels system(s).

$P(f_E/f_s)$ = Probability distribution of estimated frequency for an exciting formant of frequency f_s.

$P(f_m/f_s)$ = Probability distribution of f_m for an exciting formant of frequency f_s.

P_{IS} = Probability factor accounting for the firing probabilities of those channels which receive signal energy from f_s.

p_k = Probability of a channel being activated when one formant of k simultaneous formants is present in the channel.

PQ = Probability that all channels of a P.C.S. which do not receive signal energy from a formant (at frequency f_s) do not fire.

q = Probability of a channel not being activated when a formant is not present in the channel.

Q(,) = Q function.

$q_{n/n-1}$ = Probability of a channel (n) not firing given the presence of noise only and given that the channel directly adjacent (n - 1) does not fire.

$q_{n,n-1}$ = Probability of channels n and n - 1 not firing given the presence of noise only.

r = Dummy variable of integration.

RA = False alarm rate.

$R_{nn}( )$ = Autocorrelation function of noise in a given channel.

R_L = Integral of the probability density describing the energy contained within independent segments of extrinsic noise energy.

S = Normalized amplitude of common noise spectrum (A_c/σ_c).

SA_n = The amplitude of the input to the nth channel of the model normalized with respect to V_{th}.

$S_K(j\omega)$ = Spectral distribution of $S_K(t)$ normalized with respect to V_{th}.

$S_K(t)$ = Multiple (or K) formant signal.

SNDR = Signal-to-noise density ratio.

SNR = Signal-to-noise ratio.

STR = Signal-to-threshold (V_{th}) ratio.

t = Time.

T = Threshold detector.

T_n = Inhibitory input to the nth threshold detector of an L.H. centering process.

V_{dn} = Output of the nth threshold detector of a P.C.S.

V_f = Output of an M.F. process (V_f is proportional to the centroid of the P.C.S. inputs).

V_{in} = Excitatory input to the nth threshold detector of a P.C.S.

V_{Ln} = The output of the nth summer of the model proceeding from high to low.

V_n = Dummy variable of integration.

V_{on} = The output of the nth channel of a P.C.S. system.

V_{rn} = The output of the nth summer of the model proceeding from low to high.

V_{th} = Threshold level of detectors (noninhibited).

V_{Tn} = Inhibited threshold of the nth channel of an L.H. centering process.

W(a,V) = Integrand of the Q function.

X = Distance along basilar membrane (measured from basal end).

x = Dummy variable of integration.

X_L = Integral of multiple signal envelope function.

XN = Normally distributed random variable of variance 1/M.

Y_i = Vector representing ith possible output firing pattern of a P.C.S.

Y_i^* = Vector representing output firing pattern of a P.C.S. with lateral inhibition, after inhibition is established.

$Y_{i,n}$ = Element of Y_i corresponding to the state of the nth channel.

$Y_{i,n}^*$ = Element of Y_i^* corresponding to the state of the nth channel, after inhibition is established.

YN = Normally distributed random variable of variance 1/M.

Z = Channel SNR.

$Z_{f,n}$ = SNR in the nth channel due to a formant of frequency f.

Z_T = Channel SNR corresponding to a $P(CD/S_K) = .5$.

Z_{TD} = SNDR corresponding to a $P(CD/S_K) = .5$.

Δf = Channel bandwidth (also R.M.S. frequency discrimination).

Δf_c = Limiting value of Δf set by establishing Δt.

ΔF_s = Minimum frequency separation between individual formants of $S_K(t)$.

ΔF_{ss} = The minimum frequency spacing between simultaneous formants of $S_K(t)$.

Δf_x = Center-to-center frequency spacing between channels.

Δ_L = One of the factors accounting for edge effects in expression for $P(f_E/f_s)$ of an M.C. process.

Δ_o = One of the factors accounting for edge effects in expression for $P(f_E/f_S)$ of a M.C. process.

Δt = Formant time duration (also R.M.S. temporal discrimination).

Δ_u = One of the factors accounting for edge effects in the expression for $P(f_E/f_S)$ of a M.C. process.

$\lambda(K,t)$ = Amplitude weighting function of $S_K(t)$.

η = Single sided extrinsic noise density.

σ^2 = Noise power per channel.

σ_1^2 = Independent noise power received by one of two overlapping channels.

σ_2^2 = Independent noise power received by the other of the two overlapping channels.

σ_c^2 = Noise power common to two adjacent channels.

σ_m = Standard deviation of f_m (with respect to f_S) given the detection of f_S.

σ_f = The standard deviation of f_E (with respect to f_S) given the detection of f_S.

τ_I = Threshold inhibition time constant.

ω = Exciting frequency in radians.

$[Y_i]$ = Decimal number representing the vector Y_i.

1. INTRODUCTION

Long before the birth of electronics, attempts were being made to "explain" the physical perception of sound on the basis of the anatomic, physiological, and psychophysical study of the ear. One of the first steps toward modern concepts occurred during the middle of the nineteenth century with the statement of Ohm's Law of sound [37]. This theory states that a sound wave is perceived exclusively in terms of its individual sinusoidal components. Shortly thereafter in 1865 Helmholtz conceived of the ear functioning as a "parallel system" of resonators, each responding to a different segment of the auditory spectrum [26, 62]. When a complex tone is applied to the ear, each resonator would thus respond to energy in its particular bandpass.

Theories derived from Helmholtz's original concept are known as "place theories" because of their association of the excitation of a particular resonator (or place) with a particular frequency. Most place theories, however, suffer from a common defect, their inability to satisfy both the ear's exceptional temporal resolution and its fine frequency discrimination [60]. Gabor [21] described this conflict in his classical paper on "acoustic uncertainty" in which he showed that the product of the root mean square

time resolution (Δt) and the root mean square frequency discrimination (Δf) of a resonator is subject to a quantum condition:

$$\Delta f \, \Delta t \geq 1 \, . \tag{1}$$

The high Q resonators that must be adopted to explain the ear's frequency discrimination are totally incompatible with its excellent temporal response [39].

This restriction does not apply to all physical systems. A second set of theories known as "periodicity theories" have served as an alternate to the place theories [50, 56, 62]. These theories assume that all incoming sound signals are transmitted to the higher centers of the brain without first being analyzed. Recognition of individual frequency components is then accomplished by analysis of the periodicity of the signal waveforms.

The operation of periodicity type systems can be compared to the observation of a complex signal on a wideband oscilloscope. Both the "instantaneous" frequency and time duration of a sinusoidally modulated pulse, for instance, may be determined from the scope's trace. The analyzing ability of such systems, however, deteriorates severely when even a small amount of noise is added to the signal [8, 53]. In contrast, one of the outstanding qualities of biological signal perception systems is their ability to function in the presence of noise [27, 40, 47].

Physiological and anatomical data support a "parallel

form" of sound processing on the part of the ear [2]. In particular, electro-physiological studies of the firing patterns of auditory nerve fibers in the peripheral portion of the auditory system indicate the existence of a large number of channels, each exhibiting sharp spectrally selective characteristics [32, 33, 36].

Recent studies indicate that "lateral inhibition," a mechanism by which each channel in a group of parallel channels can influence the response of channels laterally adjacent to it, may be involved in the ear's fine frequency discrimination [32, 51]. If such a type of processing can function well in the presence of noise, it would have value as a model of the hearing system, not to mention possible applications in practical information transmission and pattern identification systems [7, 16]. Obviously, more knowledge is needed of the basic properties of such parallel channel signal processing systems.

Until recently communications theory researchers were not especially interested in parallel channel transmission systems. One exception is the area of multipath fading and diversity transmission [3, 54]. The transmission of information by frequency shift keying (FSK) may also be considered a form of parallel channel process. Until 1955 "two channel" (binary) FSK was of primary concern. Since then the advantages of M'ary (multiple channel) FSK have been made apparent [5, 30].

M'ary FSK systems and the simple parallel channel

systems which have been used as models of the ear in the past possess some similarities. Both may be considered to be processing signals through a set of multiple channels. And both systems can be represented by a bank of contiguous (ideally) orthogonal bandpass filters followed by a parallel set of detectors [8, 54]. The conventional M'ary FSK input signal, however, can have only "one" spectral peak existing at any instant of time. On the other hand, a "general" signal processing system such as the auditory system must be able to process a varying number of simultaneous spectral peaks including the important condition of zero spectral peaks or the total absence of a signal (rarely considered in the analysis of M'ary FSK systems) [59].

The first part of this thesis is concerned with the signal processing ability of parallel channel systems (P.C.S.). The ability of such systems to process signals which possess both time and frequency information is of primary interest. Consequently, for the purposes of this analysis, it is expedient to consider some especially convenient types of input signals. In the following section the concept of a "single formant" signal (a signal for which only one spectral peak can exist at any instant of time) is developed from the standard M'ary FSK signal. This type of signal is then generalized to a "multiple formant," or "K formant," signal (a signal for which a maximum of K spectral peaks can exist at any instant of

time), which has characteristics comparable to those of speech [10, 14, 16].

In order to establish a point of reference to which the performance of more interesting P.C.S. may be compared, the performance of a basic P.C.S. (consisting of a number of contiguous nonoverlapping channels of equal bandwidth) in response to multiple formant signals and noise is investigated in section 3. In particular, the relationship between the input signal-to-noise density ratio (or channel signal-to-noise ratio) and the probability of correctly detecting a signal element as a function of such system parameters as maximum information rate, overall system bandwidth, and system complexity are presented. Consideration is also given to the different effects produced by noise which originated inside the system and by noise originated outside the system.

If the bandwidths of the individual channels of which a P.C.S. is composed are wide in comparison to their center-to-center frequency spacing, these bandpasses will overlap, and the presence of a given spectral peak will excite several channels. A centering process is a mechanism by which the specific channel or channels in which a spectral peak or peaks are centered is estimated on the basis of the positions of the excited channels of the system. The frequency discrimination of a system employing such a process (and hence, the $\Delta f\ \Delta t$ product of equation (1) is not limited by the individual channel bandwidths.

In section 4, the concept of an ideal centering process is introduced. The estimate of formant center frequency made by this process is based on a multidimensional application of Bayes's decision rule [9, 54]. The performance of a P.C.S. with this ideal centering process in the presence of noise is analyzed and compared to the performance of the basic P.C.S. The two distinct abilities of the system to discriminate a change in frequency of a spectral peak and to resolve a multiple formant signal into its individual spectral components are considered.

Unfortunately, the ideal centering process considered in section 4 is not easily implemented in practice. In the next section, the performance of two more practical centering processes is considered and compared to the more complex ideal centering process. The first of these, a median channel (M.C.) centering process, can be implemented entirely in terms of logic gates. The second, a mean frequency (M.F.) centering process, requires somewhat more complicated hardware, but is still considerably simpler than the ideal process.

In section 6, lateral inhibition is investigated as a form of centering process. Lateral inhibition is of interest not only because it has often been observed in biological systems [19], but also because of its structural simplicity based on the repeated use of relatively simple elements. Lateral inhibitory (L.H.) processes appear very attractive for the practical implementation of centering

circuitry. A general mathematical model of lateral inhibition is developed and the "centering-like" characteristics of a number of specific L.H. processes are discussed. The performance in the presence of noise of the more promising of these L.H. processes are analyzed and compared to other forms of centering. The action of a L.H. process as an adaptive centering process is also considered.

The second part of this thesis is concerned with providing an understanding of the signal processing of the peripheral auditory system. Toward this end, a model of the peripheral auditory system is developed in terms of the P.C.S. analyzed in earlier sections. In order to satisfy the ear's excellent temporal resolution and its fine frequency discrimination, a centering process is employed. The particular centering process considered is implemented by means of lateral inhibition.

For the purposes of evaluating the models' performance, a 4000 channel version of the model was simulated on the RCA Spectra 70 digital computer. In section 8 the results of this simulation are presented and compared to the known physiological and psycho-acoustic characteristics of the ear.

2. PARALLEL CHANNEL INPUT CONSIDERATIONS

In order to evaluate the performance of any system, it is necessary to characterize the input to which the system is subjected. A signal such as that of M'ary FSK can be synthesized from a sequency of truncated sinusoids of frequency (f_n), and represented in the time and frequency domains as shown in Figure 1. If the individual sinusoidal trains are all of sufficient duration (Δt), then all their energy can be assumed to be concentrated within a bandpass (Δf) no greater than the bandwidths of the individual channels of which the parallel channel systems to be considered are composed. This type of excitation will be referred to as a "single formant" signal source since only one spectral peak (formant) can exist at any instant of time. Only a small percentage of the sound stimuli which a system such as the ear can encounter falls within the above classification. Thus, for the purposes of our computation, a more general form of input signal, $S_K(t)$, has been considered.

$S_K(t)$ is assumed to be composed of a combination of truncated sinusoids of frequency (f_n) and duration (Δt) such that

$$S_K(t) = \sum_{n=1}^{N} \lambda_n (K,t) \sin 2\pi f_n t \ , \qquad (2)$$

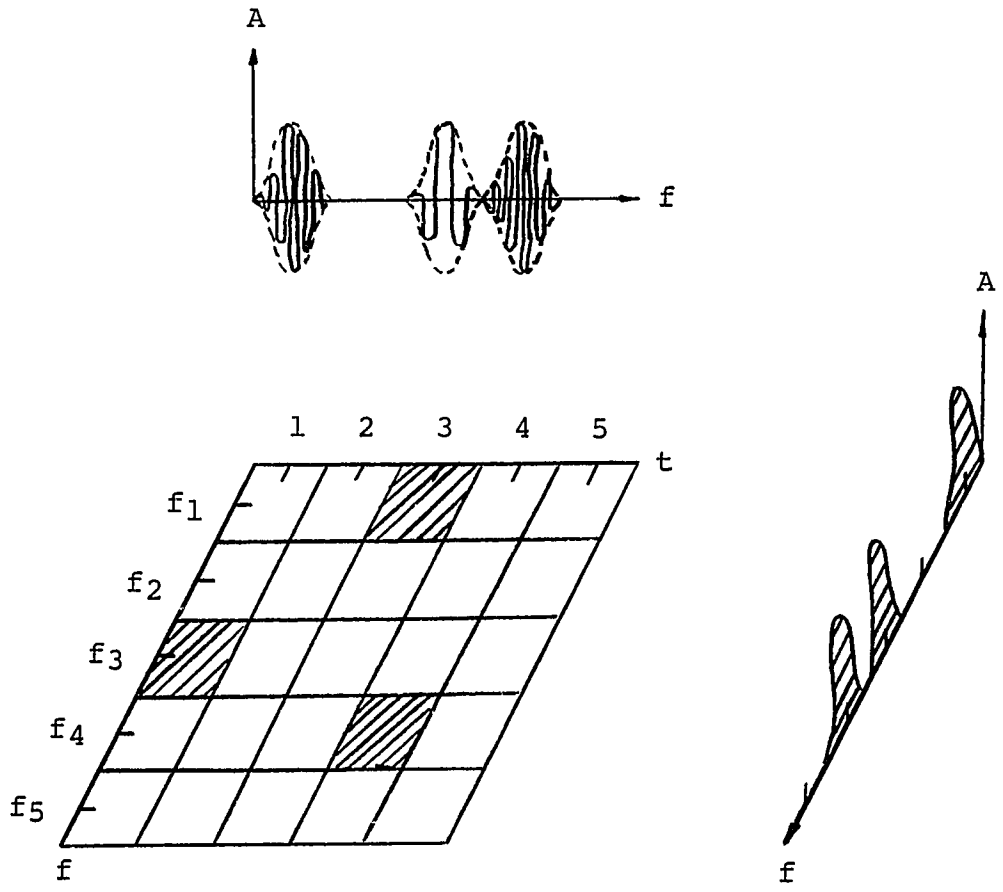


Figure 1. Time-frequency domain representation of a single formant signal indicating the quantum characteristics of this form of excitation.

where N is the number of possible frequencies f_n. λ_n (K,t) is an amplitude function establishing the following signal properties: (a) no more than K sinusoidal trains (spectral peaks) are simultaneously present at any instant of time; (b) all the energy of each individual sinusoidal train is concentrated within a bandwidth Δf, with

$$\Delta f \approx 1/\Delta t \ ; \tag{3}$$

(c) the total signal power remains constant with time, whatever the number of spectral peaks present, i.e., the given power is distributed uniformly among the spectral peaks present.

The last restriction (c) on signal power makes $S_K(t)$ analogous to the type of signal which would be provided by a number of separate oscillators of frequency (f_n) driving a constant power amplifier--as illustrated schematically in Figure 2a. Irrespective of the number of oscillators in operation (1, 2, or n), the signal power remains constant. Of course, this power restriction is not the only one that might be considered. Signal power, for instance, could have been assumed proportional to the number of spectral peaks present at any instant of time. Such a signal would be provided by the circuit of Figure 2b. The significance of alternate power restrictions shall also be considered.

This type of excitation shall be referred to as a "K formant" source since a maximum of "K" simultaneous

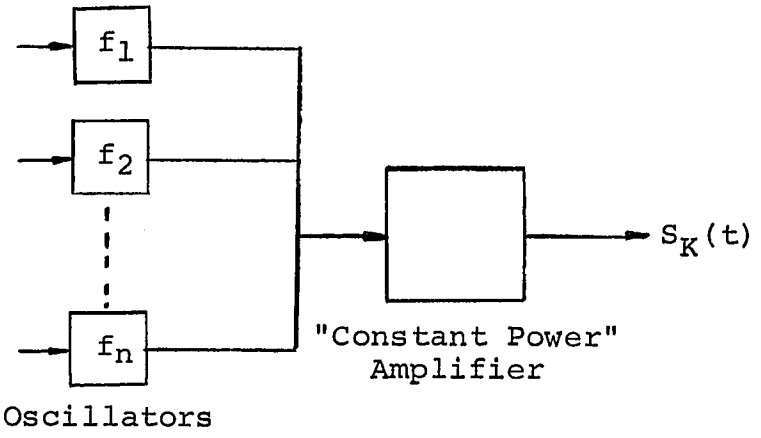


Figure 2a. Possible generator of a constant power multiple formant signal.

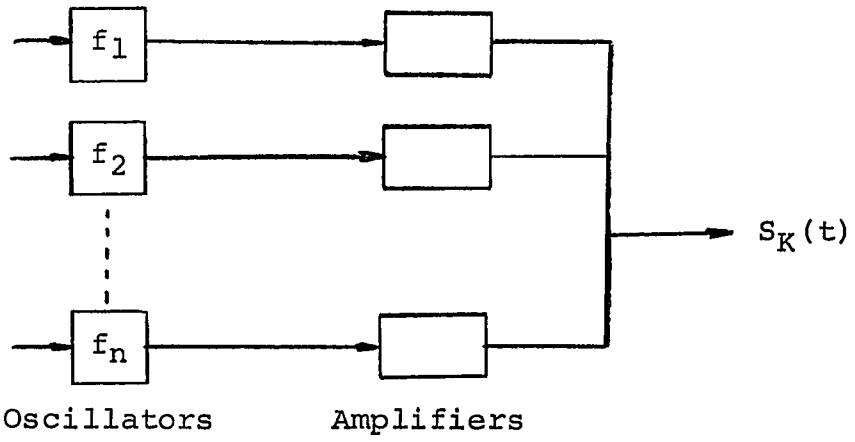


Figure 2b. Possible generator of a multiple formant signal in which signal power is proportional to the number of formants present.

spectral peaks can exist at any instant of time. It has an advantage over time or frequency domain signal representations in that it stresses the quantum characteristics of the exciting signal. A sample three-formant signal is illustrated in Figure 3. The simpler single-formant signal of M'ary FSK is a special case of this signal with K = 1.

The probability of a signal being present during any (quantum) instant of time is assumed to be unknown. This condition is essentially the one under which a "general" signal receiver such as the ear must normally function. If and when a signal, $S_K(t)$, is present, all possible formant combinations are assumed to have equal probabilities of occurrence. As a consequence, the number of possible spectral peak distributions at any instant of time depend on the number of ways K elements or less can be chosen out of a total of N.

The optimum possible performance of any physical system is limited by the presence of noise [8, 53]. In all physical systems, including biological systems (as the auditory system), noise is introduced at several different stages. Of particular interest is the difference between the effect of noise which enters a system together with the signal (extrinsic noise) and that which is generated at some point along the signal path within the system (intrinsic noise). In the case of the ear it is probably intrinsic noise which establishes the minimum detectable signal level or absolute threshold of hearing [25]. The ear's

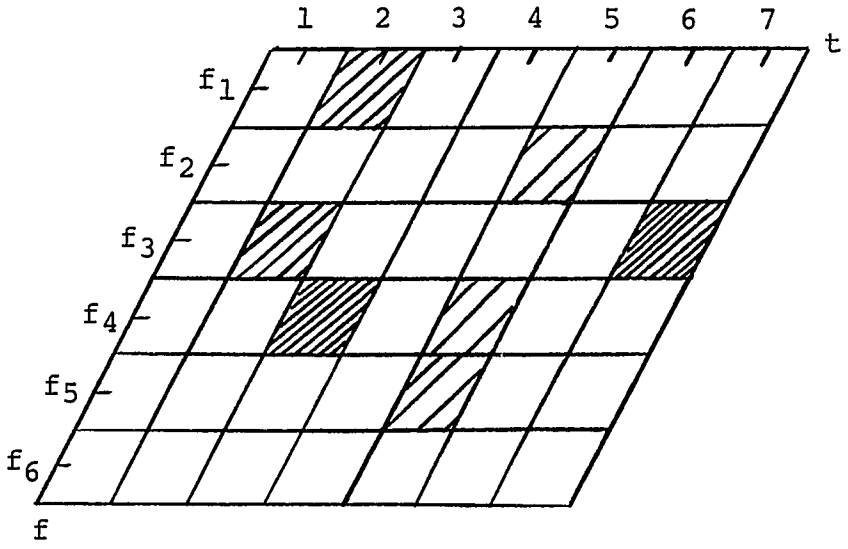


Figure 3. Time-frequency domain representation of a (K = 3) multiple formant signal. Individual formant energy is symbolized by cross-hatch density.

minimum detectable signal level in the presence of external (white Gaussian) noise is measured by means of masking tests and is known as the masked threshold of hearing [27]. All noise sources considered in this thesis are assumed to be white and normally distributed unless specifically noted as otherwise.

The performance of simple single channel systems is usually evaluated in terms of the probability of correctly detecting the presence (or absence) of a signal within the channel at any instant of time, P(CD) [22]. In the N channel case, because of the increased signal complexity, P(CD) is assumed to designate the probability of correctly detecting the presence and spectral composition of a signal at any instant.

Since P(CD) depends on the a priori probability of a signal being present, it is not possible to evaluate P(CD) directly for the assumed signal source. System performance, however, can be evaluated in terms of the probability of correctly detecting a signal given the presence of a signal, $P(CD/S_K)$. Such a performance measure is commonly used in the evaluation of radar systems [58].

In this thesis a constant false alarm rate (i.e., a fixed probability of erroneously interpreting that a signal is present, when in reality it is absent) is assumed. It is interesting to note that a number of physiological processes can be interpreted as displaying a fixed false alarm condition. Neural adaptation [63] is one important

example. For instance, the information provided by the presence of an invariant signal can be considered a function of its duration. If the given signal remains on forever, it provides no additional information since the system already knows of its existence. The signal in effect becomes a form of "noise" since it masks the detection of other changing phenomena which do convey information. The gradual reduction of neural firing of the ear in response to a continuous tone [32] can thus be thought of as an attempt by the hearing system to readjust its firing rate (in response to "noise") back to a fixed level.

The following analysis of parallel channel systems is based on the assumption of a fixed false alarm rate. The relationship between the channel signal-to-noise ratio (SNR) or signal-to-noise density ratio (SNDR) and $P(CD/S_K)$ is used as a measure of system performance. The lower the SNR (or SNDR) necessary to achieve a particular $P(CD/S_K)$, the better the system shall be assumed to be performing.

3. BASIC PARALLEL CHANNEL SYSTEM

As a point of reference, a basic parallel channel system without lateral interaction between channels or "centering" will first be analyzed. The performance of this system will then be compared to the more complex parallel channel systems (P.C.S.) utilizing centering.

A basic P.C.S. consisting of N contiguous nonoverlaping bandpass filters of equal bandwidths Δf is shown in Figure 4.[1] Each of the outputs of these filters is connected to a threshold detector. The assumption of some form of decision process to differentiate between the presence and absence of a signal is necessary. Threshold detector elements were chosen for this purpose because of their analytic simplicity and because of the similarity of their "all or nothing" characteristics with the neural signal patterns observed in the peripheral auditory system [32]. Whenever the input to one of these detectors exceeds some fixed level (V_{th}), the detector indicates the presence of a signal (spectral peak).

For the above system, in response to a single formant signal (K = 1):

---

[1] Δf for the purposes of this dissertation (when otherwise not obvious) shall be defined as the band between the 10 db points.

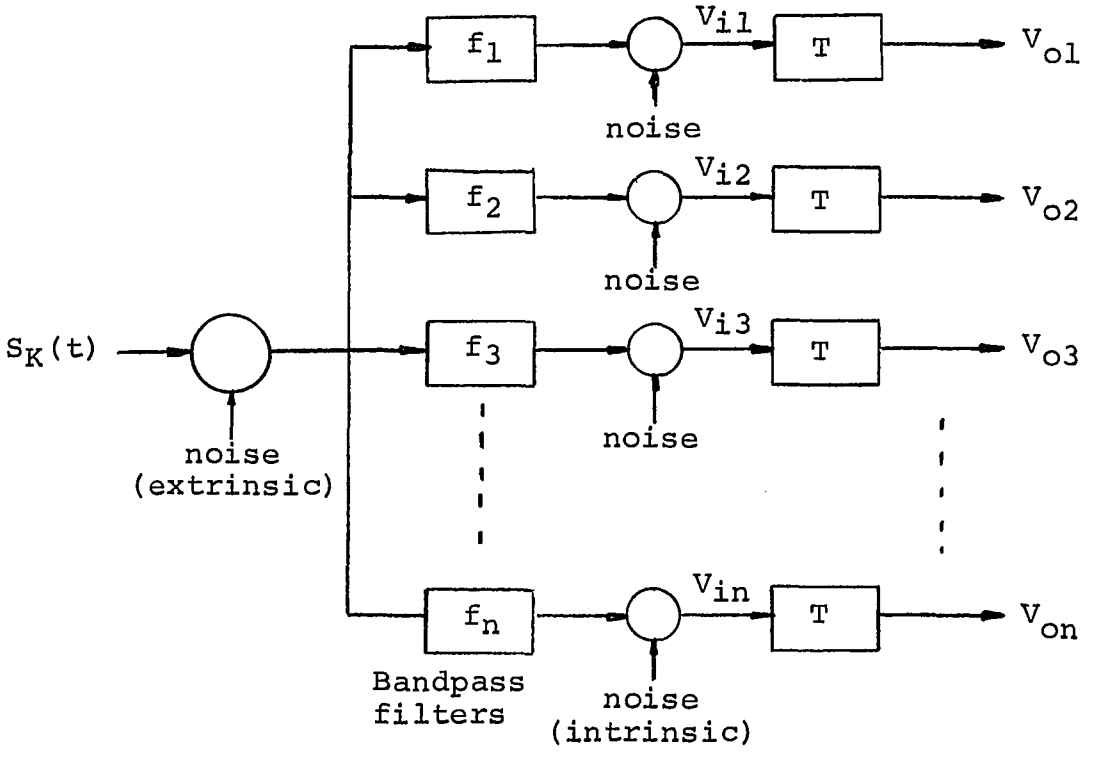


Figure 4. Basic parallel channels system (T = threshold detector).

$$P(CD/S_1) = pq^{(N-1)} \quad , \tag{4}$$

where p is the probability of the signal plus noise exceeding V_{th} in a single channel when a single formant is present and q is the probability of the noise not exceeding V_{th} in any channel in which a formant is not present.

Since the false alarm rate RA is

$$RA = 1 - q^N \tag{5}$$

or

$$q = (1 - RA)^{1/N} \quad , \tag{6}$$

$P(CD/S_1)$ may be expressed as

$$P(CD/S_1) = p(1 - RA)^{(1-1/N)} \quad . \tag{7}$$

The p and q terms are functions, respectively, of the Rayleigh-Rice and Rayleigh distributions [38, 46]. They can be conveniently evaluated by means of the Q function (well known in communications theory) [35, 54]. In terms of the Q function

$$p = Q(\sqrt{2Z},b) \tag{8}$$

$$q = 1 - Q(0,b) = 1 - e^{-b^2/2} \quad , \tag{9}$$

where

$$Q(\sqrt{2Z},b) = \int_{x=b}^{\infty} \exp\left(- Z - \frac{x^2}{2}\right) I_o(\sqrt{2Z}x)\,x\,dx \quad . \tag{10}$$

Z is the channel SNR defined on the basis of signal and

noise power; b is the detection threshold normalized by the square root of the noise power, σ, in a given channel (V_{th}/σ), and I_o is the Bessel function of zero order and imaginary argument.

In the more general K formant case,

$$P(CD/S_K) = \text{Prob.}(k = 1)p_1 q^{N-1} + \text{Prob.}(k = 2)(p_2)^2 q^{N-2} + \ldots + \text{Prob.}(k = K)(p_K)^K q^{N-K} \quad (11)$$

$$P(CD/S_K) = \left[\binom{N}{1}/N^K\right] p_1 (1 - RA)^{(N-1)/N} + \left[\binom{N}{2}/N^K\right] (p_2)^2 (1 - RA)^{(N-2)/N} + \ldots + \left[\binom{N}{K}/N^K\right] (p_K)^K (1 - RA)^{(N-K)/N} \quad (12)$$

or

$$P(CD/S_K) = \sum_{k=1}^{K} \binom{N}{k} (p_k)^k (1 - RA)^{(1-k/N)} , \quad (13)$$

where p_k is the probability of the signal plus noise exceeding V_{th} in a single channel when "k" formants are simultaneously present (i.e., the signal power per channel is reduced by k).

Equations (4) through (13) allow the determination of $P(CD/S_K)$ as a function of the mean channel SNR and the number of channels N. This calculation was carried out with the aid of an RCA Spectra 70 digital computer for a number

of different cases. Figure 5 shows the dependence of $P(CD/S_1)$ on Z for a single-formant source with N equal to 1, 10, and 100 and RA equal to .02. Figure 6 shows how the quantity Z_T, defined as the channel SNR necessary to achieve a 50 percent $P(CD/S_K)$, varies with N. Again a single-formant source has been assumed with values of RA equal to .01, .02, and .1.

As discussed earlier, noise can be thought of as being either intrinsic or extrinsic in nature. If the noise power (equal to σ^2) is intrinsic (originating after the initial filtering of $S_K(t)$ into individual channels), the two previous figures indicate that there is no advantage in conveying information over a number of channels rather than over one. The channel SNR and hence the input signal power necessary to achieve a given $P(CD/S_K)$ (error rate) increases as the number of channels in a P.C.S. is enlarged. This penalty in required signal power diminishes with decreasing values of false alarm rate and asymptotically approaches zero.

If the noise under consideration is extrinsic in origin, however, the situation can be reversed. Consider a fixed (overall) bandwidth system. As the number of channels is increased, the bandwidth (Δf) of the individual filters which compose the system must decrease. But now since the noise is generated externally (before the filtering of the signal into individual channels), σ^2 is proportional to the individual channel bandwidths;

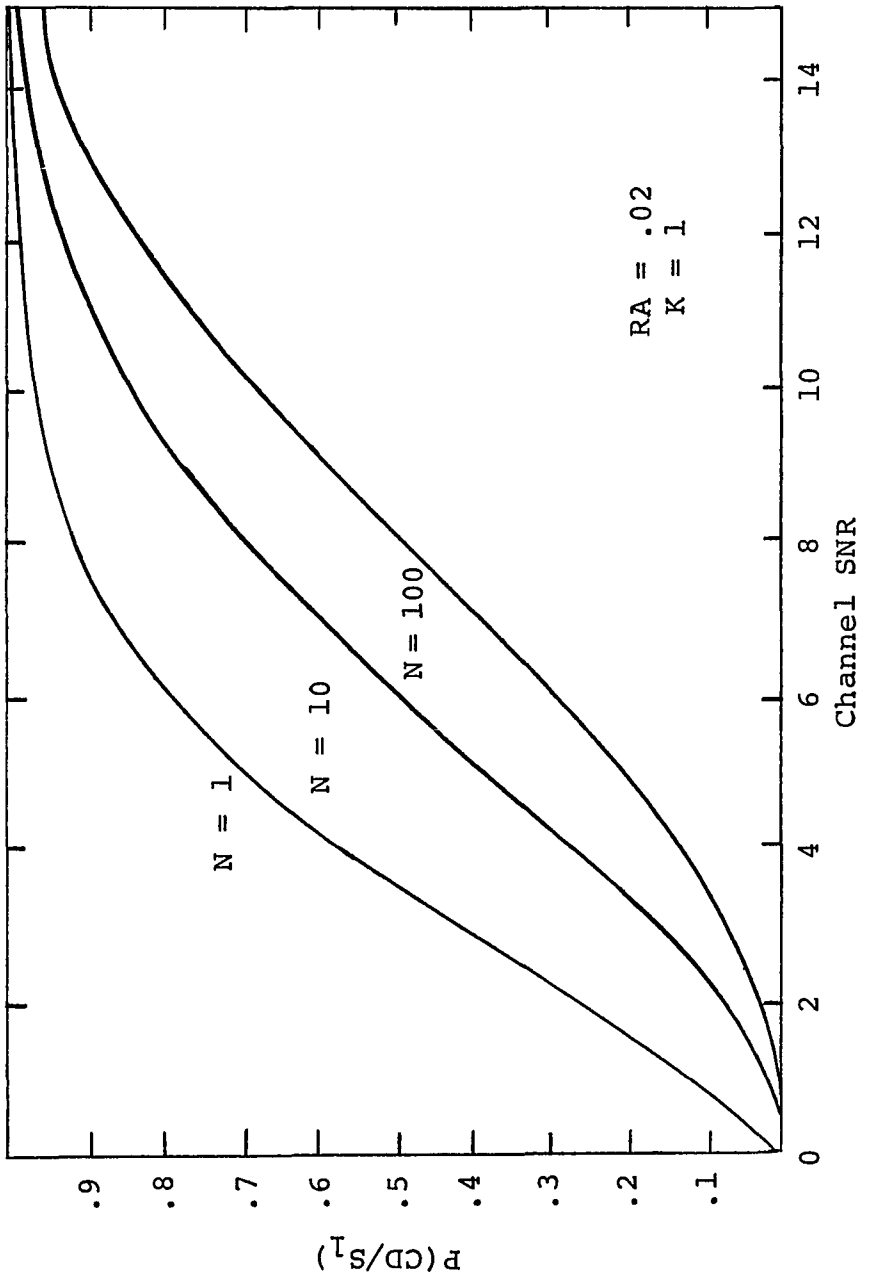


Figure 5. Probability of correct detection given the presence of a single formant as a function of channel signal-to-noise ratio for a basic parallel channels system with a false alarm rate of .02.

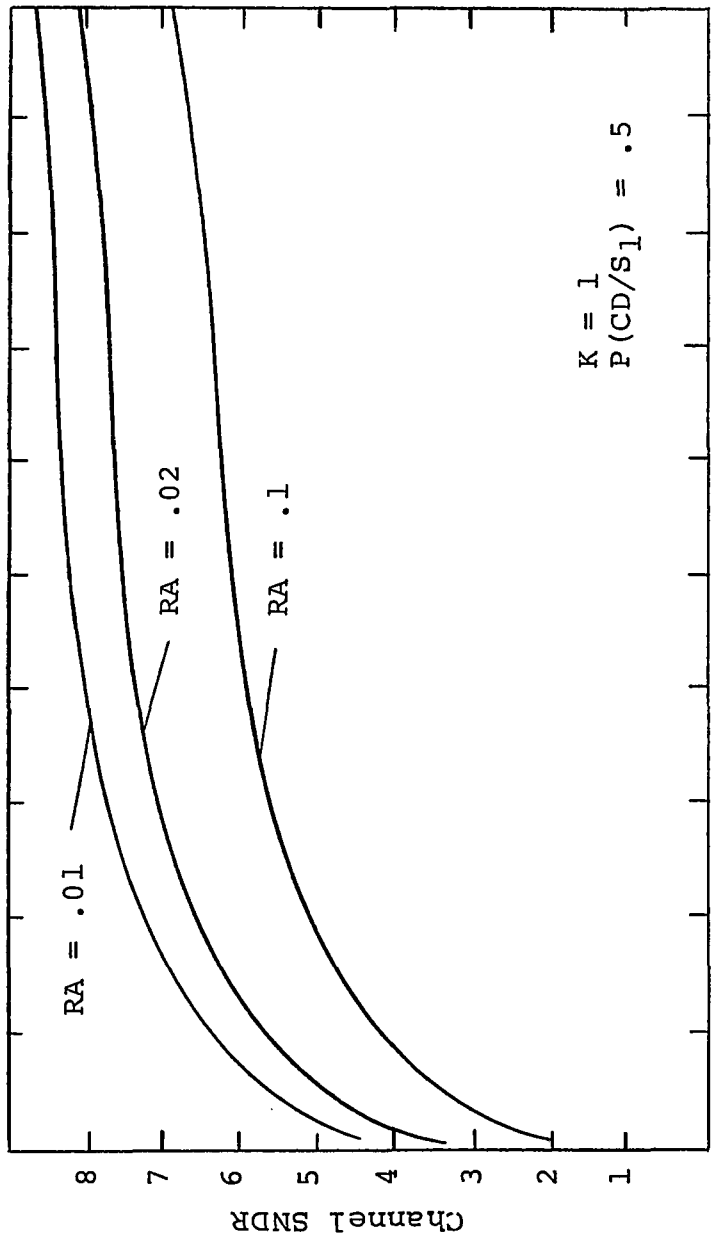


Figure 6. Channel SNR corresponding to a $P(CD/S_1) = .50$ as a function of number of channels and false alarm rate for a basic P.C.S. excited by a single formant source.

$$\sigma^2 = \eta\Delta f \tag{14}$$

where η is the (single sided) extrinsic noise density. Consequently, for a given noise density the noise power input to each channel must decrease and the mean signal-to-noise density ratio (SNDR) or input signal power necessary to achieve a given $P(CD/S_K)$ must decrease with increasing N. As a measure of the possible reduction in signal power (while maintaining system performance), the quantity

$$MTR = 10 \log Z_{TD}(1,1)/Z_{TD}(N,K) \tag{15}$$

(measure of threshold reduction) has been adopted. Z_{TD} is the SNDR necessary to achieve a $P(CD/S_K)$ of .50 and will be referred to as the threshold SNDR. $Z_{TD}(1,1)$ is the threshold SNDR for a single formant, single channel system, and $Z_{TD}(N,K)$ is the threshold SNDR for a K formant, N channel system. In Figure 7 the relationship between MTR and N is shown for a constant bandwidth P.C.S. with K equal to 1, 2, 4, and 6 and RA equal to .02. The figure indicates the substantial reduction in threshold signal power obtainable by implementation of a basic P.C.S. when dealing with extrinsic noise.

For signal elements of duration Δt, reduction of the channel bandwidths below a limiting value (Δf_c) established by the quantum condition of equation (1) will produce little or no improvement in $P(CD/S_K)$ and a loss of signal resolution

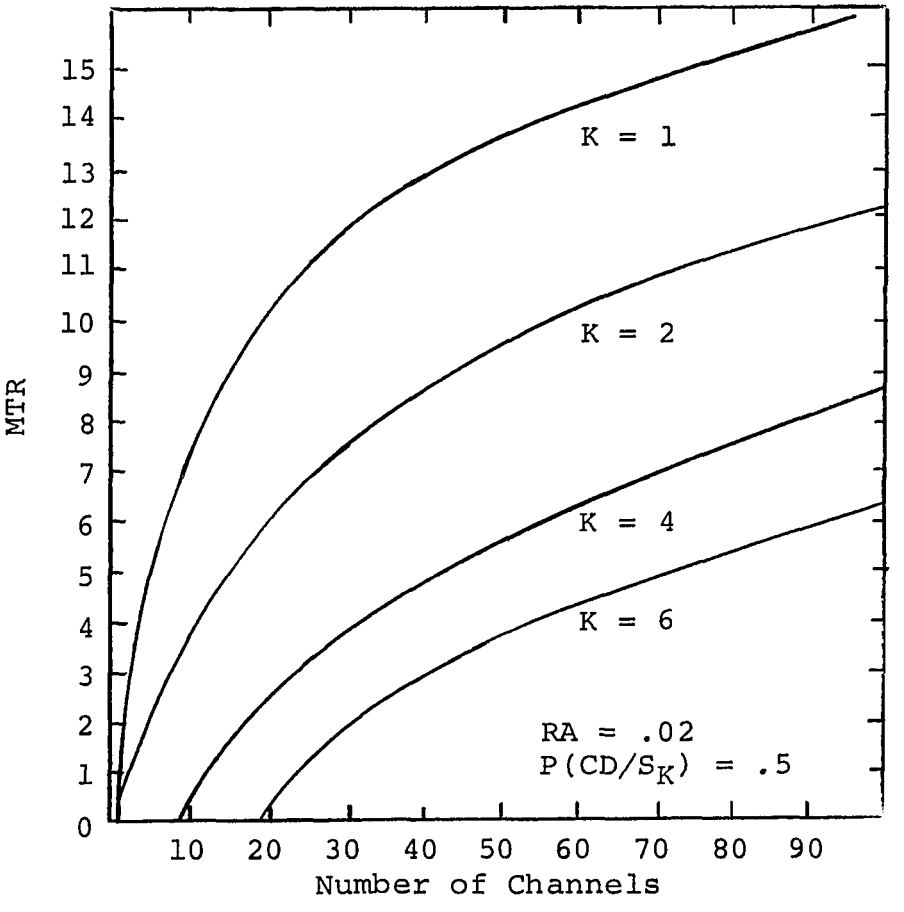


Figure 7. Reduction of SNDR (in db) necessary to achieve a $P(CD/S_K) = .50$ as the number of channels of a "fixed bandwidth" P.C.S. is increased. System false alarm rate was maintained at constant .02.

in the time domain. As a consequence, for an N channel fixed bandwidth P.C.S., an upper bound exists on the number of signal elements which can be efficiently processed in a given period of time. This limitation places a restriction on the maximum rate at which a P.C.S. can process information. This maximum information processing rate may be expressed in terms of the number of channels as

$$I = [1/\Delta t] \log_2 \sum_{k=0}^{K} \binom{N}{k} . \tag{16}$$

This relation and the restriction that the total system bandwidth of the basic P.C.S.

$$BW = N\Delta f \tag{17}$$

place the additional restriction on the system bandwidth that

$$BW \geq NI/\log_2 \sum_{k=0}^{K} \binom{N}{k} ; \tag{18}$$

from which the limiting channel bandwidth for a given information handling capability may be determined,

$$\Delta f_c = I/\log_2 \sum_{k=0}^{K} \binom{N}{k} . \tag{19}$$

With the aid of these last few equations, the relationship between MTR and N for a fixed information handling capability may be easily determined. Figure 8 shows how MTR is affected by N for this condition. From Figure 8 it can be seen that an improvement in system performance is always obtained by increasing N. However, the advantage

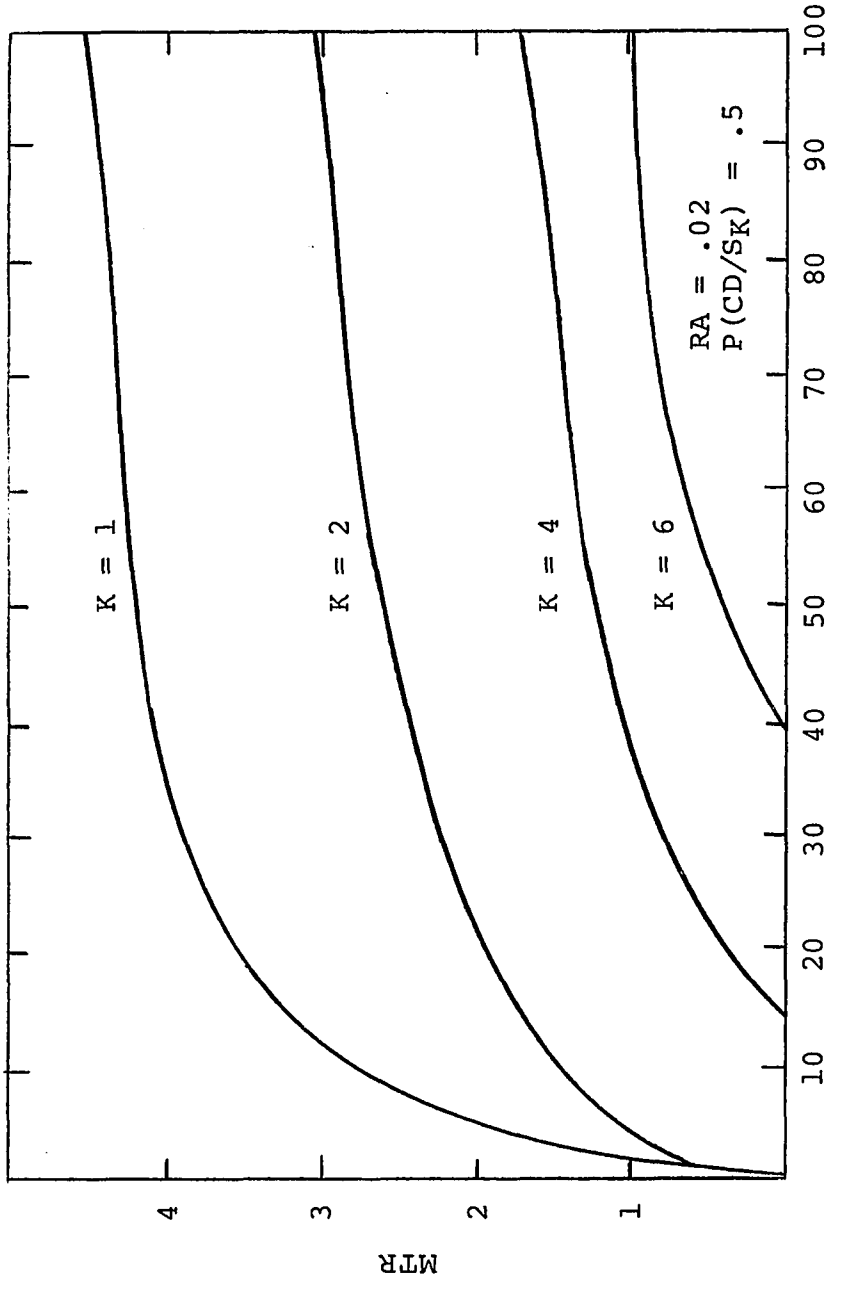


Figure 8. Reduction of SNDR (in db) necessary to achieve a $P(CD/S_K) = .50$ as the number of channels of "fixed information rate" P.C.S. is increased. System false alarm rate was maintained at a constant .02.

of increasing N becomes diminishingly small for N much greater than 30.

As N is increased, the overall system bandwidth must also increase since

$$BW/I \geq N/\log_2 \sum_{k=0}^{K} \binom{N}{k} \tag{20}$$

which explains why the MTR obtained in the fixed bandwidth case is larger than that obtained when the system's information handling capability is held constant and the number of channels is increased. In Figure 9 the relationship between MTR and normalized system bandwidth (BW/I) is shown. The amount of spectrum which must be sacrificed in order to obtain a specific MTR may be determined from this figure. It will be shown later that a P.C.S. employing centering is not restricted by equation (20) and hence an increase in N for such a system (while maintaining a constant information processing capability) need not be necessarily accompanied by an increase in system bandwidth.

Figures 7 and 8 both indicate that, for a fixed value of N, the best system performance is always obtained when the source is of the single formant type. It should be observed, however, that this conclusion is necessarily valid only for the assumed input conditions of white Gaussian noise and equiprobable (equal energy) formant signals. The assumption of different probability and energy distributions may provide improved performance for excitation by

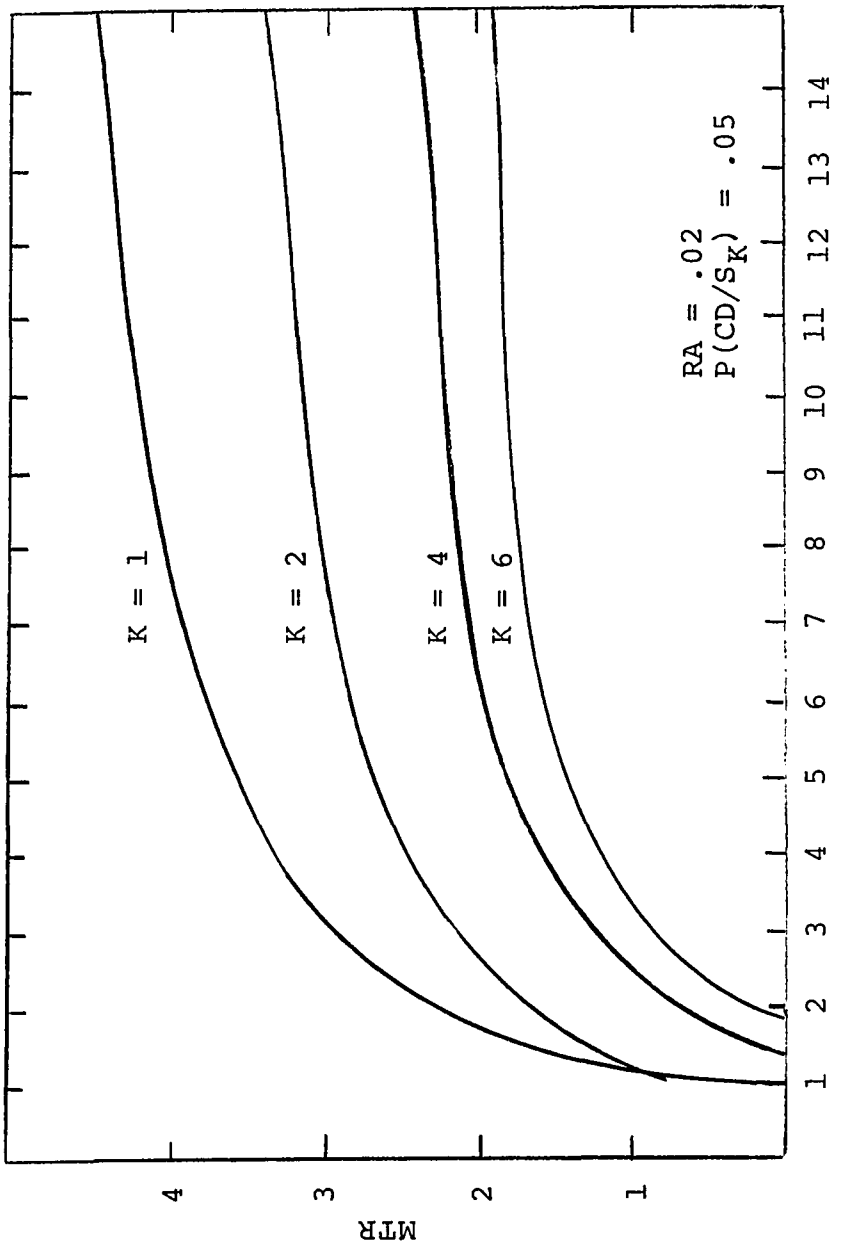


Figure 9. Reduction of SNDR (in db) necessary to achieve a $P(CD/S_K) = .50$ as function of normalized system bandwidth for a "fixed information rate" P.C.S. with a false alarm rate of .02.

multiple formant signal sources [13].[2]

---

[2]The excitation of the basic P.C.S. system with a redundantly encoded multiple formant signal (i.e., a signal where the detection of the presence of any one of a number of simultaneous formants indicates the same item of information) was investigated. This form of excitation always yielded poorer performance (for a given channel SNR) than the nonredundant form of excitation in the presence of white Gaussian noise. In the case of colored noise, however, where the mean noise energy is not stationary from channel to channel, this form of excitation can yield better performance than the nonredundant form [13].

4. PARALLEL CHANNEL SYSTEMS WITH IDEAL CENTERING

The relative performance of a parallel channel system employing an ideal centering process will now be evaluated. A P.C.S. with centering is illustrated by means of a block diagram in Figure 10a. It consists of a parallel set of filters, threshold detectors, and a centering network. This system differs from the previous system not only for the addition of the centering circuit but also because the bandpasses (Δf) of the individual filters are not necessarily orthogonal to each other but may overlap as shown in Figure 10b. Such an overlapping configuration is found in many physical systems both man man-made and biological [1].

The ratio between the channel bandwidths and their center-to-center frequency spacing (Δf_x) will be defined as:

$$M = \Delta f / \Delta f_x \ . \tag{21}$$

M shall be referred to as the channel overlap factor. This quantity is related to the number of filters which share common spectra (illustrated in Figure 10b) as follows:

$$NC = (2M - 1) \ . \tag{22}$$

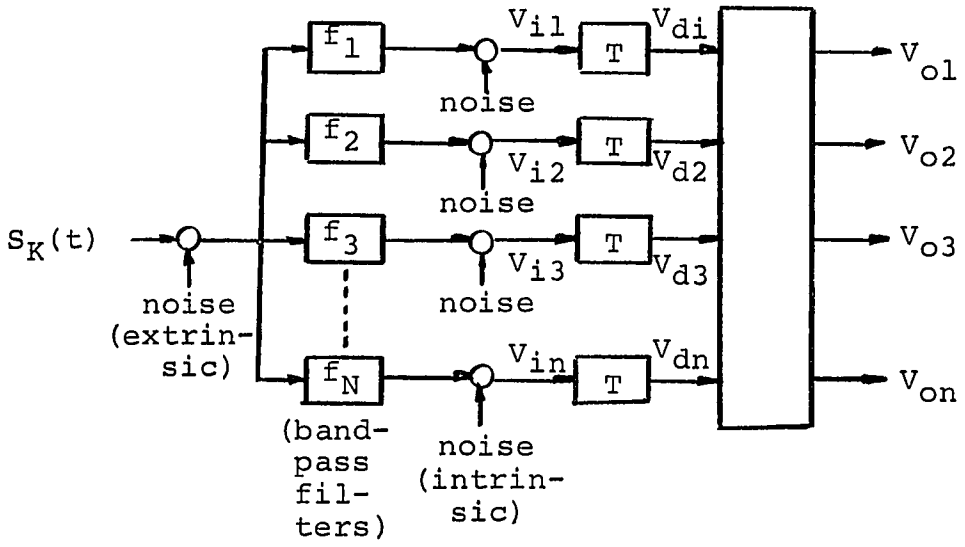


Figure 10a. Parallel channel system with centering. (T = threshold detector)

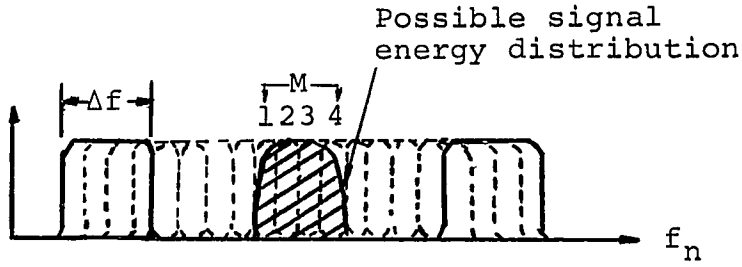


Figure 10b. Sample channel bandpass characteristics of a P.C.S. with centering for an overlap factor (M) = 4.

Since the spectral width of each formant of $S_K(t)$ is approximately equal to Δf (the bandwidth of the individual channels) and the filter bandwidths overlap, the presence of a single formant will cause a number of threshold detectors to fire simultaneously. The centering network, however, can estimate the specific channel in which a formant is centered on the basis of the outputs of the activated detectors. Thus, this system's frequency discrimination is not limited by the individual channel bandwidths.

An optimum estimate of the channel in which a formant is centered can be made from Bayes's decision rule [9]. This rule considers the likelihoods that a given pattern of activated channels (firing pattern) be caused by each of the possible input conditions. The particular input (from the set of all such possible input conditions) having the highest probability of causing the given output condition is considered to be the best estimate.

An N channel P.C.S. with centering can thus be used to estimate the frequency of formants which are not centered in any of the N channels of the given system but located somewhere between the center frequencies of two channels. In such an instance the set of allowable input frequencies must include frequencies which are not located at channel centers. The most sophisticated system would be one which considered the possibilities of an input formant of any frequency contained within its system bandpass. In this case the probability distribution of the estimated

frequency would be described by a density function. However, because of the positioning of the filters (which favors the detection of formants at their individual center frequencies), this density function would actually be composed of a number of peaks located primarily at channel centers and points midway between channel centers. (Even numbered clusters of activated channels have the highest likelihood of being caused by formants midway between channel centers; and odd numbered clusters have the highest likelihood of being caused by formants at channel centers.) Consequently, an approximate picture of the centering performance of a system which considers the possibility of many formant input frequencies (including frequencies between channel centers) can be obtained by evaluating the probability of estimated frequency distribution of a system which considers the possibility of formant frequencies only at channel centers and points midway between channel centers as inputs. In this thesis the ensemble of possible exciting frequencies shall be limited to channel centers and frequencies midway between channel centers.

All possible firing patterns of an N channel system can be represented by a set of 2^N "N" dimensional vectors (Y). (Each vector corresponds to one of the 2^N possible firing patterns.) The component element $y_{i,n}$ of a vector Y_i corresponding to the ith firing pattern would be equal to one if the nth channel is activated and equal to zero if it is not.

Consider initially a single formant signal only. The most likely channel (in terms of center frequency) of excitation (f_E) corresponds to a frequency f_j such that the conditional probability of a firing pattern Y_i given a formant at frequency f_j is

$$P(Y_i/f_j) > P(Y_i/f) \text{ : for all channels of } f \neq f_j \ . \quad (23)$$

Noise power is statistically independent from channel to channel in the case of intrinsic noise (as defined here) [53]. For the presence of intrinsic noise the conditional probability of a particular firing pattern can therefore be expressed as

$$P(Y_i/f) = \sum_{n=1}^{N} P(y_{i,n}/f) \ , \quad (24)$$

where the conditional probability of an individual channel firing or not given a formant of frequency f is

$$P(y_{i,n}/f) = Q(\sqrt{2Z_{f,n}},b) \qquad \text{if } y_{i,n} = 1$$

and

$$P(y_{i,n}/f) = 1 - Q(\sqrt{2Z_{f,n}},b) \text{ if } y_{i,n} = 0 \ . \quad (25)$$

The signal-to-noise ratio in the nth channel caused by a formant of frequency f is

$$Z_{f,n} = Z \cdot g(n,f) \ , \quad (26)$$

where g(n,f) is an amplitude function indicating the amount of signal energy which enters the nth channel for a given filter characteristic, formant frequency, and formant

spectral shape. Z is the SNR in the channel on which the given formant is centered (defined earlier).

The probability distribution of the estimated frequency of a given formant, $P(f_E/f_S)$, may be determined by summing the $P(Y_i/f_S)$ terms [equation (24)], for those Y_i for which f_E corresponds to the most likely excitation frequency [equation (23)]:

$$P(f_E/f_S) = \sum P(Y_i/f_S) \text{ over all } Y_i \text{ corresponding to the most likely } f_E \ . \qquad (27)$$

The $P(f_E/f_S)$ distribution has been evaluated for a number of different system constraints for formants of cosine squared spectral shape and square filter bandpass characteristics. These calculations, computational details of which are shown in Appendix A, were carried out on the RCA Spectra 70 computer.

In performing the above analysis, the question arises as to the evaluation of $P(f_E/f_S)$ when a unique f_E corresponding to a given Y_i does not exist. In such an instance the system was assumed to make a purely random choice among the several possible f_E frequencies.

In Figures 11 and 12 some of the computed results are shown. In Figure 11 the probability distribution of estimated frequency achievable by a 100 channel system with a false alarm rate of .02 and channel overlap factors of 3, 5, and 7 for different Z is shown. In Figure 12 the standard deviation (σ_f) of the estimated frequency (given

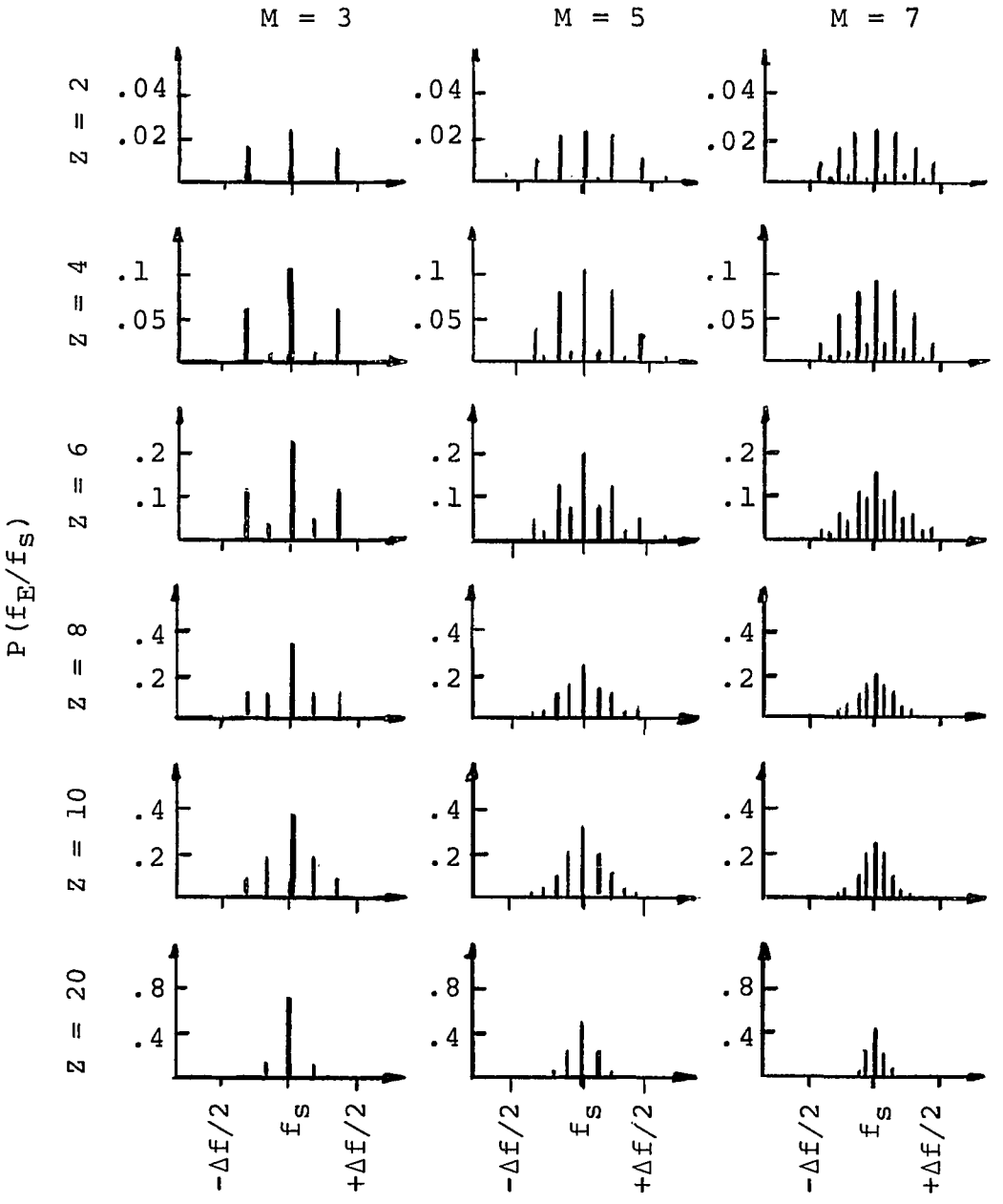


Figure 11. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with maximum likelihood centering and a false alarm rate of .02 in the presence of intrinsic noise.

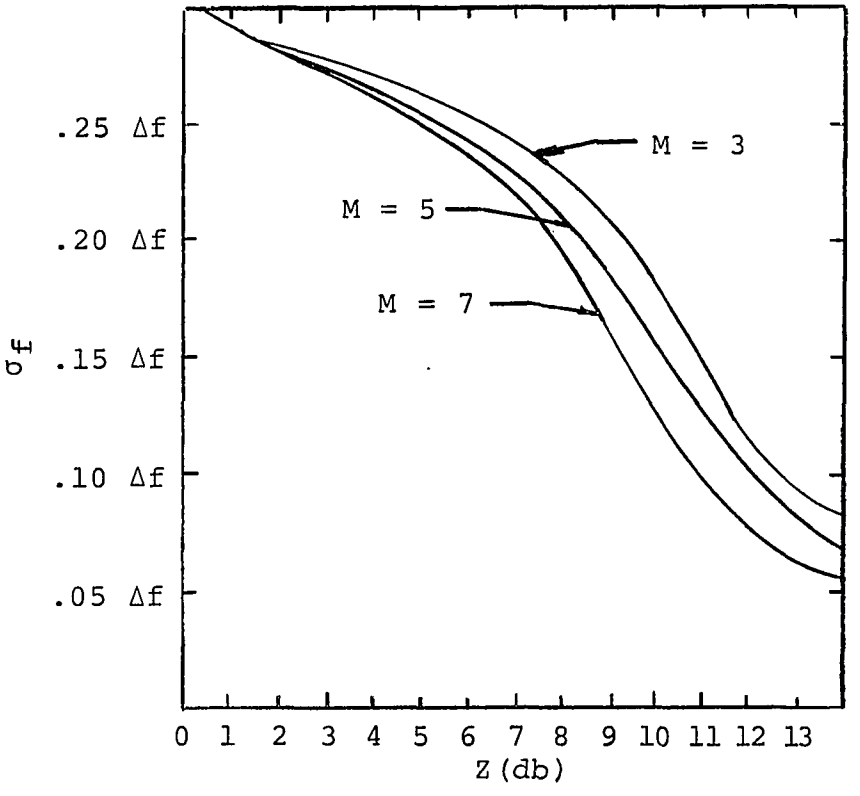


Figure 12. Relation between the standard deviation of the estimated frequency (given detection of the signal) and channel SNR for a 100 channel P.C.S. with maximum likelihood centering and a false alarm rate of .02 in the presence of intrinsic noise.

detection of the signal) is displayed as a function of channel SNR and overlap factor. It is significant that even at a channel SNR as low as 3 db a substantial amount of centering is possible ($\sigma_f = .27\ \Delta f$).

The probability distribution of the expected frequency only describes the frequency discriminating ability of a system. $P(f_E/f_s)$ does not, however, describe the overall ability of a system to process information. More insight into the performance of P.C.S. with centering can be obtained from the probability of correct detection resulting from the system's excitation by signals with different minimum frequency separation (ΔF_s) between individual formants. $P(CD/S_1)$ may be determined by summing the $P(f_E/f_s)$ terms within the range of ΔF_s.

$$P(CD/S_1) = \sum_{f_E=f_s-\Delta F_s/2}^{f_E+\Delta F_s/2} P(f_E/f_s) \ . \qquad (28)$$

From the geometry of the overlapping of individual filter bandpasses (Figure 10b) it can be seen that there will be no degradation of $P(CD/S_1)$ for values of ΔF_s greater than $\Delta f(2 - 1/M)$. This is true because, under this condition, the spread of signal energy is limited to the NC channels centered around f_s and because

$$(NC)\Delta f_x = (2 - 1/M)\Delta f \ . \qquad (29)$$

For the 100 channel system, whose frequency discrimination characteristics are displayed in Figure 11, $P(CD/S_1)$ varies

with ΔF_s as illustrated in Figure 13. The overlap factor, M equal to one, in Figure 13 represents the nonoverlapping case, i.e., the performance of the basic P.C.S. For formant separations greater than $2\Delta f/M$ the P.C.S. with centering generally yields a higher probability of correct detection than the basic P.C.S. However, for formant spacing less than $2\Delta f/M$, the performance of the P.C.S. with centering falls below that of the basic system.

In the case of extrinsic noise the mathematics describing the performance of P.C.S. with centering are more complex since not only is the signal spectrum common to a number of channels but also the noise spectrum is common to a number of channels [59]. Thus, the firing of any channel is dependent on the firing of the channels around it. The key conditional probability of a particular firing pattern given a formant of frequency f, $P(Y_i/f)$, may be expressed in terms of the joint probability densities of the energy within individual channels as follows:

$$P(Y_i/f) = X_1 \cdot \prod_{n=2}^{N-1} X_n \cdot \prod_{n=N}^{N+M-1} R_n \cdot X_N \,. \qquad (30)$$

The X_n terms represent the integration of multiple signal envelope functions [46]. The X_1 and X_N terms are separated since they must take into account edge effects. The R_n factors represent the integration of the probability density of the energy contained within independent spectral segments of noise energy. Details of the derivation of

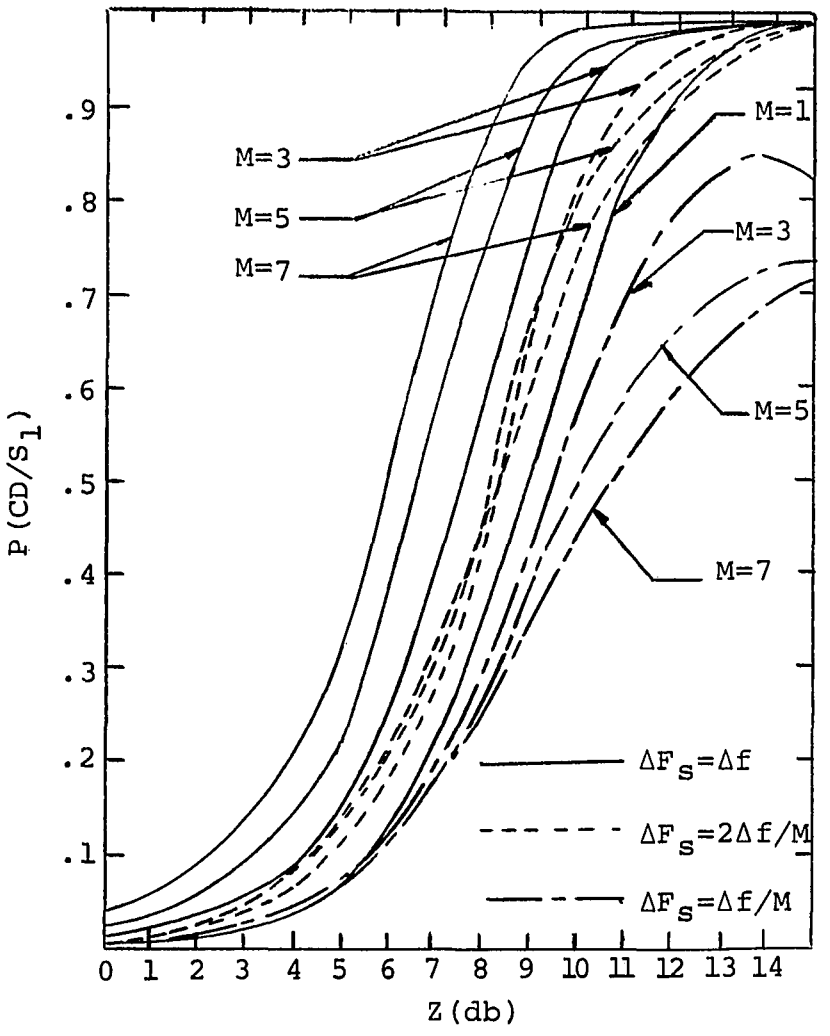


Figure 13. Probability of correct detection of a single formant provided by a 100 channel P.C.S. with maximum likelihood centering, intrinsic noise, and .02 false alarm rate as function of channel SNR, overlap factor, and minimum frequency separation between individual formants.

equation (30) are given in Appendix B.

Because of its many nested multiple integrals, it is impractical to evaluate equation (30) directly. The expression for $P(Y_i/f)$ may be simplified if only the firing patterns of those channels which receive energy from the exciting formant, f_s, are considered. Other firing patterns will have a negligible effect on the $P(f_E/f_s)$ distribution because of their small probability of occurrence and the position of their activated channels in relation to f_s. In such a case $P(Y_i/f_s)$ can be broken up into two factors:

$$P(Y_i/f_s) = P_{IS} \cdot PQ \ . \tag{31}$$

The first factor P_{IS} accounts for the firing probabilities of those channels which receive signal energy from the assumed formant. The other factor PQ accounts for the probability that all other channels do not fire, i.e., the probability that all channels which do not receive any energy from f_s will not fire.

Because of the correlation of noise energy between channels, PQ cannot be expressed as a simple function of the false alarm rate. PQ can be expressed approximately as

$$PQ \approx q(q_{n/n-1})^{N-NC-1} \ , \tag{32}$$

where q is the probability of an individual channel not firing given the presence of noise only [see equation (9)].

The $q_{n/n-1}$ term represents the probability of a channel not firing (given the presence of noise only) given that the channel directly adjacent (n-1) did not fire.

$$q_{n/n-1} = \{1/(M-1)[1 - Q(0,b)]\} \int_{S=0}^{\infty} \{1 - Q(S,\sqrt{M}b)\}^2 \, Se^{-S^2/2(M-1)} dS \; . \tag{33}$$

The derivation of equations 32 and 33 are discussed in Appendix C.

The P_{IS} factor, although containing significantly fewer terms, is still essentially described by equation (30). It is thus subject to the same limitations on direct evaluation. It is, however, practical to evaluate P_{IS} by Monte Carlo techniques [24]. In such an approach the system or part of the system of interest is mathematically simulated--usually by means of a digital computer. The response of the system is repeatedly determined a large number of times. From the results of these sample runs, various statistical parameters of the system may be determined--the larger the number of samples, the more accurate the statistical results.

The $P(f_E/f_S)$ distributions for a 100 channel system were evaluated by the above means for conditions identical to those assumed in the intrinsic noise case. Because of the adoption of the simplifications leading to equation (31), it was necessary to simulate only NC channels of the N total channels. Details of these computations are

presented in Appendix C. Figure 14 shows the resultant $P(f_E/f_S)$ distributions. Five thousand samples were taken for each input condition considered. This sample number yields a maximum error in the $P(f_E/f_S)$ distribution of less than ±2.12 percent with a 99.73 percent certainty [26]. In Figure 15 the relationship between the standard deviation of the estimated frequency σ_f and the channel SNR is shown. A comparison of these results with those of Figures 11 and 12 indicate that in the presence of extrinsic noise the system provides approximately equivalent centering performance to that provided when the same amount of intrinsic noise is present.

Following the same procedure as outlined in the intrinsic noise case [equation (28)] $P(CD/S_1)$ as a function of Δf_S, channel SNR, and M was calculated from the extrinsic noise $P(f_E/f_S)$ distributions. The results of these computations are shown in Figure 16. A comparison of these results with those presented in Figure 13 indicates that the absolute probability of detecting a formant (for a fixed false alarm rate) is lower in the presence of extrinsic noise.

Thus far our discussion has been limited to the special case of a single formant signal. In the case of the more general multiformant signal, the required performance of the centering mechanism is complicated by the possibility of the simultaneous presence of more than one formant. When a P.C.S. with centering is excited by several

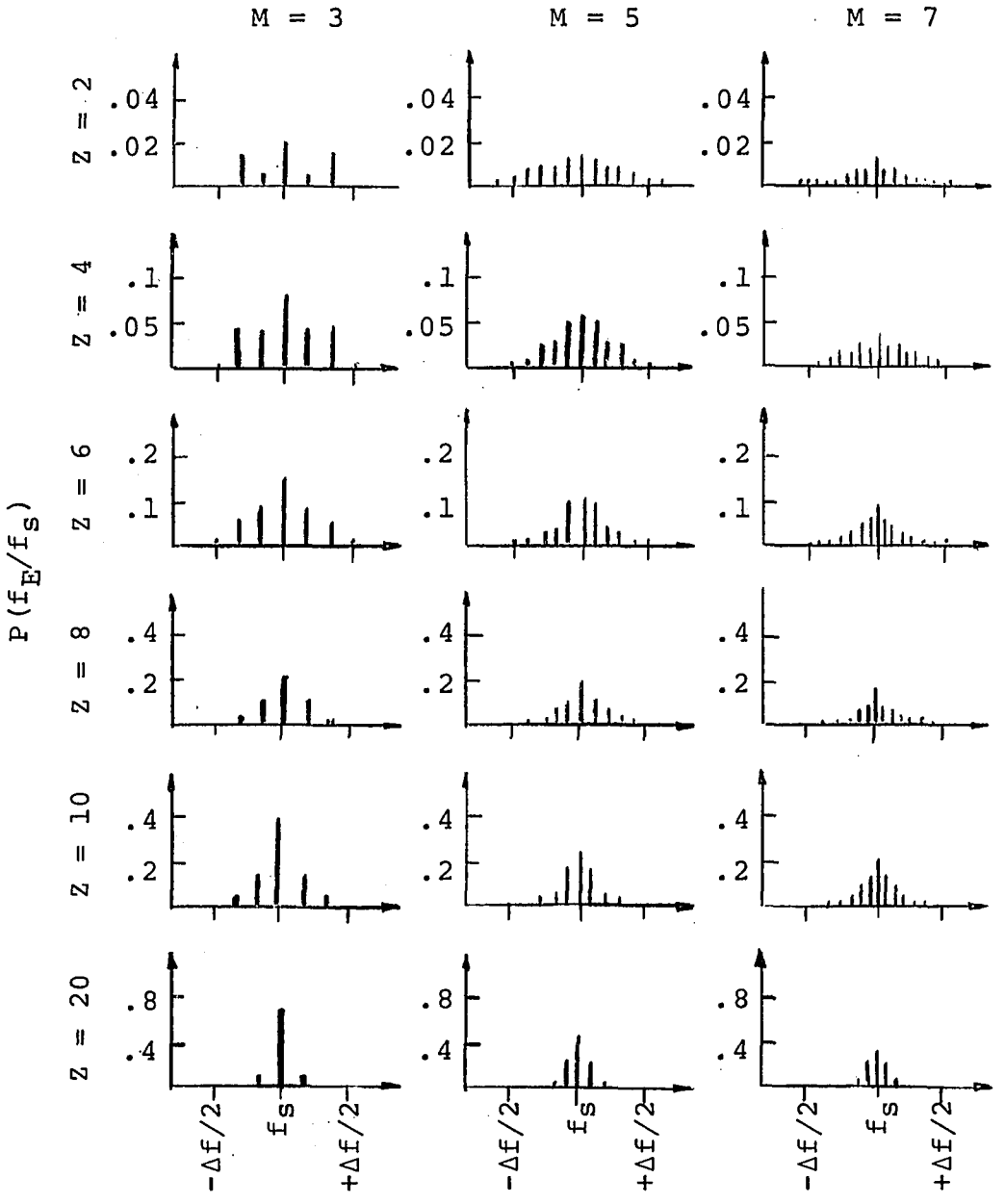


Figure 14. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with maximum likelihood centering and false alarm rate of .02 in the presence of extrinsic noise.

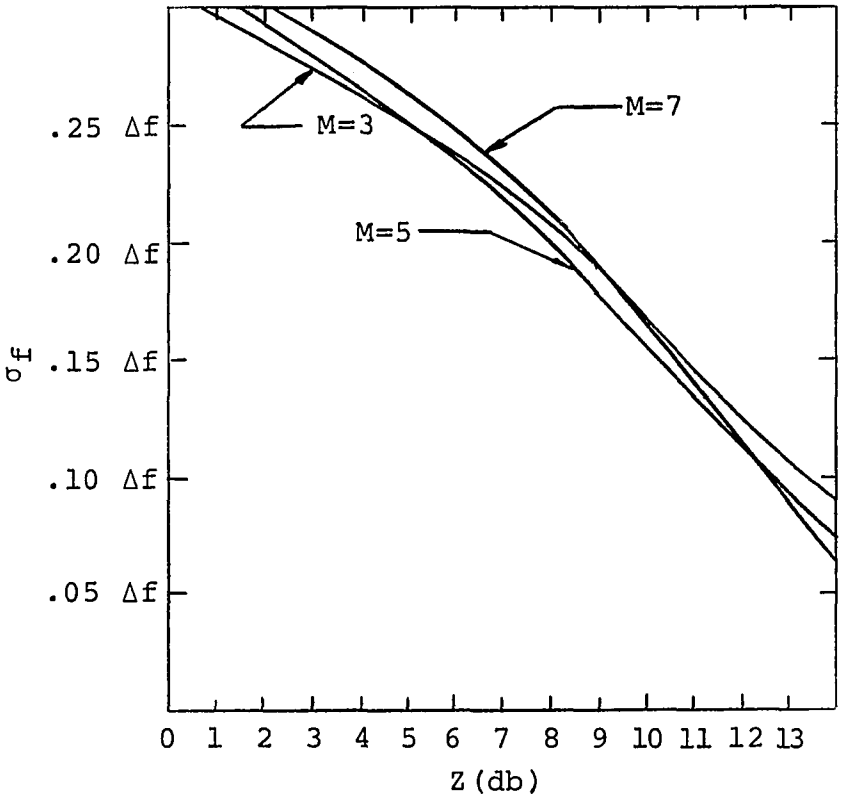


Figure 15. Relation between the standard deviation of the estimated frequency (given detection of the signal) and channel SNR for a 100 channel P.C.S. with maximum likelihood centering and a false alarm rate of .02 in the presence of extrinsic noise.

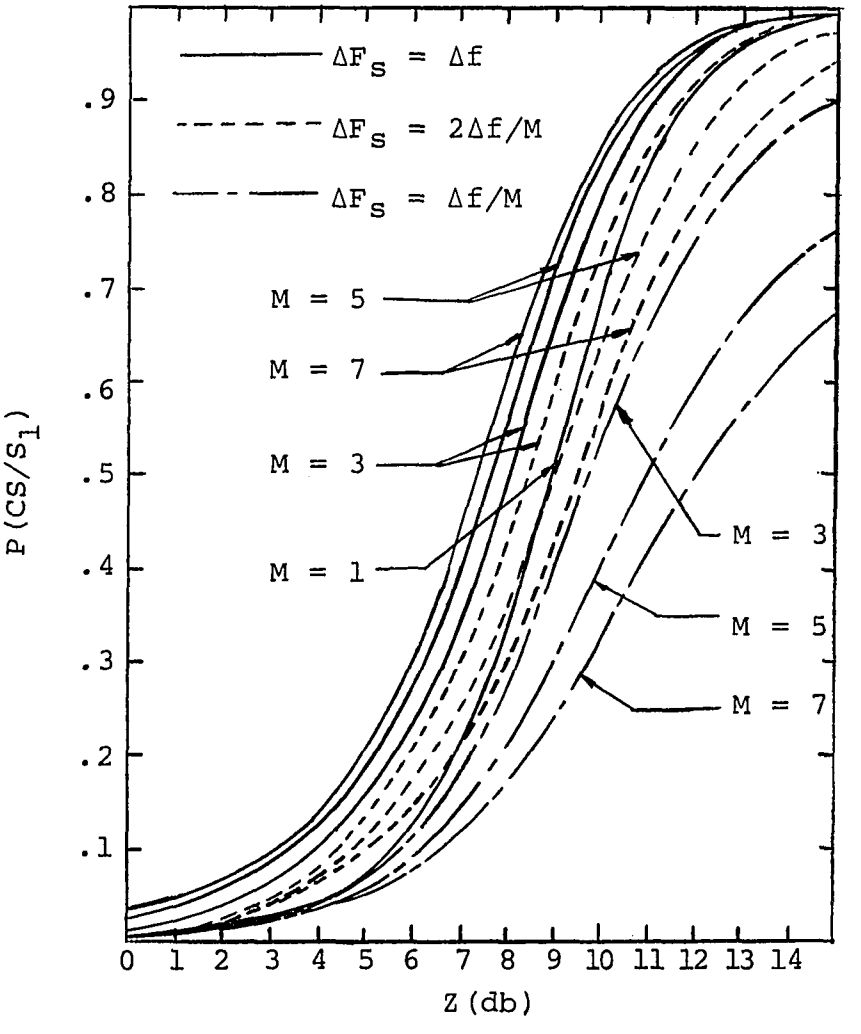


Figure 16. Probability of correctly detecting a single formant in extrinsic noise as a function of channel SNR, overlap factor, and minimum frequency separation between individual formants for a 100 channel P.C.S. with maximum likelihood centering and a .02 false alarm rate.

simultaneous spectral peaks, the performance of the system may be thought of as depending on two different processing abilities. The first is the ability of the system to determine the center frequency of the individual formants once they have been separated. This ability has been shown to be primarily a function of the spacing between the centers of the individual filters (this is the only ability required in the case of a single formant signal). The second is the ability of the system to resolve a signal into separate regions of excitation. This ability will be seen to be primarily a function of the bandpass characteristics of the individual channels. In the preceding discussion a distinction has been made between the ability of a system to discriminate frequency (detect a change in the frequency of a spectral peak) and a system's frequency resolving ability (ability to resolve a signal into its component spectral peaks).

By analyzing the geometry of filter overlap and the resultant crosstalk between channels (see Figures 10b and 17), it can be seen that if the spacing between the nearest (two) activated channels of (two) simultaneous spectral peaks is $\geq (NC)\Delta f_x$, then the resultant (local) firing patterns will not possess ambiguous centers. In order to always maintain this separation, the minimum spacing between (the centers) simultaneous formants (ΔF_{ss}) must be so restricted that

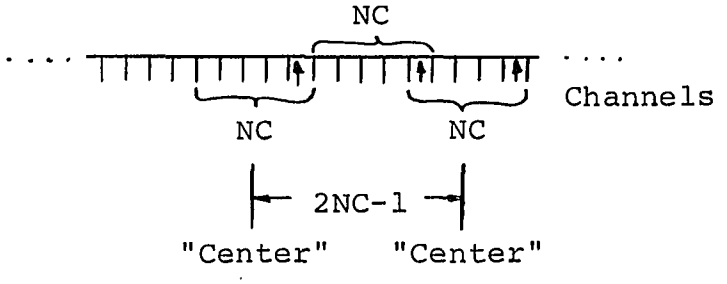


Figure 17. Diagrammatic display of the minimum spacing between simultaneous formants (ΔF_{ss}) necessary for the unambiguous resolution of their individual firing patterns.

$$\Delta F_{ss} \geq (2NC - 1)\Delta f_x = (4 - 3/M)\Delta f \quad . \tag{34}$$

When ΔF_{ss} is greater than $(4 - 3/M)\Delta f$, the overall centering process simply becomes that of simultaneously estimating the centers of a number of regional patterns of firing. In such a case, the single formant performance can be used to predict the performance of a system in response to multiple formant signals, i.e.,

$$P(CD/S_K) = \sum_{k=1}^{K} \binom{N}{k} P(CD/S_1)^k \, (1 - RA)^{(1-k)\frac{NC}{N}} \quad . \tag{35}$$

On the other hand, if ΔF_{ss} is less than $(4 - 3/M)\Delta f$, the decision process must consider the possibility that a local firing pattern may have been caused by more than one spectral peak. Such a condition greatly complicates the maximum likelihood decision process and will not be discussed quantitatively here.

5. PRACTICAL CENTERING PROCESSES

In the last section the performance of a P.C.S. employing an "ideal" centering process was analyzed. It was shown that, by the addition of this centering network, superior frequency discrimination in the presence of both extrinsic and intrinsic noise could be obtained. It was further shown that over a wide range of input conditions a P.C.S. with ideal centering (when considered as an information transmission system) could yield the highest $P(CD/S_K)$ for a given channel SNR. Unfortunately, the maximum likelihood process assumed in these previous calculations is not easily implemented in practice. An idea of the difficulties involved in putting an ideal centering process into practice can be obtained from Appendix C, describing the simultation of a maximum likelihood centering process on the digital computer.

There are a number of other types of center estimators which, while not offering as good an estimate of formant frequency as the maximum likelihood process, do offer relatively easy implementation. One such process is shown (in terms of digital logic elements) in Figure 18. This system finds the median of a cluster of activated channels. For clarity this three-dimensional network is illustrated by

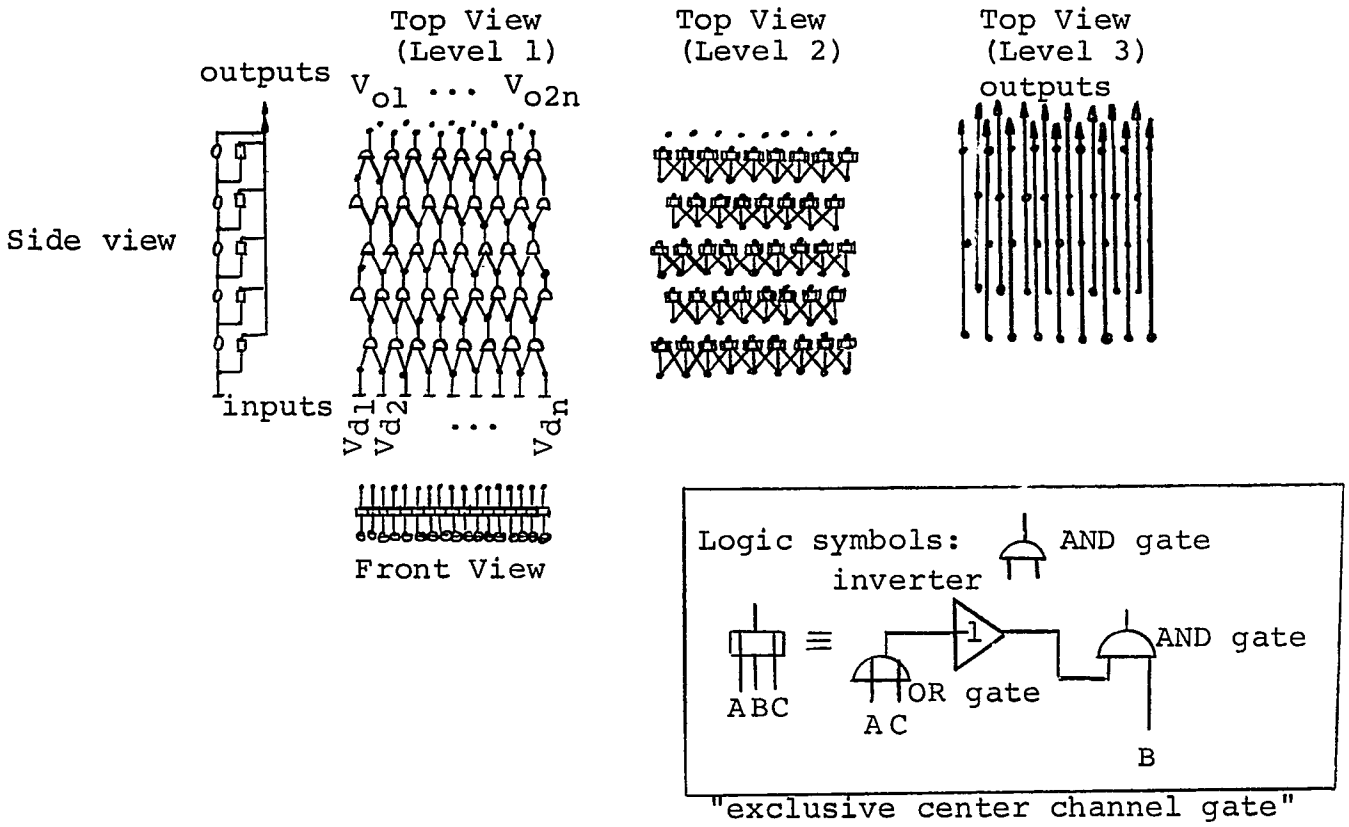


Figure 18. Median channel centering network employing digital logic elements.

means of orthographic projection. The first layer (composed of a number of AND gates) determines the center channel of a cluster of activated channels. The second layer (composed of a number of NOR and AND gates) decides when a center has been found at the first level and transfers that "center" to the proper output channel on the third level.

One of the problems in developing a satisfactory centering algorithm is the finding of a circuit which can accurately estimate (discriminate) the frequency of individual formants and yet be able to separate (resolve) a number of simultaneous formants into their respective spectral peaks. The median channel (M.C.) centering process has the advantage of always finding a unique center for a given cluster of activated channels; and thus under special conditions it is possible that it would resolve simultaneous spectral peaks even if their centers were separated by only $2\Delta f_x$. (The formant frequencies thus estimated will always be located at channel centers or at points midway between channel centers.) This system, however, has the inherent disadvantage of easily mistaking a number of formants for a single formant.

The estimated frequency of a formant (f_E) provided by an M.C. centering process when only a single cluster is present will be

$$f_E = (f_a + f_b)/2 \ , \tag{36}$$

where f_a and f_b are, respectively, the lower and upper

center frequencies of the two channels which bound the cluster. The probability distribution of the estimated frequency given a formant, $P(f_E/f_S)$, achievable by a M.C. centering process can be calculated by summing the probabilities of activation of each and all firing patterns which correspond to a given estimated frequency [equation (27)]. Since noise power in the case of intrinsic noise is statistically independent from channel to channel, the probability of activation of a given firing pattern ($P(Y_i/f_S)$ can be expressed simply as the product of the probabilities of firing of the individual "on" channels in the pattern times the product of the probabilities of not firing of the "off" channels in the pattern. The $P(f_E/f_S)$ distribution may thus be expressed as follows:

$$P(f_E/f_S) = \sum_{k=1}^{(N+1)/2-\Delta_o} \left\{ \prod_{n=\Delta_L}^{(N-2k+1)/2-D} [1 - Q(\sqrt{2 \cdot z_{f,n}}, b)] \right.$$

$$\cdot \prod_{n=(N-2k+3)/3-D}^{(N+2k-1)/2-D} [Q(\sqrt{2 \cdot z_{f,n}}, b)]$$

$$\left. \cdot \prod_{n=(N+2k+1)/2-D}^{N-\Delta_u} [1 - Q(\sqrt{2 \cdot z_{f,n}}, b)] \right\} \qquad (37)$$

when f_E is a channel center and as

$$P(f_E/f_s) = \sum_{k=1}^{(N+1)/2-\Delta_o} \left\{ \prod_{n=\Delta_L}^{(N-2k+1)/2-D} [1 - Q(\sqrt{2 \cdot Z_{f,n}}, b)] \right.$$

$$\cdot \prod_{n=(N-2k+3)/3-D}^{(N+2k-1)/2-D-1} [Q(\sqrt{2 \cdot Z_{f,n}}, b)]$$

$$\left. \cdot \prod_{n-(N+2k+1)/2-D-1}^{N-\Delta_u} [1 - Q(\sqrt{2 \cdot Z_{f,n}}, b)] \right\} \qquad (38)$$

when f_E is a point midway between channel centers. D is a factor establishing the effect of the position of f_E on the probability distribution. If f_s is assumed to be at the center channel of an N channel system and the estimated frequency, f_E, is at a channel center, D equals the number of channels separating f_E and f_s. D is positive if f_E is less than f_s and D is negative if f_E is greater than f_s. When f_E is a frequency located midway between channel centers, D is equal to the number of channels separating f_E and f_s plus 1/2. The Δ factors account for edge effects. The Δ_o's always equal D, while the Δ_u's equal zero and the Δ_L's equal D when f_E is greater than f_s, and the Δ_u's equal D and the Δ_L's equal zero when f_E is smaller than f_s. The Q function of equations (37) and (38) were defined in equation (10), and the $Z_{f,n}$ terms were defined as a function of Z in equation (26).

From equations (37) and (38) the relationship between the $P(f_E/f_s)$ distribution and the channel SNR (Z) can be

determined. Using these two equations, $P(f_E/f_s)$ was evaluated by computer techniques for formants of cosine squared spectral shape and square filter bandpass characteristics. In this evaluation firing patterns which correspond to the presence of more than one formant were assumed to add nothing to the probability of estimating a single formant frequency (f_E). Details of these calculations are given in Appendix A.

In Figure 19, the resultant relationship between $P(f_E/f_s)$ and Z is shown for a 100 channel P.C.S. system with a false alarm rate of .02 and channel overlap factors of 3, 5, and 7. In Figure 20, the (standard deviation of f_E) σ_f as function of Z (with M as a parameter) is presented.

The probability of correct detection for a P.C.S. with M.C. centering was also determined. $P(CD/S_1)$ was calculated from the above $P(f_E/f_s)$ distributions by the method outlined in equation (28). For the illustrative 100 channel system, $P(CD/S_1)$ varies with Z, ΔF_s, and M as shown in Figure 21.

Inspection of these last three figures reveals that the ability of the M.C. process to discriminate small changes in formant frequency is surprisingly good when compared to that of the maximum likelihood process. However, the overall ability of the process to correctly detect a formant (and hence process information) is inferior to the maximum likelihood process. Another property

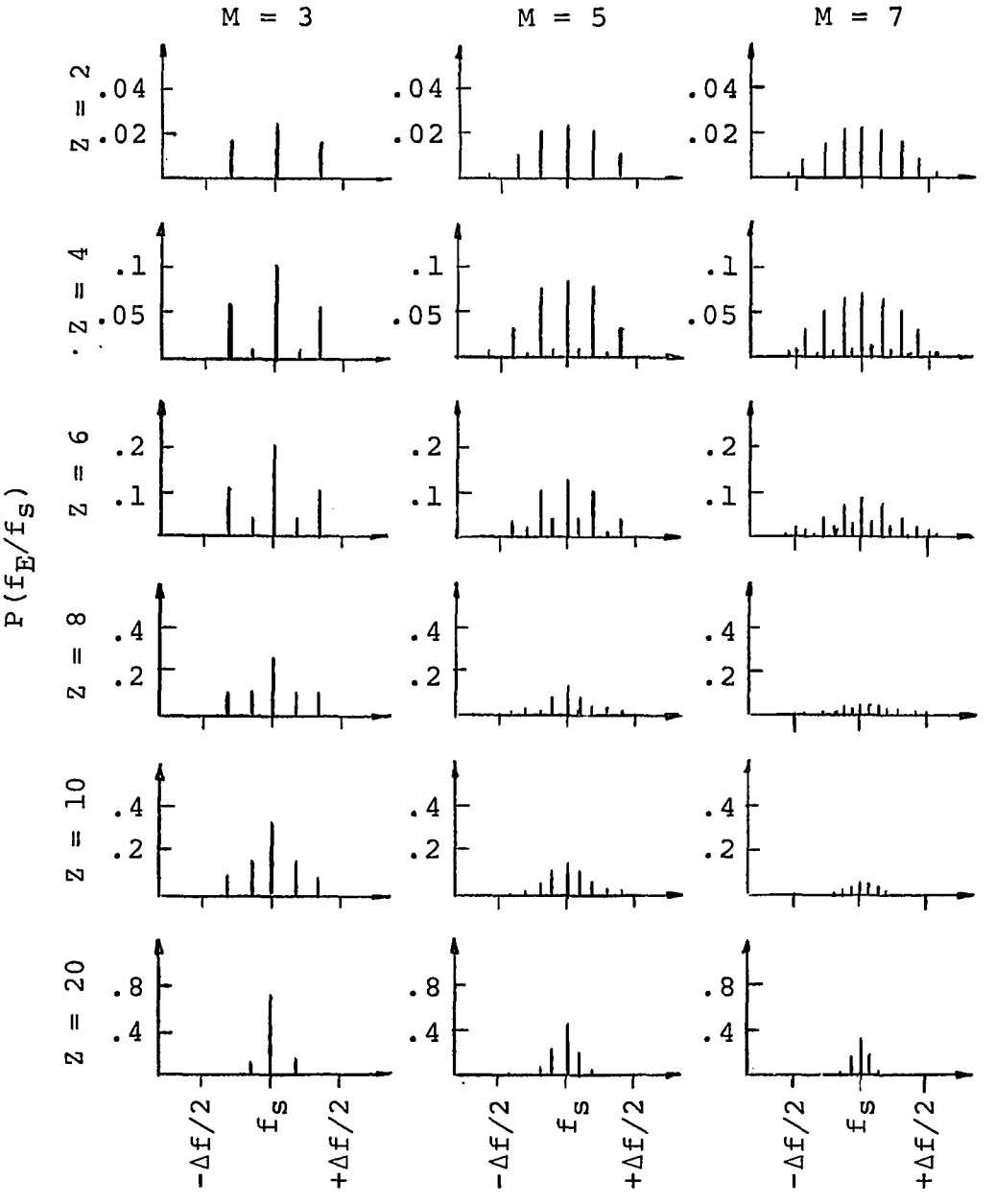


Figure 19. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with median channel centering and a false alarm rate of .02 in the presence of intrinsic noise.

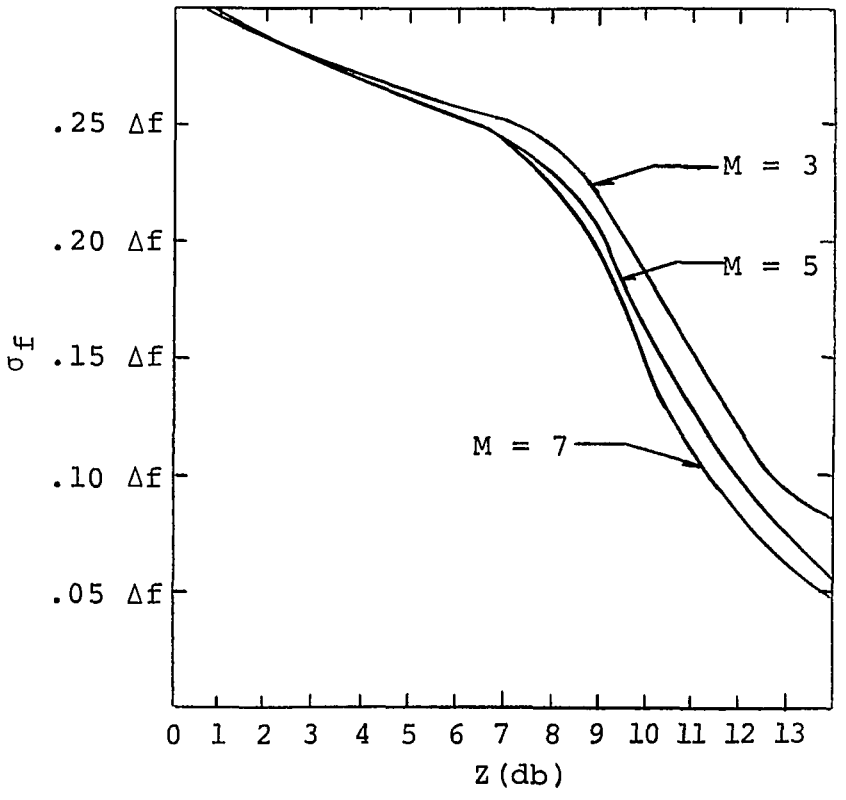


Figure 20. Standard deviation of the estimated frequency (given detection of the signal) for a 100 channel P.C.S. with median channel centering and a false alarm rate of .02 in the presence of intrinsic noise as a function of channel SNR and overlap factor.

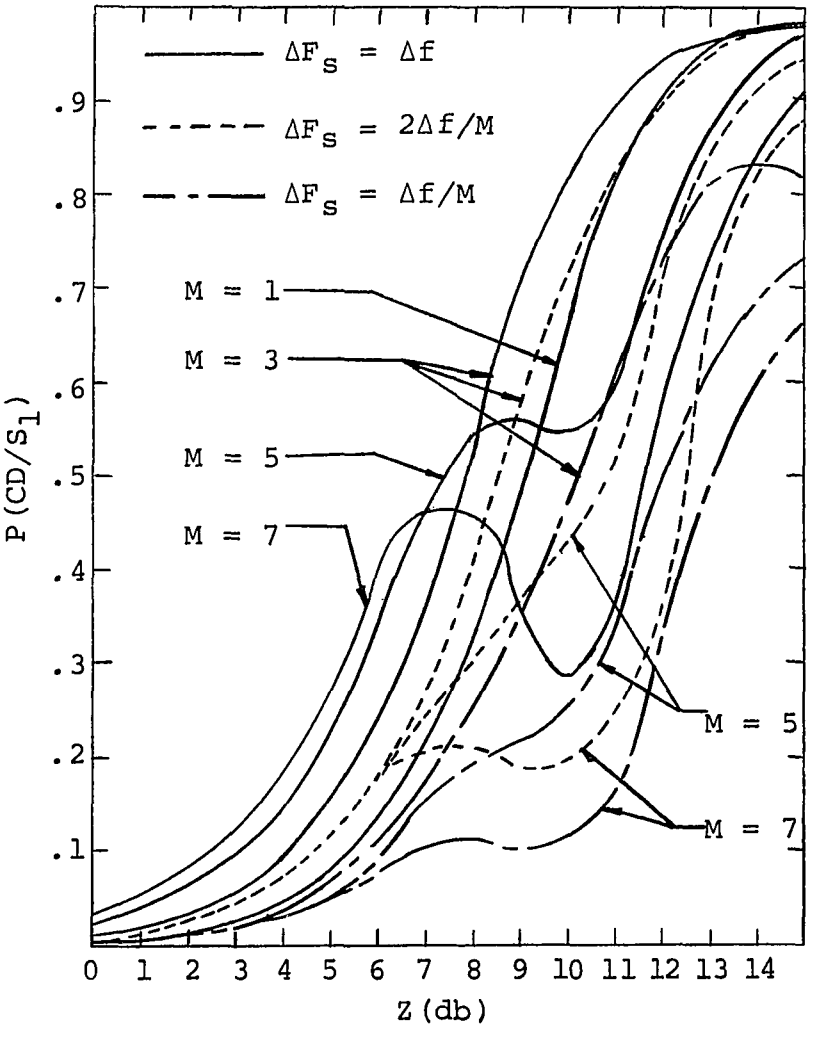


Figure 21. Probability of correct detection of single formant provided by a 100 channel P.C.S. with median channel centering, intrinsic noise, and a false alarm rate as a function of channel SNR, overlap factor, and minimum frequency separation between individual formants.

of the M.C. centering process revealed in Figure 21 is the non-monotonic relationship between $P(CD/S_1)$ and Z. For certain signal levels (and a given channel noise level) an attenuation of the signal will increase $P(CD/S_1)$.

In the case of extrinsic noise, the mathematical evaluation of system performance again becomes complicated by the correlation of noise energy between channels as it did for the maximum likelihood centering process. In this case, the approach of equations (37) and (38) must be abandoned and an approach similar to that used in evaluating the performance of the maximum likelihood process adopted. $P(Y_i/f_s)$ can be broken up into two terms [equation (31] and a Monte Carlo simulation used to evaluate the first term (P_{IS}). The other term (PQ), which accounts for the probability of not firing for all channels which do not receive signal energy, will be identical to the PQ term calculated for the maximum likelihood centering process as the value of this term does not depend on the type of centering process employed [see equations (32) and (33)].

The $P(f_E/f_s)$ distribution was evaluated, by the above means, for a 100 channel system operating under the same conditions assumed previously, except for the assumption of the presence of extrinsic noise. Details of the Monte Carlo simulation used are given in Appendix C. Figures 22 and 23 show the results of these calculations. Five thousand samples were again taken for each input condition. Figure 22 presents the relationship between probability

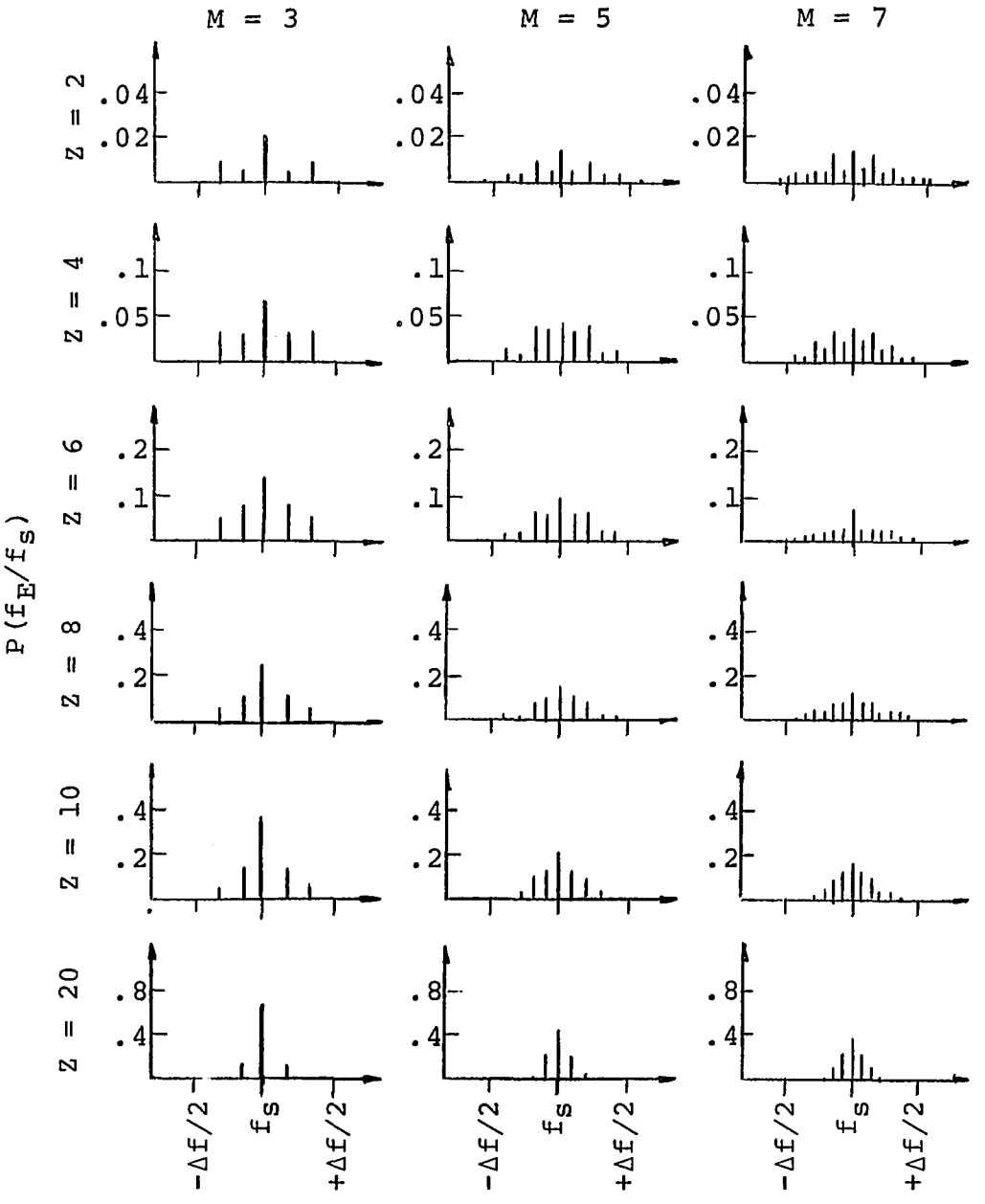


Figure 22. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with median channel centering and a false alarm rate of .02 in the presence of extrinsic noise.

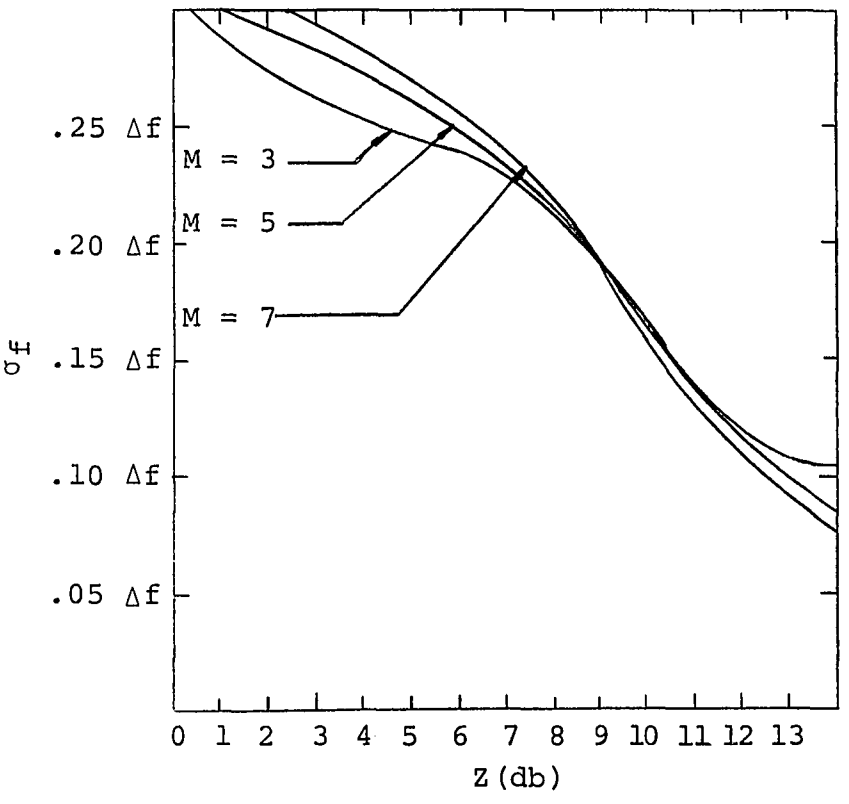


Figure 23. Standard deviation of the estimated frequency (given detection of the signal) for a 100 channel P.C.S. with median channel centering and a false alarm rate of .02 in the presence of extrinsic noise as a function of channel SNR and overlap factor.

distribution of the estimated frequency and channel SNR for the M.C. centering process operating in the presence of extrinsic noise. In Figure 23 the relationship between σ_f and the channel SNR is shown. From the $P(f_E/f_s)$ distributions, $P(CD/S_1)$ was calculated [equation (28)]. Figure 24 presents the relationship between $P(CD/S_1)$ and Z, F_s, and M.

As in the case of intrinsic noise, the ability of the M.C. process to discriminate small changes in formant frequency (in the presence of extrinsic noise) is very near to the performance achieved by the maximum likelihood process. The overall information processing ability of the M.C. process in the case of extrinsic noise, as indicated by $P(CD/S_1)$ in Figure 24, is also close to that displayed by the maximum likelihood process, and does not exhibit the large fluctuations with Z displayed by the M.C. process in the presence of intrinsic noise.

Another method for estimating the frequency of an exciting formant from a pattern of activated channels is shown in Figure 25a. The system shown estimates the frequency of an input formant on the basis of the first moment (or mean frequency) of the center frequencies (f_n) of the activated channels.[1] A similar system has been employed to identify the principal formant frequencies of speech in a noise-free environment [52]. Because of the assumption of

---

[1]In this process, f_E is not displayed by means of a number of parallel outputs. It is displayed by a single output voltage (V_f), whose amplitude is proportionally related to the position of the center of excitation.

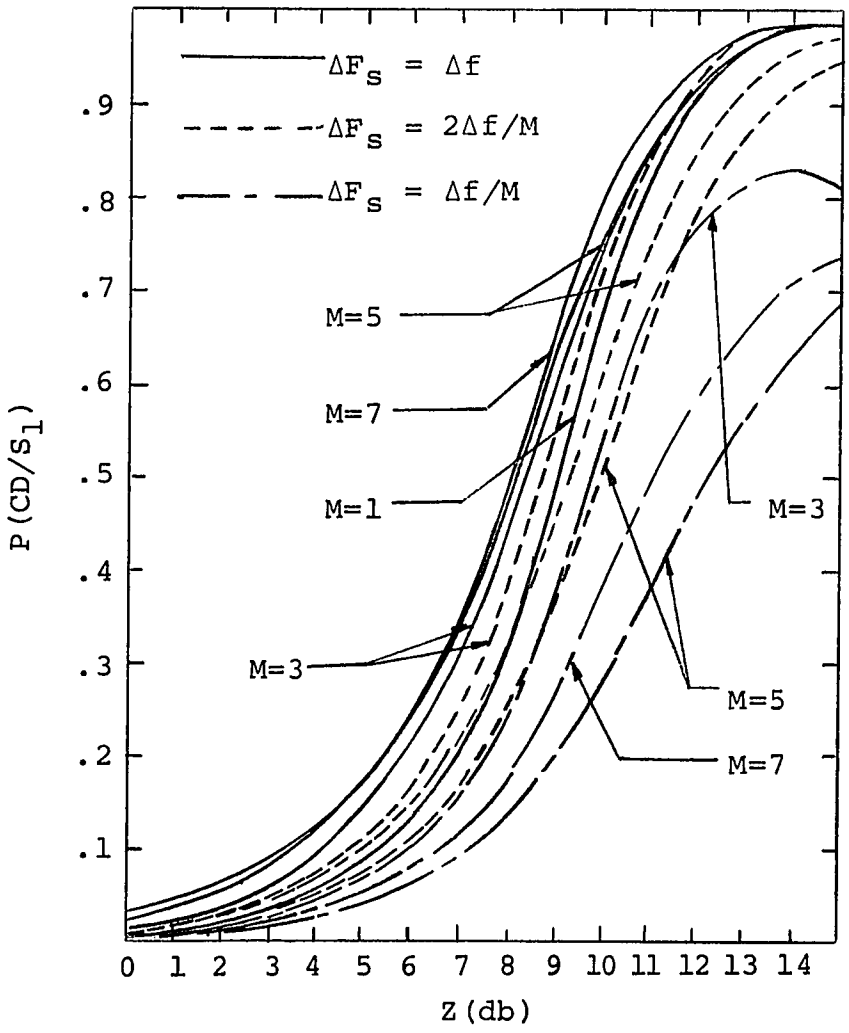


Figure 24. Probability of correctly detecting a single formant in extrinsic noise as a function of channel SNR, overlap factor, and minimum frequency separation between individual formants for a 100 channel P.C.S. with median channel centering and a .02 false alarm rate.

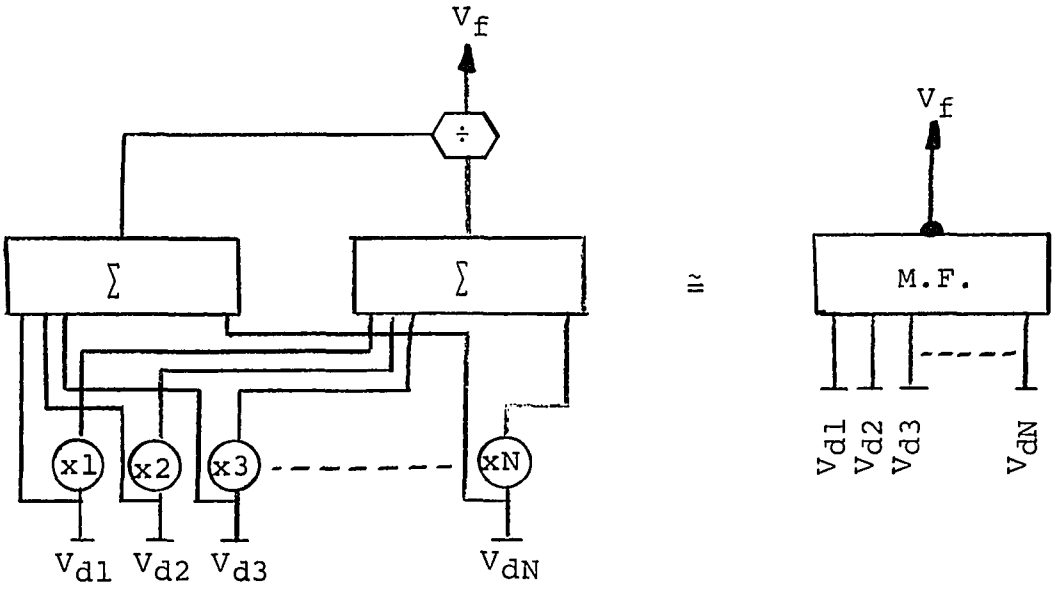


Figure 25a. Black diagram of a mean frequency centering process. Amplitude of the voltage V_f is proportionally related to the position of the center of excitation.

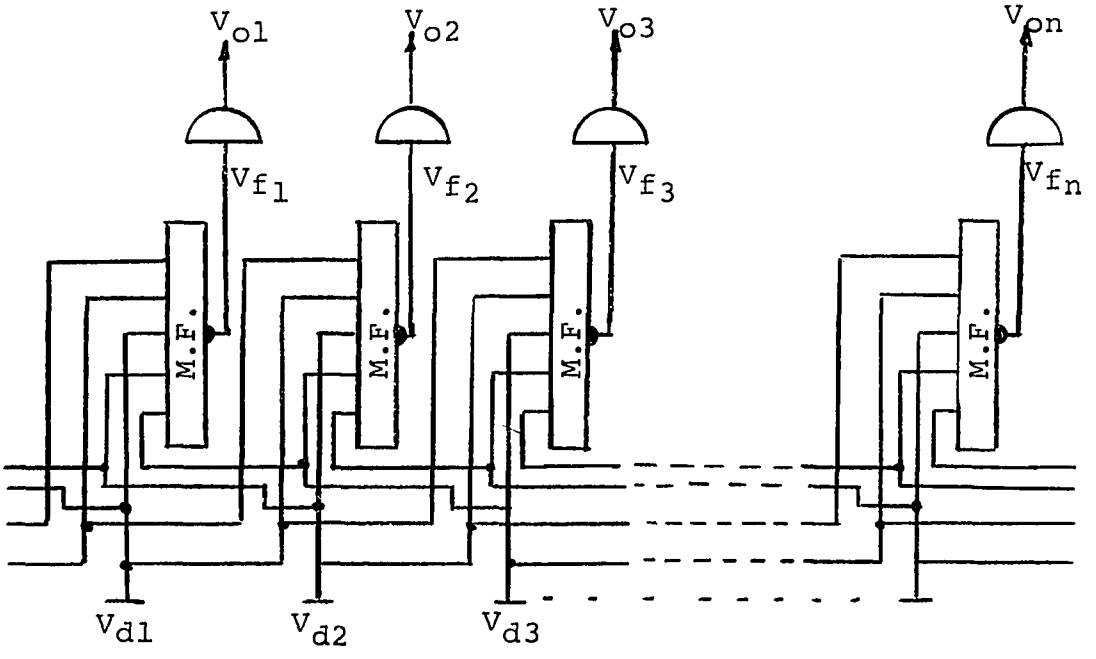


Figure 25b. Mean frequency centering network with parallel channels at both input and output. $V_{on} = 1$ if $n - 1/2 < V_{f_n} \leq n + 1/2$, otherwise $V_{on} = 0$.

threshold detection in the P.C.S. being considered, the output of each activated channel must have equal weight no matter what the level of its input. The estimated formant frequency provided by this process can thus be expressed as

$$f_E = \left[\sum_{n=1}^{N} n y_{in} / \sum_{n=1}^{N} y_{in} - 1 \right] \Delta f_x + f_1 \ , \qquad (39)$$

where each y_{in} factor corresponds to an element of the firing vector Y_i. y_{in} is equal to one if the nth channel of the ith firing pattern is activated and y_{in} is equal to zero if it is not. f_1 corresponds to the center frequency of the lowest frequency channel, and Δf_x is the center-to-center frequency spacing between channels.

In the case of corruption of a P.C.S. by intrinsic noise, the conditional probability of a firing pattern (Y_i) given a single exciting formant of frequency f_s, $P(Y_i/f_s)$, is given by equation (24) described in the analysis of the maximum likelihood centering process in section 4. The probability distribution of the estimated frequency provided by the system under consideration, which will be referred to as a mean frequency (M.F.) centering process, may be determined by summing the probabilities of those firing patterns which correspond to a particular f_E as described by equation (39), i.e.,

$$P(f_E/f_s) = \sum P(Y_i/f_s) \text{ over all } Y_i \text{ whose mean value corresponds to } f_E. \qquad (40)$$

The $P(F_E/f_s)$ distribution of a 100 channel P.C.S. utilizing the M.F. centering algorithm for an exciting formant of cosine squared spectral shape and square filter bandpass characteristics has been evaluated. In carrying out this evaluation, possible values of input frequency were assumed to exist only at channel center frequencies or frequencies midway between channel centers. When the estimated frequency (f_E) provided by a given firing pattern did not correspond exactly to a possible input frequency, it was rounded off to the nearest allowable input frequency. Details of the techniques used in the above calculations are presented in Appendix A. In Figure 26 the resultant relationship between $P(f_E/f_s)$ and Z is shown for a 100 channel P.C.S. with a false alarm rate of .02 and overlap factors of 3, 5, and 7. In Figure 27 σ_f, as a function of Z, is presented. Like the estimate of formant frequency provided by the median channel process, the $P(f_E/f_s)$ distribution and σ_f provided by the mean frequency process is very close to that provided by the maximum likelihood process.

The probability of correct detection achievable by the mean frequency centering process was also calculated. Equation (28) was again used for these computations. For the illustrative 100 channel system, $P(CD/S_1)$ was determined to vary with Z, ΔF_s, and M as shown in Figure 28. $P(CD/S_1)$ does not exhibit the large fluctuation with channel SNR produced by the median channel process, but displays an

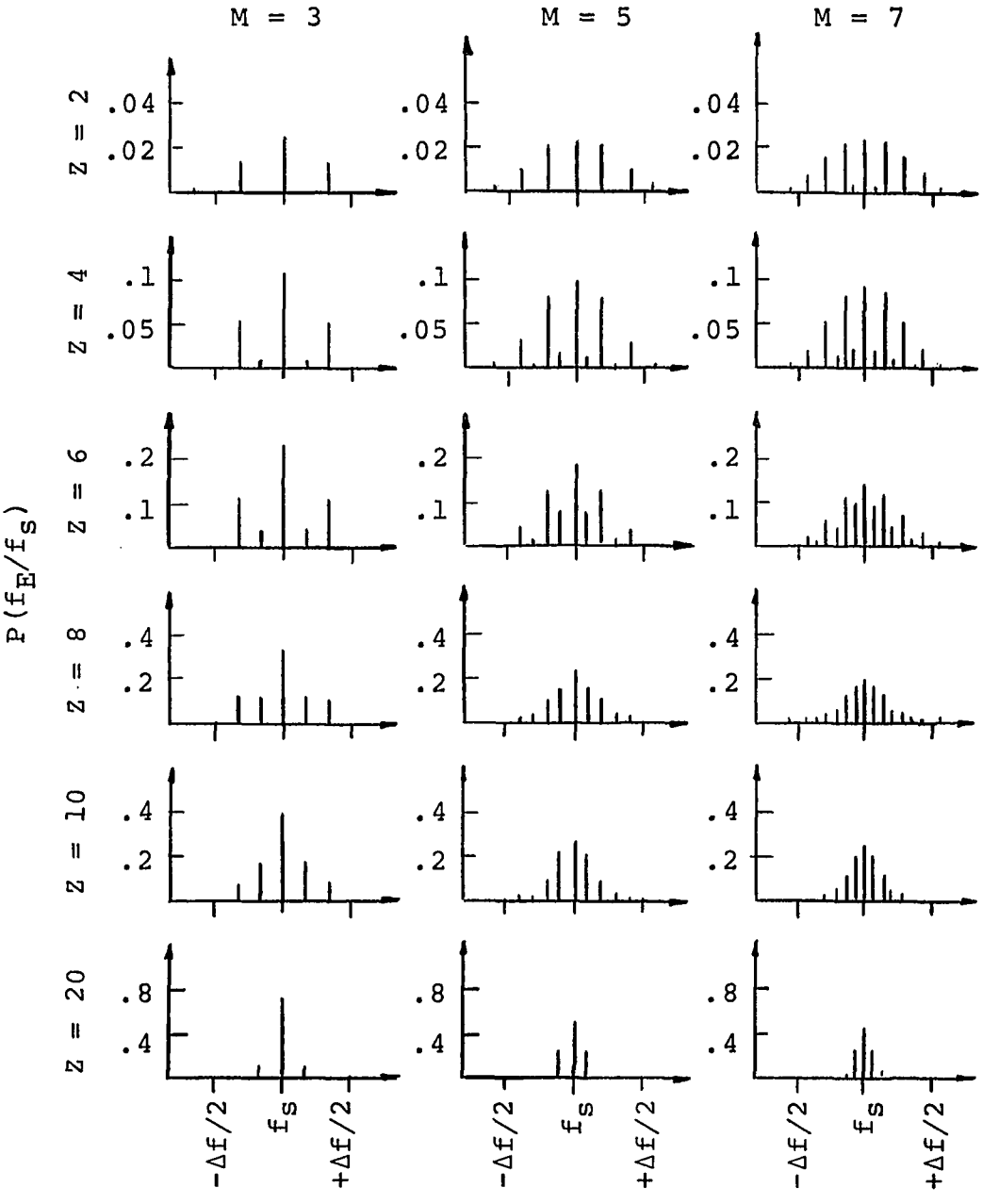


Figure 26. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with mean frequency centering and a false alarm rate of .02 in the presence of intrinsic noise.

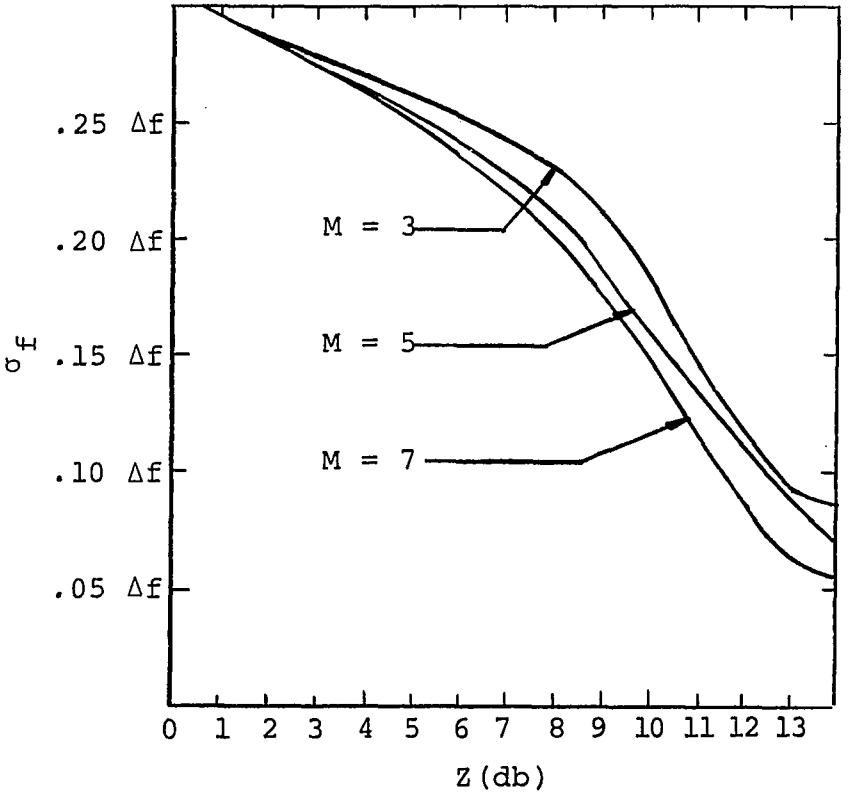


Figure 27. Standard deviation of the estimated frequency (given the detection of the signal) for a 100 channel P.C.S. with mean frequency centering and a false alarm rate of .02 in the presence of intrinsic noise as a function of channel SNR and overlap factor.

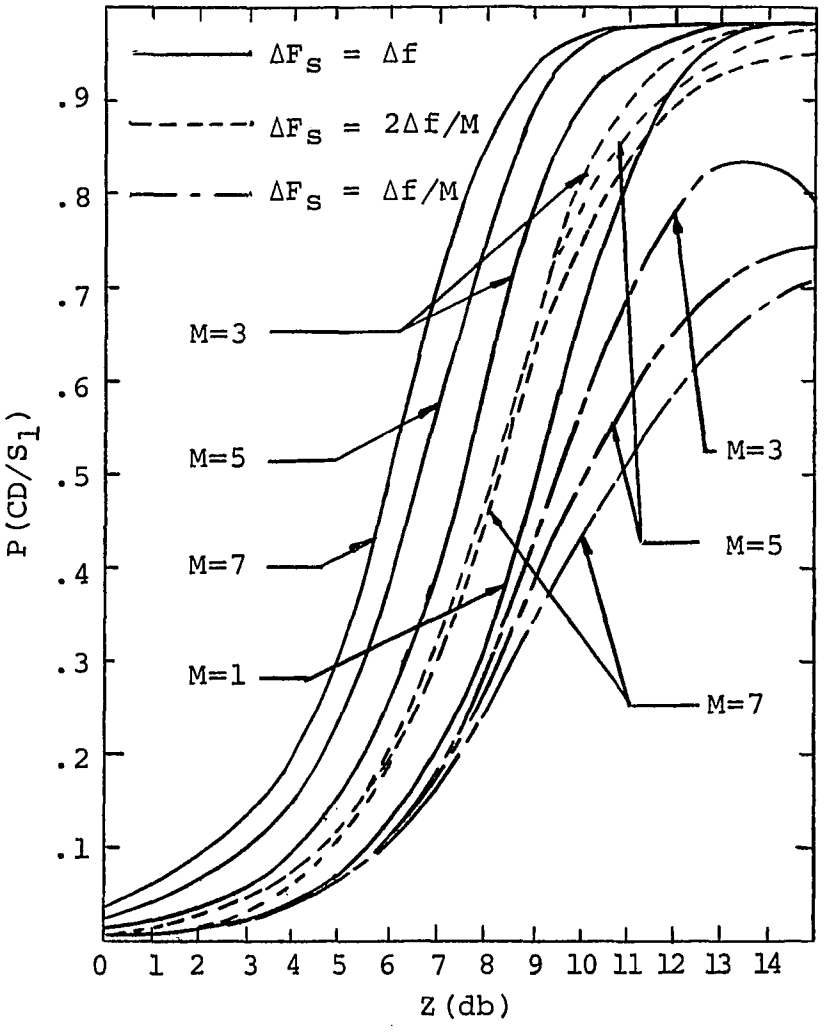


Figure 28. Probability of correct detection of a single formant provided by a 100 channel P.C.S. with mean frequency centering, intrinsic noise, and .02 false alarm rate as a function of channel SNR, overlap factor, and minimum frequency separation between individual formants.

almost monotonic increase with respect to increasing Z. This performance is almost identical to that displayed by the maximum likelihood process. Yet, in the case of the mean frequency process, this performance can be obtained at a considerable saving in system hardware. The similarity in performance provided by the mean frequency and maximum likelihood processes is partially due to the low false alarm rate assumed. In general, the lower the false alarm rate, the closer the mean frequency process will approach ideal performance.

In the case of corruption of a P.C.S. by extrinsic noise, it is not practical to evaluate the $P(Y_i/f_s)$ probabilities by means of a direct approach. The complications created by the introduction of extrinsic noise were discussed in section 4. To evaluate the performance of the mean frequency centering process in the presence of extrinsic noise, an approach similar to that used in evaluating the other centering processes in the presence of extrinsic noise may be adopted. The $P(Y_i/f_s)$ probabilities may be split into two terms [equation (31)], and a Monte Carlo simulation used to evaluate the first term (P_{IS}). The other term (PQ) will be identical to the PQ term calculated for the previous centering processes, as its value does not depend on the type of centering process employed but only on the false alarm rate. For further details on the evaluation of this term, see equations (32), (33), and Appendix C. The $P(f_E/f_s)$ distribution of a 100 channel

P.C.S. operating under conditions identical to those assumed previously except for the introduction of extrinsic noise has been evaluated by these means. Details of the Monte Carlo simulation are also presented in Appendix C. The resultant relationships are shown in Figures 29 and 30. Five thousand samples were taken for each input condition considered. Figure 29 shows the relationship between estimated frequency and channel SNR. In Figure 30 the relationship between the standard deviation of the estimated frequency and Z is shown. The $P(CD/S_1)$ characteristics were also calculated [see equation (28)]. The relationship between $P(CD/S_1)$ and Z, ΔF_s, and M is shown in Figure 31. A comparison of these last three figures with Figures 14, 15, 16, 22, 23, and 24 shows that in the presence of extrinsic noise the performance of the mean frequency centering process is even closer to that of the maximum likelihood process than that of the median channel centering process.

One of the difficulties in applying the mean frequency centering process as described in Figure 25a is its lack of parallel outputs. The other centering systems considered have parallel channels at both input and output. Neither can the mean frequency process in its present form separate two or more simultaneous formants. The system as shown will always find the center of mass of the activated channels no matter how many simultaneous formants are present in the excitations. Both these limitations may be

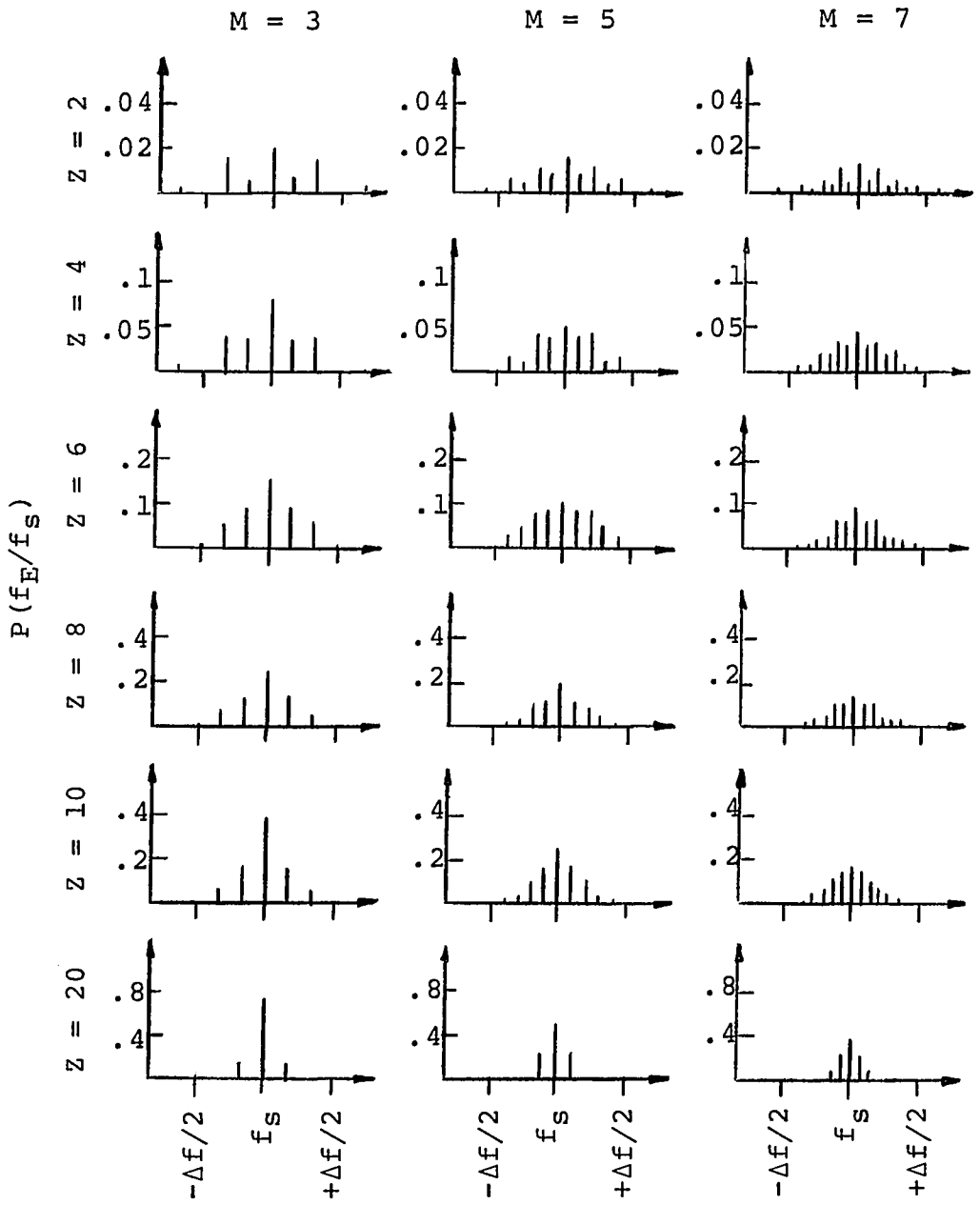


Figure 29. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with mean frequency centering and a false alarm rate of .02 in the presence of extrinsic noise.

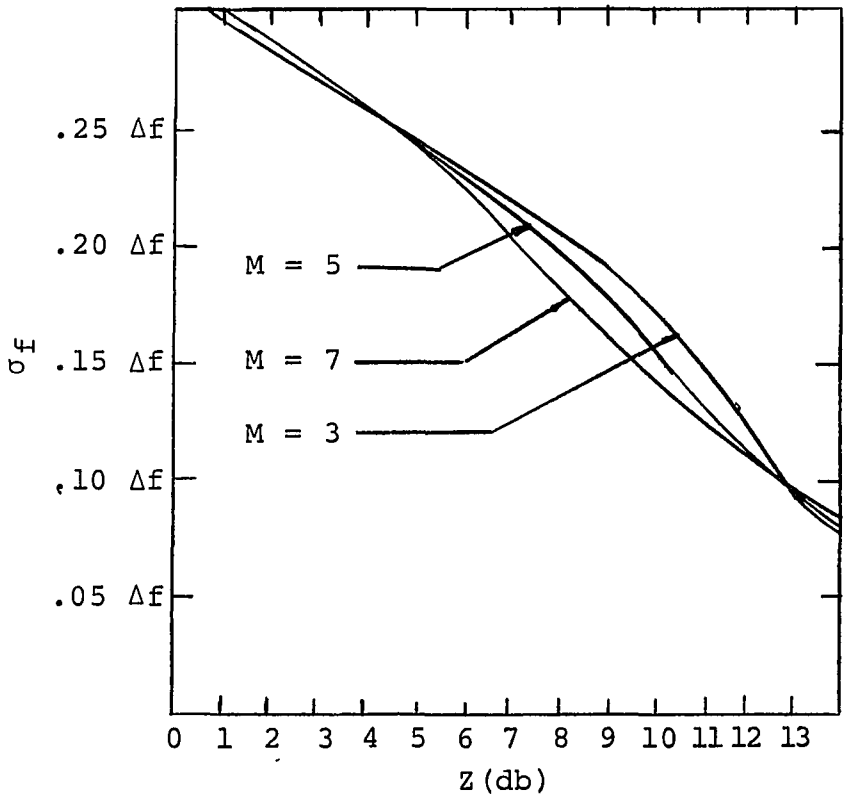


Figure 30. Standard deviation of the estimated frequency (given detection of the signal) for 100 channel P.C.S. with mean frequency and a false alarm rate of .02 in the presence of extrinsic noise as function of channel SNR and overlap factor.

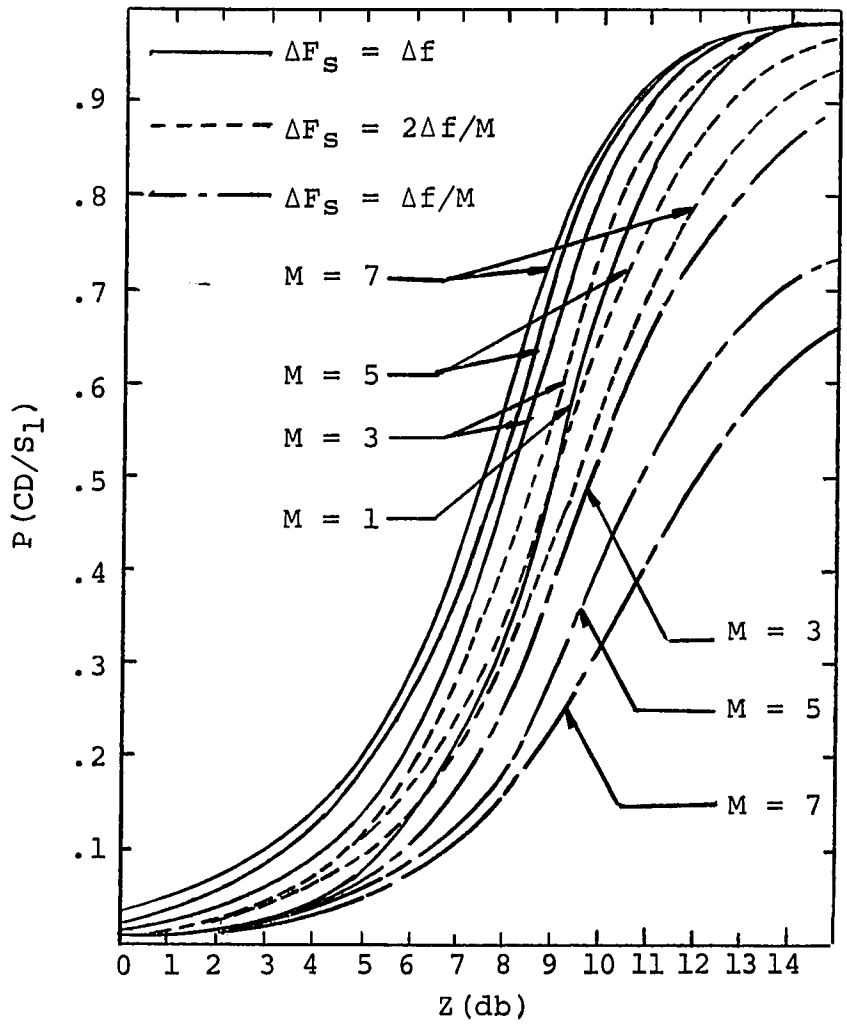


Figure 31. Probability of correctly detecting a single formant in extrinsic noise as a function of channel SNR, overlap factor, and minimum frequency separation between individual formants for a 100 channel P.C.S. with mean frequency centering and .02 false alarm rate.

circumvented by taking the center of mass (first moment) repeatedly over a limited number of channels (2N'+1) as illustrated in Figure 25b. An output will appear at any one of the many local centering processes when the output of that process

$$V_{f_n} = \sum_{n'=n-N'}^{n+N'} n' \, y_{in'} \Big/ \sum_{n'=n-N'}^{n+N'} y_{in'} \qquad (41)$$

is

$$n - 1/2 < V_{f_n} \leq n + 1/2 \ , \qquad (42)$$

where f_n is the center frequency of the center channel of the local set of channels (n-N' through n+N'), over which the center of mass is being calculated.

If the allowable range of V_f is enlarged, such that an output will occur in the nth channel, when

$$n - 1 < V_{f_n} \leq n + 1 \ , \qquad (43)$$

formants can be located at the center of an input channel or midway between channel centers. When two adjacent local centering processes are activated, a formant is assumed to be located midway between their respective channel centers.[2]

To eliminate the possibility that the above system may incorrectly interpret a single exciting formant as two or more formants, it is necessary that

---

[2]A single layer M.C. (Figure 18) following the M.F. process of Figure 25b might be used for this purpose.

$$2N' + 1 \geq 2NC - 1 \ . \tag{44}$$

In this case the spacing (ΔF_{ss}) between two simultaneous spectral peaks required for their resolution as individual formants (without the possible introduction of spurious intermediate formants) becomes:

$$\Delta F_{ss} \geq (3NC - 2)\Delta f_x = (6 - 5/M)\Delta f \tag{45}$$

These limiting conditions are illustrated in Figure 32 for a P.C.S. with an overlap factor of 3. From equations (34) and (45), it can be seen that the frequency resolution obtainable by the mean frequency centering system (Figure 25b) will be approximately 1/3 coarser ($NC >> 1$) than that obtainable by the maximum likelihood process.

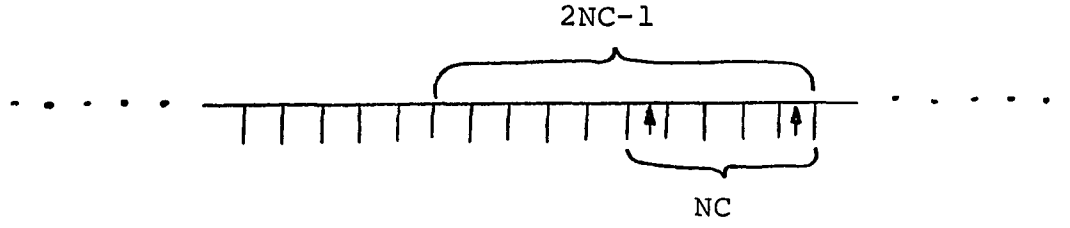


Figure 32a. The diagram shows the minimum number of channels over which the first moment must be calculated in order for the M.F. centering process described in Figure 25b to properly interpret the firing pattern provided by a single formant. If the first moment is calculated over less than 2NC-1 channel, the firing pattern in the above example would be interpreted as representing three separate formants--two spurious formants located at the activated channels (arrows) and a third formant (the correct response not shown in the figure) located midway between the activated channels.

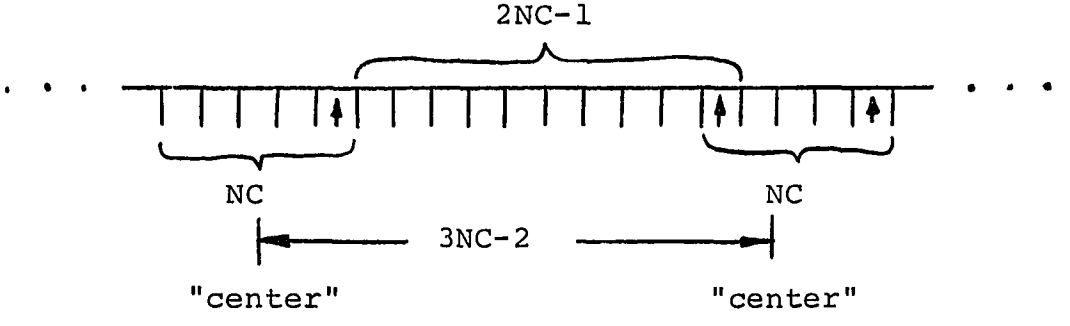


Figure 32b. The diagram shows the minimum spacing between simultaneous formants (ΔF_{SS}) necessary for their unambiguous resolution by the M.F. centering process of Figure 25b.

6. CENTERING BY LATERAL INHIBITION

Lateral inhibition (or more generally lateral interaction) occurs in a P.C.S. when the input to or output from one channel can affect the response of other channels "laterally" adjacent to it [45]. Lateral inhibitory phenomena have been observed in a number of biological systems and are believed by some researchers to be responsible for the "centering-like" characteristics displayed by both the eye and the ear [19, 48].

The output characteristics of a generalized "N" channel system employing lateral inhibition can be described by 2N dependent equations of the form [20]:

$$V_{on} = g_n \left(\sum_{m=1}^{N} k_{nm} V_{im} , \sum_{m=1}^{N} [j_{nm} V_{im} + \ell_{nm} V_{om}] \right) , \qquad (46)$$

where g_n represents the functional relationship between the output of the nth channel (V_{on}), and its inputs which, in formula (44), have been divided into two categories: excitation type inputs (first summation) and inhibitory inputs (second summation). V_{in} is the input to the nth channel and k_{nm}, j_{nm}, and ℓ_{nm} are coefficients establishing the weight to be assigned to the individual inputs. For the P.C.S. being considered in this thesis, g_n represents the

characteristics of a threshold detector for which

$$V_{on} = 1 \text{ if } \sum_{m=1}^{N} k_{nm} V_{im} > V_{th} + \sum_{m=1}^{N} (j_{nm}V_{im} + \ell_{mn}V_{om})$$

and

$$V_{on} = 0 \text{ if } \sum_{m=1}^{N} k_{nm} V_{im} \leq V_{th} + \sum_{m=1}^{N} (j_{nm}V_{im} + \ell_{nm}V_{om}) . \quad (47)$$

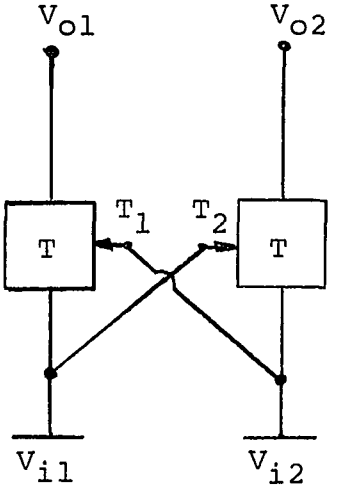
All lateral inhibitory (L.H.) processes to be considered in the following discussion will be assumed to have nonlinearities of this form.[1]

Lateral inhibitory processes are generall divided into two classes: recurrent and nonrecurrent processes [45]. Recurrent lateral inhibition occurs when the lateral control signal is obtained from the output of the nonlinear element (threshold detector); lateral inhibition of the nonrecurrent form occurs when the lateral control signal is derived before (i.e., from the input of) the nonlinear element. Figure 33 shows two different 2 channel P.C.S. with lateral inhibition along with the equations which describe their output characteristics. The system shown in part a employs a nonrecurrent process, while the system in part b employs a recurrent process.

By proper choice of the values of k_{nm}, j_{nm}, and ℓ_{nm}, a L.H. network can be made to produce centering action. A simple nonrecurrent network (all $\ell_{nm} = 0$) is shown in

---

[1]The "hard limiter" characteristic of equation (47) is a reasonable model of the response of many neural fibers. This relationship is discussed in more detail in sections 7 and 8.

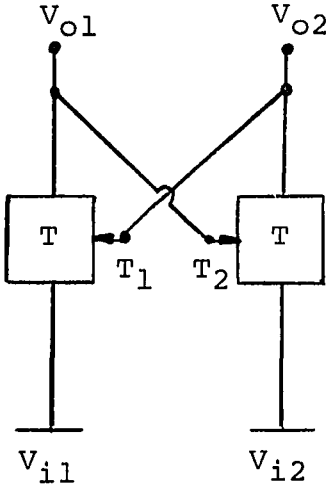


$$\begin{cases} V_{o1} = g(V_{i1}, V_{i2}) \\ V_{o2} = g(V_{i2}, V_{i1}) \end{cases}$$

In general
detector characteristic
$V_o = g(V_i, T_n)$
for
threshold detector:

$$\begin{cases} V_o = 1 & \text{if} \quad V_i > V_{th} + f(T) \\ V_o = 0 & \text{if} \quad V_i \leq V_{th} + f(T) \end{cases}$$

Figure 33a. Two channel nonreccurrent lateral inhibitory process.



$$\begin{cases} V_{o1} = g(v_{i1}, V_{o2}) = g[V_{i1}, g(V_{i2}, V_{o1})] \\ V_{o2} = g(V_{i2}, V_{o1}) = g[V_{i2}, g(V_{i1}, V_{o2})] \end{cases}$$

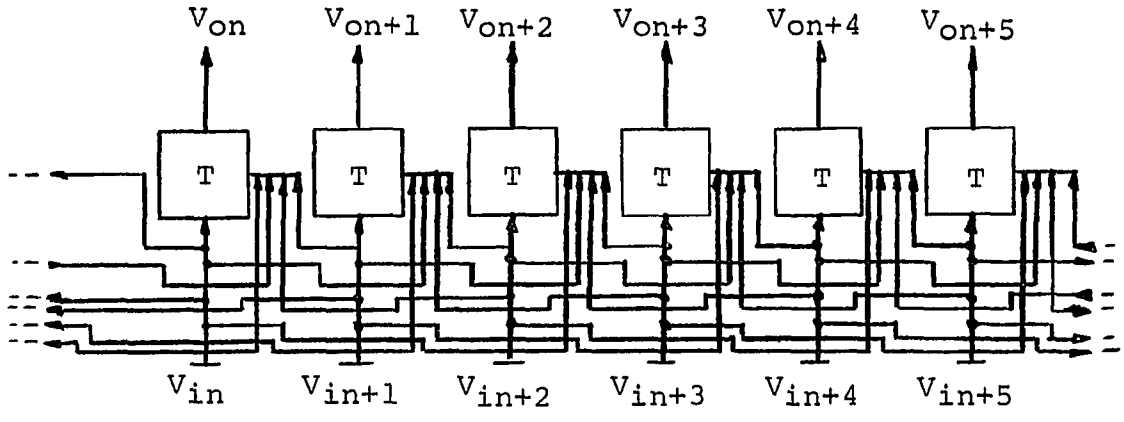
Figure 33b. Two channel recurrent lateral inhibitory process.

Figure 34. In Figure 34 and in other inhibitory networks drawn in this thesis, the connections are assumed to be unidirectional proceeding in the direction indicated by the arrowheads (diodes have been omitted to simplify the circuits). Here the input to each channel inhibits the two adjacent channels on either side.

$$j_{nm} = 0 \text{ for } m > n + 2, \text{ or } m < n - 2, \text{ or } m = n;$$
$$\text{otherwise } j_{nm} = .25 \ . \qquad (48)$$

Directly below this network is drawn a curve (heavy line) representing (as an example) a possible distribution of exciting energy as a function of frequency. The actual distribution of exciting energy to the various inputs of the network, being a function of the individual channel bandpass characteristics and the spectral shape of the exciting formants, is discontinuous. However, for convenience of visualization, it is presented here as a series of dots connected by solid lines. On the same diagram is shown the assumed noninhibitory threshold of the threshold detectors ($V_{th} = 1$) (dotted line) along with the resultant detector thresholds due to inhibition (thin line) and the resultant firing pattern (arrows shown below). From these curves it can be seen that the effect of the inhibition is to narrow the firing pattern (from 7 channels to 5 channels).

It would appear that, if lateral inhibition is to be used to sharpen the output firing pattern of a P.C.S., only



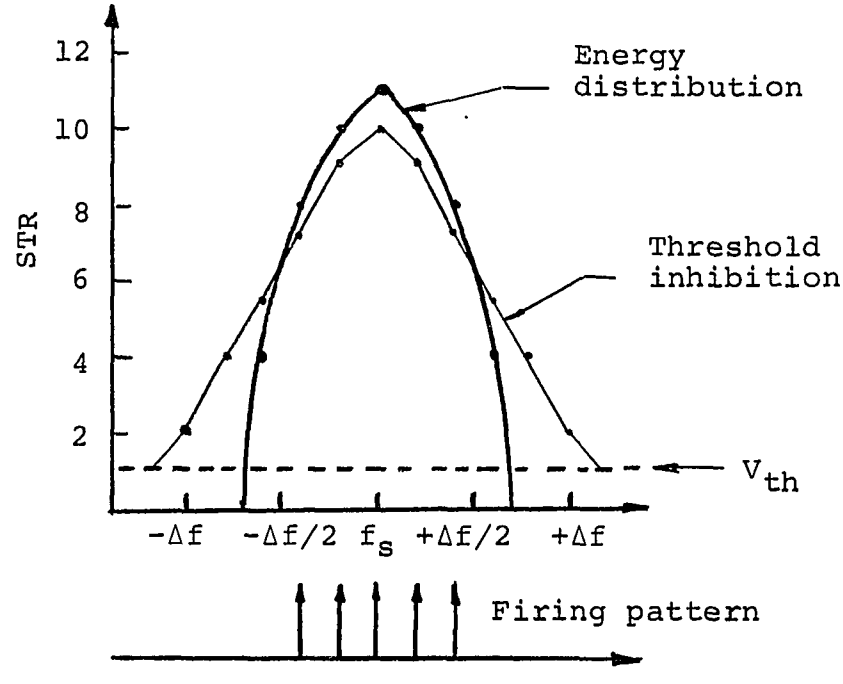


Figure 34. Sample nonrecurrent lateral inhibitory network and response characteristics

one input should excite each threshold detector, i.e., all $k_{nm} = 0$ for $m \neq n$.[2] In fact, addition of other excitatory inputs (from filters other than the one corresponding to the detector's own channel) can only broaden the distribution of the exciting energy. Summation of inhibitory inputs, on the other hand, tends to have the opposite effect. The wider the range of channels (within the range NC) over which the inhibitory inputs to a detector are summed, the shallower the slope of the threshold detection locus around a given center of excitation; and hence, the sharper the centering action of the process. However, when increasing the number of channels over which the inhibitory inputs are summed, one must be careful to scale the j_{nm} terms in such a manner that the system cannot be totally inhibited by the presence of a strong spectral peak.

Characteristics similar to those of nonrecurrent inhibitory processes can be produced by recurrent processes. Figure 35 shows the recurrent equivalent ($\ell_{nm} = 2$, or 0) of the nonrecurrent process of Figure 34. The output of each channel now inhibits the two adjacent channels on either side. With a recurrent process, since the lateral control signals depend on the outputs of the threshold detectors, a transient period must exist, during which a stable inhibitory threshold is established. For the system shown, the

---

[2]This conclusion is not necessarily valid in the case of detection in the presence of noise, as shall be indicated later.

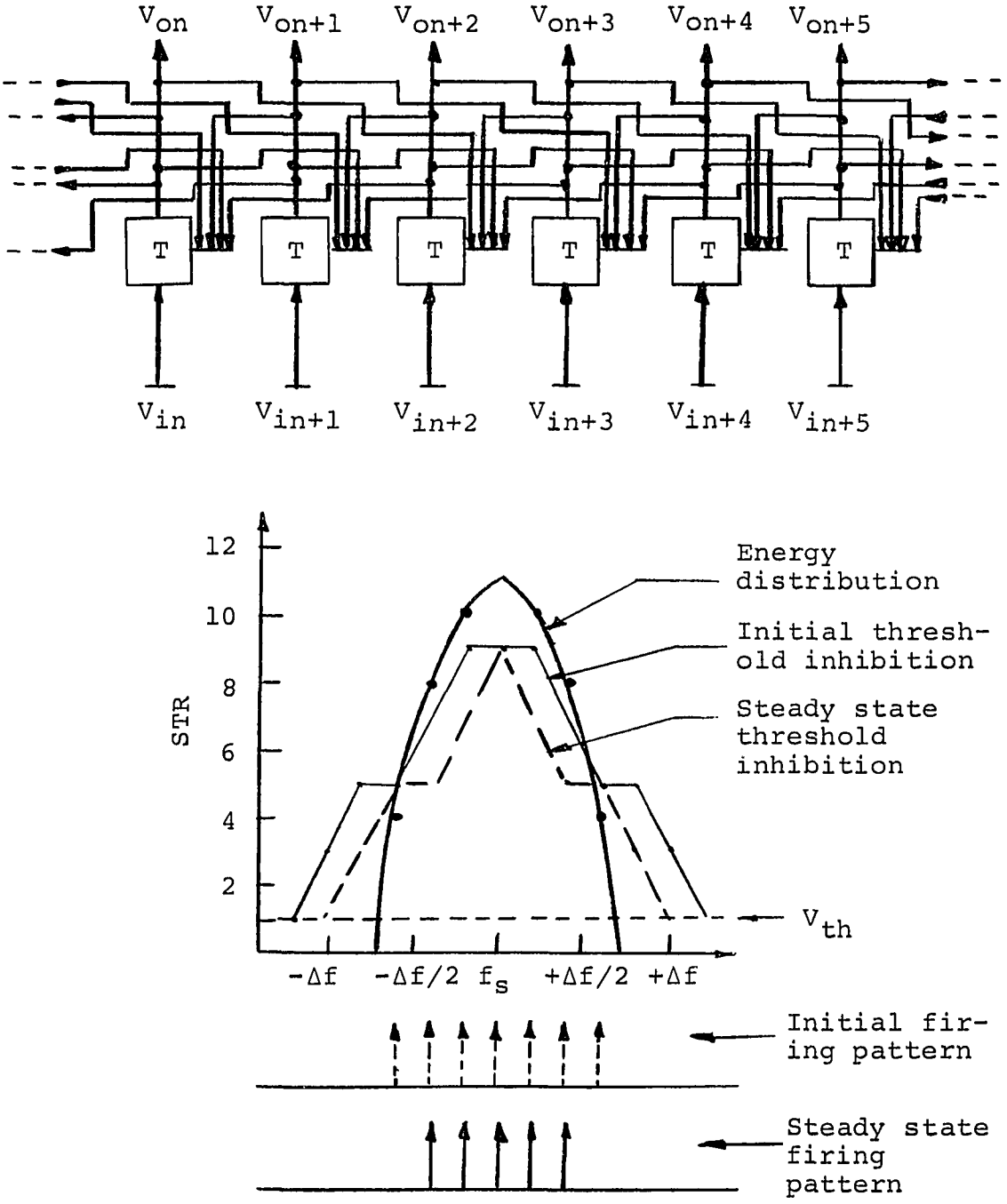


Figure 35. Sample recurrent lateral inhibitory network and response sequence.

sequence of events leading to this stable or steady state condition is as follows. Initially, the seven channels (which are excited beyond V_{th}) are activated--see initial firing pattern and response characteristics in lower half of Figure 35. The activation of these channels in turn raises the detector thresholds to the levels indicated by the thin line. This new threshold distribution allows the firing of only five channels, whose outputs in turn establish the detector thresholds shown by the thick dashed line. These threshold levels again allow the same five channels to fire. The output firing pattern does not change and, consequently, the threshold distribution does not change; therefore, a stable operating condition is reached. Whether steady state is ever reached (during the time period an exciting signal is present) and, if so, how fast it is reached depends on the inhibition time constant. In most cases, a very fast (instantaneous) rise time and a slow decay time of the inhibited threshold appears desirable.[3] In any case, the threshold inhibition time constant (τ_I) should not be longer than $1/\Delta f$, or the information capacity of the system will be impaired.

Just as in the case of a nonrecurrent process, an improper choice of the inhibitory coefficients of a recurrent process can result in the output being inhibited out

---

[3] A detailed study of the effect of different inhibitory time constants by means of analogue simulation is currently in progress.

of existence. In the case of a recurrent process, however, when such an extreme condition is created, or when the inhibiting process has progressed to the point where its centering ability is impaired, it is not possible for the system to reach a stable operating condition. The system then goes into an oscillatory mode in which the transient sequence (or some portion of it) is repeated. The significance of the excitation of an oscillatory mode depends on the time constant of the threshold inhibition. If τ_I is of the order of the formant time duration (ΔT), the initiation of an oscillatory sequence will be of no consequence as the excitation will have changed before the transient sequence can repeat itself. And even if $\tau_I << \Delta T$, so long as the rise time of the threshold inhibition is substantially shorter than the decay time, the output firing pattern will remain in its most sharpened form for the majority of the time the exciting formant is present.

Recurrent lateral inhibitory processes are favored in models of biological centering, because of their ability to produce disinhibition [45], i.e., the ability to reactivate a channel (A) which has been inhibited by a formant from another channel (B). The reactivation is obtained by the introduction of an additional formant. The introduction of a new formant can produce this effect by inhibiting the activity of channel B, and thus, terminating its inhibition of the first channel.

L.H. processes appear very attractive for practical

implementation of centering circuitry because of their structural simplicity and their repeated use of relatively simple elements. However, for a given centering process to be effective, it must be able to function in the presence of noise. The inhibited threshold levels of the L.H. processes considered thus far tend to be highest at the "center" of the excitation pattern. Such centering action has the effect of completely inhibiting firing patterns derived from formants of low exciting energy and, hence, does not provide desirable centering performance in the presence of noise. L.H. processes of this class will be referred to as M.I.C. (maximum inhibition at center) L.H. processes.

If, however, the inhibitory coefficients (j_{nm} and ℓ_{nm}) of a lateral inhibitory network are made to be positive constants for $m > n$, zero for $m = n$, and negative constants for $m < n$, and the individual inhibited threshold levels are assumed to be equal to

$$V_{Tn} = V_{th} + \left| \sum_{m=1}^{N} (j_{nm}V_{im} + \ell_{nm}V_{om}) \right| , \qquad (49)$$

then the inhibited threshold distribution will tend to its lowest levels at the "center" of excitation.[4] This class of L.H. process will be referred to as an L.I.C. (least

---

[4]An inhibited threshold distribution where the threshold inhibition tends to its lowest level at the center of excitation can be produced by a M.I.C. class L.H. process (at low levels of excitation), if activation of a given channel does not inhibit channels directly adjacent to it but only channels a number of channel widths away on either side.

inhibition at center) L.H. process.

Under these conditions the lateral inhibitory network of Figure 34, for example, would display for the same excitation distribution the inhibited threshold levels shown in Figure 36. The transient sequence of the recurrent inhibitory process described by Figure 36 ($|\ell_{nm}| = 4$ or 0) reaches steady state in only three steps, the firing pattern changing from 7 to 5 to 3 activated channels. As noted earlier, it is not necessarily detrimental to the centering performance of a lateral inhibitory process if a stable condition is not reached. By setting $|\ell_{nm}| = 5$ (for $\ell_{mn} \neq 0$) the above network would sharpen the output firing pattern to only one channel, but not reach a stable condition.

If the lateral control signal inhibits a larger number of adjacent channels, the shape of the detector threshold locus approximates more closely a triangle (rather than a trapezoid) and the centering action is intensified. The locus of inhibited threshold also spreads over a wider range of channels; the probability of an output firing pattern splitting up into a number of clusters of activated channels (particularly at the edges) is then decreased.

The above improvement in centering action (discrimination) is obtained at the expense of signal resolution. In general, when designing a centering network, a compromise must be made between signal resolution and signal

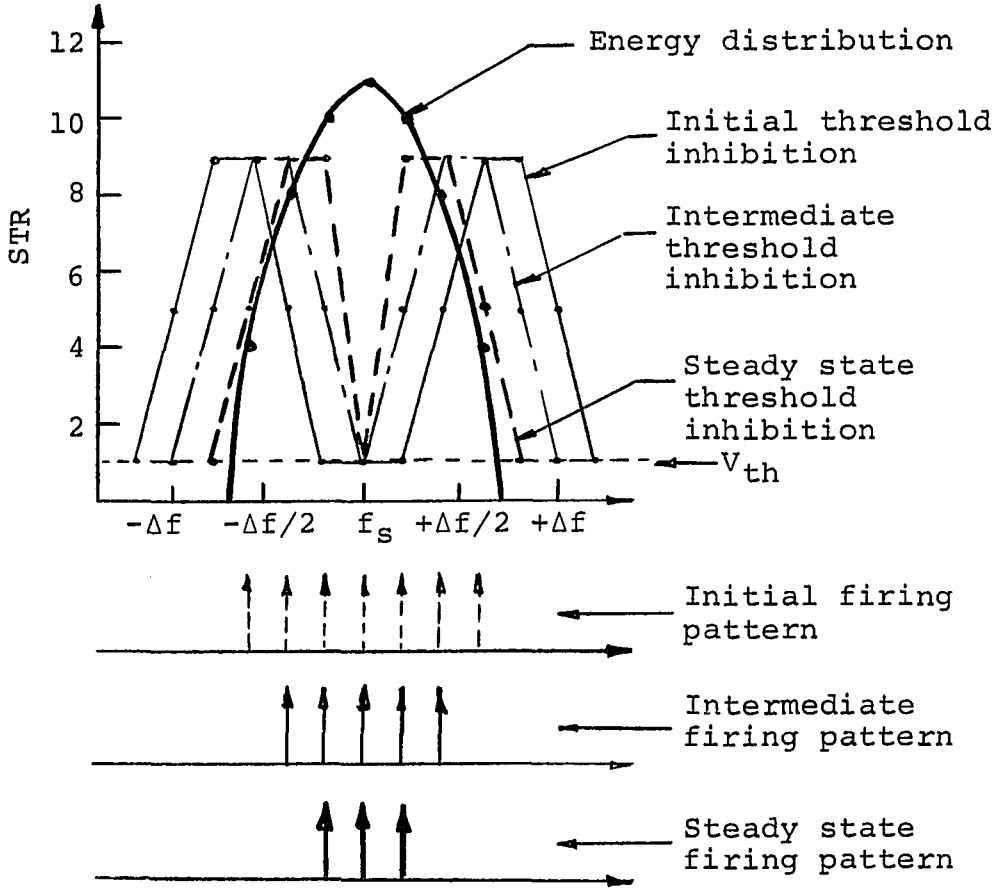


Figure 36. An example of a possible sequence of inhibited threshold levels and firing patterns generated by a recurrent L.I.C. lateral inhibitory network.

discrimination. A reasonable compromise in the case of lateral inhibition is to have the lateral control signal inhibit only those channels which share common spectra with the channel from which the control signal emanates. In this case two simultaneous formants may be separated by

$$\Delta F_{ss} = (4 - 2/M)\Delta f \tag{50}$$

without degradation to their individual inhibition.

In the case of corruption of P.C.S. by intrinsic noise, the conditional probability of a firing pattern, Y_i (without inhibition), given a single exciting formant f_s, $P(Y_i/f_s)$, is given by equation (24) (section 4). The conditional probability of the resultant firing pattern (Y_i*) after the inhibited threshold is established (same noise condition), $P(Y_i^*/f_s)$, may be expressed in terms of $P(Y_i/f_s)$ as follows:

$$P(Y_i^*/f_s) = \sum_{Y_i=1}^{(2^N-1)} P(Y_i^*/Y_i)\ P(Y_i/f_s) \tag{51}$$

but the probability of the inhibited firing pattern given the noninhibited firing pattern is

$$P(Y_i^*/Y_i) = \sum_{n=1}^{N} \{P(y_{in}^*\ ,\ y_{in})\} \Big/ P(Y_i/f_s)\ . \tag{52}$$

Each y_{in} (or y_{in}^*) factor corresponds to an element of the firing vector Y_i (or Y_i^*). y_{in} (or y_{in}^*) is equal to one if the nth channel of the ith firing pattern is activated

and y_{in} (or y_{in}^*) is equal to zero if it is not.

$$P(y_{in}^* , y_{in}) = P(y_{in}^*/f_s) \quad \text{if } y^*_{in} = y_{in} = 1$$

$$P(y_{in}^* , y_{in}) = 1 - P(y_{in}/f_s) \quad \text{if } y_{in}^* = y_{in} = 0$$

$$P(y_{in}^* , y_{in}) = P(y_{in}/f_s) - P(y_{in}^*/f_s) \quad \text{if } y_{in}^* = 0 , y_{in} = 1$$

and

$$P(y_{in}^* , y_{in}) = 0 \quad \text{if } y_{in}^* = 1 , y_{in} = 0 \qquad (53)$$

The $P(y_{in}/f_s)$ and $P(y_{in}^*/f_s)$ probabilities are described in equation (25), section 4, in terms of the Q function. In the case of $P(y_{in}^*/f_s)$, the normalized threshold (b) in equation (25) is not a constant but varies as a function of the threshold inhibition [see equation (49)].

Since the centering action of a lateral inhibitory network does not always sharpen its output firing pattern down to a single channel, some form of "secondary" decision process must be assumed to follow a lateral inhibitory network if its performance is to be quantitatively evaluated. The presence of such a secondary decision process seems highly likely in the instance of biological processing [48], and may be desirable in the practical implementation of centering circuitry. In the analysis to be performed in this thesis a median channel centering process shall be assumed to follow the lateral inhibition. Although the median channel process only represents one of a number of possible alternatives, it does represent a practical

alternative.[5] The probability distribution of the estimated frequency of a lateral inhibitory network may then be determined by summing the probabilities of those inhibited firing patterns which correspond to a particular f_E as decided by the median channel process in terms of equations (37) and (38), i.e.,

$$P(f_E/f_S) = \sum P(Y_i^*/f_S) \text{ over all } Y_i \text{ corresponding to } f_E \; . \tag{54}$$

The $P(f_E/f_S)$ distribution of a 100 channel P.C.S. employing lateral inhibition for an exciting formant of cosine squared spectral shape and square filter bandpass characteristics has been investigated. In Figure 37 the relationship between $P(f_E/f_S)$ and channel SNR (Z) is shown for a L.I.C. lateral inhibitory process (with $|\ell_{mn}| = 1$ or 0) in which the lateral control signals inhibit those channels which share common spectra with the channels from which they emanate. The P.C.S. system was assumed to have a fixed false alarm rate of .02 and overlap factors (M) of 3, 5, and 7. In Figure 38, σ_f, the standard deviation of f_E, is presented as a function of Z. The centering performance displayed by the L.H. process in these figures is

---

[5]When a L.H. process proceeds a M.C. centering process, only a very minimal number of layers of logic gates (Figure 18) are needed to implement the M.C. process. The combining of L.H. and M.C. centering processes can thus result in a considerable savings in overall system hardware, when compared to the equivalent (independent) process.

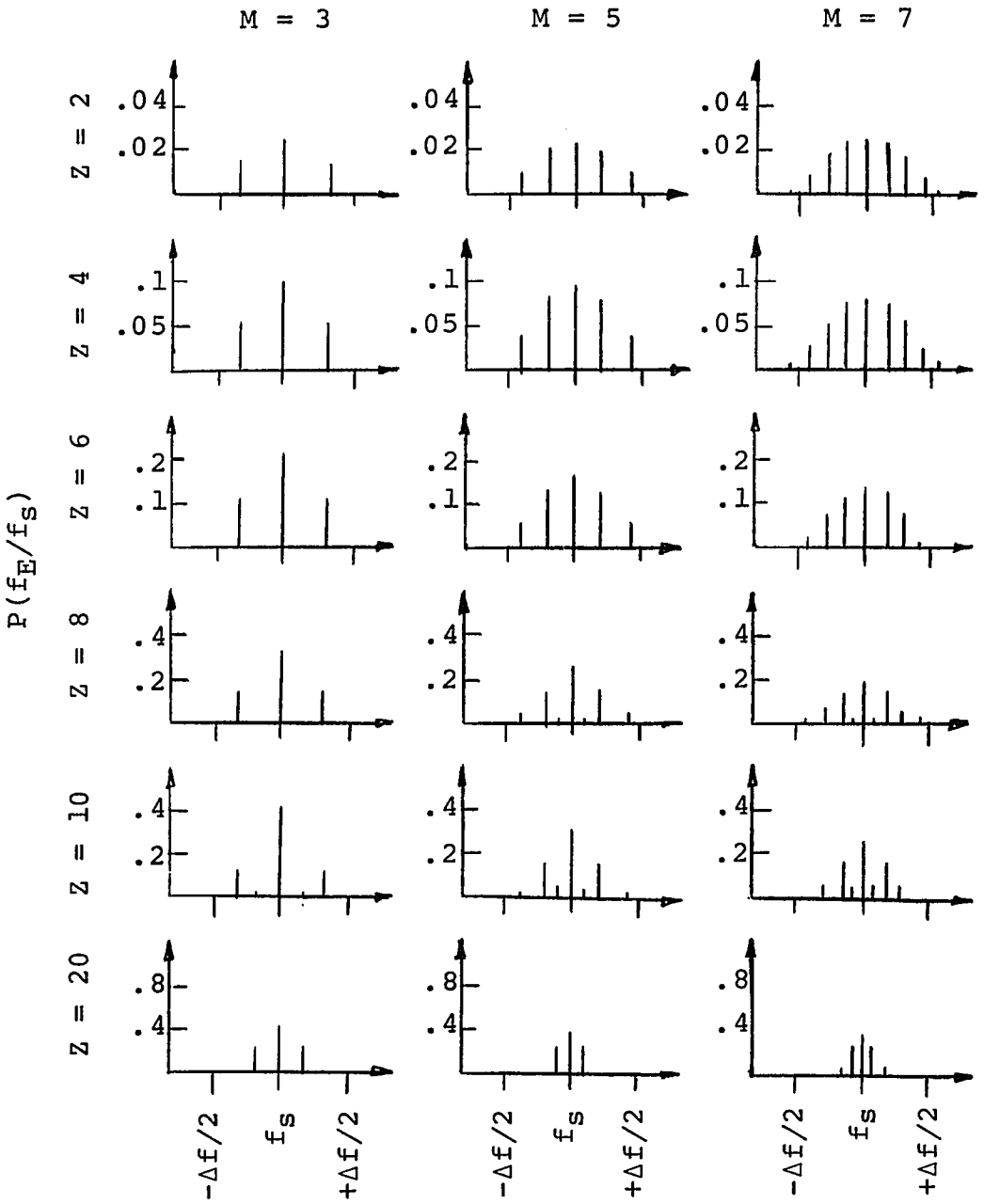


Figure 37. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of intrinsic noise.

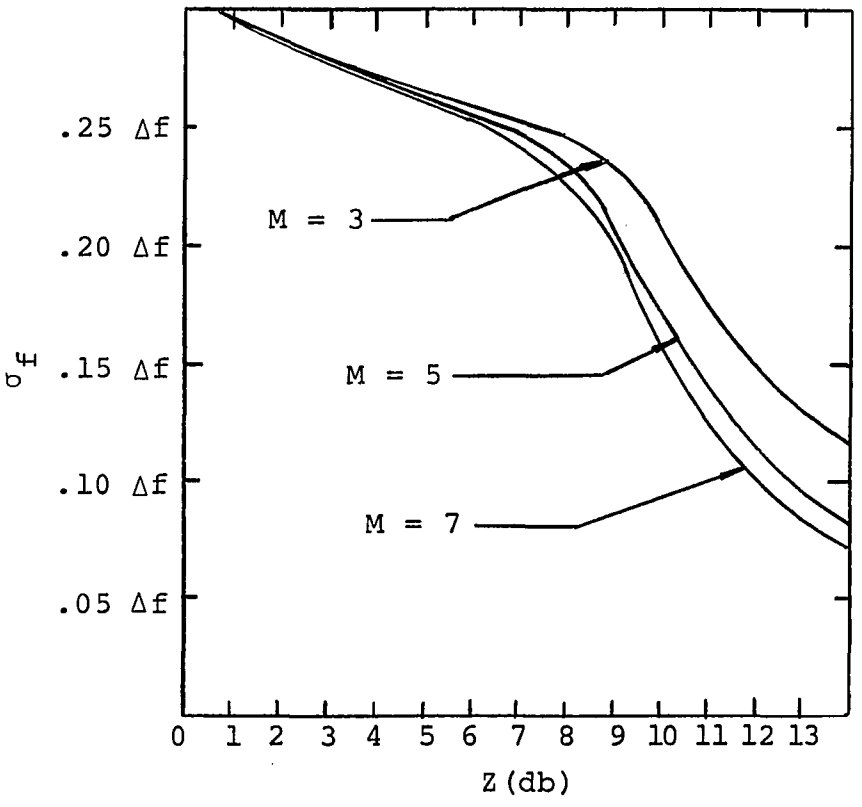


Figure 38. Standard deviation of the estimated frequency (given detection of the signal) provided by a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of intrinsic noise.

comparable to the centering processes studied earlier. Details of the techniques used in the above calculations are presented in Appendix A.

The probability of correct detection achievable by a lateral inhibitory process was also calculated. Equation (28) was again used for these computations. For the illustrative 100 channel system, $P(CD/S_1)$ was determined to vary with Z, ΔF_s, and M as shown in Figure 39. The addition of the lateral inhibition virtually eliminates the fluctuation with channel SNR displayed by the median channel process alone and (depending on ℓ_{mn}) generally increases the overall probability of correct detection in comparison to the M.C. process acting alone.

In the case of extrinsic noise, it is not practical to evaluate $P(Y_i/f_s)$ and $P(Y_i^*/f_s)$ probabilities directly. The problems created by the introduction of extrinsic noise have been discussed earlier. A Monte Carlo simulation can again be used to evaluate the $P(Y_i/f_s)$ probabilities (see section 4). The $P(Y_i^*/f_s)$ probabilities may be evaluated by incorporating into the simulation the effects of the lateral inhibition. Details of this modification in the basic Monte Carlo simulation computer program for the L.H. process are presented in Appendixes A and C.

The $P(f_E/f_s)$ distribution of a 100 channel P.C.S. system operating under conditions identical to those assumed previously except for the introduction of extrinsic noise has been evaluated by the above means. The resultant

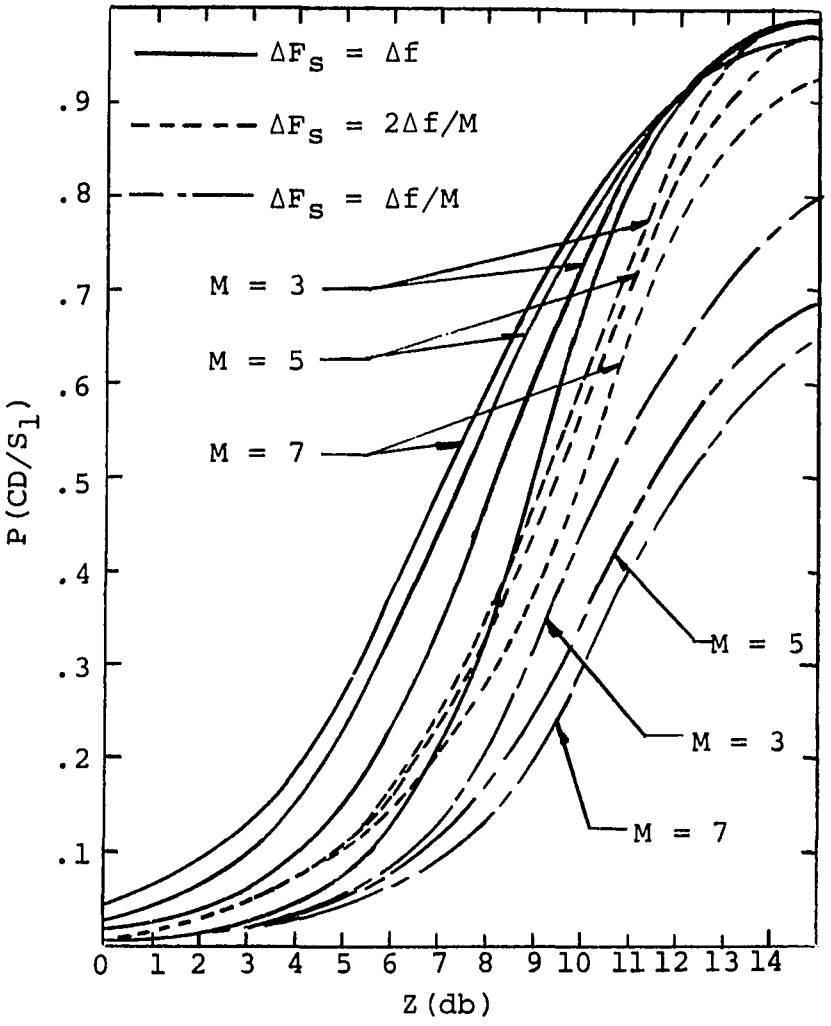


Figure 39. Probability of correctly detecting a formant provided by a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of intrinsic noise.

relationships are shown in Figures 40 and 41. Figure 40 shows the relationship between $P(f_E/f_S)$ and Z, while in Figure 41 the relationship between the standard deviation of f_E and Z is shown. The $P(CD/S_1)$ characteristics were also calculated. The relationship between $P(CD/S_1)$ and Z, ΔF_S, and M is shown in Figure 42. The centering performance in the presence of extrinsic noise, illustrated in these three figures, is in general agreement with the performance exhibited by L.H. processes in the presence of intrinsic noise, i.e., while not performing as well as the maximum likelihood process, the L.H. process can substantially improve upon the performance of the M.C. process.[6]

Placing the lowest levels of the inhibited threshold distribution at the "center" of excitation greatly improves the centering performance of lateral inhibitory networks in the presence of noise. The centering performance of lateral inhibitory processes in the presence of noise is still, however, inferior to that of the maximum likelihood or mean frequency centering processes, since a lateral inhibitory process cannot fuse two separate clusters of activated channels into one peak of activity. This handicap may be partially overcome (in the case of extrinsic noise) by the intentional cross-coupling of exciting energy between

---

[6]The improvement in overall system performance, provided by proceeding a M.C. process by a L.H. process, is dependent upon the (non-zero) inhibitory coefficients ℓ_{nm}. In general, the optimum value of $|\ell_{mn}|$ varies with input channel SNR.

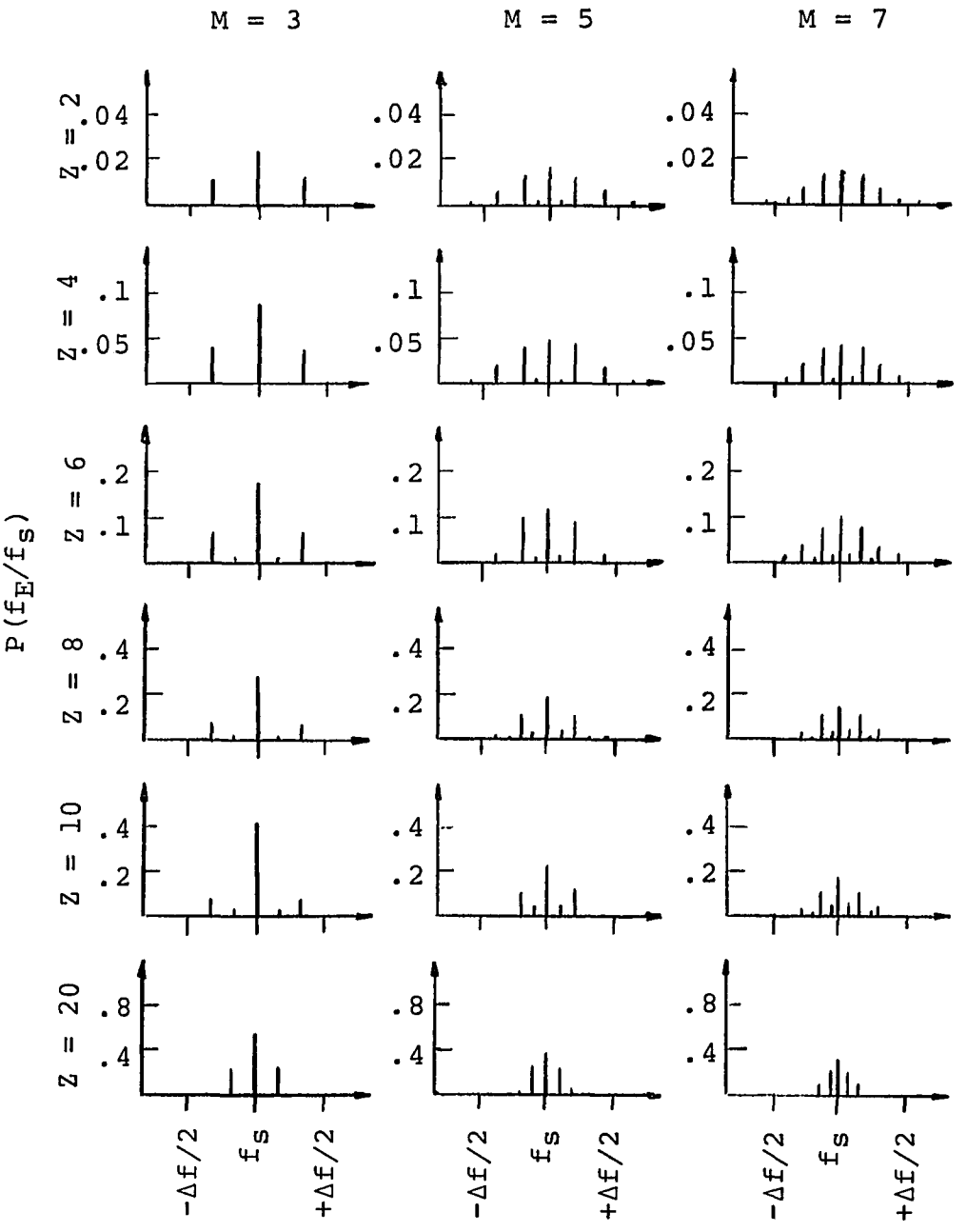


Figure 40. Probability distribution of estimated frequency provided by a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of extrinsic noise.

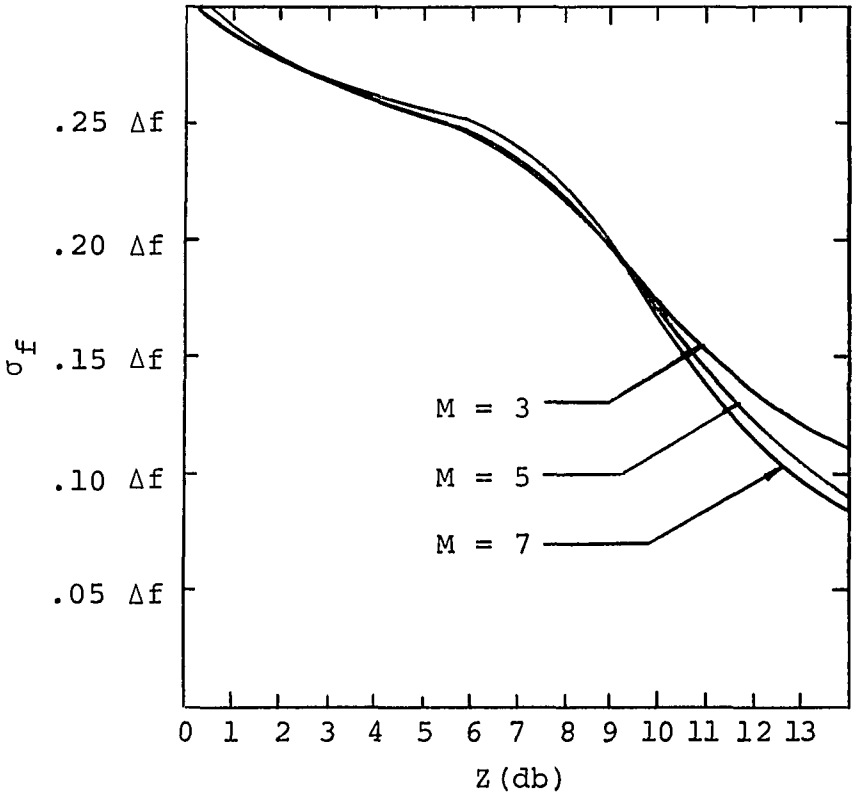


Figure 41. Standard deviation of the estimated frequency (given detection of the signal) for a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of extrinsic noise as a function of channel SNR and overlap factor.

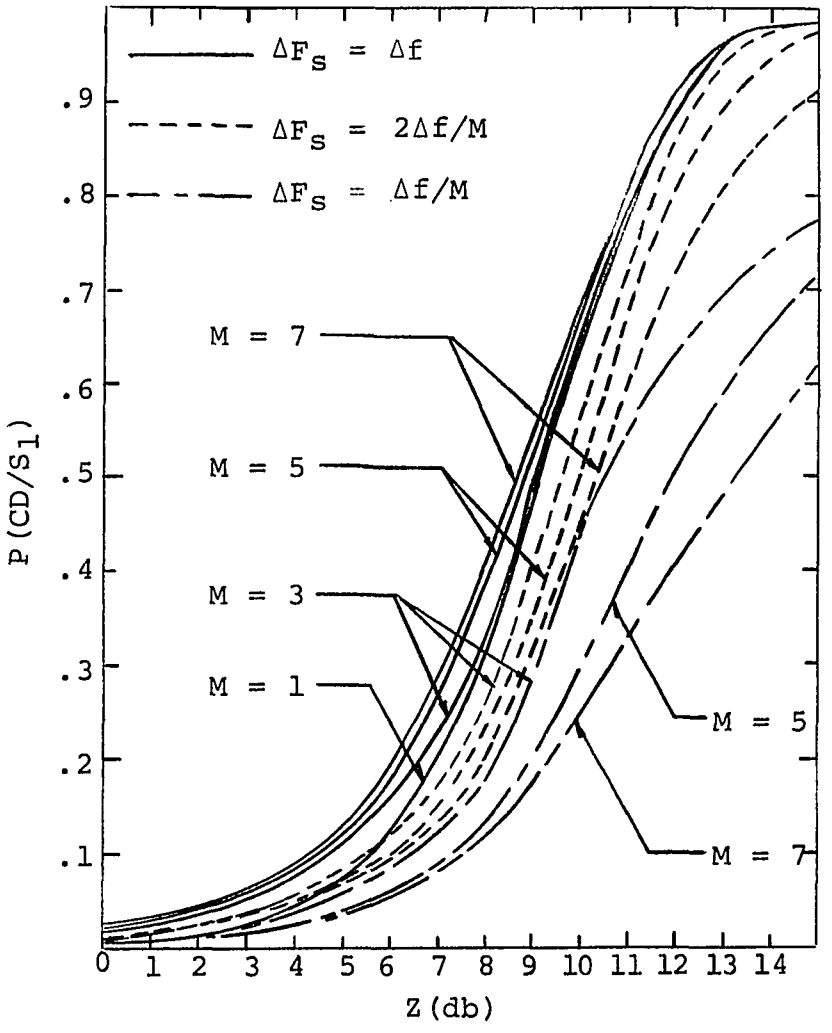


Figure 42. Probability of correctly detecting a single formant in extrinsic noise as a function of channel SNR, overlap factor, and minimum frequency separation between individual formants for a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and .02 false alarm rate.

channels, i.e., k_{nm} = constant < 1 for m ≠ n. The cross-coupling energy between detector inputs can, of course, only be achieved at the expense of signal resolution.

When a P.C.S. system is operated in the presence of intrinsic noise, the time interval between independent noise samples is not limited by the bandwidth of the individual channel filters; i.e., it is possible for the auto-correlation function of the noise in a given channel

$$R_{nn}(t) = 0 \text{ even if } t << 1/\Delta f \ . \qquad (55)$$

Under such circumstances a number of independent distributions of noise energy spectra may occur over the time interval during which the input formant spectrum remains constant.[7] As a consequence, the effect of the threshold inhibition will be to modify the individual channel firing probabilities, $P(Y_i/f_s)$, on the basis of the previous firing patterns. When the detector thresholds of a lateral inhibitory process tend to be highest at those channels where the exciting formant is most likely to be centered, the probability of correctly finding the formant center frequency in succeeding intervals of time decreases. If, however, the inhibited threshold levels tend to be lowest at those channels where the exciting formant is most likely to be centered, the probability of correctly finding the

---

[7]This condition can also exist in the case of extrinsic noise, when a P.C.S. is not being utilized at its maximum information rate; i.e., the exciting formant is maintained constant for a time period $\Delta t > 1/\Delta f$.

formant center frequency in succeeding intervals of time is increased.[8] This centering action still cannot fuse two clusters of activated channels into one, but it can produce inhibitory conditions which favor the excitation of a single cluster of activity.

The probability of a particular output firing pattern given the presence of a single exciting formant, $P(Y_I^*/f_s)$, now is described by a first order Markoff process since this probability depends on the previous firing pattern. A rough measure of this centering action can be obtained by determining the probability distribution of the frequency of the lowest level of the threshold inhibition, $P(f_m/f_s)$, the location of this lowest level being the determining factor of the succeeding $P(Y_I^*/f_s)$ probability. The location of f_m for a given firing pattern (Y_i) can be determined from equation (46). The $P(f_m/f_s)$ distribution can then be determined (in the case of intrinsic noise) by summing the probabilities of those firing patterns which correspond to a particular frequency of minimum inhibition:

$$P(f_m/f_s) = \sum P(Y_i/f_s) \text{ over all } Y_i \text{ corresponding to } f_m \,. \qquad (56)$$

The $P(f_m/f_s)$ distributions of the sample 100 channel P.C.S. for an exciting formant of cosine squared spectral

---

[8]The action of the lateral inhibitory network in this instance can be thought of as an adaptive filter.

shape and square filter bandpass characteristics were determined. In the instances of firing patterns where the lowest level of threshold inhibition was not limited to a single channel, the median channel of the cluster of channels having the lowest level of inhibited threshold was assumed to correspond to f_m. In Figure 43 the resultant relationship between $P(f_m/f_s)$ and Z is shown for a fixed false alarm rate of .02 and overlap factors (M) of 3, 5, and 7. In Figure 44 the standard deviation (σ_m) of f_m as a function of Z is presented. A comparison of these last two figures with Figures 11 and 12 of the maximum likelihood process shows that these two sets of data are almost identical. This result indicates that (for the assumed conditions) the adaptive centering performance of a lateral inhibitory process can be expected to approach that of the maximum likelihood process.

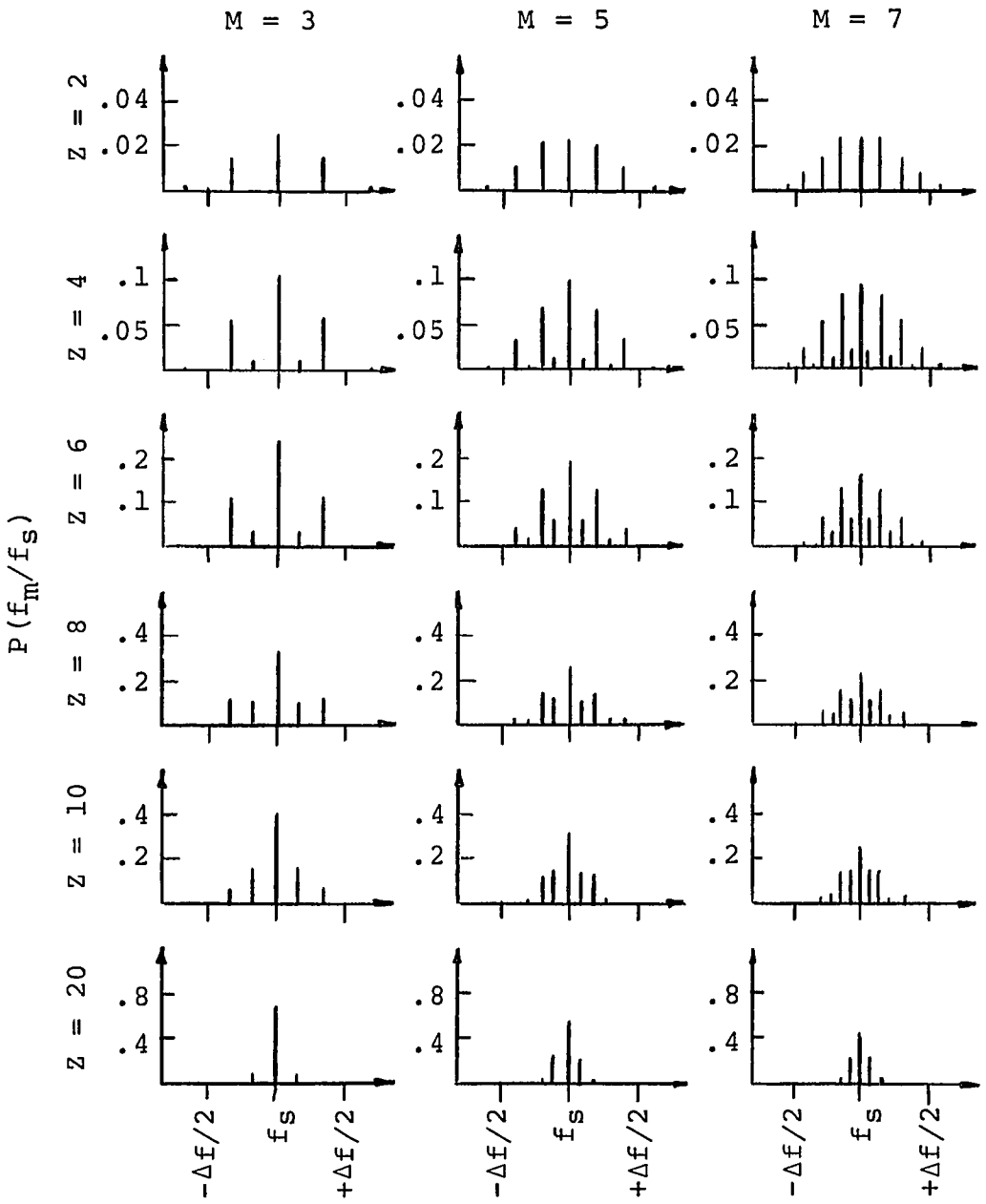


Figure 43. Probability distribution of the frequency of minimum inhibition of a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of intrinsic noise.

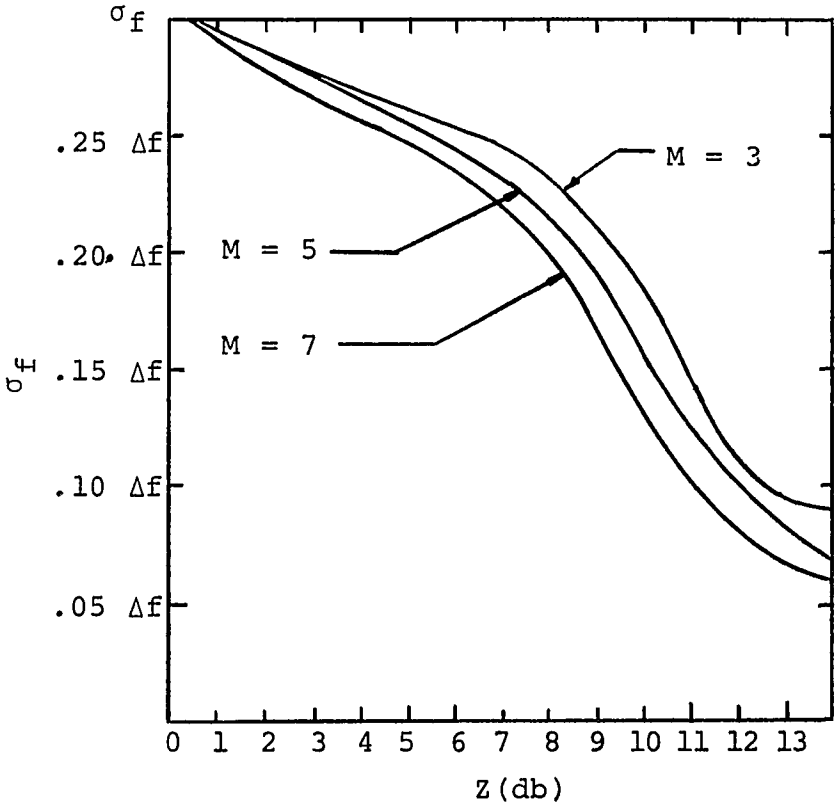


Figure 44. Standard deviation of the frequency of minimum inhibition (given detection of a formant) of a 100 channel P.C.S. with (L.I.C.) lateral inhibitory centering and a false alarm rate of .02 in the presence of intrinsic noise.

7. A REVIEW OF PERTINENT ANATOMICAL AND PHYSIOLOGICAL PROPERTIES OF THE EAR

This section is presented as background material for the discussion of the application of centering in P.C.S. to a model of the peripheral auditory system. The anatomy and physiology of the ear are covered in detail in a number of texts and will not be discussed in depth here except as they apply specifically to the problem being considered [34, 43, 62, 63]. Some general knowledge of the auditory system, however, is essential to the understanding of the model to be developed. It will be particularly helpful to consider the auditory system from the point of view of system theory representing the entire perceptory system as a chain of transmission elements of known transfer characteristics [11]. Even though such an analysis cannot as yet be performed without introducing several speculative hypotheses, it is possible to reach a reasonably reliable representation on the basis of the anatomical, physiological, and psychophysical data gathered by a large number of researchers during the past two centuries [61].

In Figure 45, a simplified representation of the anatomy of the ear is compared with a functional block diagram of the peripheral auditory system.

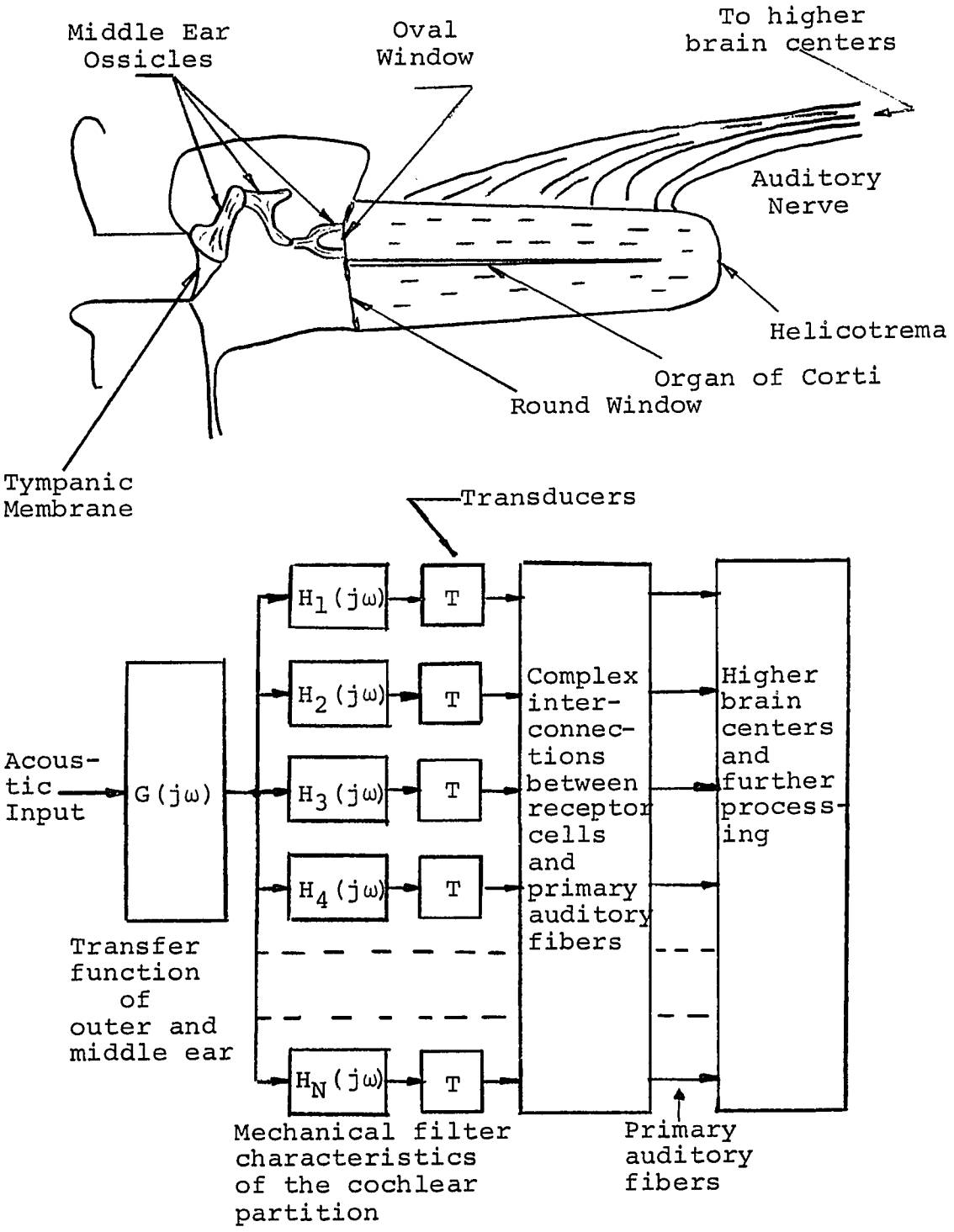


Figure 45. (a) Top diagram: schematic representation of anatomical chain; (b) bottom diagram: functional block diagram of peripheral auditory system.

Airborne sound enters the outer ear past the auricle (the externally visible ear structure) by way of the external auditory meatus (auditory canal) and impinges on the tympanic membrane (eardrum). The sound pressure sets the tympanic membrane into motion, which in turn excites a set of three small bones (the ossicles). The ossicle chain (malleus, incus, and stapes) transmits the motion of the tympanic membrane to another membrane covering the oval window, which is the mechanical input for that portion of the auditory system known as the inner ear. The primary role of this initial portion of the auditory system, referred to as the middle ear, appears to be the matching of the relatively low acoustic impedance of air to the higher impedance of the inner ear [63]. The properties of the middle ear are represented by the first block in Figure 44 and are essentially that of a frequency selective impedance matching network. The transfer characteristic of this block as determined by Flanagan [17] is shown in Figure 46.

The section of the ear of greatest interest is the inner ear, for it is here that much of the processing of sound is believed to occur. The most complex component of the inner ear is the cochlea. The cochlea is shown in straightened form in Figure 45a (in reality in man it exists as a 2-3/4 turn spiral of about 35 mm in length). Figure 47 shows two more detailed views of the cochlea. As can be seen from these diagrams, the cochlea is composed of two fluid-filled tubes, the Scala Vestibuli and Scala

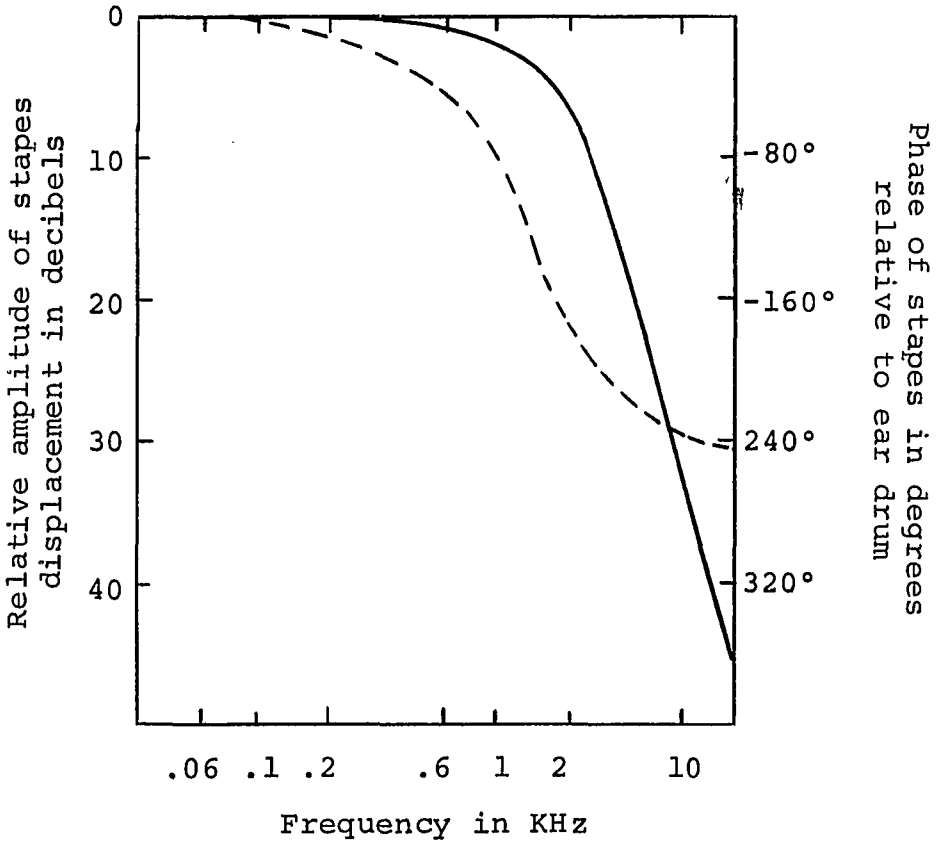


Figure 46. Approximate transmission characteristics of the middle ear based on Flanagan's function.

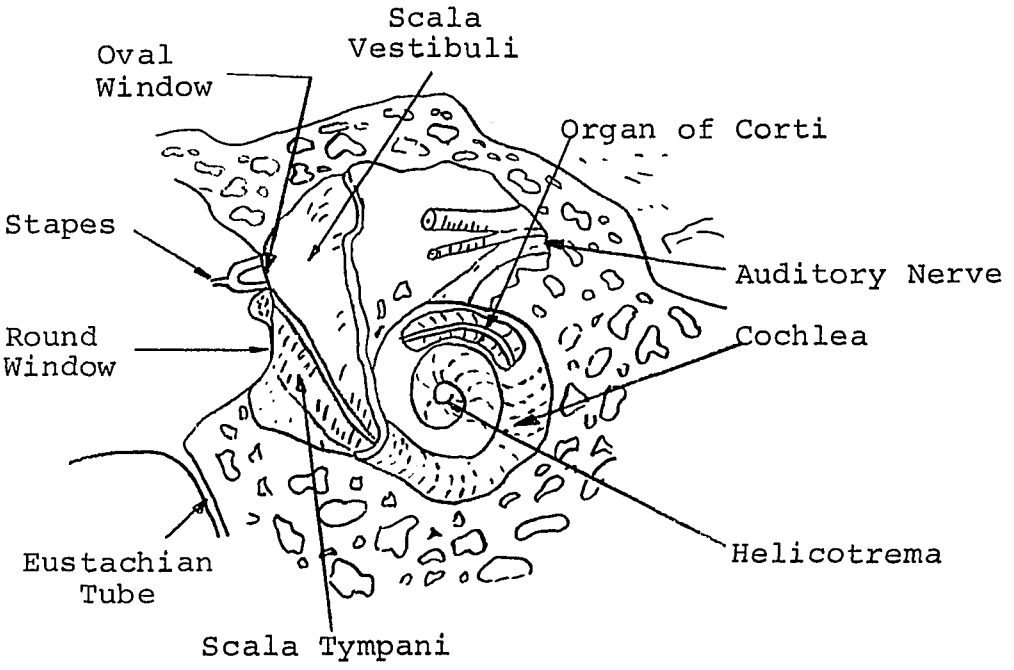


Figure 47a. The cochlear portion of the inner ear.

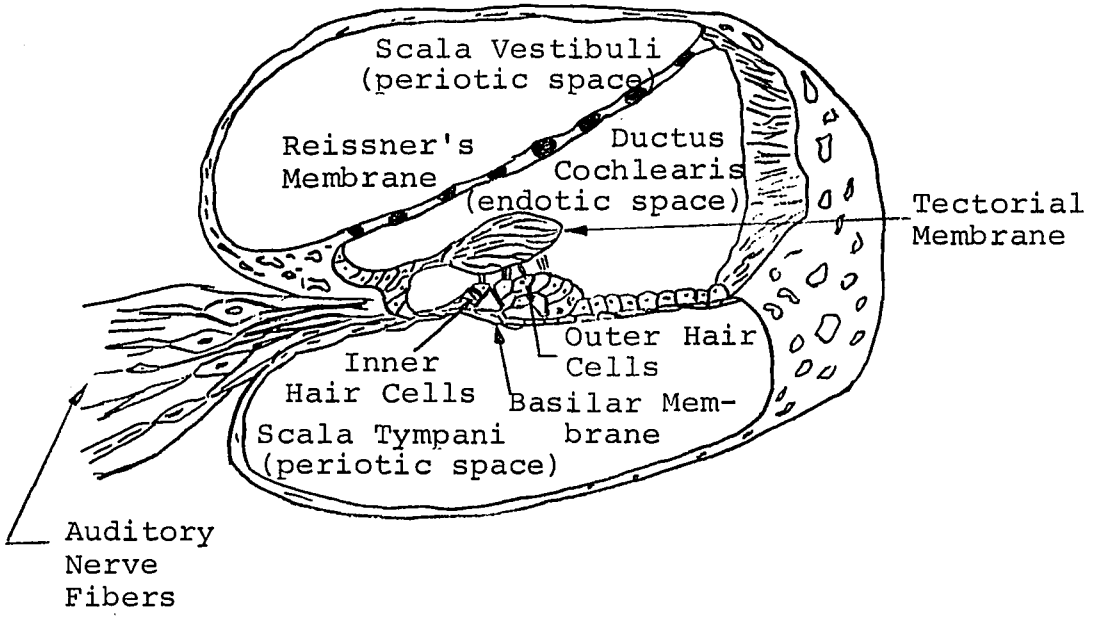


Figure 47b. Cross section of unrolled cochlea.

Tympani, which meet at the helicotrema. At the base of the cochlea these two tubes are sealed by the oval and round windows, respectively. The two scalae are separated from each other by the multimembraneous cochlear partition.

When vibrations are introduced into the cochlea's fluid by the motion of the stapes at the oval window, a traveling wave is established along the cochlear partition. This wave is due to the difference in pressure between the two scalae.

At present the universally accepted theory of the dynamics of the cochlear partition is due almost entirely to Von Békésy, who observed the motion of the cochlear partition by cutting away various portions of the cochlea and replacing these portions with transparent surfaces [2]. For sinusoidal stimuli the relationship between displacement and time of each point of the cochlear partition is sinusoidal and at steady state the envelope of the displacement distribution in space exhibits a maximum whose position depends on the excitation frequency. Figure 48 shows a number of graphs depicting the relative amplitude of the envelope of vibrations along the cochlear partition for various frequencies of excitation. The maximum value of displacement of individual points along the cochlear partition varies as a function of frequency [2, 34], as shown in Figure 49a. Figure 49b shows the position of the maximum displacement of the cochlear partition as a function of the excitatory frequency.

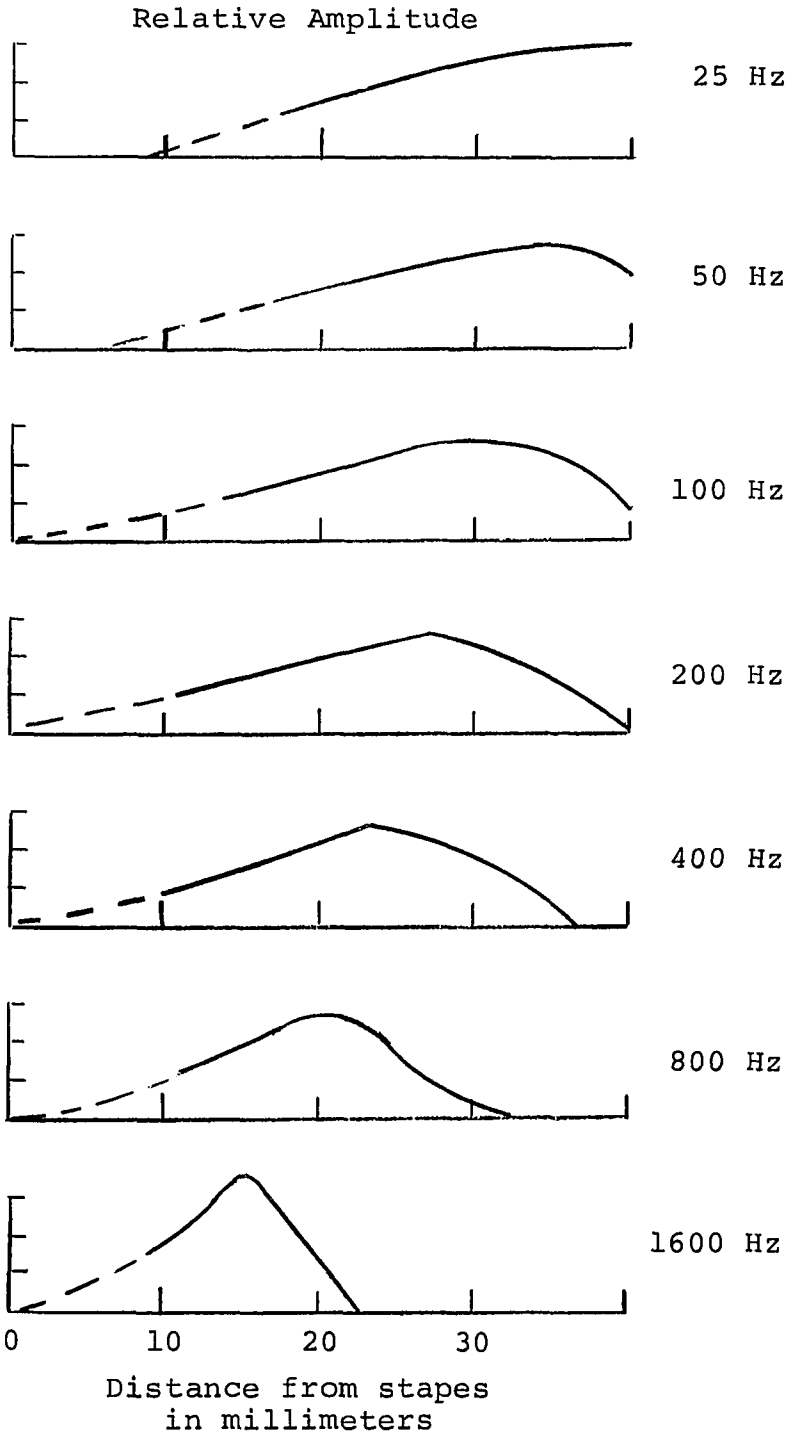


Figure 48. Relative amplitude of the envelope of vibration along the cochlear partition for various frequencies of excitation (after Békésy [2]).

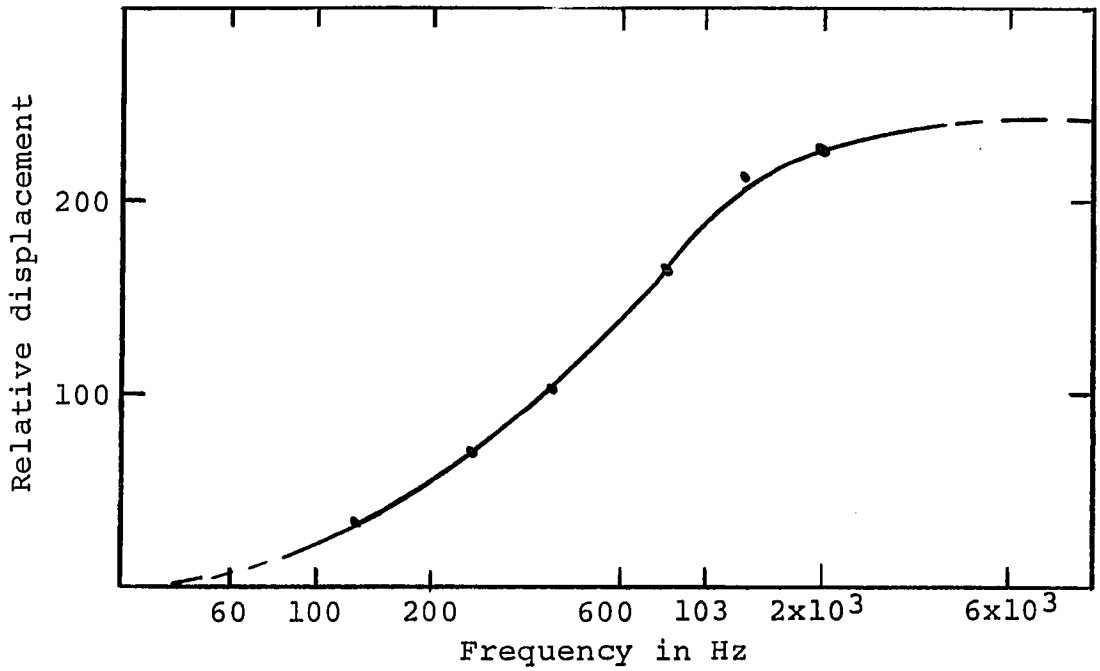


Figure 49a. Maximum value of displacement of individual points along the cochlear partition as a function of frequency [2].

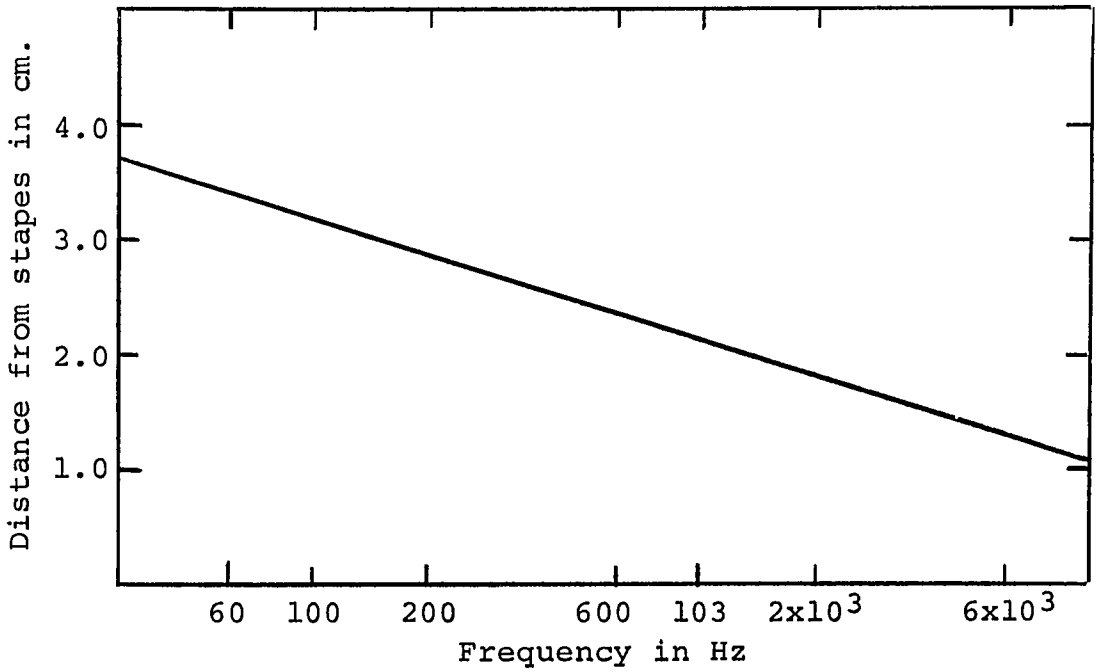


Figure 49b. Position of maximum displacement of cochlear partition as a function of frequency of excitation [2].

The cochlear partition can thus be considered a kind of mechanical filter relating position to frequency. The first column of small blocks in Figure 45b represents the filter action of individual points along the cochlear partition. The filter responses of these points overlap and have the characteristics shown in Figure 50, in which the amplitudes of displacement of various points along the cochlear partition are presented as a function of exciting frequency. Note that all the curves of Figure 50 have the same general shape (when plotted on logarithmic coordinates). These curves all show a relatively broad response with a Q of about 1.6 (when Q is defined as the ratio of the resonant frequency to the 3 db bandwidth).

The cochlear partition is bounded by and includes Reissner's membrane and the basilar membrane. It is the approximately 100:1 variation of elasticity of the basilar membrane (in the longitudinal direction) which is believed to be chiefly responsible for the cochlear partition's dynamic properties. Supported on the basilar membrane is the Organ of Corti containing (in man) approximately 2.8×10^4 receptor cells [43] known as hair cells because of the cilia which extend from their apices into the tectorial membrane.

It is generally believed that, when the cochlear partition vibrates, the cilia are forced to bend, and that this bending is somehow translated into the electrical impulses (action potentials) observed at the individual

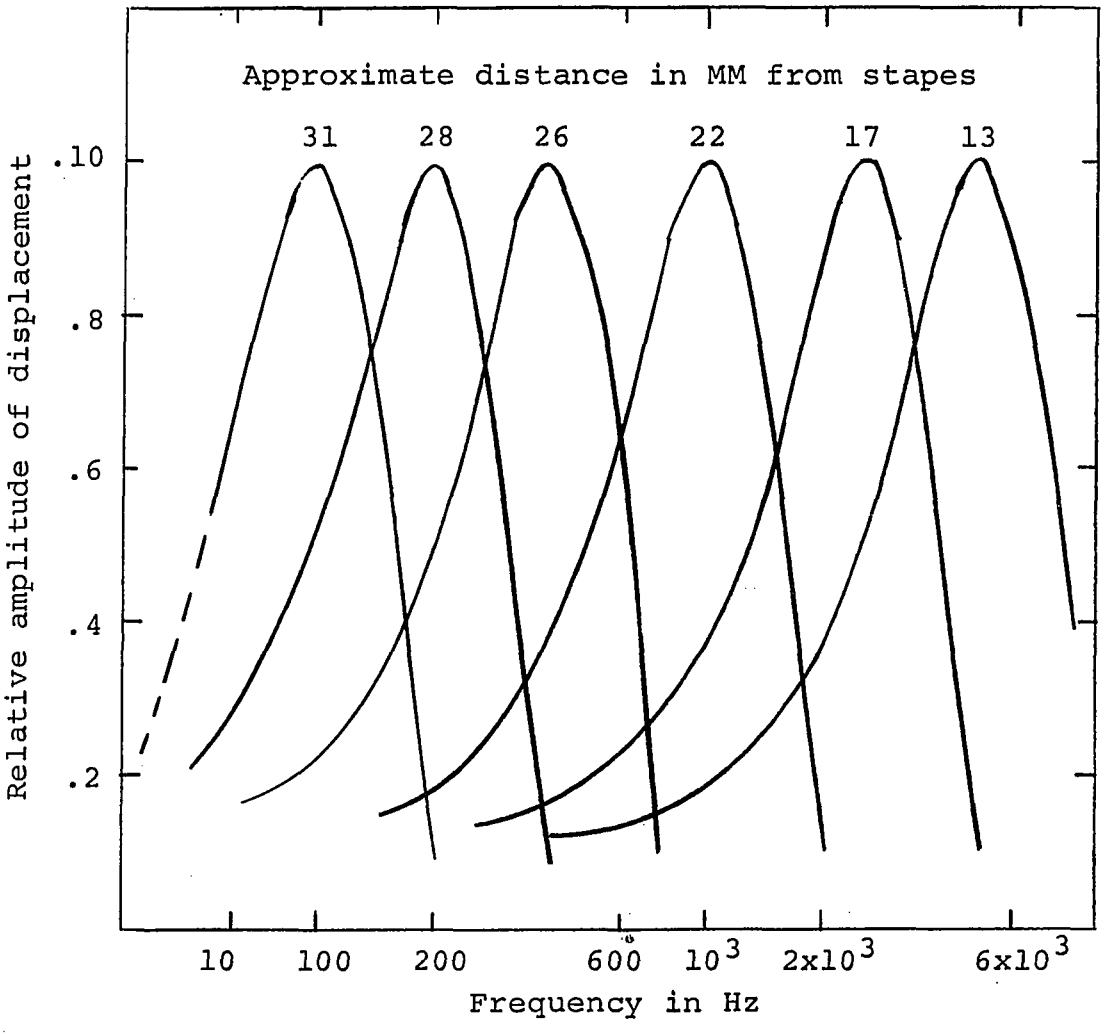


Figure 50. Amplitude of displacement of the cochlear partition as a function of frequency of excitation for several positions along the cochlear partition [2].

fibers of the auditory nerve. There is, however, no agreement on how this transformation actually occurs. The second column of smaller blocks in Figure 45b represents this "supposed" transducer action.

Among the receptor cells there are at least two distinct types: a single row of inner hair cells and three or four rows of outer hair cells. How the hair cells are innervated (connected together and to the primary auditory nerve) is not clearly understood. There appear to be at least two kinds of innervation: a local innervation provided by the radial nerve fibers and a more diffuse innervation provided by the spiral nerve fibers [15]. The inner hair cells appear to be predominantly innervated by radial fibers, while the outer hair cells seem to be innervated by spiral fibers as well as by radial fibers. Furthermore, the nerve fiber endings, adjacent to the hair cells, show some differentiable features. Some of these endings show simple knoblike structures, while others form a large area of contact between the hair cell and the neuron [12]. The second large block in Figure 45b represents the complex interconnections between the hair cells and the primary auditory nerve fibers of the VIII nerve.

The afferent nerve fibers (nerve fibers proceeding toward the brain) which innervate the receptor cells have their bodies in the spiral ganglion, while their axions end at the cochlear nucleus in the medulla oblongata (central

nervous system).[1] The cells of the spiral ganglion are bipolar; and it is their aggregate which primarily forms the VIII nerve (represented in Figure 45b by the many lines interconnecting the third and fourth large blocks). The histology of a cross section of the primary auditory nerve reveals a remarkable uniformity of fiber diameters (3-5 μ) [49]. Furthermore, fibers originating in the basal turns (high frequency end) of the cochlea wrap around the central core fibers originating in the more apical turns (lower frequency sections) of the cochlea. The VIII nerve, therefore, is composed of a regular array of fibers; and spatial contiguity of adjacent fibers is thus preserved [43].

Measurements of the electrical activity in the cochlea reveal the presence of several distinguishable electrical potentials. In the absence of stimulation, the endolymph (fluid contained within the cochlear partition) has been observed to have a potential of +50 MV (referred to as the EP or endocochlear potential) with respect to the perilymph (fluid contained within the two scalae) [2]. If the ear is stimulated with sound, several other "gross" potentials can be observed.[2] These are the cochlear microphonic potential (CM), the summation potential (SP), and the gross action

---

[1]There is also evidence that a number of efferent fibers (nerve fibers proceeding from the brain) terminates in the Organ of Corti [75].

[2]Gross potentials are potentials observed with electrodes of large diameter, when compared with the dimensions of the microstructures which are generating the electrical events.

potential (N_1). The CM and the SP potentials are believed to be involved in the initiation of the action potentials in the individual nerve fibers of the auditory nerve [2, 34]. Local recordings of these potentials, however, are difficult to obtain. The N_1 potential is thought to represent a spatio-temporal summation of the electrical activity in the individual auditory nerve fibers [63].

Of particular interest is the electrical activity of auditory nerve fibers on an individual basis. In Figure 51 the relationship between the number of electrical impulses per second (spike rate) generated by individual auditory fibers (of a cat) and the minimum intensity of acoustic excitation (at the stapes) required to produce this rate is presented [32]. This figure may be thought of as representing the transfer characteristics of the transducer blocks of Figure 45b. Although there are great variations in spike rates from fiber to fiber, some generalizations can still be made. All fibers appear to exhibit a spontaneous firing rate; i.e., action potentials are observed in the absence of acoustic stimulation. As acoustic stimulation is introduced, the firing level remains constant until a critical level is reached, at which point the spike rate appears to rise sharply. Above this level the spike rate increases in an "approximately linear" fashion, as a function of stimulation until a second critical level is reached. Increasing the stimulation beyond this second level results in little or no increase in spike rate, and

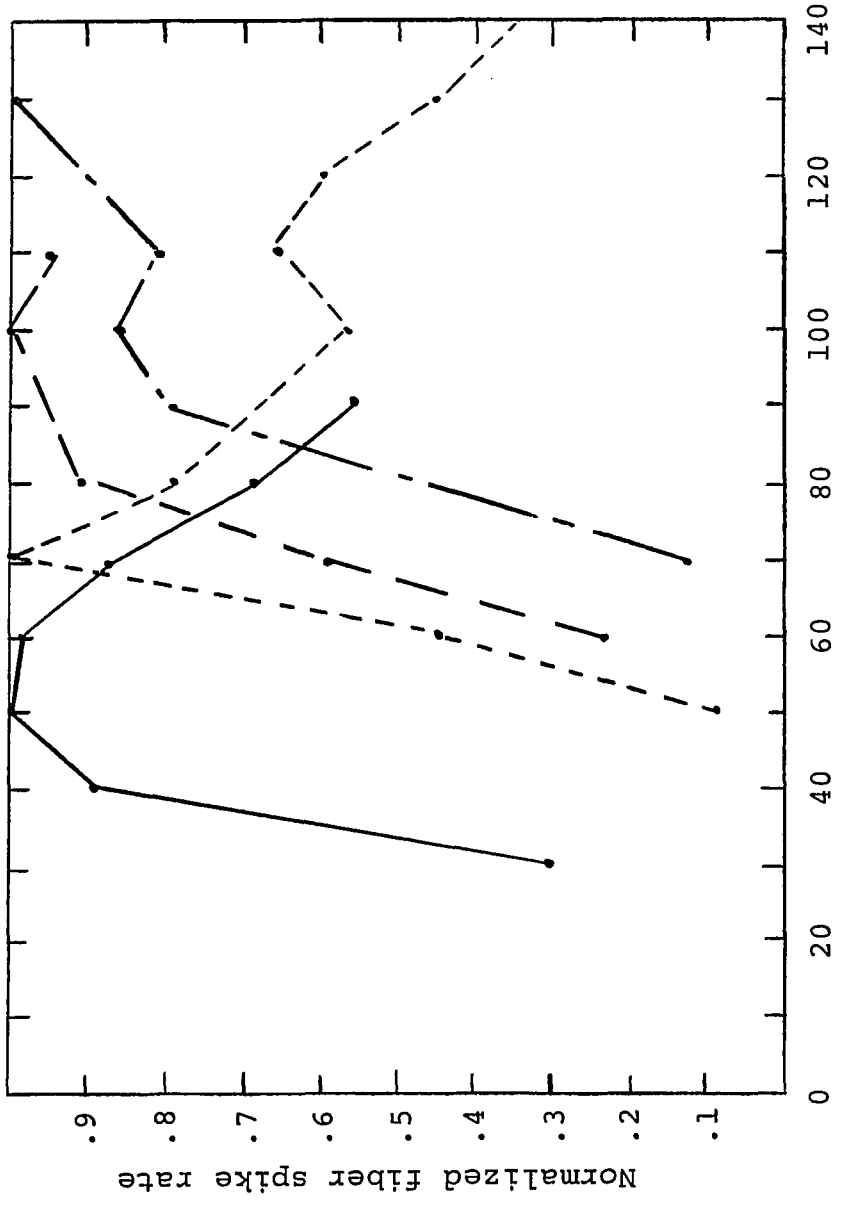


Figure 51. Normalized spike rate of several individual auditory fibers (of a cat) as a function of the minimum acoustic excitation. The normalized spike rate equals the actual spike rate minus the spontaneous spike rate divided the maximum spike rate minus the spontaneous spike rate (after Kiang [32]).

in some instances even a decrease in spike rate. The first critical level may be thought of as a kind of threshold of excitation, and the second as a threshold of saturation. The range of stimulation between these two extremes is 20 db or less for many fibers and seldom exceeds 30 db.

The minimum acoustic intensity required to produce a given spike rate in a given fiber is very much a function of the frequency of the acoustic stimulation [32, 33]. The generation of spikes for a minimum amount of acoustic energy occurs at a sinusoidal exciting frequency, referred to as the characteristic frequency (CF) of the fiber in question. A measure of the spectral selective properties of several fibers observed by Kiang [33] in one cat is displayed in terms of the "tuning" curves shown in Figure 52. Tuning curves are graphs relating the threshold of firing of a fiber (the intensity at which an observable increase in a fiber's spike rate occurs when compared to its spontaneous rate of discharge) to the frequency of sinusoidal sound stimulation. The relative sharpness of these tuning curves does not change significantly up to a frequency of about 2 KHz. Above 2 KHz the relative sharpness of these tuning curves increases markedly as a function of frequency. When wideband noise is added to the sinusoidal stimulation, the overall sensitivity decreases, but the general shape of the tuning curves remains the same.

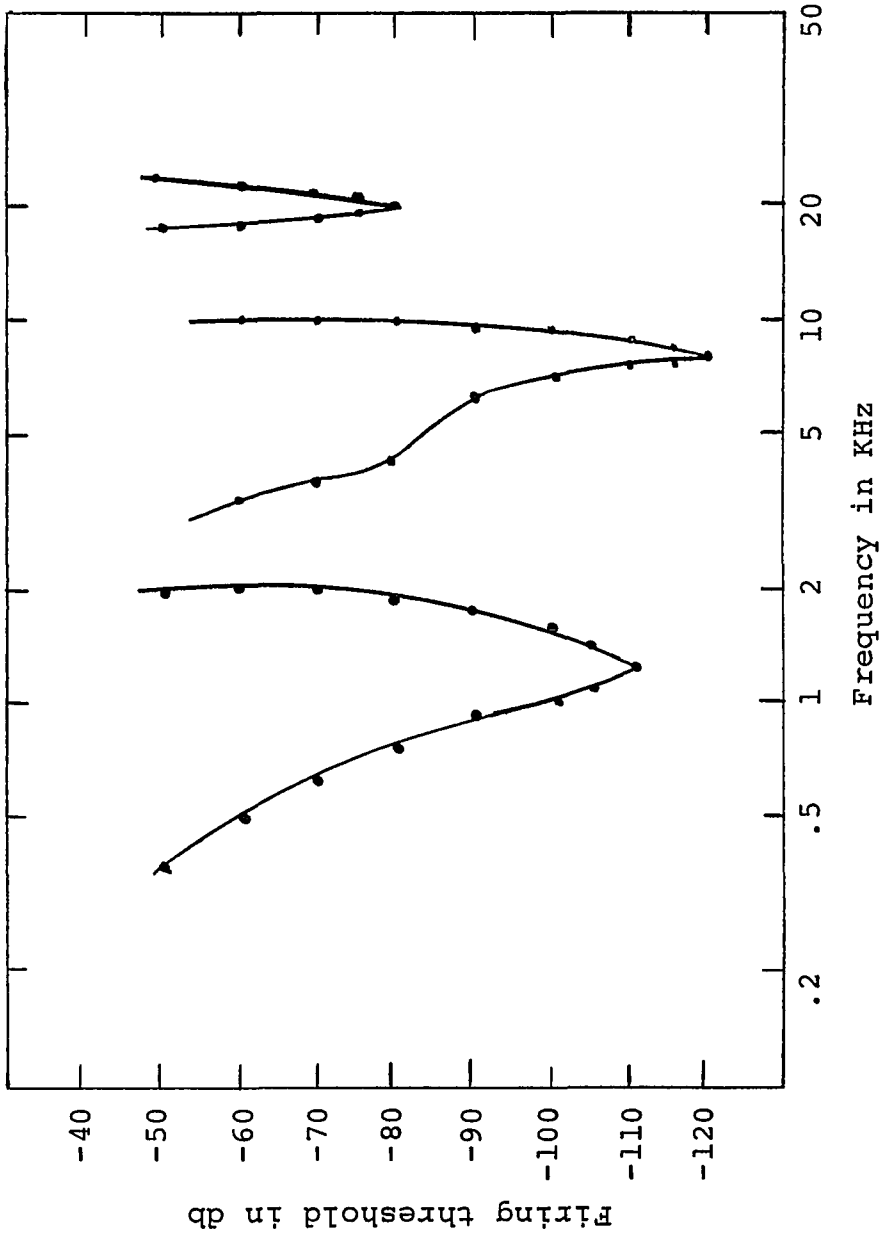


Figure 52. Tuning curves of three primary auditory nerve fibers of the cat [33].

8. A MODEL OF CENTERING IN THE PERIPHERAL AUDITORY SYSTEM

There is much indirect evidence that the frequency-selective properties displayed by individual auditory-nerve fibers are at least partially a function of the "resonant" properties of the cochlear partition. Fibers which respond to higher frequencies are found further from the center of the primary auditory nerve than fibers which respond to low frequencies [32]. The skirts of the tuning curves are sharper for frequencies above their characteristic frequency than for frequencies below it (as are the properties of the cochlear partition).

Studies of the time domain characteristics of the action potentials generated in response to an acoustic click (impulse) indicate that the onset of spike activity (on the average) occurs later for fibers with low CF than for fibers with high CF [32]. This delay does not vary significantly with changes in intensity of stimulation. There is also a difference in the time of occurrence of the first spike activity for clicks of different polarity (rarefaction vs. condensation--which can be related to the dynamics of the cochlear partition) [32].

The frequency selective characteristics displayed by

the auditory-nerve fibers, however, are much sharper than would be expected as a result of their innervation by the (broadly selective) cochlear partition. (This condition exists even for fibers with CF below 2 KHz.)

In Figure 53, curve A, the relative firing thresholds expected of individual neural fibers, exclusively on the basis of their excitation by the motion of the cochlear partition in response to a tone, are plotted as a function of their normalized characteristic frequency (f_{CF}).

$$f_{CF} = CF/f_n \ , \tag{57}$$

where f_n is the frequency of the exciting tone. The actual frequency of the exciting tone is unimportant as the Q of the cochlear partition is believed to be independent of frequency.[1] For the purposes of calculating curve A, the following assumptions were made: (1) the hair cells (transducers) were assumed to be uniformly distributed along the basilar membrane;[2] (2) the CF of any given fiber was assumed equal to the frequency of stimulation which would cause a peak in the envelope of the displacement of the cochlear partition at the position along the basilar membrane of the

---

[1]Actually, Békésy's measurements of the dynamics of the cochlear partition were restricted to frequencies below 2 KHz.

[2]Although the hair cells are uniformly distributed along the length of the basilar membrane, the nerve fibers which innervate them are not. The primary nerve fibers are more heavily concentrated at the basal end of the cochlear partition [23].

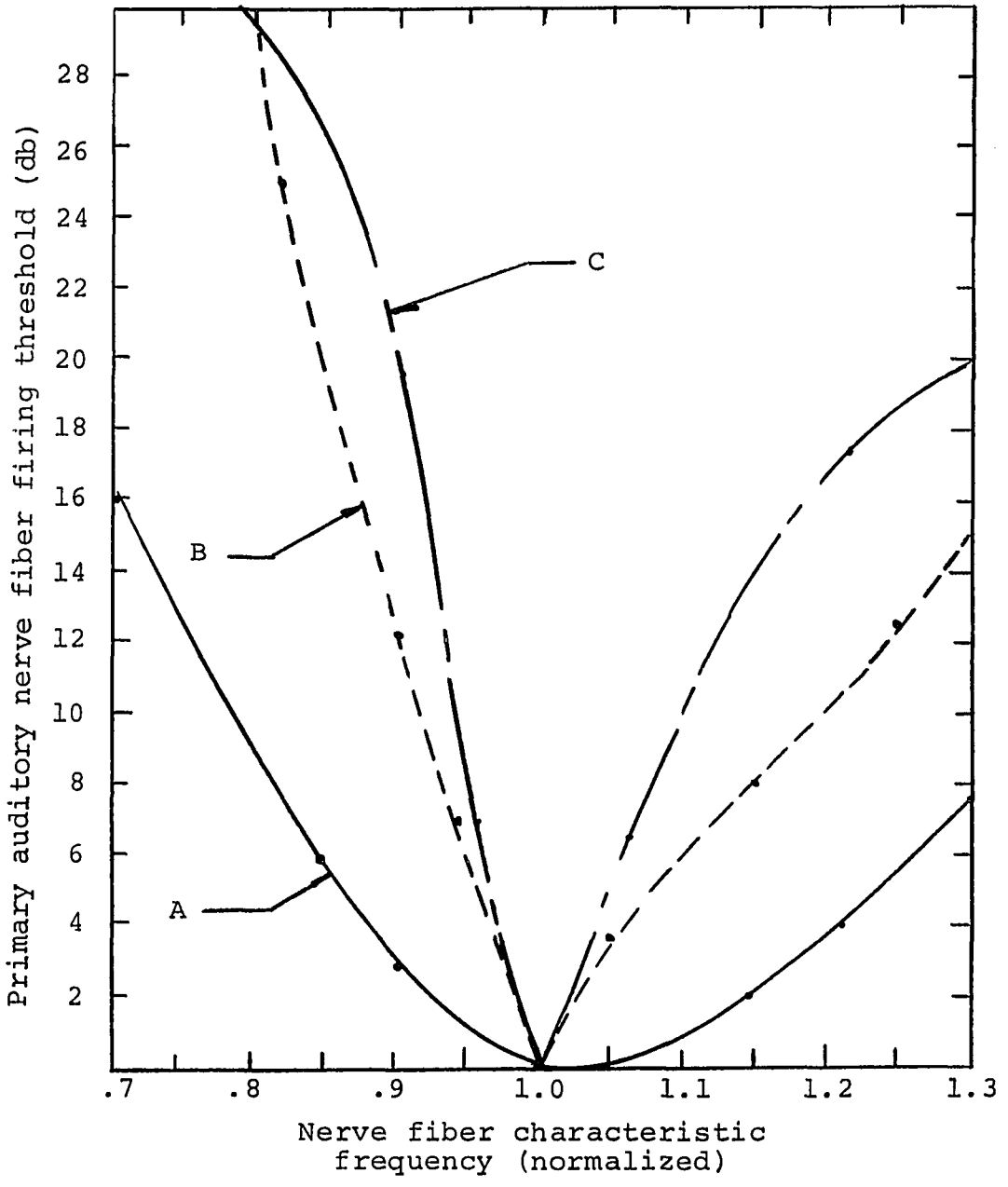


Figure 53. Expected and experimental firing thresholds of individual auditory nerve fibers as a function of their normalized characteristic frequencies. Curve A: Firing thresholds of fibers expected exclusively on the basis of their excitation by the cochlear partition. Curve B: Idealized firing thresholds extrapolated from experimental results of Kiang [32, 33]. Curve C: Firing thresholds exhibited by the model.

hair cell it innervates; (3) the position of the peak in the envelope of the vibration of the cochlear partition was assumed to be logarithmically related to frequency of stimulation; (4) the receptor cells were assumed to be sensitive to the vertical displacement of the cochlear partition;[3] (5) the variation in the absolute value of displacement of the cochlear partition for frequencies below 1 KHz (see Figure 49a) was neglected as well as the effects of phase delay along the partition.[4] The frequency response of individual points along the cochlear partition was modeled after Flanagan's [17] expression for the transfer characteristic of the basilar membrane,

$$H_X(j\omega) = A_X \left\{ \frac{j\omega}{2\pi CF} \Big/ \left[\left(\frac{j\omega}{2\pi CF} + 1 \right) \left(\left[\frac{j\omega}{2\pi CF} + .5 \right]^2 + 1.0 \right)^2 \right] \right\} , \tag{58}$$

where X represents distance along the basilar membrane and ω is the exciting frequency in radians. Distance along the basilar membrane is related to CF (for the above

---

[3]It has been suggested that the hair cells respond to the longitudinal motion of the cochlear partition [31] rather than its vertical displacement. Such an assumption would provide a narrower spectral response of individual points along the cochlear partition than assumed here. This sharpened response is still, however, inadequate to explain the tuning characteristics displayed by the individual neural fibers.

[4]These assumptions seem reasonable in light of our prime concern with the relative frequency selective properties at a point along the basilar membrane.

assumptions) by the equation

$$X = C - KI \log (CF) , \tag{59}$$

where

$$C = \frac{DI}{1 - \log (CF_{min})/\log (CF_{max})} . \tag{60}$$

DI is the length of the basilar membrane in the same units as X; CF_{min} is the CF corresponding to the most apical point of the basilar membrane; and CF_{max} is the CF corresponding to the most basal point of the basilar membrane.

$$KI = C/\log (CF_{max}) . \tag{61}$$

Curve B (Figure 54) presents the idealized firing thresholds of primary auditory fibers as a function of normalized CF. This curve was extrapolated from data gathered by Kiang [32, 33] on the tuning characteristics of auditory nerve fibers.[5] The curve shown is based on fibers with CF around 1 KHz. Fibers with higher CF display even sharper frequency selectivity. Again differences in the absolute threshold of firing have been ignored.

---

[5]For the purposes of calculating curve B, assumptions similar to those used in the calculation of curve A were made. However, while in the calculation of curve A use was made of the frequency response of the cochlear partition, in the case of curve B, a "typical fiber tuning curve" (CF about 1000 Hz), as determined by Kiang, was used to evaluate the pure tone levels of a given frequency, necessary to excite activity in different nerve fibers.

From these two curves, the magnitude of the disparity between auditory fiber firing thresholds, expected on the basis of the spectral response of the cochlear partition, and those actually observed, can be seen. This difference must be corrected for, by some physical process.

Any mechanism introduced in the model to sharpen the system's frequency discrimination must not limit its excellent temporal resolution. The excellent temporal resolution and spectral discrimination displayed by the ear has been measured by a number of different psycho-acoustical tests and are well documented in the literature [27, 34, 63].

A measure of the restrictions on the temporal resolution of acoustic signals imposed by the dynamics of the cochlear partition can be obtained from its impulse response. If the Q of the cochlear partition is approximately constant with frequency, its temporal resolving ability should improve at higher frequencies. The approximate impulse response of the cochlear partition has been calculated by both Flanagan and Siebert [18, 57] using different simplifying assumptions. The impulse responses obtained agree both in general shape and order of magnitude.

The temporal resolution at the individual auditory fibers is harder to interpret due to the apparent nonlinear transducer action of the hair cells. If the hair cells' nonlinear characteristics are assumed to be simple "go--no go" threshold functions, then the impulse response of

the cochlear partition can be approximately related to the time duration of spike generation. Such an assumption seems reasonable in view of the transfer characteristics displayed in Figure 51.

The impulse response of the basilar membrane as calculated by Flanagan [16] is

$$h(x,t) = A_x \{(.033 + .360t)e^{-t/2} \sin t + (.575 - .320t)e^{-t/2} \cos t - .575\, e^{-t}\} \quad (62)$$

for a normalized frequency $f_{CF} = 1$ and zero phase delay. The approximate duration of spike generation expected for this response, due to a threshold type nonlinearity can be estimated from Figure 54 which related impulse intensity (with respect to the firing threshold) in db to time. A comparison of this figure with the experimental results of Kiang [32] shows that even though the frequency response measured at the auditory fibers is sharper than the frequency response measured at corresponding points along the cochlear partition, the temporal response of the system has not been significantly degraded.

One way in which the spectral discrimination of a parallel channels system (assuming that the auditory system is a P.C.S.) may be sharpened without sacrificing temporal discrimination is by means of the centering processes considered in the past sections. In particular, the

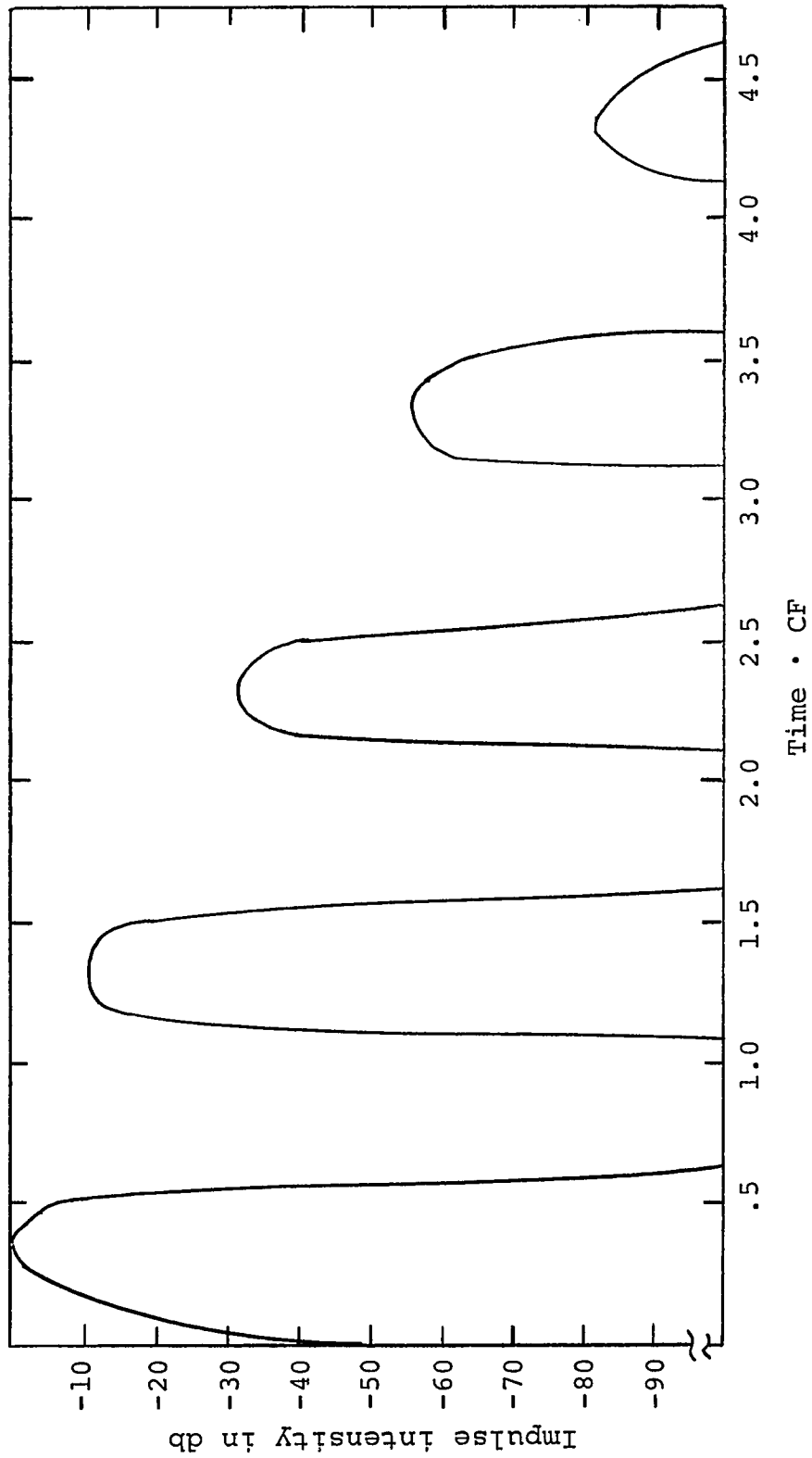


Figure 54. Impulse response of cochlear partition (normalized with respect to its maximum intensity) as a function of time (normalized with respect to 1/CF).

complex interconnections between the hair cells and the primary auditory fibers (second large block in Figure 45b) might be interpreted as the anatomical elements responsible for the introduction of centering by means of lateral inhibition.

Significantly, the existence of two-tone inhibition (the ability of a second tone to reduce the rate of spike generation produced by another tone) has been observed by Sacks and Kiang at the level of the primary auditory nerve fibers [51]. In Figure 55 the idealized response and inhibitory areas are shown for a typical auditory nerve fiber. The response area (tuning curve) is bounded by the dotted line and represents the region where an above threshold spike rate is generated by the presence of a single tone in terms of tone frequency and intensity. The inhibitory area has been drawn in cross-hatch and represents the values of intensity and frequency that a second tone must possess in order to inhibit the spike activity of a first tone at the fiber's CF. The small triangle represents the level of the first tone.

As an extension of two-tone inhibition it appears plausible to conceive that the spike generation at a given fiber due to the presence of a single tone may be inhibited by the action of the same tone at other fibers. Such conditions are precisely the ones created in the case of lateral inhibition. In order to investigate the validity of "centering" in general and lateral inhibition specifically,

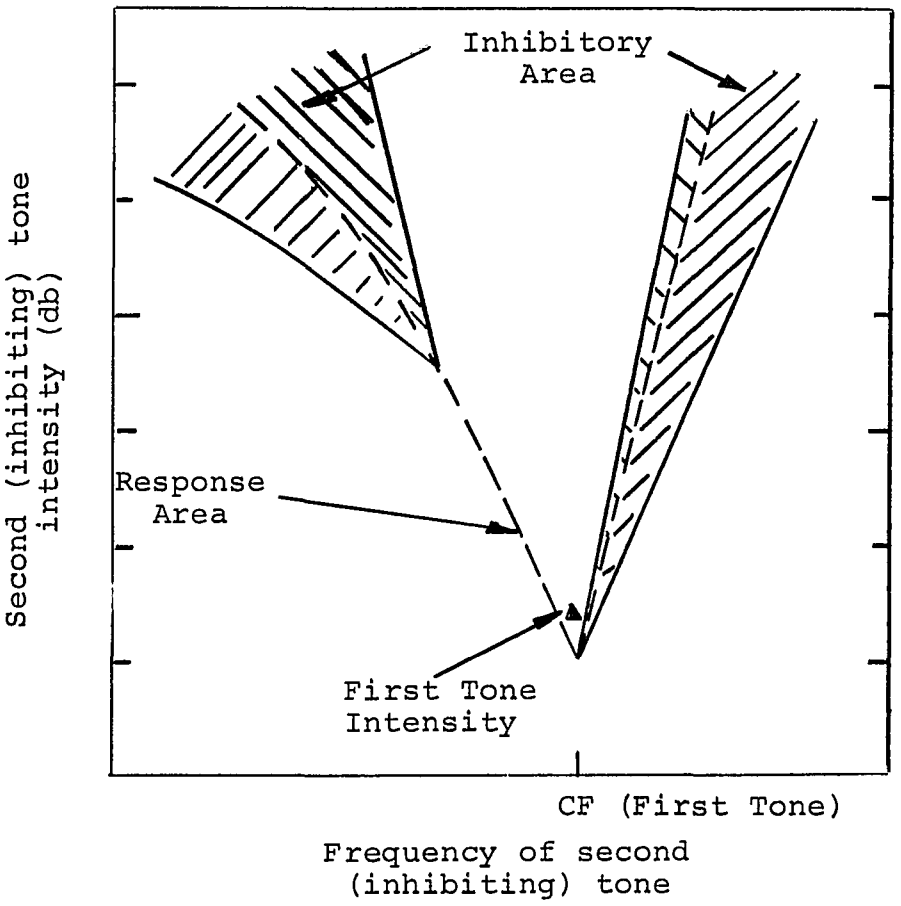


Figure 55. Idealized inhibitory and response areas for a typical auditory nerve fiber (after Sacks and Kiang [55]).

as possible mechanisms for sharpening the frequency discrimination of the peripheral auditory system, a multichannel model of the auditory system was simulated on the digital computer. The model was based on the schematic of the peripheral auditory system shown in Figure 45b. The spectral characteristics of the middle ear were neglected as most of the experimental data by which the performance of the model could be evaluated are based on acoustical stimuli introduced after this portion of the ear.

It was decided to simulate the system in terms of 4000 parallel channels, a compromise between the approximately 28,000 auditory fibers suggested by anatomical evidence and the time availability of the digital computer. Each channel consisted of a bandpass filter modeled after Flanagan's expression for the transfer characteristics of the basilar membrane [equation (58)] followed by a threshold type detector (transducer) which was coupled into a simple recurrent lateral inhibitory network. Assumptions identical to those used in deriving Figure 53, curve A, were made in order to specify the characteristics of each channel.

Of the many types of lateral inhibitory networks considered in section 6, a recurrent process which tends to locate its lowest level of threshold inhibition at the center of excitation [equation (46) and Figure 38] was chosen for use in the model because of its superior performance in the presence of both extrinsic and intrinsic noise.

In section 6, it was pointed out that a reasonable compromise between signal resolution and signal discrimination is obtained if the lateral control signal inhibits only those channels which share common spectra with the channel from which the control signal emanates. In the case of the P.C.S. considered in section 6, with their sharp (square) filter characteristics, there is no question as to the locations of the boundaries of this common region. In the case of the broad nonsymmetrical response of points along the cochlear partition, there is considerable uncertainty as to the location of the boundaries of this region. Even if the lateral control signals only inhibit those channels which share common spectra within 3 db of the signal level at the channels from which they originate, hundreds of separate lateral connections per channel and possibly millions of interconnections on a system level would be required.

Although such a mass of interconnections is not inconceivable in biological systems, it appears fruitful (if only from the point of view of simplifying the model simulation) to consider how the number of interconnections in a lateral inhibitory network may be reduced. One very promising circuit is shown in Figure 55. In this circuit the outputs of all threshold detectors are summed serially from low to high frequency and simultaneously from high to low frequency. The output of the nth summer proceeding from low to high (V_{rn}) can be considered to correspond to

the sum of all inhibitory coefficients for which $m \leq n$ [see equation (44)], while the output of the nth summer going from high to low (V_{Ln}) can be considered to correspond to the sum of all inhibitory coefficients for which $m \geq n$. Quantitatively,

$$V_{rn} = \sum_{m=1}^{n} (\ell_r)^{n-m+1} V_{om} \tag{63}$$

and

$$V_{Ln} = \sum_{e=1}^{N-n} (\ell_L)^{m-n} V_{om} , \tag{64}$$

where

$$m = N - e + 1 , \tag{65}$$

N is the total number of channels, and ℓ_r and ℓ_L are the summer gains proceeding, respectively, from low to high and from high to low frequency.

For the threshold detectors shown in Figure 56,

$$V_{Tn} = V_{th} + |K_n(V_{rn} - V_{Ln})| , \tag{66}$$

where K_n is a constant. The threshold inhibition generated by this network will be very similar to that described by equation (46), i.e., the lowest level of threshold inhibition will tend to occur at the center of excitation.

If the summer gains are set equal to one, the inhibited threshold levels generated by the above process will

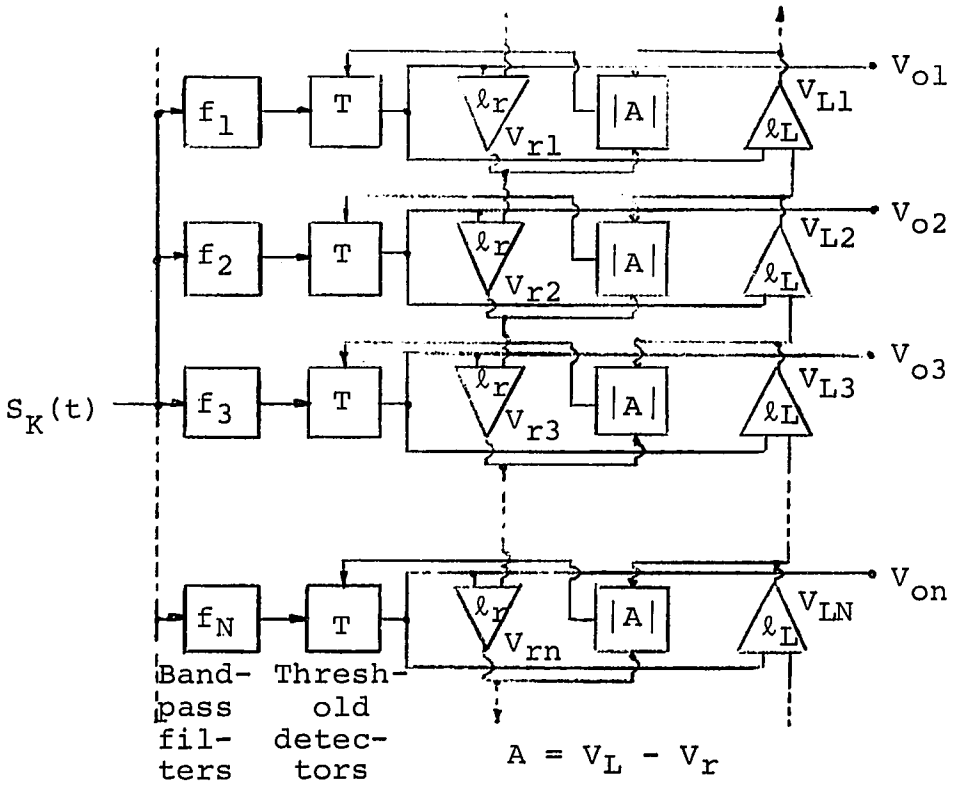


Figure 56. Lateral inhibitory process without multiple interconnections. The input signal passes through the bandpass filters (f_1 - f_n) activating the threshold detectors (T). The outputs of the threshold detectors in the form of a train of pulses or a D.C. signal in turn activate two series of summing amplifiers traveling upward and downward. The magnitude of the difference of individual summing amplifier outputs is then used as a reference signal which controls the level of the individual threshold detectors.

be identical to those established by the L.I.C. lateral inhibitory networks studied in section 6 in which the lateral control signal inhibits all adjacent channels. When the summer gains are set equal to values less than one, the inhibited threshold levels generated will approximate the levels established by the L.H. networks of section 6, in which the lateral control signals inhibit fewer than all adjacent channels. (When the ℓ_L's and ℓ_r's are small, then a signal in a given channel will sizably inhibit only a few adjacent channels.)

It is upon this process (Figure 56) that the lateral inhibition of the model of the auditory system investigated was based. Details of the computer programs used in the simulation and investigation of the model are presented in Appendix D.

In Figure 57 are shown a number of graphs depicting the centering action produced by the model for a signal 20 db (STR) above the noninhibited threshold of the detectors for various values of summer gain ($\ell_r = \ell_L = 1$, .999, .995, and $K_n = .35$). The exciting signal was assumed to be on continuously (sinusoidal tone); however, the graphs display the response of the centering process after only a few recurrent cycles.[6] The solid lines represent the output sampled at every 40th filter and are plotted as a function

---

[6] In the case of processes with $\ell_L = \ell_r = 1.0$ and .999, the response illustrated essentially represents the steady state inhibitory condition. In the case of $\ell_r = \ell_L = .995$, a steady state inhibitory condition is never achieved.

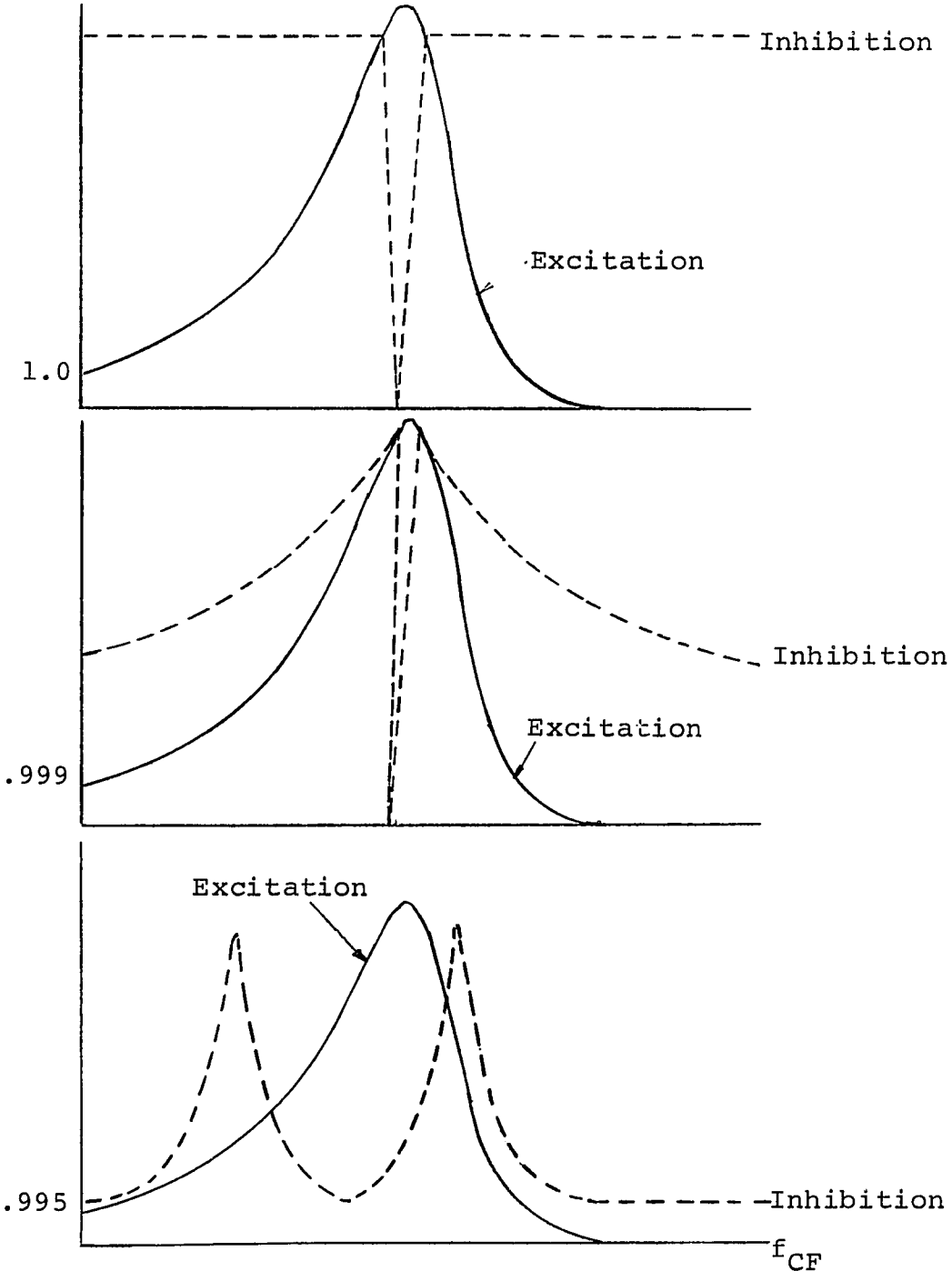


Figure 57. Centering action produced by the model in response to a 20 db STR signal for summer gains ($\ell_r = \ell_L$) of 1.0, .999, and .995 and K_n = .35.

of f_{CF}. The dotted lines represent the threshold inhibition generated in response to the signal. Wherever the filter output is greater than the threshold inhibition, an output signal is generated.

The threshold inhibition thus produced displays in all instances a characteristic wedge shape about the center of excitation. The sides of this wedge become increasingly concave for smaller values of summer gain. Of greater importance is the decrement in the threshold inhibition (for summer gains of less than unity) after the inhibition has reached its maximum at the top edges of the wedge. The rate at which this decrement occurs is a function of summer gain--increasing for smaller values.

The graphs of Figure 58 display the centering action of the model for conditions similar to those just assumed, except that the centering action here is displayed as a function of signal level (10, 20, and 30 db STR) for summer gains of .999. As the signal intensity is increased, the sides of the inside wedge become less steep while the slope of the threshold distribution curve outside the central wedge stays essentially invariant. Thus, as the signal intensity is increased, the number of activated channels increases. It is conceivable that the brain could interpret this greater number of adjacent activated channels as representing a louder signal [23] providing a mechanism for amplitude sensing.

The firing thresholds of curve B, Figure 53, are based

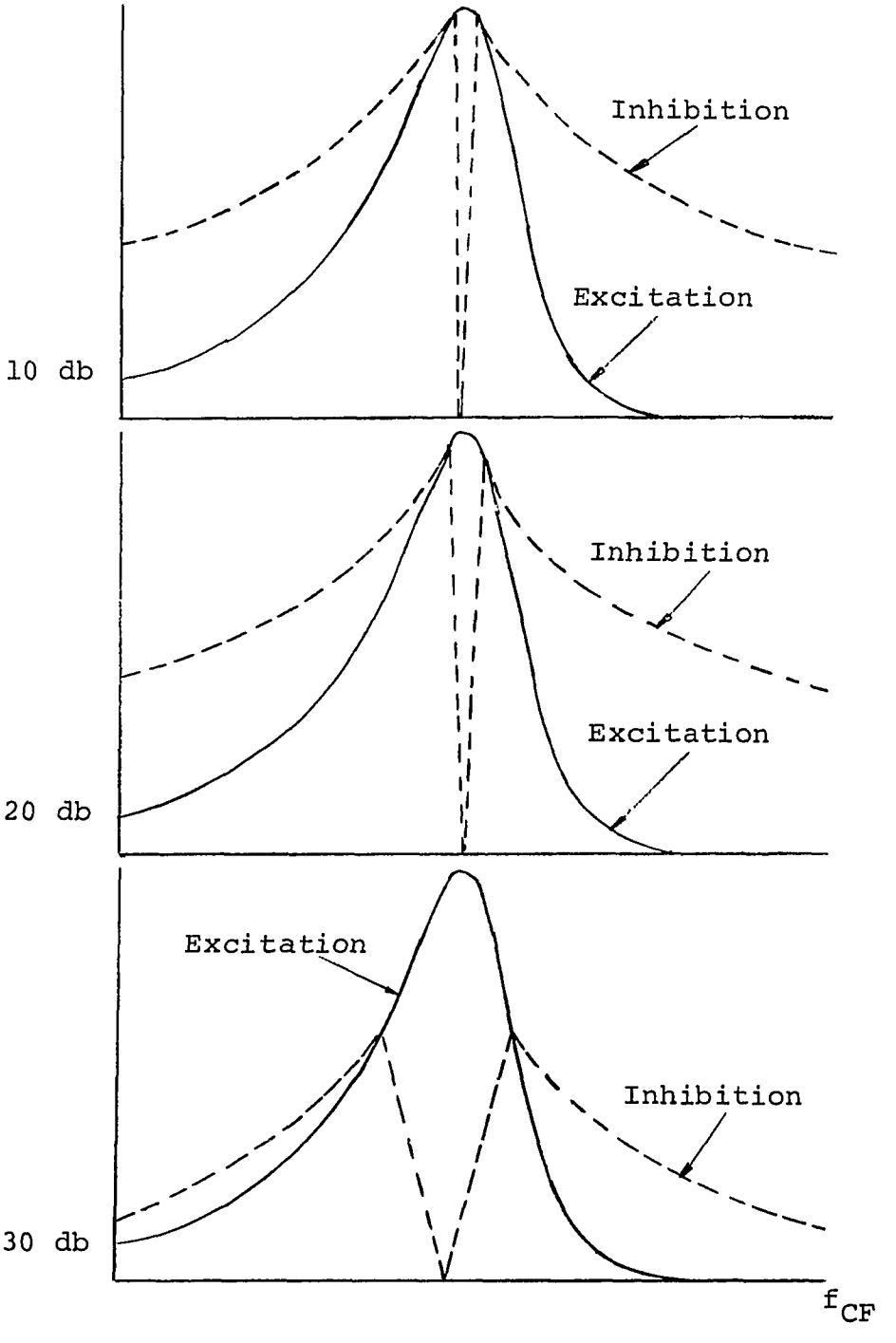


Figure 58. Centering action of the model (with $\ell_L = \ell_r = .999$ and $K_n = .35$) in response to signal levels of 10, 20, and 30 db STR.

on average nerve firing rates occurring presumedly after the system has reached a steady or "quasi" steady firing state. The steady state firing conditions of a lateral inhibitory process are dependent on the inhibited threshold time constant as discussed in section 6. An inhibitory characteristic with a fast rise time and slow decay time appears to be the most desirable.

The sequence of graphs of Figure 59 depict successive threshold distributions due to the centering action of the model as a function of time for a tone of 20 db STR and of sustained duration. The centering characteristics shown are for a system with summer gains of .999 and inhibitory time constants of the order of the time period between successive samples. In this example the system reaches a steady state firing condition in about 7 or 8 time constants. From the steady state centering characteristics the relationship between firing threshold (FT) and f_{CF} may be calculated.

$$FT\ (f_{CF}) = STR_{min}\Big|_{f_{CF=1}} \quad \text{for excitation at } f_{CF}\ . \tag{67}$$

Figure 53, curve C, shows the locus of firing thresholds produced by the above system. The firing thresholds illustrated are too high for fibers close to the frequency of excitation, but not high enough for fibers far from the frequency of excitation. The tuning curve of a 1 KHz fiber (channel) is displayed in Figure 60 together with the

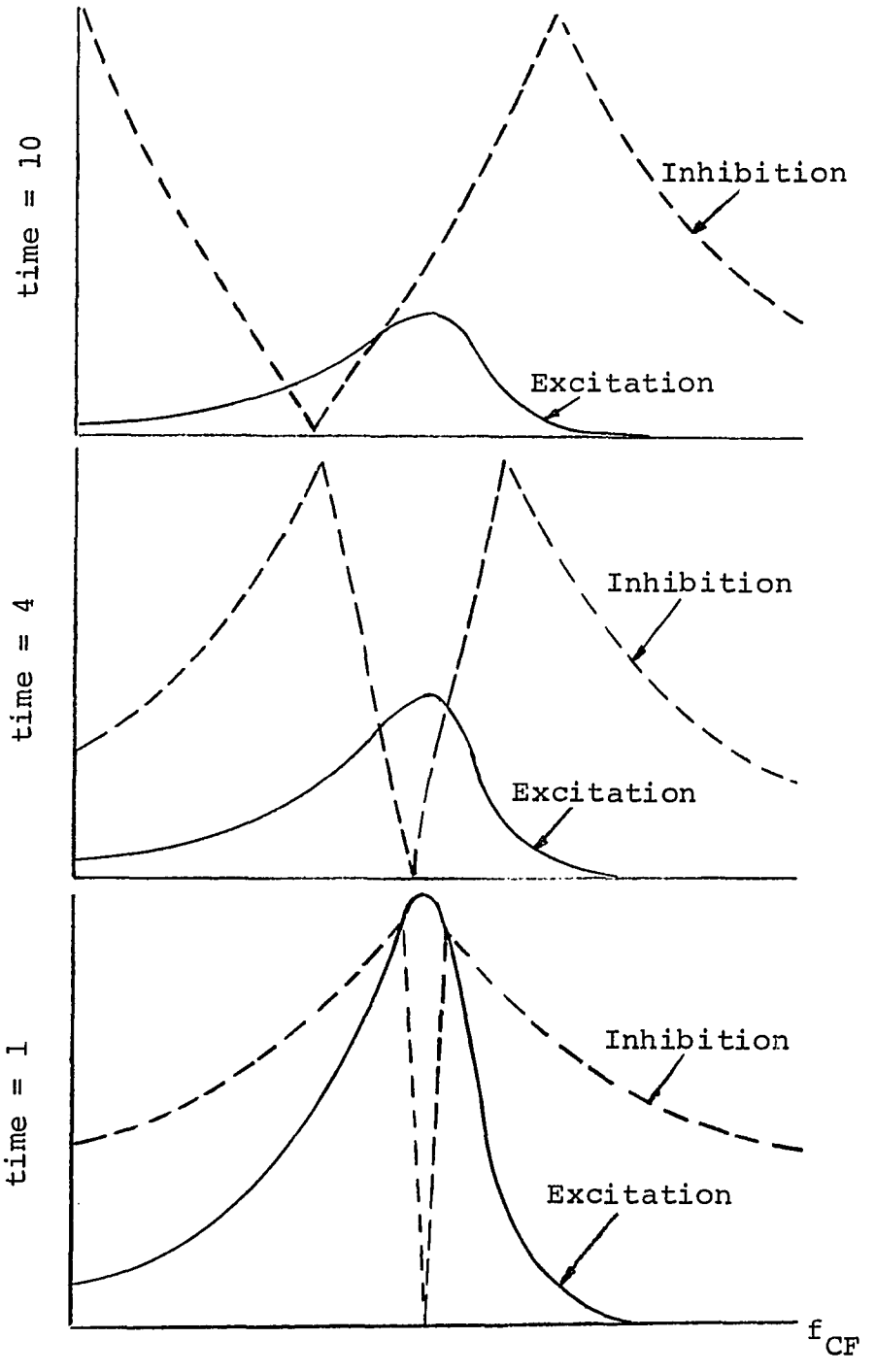


Figure 59. Successive threshold distributions of model as a function of time (recurrent cycle) for a 20 db STR tone, $\ell_L = \ell_r = .999$ and $K_n = .35$.

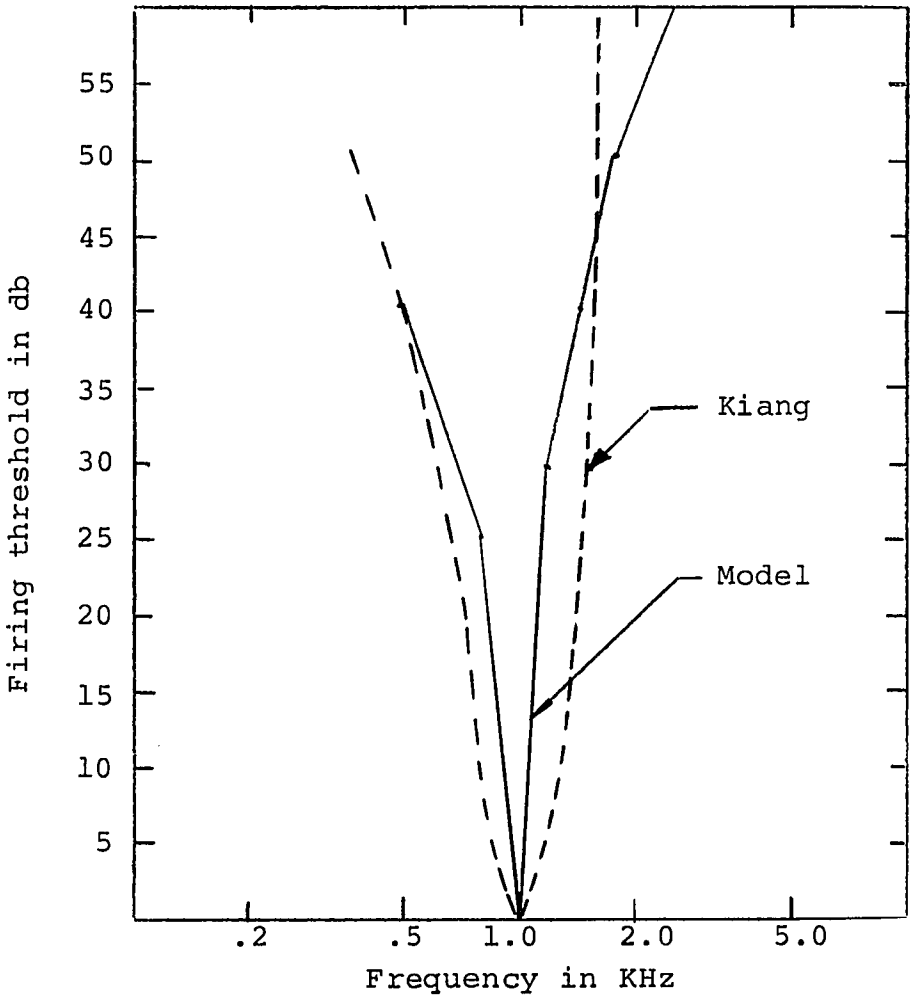


Figure 60. Tuning curve of 1.0 KHz channel of model (solid line) compared with a typical tuning curve [33] of a primary auditory fiber with CF of about 1 KHz (dotted line).

tuning curve of an approximately 1 KHz fiber observed by Kiang [33].

Differences between the firing thresholds of the model and those deduced from experimental data may be due to any of a number of simplifying assumptions. The assumption of simple "go or no go" threshold detectors eliminates the possibility of increased inhibition from a given channel at higher levels of simulation. Also, no attempt has yet been made to evaluate the effects of the (40 to 60 db) variation in absolute threshold of firing of individual neural fibers. As signal level is increased, more and more fibers (over a given segment of stimulation along the cochlear partition) should be activated. This increased number of channels activated per unit length should proportionately increase the threshold inhibition, sharpening curve C at higher levels of excitation.

By changing the distribution of signal energy among different channels, the performance of the model in response to tones of different duration was investigated. The period of time necessary for the centering process of the model to reach a stable state was in all instances assumed to be much shorter than the duration of the tone. Figure 61 shows the steady state centering of the model for a tone burst of approximately 2 msec. (at 1000 Hz). A comparison of the resultant threshold inhibition with that produced by a continuous tone (last graph in Figure 58) shows only a minor shift of the center of the activated

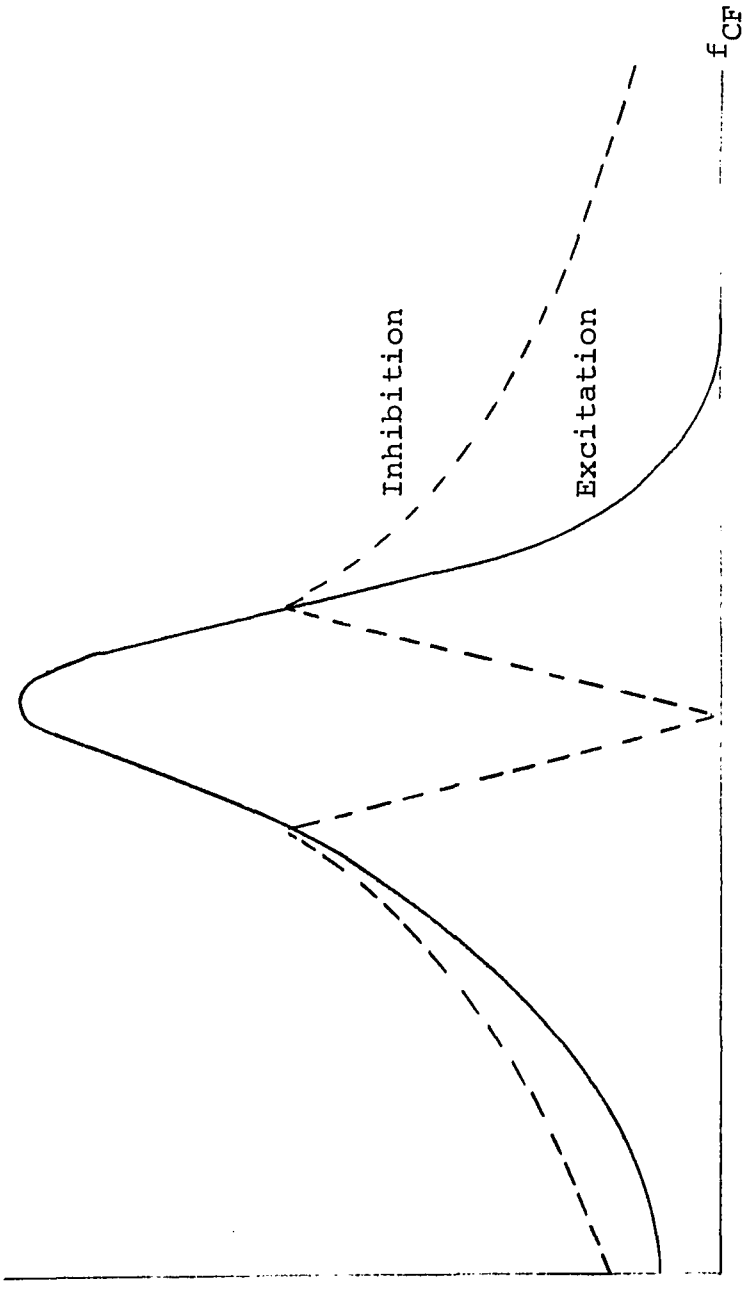


Figure 61. Centering action of the model in response to a tone burst of approximately 2 msec. duration (at 1000 Hz).

channels. Interestingly, this shift is in the same direction as observed in psycho-acoustic experiments [27] (toward higher frequencies).

The centering action of the model when excited by two-tone stimulation was also investigated. Figure 62 shows the centering action of the model in response to two continuous tones separated by .5 f_{CF} (with respect to the higher frequency) as a function of time. For this example, the rise and fall times of the threshold inhibition were assumed to be equal and of a total duration on the order of the sampling period. From the curves shown, it can be seen that the threshold distribution tends to center on the individual peaks but never quite reaches a steady state condition.

The "two-tone inhibition" created by the model, however, is qualitatively similar to that observed at the primary auditory fibers of the cat.

For tones of about the same energy level, the higher frequency tone will always dominate, causing the lower frequency tone to be inhibited. On the other hand, for the lower frequency tone to dominate the firing pattern and inhibit the higher frequency tone, it must be substantially stronger than the higher frequency tone. Again these characteristics are in agreement with the results of psycho-acoustic masking tests [27].

From the discussion of the frequency resolving ability of a lateral inhibitory process in section 6, it can be

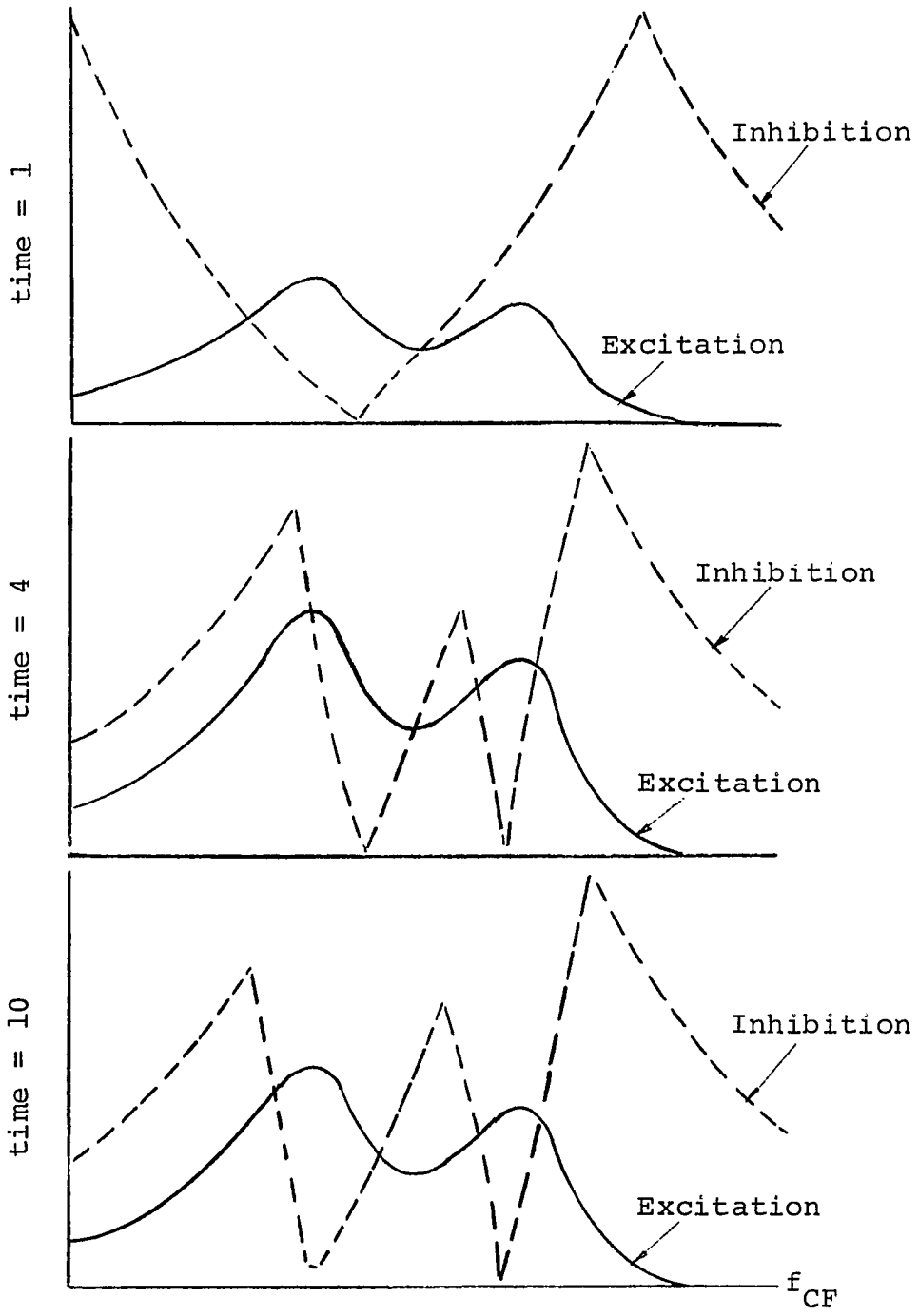


Figure 62. Successive threshold distributions of the model as function of time (recurrent cycle) in response to two continuous tones separated by .5 f_{CF} (with respect to the higher frequency), $\ell_L = \ell_r = .999$ and $K_n = .35$.

seen that a lateral inhibitory network which locates its lowest level of inhibition at the center of excitation can never "stably" resolve two formants which do not have at least one unexcited channel between their uninhibited excitation centers. This restriction suggests that the model in its present form is not adequate to explain the spectral resolving ability displayed by the ear for signals of higher STR. It is possible to conceive of a number of extensions of the present model which should not be subject to this limitation. One possible extension consists of the combination of a L.H. process in which the lowest level of inhibition tends to occur at the center of excitation, with a L.H. process in which the highest level of inhibition tends to occur at the center of excitation. If the first (L.I.C.) form of inhibition dominates at low stimulation levels while the second (M.I.C.) form of inhibition becomes significant at high levels of stimulation, this process should be able to provide improved frequency resolution without sacrificing the present model's excellent performance in the presence of noise. Such a combination of inhibitory processes might also produce increased threshold inhibition at high signal levels and thus provide an improved model in the single tone case. This promising line of investigation should furnish the basis for future research.

9. CONCLUSION

The signal processing ability of a number of different parallel channel systems was investigated. The performance of these systems was compared to the characteristics of the ear, and a model of the peripheral auditory system was developed on the basis of the most promising of these systems.

The performance of a basic P.C.S. (no overlapping of channel bandpasses) was first analyzed. The analysis was next extended to consider the effect of centering (a process by which the particular channel a spectral peak is centered is estimated on the basis of the outputs of a number of overlapping channels).

Four different types of centering processes were investigated: maximum likelihood centering which represents an "optimum performance" process; median channel centering which bases its estimate of formant frequency on the determination of the median channel of any cluster of activated channels; mean frequency centering which estimates the formant frequency to coincide with the centroid of the outputs of a number of activated channels; and lateral inhibitory centering in which the estimate of formant frequency involves raising the detection thresholds

of channels in and around the region of excitation.

The performance of these systems when excited by single and multiple formant signals in the presence of both intrinsic and extrinsic noise for a fixed false alarm rate was investigated. In particular, the signal processing abilities of these systems have been related to (1) channel SNR or SNDR, (2) maximum information processing rate, (3) overall system bandwidth, and (4) system complexity.

The above investigation revealed the following basic conclusions: In the basic P.C.S., when corrupted by intrinsic noise, no advantage is obtained by conveying information over a number of identical (noisy) channels. As the number of channels over which information is being processed is increased, the channel SNR must also be increased in order to maintain the error rate (measured in terms of $P(CD/S_K)$ at the same level achieved by a single channel system.

On the other hand, in the case of extrinsic noise (where the noise enters the system along with the signal), the analysis shows that there is a significant advantage in conveying information over a number of channels. As the number of channels over which information is being processed is increased, the SNDR necessary to maintain a given error rate decreases. As N is increased, the rate of decrement of SNRD (d(SNDR)/dN) decreases so that there is no advantage in enlarging N beyond a critical value. The critical size of N is an increasing function of the

allowable number of simultaneous formants (K). However, the system, in the presence of white normally distributed noise processes information most efficiently when excited by a single formant source.

The analysis also indicates that for a fixed information rate the basic P.C.S. displays a definite relationship between total system bandwidth and system complexity (as measured by N). As system complexity is increased, overall system bandwidth must also increase. Consequently, in order to obtain an improvement in system performance (for a fixed SNDR and maximum information rate), both system complexity and bandwidth must be increased.

On the other hand, P.C.S. with centering are not necessarily similarly restricted. The analysis indicates that, with centering, it is possible to obtain an improvement in system performance without having to increase system bandwidth. Application of centering also relaxes the requirement that a formant's frequency be near the center of a channel for detection. When the channel overlap factor, M, is large, a spectral peak has a high likelihood of detection no matter where its specific frequency is located, within the overall system bandpass.

In almost all cases, when P.C.S. with centering are corrupted with intrinsic noise, the analysis indicated that performance improves with greater overlapping of channel bandpasses. This result implies that the deterioration in system performance, due to the processing of information

over a number of internally noisy channels, can at least be partially compensated for, by using a large overlap factor.

In the presence of extrinsic noise, system performance also improves with larger values of overlap factor (but to a lesser extent than in the intrinsic noise case). A consequence of this performance improvement is that the bandwidth required to process a given amount of information in a given period of time and at a given SNDR and error rate is reduced.

The maximum likelihood type of centering (representing an "optimum performance" process) provides the best performance of all the centering processes investigated in the presence of noise. The bandwidth savings obtainable by application of a centering process to a P.C.S. can be predicted by means of the performance characteristics of P.C.S. with this form of centering process (presented in section 4). Figure 63 shows the approximate percent bandwidth reduction achievable (for a $P(CD/S_1) = .50$) by a maximum likelihood process as a function M.

The practical implementation of a maximum likelihood process is rather complex (probably requiring a small digital computer). Fortunately, the other forms of centering processes investigated are physically much easier to realize.

The M.F. centering process can be implemented by means of a number of summing and multiplying elements. Its performance is very close to that of the maximum likelihood

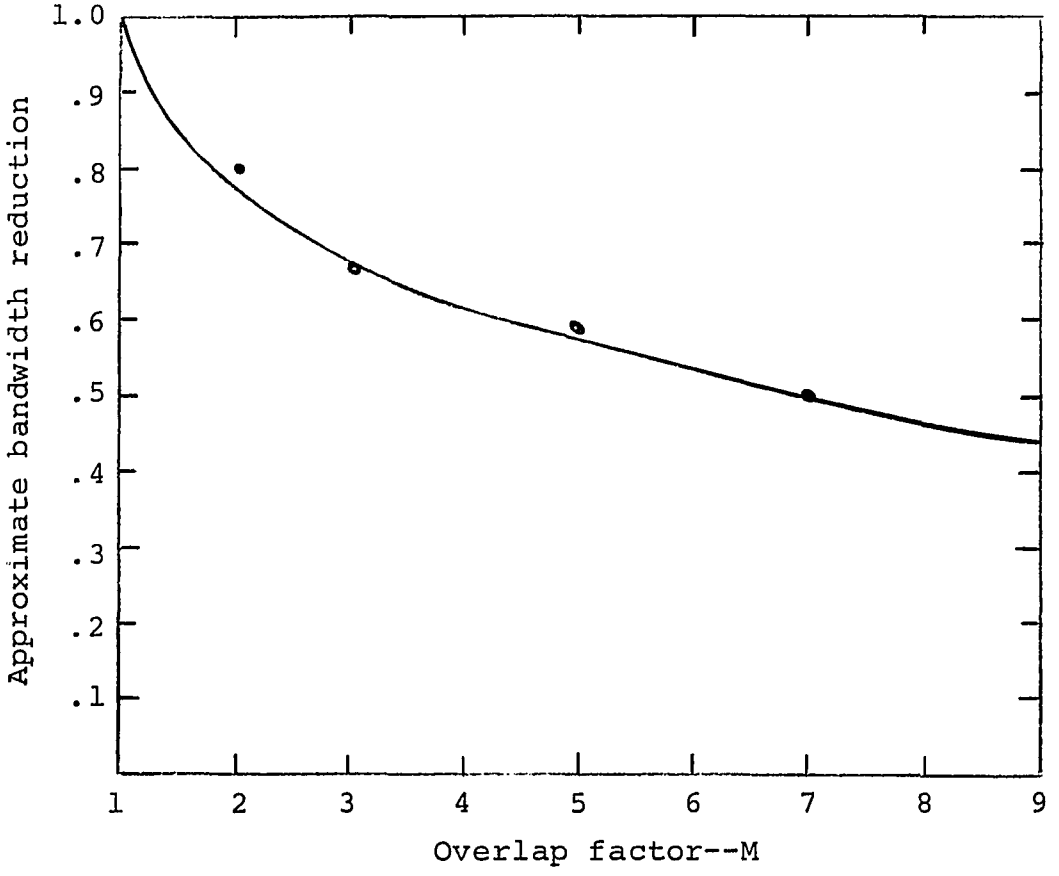


Figure 63. Approximate system bandwidth reduction achievable (for $P(CD/S_1) = .50$) by a maximum likelihood centering process as a function of overlap factor.

process in almost every respect. However, the minimum frequency separation between two simultaneous formants for their (independent) resolution must be at least 1/3 greater than with the maximum likelihood centering process.

The M.C. process can be implemented in terms of simple digital logic elements--primarily AND gates. Although the overall signal processing performance of the M.C. is slightly inferior to that of the M.F. process, its ability to resolve simultaneous formants is superior to both the M.F. and maximum likelihood centering processes. One of the interesting properties of the M.C. process is that, in the case of intrinsic noise, its $P(CD/S_K)$ does not rise monotonically with channel SNR, but displays a number of sharp dips, the depth of which increases with larger values of M. Similar characteristics are not displayed by the other centering processes although some nonmonotonic characteristics are displayed by the standard deviation of the estimated frequency (given detection) for all centering types investigated.

The major disadvantage of the M.C. process is that the number of "layers" of AND gates required for its implementation is proportional to the overlap factor, M, and, hence, the total number of system elements is proportional to $(M)^2$.

L.H. centering processes are especially attractive from the point of view of practical implementation because of the structural simplicity of the centering circuitry

based on the repeated use of relatively simple elements. The L.H process, however, does not necessarily narrow the output firing pattern to a single channel and thus in most applications would have to be followed by a centering process of another type for the final determination of the specific channel in which a formant is centered. A system consisting of a L.H. process followed by a M.C. process can be synthesized with a total number of elements proportional to M.

Of the two basic forms of L.H., recurrent and nonrecurrent, the recurrent form appears to be the most useful because of its ability to partially adapt its characteristics to the dynamics of the excitation. The performance displayed by these processes in the presence of noise is dependent on the class and amount of lateral inhibition employed. L.I.C. processes (i.e., L.H. processes where the least inhibition tends to occur at the center of excitation) provide the best performance in the presence of noise and were the only class of L.H. process investigated quantitatively. The analysis shows that the performance displayed by the L.I.C. L.H. processes in the presence of both intrinsic and extrinsic noise lies between those of the M.F. and M.C. centering processes (depending on choice of inhibitory coefficients). The optimum value of (non-zero) inhibitory coefficients vary with channel SNR. Possibly some form of L.H. process other than the one considered may yield an optimum value of inhibitory

coefficient, which is independent of the excitation level.

The analysis also shows the functional similarity between the characteristics of P.C.S. with centering and those of the peripheral auditory system. The outputs of the many overlapping channels may be compared to the excitations provided to different hair cells of the Organ of Corti by the basilar membrane. The action of the centering process may be related to the lateral interactions of the parallel neural fibers of the ear. It was shown that P.C.S., as the ear, are not restricted by the acoustic uncertainty principle.

In their behavior with respect to noise, these systems display several other similarities to the ear. The analysis shows that these systems are capable of functioning efficiently in the presence of both intrinsic and extrinsic noise and are capable of correctly estimating the frequency of a formant when its level is within 3 db of the threshold at which just the presence of the formant is first detectable.

The ear normally functions in a noisy environment. Considering intrinsic noise, the (unstimulated) primary neural channels are particularly noisy. Wiess [61] suggests that this noise may be characterized by a Gaussian noise source, which is white from about 15 Hz to 15 KHz. Yet, even at low SNR (5 db SL), the ear is capable of sharp frequency discrimination of pure tones [55]. This frequency

discrimination shows little improvement for sensation levels above 40 db (SL). Similarly, in the case of P.C.S. with centering, little or no improvement in frequency discriminating ability is achieved beyond a critical value of SNR (about 30 db for M = 7). The channel SNR corresponding to this value is a function of channel overlap factor, increasing with larger values of M. With regard to extrinsic noise, more psycho-physical data are needed on the ear's frequency discrimination in the presence of masking noise before any comparisons can be drawn. In such tests, the effect of the length of the period of observation should be more fully investigated. Although the interval of time between two successive tones is not crucial in the comparison of pitch according to Postman [44], it should be noted that the duration of the tone pulse does affect results. Not only do tone pulses of longer duration show a narrower spectral spread, but their extended duration may also allow the acoustic sensory system to perform several estimates of their center frequency. In such a case, the aggregate sample would provide a better estimate than a single sample. This effect would be especially important at low SNR and can be expected to be related to the hearing system's temporal summation period [64].

In the case of complex tones, composed of two or more spectral peaks, the ear's ability to resolve such signals into their individual components is much coarser than its ability to discriminate slow frequency changes in single

formant responses [41, 42]. Similarly, the spectral resolving ability of the centering processes investigated is much coarser than their frequency discriminating ability (this ratio for large values of M is greater than $4\Delta f$).

In this dissertation, system performance was evaluated on the basis of the probability of correctly interpreting a signal, given its presence, for a fixed false alarm condition. The ear cannot be expected to optimize its performance on the basis of the probability of correct detection since it does not normally know the a priori characteristics of the sound stimuli it receives. Performance could be optimized on the basis of a fixed probability of erroneously detecting a signal when none is present (a condition corresponding to the assumed performance criterion). Such a mechanism requires only the assumption that the system can utilize periods of time when the detection of a signal is highly unlikely.

In the light of the neural channels' adaptive characteristics, this appears a reasonable possibility at least in the case of the peripheral auditory system.

Of the several forms of centering processes investigated, the L.H. process appears the most likely centering mechanism in the peripheral auditory system. To investigate further the possible existence of L.H. centering in the peripheral auditory system, a 4000 channel model of the peripheral auditory system employing L.H. centering was simulated on the digital computer. Because of its superior

performance in the presence of noise, a recurrent L.I.C. form of L.H. process was used in the simulation.

The response of this model compared favorably in many respects with the characteristics of the peripheral auditory system observed by a number of researchers [32, 33, 36, 51]. The spectral response curves of individual channels were very similar to the neural fiber tuning curves described by Kiang [32, 33]. These responses displayed very sharp skirts for frequencies above their "resonant" frequency (CF), as well as for frequencies below (but to a lesser degree). Despite the nonsymmetrical bandpass characteristics of individual channels and their variable bandwidth, the L.H. process "adaptively" located the center of excitation within a very few (less than 10) recurrent cycles.

The distribution of firing thresholds was also qualitatively similar to data extrapolated from the experimental work of Kiang [32, 33]. For the inhibitory coefficients chosen, the firing thresholds produced by the model tend to be too high for fibers close to the exciting frequency, but not high enough for fibers far from the frequency of excitation. This deviation may be due to any of a number of simplifying assumptions. When the model was excited by tone bursts of short duration, its response did not change significantly, displaying only a minor shift of the center frequency of the activated channels. This shift, although small, was in the same direction as observed in

psycho-acoustic experiments [27], i.e., toward higher frequencies.

When the model was excited by two simultaneous tones, the inhibition produced was similar to that observed at the primary auditory fibers of the cat [51]. For tones of approximately the same energy level, the higher frequency tone will always dominate, causing the lower frequency tone to be inhibited. On the other hand, for the lower frequency tone to dominate the firing pattern and inhibit the higher frequency tone, it must be substantially stronger than the higher frequency tone. The threshold inhibition produced by the model, however, never reached a steady state condition (except for tones of very low level).

This result suggests that the model in its present form is not yet fully adequate to "explain" the spectral resolving ability display by the ear. A number of extensions of the present model exist which should not be subject to this limitation and should furnish the basis for future research.

10. RECOMMENDATIONS

In this dissertation, the performance characteristics of P.C.S. employing centering processes have been analyzed. In performing this analysis, it was necessary to restrict the conditions under which these systems functioned. Sources of noise were all assumed to be white and normally distributed. The signal sources considered were assumed to provide a uniform distribution of formant frequency combinations, with the source energy remaining constant irrespective of the number of formants present.

In practice, other forms of noise corruption and signal sources are common. In the perception of speech by the auditory system, for example, it is rarely pure white noise which represents the disturbing source. Similarly, a constant energy signal source as was considered in this dissertation is a poor model of speech [16]. Articulation is better represented by a signal source in which the signal energy is proportional to the number of formants present. The consideration of different forms of excitation should aid in evaluating the utility of P.C.S. as practical information processing systems.

In some cases, our definition of correct detection might be too restrictive. Consideration should also be

given to the manner in which signal sources are encoded. For some coding, for instance, the detection of any one of a number of simultaneous spectral peaks might convey the information, and so correspond to a correct detection--speech appears to possess such redundant characteristics [16].

In the analysis, four different types of centering processes were considered. Each of these processes displayed its own individual advantages and disadvantages. Of special interest was the relationship between a centering process's ability to resolve a spectrally complex signal into its component formants, and its ability to precisely estimate the frequency of formants which compose the signal. The identification and clarification of centering processes possessing both these abilities in various degrees should be of real value.

Of the centering processes considered, the L.H. process appears the most promising. This type of centering process can be realized in a number of different forms. The performance of only one of these forms was considered in the presence of noise. The evaluation of the performance of other forms should also be of value.

Recurrent L.H. networks can exhibit adaptive characteristics. Such properties can be of considerable value when used in conjunction with a P.C.S., whose channels have nonsymmetrical filter responses. The effect of noise on these adaptive properties was only touched upon in this

dissertation and should represent an important area of future analysis.

In performing this analysis, the threshold characteristics of the L.H. process should not be limited to the "all or nothing" characteristics considered here. Characteristics such as the "linear or nothing" characteristics suggested by Furman and Frishkopf [20] should also be investigated.

Of the many possible extensions of the work presented here, the most interesting are probably in the area of applications. Ten years ago the use of a multitude of parallel systems (each of possibly limited reliability) for processing signals might have seemed highly impractical, but today with the rapid development in integrated and large-scale integrated circuits, such a concept seems technologically feasible, if not an eventual necessity.

The direct application of P.C.S. with centering to information transmission, as in digital signaling, is unlikely, as much of this form of communication is done over a coherent channel, where signal phase is known exactly [54]. However, P.C.S. with centering should be of value in the area of signal identification--as required in speech processing, radio astronomy, cardiology, or even FM detection, etc.

Another application of P.C.S. is in the development of prosthetic devices for the sensory handicapped. The body appears to process sensory information in a parallel form.

The expanded knowledge of the characteristics of P.C.S. should aid in the search for substitute sensory channels.

In the latter part of this dissertation, a P.C.S. with L.H. centering is used in a model of the peripheral auditory system. In particular, this model does not adequately describe the spectral resolving ability of the ear. There are a number of possible extensions and refinements of this model, which should enable it to match more closely the characteristics of the peripheral auditory system. It should also be of interest to investigate the performance of the present model in the presence of both signal and noise.

The concepts upon which the model of the auditory system investigated here are based are also applicable to other biological sensory processes, such as vision and touch. The employment of P.C.S. with centering in models of these processes offers still another interesting avenue of research.

APPENDIX A

CALCULATION OF THE DISTRIBUTION OF ESTIMATED FREQUENCY OF P.C.S. WITH CENTERING IN THE PRESENCE OF INTRINSIC NOISE

The probability distributions of estimated frequency in the presence of intrinsic noise of P.C.S. with the following forms of centering processes have been calculated by digital computer techniques:

1. Maximum Likelihood Centering
2. Median Channel Centering
3. Mean Frequency Centering
4. Lateral Inhibitory Centering

In all cases, the computations took the general form outlined in Figure A-1. All possible firing patterns were generated. In most instances, this generation took the form of consecutively converting the decimal numbers 1 through 2^{N-1} into their binary form. The probability of each of these firing patterns, $P(Y_i/f_S)$, was calculated by taking the product of the probabilities of the individual channels firing or not firing [equation (24)]. A decision was next made (depending on the particular centering process under consideration) associating a particular frequency of excitation, f_E, with the given firing pattern.

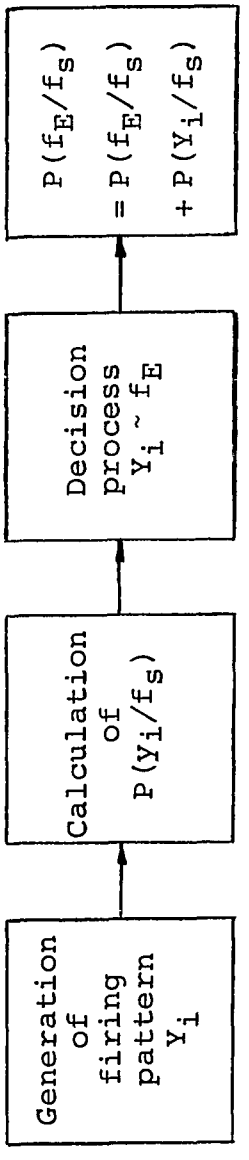


Figure A-1. Machine computation of $P(f_E/f_S)$ in intrinsic noise.

The probability of occurrence of that firing pattern was then assigned to the probability of the corresponding frequency, $P(f_E/f_s)$, and the process repeated for the next firing pattern.

In performing the actual machine computation, only the firing patterns of the NC channels which receive energy from the given exciting formant (f_s) were considered. The firing of any other channel was assumed to correspond to an error in detection.

The decision process used for the maximum likelihood process is outlined in Figure A-2a. The operations necessary to associate a particular f_E with a given firing pattern (Y_i) are more complex for this centering process than for those of the other centering processes, since this decision depends on the probabilities of occurrence of the given Y_i pattern provided by all possible exciting formants. Consequently, it was necessary to calculate the probability of occurrence of the given firing pattern for all possible input formant frequencies within the NC channels common to the given exciting formant. (Other input formant frequencies have a negligibly small probability of exciting any of the firing patterns considered.) The values of these probabilities, $P(Y_i/f)$, were compared to each other and the largest determined. The probability of occurrence of the given firing pattern, $P(Y_i/f_s)$, was then assigned to input frequency achieving the highest probability.

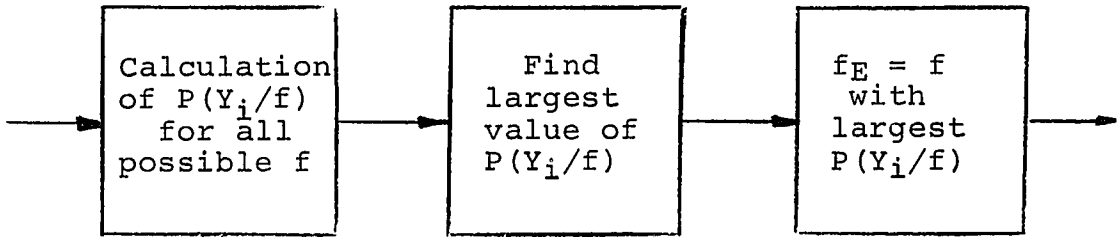


Figure A-2a. Maximum likely decision process.

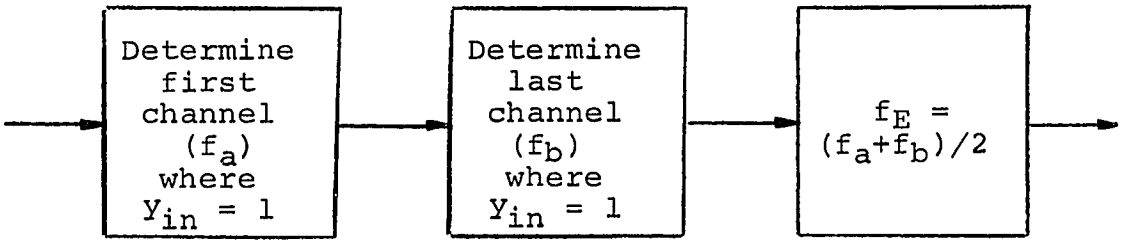


Figure A-2b. Median channel decision process.

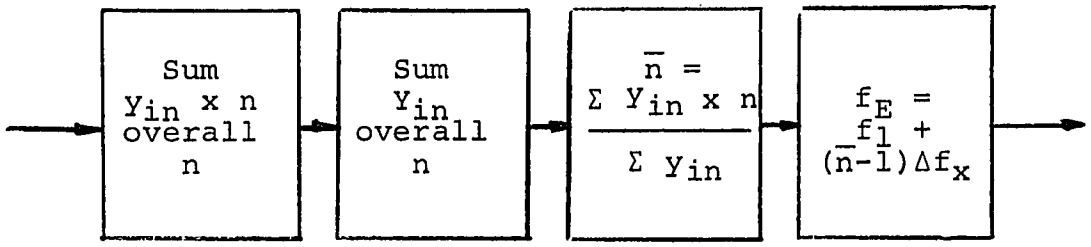


Figure A-2c. Mean frequency decision process.

In the case of M.C. centering, any firing pattern containing a "hole" as for example:

$$Y_I = 1\ 1\ 1\ 0\ 1$$

or

$$Y_I = 1\ 0\ 1\ 0\ 1 \qquad \text{(A-1)}$$

will be erroneously interpreted as two or more separate formants and constitute an error. Hence, it was only necessary to initially generate those firing patterns which do not contain holes [see equations (37) and (38) in section 4]. When the ensemble of possible firing patterns is thus limited, the decision process for the M.C. centering process simply involves determining the median frequency of the given firing pattern as shown in Figure A-2b. The $P(Y_i/f_s)$ was then associated with this frequency.

The decision process for the M.F. centering process is also easily implemented on the digital computer. The first moment of the firing vector Y_I is calculated as outlined in Figure A-2c. Since the number of possible input frequencies is limited, the resultant first moment is simply associated with the closest allowable input frequency and the $P(Y_I/f_s)$ assigned to this frequency.

The decision process used for the L.H. centering process is somewhat different from those used for the other cases. Immediately after a firing pattern has been generated by an L.H. process (depending on the type of L.H.

process), the distribution of detector-thresholds (V_{th}) is modified and a new firing pattern (Y_I*) produced. A large number of new firing patterns are possible (depending on Y_I). In order to evaluate the performance of an L.H. process, the estimated frequency which corresponds to each of these new firing patterns must be considered, and the probability of occurrence of each of these patterns, $P(Y_i, Y_i^*)$, associated with each f_E.

The decision process used for evaluating the performance of L.H. processes is shown in block form in Figure A-3. The threshold inhibition (generated in response to Y_i) was first calculated. For the recurrent L.H. processes of the L.I.C. class, the new distribution of detector thresholds can be expressed as

$$V_{Tn} = V_{th} + \ell_{nm} \left| \sum_{m=n-NC}^{n-1} Y_{im} - \sum_{m=n+1}^{n+NC} Y_{im} \right| , \qquad (A-2)$$

if the lateral control signals inhibit those channels which share common spectra with the channels from which they emanate. ℓ_{nm} is the recurrent inhibitory coefficient [see equation (49) in section 6]. A set of new firing patterns were then generated. This set included only those firing patterns which contain no "holes" (since the existence of a hole indicates the presence of more than one spectral peak), and were generated by the same means that the initial set of firing patterns for the M.C. centering process were generated. Each of these Y_i* were next tested to see

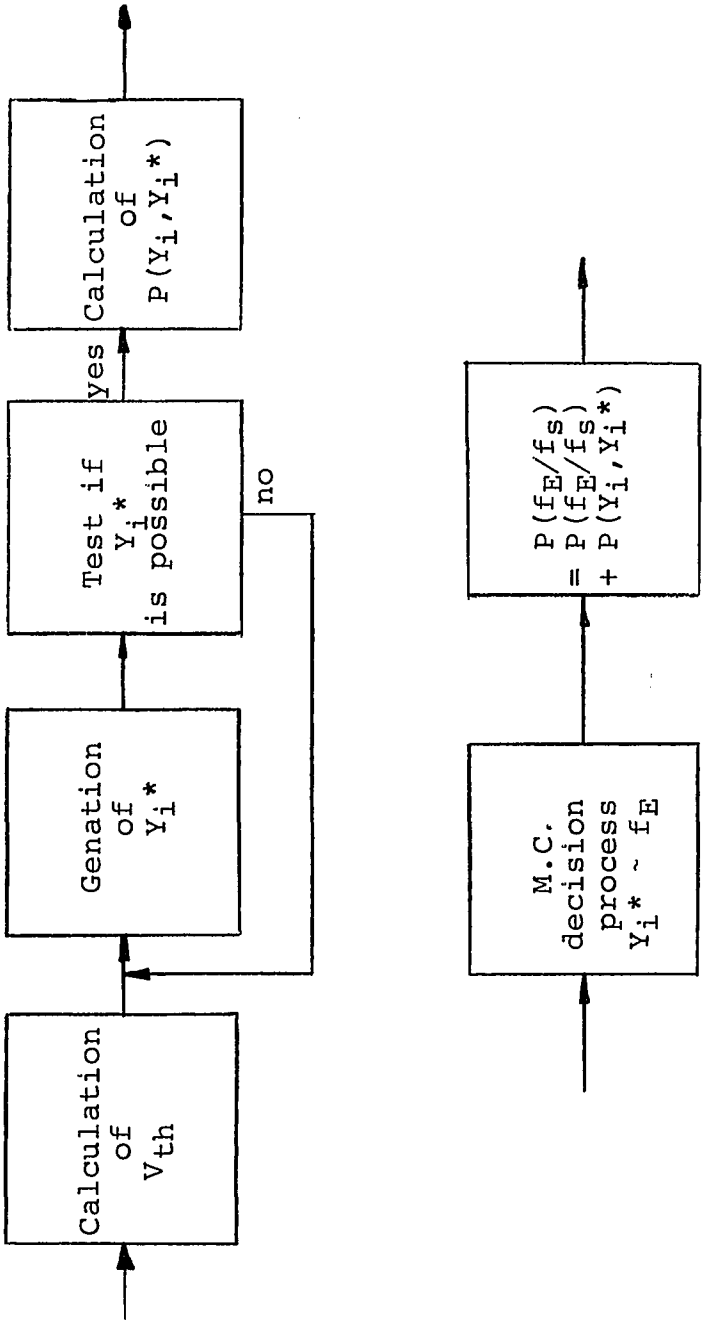


Figure A-3. Lateral inhibitory decision process.

if they represent a possible new firing pattern; for example, if

$$Y_{in} = 0 \; ; \; Y_{in}^* \neq 1 \; . \tag{A-3}$$

If Y_i^* did represent a possible new firing pattern, the joint probability of its occurrence and that of the initial firing pattern, $P(Y_i, Y_i^*)$, were calculated [equation (53) in section 6] on the base of the new distribution of detector thresholds (V_{Tn}). The frequency corresponding to Y_i^* was then determined by means of the M.C. decision process shown in Figure A-1c and $P(Y_i, Y_i^*)$ associated with this frequency, i.e.,

$$\sum_{\text{all } Y_i^*} P(Y_i, Y_i^*) = P(Y_i/f_s) \; . \tag{A-4}$$

APPENDIX B

PROBABILITY OF OCCURRENCE OF FIRING PATTERN Y_i IN THE CASE OF EXTRINSIC NOISE

In the presence of extrinsic noise, the conditional probability of occurrence of a firing pattern Y_i given the presence of a formant of frequency f_s may be expressed as

$$P(Y_i/f_s) = X_1 \cdot \prod_{n-2}^{N-1} X_n \cdot \prod_{n=N}^{N+M-1} R_n \cdot X_N \, . \qquad \text{(B-1)}$$

The X_1 and X_N terms represent the conditional probability that the envelope of the input to a boundary (edge) channel of the set of N overlapping adjacent channels exceeds V_{th}, if this channel is excited ($y_{in} = 1$), and does not exceed V_{th}, if it is not excited ($y_{in} = 0$), given the firing pattern of the other channels. These two factors may be expressed in terms of their input signal energy ($Z_{f,n}$), a Gaussian noise energy term uncorrelated with any other channel (σ^2/M), plus M - 1 noise terms (correlated with other channels) of known amplitude (A_n) and random phase as follows:

$$X_1 = \int_{V_1=V_{th}}^{\infty} V_1 \int_{r=0}^{\infty} r\, J_o(V_1 r)\, J_o(\sqrt{2Z_{f,1}}\, r\sigma^2)$$

$$J_o(A_1 r) \cdot \cdot \cdot J_o(A_{1+M-2} r) e^{-\sigma^2 r^2/2M} dr dV_1$$

$$\text{if } y_{i1} = 1$$

$$X_1 = \int_{V_1=0}^{V_{th}} V_1 \int_{r=0}^{\infty} r\, J_o(V_1 r)\, J_o(\sqrt{2Z_{f,1}}\, r\sigma^2)$$

$$J_o(A_1 r) \cdot \cdot \cdot J_o(A_{1+M-2} r) e^{-\sigma^2 r^2/2M} dr dV_1$$

$$\text{if } y_{i1} = 0 \qquad \text{(B-2)}$$

and

$$X_N = \int_{V_N=V_{th}}^{\infty} V_N \int_{r=0}^{\infty} r\, J_o(V_N r)\, J_o(\sqrt{2Z_{f,N}}\, r\sigma^2)$$

$$J_o(A_{N-1} r) \cdot \cdot \cdot J_o(A_{N+M-3} r) e^{-\sigma^2 r^2/2M} dr dV_N$$

$$\text{if } y_{iN} = 1$$

$$X_N = \int_{V_N=0}^{V_{th}} V_N \int_{r=0}^{\infty} r\, J_o(V_N r)\, J_o(\sqrt{2Z_{f,N}}\, r\sigma^2)$$

$$J_o(A_{N-1} r) \cdot \cdot \cdot J_o(A_{N+M-3} r) e^{-\sigma^2 r^2/2M} dr dV_N$$

$$\text{if } y_{iN} = 0 \ . \qquad \text{(B-3)}$$

The other X terms each represent the conditional probability that the envelope of the input to one of the interior channels of the set of overlapping adjacent channels exceeds V_{th}, if the channel is excited, and does not

exceed V_{th}, if it is not, given the firing pattern of the other channels. Since, for interior channels, there is no portion of the spectrum which is not correlated with another channel, these X factors may be expressed in terms of their input signal energy ($Z_{f,n}$) plus M noise terms of known amplitude and random phase as follows:

$$X_n = \int_{V_n=V_{th}}^{\infty} V_n \int_{r=0}^{\infty} r\, J_o(V_n r)\, J_o(\sqrt{2Z_{f,n}}r\sigma^2)$$

$$J_o(A_n r) \cdot \cdot \cdot J_o(A_{n+M-1} r)\, dr dV_n$$

$$\text{if } y_{in} = 1$$

$$X_n = \int_{V_n=0}^{V_{th}} V_n \int_{r=0}^{\infty} r\, J_o(V_n r)\, J_o(\sqrt{2Z_{f,n}}r\sigma^2)$$

$$J_o(A_n r) \cdot \cdot \cdot J_o(A_{n+M-1} r)\, dr dV_n$$

$$\text{if } y_{in} = 0 \,. \qquad \text{(B-4)}$$

The R_n factors each represent the integral of a single noise term's amplitude (A_n) probability density function evaluated over all possible values of amplitude. Individual noise terms can be represented by independent narrow bands of Gaussian noise of power level σ^2/M. Their amplitude probability can thus be expressed in terms of the Rayleigh distribution as follows:

$$R_n = \int_{A_n=0}^{\infty} (MA_n/\sigma^2) e^{-MA_n^2/2\sigma^2} dA_n \,. \qquad \text{(B-5)}$$

APPENDIX C

CALCULATION OF THE PROBABILITY DISTRIBUTION OF THE ESTIMATED FREQUENCY OF A P.C.S. WITH CENTERING IN THE PRESENCE OF EXTRINSIC NOISE

The probability distribution of estimated frequency in the presence of extrinsic noise of a P.C.S. with the following forms of centering have been investigated by means of Monte Carlo simulation on the digital computer:

1. Maximum Likelihood Centering
2. Median Channel Centering
3. Mean Frequency Centering
4. Lateral Inhibitory Centering

In all cases, the simulation can be thought of as taking the general form outlined in Figure C-1.

Sample inputs for each of the NC channels which receive energy from the given exciting formant f_s are first generated. Since the noise energy in any channel is common to a number of the channels (see Figure 10b), the noise spectrum (of the NC channels) is divided into M + NC - 1 independent segments. The noise intensity in each of these segments (CN_L) can be represented by a random vector whose amplitude is normally distributed and phase displays a uniform distribution. Each of these

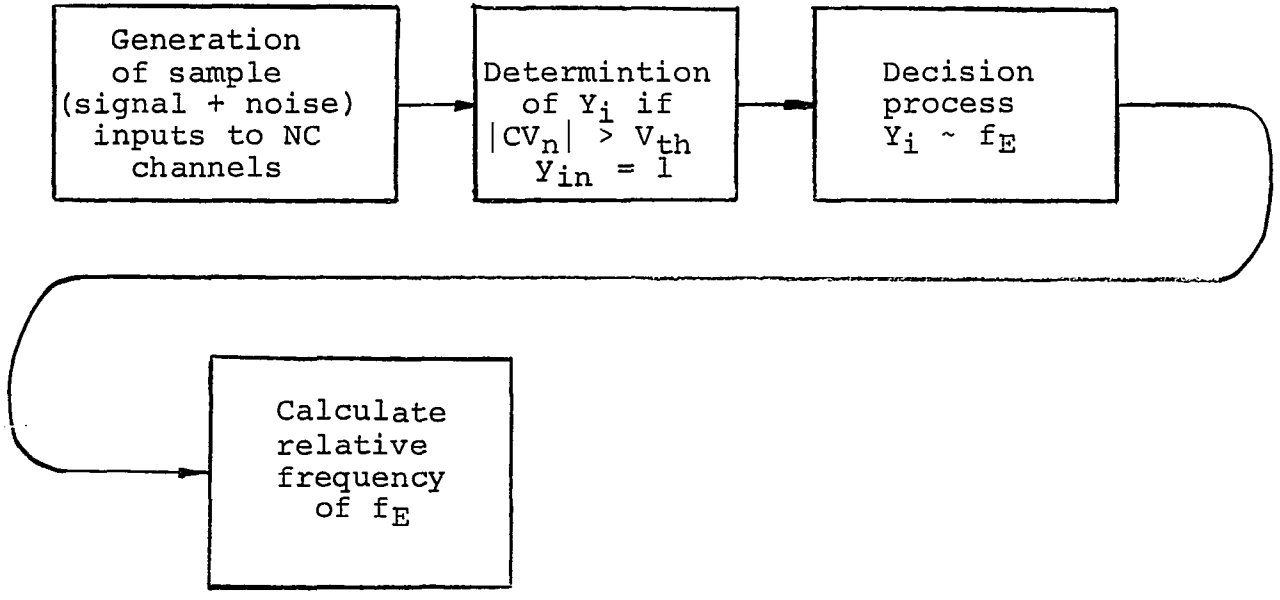


Figure C-1. Machine computation of $P(f_E/f_s)$ in presence of extrinsic noise by means of Monte Carlo simulation.

vectors is constructed from two independent Gaussian variables (XN,YN) of variance 1/M (generated by means of a normally distributed random number subroutine [4]) such that

$$CN_L = XN + jYN \ . \qquad (C\text{-}1)$$

The sample input to each channel (CV_n) is represented by the complex sum of M - 1 of these vectors plus a vector (CS_n) representing the signal intensity in each channel.

$$CS_n = (2Z_{f,n})^{1/2} + jO \qquad (C\text{-}2)$$

and

$$CV_n = CS_n + \sum_{L=n}^{n=M-1} CN_L \ . \qquad (C\text{-}3)$$

The firing pattern corresponding to a given set of sample inputs is determined by comparing the magnitude of CV_n with the detector thresholds;

$$\text{if} \quad |CV_n| > V_{th}/\sigma \qquad y_{in} = 1$$

$$\text{if} \quad |CV_n| \leq V_{th}/\sigma \qquad y_{in} = 0 \ .$$

Because of the correlation of noise between channels, V_{th} can no longer be simply related to the false alarm rate. The probability of no channel firing (given the absence of a signal) can be expressed in terms of the joint probability of the individual channels not firing:

$$1 - RA = P(y_{i1} = 0, y_{i2} = 0, \ . \ . \ ., y_{iN} = 0) \ . \qquad (C\text{-}4)$$

Although this joint probability cannot be expressed in terms of the product of the individual channels not firing, it can be expressed in terms of a product of conditional probabilities:

$$P(y_{i1} = 0, y_{i2} = 0, \ldots, y_{iN} = 0) = P(y_{i1} = 0) \cdot P(y_{i2} = 0/y_{i1} = 0)\ P(Y_{i3} = 0/y_{i2} = 0, y_{i1} = 0) \times \cdots P(y_{iN} = 0/y_{iN-1} = 0, y_{iN-2} = 0, \ldots, y_{i1} = 0) . \quad \text{(C-5)}$$

The probability that channel n does not fire ($y_{in} = 0$) now depends on that of channel n - m not firing ($y_{in-m} = 0$). This dependence is a function of the number of channels between channel n and channel n - m, i.e.,

$$P(y_{in} = 0/y_{in-1} = 0, y_{in-2} = 0, \ldots, y_{i1} = 0) \approx P(y_{in} = 0/y_{in-1} = 0, y_{in-2} = 0, \ldots, y_{i2} = 0) \approx \ldots \approx P(y_{in} = 0/y_{in-1} = 0, y_{in-2} = 0, \ldots, y_{in-m} = 0) . \quad \text{(C-6)}$$

Because of the configuration of the overlapping of the bandpasses of a P.C.S. (see Figure 10b), knowing that channel n - 2 does not fire provides very little new information once it is known that channel n - 1 did not fire. Consequently, the conditional probability of the nth channel not firing, given that all channels less than n did not

fire, can be approximated as

$$P(y_{in} = 0/y_{n-1} = 0,\ y_{n-2} = 0,\ \ldots,\ y_{i1}) \approx P(y_{in} = 0/y_{in-1} = 0)\ , \tag{C-7}$$

and

$$P(y_{i1} = 0,\ y_{i2} = 0,\ \ldots,\ y_{in} = 0) \approx P(y_{i1} = 0)\ P(y_{i2} = 0/y_{i1} = 0) \cdot P(y_{i3} = 0/y_{i2} = 0) \cdot \cdots P(y_{in} = 0/y_{in-1} = 0)\ . \tag{C-8}$$

For the symmetrical P.C.S. systems being considered,

$$P(y_{i2} = 0/y_{i1} = 0) = P(y_{i3} = 0/y_{i2} = 0) = \ldots = P(y_{in} = 0/y_{n-1} = 0) \equiv q_{n/n-1}\ . \tag{C-9}$$

Then, when no signal is present,

$$P(y_{i1} = 0,\ y_{i2} = 0,\ \ldots,\ y_{iN} = 0) \approx q(q_{n/n-1})^{N-1}\ . \tag{C-10}$$

q can be evaluated by means of equation (9) in section 3.

$$q = 1 - e^{-(b^2/2)} = 1 - Q(0,b)\ , \tag{C-11}$$

where b is the normalized detector threshold (V_{th}/σ). $q_{n/n-1}$ can be evaluated by means of equation (30) (discussed in Appendix B) by using the relation:

$$q_{n/n-1} = q_{n,n-1}/q\ , \tag{C-12}$$

where

$$q_{n,n-1} = P(y_{in} = 0,\ y_{in-1} = 0) \qquad \text{(C-13)}$$

when no signal is present. Referring to Figure C-2, if σ^2 represents the noise power input to each channel, then the noise power common to adjacent channels is

$$\sigma_c^2 = (1 - 1/M)\sigma^2 \ , \qquad \text{(C-14)}$$

where M is the overlap factor. The uncorrelated noise power in each channel is

$$\sigma_1^2 = \sigma_2^2 = \sigma^2/M \ . \qquad \text{(C-15)}$$

By equation (30) for the case of two overlapping channels:

$$q_{n,n-1} = \int_{Ac=0}^{\infty} \int_{V_1=0}^{V_{th}} \int_{V_2=0}^{V_{th}}$$

$$W\left(\frac{A_c}{\sigma_1}, \frac{V_1}{\sigma_1}\right) W\left(\frac{A_c}{\sigma_2}, \frac{V_2}{\sigma_2}\right) \frac{A_c}{\sigma_c^2}\, e^{-(A_c/2\sigma_c)^2}$$

$$dV_1 dV_2 dA_c \ , \qquad \text{(C-16)}$$

where

$$W(a,V) = \exp\left(-\frac{a^2 - V^2}{2}\right) I_0(aV)VdV \qquad \text{(C-17)}$$

and is the integrand of the Q function. Letting

$$S = A_c/\sigma_1 = A_c/\sigma_2 \ , \qquad \text{(C-18)}$$

$$S/(M - 1)^{1/2} = A_c/\sigma_c \ , \qquad \text{(C-19)}$$

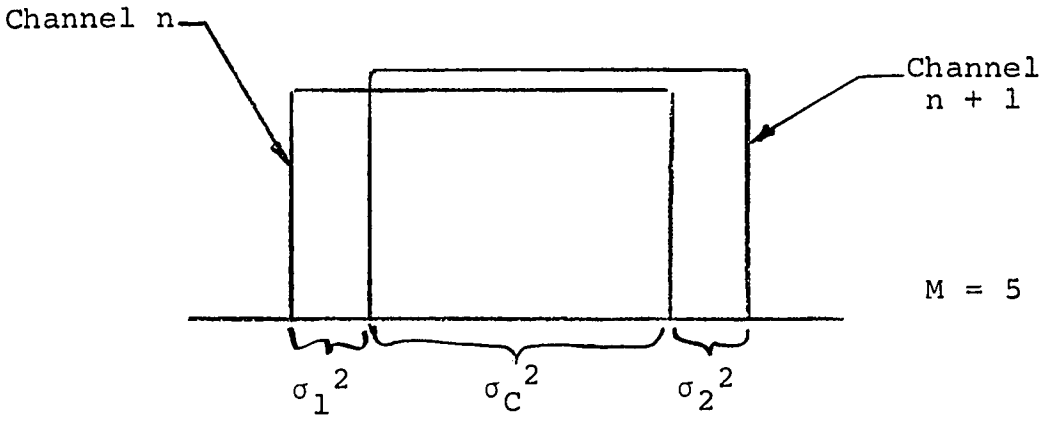


Figure C-2. Relationship of extrinsic noise energies of two adjacent channels.

and

$$\frac{V_{th}}{\sigma_1} = \frac{V_{th}}{\sigma_2} = \sqrt{M}\,\frac{V_{th}}{\sigma} = \sqrt{M}b \ ; \qquad \text{(C-20)}$$

and recognizing that the integration with respect to V_1 is independent of the integration with respect to V_2, equation (C-16) can be rewritten in terms of the Q function:

$$q_{n,n-1} = \int_{S=0}^{\infty} [1 - Q(S,\sqrt{M}b)]^2 \, \frac{Se^{-S^2/2(M-1)}}{(M-1)^{1/2}} \, dS \ , \qquad \text{(C-21)}$$

since

$$Q_1(S,\sqrt{M}b) = Q_2(S,\sqrt{M}b) \ .$$

and

$$q_{n/n-1} = \{(M-1)^{1/2}\,[1 - Q(0,b)]\}^{-1} \int_{S=0}^{\infty} [1 - Q(S,\sqrt{M}b)]^2 \, Se^{-S^2/2(M-1)} dS \ . \qquad \text{(C-22)}$$

From this last equation, the relationship between V_{th} and RA may be determined directly.

When N >> NC, this last expression may also be used to approximate the probability of all channels which do not receive signal energy from f_s not firing (PQ); in which case

$$PQ \approx q(q_{n/n-1})^{N-NC-1} \ . \qquad \text{(C-23)}$$

Once Y_i is determined, a decision is made relating the resultant Y_i to an estimated input frequency f_E. In the

case of the MC, MF, and LH centering processes, the decision mechanisms are essentially the same as those used in evaluating $P(f_E/f_s)$ in the case of intrinsic noise and have already been described in Appendix A. The relative frequencies of the resultant f_E's for a large number of sets of inputs are recorded. From these relative frequencies, the $P(f_E/f_s)$ distribution is deduced.

In the case of maximum likelihood centering, the decision process is similar to that of the intrinsic noise case. However, for this process the association of f_E with Y_i is based on the probabilities that Y_i be generated by each of all possible input frequencies (in the presence of extrinsic noise). To determine these probabilities, using Monte Carlo techniques requires that the ensemble of $P(Y_I/f_s)$ first be calculated. This ensemble can be calculated by generating a large number of sample inputs, and from these inputs determining the corresponding Y_i's. Instead of immediately associating with each Y_i an f_E (as was done in the case of the other centering processes), the relative frequency of the different Y_i patterns is first recorded. Since the P.C.S. being considered are symmetrical, the probabilities of generating Y_i for frequencies other than f_s may be determined from the $P(Y_I/f_s)$ distribution.

Each of the Y_i vectors can be represented in terms of its components as a 1 x N matrix. As each component y_{in} can assume a value of either 0 or 1 (see section 4), this

array has the appearance of an N-digit binary number. This number is expressed by

$$[Y_i] = \sum_{n=1}^{N} (y_{in})^{2^{n-1}} , \qquad \text{(C-24)}$$

and uniquely determines Y_i. In the following we shall designate Y_i by this number $[Y_i]$. This permits easy handling of the pattern vectors by means of computers. The probability of generating firing pattern Y_i by a formant centered in a channel s + m can be expressed as:

$$P(Y_i/f_{s+m}) = P([Y_i] \cdot 2^{-m}/f_s) . \qquad \text{(C-25)}$$

Using this procedure, the input frequency having the highest likelihood of generating a given firing pattern may be determined, and the relative frequency of each possible f_E becomes equal to the sum of the relative frequencies of all Y_i's which are associated with f_E.

APPENDIX D

COMPUTER SIMULATION OF THE PERIPHERAL AUDITORY SYSTEM

The characteristics of the P.C.S. model of the peripheral auditory system, illustrated in Figure 45b, have been investigated for a number of different excitatory and inhibitory conditions, by means of simulation on the digital computer. In all cases, the simulation took the general form outlined in Figure D-1. Since the P.C.S. simulated consists of 4000 channels (each having spectral properties corresponding to a different point along the cochlear partition of the ear), it was necessary to first determine the normalized amplitude

$$SA_n = V_{in}/V_{th} \qquad (D\text{-}1)$$

of the assumed input signal in each channel. This operation was accomplished by calculating the spectral distribution of the input, normalized with respect to V_{th}, $S_K(j\omega)$ (usually assumed to be a narrow spectral peak of cosine squared energy distribution), taking the square of the magnitude of the product of $S_K(j\omega)$ and the transfer characteristics of the individual channels, $H_n(j\omega)$ [see equation (58)], integrating with respect to $d\omega$, and finally

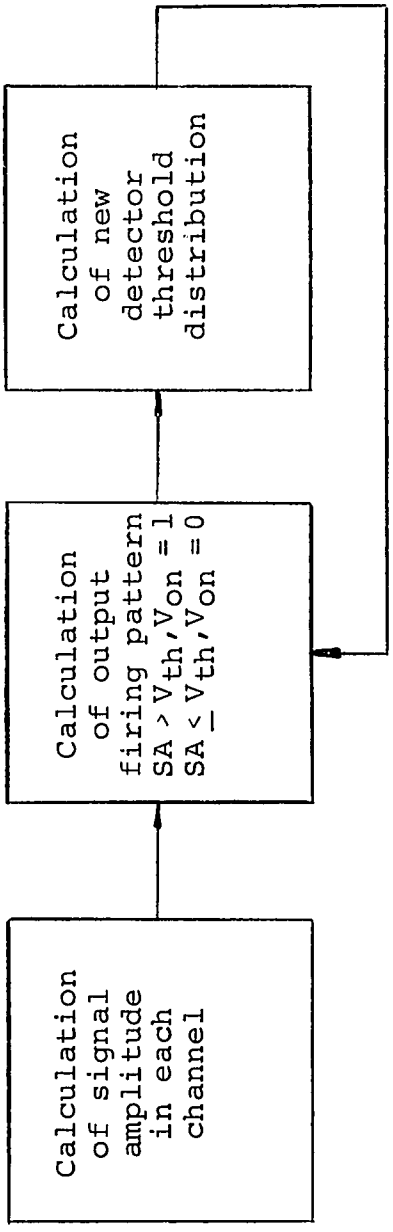


Figure D-1. Machine simulation P.C.S. model of peripheral auditory system with L.I.C. lateral inhibition.

taking the square root.

$$SA_n = \left\{ \int \left| S_K(j\omega)\ H(j\omega) \right|^2 d\omega \right\}^{1/2} \qquad \text{(D-2)}$$

SA_n was then compared to the detector thresholds, V_{Tn}, in each channel. The output, V_{on}, was assumed

$$V_{on} = 1 \text{ if } SA_n > V_{Tn}$$

$$V_{on} = 0 \text{ if } SA_n \geq V_{Tn} \ . \qquad \text{(D-3)}$$

In this manner the output firing pattern of the entire system was determined.

From the output firing pattern, the modification of the detector threshold distribution, due to the lateral inhibition may be calculated. The L.H. process described in Figure 56 (of section 8) was chosen for the simulation. For this process the detector thresholds are related to the output firing pattern by equations (63), (64), and (66). The summing operations described by these equations are easily carried out on the computer.

Once the change in detector threshold distribution has been calculated, the corresponding output pattern can be recalculated and the change in threshold inhibition determined for the new output firing pattern. This process was repeated (in the simulation) until the system achieved a stable output pattern or some repetitive firing sequence was observed. The effect of an inhibitory time constant (τ_i) was investigated by not allowing V_{Tn} to change

instantaneously (change totally between two samples). Because of the many channels (4000), it was not practical to display the total output pattern. As a compromise, the firing condition of every 40th channel was displayed for each sample period.

APPENDIX E

COMPUTER PROGRAM LISTINGS

The listings of several of the more important computer programs used in this dissertation are reproduced in this appendix. The programs listed here include:

MAXCEN--Evaluation of the performance of a maximum likeli-centering process in the presence of intrinsic noise.

CORLAT--Evaluation of the performance of a lateral inhibitory centering process in the presence of extrinsic noise.

CORMAX--Evaluation of the performance of a maximum likelihood centering process in the presence of extrinsic noise.

BASCEN--Simulation of the peripheral auditory system.

```
      PROGRAM MAXCEN
C     EVALUATION OF PERFORMANCE OF MAXIMUM LIKELIHOOD
C     CENTERING PROCESS
C     IN INTRINSIC NOISE
      COMMON N,AN,M,AM,NA,NC,VT,PCST
      DIMENSION Q(40), IP(20), PCO(40),PT(40)
    1 READ 2,N,M,LC
C     N=NUMBER OF CHANNEL, M=OVERLAP FACTOR
    2 FORMAT (3I4)
C     FA=FALSE ALARM RATE
      FA=.02
      NC=2*M-1
      AN=N
      LC=NC
C     Z=CHANNEL SNR
      Z=4.0
      PQ=(1.0-FA)**(1.0/AN)
C     VT= DETECTOR THRESHOLD
      VT=VTH(PQ)
    3 PRINT 4,Z
    4 FORMAT(1H1,3H Z=F12.7)
      NT=2*NC+1
C     CALC. OF ENERGY AND PROB. OF EXCITATION
      NAG=0.0
C     NAG=TEST FOR EQUI-LIKELY INPUTS
      DO 5 L=1,NT
      PCO(L)=0.0
      PF=PWF(NC,L)
C     PF=SIGNAL ENERGY IN N CHANNEL
      ZA=PF*Z
    5 Q(L)=QF(ZA)
C     Q( )=PROB. OF N CHANNEL BEING EXCITED
C     QF( )=Q FUNCTION SUBROUTEEN
      Q(NT+1)=QF(0.0)
      NTT=NT+1
      NN=2**NC-1
C     GENERATION OF FIRING MATRIX IP( )
      DO 30 J=1,NN
      NB=J
      DO 6 I=1,NC
      IP(I)=0
      IF(NB-2*(NB/2).EQ.1)  IP(I)=1
    6 NB=NB/2
      PY=0.0
C     DET. OF MOST LIKELY INPUT JF OF ALL POSSIBLE INPUTS NF
      DO 23 NF=1,NT
      PT(NF)=1.0
      DO 9 I=1,NC
      L=2*I+2*M-NF
      IF(L.LE.0) L=NT+1
      IF(L.GT.NT) L=NT+1.
```

```
      Q1=1.0-Q(L)
      IF(IP(I).LT.1) QT=Q(L)
    9 P1(NF)=PT(NF)*Q1
      DO 20 K=1,M
      NF=NF
      IF(NF.GT.2*M) NF=NT-NF
      L=2**K-2**(K-1)+2-NF
      IF(L.LT.0) L=NT+1
   20 P1(NF)=PT(NF)*(1-Q(L))
      IF (P1(NF)-PY) 23,22,21
   22 AAG=AAG+1.0
      GO TO 23
   21 PY=PT(NF)
      IF=NF
      AAG=0.0
   23 CONTINUE
      IF(AAG.EQ.0.0) GO TO 25
      DO 24 NF=1,NT
      IF(PY.GT.PT(NF))  GO TO 24
      PCD(NF)=P1(2*M)/(AAG+1.0) +PCD(NF)
   24 CONTINUE
      GO TO 30
   25 PCD(IF)=P1(2*M)+PCD(IF)
   30 CONTINUE
      PDET=0.0
C     CALC. OF PROB. DIST. OF ESTIMATED FREQ. PCD( )
C     PDET=OVERALL PROB. OF DET. A FORMANT
      DO 40 L=1,NT
      PCD(L)=PCD(L)*PQ**(N-M-NC)
      PDET=PDET+PCD(L)
      L2=L
      CL=L
      F=CL/2.0-0.5
      PRINT 42,NC,N,M,L2,F,PCD(L2)
   42 FORMAT(4H NC=I5,3H N=I5,3H M=I5,4H L2=I5,5H PCD(F5.2,
     2,2H)=F12.7)
   40 CONTINUE
C     GENERATION OF GRAPHICAL OUTPUT
      PRINT 71
   71 FORMAT(1H1,///,'      PROBABILITY DISTRIBUTION OF F')
      PRINT 72
   72 FORMAT(1H1,18,'P(F)='T17,'0'T26,'.1'T36,'.2'T46,'.3'T
     256,'.4'T66,'.5'T76,'.6'T86,'.7'T96,'.8'T106,'.9'T116,
     2'1.'
      DO 73 I1=1,131
   73 CALL PLOT(I1,'_N')
      CALL PLOT(132,'_ ')
      FF=0.0
      FF2=0.0
      DO 60 L2=1,NT
      CL=L2
```

```
      F=CL/2.0-0.5
      FN=F/AMC
      IF(PDET)90,90,91
   90 PFN=0.0
      GO TO 92
   91 PFN=PCD(I2)/PDET
   92 IPF=PFN*100+10
      CALL PLOT(10,'FN',IPF,'*N',110,'F ')
      PRINT 75,FN,PCD(I2)
   75 FORMAT(1H1,5H F/N=F10.6,10X,F10.6)
      YF=FN*PFN
      YF2=FN**2*PFN
      EF=EF+XF
C     EF=FIRST MOMENT OF EST. FREQ.
   80 EF2=EF2+YF2
C     EF2=SECOND MOMENT OF EST. FREQ.
      PRINT 81
   81 FORMAT(10X,/)
      PRINT 82,N,Z,PDET,EF,EF2
   82 FORMAT(5H N=I4,5H Z=F7.2,6H PDET=F12.7,7H E(FN)=F12.7
     2,9H VAR(FN)=F12.7)
   50 IF(LC-1)1,52,52
   52 STOP
      END
*EOF
LIST
```

```
      PROGRAM CORLAT
C     CENTERING IN THE PRESENCE OF EXTRINSIC NOISE
C     EVALUATION OF PERFORMANCE OF LATERAL INHIBITORY
C     BY MONTY CARLO SIMULATION
      COMPLEX CN,CV
      DIMENSION PX(40), PCD(40), CN(60), IP(40)
      DIMENSION LSUM(30), RSUM(30), IP(30), VN(30)
C     N=NUMBER OF CHANNEL, M=OVERLAP FACTOR
      N=7
      AN=N
      NC=2*N-1
      ANC=NC
C     VI=RELATIVE LEVEL OF INHIBITION
      VI=1
C     SUM=NORMALIZED STANDARD DEVIATION OF NOISE
      SUM=1.0/SQRT(AM)
      NN=N+NC-1
      NT=2*NC+1
      M=100
C     PA=FALSE ALARM RATE
      PA=.02
      PQ=1.0-PA
C     VT= DETECTOR THRESHOLD
      VT=4.0121
      PC=1.0-EXP(-VT**2/2.0)
C     PQ0= PROBABILITY OF CHANNELS WHICH RECEIVE NO SIGNAL
C     BUT FIRING
      PQ0=PC*(PQ/PC)**(1.0-NC/(N-1.0))
C     IN=INITIAL VALUE FOR RANDOM NOISE SUBROUTEEN
      IN=786395481
C     Z=CHANNEL SNR
      Z=20.0
C     CALC. OF ENERGY IN NC CHANNELS
      DO 3 I=1,NC
C     PX=SIGNAL ENERGY IN CHANNEL N
    3 PX(I)=PWF(NC,I)
    2 DO 6 I=1,NT
    6 PCD(I)=0.0
      PDET=0.0
C     PCD(INF)=PROBABILITY OF INPUT NF
C     PDET=OVERALL PROB. OF DET. A FORMANT
      IMF=1
      PRINT 20,N,M,Z
   20 FORMAT(101,3H N=I4,3H M=I4,3H Z=F7.4)
C     GENERATION OF RANDOM INPUTS (SIGNAL + NOISE)
    4 DO 5 J=1,NN
C     NORMAL=NORMALLY DIST. RANDOM VAR. SUBROUTEEN
      CALL NORMAL (IN,SUM,XN,YN)
C     XN AND YN=NORMALLY DISTRIBUTED RANDOM NUMBERS
C     CN( )=COMPLEX NOISE VOLTAGE IN INDEP. BAND
    5 CN(J)=CMPLX(XN,YN)
```

```
      SUM=0.0
      SUM=0.0
      TT=0
      T1=0
      TM(2)=NC
      DO 50 L=1,NC
      ZP=PX(L)*Z
      ZS=SQRT(2*ZP)
C     CV=COMPLEX CHANNEL IN PUT (CORRELATED SIGNAL + NOISE)
      CV=CMPLX(ZS,0.0)
      DO 9 K=1,M
      KL=L-1+K
    9 CV=CV+CN(KL)
C     VN=MAGNITUDE OF CV
      VN(L)=CABS(CV)
C     IP(N)=FIRING CONDITION IN CHANNEL N (I.E. 1 OR 0)
      IP(L)=0
   58 IF(VN(L).GT.VT) IP(L)=1
C     GENERATION OF INHIBITION IN
      DO 8 L=1,NC
      SUM=SUM+IP(L)
      LSUM(L)=SUM
      LR=NC+1-L
      SUM=SUM+IP(LR)
    8 RSUM(LR)=SUM
C     DETERMINATION OF NEW FIRING PATTERN RESULTING FROM
C     INHIBITION
      DO 10 L=1,NC
      TH=ABS(LSUM(L)-RSUM(L))
      IX=0
      IF(VN(L).GT.K1*TH+VT) IX=1
C     TT=TEST FOR THE PRESENCE OF MORE THAN ONE FORMANT
      IF(IX-TT)11,10,12
   11 T1=JT+1
      TM(JT)=L
      TT=0
      GO TO 10
   12 T1=JT+1
      TM(JT)=L
      TT=1
   10 CONTINUE
C     ESTIMATION OF EXCITATION FREQ. (IC) BY MEDIAN PROCESS
      IF(JT.EQ.0) GO TO 15
      IF(JT.GT.2) GO TO 15
      IC=TM(1)+TM(2)-1
      IF(JT.EQ.1) IC=TM(1)+NC
      PCD(IC)=PCD(IC)+1.0
   15 IF(TME-2500)16,17,17
   16 TME=TME+1
      GO TO 4
   17 AME=TME
```

```
C       CALC. OF DIST. OF ESTIMATED FREQ.   PCD( )
        DO 40 L=1,NT
        PCD(L)=PCD(L)/AME*PWO
        PDET=PDET+PCD(L)
        L2=L
        CL=L
        F=CL/2.0-0.5
        PRINT 42,NC,N,M,L2,F,PCD(L2)
   42   FORMAT(4H NC=I5,3H N=I5,3H M=I5,4H L2=I5,5H PCD(F5.2,
       2?H)=F12.7)
   40   CONTINUE
C       GENERATION OF GRAPHICAL OUTPUT
        PRINT 71
   71   FORMAT('0',///,'      PROBABILITY DISTRIBUTION OF F')
        PRINT 72
   72   FORMAT('0',18,'P(F)='17,'0'T26,'.1'T36,'.2'T46,'.3'T
       2?56,'.4'166,'.5'176,'.6'186,'.7'T96,'.8'T106,'.9'T116,
       2'1.'
        DO 73 I1=1,131
   73   CALL PLOT(I1,'-N')
        CALL PLOT(132,'- ')
        EF=0.0
        EF2=0.0
        DO 80 L2=1,NT
        CL=L2
        F=CL/2.0-0.5
        FN=F/AMC
        IF(PDET)90,90,91
   90   PFN=0.0
        GO TO 92
   91   PFN=PCD(L2)/PDET
   92   IPF=PFN*100+16
        CALL PLOT(16,'|N',IPF,'*N',116,'| ')
        PRINT 75,FN,PCD(L2)
   75   FORMAT('+',5H F/N=F10.6,102X,F10.6)
        XF=FN*PFN
        XF2=FN**2*PFN
        EF=EF+XF
C       EF=FIRST MOMENT OF EST. FREQ.
   80   EF2=EF2+XF2
C       EF2=SECOND MOMENT OF EST. FREQ.
        PRINT 81
   81   FORMAT('0',/)
        PRINT 82,N,Z,PDET,EF,EF2
   82   FORMAT(3H N=I4,3H Z=F7.2,6H PDET=F12.7,7H E(FN)=F12.7
       2,9H VAR(FN)=F12.7)
        IF(Z-10.0)44,45,45
   44   Z=Z+2.0
        GO TO 2
   45   IF(Z-100.0)46,50,50.
   46   Z=Z+10.0
```

```
GO TO 2
50 STOP
END
*EOF
LIST
```

```
      PROGRAM CURMAX
C     EVALUATION OF PERFORMANCE OF MAXIMUM LIKELIHOOD
C     CENTERING PROCESS
C     IN EXTRINSIC NOISE
C     BY MONTY CARLO SIMULATION
      COMPLEX CN, CV
      DIMENSION PX(60), ICD(10500,2), PCD(60), CN(60),
     2 PT(60)
      INTEGER*2 ICD
C     N=NUMBER OF CHANNEL, M=OVERLAP FACTOR
      N=7
      AN=N
      NC=2*N-1
      ANC=NC
C     SGM=NORMALIZED STANDARD DEVIATION OF NOISE
      SGM=1.0/SQRT(AM)
      NN=N+NC-1
      M=100
C     PA=FALSE ALARM RATE
      PA=.02
      PQ=(1.0-PA)
C     VT= DETECTOR THRESHOLD
      VT=4.0121
      PC=1.0-EXP(-VT**2/2.0)
C     PQQ= PROBABILITY OF CHANNELS WHICH RECEIVE NO SIGNAL
C     NOT FIRING
      PQQ=PC*(PQ/PC)**(1.0-NC/(N-1.0))
C     IN=INITIAL VALUE FOR RANDOM NOISE SUBROUTEEN
      IN=963852147
C     Z=CHANNEL SNR
      Z=8.0
      NT=2*NC+1
C     CALC. OF ENERGY IN NC CHANNELS
      DO 3 L=1,NT
C     PX=SIGNAL ENERGY IN CHANNEL N
    3 PX(L)=PWF(NC,L)
C     GENERATION OF ENSEMBLE OF FIRING VECTORS ICD( , )
C     ID=1 IF SIGNAL  AT CHANNEL CENTER, ID=2 IF SIGNAL
C     INBETWEEN CHANNELS
      NL=2**NC-1
    2 DO 13 NB=1,NL
      DO 13 ID=1,2
   13 ICD(NB,ID)=0
      DO 14 NF=1,NT
   14 PCD(NF)=0.0
      PDET=0.0
C     PCD(INF)=PROBABILITY OF INPUT NF
C     PDET=OVERALL PROB. OF DET. A FORMANT
      IME=1
      PRINT 20,N,M,Z
   20 FORMAT(101,3H N=I4,3H M=I4,3H Z=F7.4)
```

```
      4 DO 15 ID=1,2
        NB=0
        DO 5 J=1,NN
C       NORMAL=NORMALLY DIST. RANDOM VAR. SUBROUTEEN
        CALL NORMAL (IN,SGM,XN,YN)
C       XN AND YN=NORMALLY DISTRIBUTED RANDOM NUMBERS
C       CN( )=COMPLEX NOISE VOLTAGE IN INDEP. BAND
      5 CN(J)=CMPLX(XN,YN)
        DO 10 L=1,NC
        LX=2*L-1+ID
        ZP=PY(LX)*Z
        ZS=SQRT(2*ZP)
C       CV=COMPLEX CHANNEL IN PUT (CORRELATED SIGNAL + NOISE)
        CV=CMPLX(ZS,0.0)
        DO 9 K=1,M
        KL=L-1+K
      9 CV=CV+CN(KL)
C       VN=MAGNITUDE OF CV
        VN=CABS(CV)
        JX=L-1
C       DETERMINATION OF FIRING PATTERN REP. BY NB
     10 IF(VN.GT.VT) NB=NB+2**JX
        IF(NB.EQ.0) GO TO 15
        ICD(NB,ID)=ICD(NB,ID)+1
     15 CONTINUE
        IF(TME-2500)16,17,17
     16 TME=TME+1
        GO TO 4
     17 AME=TME
C       DETERMINATION OF MOST LIKELY INPUT FREQ. OF CORR. TO
C       PATTERN NB (OF ALL POSSIBLE INPUTS NF)
        DO 30 NB=1,NL
        IF(ICD(NB,1).GT.0) GO TO 18
        IF(ICD(NB,2).GT.0) GO TO 18
        GO TO 30
     18 PY=0.0
        AAC=0.0
C       AAC=TEST FOR EQUI-LIKELY INPUTS
        DO23 NF=1,NT
        ID=1
        NFF=(NF-1)/2
        IF(2*NFF.EQ.NF-1) ID=2
        ANB=NB
        ANX=ANB*2.0**(NC-M-NFF)
        NX=ANX
        NXX=100.0*ANX
        IF(NX.LE.0) GO TO 19
        IF(NXX.GT.100*NX) GO TO 19
        IF(NX.GT.NL) GO TO 19
        PT(NF)=ICD(NX,ID)
        GO TO 200
```

```
      19 PT(NF)=0.0
     200 IF(PT(NF)-PY) 23,22,21
      22 AAG=AAG+1.0
         GO TO 23
      21 PY=PT(NF)
         JF=NF
         AAG=0.0
      23 CONTINUE
         IF(AAG.EQ.0.0) GO TO 25
         DO 24 NF=1,NT
         IF(PY.GT.PT(NF))  GO TO 24
         PCD(NF)=PT(2*M)/(AAG+1.0) +PCD(NF)
      24 CONTINUE
         GO TO 30
      25 PCD(JF)=PT(2*M)+PCD(JF)
      30 CONTINUE
C        CALC. OF PROB. DIST. OF ESTIMATED FREQ. PCD( )
         DO 40 L=1,NT
         PCD(L)=PCD(L)/AMF*PWO
         PDET=PDET+PCD(L)
         L2=L
         CL=L
         F=CL/2.0-0.5
         PRINT 42,NC,N,M,L2,F,PCD(L2)
      42 FORMAT(4H NC=I5,3H N=I5,3H M=I5,4H L2=I5,5H PCD(F5.2,
        2PH)=F12.7)
      40 CONTINUE
C        GENERATION OF GRAPHICAL OUTPUT
         PRINT 71
      71 FORMAT('0',//,'      PROBABILITY DISTRIBUTION OF F')
         PRINT 72
      72 FORMAT('0',18,'P(F)='T17,'0'T26,'.1'T36,'.2'T46,'.3'T
        256,'.4'T66,'.5'T76,'.6'T86,'.7'T96,'.8'T106,'.9'T116,
        2'1.'
         DO 73 I1=1,131
      73 CALL PLOT(I1,'-N')
         CALL PLOT(132,'-')
         PF=0.0
         PF2=0.0
         DO 80 L2=1,NT
         CL=L2
         F=CL/2.0-0.5
         FN=F/ANC
         IF(PDET)90,90,91
      90 PFN=0.0
         GO TO 92
      91 PFN=PCD(L2)/PDET
      92 IPF=PFN*100+16
         CALL PLOT(10,'IN',IPF,'*N',116,'I ')
         PRINT 75,FN,PCD(L2)
      75 FORMAT('+',5H F/N=F10.6,102X,F10.6)
```

```
      XF=FN*PFN
      XF2=FN**2*PFN
      EF=EF+XF
C     EF=FIRST MOMENT OF EST. FREQ.
   80 EF2=EF2+XF2
C     EF2=SECOND MOMENT OF EST. FREQ.
      PRINT 81
   81 FORMAT(10I,/)
      PRINT 82,N,Z,PDET,EF,EF2
   82 FORMAT(3H N=I4,3H Z=F7.2,6H PDET=F12.7,7H E(FN)=F12.7
     2,9H VAR(FN)=F12.7)
      IF(Z-10.0)44,45,45
   44 Z=Z+2.0
      GO TO 2
   45 IF(Z-20.0)46,50,50
   46 Z=Z+10.0
      GO TO 2
   50 STOP
      END
*EOF
F LIST
```

```
      PROGRAM BASCEN
C     VERSION III
C     SIMULATION OF CENTERING OF BASILAR MEMBRANE
C     FOR COS SQUARE SPECTRAL PULSE SIGNAL
      FUNC(X)=10.0**((2000.0-X)/2000.0)
      INTEGER*2 IP(4000)
      REAL*4 LSUM(4000), RSUM(4000), INH(4000), JNH(4000)
     2 PW(4000)
      COMMON DF,F0,F1
C     DF=SPECTRAL WIDTH OF EXCITING PULSE (NORMALIZED)
C     F0=CENTER FREQ. OF EXCITING PULSE, F1=CENTER FREQ OF
C     (OPTIONAL) SECOND EXCITING FORMANT
C     N=TOTAL NUMBER OF (FIBERS) CHANNELS
      DF=.001
      F0=1.0
      N=4000
C     STH=DET. THRESHOLD LEVEL (NO INHIBITION)
      STH=1.0
      DO 250 L=1,N
      X=L
  250 PW(L)=PWF(X)
C     PWF( )= FUNCTION RELATING AMPLITUDE OF VIBRATION
C     ALONG BASILAR MEMBRANE (FOR GIVEN INPUTS) TO POSI-
C     TION  X (CHANNEL NUMBER)
    1 READ 150,E,EC,LC
  150 FORMAT(2F7.4,I3)
C     E AND EC ARE INHIBITIARY CONSTANTS, EC DET. MAGNITUDE
C     OF INHIBITION (INH), EC DET. SPACIAL DECAY (=SUMMER
C     CATN LEVELS)
C     S=INPUT SIGNAL LEVEL, DB=SINAL LEVEL IN DB
      DB=15.0
    2 S=10.0**(DB/20.0)
      TME=0
C     TME INDICATES RECURRENT CYCLE NUMBER (TIME)
      PRINT 90
   90 FORMAT(1H1)
      DO 200 L=1,N
  200 INH(L)=0.0
    3 JAG=0
C     GENERATION OF (RECURRENT) FIRING PATTERN
      DO 15 L=1,N
      SA=PW(L)*S
C     IP( )=FIBER FIRING CONDITION (1 OR 0)
      IP(L)=0
      IF(SA.GT.STH+(EC*INH(L))) IP(L)=1
      IF(JAG.EQ.1) GO TO 9
      IF(IP(L).EQ.0) GO TO 15
      JAG=1
      GO TO 11
    9 IF(IP(L).EQ.1) GO TO 15
      JAG=0
```

```
   11 Y=L
      FX=FUNC(Y)
C     FUNC(X)=FUNCTION RELATING POSITION ALONG MEMBRANE (X)
C     TO PEEK CHARACTERISTIC FREQ. (F)
      PRINT 12,L,FX,IP(L)
   12 FORMAT(3H L=I5,3H F=F10.4,4H IP=I4)
   15 CONTINUE
      SUM=0.0
      SUM=0.0
C     GENERATION OF INHIBITION INH
      DO 20 I=1,N
      SUM=F*(SUM+IP(L))
      LSUM(L)=SUM
      LK=N+1-L
      SUM=F*(SUM+IP(LK))
   20 RSUM(LR)=SUM
      AX=0.01
C     AX=VARIBLE USED IN NORMALIZING GRAPHICL OUTPUT
      DO 22 L=1,N
      INH(L)=ABS(LSUM(L)-RSUM(L))
      IF(INH(L)-AX)22,22,21
   21 AX=INH(L)
   22 CONTINUE
      IMF=IMF+1
      AX=AX*EC+1.0
C     GENERATION OF GRAPHICAL OUTPUT
      PRINT 30,IME,S,DR,AX
   30 FORMAT(7H TIMES=I3,6H SIGNAL=F6.2,4H DR=F6.2,5H MAX=F
     212.5)
      PRINT 71
   71 FORMAT('0',!                                    INHIBITION!)
      PRINT 72
   72 FORMAT('0',T8,'P(F)='T17,'0'T26,'.1'T36,'.2'T46,'.3'T
     256,'.4'T66,'.5'T76,'.6'T86,'.7'T96,'.8'T106,'.9'T116,
     2'1.'
      DO 73 I1=1,131
   73 CALL PLOT(I1,'-N')
      CALL PLOT(132,'-')
      IF(S .GT. AX) AX=S
      DO 80 L=1,100
      LP=L*40
      HIP=(INH(LP)*EC+1.0)/AX
      IH=HIP*100+16
      SA=(PW(LP)*S)
      SN=SA/AX
      IS=SN*100+16
      Y=LP
      F=FUNC(X)
      CALL PLOT(16,'IN',IH,'*N',IS,'XN',116,'|')
      XNH=INH(LP)*EC+1.0
   80 PRINT 75,LP,F,XNH,SA
```

```
      75 FORMAT(1H1,15,3H F=F7.3,102X,2F7.3)
      31 IF(IME-10)3,32,32
      32 IF(UP-15.0)33,44,44
      33 UP=UP+5.0
         GO TO 2
      44 IF(LC-1)1,45,45
      45 STOP
         END
*EOF
F LIST
```

REFERENCES

[1]. Békésy, G. V. (1963). "Hearing Theories and Complex Sounds," J. Acoust. Soc. Amer., 35, 588-592.

[2]. Békésy, G. V. (1960). Experiments in Hearing (McGraw-Hill, New York).

[3]. Beverage, H. H., and Peterson, O. H. (1931). "Diversity Receiving System of R.C.A. Communications, Inc.," Proc. IRE, 19, 562-584.

[4]. Box, G. E. P., and Miller, M. E. (1958). "Normal Random Distribution," Ann. Mathematical Statistics, 29, 610-611.

[5]. Brookner, E. (1965). "The Design of Parallel-FSK Systems for Use in Multipath Channels," Proc. IEEE (Correspondence), 53, 1162-1163.

[6]. Campbell, R. A. (1966). "Auditory Perception and Neural Coding," J. Acoust. Soc. Amer., 39, 1030-1035.

[7]. Corlis, E. (1967). "Mechanicalistic Aspects of Hearing," J. Acoust. Soc. Amer., 41, 1500-1516.

[8]. Corlis, E. (1963). "Resolution Limits of Analyzers and Oscillatory Systems," J. Res. N.B.S., 67A, 461-474.

[9]. Davenport, W. B., and Root, W. L. (1958). Random Signals and Noise (McGraw-Hill, New York), Chapter 14.

[10]. Denes, P. B., and Pinson, E. N. (1963). The Speech Chain (Bell Telephone Laboratories, New Jersey).

[11]. Deutsch, S. (1967). Models of Nervous Systems (John Wiley and Sons, Inc., New York).

[12]. Engström, H., Ades, H. W., and Hawkins, J. E. (1962). "Structure and Functions of the Sensory Hairs of the Inner Ear," J. Acoust. Soc. Amer., 34, 1356-1363.

[13]. Fano, R. M. (1961). Transmission of Information (The M.I.T. Press, Cambridge, Massachusetts).

[14]. Fant, G. M. C. (1960). Acoustic Theory of Speech Production (Mouton, the Hague).

[15]. Fernandez, D. (1961). "The Innervation of the Cochlea (Guinea Pig)," Laryngoscope, 61, 1152-1172.

[16]. Flanagan, J. L. (1965). Speech Analysis and Perception (Academic Press, New York).

[17]. Flanagan, J. L. (1962). "Models for Approximating Basilar Membrane Displacement--Part II," Bell System Tech. J., 41, 959-1009.

[18]. Flanagan, J. L. (1960). "Computational Model for Approximating Basilar Membrane Displacement," Bell System Tech. J., 39, 1163-1192.

[19]. Florey, E. (Ed.). (1961). Nervous Inhibition (Pergamon Press, New York).

[20]. Furman, G. G., and Frishkopf, L. S. (1964). "Models of Neural Inhibition in the Mammalian Cochlea," J. Acoust. Soc. Amer., 36, 2194-2201.

[21]. Gabor, D. (1947). "Acoustic Quanta and the Theory of Hearing," Nature, 159, 591-594.

[22]. Green, D. M., and Swets, J. A. (1966). Signal Detection Theory and Psychophysics (John Wiley and Sons, Inc., New York).

[23]. Guild, S. R. (1932). "Correlation of Histological Observations and the Acuity of Hearing," Acta. Oto-Larying. Stockh., 17, 207-210.

[24]. Hamming, R. W. (1962). Numerical Methods for Scientists and Engineers (McGraw-Hill, New York).

[25]. Harris, G. G. (1966). "Brownian Motion in the Cochlea," J. Acoust. Soc. Amer., 40, 1264.

[26]. Helmholtz, H. L. F. (1863). Die Lehre Vonder Tonempfindungen als physiologische Grundluge für die Theorie der Musik, 1st Ed. English Translation, Ellis, A. J. (1875). Sensations of Tone (London).

[27]. Hirsh, I. J. (1952). The Measurement of Hearing (McGraw-Hill, New York).

[28]. Hoel, P. G. (1962). Introduction to Mathematical Statistics (John Wiley and Sons, Inc., New York).

[29]. Jeffress, L. A. (1965). "Stimulus Oriented Approach to Detection Re-examined," J. Acoust. Soc. Amer., 14, 802-806.

[30]. Jordan, D. B., Greenber, H., Eldrege, E. E., and Serniuk, S. (1955). "Multiple Frequency Shift Teletype Systems," Proc. IRE, 43, 1647-1655.

[31]. Khanna, S. M., Sears, S. M., and Tonndorf, J. (1968). "Some Properties of Longitudinal Shear Waves: A Study of Computer Simulation," J. Acoust. Soc. Amer., 43, 1077-1087.

[32]. Kiang, N. Y. S., with the assistance of Watanabe, W., Thomas, E. C., and Clark, L. F. (1965). Discharge Patterns of Single Fibers in the Cats' Auditory Nerve (The M.I.T. Press, Cambridge, Massachusetts).

[33]. Kiang, N. Y. S., and Peak, W. T. (1967). "Shape of Tuning Curves for Single Auditory Nerve Fibers," J. Acoust. Soc. Amer. (Correspondence), 42, 1341-1342.

[34]. Littler, T. S. (1965). The Physics of the Ear (Pergamon Press, London).

[35]. Marcum, J. I. (1960). "A Statistical Theory of Target Detection by Pulsed Radar," IRE Trans. Inform. Theory, IT-6, 59-267.

[36]. Nomoto, N., Suga, N., and Katsuki, Y. (1964). "Discharge Patern and Inhibition of Primary Auditory Nerve Fibers in the Monkey," J. Neurophysiol, 27, 768-787.

[37]. Ohm, G. S. (1843). "Ueber die Definition des Tones, nebst daran geknüpfter Theorie der sirene und ähnlicher tonbildener Vorrichtungen," Ain. d. Phys., 59, 497-565.

[38]. Panter, P. F. (1965). Modulation, Noise, and Spectral Analysis (McGraw-Hill, New York).

[39]. Papoulis, A. (1962). The Fourier Integral and Its Applications (McGraw-Hill, New York), 53-75.

[40]. Perfel, D. H., and Bullock, T. H. (1968). "Neural Coding," Neuroscience Res. Prog. Bull., 6, 236.

[41]. Plomp, R. (1964). "The Ear as a Frequency Analyzer," J. Acoust. Soc. Amer., 36, 1628-1636.

[42]. Plomp, R., and Mimpen, A. M. (1968). "The Ear as a Frequency Analyzer II," J. Acoust. Soc. Amer., 43, 764-767.

[43]. Polyak, S. L. (1946). The Human Ear in Anatomical Transparencies (Sonotone Corporation, Elmsford, New York).

[44]. Postman, L. J. (1946). "Time-Error in Auditory Perception," Amer. J. Psychol., 59, 193-219.

[45]. Ratliff, F., Hartline, H. K., and Miller, W. H. (1963). "Spatial and Temporal Aspects of Retinal Inhibitory Interaction," J. Optical Soc. Amer., 53, 110-120.

[46]. Rice, S. O. (1942-1944). "Mathematical Analysis of Random Noise," Bell System Tech. J., 23, 282-332, 24, 46-156. Reprinted in Wax, N., Ed. (1954). Noise and Stochastic Processes (Dover, New York).

[47]. Rodieck, R. W., Kiang, N. Y. S., and Gerstein, G. L. (1962). "Some Quantitative Methods for the Study of Spontaneous Activity of Single Neurons," Biophys. J., 2, 351-368.

[48]. Rosenblith, W. A. (Ed.). (1961). Sensory Communications (The M.I.T. Press, Cambridge, Massachusetts).

[49]. Ruben, R. J., Fisch, U., and Hudson, W. (1962). "Properties of the Eight Nerve Action Potential," J. Acoust. Soc. Amer., 34, 99-102.

[50]. Rutherford, W. (1886). "A New Theory of Hearing," J. of Anat. Physiol., 21, 166-168.

[51]. Sachs, M. B., and Kiang, N. Y. S. (1968). "Two-Tone Inhibition in Auditory Nerve Fibers," J. Acoust. Soc. Amer. 43, 1120-1128.

[52]. Schafer, R. W., and Rabiner, L. R. (1970). "Systems for Automatic Formant Analysis of Voiced Speech," J. Acoust. Soc. Amer., 47, 634-648.

[53]. Schwartz, M. (1959). Information Transmission, Modulation and Noise (McGraw-Hill, New York).

[54]. Schwartz, M., Bennett, W. R., and Stein, S. (1966). Communication Systems and Techniques (McGraw-Hill, New York).

[55]. Shower, E. G., and Biddulph, R. (1931). "Differential Pitch Sensitivity of the Ear," J. Acoust. Soc. Amer., 3, 275-287.

[56]. Siebert, W. M. (1970). "Frequency Discrimination in the Auditory System: Place or Periodicity Mechanism?" Proc. IEEE, 58, 723-730.

[57]. Siebert, W. M. (1962). "Models for the Dynamic Behavior of the Cochlear Partition," M.I.T. Research Lab. of Elect. Tech. Report No. 64, Jan., 242-248.

[58]. Skolnik, M. I. (1962). Introduction to Radar Systems (McGraw-Hill, New York).

[59]. Stein, S. (1964). "Unified Analysis of Certain Coherent and Non-Coherent Binary Communications Systems," IEEE Trans. Inform. Theory, IT-10, 43-51.

[60]. Stewart, G. W. (1931). "Problems Suggested by an Uncertainty Principle in Acoustics," J. Acoust. Soc. Amer., 2, 325.

[61]. Weiss, T. F. (1964). "A Model for Firing Patterns of Auditory Nerve Fibers," M.I.T. Research Lab. of Elect. Tech. Report No. 418, March.

[62]. Wever, E. G. (1949). Theory of Hearing (John Wiley and Sons, Inc., New York).

[63]. Wever, E. G., and Lawrence, M. (1952). Physiological Acoustics (Princeton University Press, Princeton, New Jersey).

[64]. Zwislocki, J. (1962). "Theory of Temporal Auditory Summation," J. Acoust. Soc. Amer., 32, 1046-1060.

VITA

The author was born in , . He received the degree of Bachelor of Science in Electrical Engineering from Newark College of Engineering in 1964.

From 1964 to 1966 he was a teaching assistant in electrical engineering at Rutgers University, New Brunswick, New Jersey, receiving the degree of Master of Science, major in Electrical Engineering from Rutgers University in 1966.

In 1966 he became a teaching fellow at Newark College of Engineering, remaining in this position until 1968 when he became a special lecturer in the electrical engineering department.

In 1970 he became an instructor in the mathematics department of Newark College of Engineering.

His industrial experience includes mostly summer employment: with I.T.T. Communications Corporation, Nutley, New Jersey, in 1966 and 1967 (microwave systems and communications theory); RKO General, Inc., New York in 1968 and 1969 (broadcasting).

From 1964 to the present he has been contributing V.H.F. editor of CQ-Radio Amateur Journal for which he has authored several papers.

The author is a member of Tau Beta Pi and Eta Kappa Nu, has a First Class Radio Telephone License and Advanced Amateur Radio Operator's License, and is a member of the Institute of Electrical and Electronic Engineers and of the Acoustical Society of America.

A list of papers generated in conjunction with and during the preparation of the dissertation follows:

Simulation of Centering in the Presence of Extrinsic noise with application to Audition by A. Katz and M. Zambuto, accepted for presentation and publication in Proceedings of 1971 Summer Computer Simulation Conference, July 19-21, Boston, Massachusetts.

Analysis of Parallel Channel Signal Processing Systems with Application to Audition by A. Katz and M. Zumbuto. Paper submitted for publication in the Journal of Acoustical Society of America, September 1970.

A centering Model of the Peripheral Auditory System by A. Katz and M. Zambuto. Abstract published in the 1970 Proceedings of the Annual Conference on Engineering in Medicine and Biology, p. 233 (1970).

Parallel Channels Models of Auditory Detection by A. Katz and M. Zambuto presented April 21, 1970, at 79th Meeting of Acoustical Society of America. Abstract published in the Journal of Acoustical Society of America, Vol. 48, July, p. 71 (1970).

Signal-Noise Model of Hearing by M. Rosenstein and A. Katz, presented at 69th Meeting of Acoustical Society of America. Abstract published in the Journal of Acoustical Society of America, Vol. 44, p. 384 (1968).